
Manual do Usuário

StatTools

*Add-In de Estatística para o
Microsoft® Excel*

Versão 5.7
1 de setembro de 2010

Palisade Corporation
798 Cascadilla St.
Ithaca, NY 14850
EUA
+1-607-277-8000
<http://www.palisade.com>

Direitos autorais

Copyright © 2010, Palisade Corporation.

Reconhecimento de marcas comerciais

Microsoft, Excel e Windows são marcas registradas da Microsoft Corporation.

IBM é marca registrada da International Business Machines, Inc.

Palisade, TopRank, BestFit e RISKview são marcas registradas da Palisade Corporation.

Bem-vindo ao StatTools para o Excel

Bem-vindo

O StatTools coloca à sua disposição um novo conjunto avançado de ferramentas de estatística para uso com o Microsoft Excel – a ferramenta de modelagem e análise de dados padrão do setor! O StatTools é um suplemento (add-in) de estatística para o Microsoft Excel, que possibilita analisar dados em planilhas de Excel e trabalhar no ambiente do Microsoft Office que você já conhece. Combinando um gerenciador de dados poderoso com análises que rivalizam com os melhores pacotes de estatísticas disponíveis, o StatTools oferece o melhor de dois mundos: os relatórios e a facilidade de uso do Microsoft Office e recursos avançados de estatística.

Trabalhe em um ambiente conhecido

Se você sabe usar o Excel, sabe como trabalhar com o StatTools! O StatTools funciona exatamente como o Excel, com barras de ferramentas, menus e funções de planilhas personalizadas – tudo no próprio Excel. Ao contrário dos softwares autônomos de estatísticas, o StatTools não demanda um longo período de aprendizagem nem altos custos antecipados de treinamento, pois você trabalha da forma como já está acostumado a fazer no Excel. Os seus dados e variáveis estão em planilhas do Excel. Você pode usar fórmulas padrão do Excel para fazer cálculos e transformações, usar recursos de classificação e tabelas dinâmicas. Relatórios e gráficos de análises estatísticas são gerados em formato padrão do Excel, aproveitando todos os recursos de formatação integrados no Excel.

Estatísticas sólidas no Excel

O StatTools substitui as funções estatísticas incorporadas no Excel pelos seus próprios cálculos sólidos e rápidos. A exatidão dos cálculos estatísticos incorporados no Excel tem sido muitas vezes questionada – o StatTools não usa nenhum deles! Até mesmo as funções estatísticas padrão do Excel – como DESVPAD() – são substituídas por versões novas e mais robustas do StatTools, como a StatSTDEV(). Os cálculos estatísticos do StatTools satisfazem os mais rigorosos testes de exatidão, e apresentam um desempenho otimizado pelo uso de .DLLs de C++, e não cálculos por macros.

Análises do StatTools

O StatTools abrange a gama de procedimentos estatísticos mais utilizados, e oferece recursos inigualáveis para a adição de análises novas e personalizadas. Um total de 36 procedimentos estatísticos abrangentes e 8 utilidades de dados incorporados cobrem as análises estatísticas mais amplamente utilizadas. As funções estatísticas oferecidas incluem estatísticas descritivas, testes de normalidade, comparações de grupos, correlação, análise de regressão, controle de qualidade, previsões e outras. Acrescente a isso uma biblioteca de procedimentos personalizados (escritos por seu pessoal ou outros especialistas na área) e você terá um conjunto de ferramentas completo e personalizável, que pode ser usado diretamente dentro do Excel!

O StatTools oferece cálculos estatísticos em tempo real com “vínculos ativos”! Ao alterar um valor no Excel, você quer que a sua planilha recalcule e forneça uma nova resposta. Bom, o mesmo acontece com o StatTools! Altere um valor no seu conjunto de ferramentas e o seu relatório estatístico é automaticamente atualizado. O StatTools usa um conjunto poderoso de funções personalizadas de planilha para garantir que as estatísticas exibidas nos relatórios estejam sempre atualizadas com os dados mais recentes.

Gerenciamento de dados no StatTools

O StatTools oferece um conjunto de ferramentas abrangente e um gerenciador de variáveis dentro do Excel, como você esperaria de um pacote de estatísticas autônomo. Você pode definir quantos conjuntos de dados quiser, cada um com as variáveis que deseja analisar, diretamente com os dados já contidos no Excel. O StatTools analisa os blocos de dados de forma inteligente, sugerindo nomes e tipos de variáveis, além de locais para os dados. Os conjuntos de dados e as variáveis podem residir em diferentes planilhas e pastas de trabalho, o que permite organizá-los de acordo com a sua preferência. Assim, você pode executar análises estatísticas relacionadas às suas variáveis, em vez de ter de selecionar repetidamente os dados no Excel. Além disso, as variáveis do StatTools não se limitam a uma única coluna de dados na planilha do Excel; você pode usar a mesma coluna para uma única variável em até 255 planilhas! (65.535 X 255, ou mais de 16 milhões de casos no StatTools Industrial; 10.000 casos no StatTools Professional Edition)

Relatórios do StatTools

O Excel é excelente para criar relatórios e gráficos, e o StatTools tira máximo proveito disso. O aplicativo usa gráficos em formato Excel que podem ser facilmente personalizados com novas cores, fontes e texto. Títulos de relatórios, formatos numéricos e texto podem ser alterados da mesma forma que em uma planilha Excel padrão. Arraste e solte tabelas e gráficos de relatórios do StatTools diretamente em documentos de outras aplicações. Gráficos e tabelas permanecem vinculados aos seus dados no Excel, assim, sempre que os relatórios de análises forem alterados, o documento é automaticamente atualizado.

Acesso e compartilhamento de dados

O Excel conta com ótimos recursos para importação de dados, o que facilita muito a transferência de seus dados para o StatTools! Use os recursos padrão do Excel para ler dados do Microsoft SQL Server, Oracle, Microsoft Access ou qualquer outro banco de dados compatível com ODBC. Carregue dados de arquivos de texto ou outras aplicações – se os dados podem ser lidos no Excel, podem ser usados no StatTools!

O StatTools salva todos os resultados e dados em pastas de trabalho do Excel. Da mesma forma que com qualquer arquivo do Excel, é possível enviar resultados e dados do StatTools a colegas em qualquer lugar. O compartilhamento não poderia ser mais fácil!

StatTools Industrial

O StatTools Industrial inclui uma interface de programação completa, orientada a objeto, em que os procedimentos estatísticos podem ser adicionados utilizando-se a linguagem de programação VBA incorporada no Excel. Os procedimentos personalizados podem utilizar as ferramentas de gerenciamento de dados, geração de gráficos e relatórios incorporadas no StatTools, todas acessíveis pelos controles, funções e métodos personalizados do StatTools. Os procedimentos personalizados podem até mesmo ser exibidos no menu do StatTools, para acesso mais fácil!

E se você não pretender escrever os próprios procedimentos estatísticos, também não tem problema. O StatTools Industrial proporciona uma grande vantagem mesmo assim, pois permite personalizar procedimentos escritos por terceiros, diretamente no menu padrão do StatTools! Especialistas na área constantemente acrescentam à biblioteca procedimentos novos e personalizados construídos com o StatTools. É só copiar uma pasta de trabalho com um novo procedimento na pasta do StatTools do seu computador e ele será exibido instantaneamente no menu do StatTools. Execute-o e você verá todas as ferramentas de gerenciamento de dados do StatTools, combinadas com as novas análises estatísticas de que você necessita!

Índice

Capítulo 1: Primeiros Passos	1
Introdução	3
Verificação do pacote.....	3
O que está incluído no pacote.....	3
Sobre esta versão	3
Funcionamento com o seu ambiente operacional	4
Se necessitar de ajuda.....	4
Requisitos de sistema do StatTools	6
Instruções de instalação	7
Instruções gerais de instalação	7
Instalação de ícones ou atalhos do StatTools.....	8
O DecisionTools Suite	8
Ativação do software	9
Capítulo 2: Visão Geral do StatTools	13
Visão geral	15
Menu e barra de ferramentas do StatTools	15
Conjuntos de dados e Gerenciador de conjunto de dados	16
Relatórios e gráficos do StatTools.....	18
Capítulo 3: Guia de Referência do StatTools	21
Introdução	25
Linguagem de macro VBA do StatTools e kit de ferramentas do desenvolvedor.....	25
Referência: ícones do StatTools	27
Barra de ferramentas do StatTools	27
Referência: comandos do menu StatTools	31
Introdução	31

Pacotes de add-in de análise.....	31
Lista de comandos.....	33
Menu StatTools – Conjunto de dados.....	41
Comando Gerenciador de conjunto de dados.....	41
Menu Utilidades de dados.....	47
Comando Empilhar.....	47
Comando Desempilhar.....	49
Comando Transformar.....	50
Comando Defasagem.....	52
Comando Diferença.....	54
Comando Interação.....	56
Comando Combinação.....	58
Comando Fictício.....	60
Comando Amostra aleatória.....	62
Menu Resumo estatístico.....	65
Comando Resumo de 1 variável.....	65
Comando Correlações e covariância.....	68
Menu Gráficos de resumo.....	71
Comando Histograma.....	71
Comando Gráfico de dispersão.....	74
Comando Gráfico Box-Whisker.....	76
Menu Inferência estatística.....	79
Comando Intervalo de confiança de média/desvio padrão.....	79
Comando Intervalo de confiança – Proporção.....	82
Comando Teste de hipótese - média/desvio padrão.....	85
Comando Teste de hipótese - Proporção.....	88
Comando Seleção de tamanho amostral.....	91
Comando ANOVA unidirecional.....	93
Comando ANOVA bidirecional.....	96
Comando Teste de independência de qui-quadrado.....	98
Menu Testes de normalidade.....	101
Comando Teste de normalidade de qui-quadrado.....	101
Comando Teste de Lilliefors.....	105
Comando Gráfico normal Q-Q.....	108
Menu Série temporal e previsão.....	111
Comando Gráfico de série temporal.....	111
Comando Autocorrelação.....	113
Comando Teste de execuções para verificar aleatoriedade.....	115
Comando Previsão.....	117

Menu Regressão e classificação	123
Comando Regressão	123
Comando Regressão logística	128
Comando Análise discriminante.....	132
Menu Controle de qualidade.....	137
Comando Gráfico de Pareto	137
Comando Gráficos X/R.....	140
Comando Gráfico P.....	144
Comando Gráfico C	148
Comando Gráfico U.....	151
Menu Testes não-paramétricos.....	155
Comando Teste de sinal.....	157
Comando Teste dos postos sinalizados de Wilcoxon.....	160
Comando Teste de Mann-Whitney	163
Menu Utilidades	167
Comando Configurações da aplicação	167
Comando Excluir conjuntos de dados.....	173
Comando Limpar memória da caixa de diálogo.....	173
Comando Descarregar o add-in StatTools.....	173
Menu Ajuda.....	175
Ajuda do StatTools	175
Manual on-line	175
Comando Ativação da licença.....	175
Comando Sobre	175
Referência: Funções do StatTools	177
Introdução	177
Funções do StatTools vs funções do Excel.....	178
Funções de distribuição	179
Relatórios “em tempo real”	180
Referência: listagem das funções estatísticas.....	183
Tabela de funções disponíveis	183
Descrições detalhadas das funções	187
Índice	203

Capítulo 1: Primeiros Passos

Introdução	3
Verificação do pacote.....	3
O que está incluído no pacote.....	3
Sobre esta versão	3
Funcionamento com o seu ambiente operacional	4
Se necessitar de ajuda.....	4
Requisitos de sistema do StatTools	6
Instruções de instalação.....	7
Instruções gerais de instalação	7
Instalação de ícones ou atalhos do StatTools.....	8
O DecisionTools Suite	8

Introdução

Esta introdução descreve o conteúdo da embalagem do StatTools e explica como instalá-lo e vinculá-lo ao Microsoft Excel 2000 para Windows 2000 ou versão mais recente.

Verificação do pacote

O pacote do StatTools deve conter:

Learning Statistics with StatTools, um livro de estatísticas que utilizam o StatTools, escrito pelo Dr. S. Christian Albright, da Universidade de Indiana

CD-ROM do StatTools ou DecisionTools Suite com:

- *Programa StatTools*
- *Tutorial do StatTools*
- *Manual do Usuário do StatTools (este livro) em formato .PDF*

O contrato de licença do StatTools

Se algo estiver faltando, entre em contato com o seu revendedor ou distribuidor do StatTools, ou ligue diretamente para a Palisade Corporation pelo número +1-607-277-8000.

O que está incluído no pacote

O StatTools pode ser adquirido individualmente ou com as versões Professional e Industrial do DecisionTools Suite. O CD-ROM do StatTools contém o add-in StatTools Excel, vários exemplos para o StatTools, e um sistema de ajuda eletrônica totalmente indexado. As versões Professional e Industrial do DecisionTools Suite contêm todos os componentes mencionados acima, além de aplicações adicionais.

Sobre esta versão

Esta versão do StatTools pode ser instalada como programa de 32 bits para o Microsoft Excel 2000 ou versão mais recente.

Funcionamento com o seu ambiente operacional

Este Manual do Usuário pressupõe que o usuário saiba, de modo geral, usar o sistema operacional Windows e o Excel. Especificamente:

- *O usuário sabe usar o computador e o mouse.*
- *O usuário conhece termos como: ícones, clique, duplo-clique ou clicar duas vezes, menu, janela, comando, objeto.*
- *O usuário tem uma compreensão de conceitos básicos como: estrutura de pastas, atribuição de nomes a arquivos.*

Se necessitar de ajuda

Fornecemos suporte técnico gratuito a todos os usuários registrados do StatTools com plano de manutenção vigente; também oferecemos suporte técnico mediante pagamento por incidente individual. Para ter certeza de estar registrado como usuário do StatTools, **faça seu registro online, no site** <http://www.palisade.com/support/register.asp>.

Ao nos contatar por telefone, tenha à mão o número de série do seu produto e o Manual do Usuário. Podemos prestar melhor suporte técnico se você estiver em frente ao seu computador, pronto para trabalhar.

Antes de nos contatar

Antes de contatar o suporte técnico, confira o seguinte:

- *Você consultou a Ajuda online?*
- *Você consultou este Manual do Usuário e assistiu ao tutorial multimídia online?*
- *Você leu o arquivo LEIAME.WRI? Ele contém informações atualizadas sobre o StatTools, que podem não estar incluídas no manual.*
- *O problema que está ocorrendo pode ser reproduzido sempre da mesma forma? É possível reproduzir o problema em outro computador ou outro modelo de computador?*
- *Você consultou nosso site na Internet? O endereço é: <http://www.palisade.com>. O site também contém respostas a perguntas frequentes (FAQ), na forma de um banco de dados pesquisável de perguntas com as respostas do suporte técnico, assim como patches para o StatTools, na seção Suporte Técnico. Recomendamos acessar o site regularmente para ver as últimas informações sobre o StatTools e outros programas da Palisade.*

Como contatar a Palisade

A Palisade Corporation recebe com satisfação perguntas, comentários e sugestões relacionados ao StatTools. Entre em contato com a nossa equipe de suporte técnico pelos seguintes meios:

- *E-mail: support@palisade.com.*
- *Telefone: +1-607-277-8000, dias úteis, das 9h às 17h (horário de Nova York). Siga as instruções telefônicas para ser conectado ao suporte técnico.*
- *Fax: +1-607-277-8001.*
- *Correspondência:*

**Technical Support
Palisade Corporation
798 Cascadilla St.
Ithaca, NY 14850
EUA**

Se quiser contatar a Palisade Europe:

- *E-mail: support@palisade-europe.com.*
- *Telefone: +44 1895 425050 (RU).*
- *Fax: +44 1895 425051 (RU).*
- *Correspondência:*

**Palisade Europe
31 The Green
West Drayton
Middlesex
UB7 7PN
Reino Unido**

Para contatar a Palisade Asia-Pacific:

- *E-mail: support@palisade.com.au.*
- *Telefone: +61 2 9252 5922 (Austrália).*
- *Fax: +61 2 9252 2820 (Austrália).*
- *Correspondência:*

**Palisade Asia-Pacific Pty Limited
Suite 404, Level 4
20 Loftus Street
Sydney NSW 2000
Austrália**

Ao nos contatar, pedimos que sempre inclua o nome do produto, a versão e o número de série. O número exato da versão pode ser encontrado no comando Sobre, da Ajuda, no menu do StatTools no Excel.

Versão Student

Não oferecemos suporte técnico por telefone para a versão de estudante do StatTools. Se necessitar de ajuda, recomendamos as seguintes alternativas:

- ◆ *Consulte o seu professor ou colega.*
- ◆ *Acesse <http://www.palisade.com> e veja as respostas às perguntas mais comuns.*
- ◆ *Contate nosso departamento de suporte técnico por e-mail ou fax.*

Requisitos de sistema do StatTools

Os requisitos de sistema para uso do StatTools com Microsoft Excel para Windows são os seguintes:

- *Pentium PC ou processador mais veloz, com disco rígido.*
- *Microsoft Windows 2000 SP4, Windows XP ou versão superior.*
- *Microsoft Excel 2000 ou versão superior.*

Instruções de instalação

Instruções gerais de instalação

O programa de instalação copia os arquivos de sistema do StatTools no diretório especificado do disco rígido. Para executar o programa de instalação no Windows 2000 ou versão superior:

- 1) *Insira o CD-ROM do StatTools ou do DecisionTools Suite na unidade de CD-ROM do seu computador*
- 2) *Clique no botão Iniciar; em seguida, clique em Configurações e em Painel de Controle*
- 3) *Clique duas vezes no ícone Adicionar/Remover Programas*
- 4) *Na guia Instalar/Desinstalar, clique no botão Instalar*
- 5) *Siga as instruções do programa de instalação apresentadas na tela*

Se tiver algum problema durante a instalação do StatTools, verifique se a unidade de disco na qual a instalação está sendo feita tem espaço suficiente. Após liberar espaço suficiente, tente executar a instalação novamente.

Remoção do StatTools do computador

Se quiser remover o StatTools de seu computador, use o recurso Adicionar/Remover Programas no Painel de Controle e selecione o item correspondente ao StatTools.

Instalação de ícones ou atalhos do StatTools

No Windows, o programa de instalação cria automaticamente um comando do StatTools no menu Programas\Palisade DecisionTools da barra de tarefas. Contudo, se houver algum problema durante a instalação, ou se quiser fazer isso manualmente em outra ocasião, siga estas instruções. Observe que as instruções abaixo são para o Windows XP Profissional. As instruções podem ser diferentes para outros sistemas operacionais.

- 1) *Clique no botão Iniciar e, em seguida, aponte para Configurações.*
- 2) *Clique na guia do menu Iniciar, na barra de tarefas.*
- 3) *Clique em Personalizar, Adicionar e, em seguida, Procurar.*
- 4) *Localize o arquivo StatTools.EXE e clique nele; em seguida, clique em OK.*
- 5) *Clique em Avançar e, em seguida, clique duas vezes no menu em que deseja incluir o programa.*
- 6) *Digite o nome "StatTools" e clique em Concluir.*
- 7) *Clique em OK em todas as caixas de diálogo abertas.*

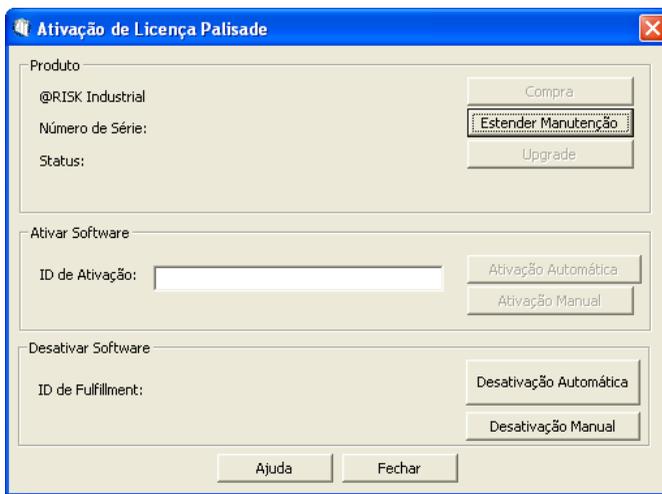
O DecisionTools Suite

O StatTools faz parte do DecisionTools Suite, um conjunto de produtos para análise de risco e decisões que pode ser adquirido da Palisade Corporation. O procedimento de instalação padrão do StatTools coloca o StatTools em uma subpasta de "Arquivos de programas\Palisade". O processo é bem semelhante ao da instalação do Excel, muitas vezes feita em uma subpasta denominada "Microsoft Office".

Um dos subdiretórios de Arquivos de Programas\Palisade será a pasta do StatTools (por definição padrão chamado de StatTools5). Este diretório contém o arquivo do suplemento StatTools (STATTOOLS.XLA), além de exemplos de modelos e outros arquivos necessários para executar o StatTools. Um outro subdiretório de Arquivos de Programas\Palisade é o diretório SISTEMA (ou SYSTEM), que contém os arquivos necessários para cada programa do DecisionTools Suite, inclusive as bibliotecas dos programas e arquivos de Ajuda comuns.

Ativação do software

A ativação é um processo de verificação de licença que é efetuado apenas uma vez e é necessário para poder executar o software StatTools como produto plenamente licenciado. A fatura impressa do produto, ou a fatura enviada por e-mail, contém um **código de ativação**, que consiste em uma sequência separada por traços, como por exemplo: “19a0-c7c1-15ef-1be0-4d7f-cd”. Se você forneceu o código de ativação durante a instalação, o software será ativado na primeira vez que for executado, e nenhuma outra ação será necessária. Para ativar o software após a instalação, selecione o comando Ativação de Licença no menu Ajuda do StatTools e digite o código de ativação na caixa de diálogo **Ativação de Licença Palisade** apresentada.



1) E se o software não for ativado?

Se o código de ativação não for fornecido durante a instalação, ou se for instalada uma versão de avaliação, o software será executado como versão de avaliação, com limitações de tempo e número de usos, e terá de ser ativado com um código de ativação para funcionar de modo pleno.

2) Durante quanto tempo posso usar o produto sem ativá-lo?

Sem ativação, o software pode ser usado por um período de 15 dias. Toda a funcionalidade estará presente, mas a caixa de diálogo de Ativação de Licença aparecerá cada vez que o programa for iniciado, para lembrá-lo de ativar a licença e para indicar o período de uso restante. Após o período de avaliação de 15 dias, o software só rodará se for ativado.

3) Como faço para verificar o status da ativação?

A caixa de diálogo de Ativação de Licença pode ser acessada através do comando Ativação de licença, no menu Ajuda do StatTools. O software ativado é indicado com o status **Ativado**; a versão de avaliação é indicada pelo status **Não Ativado**. Se o software não tiver sido ativado, será indicado o tempo restante de execução permitido.

4) Como faço para ativar o software?

Se você ainda não tem um código de ativação, clique no botão **Comprar**, na caixa de diálogo **Ativação de Licença**, para obtê-lo mediante a compra do software. Ao ser efetuada a compra online, o comprador recebe imediatamente um código de ativação e um link opcional para fazer download do instalador, caso haja necessidade de reinstalar o software. Para efetuar a compra por telefone, ligue para o escritório local da Palisade; as informações de contato são fornecidas na seção **Como contatar a Palisade**, neste capítulo.

A ativação pode ser feita pela Internet ou por e-mail:

- **Ativação pela Internet**

Na caixa de diálogo **Ativação de Licença** da Palisade, digite ou cole o código de ativação e pressione **“Ativação Automática”**. Após alguns segundos deverá aparecer uma mensagem indicando que a ativação foi satisfatória; a caixa de diálogo de **Ativação de Licença** indicará o status do software como ativado.

- **Ativação se você não tiver acesso à Internet**

A ativação automática por e-mail requer algumas etapas:

1. **Clique em “Ativação Manual”** para exibir o arquivo `request.xml`, que pode ser gravado ou copiado para a área de transferência do Windows (recomenda-se anotar o endereço do arquivo `request.xml` em seu computador).
2. **Copie ou anexe o arquivo XML** no e-mail e envie-o para `activation@palisade.com`. Você deverá receber uma resposta automática no seu endereço de retorno de e-mail em breve.
3. **Salve o anexo `response.xml`** do e-mail de resposta no disco rígido.
4. **Clique no botão `Processar`**, que agora aparece na caixa de diálogo **Ativação de licença Palisade**, e vá até o arquivo `response.xml`. Selecione o arquivo e clique em **OK**.

Deverá aparecer uma mensagem indicando que a ativação foi satisfatória e a caixa de diálogo de **Ativação de Licença** refletirá o status do software ativado.

5) Como faço para transferir a licença do software para outro computador?

A transferência de uma licença ou “**rehosting**” (mudança de host), pode ser executada na caixa de diálogo Ativação de licença Palisade como procedimento de duas etapas: *desativação* no primeiro computador e *ativação* no segundo. Uma utilização típica de rehosting é transferir a cópia do StatTools do PC do escritório para o seu laptop. Para mudar o host de uma licença do *Computador1* para o *Computador2*, certifique-se de o software foi instalado nos dois computadores e que ambos estejam conectados à Internet durante a desativação/ativação.

1. No *Computador1*, clique na opção **Desativação Automática**, na caixa de diálogo Ativação de licença. Aguarde até aparecer a mensagem indicando que a ativação foi efetuada.
2. No *Computador2*, clique na opção **Ativação Automática**. Aguarde até aparecer a mensagem indicando que a ativação foi efetuada.

Se os computadores não tiverem acesso à Internet, será necessário seguir instruções semelhantes às fornecidas acima para fazer a mudança de host pelo processo automático por e-mail.

6) Tenho acesso à Internet mas não consigo Ativar/Desativar automaticamente.

É necessário definir o firewall do seu sistema para que permita acesso TCP ao servidor de licenças. No caso de instalações de usuários individuais (não instalações de rede), o endereço é <http://service.palisade.com:8888> (porta TCP 8888 em <http://service.palisade.com>).

Capítulo 2: Visão Geral do StatTools

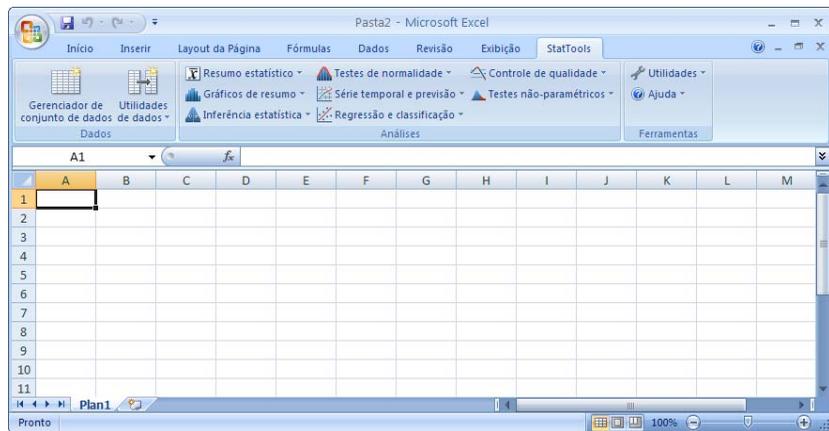
Visão geral	15
Menu e barra de ferramentas do StatTools	15
Conjuntos de dados e Gerenciador de conjunto de dados	16
Dados de múltiplos intervalos	17
Dados empilhados e desempilhados	17
Tratamento de valores ausentes	18
Relatórios e gráficos do StatTools	18
Uso de fórmulas em vez de valores	19
Uso de comentários em células	20

Visão geral

O StatTools apresenta recursos avançados de estatística em um ambiente que você já conhece: o Microsoft Excel. Os procedimentos do StatTools – como a criação de gráficos de dispersão, teste de uma variável para normalidade e execução de análise de regressão – podem ser executados com os seus dados no Excel e os relatórios e gráficos das suas análises criadas no Excel.

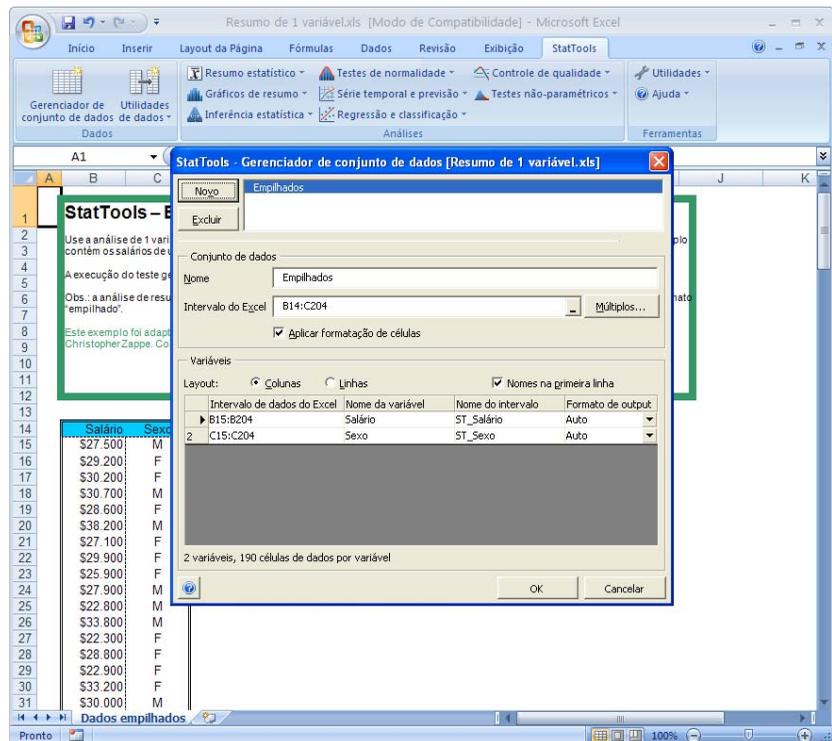
Menu e barra de ferramentas do StatTools

Após o StatTools ser instalado, o menu e os comandos do programa são incluídos na barra de menus do Excel 2003 e versões anteriores. Também é exibida uma barra de ferramentas do StatTools. No Excel 2007 e mais recentes, uma faixa do StatTools será exibida.



Conjuntos de dados e Gerenciador de conjunto de dados

O StatTools é análogo à maioria dos pacotes de software estatísticos autônomos por ser estruturados com base em variáveis. Para a maioria das análises, é necessário que você trabalhe com um conjunto de dados, ou com um conjunto de variáveis estatísticas, muitas vezes localizado em colunas contíguas, com os nomes das variáveis na primeira linha do conjunto de dados. O **Gerenciador de conjunto de dados** do StatTools permite definir conjuntos de dados, variáveis. Subsequentemente, essas variáveis predefinidas podem ser usadas nas suas análises estatísticas, sem necessidade de selecionar novamente os dados a serem analisados.



Cada variável contida no conjunto de dados tem um nome e um intervalo de células correspondente no Excel. Um layout de variável típica é **1 variável por coluna**, porém, as variáveis também podem ser dispostas em linhas. O conjunto de dados pode incluir vários blocos de células, o que permite colocar dados em diferentes planilhas de uma mesma pasta de trabalho.

Durante a definição de um conjunto de dados, o StatTools tenta identificar as variáveis contidas em um bloco de células circundante à seleção feita no Excel. Isso torna mais rápido e fácil definir um conjunto de dados com os nomes das variáveis na linha superior e as variáveis dispostas em colunas.

Os comprimentos das colunas no conjunto de dados não precisam necessariamente ser iguais. Por exemplo, é possível ter duas variáveis, *Peso_Homens* e *Peso_Mulheres*, com diferentes números de observações. Contudo, em muitas análises, o StatTools trata as células em branco nas colunas mais curtas como dados ausentes.

No Excel 2003 ou versões anteriores, uma coluna individual da planilha pode conter até 65536 pontos de dados correspondentes a uma variável. Se as suas variáveis tiverem mais valores que isso e você não usar o Excel 2007, o StatTools permite atribuir múltiplos intervalos de células a um mesmo conjunto de dados. Por exemplo, é possível “repetir” um conjunto de dados em diversas planilhas, designando as mesmas colunas em diferentes planilhas para conter todos os valores de um determinado conjunto de dados. Você também poderia usar esta capacidade para atribuir diferentes blocos de células na mesma planilha a um único conjunto de dados. Isto é útil se os seus dados estiverem dispersados em uma única planilha, e você quiser combiná-los todos em um único conjunto de dados.

O StatTools é compatível com dados empilhados e desempilhados. Com alguns procedimentos estatísticos, é fácil trabalhar com dados empilhados, e, com outros, dados desempilhados. Por exemplo, se compararmos as rendas médias familiares de bairros diferentes, na forma desempilhada haveria uma variável *Renda* (ou coluna) para cada bairro. No formato empilhado haveria uma variável de valor *Renda* e uma variável de categoria *Bairro* indicando o bairro em que está cada família.

A utilidade de empilhamento de variáveis do StatTools permite o “empilhamento” das suas variáveis em duas colunas: uma coluna de valor, *Renda*, e uma coluna de categoria, *Bairro*. Dependendo do tipo de análise, o conjunto de dados empilhado pode ser mais fácil de usar do que a versão de dados desempilhados.

Dados de múltiplos intervalos

Dados empilhados e desempilhados

Tratamento de valores ausentes

Se o seu conjunto de valores tiver dados ausentes (uma ocorrência comum em análises estatísticas), o StatTools lida com eles da forma apropriada, dependendo da tarefa. Por exemplo, indicadores de resumo, como médias e desvios padrões, ignoram valores ausentes. Um outro exemplo seria uma análise de regressão envolvendo três variáveis, que usa apenas as linhas dos conjuntos de dados que não tenham valores ausentes para nenhuma das três variáveis. (Essa técnica de exclusão é chamada de “listwise” ou “casewise”.) Um terceiro exemplo: um gráfico de dispersão de duas variáveis traça apenas os pontos em que ambas as variáveis não tenham valores ausentes.

Nota: nem todos os procedimentos do StatTools permitem valores ausentes. Verifique a seção Referência deste manual para saber quais procedimentos lidam com valores ausentes.

Relatórios e gráficos do StatTools

Sempre que o StatTools cria um output numérico, como um relatório de análise de regressão ou um tabela de resumo de estatísticas, ele fornece um conjunto de opções para a colocação do relatório. As opções são:

- **Nova pasta de trabalho** – é criada uma nova pasta de trabalho (se necessário) e cada relatório é colocado em uma planilha da pasta de trabalho.
- **Nova planilha da pasta de trabalho ativa** – cada novo relatório é colocado em uma nova planilha da pasta de trabalho ativa.
- **Depois da última coluna usada na planilha ativa** – cada relatório é colocado na planilha ativa, à direita da última coluna utilizada.
- **Célula inicial** – você seleciona a célula e o relatório ou gráfico é colocado no canto superior esquerdo da mesma.

Sempre que o StatTools cria um ou mais gráficos, eles são incluídos nos relatórios. Os gráficos são criados em formato Excel e podem ser personalizados usando-se os comandos padrão de gráficos do Excel.

Uso de fórmulas em vez de valores

Por padrão, o StatTools tenta fazer com que os resultados sejam o mais “em tempo real” possível. Ou seja, sempre que possível, os relatórios contêm fórmulas vinculadas aos dados originais. Por exemplo, suponha que haja uma variável *Peso* e você queira indicadores de resumo sobre *Peso*, como média e desvio padrão. O procedimento Resumo estatístico atribui à gama de pesos o nome *Peso*, e em seguida insere as fórmulas nas células de output: **=StatMean(Peso)** e **=StatStdDev(Peso)**. StatMean e StatStdDev são incorporadas nas funções do StatTools para o cálculo da média e do desvio padrão. Elas substituem as funções incorporadas no Excel para as mesmas estatísticas.

Resumo de 1 variável	Salário (F)	Salário (M)
Média	\$29441,67	\$30089,36
Variância	14737192,98	12661821,09
Desvio padrão	\$3838,91	\$3558,35
Assimetria	-0,3034	0,0493
Curtose	3,1872	2,4343
Mediana	\$29700,00	\$29900,00
Desvio médio absoluto	\$3050,35	\$2893,16
Moda	\$25100,00	\$26900,00
Mínimo	\$17100,00	\$22400,00
Máximo	\$36900,00	\$38200,00
Intervalo	\$19800,00	\$15800,00
Contagem	96	94
Soma	\$2826400,00	\$2828400,00
1º quartil	\$26700,00	\$27500,00
3º quartil	\$31900,00	\$32500,00
Intervalo interquartilico	\$5200,00	\$5000,00
1,00%	\$17100,00	\$22400,00
2,50%	\$22300,00	\$23400,00
5,00%	\$22800,00	\$23900,00
10,00%	\$24500,00	\$25200,00
20,00%	\$26500,00	\$26900,00
30,00%	\$32400,00	\$33800,00
40,00%	\$34600,00	\$35000,00
50,00%	\$36200,00	\$35900,00
60,00%	\$36400,00	\$36800,00
70,00%	\$36900,00	\$38200,00

As fórmulas são usadas em relatórios por dois motivos. Primeiro, convém aprender os procedimentos estatísticos e as funções do StatTools no Excel. Você não vê apenas um resultado numérico: você vê como ele foi formado. Segundo, há ainda a vantagem prática de que se os dados mudarem, os resultados mudarão automaticamente, assim não será necessário executar o procedimento outra vez.

Às vezes não é prático fazer isso. O melhor exemplo é o da regressão. O StatTools não fornece as fórmulas utilizadas para criar o output da regressão, apenas os resultados numéricos. Nesses casos, se os seus dados forem alterados, os procedimentos deverão ser executados novamente.

O StatTools também oferece a opção de desativar a atualização em tempo real. Isto é útil se o tempo de recálculo do Excel se tornar um problema cada vez que os dados forem alterados.

Uso de comentários em células

Um dos recursos do Excel é a capacidade de incluir um comentário instantâneo, na forma de pop-up, em qualquer célula. É possível saber se há um comentário em uma célula pela presença de um pequeno triângulo vermelho no canto superior direito da célula. Você pode ler o comentário colocando o cursor sobre a célula. O StatTools aproveita esses comentários para inserir ajuda sensível ao contexto. Esses comentários podem ser considerado a ajuda mais “online” de todas as ajudas online!

A propósito, se já lhe ocorreu o caso de uma planilha em que os comentários não desaparecem, ou seja, ficam o tempo todo na frente dos dados, selecione Opções, no menu Ferramentas, clique na guia Exibir, e clique no botão “Somente indicador de comentário”.

Capítulo 3: Guia de Referência do StatTools

Introdução	25
Linguagem de macro VBA do StatTools e kit de ferramentas do desenvolvedor	25
Referência: ícones do StatTools	27
Barra de ferramentas do StatTools	27
Referência: comandos do menu StatTools	31
Introdução	31
Pacotes de add-in de análise.....	31
Lista de comandos.....	33
Menu StatTools – Conjunto de dados.....	41
Comando Gerenciador de conjunto de dados.....	41
Menu Utilidades de dados	47
Comando Empilhar	47
Comando Desempilhar	49
Comando Transformar	50
Comando Defasagem.....	52
Comando Diferença	54
Comando Interação	56
Comando Combinação	58
Comando Fictício.....	60
Comando Amostra aleatória.....	62
Menu Resumo estatístico	65
Comando Resumo de 1 variável	65
Comando Correlações e covariância	68

Menu Gráficos de resumo	71
Comando Histograma.....	71
Comando Gráfico de dispersão	74
Comando Gráfico Box-Whisker.....	76
Menu Inferência estatística	79
Comando Intervalo de confiança de média/desvio padrão	79
Comando Intervalo de confiança - Proporção	82
Comando Teste de hipótese - média/desvio padrão	85
Comando Teste de hipótese - Proporção.....	88
Comando Seleção de tamanho amostral	91
Comando ANOVA unidirecional	93
Comando ANOVA bidirecional.....	96
Comando Teste de independência de qui-quadrado	98
Menu Testes de normalidade	101
Comando Teste de normalidade de qui-quadrado	101
Comando Teste de Lilliefors.....	105
Comando Gráfico normal Q-Q.....	108
Menu Série temporal e previsão	111
Comando Gráfico de série temporal.....	111
Comando Autocorrelação	113
Comando Teste de execuções para verificar aleatoriedade	115
Comando Previsão	117
Menu Regressão e classificação	123
Comando Regressão	123
Comando Regressão logística	128
Comando Análise discriminante.....	132
Menu Controle de qualidade	137
Comando Gráfico de Pareto	137
Comando Gráficos X/R.....	140
Comando Gráfico P.....	144
Comando Gráfico C	148
Comando Gráfico U	151
Menu Testes não-paramétricos	155
Comando Teste de sinal.....	157
Comando Teste dos postos sinalizados de Wilcoxon.....	160
Comando Teste de Mann-Whitney.....	163

Menu Utilidades.....	167
Comando Configurações da aplicação.....	167
Comando Excluir conjuntos de dados.....	173
Comando Limpar memória da caixa de diálogo.....	173
Comando Descarregar o add-in StatTools.....	173
Menu Ajuda	175
Ajuda do StatTools	175
Manual on-line	175
Comando Ativação da licença	175
Comando Sobre	175
Referência: Funções do StatTools	177
Introdução	177
Funções do StatTools vs funções do Excel.....	178
Funções de distribuição.....	179
Relatórios “em tempo real”	180
Referência: listagem das funções estatísticas	183
Tabela de funções disponíveis.....	183
Descrições detalhadas das funções	187

Introdução

O capítulo Guia de Referência do StatTools descreve os ícones, comandos e funções estatísticas do StatTools. Este capítulo é dividido em quatro seções:

- 1) *Referência: ícones do StatTools*
- 2) *Referência: resumo de procedimentos do StatTools*
- 3) *Referência: comandos do StatTools*
- 4) *Referência: funções do StatTools*

Linguagem de macro VBA do StatTools e kit de ferramentas do desenvolvedor

O StatTools também contém uma poderosa linguagem de macro baseada em VBA, que pode ser utilizada para:

- 1) **Automatizar análises do StatTools.**
- 2) **Desenvolver novas análises estatísticas que usam o Gerenciador de conjunto de dados do StatTools, relatórios e gráficos.** Esses cálculos personalizados possibilitam análises não encontradas nos procedimentos incorporados do StatTools. Essas análises podem aparecer no menu e na barra de ferramentas do StatTools.

Para obter mais informações sobre a linguagem de macro e o kit de ferramentas do desenvolvedor do StatTools, consulte a documentação online fornecida com o produto.

Referência: ícones do StatTools

Barra de ferramentas do StatTools

Os ícones do StatTools são usados para definir conjuntos de dados e variáveis, e executar procedimentos estatísticos com essas variáveis. Os ícones do StatTools aparecem na barra de ferramentas do Excel (como uma barra de ferramentas personalizada do Excel), na versão 2003 e anteriores; no Excel 2007, aparece como uma faixa de opções. Esta seção descreve brevemente cada ícone e as respectivas funções, e os comandos de menu correspondentes.

Os ícones a seguir aparecem na barra de ferramentas do StatTools no Excel 2003 e anteriores.

Ícone **Comandos e funções correspondentes**



Define um conjunto de dados e variáveis, edita ou exclui um conjunto de dados e variáveis existentes

Comando equivalente: comando Gerenciador de conjunto de dados



Executa uma utilidade de dados

Comando equivalente: comando Utilidade de dados



Executa um procedimento de resumo estatístico

Comando equivalente: comando Resumo estatístico



Cria gráficos resumidos para variáveis

Comando equivalente: comando Gráficos de resumo



Executa um procedimento de resumo estatístico

Comando equivalente: comando Inferência estatística



Executa um teste de normalidade em variáveis

Comando equivalente: comando Testes de normalidade



Executa uma série temporal ou procedimento de previsão

Comando equivalente: comando Série temporal e previsão



executa um procedimentos de regressão ou classificação

Comando equivalente: comando Regressão e classificação



Executa um procedimento de controle de qualidade

Comando equivalente: comando Controle de qualidade



Executa uma teste não-paramétrico

Comando equivalente: comando Testes não-paramétricos



Exibe utilidades do StatTools

Comando equivalente: comando Utilidades



Exibe o arquivo de ajuda do StatTools

Comando equivalente: comando Ajuda

Os ícones a seguir aparecem na faixa de opções do StatTools no Excel 2007.

Ícone

Comandos e funções correspondentes



Define um conjunto de dados e variáveis, edita ou exclui um conjunto de dados e variáveis existentes

Comando equivalente: comando Gerenciador de conjunto de dados



Executa uma utilidade de dados

Comando equivalente: comando Utilidade de dados



Executa um procedimento de resumo estatístico

Comando equivalente: comando Resumo estatístico



Cria gráficos resumidos para variáveis

Comando equivalente: comando Gráficos de resumo

 Inferência estatística -

Executa um procedimento de resumo estatístico

Comando equivalente: comando Inferência estatística

 Testes de normalidade -

Executa um teste de normalidade em variáveis

Comando equivalente: comando Teste de normalidade

 Série temporal e previsão -

Executa uma série temporal ou procedimento de previsão

Comando equivalente: comando Série temporal e previsão

 Regressão e classificação -

Executa um procedimentos de regressão ou classificação

Comando equivalente: comando Regressão e classificação

 Controle de qualidade -

Executa um procedimento de controle de qualidade

Comando equivalente: comando Controle de qualidade

 Testes não-paramétricos -

Executa uma teste não-paramétrico

Comando equivalente: comando Testes não-paramétricos

 Utilidades -

Exibe utilidades do StatTools

Comando equivalente: comando Utilidades

 Ajuda -

Exibe o arquivo de ajuda do StatTools

Comando equivalente: comando Ajuda

Referência: comandos do menu StatTools

Introdução

Esta seção do Guia de Referência detalha os comandos do StatTools mostrados no menu StatTools no Excel 2003 ou anteriores, e na faixa StatTools no Excel 2007 e posteriores. Os comandos são explicados na ordem em que aparecem no menu, começando pelo Gerenciador de conjunto de dados. Os ícones do StatTools podem ser usados para executar vários comandos. A seção **Referência: ícones do StatTools** deste capítulo apresenta os comandos correspondentes a cada ícone do **StatTools**.

Diversos comandos do **StatTools** também estão disponíveis em um menu pop-up flutuante que é exibido quando o botão direito do mouse é clicado no Excel.

Pacotes de add-in de análise

O StatTools Industrial pode usar “pacotes de análise”, ou pacotes de add-in de novas análises encontrados no produto central do StatTools. Os comandos encontrados nesses pacotes serão adicionados ao menu StatTools. Assim, ao usar pacotes de análise de add-in, o menu pode não ser exatamente igual ao mostrado aqui. Para obter mais informações sobre o desenvolvimento de procedimentos suplementares do StatTools, consulte a documentação online da linguagem de macro VBA e o kit de ferramentas do desenvolvedor do StatTools.

Lista de comandos

Os procedimentos disponíveis no StatTools são agrupados de forma natural. Para cada grupo há um item no menu StatTools. Se um grupo tiver mais de um item, há um submenu com a lista dos itens. Esta seção fornece uma breve descrição de cada procedimento, em cada grupo. Informações detalhadas sobre cada procedimento são fornecidas na seção **Referência: Comandos do StatTools**, neste capítulo.

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de vars.
Resumo estatístico							
Comando Resumo de 1 variável	Gera um resumo estatístico, incluindo as medidas usuais, como média, mediana e desvio padrão, mais opções, como quartis e percentis.	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-100
Comando Correlações e covariância	Cria uma tabela de correlações e/ou covariâncias para um conjunto de variáveis que você selecionar.	Permitidos no início, meio e fim dos dados	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250
Gráficos de resumo							
Comando Histograma	Cria um histograma para cada variável que você selecionar. Oferece a opção de definir as categorias ou "bins" do histograma.	Permitidos no início, meio e fim dos dados	Tempo real parcial – alterações de dados atualizam o gráfico quando os dados estiverem no intervalo do eixo do X	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-100
Comando Gráfico de dispersão	Cria um gráfico de dispersão para cada par de variáveis que você selecionar.	Permitidos no início, meio e fim dos dados	Tempo real	Apenas dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1-10
Comando Gráfico Box-Whisker	Cria um boxplot simples (se você selecionar uma única variável) ou boxplots lado a lado (se você selecionar diversas variáveis).	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Não	Ignorados	1-10

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de vars.
Inferência estatística							
Comando Intervalo de confiança de média/desvio padrão	Calcula o intervalo de confiança da média e do desvio padrão de variáveis simples, ou as diferenças entre as médias de pares de variáveis. Os intervalos de confiança podem ser calculados através de uma análise uniamostrual, uma análise biamostrual ou uma análise de amostras pareadas.	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)
Comando Intervalo de confiança – Proporção	Analisa a proporção dos itens em uma amostra que pertence a uma determinada categoria (análise uniamostrual), ou compara duas amostras quanto à proporção dos itens em uma determinada categoria (análise biamostrual).	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)
Comando Teste de hipótese - média/desvio padrão	Executa testes de hipótese da média e do desvio padrão de variáveis simples, ou as diferenças entre as médias de pares de variáveis. Os testes de hipótese podem ser executados através de uma análise uniamostrual, uma análise biamostrual ou uma análise de amostras pareadas.	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)
Comando Teste de hipótese - Proporção	Analisa a proporção dos itens em uma amostra que pertence a uma determinada categoria (análise uniamostrual), ou compara duas amostras quanto à proporção dos itens em uma determinada categoria (análise biamostrual).	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de vars.
Comando Seleção de tamanho amostral	Determina o tamanho da amostra (ou tamanhos de amostra) necessário para se obter um intervalo de confiança com um pontos central prescrito. Faz isso para intervalos de confiança para uma média, uma proporção, a diferença entre duas médias e a diferença entre duas proporções.	N/D	N/D	N/D	N/D	N/D	N/D
Comando ANOVA unidirecional	Uma extensão de análise biomostrol, comparando duas médias de população. Testa se duas ou mais médias são iguais.	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	2-50
Comando ANOVA bidirecional	Executa uma análise bidirecional da variância. Isto é normalmente feito no contexto de um projeto experimental, em que há "fatores", cada qual definido em diversos níveis de "tratamento.	Não permitidos	Tempo real	Dados empilhados Até 16 milhões de casos O experimento deve ser balanceado	Sim	Não permitidos	2 variáveis de categoria, 1 variável de valor
Comando Teste de independência de qui-quadrado	Usa um teste de qui-quadrado para verificar se os atributos de uma linha e uma coluna em uma tabela de contingência são independentes estatisticamente.	Não	Tempo real (contanto que o tamanho da tabela não seja alterado)	N/D	N/D	N/D	N/D
Testes de normalidade							
Comando Teste de normalidade de qui-quadrado	Executa um teste de normalidade qui-quadrado para qualquer variável que você selecionar.	Permitidos no início, meio e fim dos dados	Tempo real parcial (posicionamento do bin não é alterado, mas ocupação e gráficos, sim)	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1
Comando Teste de Lilliefors	Proporciona um teste mais poderoso de normalidade que o teste de adequação de ajuste qui-quadrado (mais poderoso significa que há maior probabilidade de detectar não normalidade, se houver).	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-10

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de vars.
Comando Gráfico normal Q-Q	Cria um gráfico quantil-quantil (Q-Q) para uma variável selecionada. Proporciona um teste informal de normalidade.	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Não	Ignorados	1
Série temporal e previsão							
Comando Gráfico de série temporal	Cria um gráfico de série temporal de uma ou mais variáveis de séries temporais, todas na mesma tabela.	Permitidos no início, meio e fim dos dados	Tempo real	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1-100
Comando Autocorrelação	Calcula qualquer número de autocorrelações de uma variável de série temporal, indica quais (se houver) são significativamente não zero, e (opcionalmente) oferece um gráfico de barras (chamado correlograma) das autocorrelações.	Permitidos no início e fim dos dados	Tempo real	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1-10
Comando Teste de execuções para verificar aleatoriedade	Executa testes de execuções para verificar se uma variável (usualmente a variável em uma série temporal) é aleatória.	Permitidos no início e fim dos dados	Tempo real	Dados desempilhados Até 16 milhões de casos	Sim	Não permitidos	1 ou mais
Comando Previsão	Prevê dados de séries temporais usando método de médias móveis, suavização exponencial simples, método de suavização exponencial de Holt para a captura de tendências, e método de suavização exponencial de Winters para a captura da sazonalidade.	Permitidos no início dos dados apenas	Tempo real	Dados desempilhados Até 32.000 casos permitidos	Sim	Não permitidos	1 ou mais
Regressão e classificação							
Comando Regressão	Executa uma variedade de análises de regressão, como simples múltipla, passo a passo, prospectiva, retrospectiva e de bloco.	Permitidos no início, meio e fim dos dados	Estática	Dados desempilhados Até 16 milhões de casos	Sim	Não permitidos	1 dependente; 1-250 independent e

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de vars.
Comando Regressão logística	Executa uma análise de regressão logística em um conjunto de dados. Trata-se basicamente de um tipo não linear de análise de regressão em que a variável de resposta é binária: 0 ou 1.	Permitidos no início, meio e fim dos dados	Estática	Dados empilhados Até 16 milhões de casos	Sim	Não permitidos	1 dependente; 1-250 independent e
Comando Análise discriminante	Executa uma análise discriminante em um conjunto de dados. Deve haver uma variável de "categoria" que especifica em quais dois ou mais grupos cada observação está, mais uma ou mais variáveis explicativas que podem ser usadas para prever a associação ao grupo.	Permitidos no início, meio e fim dos dados	Estática	Dados empilhados Até 16 milhões de casos	Sim	Não permitidos em variáveis dependentes	1 dependente; 1-250 independent e
Controle de qualidade							
Comando Gráfico de Pareto	Produz um gráfico de Pareto que permite ver a importância relativa de dados categorizados.	Permitidos no início, meio e fim dos dados	Estática	Dados desempilhados	Sim	Ignorados	1 categoria, ou 1 valor e 1 categoria
Comando Gráficos X/R	Produz gráficos X e R, que permitem que você veja se um processo está em controle estatístico.	Não permitidos	Estática	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	2-25
Comando Gráfico P	Produz gráficos P, que permitem que você veja se um processo está em controle estatístico.	Não permitidos	Estática	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1 variável 1 variável de tamanho
Comando Gráfico C	Produz gráficos C, que permitem que você veja se um processo está em controle estatístico.	Não permitidos	Estática	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1
Comando Gráfico U	Produz gráficos U, que permitem que você veja se um processo está em controle estatístico.	Não permitidos	Estática	Dados desempilhados Até 32.000 casos permitidos	Não	Não permitidos	1
Testes não-paramétricos							
Comando Teste de sinal	Executa testes de hipótese da mediana de uma variável simples ou de uma mediana de diferenças de um par de variáveis	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrai) Exatamente 2 (análise biamostrai); análise de amostras pareadas)

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de var.
Comando Teste dos postos sinalizados de Wilcoxon	Executa testes de hipótese assim como o Teste de sinal, mas pressupõe que a distribuição da probabilidade é simétrica	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)
Comando Teste de Mann-Whitney	Executa testes de hipótese em duas amostras	Permitidos no início, meio e fim dos dados	Tempo real	Dados empilhados e desempilhados Até 16 milhões de casos	Sim	Ignorados	1-250 (análise uniamostrual) Exatamente 2 (análise biamostrual; análise de amostras pareadas)
Utilidades de dados							
Comando Empilhar	Pega um conjunto de dados com variáveis separadas para cada grupo em colunas separadas, e permite que sejam "empilhados" em duas colunas: uma coluna "valor" e uma coluna "categoria". Dependendo do tipo de análise, o conjunto de dados empilhado pode ser mais fácil de trabalhar que a versão desempilhado.	Sim – em qualquer parte da variável	Estática	Apenas dados desempilhados Até 65.535 casos permitidos	Não	N/D	1-100
Comando Desempilhar	Faz exatamente o oposto do procedimento Empilhar.	Sim – em qualquer parte da variável	N/D	Apenas dados empilhados Até 16 milhões de casos	Sim	N/D	1-32
Comando Fictício	Cria variáveis fictícias (0-1) baseadas em variáveis existentes.	Sim – em qualquer parte da variável	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	N/D	1
Comando Combinação	Cria uma nova variável de um par de variáveis numéricas, uma variável de categoria e uma numérica ou duas variáveis de categoria	Sim – em qualquer parte da variável	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	Não permitidos	2-32 do mesmo conjunto de dados

Procedimento	Descrição	Dados ausentes?	Relatórios em tempo real ou estát.	Requisitos de dados	Dados de múlt. interv.?	Dados inválidos	Núm. de var.
Comando Interação	Cria uma nova variável considerando o produto, soma, média, mín., máx. ou intervalo mín-máx de uma ou mais variáveis.	Sim – em qualquer parte da variável	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	N/D	2 variáveis de valor, ou 1 variável de valor e 1 variável de categoria, ou 2 variáveis de categoria
Comando Defasagem	Cria uma nova variável defasada com base em uma variável existente. Uma variável defasada é simplesmente uma versão da variável original, "forçada para baixo" por um número de linhas igual à defasagem.	Sim – em qualquer parte da variável	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	Ignorados	1
Comando Transformar	Aplica-se a qualquer uma de quatro transformações não lineares que você selecionar – logaritmo natural, quadrada, raiz quadrada ou recíproca – para criar uma nova variável	Sim – em qualquer parte da variável	Tempo real ou estática	Apenas dados desempilhados Até 16 milhões de casos	Sim	Ignorados	1-100
Comando Diferença	Cria qualquer número de variáveis de diferença de uma variável original.	Sim – em qualquer parte da variável	Tempo real	Apenas dados desempilhados Até 16 milhões de casos	Sim	N/D	1
Comando Amostra aleatória	Permite que você gere qualquer número de amostras aleatórias de um determinado conjunto de dados, em que uma amostragem se dá com ou sem substituição.	Sim – em qualquer parte da variável	Estática	Apenas dados empilhados Até 16 milhões de casos	Sim	Ignorados	1-32

Menu StatTools – Conjunto de dados

Comando Gerenciador de conjunto de dados

Permite definir conjuntos de dados e variáveis do StatTools, ou editar ou excluir conjuntos de dados e variáveis existentes

O comando **Gerenciador de conjunto de dados** do permite definir conjuntos de dados, variáveis. Uma vez que os conjuntos de dados e variáveis são definidos, podem ser analisados nos procedimentos StatTools. A caixa de diálogo Gerenciador de conjunto de dados permite adicionar ou remover conjuntos de dados, atribuir nomes aos mesmos, especificar o layout e atribuir nomes às variáveis de determinado conjunto de dados.

O que são conjunto de dados e variáveis?

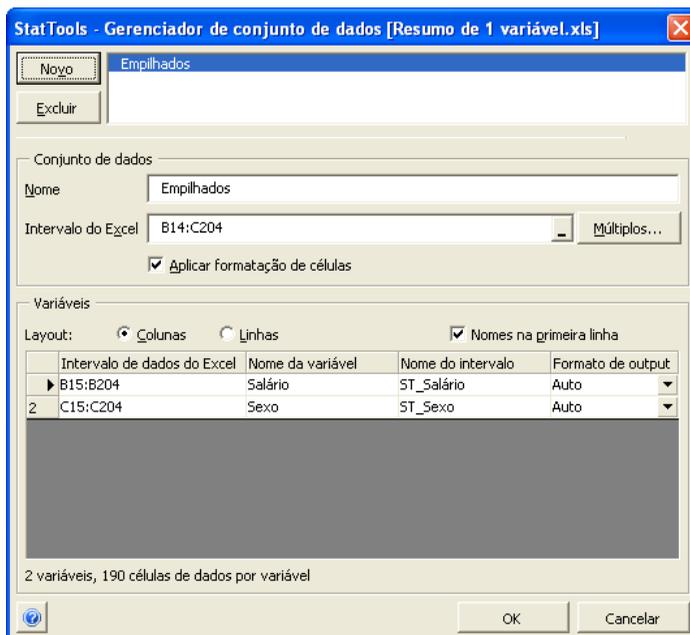
O StatTools é análogo à maioria dos pacotes de software estatísticos autônomos por ser estruturados com base em variáveis. Para a maioria das análises, é necessário que você trabalhe com um conjunto de dados ou com um conjunto de variáveis estatísticas, muitas vezes localizado em colunas contíguas, com os nomes das variáveis na primeira linha do conjunto de dados. Subsequentemente, essas variáveis predefinidas podem ser usadas nas suas análises estatísticas, sem necessidade de selecionar novamente os dados a serem analisados.

Cada variável contida no conjunto de dados tem um nome e um intervalo de células correspondente no Excel. O **Layout** selecionado especifica como as variáveis estão localizadas em um conjunto de dados. Um layout de variável típica é **Colunas**, com uma variável por coluna, porém, as variáveis também podem ser dispostas em **Linhas**. O conjunto de dados pode incluir vários blocos de células, o que permite colocar dados em diferentes planilhas de uma mesma pasta de trabalho.

Durante a definição de um conjunto de dados, o StatTools tenta identificar as variáveis contidas em um bloco de células circundante à seleção feita no Excel. Isso pode acelerar e facilitar a configuração de um conjunto de dados, com os nomes de variáveis na linha superior e variáveis dispostas por coluna.

Os comprimentos das colunas no seu conjunto de dados não precisam necessariamente ser iguais. Por exemplo, é possível ter duas variáveis, *Peso_Homens* e *Peso_Mulheres*, com diferentes números de observações. Contudo, para muitas análises, o StatTools tratará as células em branco nas colunas mais curtas como dados ausentes.

Caixa de diálogo Gerenciador de conjunto de dados



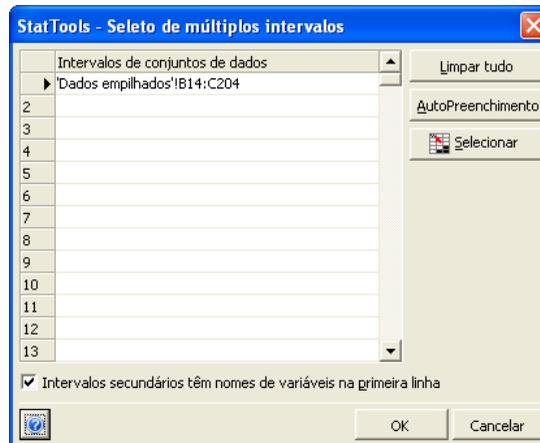
As opções do **Conjunto de Dados** na caixa de diálogo Gerenciador de conjunto de dados são:

- **Novo** e **Excluir** – para adicionar um novo conjunto de dados ou excluir um conjunto existente.
- **Nome** – especifica o nome do conjunto de dados.
- **Intervalo do Excel** – especifica o intervalo do Excel correspondente ao conjunto de dados. Se vários intervalos de células tiverem sido designados a um conjunto de dados, este item será antecedido pelo rótulo **Múltiplos**.
- **Aplicar formatação de células** – acrescenta uma grade e cores que identificam os conjuntos de dados.
- **Múltiplos** – ao clicar no botão **Múltiplos**, na caixa de diálogo Gerenciador de conjunto de dados, aparece a caixa de diálogo **Seletor de múltiplos intervalos**, que permite especificar intervalos de células individuais que formam o conjunto de dados com múltiplos intervalos de células.

Conjuntos de dados de intervalos múltiplos

O StatTools permite designar múltiplos intervalos de células a um conjunto de dados. Um conjunto de dados com múltiplos intervalos pode ser usado quando:

- 1) Cada variável do conjunto de dados tem mais de 65.536 pontos de dados (no do Excel 2003 ou versão anterior), exigindo que o conjunto de dados seja distribuído em diversas planilhas de uma mesma pasta de trabalho.
- 2) Os dados referentes a uma variável encontram-se em vários blocos dispersos em várias planilhas de uma pasta de trabalho.

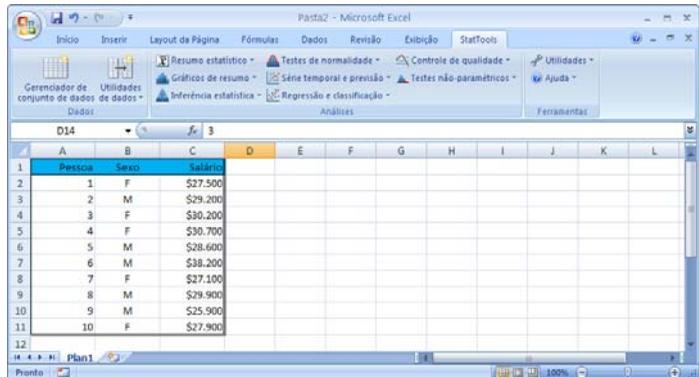


As opções da caixa de diálogo Seletor de múltiplos intervalos incluem:

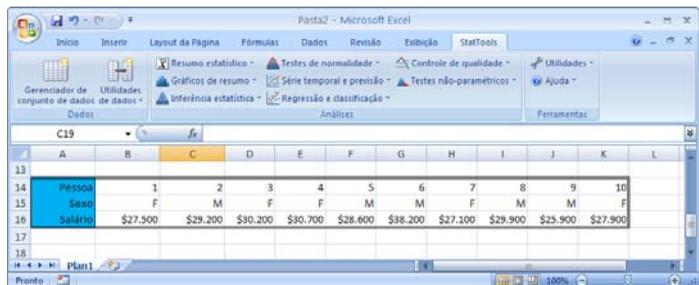
- **Limpar tudo** – apaga todos os intervalos especificados.
- **AutoPreenchimento** – aplica o primeiro intervalo especificado (na linha 1) em todas as planilhas visíveis da pasta de trabalho ativa, e insere as referências de SheetName!CellRange na grade
- **Selecionar** – exibe um seletor para destacar um bloco de células a ser usado como um intervalo do conjunto de dados.
- **Intervalos secundários têm nomes de variáveis na primeira coluna (linha)** – conjuntos de dados com múltiplos intervalos podem ter nomes de variáveis como títulos de cada coluna (ou linha, dependendo do layout de variável selecionado) em cada intervalo listado na caixa de diálogo; ou nomes de variáveis como identificação das colunas ou linhas apenas no primeiro intervalo selecionado. O primeiro intervalo selecionado é o que foi inserido na linha 1 da caixa de diálogo Seletor de múltiplos intervalos.

As opções de **Variáveis** na caixa de diálogo Gerenciador de conjunto de dados são:

- **Layout** – especifica como as variáveis são estruturadas no intervalo do Excel que contém o conjunto de dados. As opções de Layout incluem:
 - **Colunas**. Este é o layout típico em que cada coluna no intervalo do Excel do conjunto de dados tem os dados de uma variável. Muitas vezes os nomes de variáveis serão inseridos no topo de cada coluna.



- **Linhas**. Com este layout, cada linha de conjunto de dados contém os dados de uma variável. Este layout muitas vezes é utilizado para os dados de uma série temporal no Excel.

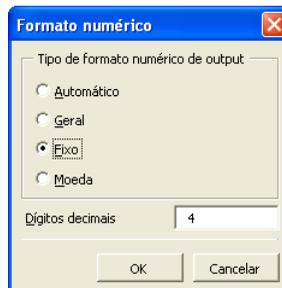


- **Nomes na primeira coluna (ou linha)** – selecione esta opção quando você incluiu os nomes das variáveis de um conjunto de dados no topo das colunas (ou nas células mais à esquerda quando o layout de variável for **Linhas**).

Opções de variáveis

Cada linha da grade da caixa de diálogo Gerenciador de conjunto de dados lista as variáveis de um conjunto de dados, incluindo o nome de cada variável, o intervalo do Excel que contém os pontos de dado de uma variável, e o nome do intervalo do Excel usado para identificar os dados da variável nas fórmulas do Excel.

- **Nome do intervalo do Excel** – O nome do intervalo mostrado será usado nas fórmulas do Excel criadas nos relatórios e gráficos do StatTools. Essas fórmulas permitem que os seus relatórios sejam “em tempo real” – quer dizer, sejam automaticamente atualizados quando os dados de uma variável são alterados. Ter nomes de intervalo significativos ajuda a tornar as fórmulas mais compreensíveis.
- **Formato de output** – especifica o formato de valores mostrados para uma variável nos relatórios das análises do StatTools. A entrada **Auto** especifica que o StatTools selecionará o “melhor” formato com base na formatação numérica aplicada às células contendo os valores das variáveis no Excel. Clicar na seta ao lado do campo de Formato de output permite selecionar alternativamente um formato específico a ser utilizado:



Geral equivale ao formato numérico Geral do Excel. **Fixo** exibe a precisão adotada com o número de **Dígitos decimais** especificado. **Moeda** equivale ao formato numérico Moeda do Excel, e exibe a precisão adotada com o número de **Dígitos decimais** especificado.

Nota: o formato de output desejado também pode ser inserido diretamente na caixa de diálogo Gerenciador de conjunto de dados, com a notação *selectedFormat(#decimalDigits)*, como *Currency(4)*.

Conjunto de dados e capacidades das variáveis

O StatTools tem capacidade para processar, em uma mesma sessão:

- Até 256 conjuntos de dados contidos em uma mesma pasta de trabalho.
- Até 256 variáveis por conjunto de dados. Todos os dados usados em um conjunto de dados precisam estar na mesma pasta de trabalho.
- Até 16.777.216 pontos de dados por variável.

A capacidade efetiva no que se refere aos dados pode ser inferior às indicadas acima, dependendo da configuração do sistema e da versão do Excel que está sendo usada. Análises específicas do StatTools podem ter limitações diferentes. As limitações de memória do Excel também podem afetar a capacidade em termos do volume de dados.

Nota: a caixa de diálogo Gerenciador de conjunto de dados relaciona todos os conjuntos de dados e variáveis contidos na pasta de trabalho ativa (ou seja, a pasta de trabalho indicada na legenda da caixa de diálogo). Nota: para incluir na lista conjuntos de dados contidos em outras pasta de trabalho, ative a pasta de trabalho desejada no Excel e exiba a caixa de diálogo Gerenciador de conjunto de dados.

Menu Utilidades de dados

Comando Empilhar

Converte um conjunto de variáveis para do formato desempilhado para o formato empilhado

O comando Empilhar permite a conversão de dados no formato “desempilhado”, em que o conjunto de dados inclui pelo menos duas variáveis de **valor**; para o formato “empilhado”, em que o conjunto de dados inclui uma variável de **categoria** e uma variável de valor. Por exemplo, se compararmos as rendas médias familiares de bairros diferentes, na forma desempilhada haveria uma variável *Renda* (ou coluna) para cada *Bairro*. Essas colunas não precisam ser do mesmo comprimento, ou seja, cada bairro poderia ter um tamanho de amostra diferente. No formato empilhado haveria uma variável de valor *Renda* e uma variável de categoria *Bairro* indicando o bairro em que está cada família.

Basicamente, este procedimento permite o “empilhamento” das suas variáveis em duas colunas: uma coluna de valor, *Renda*, e uma coluna de categoria, *Bairro*. Dependendo do tipo de análise, o conjunto de dados empilhados pode ser mais fácil de trabalhar que a versão de dados desempilhados.

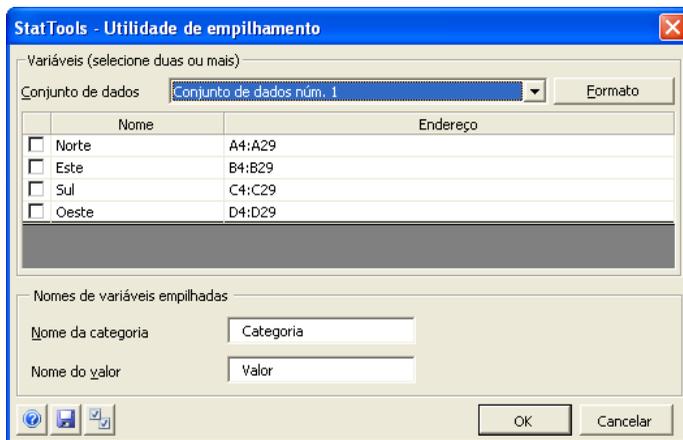
Variáveis empilhadas e desempilhadas

The screenshot shows a Microsoft Excel spreadsheet with two data tables. The first table, titled 'Renda (dados desempilhados)', is a wide format with columns for 'Bairro' (North, East, South, West) and rows of income values. The second table, titled 'Renda (dados empilhados)', is a long format with columns for 'Bairro' and 'Renda', where each row represents a single data point from the first table.

Renda (dados desempilhados):					Renda (dados empilhados):	
Bairro	Este	Sul	Oeste		Bairro	Renda
Este	546.809	543.108	573.115	528.650	Este	543.108
Este	545.812	54.524	535.747	55.046	Este	54.524
Este	549.132	555.628	5129.974	52.702	Este	555.628
Este	535.459	546.488	534.520	5175.933	Este	546.488
Este	539.503	580.831	5113.469	5126.656	Este	580.831
Este	59.547	558.703	5238.111	5129.321	Este	558.703
Este	549.880	545.000	505.385	5107.794	Este	545.000
Este	541.866	565.527	5194.381	589.853	Este	565.527
Este	53.129	558.842	588.441	584.903	Este	558.842
Este	523.115	512.067	5139.580	572.781	Este	512.067
Este	547.886	526.313	5176.853	525.464	Este	526.313
Este	522.556	538.457	5234.325	5125.344	Este	538.457
Este	515.975	58.241		582.216	Este	58.241
Este	5716	552.476		581.951	Este	552.476
Este	527.065	588.298		582.222	Este	588.298
Este	517.569	592.347		5147.837	Este	592.347
Este	52.020	559.968		5147.812	Este	559.968
Este	593.883	569.574		5175.167	Este	569.574
Este	556.984	53.055		5159.937	Este	53.055
Este	529.041	588.375		5187.547	Este	588.375
Este		594.943		5117.448	Este	594.943
Este		554.889		5115.498	Este	554.889
Este		586.583		597.200	Este	586.583
Este		556.665			Este	556.665
Este		528.078			Este	528.078
Este		571.380			Este	571.380
Oeste				528.650	Oeste	528.650
Oeste				55.046	Oeste	55.046
Oeste				52.702	Oeste	52.702
Oeste				5175.933	Oeste	5175.933
Oeste				5126.656	Oeste	5126.656
Oeste				5129.321	Oeste	5129.321
Oeste				5107.794	Oeste	5107.794

Caixa de diálogo Utilidade de empilhamento

As variáveis são empilhadas usando a caixa de diálogo **Utilidade de empilhamento**:



No mínimo duas variáveis devem ser selecionadas para o empilhamento. O conjunto de dados selecionado é sempre tratado inicialmente como dados desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções na caixa de diálogo Utilidade de empilhamento de variável incluem:

- **Nomes de variáveis empilhadas** – especifica o nome das variáveis de **categoria** e **valor** que formarão o conjunto de dados empilhados de duas variáveis. Esses nomes aparecerão no topo das colunas das variáveis de categoria e de valor.

Quando você clica em OK, as variáveis são empilhadas e um novo conjunto de dados é criado para os dados empilhados.

O que são variáveis de categoria e valor?

As variáveis **Categoria** e **Valor** são necessárias para um conjunto de dados empilhados. A variável **Categoria** (às vezes chamada de variável de “código”) é simplesmente um identificador descritivo de um conjunto de variáveis de valor. A variável **Categoria** é muitas vezes um rótulo de texto. As variáveis de valor (às vezes chamada de variáveis de “medição”), por outro lado, são variáveis numéricas padrão, que podem ser analisadas em procedimentos estatísticos.

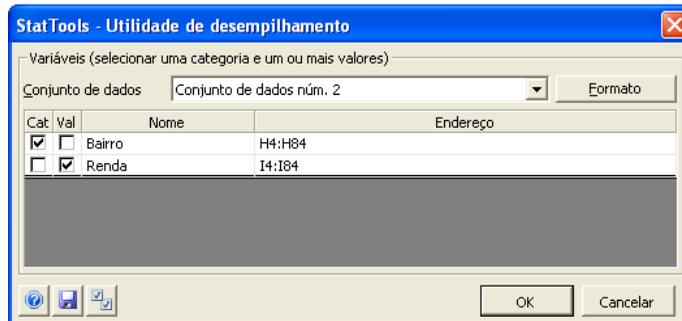
Comando Desempilhar

Converte um conjunto de variáveis do formato empilhado para o formato desempilhado

O comando Desempilhar faz exatamente o oposto do comando Empilhar. Por exemplo, se você começar com uma variável de categoria *Sexo* e uma variável de valor *Peso*, este comando as desempilhará em colunas separadas *Peso_Homens* e *Peso_Mulheres*.

As variáveis são desempilhadas usando a caixa de diálogo **Utilidade de desempilhamento**:

**Caixa de diálogo
Utilidade de
desempilhamento**



O conjunto de dados selecionado é sempre tratado inicialmente como dados empilhados. No mínimo duas variáveis devem ser selecionadas para o desempilhamento. Uma dessas variáveis é identificada como a variável de **Categoria** (marcando-se **Cat**) e uma ou mais variáveis são identificadas como variáveis de **Valor** (marcando-se **Val**). As variáveis podem pertencer a conjuntos de dados diferentes.

Quando você clica em OK, as variáveis são desempilhadas e um ou mais novos conjuntos de dados são criados para os dados desempilhados.

Comando Transformar

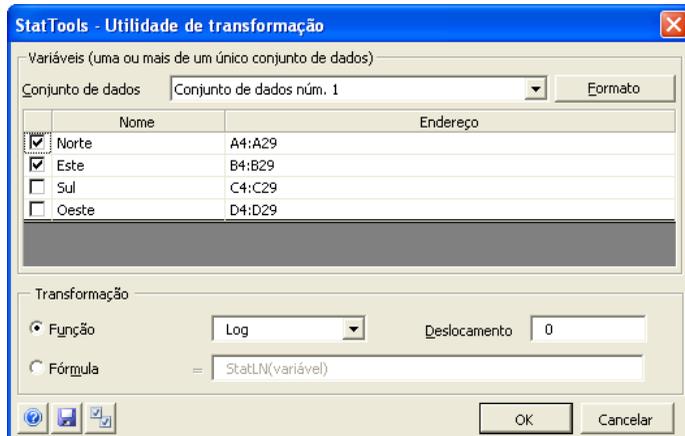
Transforma uma ou mais variáveis em novas variáveis e valores com base em uma função de transformação inserida

O comando Transformar permite que você transforme qualquer variável com uma entre quatro transformações possíveis: logaritmo natural, quadrado, raiz quadrada ou recíproca. Além disso, permite que você insira uma fórmula que será usada para calcular um valor de variável transformado.

Se houver valores ausentes na variável na qual a variável transformada se baseará, haverá valores ausentes correspondentes na variável transformada.

As variáveis são empilhadas usando a caixa de diálogo **Utilidade de transformação**:

**Caixa de diálogo
Utilidade de
transformação**



O conjunto de dados selecionado é sempre de dados desempilhados. Só podem ser transformadas variáveis de um conjunto de dados por vez.

As opções na caixa de diálogo Utilidade de transformação incluem:

- **Função da transformação** – ou a operação matemática que será executada em cada valor das variáveis selecionadas na geração do novo valor transformado. As funções de transformação incorporadas incluem logaritmo natural, quadrado, raiz quadrada ou recíproca. Um **Fórmula** personalizada também pode ser definida, para calcular um novo valor de variável com base em uma expressão matemática, como:

$$(Variable*1,5)^2$$

Observe que na equação a palavra-chave “Variable” é usada como placeholder, ou seja, é substituída pelo valor real da variável a ser transformada.

Conjunto de dados com novas variáveis transformadas

	A	B	C	D	E	F
1	Renda (dados desempilhados):					
2						
3	Norte	Este	Sul	Oeste	Log(Norte)	Log(Este)
4	\$46.809	\$43.108	\$73.115	\$28.650	10,75383077	10,67146387
5	\$45.812	\$4.524	\$35.747	\$5.046	10,73230134	8,417151837
6	\$49.132	\$55.628	\$129.974	\$2.702	10,80226583	10,92644195
7	\$15.459	\$46.488	\$34.920	\$175.693	9,645946637	10,74694949
8	\$39.503	\$80.831	\$113.469	\$126.656	10,5841319	11,30011583
9	\$9.547	\$58.703	\$238.111	\$129.321	9,163982248	10,98024611
10	\$49.880	\$45.000	\$55.385	\$107.794	10,8173754	10,71441777
11	\$41.866	\$65.527	\$104.381	\$89.853	10,64222932	11,09021755
12	\$3.129	\$58.842	\$88.441	\$84.903	8,048468744	10,98261116
13	\$23.115	\$12.067	\$139.580	\$72.781	10,04823704	9,398229733
14	\$47.886	\$26.313	\$176.853	\$25.464	10,77657847	10,17781839
15	\$22.550	\$38.457	\$234.325	\$125.144	10,02349034	10,55729601
16	\$15.975	\$8.241		\$82.216	9,678780279	9,016876975
17	\$716	\$52.476		\$81.951	6,573680167	10,8681112
18	\$27.065	\$88.298		\$82.222	10,20599666	11,38847274
19	\$17.569	\$92.347		\$147.837	9,773891264	11,4333085
20	\$2.020	\$59.968		\$147.412	7,61085279	11,00156637
21	\$53.883	\$69.574		\$175.167	10,89457031	11,15014621
22	\$56.984	\$3.055		\$159.937	10,95052581	8,024534872
23	\$29.041	\$28.375		\$187.547	10,2764639	10,25326375
24		\$94.943		\$137.448		11,46103199
25		\$54.889		\$115.696		10,91306824
26		\$86.583		\$97.260		11,36885877
27		\$56.665				10,94491202
28		\$28.078				10,24274163
29		\$71.380				11,175773

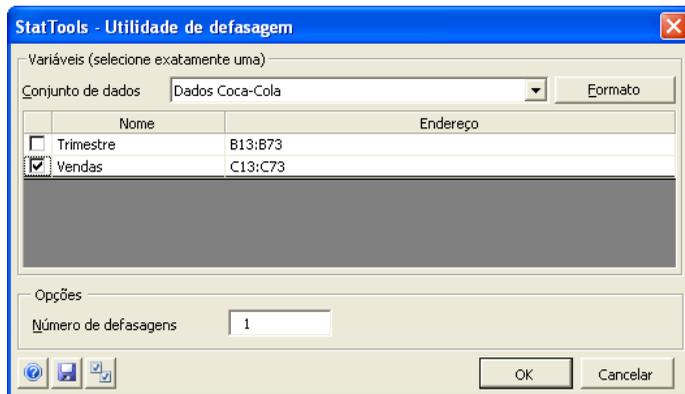
Comando Defasagem

Cria uma nova variável defasada com base em uma variável existente

O comando Defasagem permite que você crie uma nova variável defasada com base em uma variável existente. Uma variável defasada é simplesmente uma versão da variável original, “forçada para baixo” por um número de linhas igual à defasagem. Por exemplo, a versão defasagem 3 das vendas em novembro de 1998 são as vendas de três meses antes, de agosto de 1998.

Caixa de diálogo Utilidade de defasagem

As variáveis são defasadas usando a caixa de diálogo **Utilidade de defasagem**:



O conjunto de dados selecionado é sempre de dados desempilhados. Apenas uma variável por vez pode ser defasada.

As opções na caixa de diálogo Utilidade de defasagem de variável incluem:

- **Número de defasagens** – ou o número de períodos para a defasagem dos valores na criação de novas variáveis. Uma nova variável é criada para cada uma das defasagens, até o número inserido de defasagens.

Variável defasada à direita do conjunto de dados de origem

The screenshot shows the Microsoft Excel interface with the StatTools ribbon selected. The ribbon includes options for statistical analysis such as 'Resumo estatístico', 'Testes de normalidade', 'Controle de qualidade', 'Utilidades', 'Gráficos de resumo', 'Série temporal e previsão', 'Testes não-paramétricos', 'Ajuda', 'Inferência estatística', and 'Regressão e classificação'. The active worksheet is titled 'Previsão.xls' and contains a table with three columns: 'Trimestre', 'Vendas', and 'Defasagem 1(Vendas)'. The data spans from 1986 to 1993, with each year having four quarterly entries (T1 to T4). The 'Defasagem 1(Vendas)' column shows values that are one quarter ahead of the 'Vendas' column.

Trimestre	Vendas	Defasagem 1(Vendas)
T1-86	1734,83	
T2-86	2244,96	1734,83
T3-86	2533,8	2244,96
T4-86	2154,96	2533,8
T1-87	1547,82	2154,96
T2-87	2104,41	1547,82
T3-87	2014,36	2104,41
T4-87	1991,75	2014,36
T1-88	1869,05	1991,75
T2-88	2313,63	1869,05
T3-88	2128,32	2313,63
T4-88	2026,83	2128,32
T1-89	1910,6	2026,83
T2-89	2331,16	1910,6
T3-89	2206,55	2331,16
T4-89	2173,97	2206,55
T1-90	2148,28	2173,97
T2-90	2739,31	2148,28
T3-90	2792,75	2739,31
T4-90	2556,01	2792,75
T1-91	2480,97	2556,01
T2-91	3039,52	2480,97
T3-91	3172,12	3039,52
T4-91	2879	3172,12
T1-92	2772	2879
T2-92	3550	2772
T3-92	3508	3550
T4-92	3243,86	3508
T1-93	3056	3243,86
T2-93	3899	3056

Comando Diferença

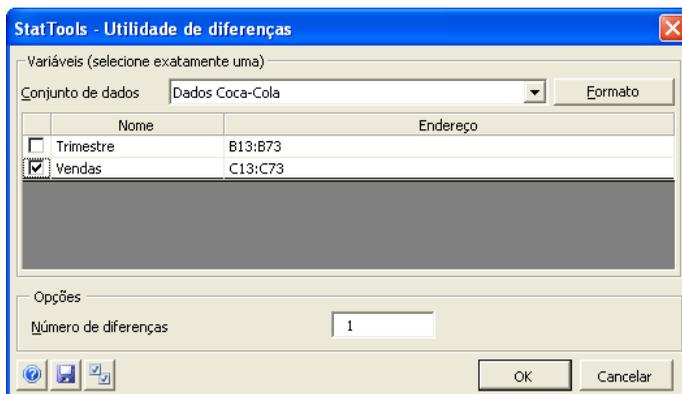
Cria qualquer número de variáveis de diferença de uma variável original

O comando Diferença permite que você crie qualquer número de variáveis de diferença a partir de uma variável original. É usado principalmente para variáveis de série temporal. Uma variável a ser diferenciada é selecionada juntamente com o número de diferenças (normalmente 1 ou 2). O procedimento em seguida cria essas muitas novas variáveis de diferença. Cada variável de diferença contém diferenças da variável selecionada. Por exemplo (para dados mensais), o valor de diferença de março de 1997 é o valor original de março de 1997 menos o valor original de fevereiro de 1997. Da mesma forma, a segunda variável de diferença (se solicitado) contém as diferenças das primeiras diferenças.

A diferenciação é muitas vezes usada em uma análise de série temporal quando a variável original não é “estacionária” ao longo do tempo. Por exemplo, uma série temporal com uma tendência ascendente não é estacionária. A diferenciação muitas vezes atinge o ponto estacionário. Às vezes uma segunda diferenciação é útil, mas é menos comum. A terceira diferenciação (ou diferenciação além da terceira) quase nunca é necessário.

As variáveis de diferença são criadas com a caixa de diálogo **Utilidade de diferença**:

Caixa de diálogo
Utilidade de
diferença



O conjunto de dados selecionado é sempre de dados desempilhados. Apenas uma variável por vez pode ser usada para a criação de variáveis de diferença.

As opções na caixa de diálogo Utilidade de diferença incluem:

- **Número de diferenças** – ou o número de diferenças que serão criadas.

Conjunto de dados com variável de diferença

The screenshot shows the Microsoft Excel interface with the StatTools add-in. The 'Diferença' dialog box is open, and a data table is visible below it. The table has three columns: 'Trimestre', 'Vendas', and 'Diferença 1(Vendas)'. The data is as follows:

Trimestre	Vendas	Diferença 1(Vendas)
T1-86	1734,83	
T2-86	2244,96	510,13
T3-86	2533,8	288,84
T4-86	2154,96	-378,84
T1-87	1547,82	-607,14
T2-87	2104,41	556,59
T3-87	2014,36	-90,05
T4-87	1991,75	-22,61
T1-88	1869,05	-122,7
T2-88	2313,63	444,58
T3-88	2128,32	-185,31
T4-88	2026,83	-101,49
T1-89	1910,6	-116,23
T2-89	2331,16	420,56
T3-89	2206,55	-124,61
T4-89	2173,97	-32,58
T1-90	2148,28	-25,69
T2-90	2739,31	591,03
T3-90	2792,75	53,44
T4-90	2556,01	-236,74
T1-91	2480,97	-75,04
T2-91	3039,52	558,55
T3-91	3172,12	132,6
T4-91	2879	-293,12
T1-92	2772	-107
T2-92	3550	778
T3-92	3508	-42
T4-92	3243,86	-264,14
T1-93	3056	-187,86
T2-93	3899	843

Comando Interação

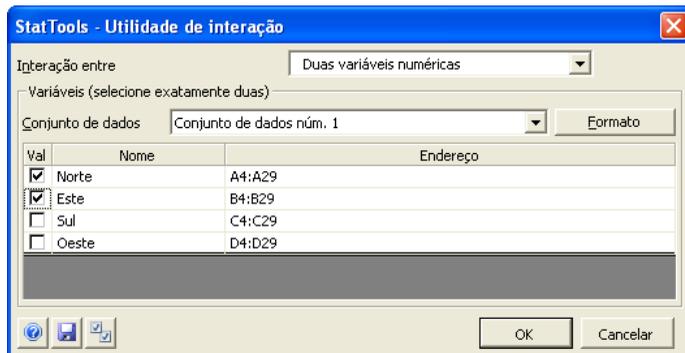
Cria uma variável de interação a partir de uma ou mais variáveis originais

O comando Interação permite a criação de uma variável de interação de uma ou mais variáveis originais. As variáveis de interação podem ser criadas com duas variáveis numéricas, uma numérica e uma de categoria ou duas de categoria.

Se ambas as variáveis forem numéricas (não de categoria), o seu produto é criado. Se uma variável for numérica e a outra de categoria, são criados os produtos da variável numérica com cada valor fictício correspondente às categorias da variável de categoria. Por fim, se ambas as variáveis forem de categoria, são criados os produtos de todos os pares fictícios a partir das duas variáveis de categoria.

As variáveis de interação são criadas com a caixa de diálogo **Utilidade de interação**:

Caixa de diálogo
Utilidade de
interação



O conjunto de dados selecionado é sempre de dados desempilhados. Uma ou mais variáveis por vez podem ser usadas para a criação de variáveis de interação.

As opções na caixa de diálogo Utilidade de interação incluem:

- **Interação entre** – seleciona o tipo de cada variável a ser selecionada; **Duas variáveis numéricas**, **Uma variável numérica e uma categórica** ou **Duas variáveis categóricas**.

Como uma variável de interação é criada?

Uma variável de interação é formada por duas variáveis selecionadas na caixa de diálogo. Há três opções básicas para essas duas variáveis. Na primeira, ambas podem ser variáveis numéricas de “medição”. Então, a variável de interação é o produto delas. Na segunda, uma variável pode ser numérica de “medição” e a outra, de categoria. Então, o StatTools cria internamente variáveis fictícias para cada categoria da variável de categoria e multiplica *cada* variável fictícia pela variável numérica. Na terceira, ambas as variáveis podem ser de categoria. Então, o StatTools cria internamente variáveis fictícias para cada categoria da variável de categoria e multiplica *cada* variável fictícia da primeira por cada fictícia da segunda. Por exemplo, se duas variáveis de categoria têm 2 e 5 categorias, respectivamente, o StatTools criará $2 \times 5 = 10$ variáveis de interação.

Conjunto de dados com variável de interação

	A	B	C	D	E	F	G	H	I
1	Renda (dados desempilhados):								
2									
3	Norte	Este	Sul	Oeste	Interação(Norte,Este)				
4	\$46.809	\$43.108	\$73.115	\$28.650	2017842372				
5	\$45.812	\$4.524	\$35.747	\$5.046	207253488				
6	\$49.132	\$55.628	\$129.974	\$2.702	2733114896				
7	\$15.459	\$46.488	\$34.920	\$175.693	718657992				
8	\$39.503	\$80.831	\$113.469	\$126.656	3193066993				
9	\$9.547	\$58.703	\$238.111	\$129.321	560437541				
10	\$49.880	\$45.000	\$55.385	\$107.794	2244600000				
11	\$41.866	\$65.527	\$104.381	\$89.853	2743353382				
12	\$3.129	\$58.842	\$88.441	\$84.903	184116618				
13	\$23.115	\$12.067	\$139.580	\$72.781	278928705				
14	\$47.886	\$26.313	\$176.853	\$25.464	1260024318				
15	\$22.550	\$38.457	\$234.325	\$125.144	867205350				
16	\$15.975	\$8.241		\$82.216	131649975				
17	\$716	\$52.476		\$81.951	37572816				
18	\$27.065	\$88.298		\$82.222	2389785370				
19	\$17.569	\$92.347		\$147.837	1622444443				
20	\$2.020	\$59.968		\$147.412	121135360				
21	\$53.883	\$69.574		\$175.167	3748855842				
22	\$56.984	\$3.055		\$159.937	174086120				
23	\$29.041	\$28.375		\$187.547	824038375				
24		\$94.943		\$137.448					
25		\$54.889		\$115.696					
26		\$86.583		\$97.260					
27		\$56.665							
28		\$28.078							
29		\$71.380							

Comando Combinação

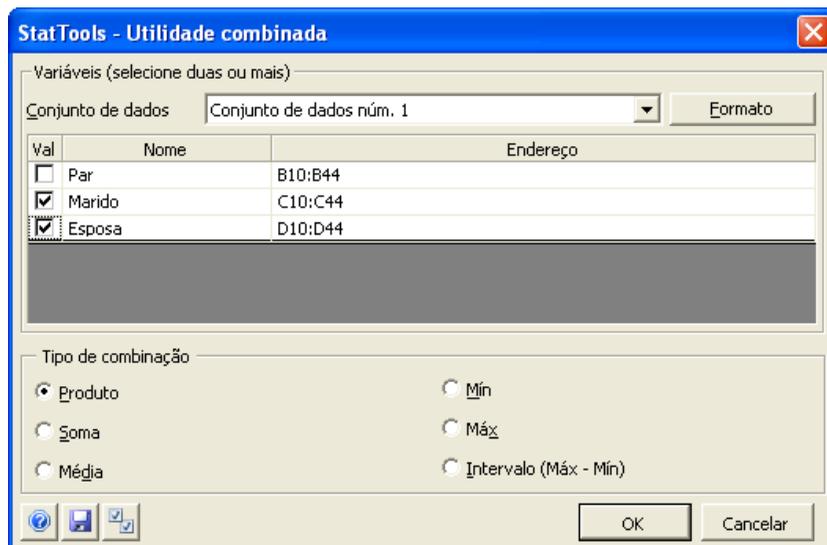
Cria uma variável de combinação a partir de uma ou mais variáveis originais

O comando Combinação permite a criação de uma variável de combinação de uma ou mais variáveis originais. Cria uma nova variável de combinação considerando o produto, soma, média, mín., máx. ou intervalo mín-máx de uma ou mais variáveis.

Caixa de diálogo Utilidade de combinação

As variáveis de combinação são criadas com a caixa de diálogo

Utilidade de combinação:



O conjunto de dados selecionado é sempre de dados desempilhados. Uma ou mais variáveis por vez podem ser usadas para a criação de variáveis de combinação.

As opções na caixa de diálogo Utilidade de combinação incluem:

- **Tipo de combinação** – ou operação matemática a ser executada nas variáveis selecionadas ao criar a variável de interação. Ela pode ser um produto, soma, média, mín., máx. ou intervalo mín-máx.

Conjunto de dados com variável de combinação

The screenshot shows a Microsoft Excel spreadsheet with the following data:

	Par	Marido	Esposa	Produto(Marido,Esposa)
10	1	6	3	18
11	2	7	8	56
12	3	8	5	40
13	4	6	4	24
14	5	8	5	40
15	6	7	6	42
16	7	8	5	40
17	8	6	7	42
18	9	7	8	56
19	10	6	5	30
20	11	5	3	15
21	12	8	4	32
22	13	7	5	35
23	14	7	8	56
24	15	7	5	35
25	16	6	6	36
26	17	5	5	25
27	18	6	4	24
28	19	9	5	45
29	20	7	10	70
30	21	9	9	81
31	22	6	6	36
32	23	6	5	30
33	24	6	4	24
34	25	8	5	40
35	26	9	5	45
36	27	7	7	49
37	28	8	5	40

StatTools menu options: Gerenciador de conjunto de dados de dados, Utilidades de dados, Resumo estatístico, Gráficos de resumo, Inferência estatística, Testes de normalidade, Série temporal e previsão, Regressão e classificação, Controle de qualidade, Testes não-paramétricos, Ferramentas, Ajuda.

Status bar: Média: 39.42857143, Contagem: 36, Soma: 1380, 100%

Comando Fictício

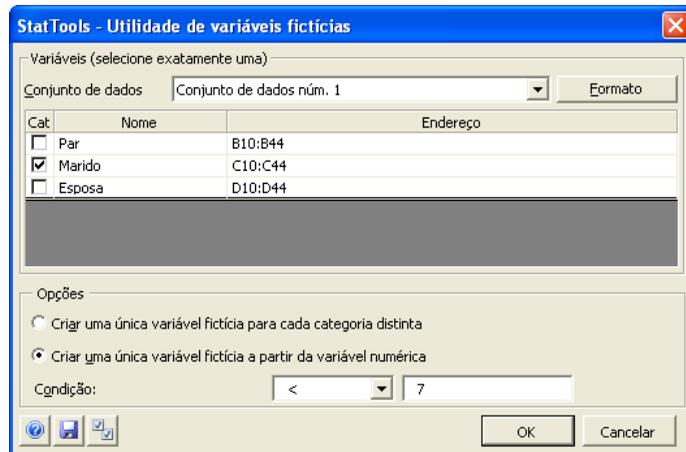
Cria variáveis fictícias (0-1) baseadas em variáveis existentes

O comando Fictício cria variáveis fictícias (0-1) baseadas em variáveis existentes. Há duas opções:

- 1) Você pode criar uma variável fictícia para cada categoria de uma variável de categoria. Por exemplo, se houver uma variável de categoria *Origem* (para automóveis) com categorias *EUA*, *Europa* e *Ásia*, três variáveis fictícias são criadas com os nomes *Origem_EUA*, *Origem_Europa* e *Origem_Ásia*.
- 2) Você pode criar uma única variável fictícia a partir de uma variável numérica, com base em um valor de corte selecionado. Por exemplo, se houver uma variável *Peso*, é possível criar variáveis fictícias para a condição $Peso \leq 160$. Neste caso, uma nova variável seria adicionada com o valor 0 quando $Peso > 160$ e 1 quando $Peso \leq 160$.

Caixa de diálogo Utilidade de variáveis fictícias

As variáveis fictícias são criadas com a caixa de diálogo **Utilidade de variáveis fictícias**:



O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. Uma variável por vez pode ser usada para a criação de variáveis fictícias.

As opções na caixa de diálogo Utilidade de variáveis fictícias incluem:

- **Opções** – seleciona a forma na qual as variáveis fictícias podem ser criadas: 1) uma variável fictícia para cada categoria de variável de categoria ou 2) uma única variável fictícia a partir de uma variável numérica. O **Valor de corte** especifica o corte a ser usado na atribuição de uma variável numérica a uma variável fictícia 0-1.

Conjunto de dados com variável fictícia

The screenshot shows an Excel spreadsheet with the following data:

Par	Marido	Esposa	Marido < 7	
1	6	3	1	
2	7	8	0	
3	8	5	0	
4	6	4	1	
5	8	5	0	
6	7	6	0	
7	8	5	0	
8	6	7	1	
9	7	8	0	
10	6	5	1	
11	11	3	1	
12	12	4	0	
13	13	7	5	0
14	14	7	8	0
15	15	7	5	0
16	16	6	6	1
17	17	5	5	1
18	18	6	4	1
19	19	9	5	0
20	20	7	10	0
21	21	9	9	0
22	22	6	6	1
23	23	6	5	1
24	24	6	4	1
25	25	8	5	0
26	26	9	5	0
27	27	7	7	0
28	28	8	5	0

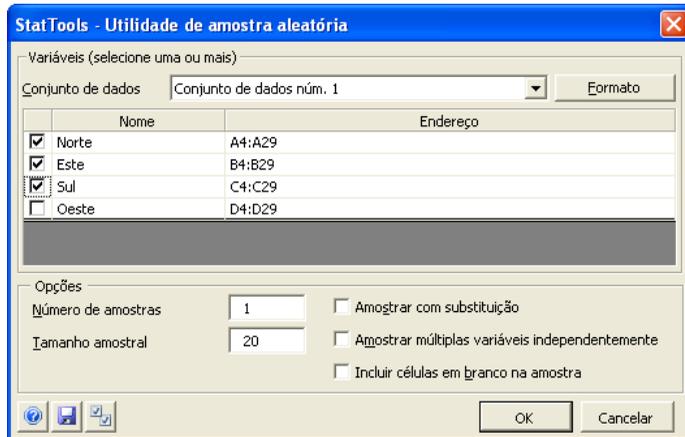
Comando Amostra aleatória

Gera qualquer número de amostras aleatórias com base em variáveis selecionadas

O comando Amostra aleatória permite a geração de qualquer número de amostras aleatórias a partir de uma ou mais variáveis selecionadas. Você especifica o número de amostras e o tamanho de cada amostra, e o StatTools gera as amostras a partir das variáveis selecionadas. Diversas variáveis podem ser amostradas de forma independente ou dependente, e a amostragem pode ser feita com ou sem substituição.

As amostras aleatórias são geradas na caixa de diálogo **Utilidade de amostra aleatória**:

**Caixa de diálogo
Utilidade de
amostra aleatória**

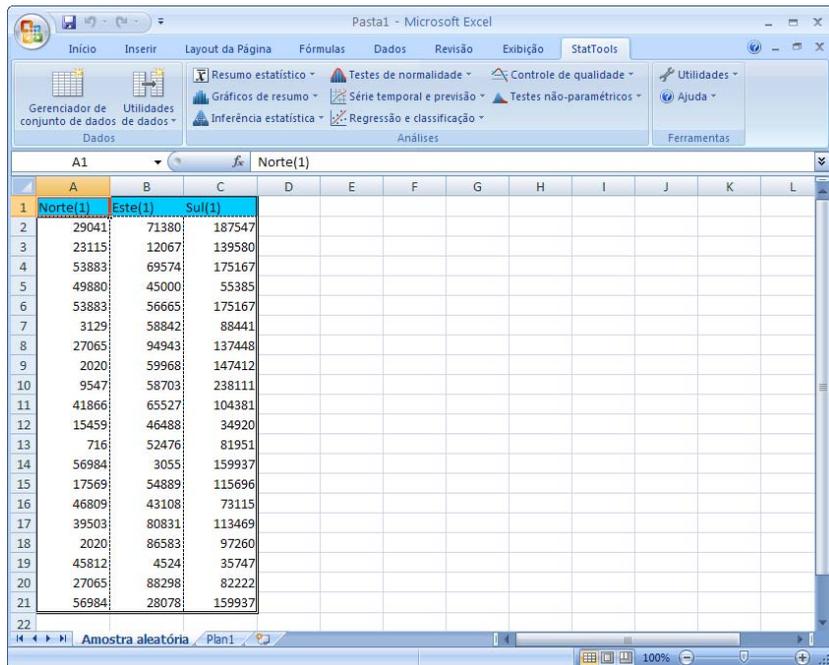


O conjunto de dados selecionado é sempre de dados desempilhados. Uma ou mais variáveis por vez podem ser usadas para a criação de amostras aleatórias.

As opções na caixa de diálogo Utilidade de amostra aleatória incluem:

- **Número de amostras e Tamanho da amostra.** O número selecionado de amostras será gerado para cada variável selecionada, e cada amostra terá um número de elementos igual ao tamanho da amostra.
- **Amostrar com substituição.** Indica que um valor “volta” à população original após ser amostrado, permitindo que seja amostrado novamente. Do contrário, se **Amostragem com substituição** não for selecionada (ou seja, se for utilizada a amostragem sem substituição), um valor não é retornado e não poderá ser amostrado novamente.
- **Amostrar múltiplas variáveis independentemente.** Opção que determina um sorteio independente usado para cada valor amostrado para cada variável. Do contrário, para cada valor amostrado, o mesmo índice amostrado (um número entre 1 e # de valores da variável) será usado para todas as variáveis.

Amostras aleatórias geradas



	A	B	C	D	E	F	G	H	I	J	K	L
1	Norte(1)	Este(1)	Sul(1)									
2	29041	71380	187547									
3	23115	12067	139580									
4	53883	69574	175167									
5	49880	45000	55385									
6	53883	56665	175167									
7	3129	58842	88441									
8	27065	94943	137448									
9	2020	59968	147412									
10	9547	58703	238111									
11	41866	65527	104381									
12	15459	46488	34920									
13	716	52476	81951									
14	56984	3055	159937									
15	17569	54889	115696									
16	46809	43108	73115									
17	39503	80831	113469									
18	2020	86583	97260									
19	45812	4524	35747									
20	27065	88298	82222									
21	56984	28078	159937									
22												

Menu Resumo estatístico

Os comandos do menu Resumo estatístico permitem que você calcule diversas medidas numéricas de resumo para variáveis simples ou pares de variáveis. Observe que não há tabelas de contingência nos procedimentos do StatTools. O Excel já oferece essa funcionalidade com as tabelas dinâmicas.

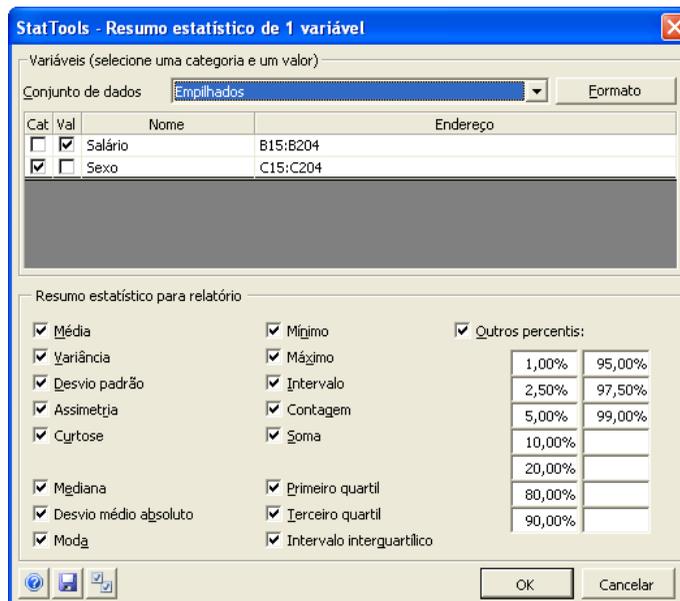
Comando Resumo de 1 variável

Calcula estatísticas de resumo para variáveis

O comando **Resumo de 1 variável** oferece dados resumidos para qualquer número de variáveis numéricas selecionadas. Elas incluem: média, mediana, desvio padrão, variância, mínimo, máximo, intervalo, primeiro quartil, terceiro quartil, intervalo interquartil, desvio absoluto da média, assimetria, curtose, contagem, soma e percentis selecionados.

Caixa de diálogo Resumo estatístico de 1 variável

Esta análise é definida com a caixa de diálogo **Resumo estatístico de 1 variável**:



Uma ou mais variáveis podem ser selecionadas para análise. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções na caixa de diálogo Utilidade de resumo de 1 variável inclui:

- **Resumo estatístico para relatório** – seleciona as estatísticas desejadas que devem ser incluídas no relatório. Os percentis podem ser adicionados especificando-se os valores desejados.

Relatório de resumo de 1 variável

O Relatório de resumo de 1 variável usa as funções Stat do StatTools (como **StatSkewness**) para permitir vínculos ativos com dados. O relatório é colocado na localização especificada usando o comando Configurações.

The screenshot shows the Microsoft Excel interface with the StatTools ribbon active. The ribbon includes options like 'Resumo estatístico', 'Gráficos de resumo', 'Inferência estatística', 'Testes de normalidade', 'Série temporal e previsão', 'Regressão e classificação', 'Controle de qualidade', 'Testes não-paramétricos', 'Utilidades', and 'Ajuda'. The main window displays a report titled 'Relatório do StatTools' with the following content:

Relatório do StatTools
 Análise: Resumo de uma variável
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Tempo real

Resumo de 1 variável	Salário (F)	Salário (M)
Empilhados	Empilhados	Empilhados
Média	\$29441,67	\$30089,36
Variância	14737192,98	12661821,09
Desvio padrão	\$3838,91	\$3558,35
Assimetria	-0,3034	0,0493
Curtose	3,1872	2,4343
Mediana	\$29700,00	\$29900,00
Desvio médio absoluto	\$3050,35	\$2893,16
Moda	\$25100,00	\$26900,00
Mínimo	\$17100,00	\$22400,00
Máximo	\$36900,00	\$38200,00
Intervalo	\$19800,00	\$15800,00
Contagem	96	94
Soma	\$2826400,00	\$2828400,00
1º quartil	\$26700,00	\$27500,00
3º quartil	\$31900,00	\$32500,00
Intervalo interquartilico	\$5200,00	\$5000,00
1,00%	\$17100,00	\$22400,00
2,50%	\$22300,00	\$23400,00
5,00%	\$22800,00	\$23900,00
10,00%	\$24500,00	\$25200,00
20,00%	\$26500,00	\$26900,00
80,00%	\$32400,00	\$33800,00
90,00%	\$34600,00	\$35000,00
95,00%	\$36200,00	\$35900,00
97,50%	\$36400,00	\$36800,00
99,00%	\$36900,00	\$38200,00

**Dados ausentes
e conectar aos
dados**

- **Dados ausentes** – este procedimento permite dados ausentes pela técnica “casewise”. Ou seja, para cada variável, os dados ausentes para aquela variável são ignorados no cálculo das medidas de resumo. (De qualquer forma, este é o método padrão do Excel. Por exemplo, se você usar a função MÉDIA em um intervalo, a média tirada será apenas dos valores *numéricos* no intervalo.)
- **Conectar aos dados** – todas as medidas de resumo são calculadas por fórmulas vinculadas com dados. Portanto, se algum dado for alterado, as medidas de resumo mudam automaticamente.

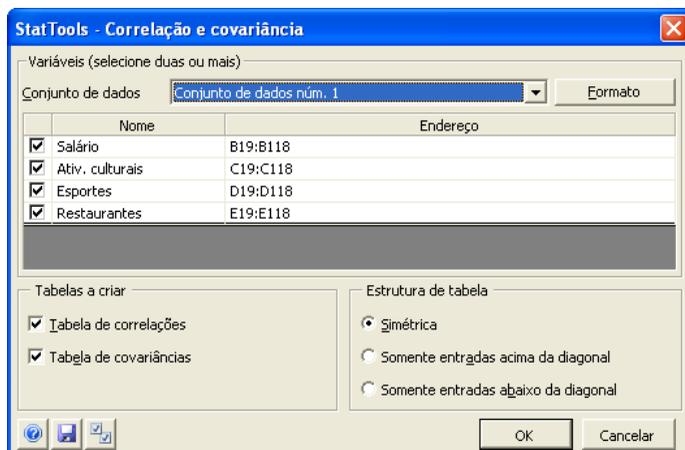
Comando Correlações e covariância

Produz uma tabela de correlações e/ou uma tabela de covariâncias entre variáveis

O comando **Correlações e covariância** produz uma tabela de correlações e/ou uma tabela de covariâncias entre qualquer conjunto de variáveis numéricas selecionadas. Como ambas as tabelas são simétricas (por exemplo, a correlação entre X e Y é a mesma que a correlação entre Y e X), você pode optar por ter: (1) apenas correlações (ou covariâncias) *abaixo* da diagonal mostrada, (2) apenas as *acima* da diagonal mostrada, ou (3) as *abaixo e acima* da diagonal mostrada.

Caixa de diálogo Correlações e covariância

Esta análise é definida com a caixa de diálogo **Correlações e covariância**:



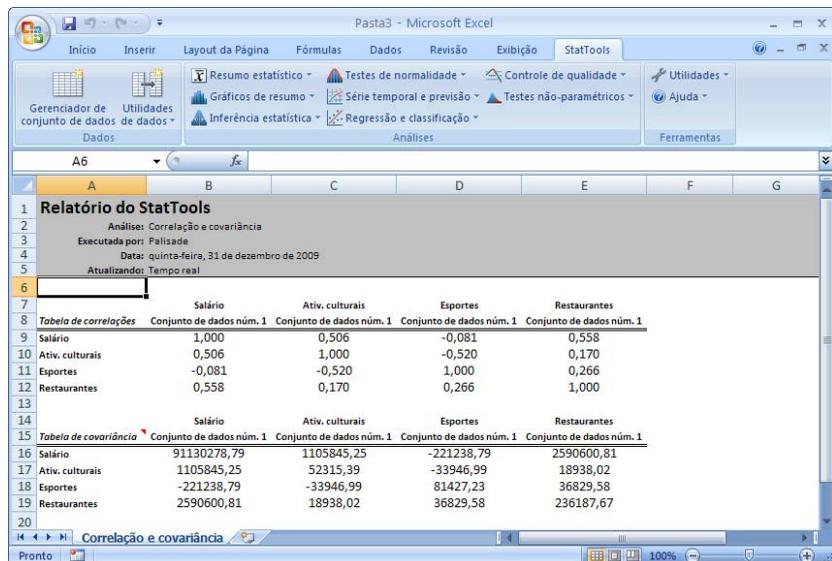
Duas ou mais variáveis devem ser selecionadas para análise. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções na caixa de diálogo Correlações e covariância incluem:

- **Tabelas a criar** – Seleciona a tabela desejada de correlação e/ou covariância
- **Estrutura da tabela** – especifica a estrutura das tabelas a serem geradas:
 - Correlações (ou covariâncias) **Simétricas** abaixo e acima da diagonal mostradas
 - Correlações (ou covariâncias) **Somente entradas acima da diagonal** (*acima* da diagonal mostrada)
 - Correlações (ou covariâncias) **Somente entradas abaixo da diagonal** (*abaixo* da diagonal mostrada)

Relatório de correlações e covariância

O Relatório de correlações e covariância usa as funções Stat do StatTools (como **StatCorrelationCoeff**) para permitir vínculos ativos com dados. O relatório é colocado na localização especificada usando o comando Configurações.



The screenshot shows the StatTools ribbon in Microsoft Excel. The report displays the following data:

	Salário	Ativ. culturais	Esportes	Restaurantes
Tabela de correlações	Conjunto de dados núm. 1			
Salário	1,000	0,506	-0,081	0,558
Ativ. culturais	0,506	1,000	-0,520	0,170
Esportes	-0,081	-0,520	1,000	0,266
Restaurantes	0,558	0,170	0,266	1,000

	Salário	Ativ. culturais	Esportes	Restaurantes
Tabela de covariância	Conjunto de dados núm. 1			
Salário	91130278,79	1105845,25	-221238,79	2590600,81
Ativ. culturais	1105845,25	52315,39	-33946,99	18938,02
Esportes	-221238,79	-33946,99	81427,23	36829,58
Restaurantes	2590600,81	18938,02	36829,58	236187,67

**Dados ausentes
e conectar aos
dados**

- **Dados ausentes** – dados ausentes são permitidos e tratados pela técnica pairwise. Ou seja, para obter a correlação (ou covariância) entre qualquer par de variáveis, todos os casos com dados ausentes em *qualquer uma* das duas variáveis são ignorados.
- **Conectar aos dados** – as correlações e covariâncias são calculadas por fórmulas vinculadas com dados. Portanto, se algum dado for alterado, as medidas de resumo mudam automaticamente.

Menu Gráficos de resumo

Os comandos do menu Gráficos de resumo permitem a criação de gráficos que são muito úteis em análises estatísticas e não muito fáceis (ou possíveis) de produzir com o assistente de gráficos do Excel. Claro, os recursos de criação de gráficos do Excel são amplos, assim, o StatTools procura não duplicar coisas que o Excel já faça bem.

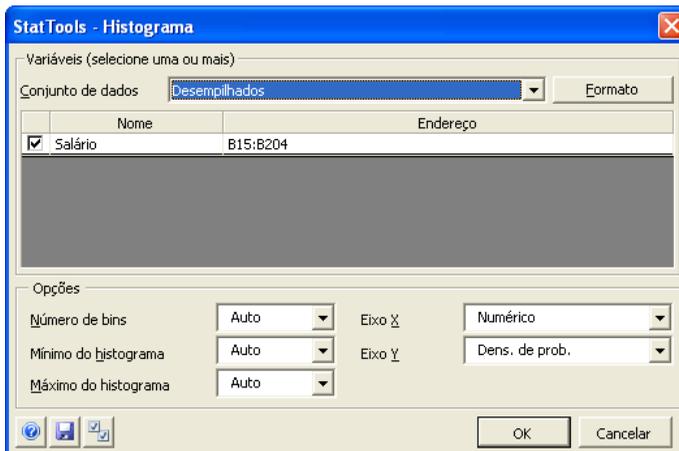
Comando Histograma

Cria histogramas de variáveis

O comando Histograma cria uma histograma para cada variável selecionada. Dá a você a opção de definir as categorias do histograma (normalmente chamadas de “bins”), e mostra-as claramente no gráfico. Também cria uma tabela de frequência em que se baseia cada histograma.

Caixa de diálogo Histograma

Este tipo de gráfico é configurado na caixa de diálogo **Histograma**:

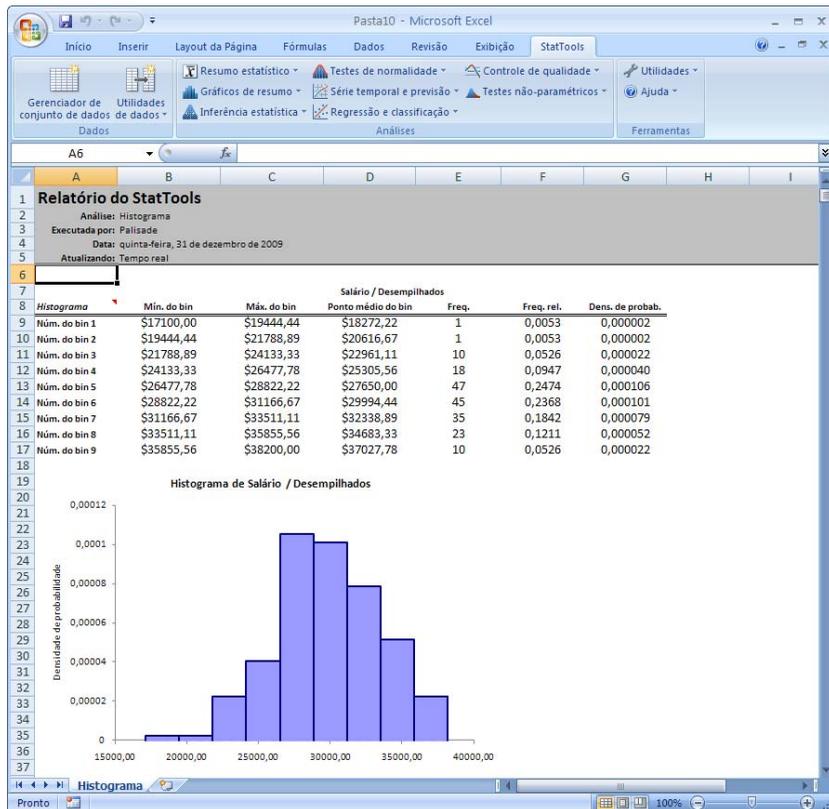


Uma ou mais variáveis podem ser selecionadas para a geração do gráfico. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Histograma** incluem:

- **Número de bins.** Define o número de intervalos do histograma calculados pelo intervalo de um gráfico. O valor entrado deve estar no intervalo de 1 a 200. A definição **Auto** calcula o melhor número de bins a usar para os seus dados com base em uma heurística interna.
- **Mínimo do histograma.** Define um valor mínimo onde os bins do histograma começam. **Auto** especifica que o StatTools iniciará os bins do histograma com base no mínimo dos dados no gráfico.
- **Máximo do histograma.** Define um valor máximo onde os bins do histograma terminam. **Auto** especifica que o StatTools terminará os bins do histograma com base no máximo dos dados no gráfico.
- **Eixo X.** Seleciona **Catégorica** ou **Numérica**. Um eixo X categórico simplesmente identifica cada bin com o ponto médio do bin. O eixo do X numérico tem um mínimo e máximo de eixo X “legíveis”, e pode ser redimensionado usando as opções de redimensionamento padrão do Excel.
- **Eixo Y.** Seleciona **Frequência**, **Frequência rel.** ou **Dens. de prob.** como a unidade de medida relatada no eixo Y. Frequência é o número real de observações contidas em um bin. A frequência relativa é a probabilidade de ocorrência de um valor no intervalo de um bin (observações em um bin/observações totais). Densidade é o valor de frequência relativa dividida pela abrangência do bin, garantindo que os valores do eixo Y permaneçam constantes à medida que o número de bins é alterado.

Histograma de uma variável simples



Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – os histogramas são parcialmente vinculados a dados. Quando os dados são alterados e novos valores se enquadrarem no intervalo dos bins do histograma original, o gráfico será automaticamente atualizado. As alterações de dados que demandem nova formação de bins não atualizam o gráfico.

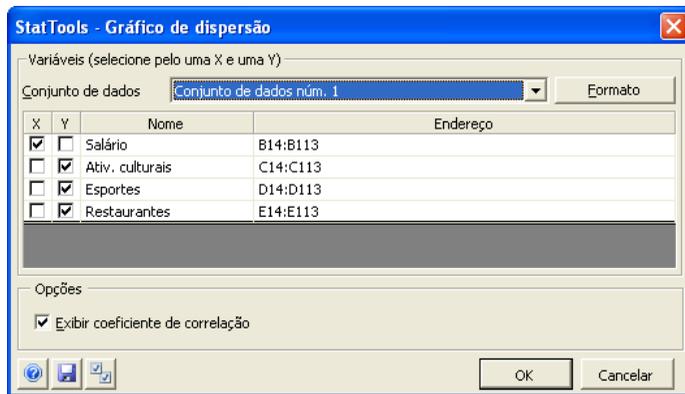
Comando Gráfico de dispersão

Cria gráficos de dispersão para pares de variáveis

O comando Gráfico de dispersão cria um gráfico de dispersão para cada par de variáveis que você selecionar. Os gráficos de dispersão podem ser criados com a opção de plotagem XY do Excel, mas o Excel automaticamente põe a primeira (ou mais à esquerda) variável no eixo horizontal, que pode não ser a desejada. O StatTools permite que você escolha qual das variáveis será colocada no eixo horizontal. Cada gráfico de dispersão mostra a correlação entre duas variáveis com o gráfico correspondente.

Caixa de diálogo Gráfico de dispersão

Este tipo de gráfico é configurado na caixa de diálogo **Gráfico de dispersão**:

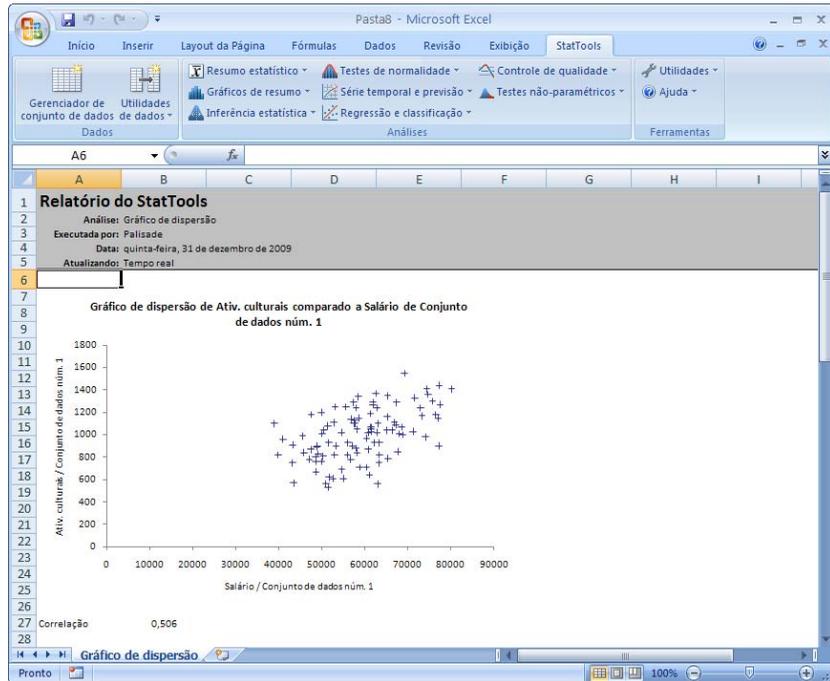


Duas ou mais variáveis podem ser selecionadas para a geração do gráfico. Pelo menos uma variável do eixo X e uma do eixo Y são necessárias. Se mais de duas variáveis forem selecionadas, múltiplos gráficos de dispersão são criados. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Gráfico de dispersão** incluem:

- **Exibir coeficiente de correlação.** Especifica que o coeficiente de correlação entre as variáveis do gráfico será exibido

Exemplo de gráfico de dispersão



Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes em uma das duas variáveis selecionadas em qualquer par são ignoradas.
- **Conectar aos dados** – os gráficos de dispersão são vinculados aos dados originais. Se os dados forem alterados, os gráficos de dispersão também serão. Contudo, as escalas dos eixos podem ter de ser atualizadas manualmente se os intervalos das variáveis selecionadas mudarem de forma significativa.

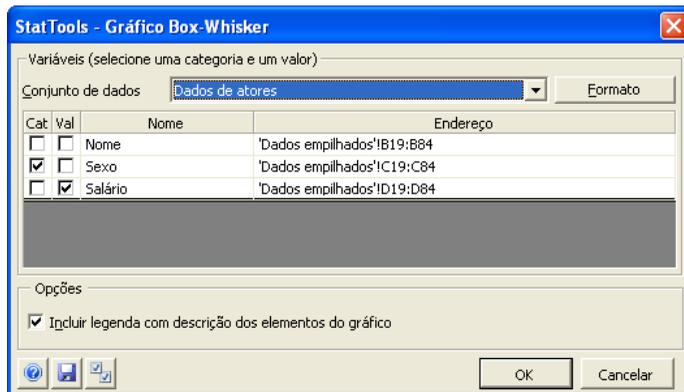
Comando Gráfico Box-Whisker

Cria gráficos Box-Whisker para variáveis

O comando gráfico Box-Whisker cria um gráfico Box-Whisker simples (se você selecionar uma única variável) ou gráficos Box-Whisker lado a lado (se você selecionar diversas variáveis). Também cria uma planilha que mostra as estatísticas de resumo (quartis, intervalo interquartil, etc.) usados para formar gráfico(s) Box-Whisker.

Caixa de diálogo Gráfico Box-Whisker

Este tipo de gráfico é configurado na caixa de diálogo **Gráfico Box-Whisker**:

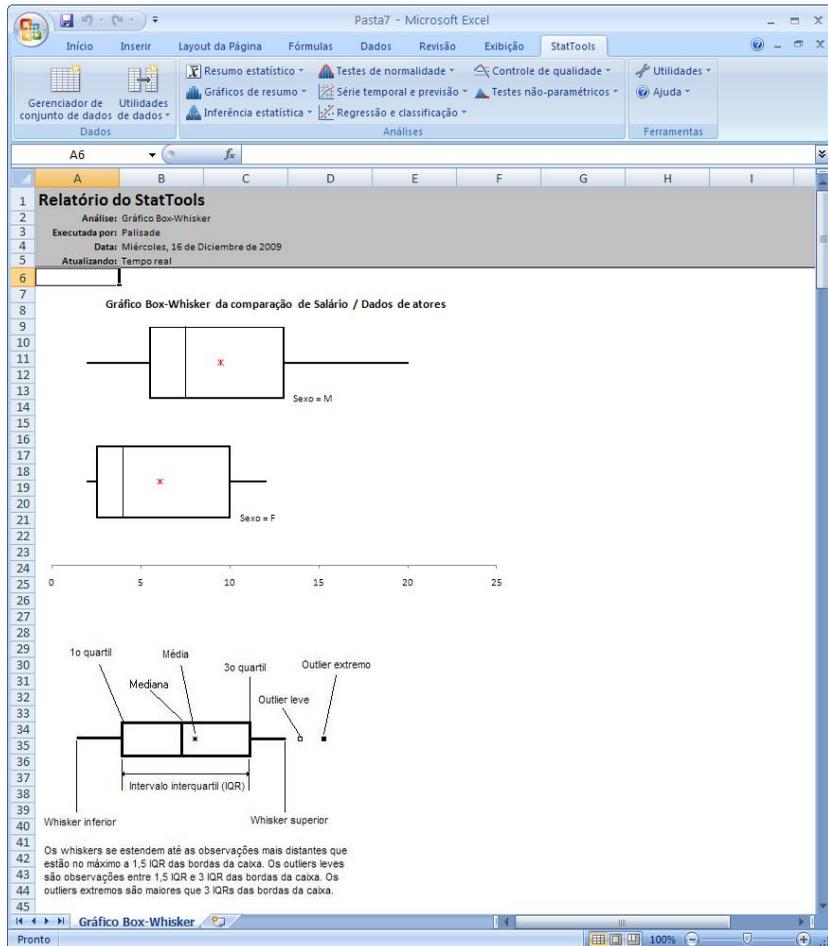


Uma ou mais variáveis podem ser selecionadas para a geração do gráfico. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Gráfico Box-Whisker** incluem:

- **Incluir legenda com descrição dos elementos do gráfico.** Especifica que uma chave separada que descreve os elementos do gráfico serão exibidos abaixo do gráfico.

Exemplo de gráfico Box-Whisker



Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes em qualquer das variáveis selecionadas para o gráfico são ignorados.
- **Conectar aos dados** – os gráficos Box-Whisker criados são vinculados com os dados originais; se os dados forem alterados, os gráficos Box-Whisker também serão. Contudo, a escala do eixo horizontal pode precisar ser atualizada manualmente se a escala dos dados alterar de forma significativa

Menu Inferência estatística

Os comandos do menu Inferência estatística executam as análises de inferência estatística mais comuns: testes de hipótese e intervalos de confiança, bem como ANOVA uni e bidirecional.

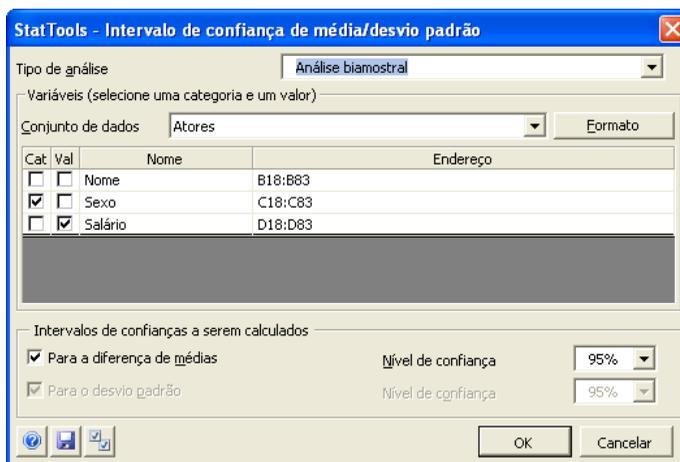
Comando Intervalo de confiança de média/desvio padrão

Calcula intervalos de confiança para média e desvio padrão de variáveis

O comando StatTools – Intervalo de confiança de média/desvio padrão calcula o intervalo de confiança da média e do desvio padrão de variáveis simples, ou as diferenças entre as médias de pares de variáveis. Os intervalos de confiança podem ser calculados através de uma **Análise uniamostrual**, uma **Análise biamostrual** ou uma **Análise de amostras pareadas**.

Essa análise é definida na caixa de diálogo **Intervalo de confiança de média/desvio padrão**:

Caixa de diálogo Intervalo de confiança de média/desvio padrão

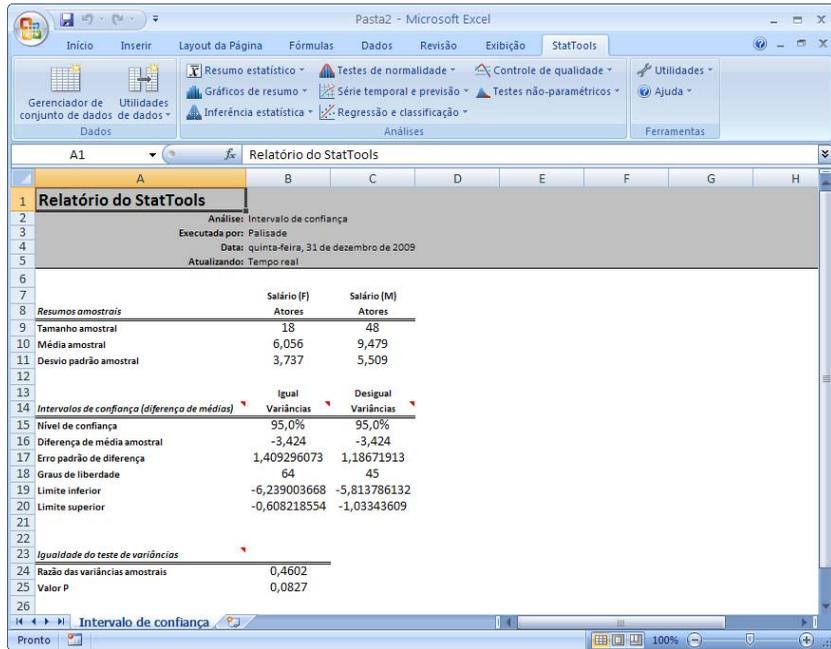


O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise uniamostrual exige uma ou mais variáveis, enquanto uma análise biamostrual e uma análise de amostra pareada exigem duas variáveis. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Intervalo de confiança** incluem:

- **Tipo de análise.** Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostrual.** Calcula intervalos de confiança para uma variável numérica simples.
 - **Análise biamostrual.** Calcula o intervalo de confiança para a diferença entre as médias de duas populações independentes.
 - **Análise de amostras pareadas.** Trata-se basicamente da mesma coisa que a análise biamostrual, mas aplica-se quando duas variáveis são naturalmente pareadas de alguma forma. Em essência, ela executa uma análise uniamostrual nas diferenças entre pares.
- **Intervalos de confianças a serem calculados.** Especifica os intervalos de confiança que serão calculados com as variáveis selecionadas. As opções mudam de acordo com o tipo de análise selecionado:
 - **Análise uniamostrual.** Define o cálculo de intervalos de confiança sobre a média e/ou desvio padrão, e o nível de confiança (0 a 100%) para cada.
 - **Análise biamostrual** ou **Análise de amostras pareadas.** Define o cálculo do intervalo de confiança pela diferença entre as médias para duas variáveis, e especifica o nível de confiança (0 a 100%) desejado.

Relatório de intervalo de confiança



The screenshot shows the Microsoft Excel interface with the StatTools ribbon active. The report is titled "Relatório do StatTools" and displays the following data:

Análise: Intervalo de confiança		
Executada por: Palisade		
Data: quinta-feira, 31 de dezembro de 2009		
Atualizando: Tempo real		
Resumos amostrais		
	Salário (F)	Salário (M)
	Atores	Atores
Tamanho amostral	18	48
Média amostral	6,056	9,479
Desvio padrão amostral	3,737	5,509
Intervalos de confiança (diferença de médias)		
	Igual	Desigual
	Variâncias	Variâncias
Nível de confiança	95,0%	95,0%
Diferença de média amostral	-3,424	-3,424
Erro padrão de diferença	1,409296073	1,18671913
Graus de liberdade	64	45
Limite inferior	-6,239003668	-5,813786132
Limite superior	-0,608218554	-1,03343609
Igualdade do teste de variâncias		
Razão das variâncias amostrais	0,4602	
Valor P	0,0827	

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

Comando Intervalo de confiança – Proporção

Calcula os intervalos de confiança de proporções

O comando Intervalo de confiança de proporções permite a análise da proporção dos itens em uma amostra que pertence a uma determinada categoria (**Análise uniamostrual**), ou a comparação de duas amostras quanto à proporção dos itens em uma determinada categoria (**Análise biamostrual**). Há três tipos de dados que podem ser usados nessa análise: **Amostra da população**, **Tabela de resumo com contagens** e **Tabela de resumo com proporções**.

Caixa de diálogo Intervalo de confiança para proporção

Esta análise é definida com a caixa de diálogo **Intervalo de confiança para proporção**:

StatTools - Intervalo de confiança para proporção

Tipo de análise: Análise uniamostrual

Tipo de dados: Tabela de resumo com contagens

Selecione uma variável com categorias e uma ou mais variáveis com contagens

Conjunto de dados: Contagens simples

Cat	Ct	Nome	Endereço
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Categoria	B18:B20
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Contagem	C18:C20

Categorias a analisar (selecionar uma ou mais):

- Democrata
- Outra
- Republicano

3 categorias na variável selecionada.

Opções

Nível de confiança: 95%

Tamanho da primeira amostra: 150

Tamanho da segunda amostra: 150

OK Cancelar

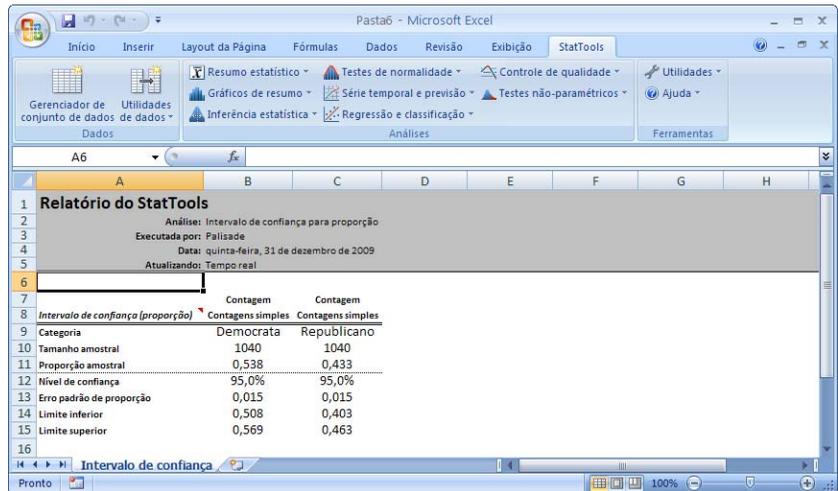
O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise **uniamostr**al exige uma ou mais amostras para análise; uma análise **biamostr**al exige duas amostras. Um define variáveis com informações de amostra usando a coluna identificada como **Ct** (contagem), % (proporção), ou coluna sem identificação para o tipo de dados **Amostra populacional**. (Se os dados da amostra populacional forem empilhados, as amostras são selecionadas nas colunas identificadas como **C1** e **C2**, em que C1 contém as categorias empilhadas.) Se os dados estiverem na forma de uma tabela com contagens ou proporções, haverá uma coluna **Cat** adicional para a seleção de uma variáveis com nomes de categoria.

Com o tipo de dados **Amostra populacional**, as variáveis podem ser de diferentes conjuntos de dados.

As opções da caixa de diálogo **Intervalo de confiança** incluem:

- **Tipo de análise.** Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostr**al. Calcula os intervalos de confiança da proporção de itens em uma amostra que pertence a uma determinada categoria.
 - **Análise biamostr**al. Calcula o intervalo de confiança de duas amostras quanto à proporção de itens em uma determinada categoria
- **Tipo de dados.** Especifica o tipo de dados a ser analisado: **Amostra populacional**, **Tabela de resumo com contagens** ou **Tabela de resumo com proporções**.
- **Opções.** As opções disponíveis mudam com o Tipo de análise e o Tipo de dados. As opções são:
 - **Nível de confiança.** Selecione o nível de confiança (0 a 100%) para a análise.
 - **Tamanho da primeira amostra** e **Tamanho da segunda amostra.** Para o tipo de dados **Tabela de resumo com proporções**, seleciona o tamanho da primeira amostra e da segunda amostra (apenas a análise biamostral).

Relatório de intervalo de confiança



The screenshot shows the Microsoft Excel interface with the StatTools ribbon active. The report is titled "Relatório do StatTools" and provides the following details:

- Análise: Intervalo de confiança para proporção
- Executada por: Palisade
- Data: quinta-feira, 31 de dezembro de 2009
- Atualizando: Tempo real

Intervalo de confiança (proporção)	Contagem	
	Contagens simples	Contagens simples
Categoria	Democrata	Republicano
Tamanho amostral	1040	1040
Proporção amostral	0,538	0,433
Nível de confiança	95,0%	95,0%
Erro padrão de proporção	0,015	0,015
Limite inferior	0,508	0,403
Limite superior	0,569	0,463

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

Comando Teste de hipótese - média/desvio padrão

Executa um teste de hipótese da média e do desvio padrão de variáveis

O comando Teste de hipótese para média/desvio padrão executa testes de hipótese da média e do desvio padrão de variáveis simples, ou as diferenças entre as médias de pares de variáveis. Os testes de hipótese podem ser executados através de uma **análise uniamostr**al, uma **análise biamostr**al ou uma **análise de amostras pareadas**.

Caixa de diálogo **Teste de hipótese** **para média/desvio** **padrão**

Essa análise é definida na caixa de diálogo **Teste de hipótese para média/desvio padrão**:

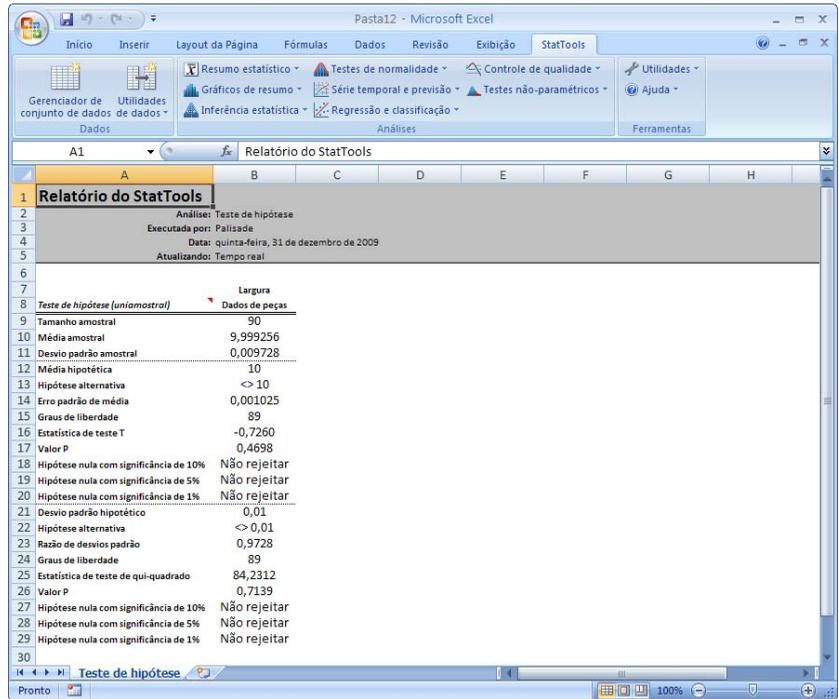
Nome	Endereço
<input type="checkbox"/> Núm. peça	B18:B107
<input checked="" type="checkbox"/> Largura	C18:C107

O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise uniamostr exige uma ou mais variáveis, enquanto uma análise biamostr e uma análise de amostra pareada exigem duas variáveis. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Teste de hipótese** incluem:

- **Tipo de análise.** Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostr.** Executa testes de hipótese para uma variável numérica simples.
 - **Análise biamostr.** Calcula testes de hipótese para a diferença entre as médias de duas populações independentes.
 - **Análise de amostras pareadas.** Trata-se basicamente da mesma coisa que a análise biamostr., mas aplica-se quando duas variáveis são naturalmente pareadas de alguma forma. Em essência, ela executa uma análise uniamostr. nas diferenças entre pares.
- **Teste de hipótese a realizar.** Especifica os testes de hipótese que serão executados com as variáveis selecionadas. As opções mudam de acordo com o tipo de análise selecionado. Uma **análise uniamostr.** seleciona testes de hipótese sobre a média e/ou desvio padrão. Uma **análise biamostr.** ou **análise de amostras pareadas** selecionam testes de hipótese para a diferença entre as médias de duas variáveis. Para cada teste de hipótese selecionado, as opções incluem:
 - **Valor de hipótese nula,** ou o valor do parâmetro de população sob hipótese nula.
 - **Tipo de hipótese alternativa,** ou a alternativa ao Valor de hipótese nula que será avaliado durante a análise. O Tipo de hipótese alternativa pode ser “**monocaudal**” (ou seja, maior ou menor que a hipótese nula) ou “**bicaudal**” (ou seja, não igual à hipótese nula).

Relatório de testes de hipótese



Relatório do StatTools

Análise: Teste de hipótese
Executada por: Palisade
Data: quinta-feira, 31 de dezembro de 2009
Atualizando: Tempo real

Teste de hipótese (uniamostr)	Largura	Dados de peças
Tamanho amostral	90	
Média amostral	9,999256	
Desvio padrão amostral	0,009728	
Média hipotética	10	
Hipótese alternativa	<> 10	
Erro padrão de média	0,001025	
Graus de liberdade	89	
Estatística de teste T	-0,7260	
Valor P	0,4698	
Hipótese nula com significância de 10%	Não rejeitar	
Hipótese nula com significância de 5%	Não rejeitar	
Hipótese nula com significância de 1%	Não rejeitar	
Desvio padrão hipotético	0,01	
Hipótese alternativa	<> 0,01	
Razão de desvios padrão	0,9728	
Graus de liberdade	89	
Estatística de teste de qui-quadrado	84,2312	
Valor P	0,7139	
Hipótese nula com significância de 10%	Não rejeitar	
Hipótese nula com significância de 5%	Não rejeitar	
Hipótese nula com significância de 1%	Não rejeitar	

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

Comando Teste de hipótese - Proporção

Executa um teste de hipótese para proporções

O comando Teste de hipótese para proporção analisa a proporção dos itens em uma amostra que pertence a uma determinada categoria (análise uniamostrual), ou compara duas amostras quanto à proporção dos itens em uma determinada categoria (análise biamostrual). Há três tipos de dados que podem ser usados nessa análise: **Amostra da população**, **Tabela de resumo com contagens** e **Tabela de resumo com proporções**.

**Caixa de diálogo
Teste de hipótese
para proporção**

Esta análise é definida com a caixa de diálogo **Teste de hipótese para proporção**:

StatTools - Teste de hipótese para proporção

Tipo de análise:

Tipo de dados:

Selecione uma variável com categorias e uma ou mais variáveis com contagens

Conjunto de dados:

Cat	Ct	Nome	Endereço
<input checked="" type="checkbox"/>		Categoria	B18:B19
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Contagem	C18:C19

Categorias a analisar (selecionar uma ou mais)

Insatisfeito

Satisfeito

2 categorias na variável selecionada.

Hipóteses sobre proporção

Valor de hipótese nula:

Tipo de hipótese alternativa:

Opções

Tamanho da primeira amostra:

Tamanho da segunda amostra:

O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise **uniamostr**al exige uma ou mais amostras para análise; uma análise **biamostr**al exige duas amostras. Um define variáveis com informações de amostra usando a coluna identificada como **Ct** (contagem), % (proporção), ou coluna sem identificação para o tipo de dados **Amostra populacional**. (Se os dados da amostra populacional forem empilhados, as amostras são selecionadas nas colunas identificadas como **C1** e **C2**, em que C1 contém as categorias empilhadas.) Se os dados estiverem na forma de uma tabela com contagens ou proporções, haverá uma coluna **Cat** adicional para a seleção de uma variáveis com nomes de categoria.

Com o tipo de dados **Amostra populacional**, as variáveis podem ser de diferentes conjuntos de dados.

As opções da caixa de diálogo **Teste de hipótese para proporção** incluem:

- **Tipo de análise**. Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostr**al. Executa o teste de hipótese da proporção de itens em uma amostra que pertence a uma determinada categoria.
 - **Análise biamostr**al. Executa o teste de hipótese de duas amostras quanto à proporção de itens em uma determinada categoria.
- **Tipo de dados**. Especifica o tipo de dados a ser analisado: **Amostra populacional**, **Tabela de resumo com contagens** ou **Tabela de resumo com proporções**.
- **Hipóteses sobre proporção**. Especifica os testes de hipótese que serão executados com a proporção selecionada. As opções mudam de acordo com o tipo de análise selecionado. As opções são:
 - **Valor de hipótese nula**, ou o valor do parâmetro de população sob hipótese nula.
 - **Tipo de hipótese alternativa**, ou a alternativa ao Valor de hipótese nula que será avaliado durante a análise. O Tipo de hipótese alternativa pode ser "**monocaudal**" (ou seja, maior ou menor que a hipótese nula) ou "**bicaudal**" (ou seja, não igual à hipótese nula).

- **Tamanho da primeira amostra e Tamanho da segunda amostra.** Para o tipo de dados **Tabela de resumo com proporções**, seleciona o tamanho da primeira amostra e da segunda amostra (apenas a análise biamostrai).

Relatório de testes de hipótese

Relatório do StatTools	
Análise: Teste de hipótese para proporção	
Executada por: Palisade	
Data: quinta-feira, 31 de dezembro de 2009	
Atualizando: Tempo real	
Contagem	
Tabela com contagens	
Teste de hipótese (proporção)	
Categoria	Insatisfeito
Tamanho amostral	400
Proporção amostral	0,058
Proporção hipotética	0,075
Hipótese alternativa	< 0,075
Erro padrão de proporção amostral	0,013
Estatística de teste Z	-1,3288
Valor P	0,0920
Hipótese nula com significância de 10%	Rejeitar
Hipótese nula com significância de 5%	Não rejeitar
Hipótese nula com significância de 1%	Não rejeitar

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

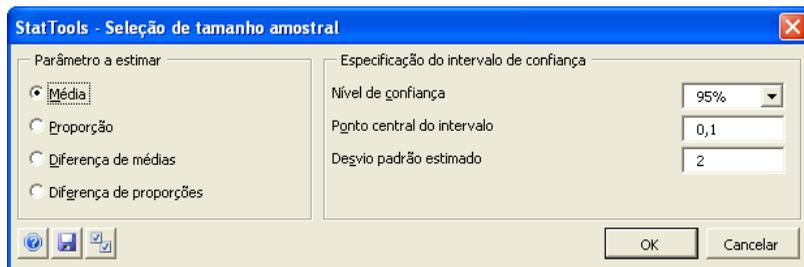
Comando Seleção de tamanho amostral

Determina o tamanho da amostra necessário para calcular intervalos de confiança

O comando Seleção de tamanho amostral determina o tamanho da amostra (ou tamanhos de amostra) necessário para se obter um intervalo de confiança com um pontos central prescrito. Faz isso para intervalos de confiança para uma média, uma proporção, a diferença entre duas médias e a diferença entre duas proporções. Não são necessários conjuntos de dados nem variáveis, uma vez que o tamanho da amostra é a informação normalmente necessária *antes* de coletar os dados. É preciso especificar o nível de confiança, o ponto central desejado e quaisquer outros parâmetros necessários para a determinação do tamanho da amostra.

Caixa de diálogo Seleção do tamanho amostral

Esta análise é definida com a caixa de diálogo **Seleção do tamanho amostral**:

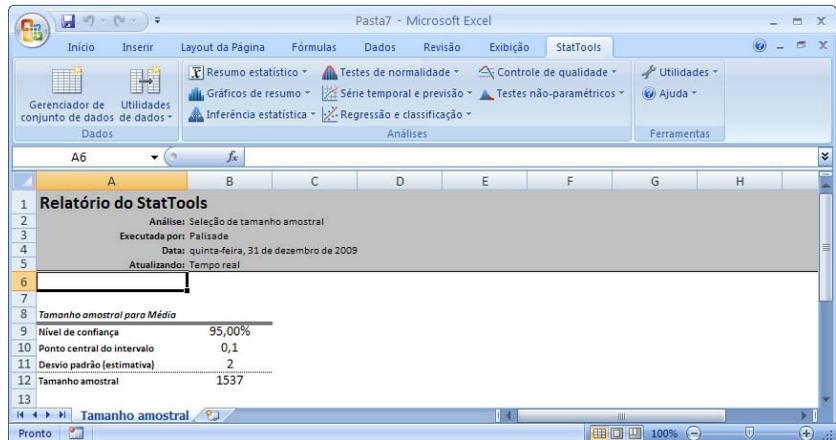


As opções na caixa de diálogo **Seleção do tamanho amostral** incluem:

- **Parâmetro a estimar** – seleciona o tipo de parâmetro que será estimado com a amostra (cujo tamanho você está determinando). As opções incluem **Média**, **Proporção** (valores entre 0 e 1), **Diferenças de médias** e **Diferenças de proporções**.

- **Especificação do intervalo de confiança** – estas opções variam de acordo com o Parâmetro a estimar selecionado, como a seguir:
 - Quando Parâmetro a estimar for **Média e Diferença de médias**. Primeiro, insira o **Nível de confiança** desejado (normalmente entre 90% e 100%), o **Ponto central do intervalo** (o componente “mais ou menos” do intervalo) e o **Desvio padrão estimado** da população. Nota: o nível de confiança e o comprimento do intervalo estão relacionados, uma vez que um nível de confiança maior exige um comprimento de intervalo de confiança maior.
 - Quando Parâmetro a estimar for **Proporção e Diferença de proporções**. Primeiro, insira o **Nível de confiança** desejado (normalmente entre 90% e 100%), o **Ponto central do intervalo** (o componente “mais ou menos” do intervalo) e o **Proporção estimada** (um valor entre 0 e 1). Se **Diferença de proporções** estiver sendo estimado, especifique uma **Proporção estimada** para cada população.

Relatório de tamanho de amostra estimado



Dados ausentes e conectar aos dados

- **Dados ausentes** – não relevante.
- **Conectar aos dados** – não relevante.

Comando ANOVA unidirecional

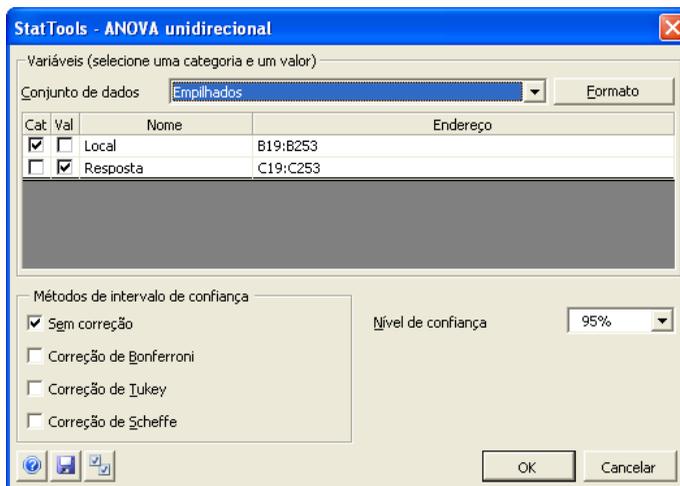
Executa uma ANOVA unidirecional sobre variáveis

O comando ANOVA unidirecional é uma generalização do procedimento biométrico para a comparação de médias entre duas populações. Com a ANOVA unidirecional, as médias de *pelo menos duas* (normalmente mais de duas) populações são comparadas. Isso é feito com uma tabela ANOVA (análise de variância). Esta tabela, na verdade, compara duas fontes de variação: a variação *dentro* de cada população em relação à variação *entre* médias amostrais de diferentes populações. Se a segunda variação for grande em relação à primeira, de acordo com a medição de um teste F, então há indicação de diferenças entre as médias das populações.

O valor-chave na tabela ANOVA é o valor p. Um valor p pequeno indica médias populacionais diferentes. Além da tabela ANOVA, é útil considerar os intervalos de confiança de todas as diferenças entre pares de médias. Os intervalos de confiança que *não* incluem 0 indicam médias que *não* são iguais. O StatTools oferece a opção de diversos tipos de intervalos de confiança, cada qual baseado em um método ligeiramente diferente.

Caixa de diálogo ANOVA

Esta análise é definida com a caixa de diálogo **ANOVA unidirecional**:



Duas ou mais variáveis devem ser selecionadas para análise. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções na caixa de diálogo ANOVA unidirecional inclui:

- **Métodos de intervalo de confiança** – selecione um ou mais dos seguintes métodos de correção de intervalos de confiança de variáveis individuais. **Sem correção**, ou a não execução de correções, e os métodos de correção de **Bonferroni**, **Tukey** e **Scheffe**.
- **Nível de confiança** – trata-se de um nível de confiança “simultâneo” dos resultados de todas as variáveis. Ou seja, é a confiança que você quer ter de que *todos* os intervalos de confiança contenham suas respectivas diferenças de médias populacionais. Por motivos técnicos, o nível de confiança geral *real* normalmente será menor que esse nível especificado para os intervalos de confiança “sem correção”. Esta é a razão dos métodos de “correção” mencionados. Eles corrigem (expandem) o comprimento dos intervalos de confiança para que o nível de confiança geral seja o especificado.

Relatório ANOVA unidirecional

Análise: ANOVA unidirecional
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Tempo real

Resumo de ANOVA

Tamanho amostral total	235
Média total	5,383
Desvio padrão combinado	1,976
Variância combinada	3,904
Número de amostras	5
Nível de confiança	95,00%

Estatísticas amostrais de ANOVA

	Resposta (Centro-oeste) Empilhados	Resposta (Nordeste) Empilhados	Resposta (Oeste) Empilhados	Resposta (Sudoeste) Empilhados	Resposta (Sul) Empilhados
Tamanho amostral	55	43	50	47	40
Média amostral	5,400	4,140	4,980	6,745	5,600
Desvio padrão amostral	2,469	1,820	1,635	1,687	2,073
Variância amostral	6,096	3,313	2,673	2,846	4,297
Peso combinado	0,2348	0,1826	0,2130	0,2000	0,1696

Tabela ANOVA unidirecional

	Soma dos quadrados	Graus de liberdade	Quadrados médios	Razão F	Valor P
Varição entre grupos	163,653	4	40,913	10,480	< 0,0001
Varição dentro do grupo	897,879	230	3,904		
Varição total	1061,532	234			

Testes de intervalo de confiança

	Diferença de médias	Sem correção	
		Inferior	Superior
Resposta (Centro-oeste)-Resposta (Nordeste)	1,260	0,467996986	2,052933246
Resposta (Centro-oeste)-Resposta (Oeste)	0,420	-0,340698646	1,180898646
Resposta (Centro-oeste)-Resposta (Sudoeste)	-1,345	-2,117991788	-0,571369914
Resposta (Centro-oeste)-Resposta (Sul)	-0,200	-1,008974445	0,608974445
Resposta (Nordeste)-Resposta (Oeste)	-0,840	-1,650132388	-0,30797844
Resposta (Nordeste)-Resposta (Sudoeste)	-2,605	-3,426674097	-1,783617838
Resposta (Nordeste)-Resposta (Sul)	-1,460	-2,315648321	-0,605281911
Resposta (Oeste)-Resposta (Sudoeste)	-1,765	-2,555607623	-0,973754079
Resposta (Oeste)-Resposta (Sul)	-0,620	-1,445829889	0,205829889
Resposta (Sudoeste)-Resposta (Sul)	1,145	0,307218982	1,98214272

No relatório ANOVA unidirecional, os resumos de estatísticas de cada população (neste caso, cada fábrica) aparecem em cima. Depois dos resumos de estatísticas, há uma tabela de amostras estatísticas de cada variável. A tabela ANOVA aparece em seguida. Neste exemplo, o valor p muito pequeno indica sem dúvida que as pontuações médias das cinco fábricas não são iguais. Para saber quais médias são diferentes de outras, consideramos os intervalos de confiança na parte inferior do relatório. Os pares com valores em negrito tem médias ligeiramente diferentes.

**Dados ausentes
e conectar aos
dados**

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

Comando ANOVA bidirecional

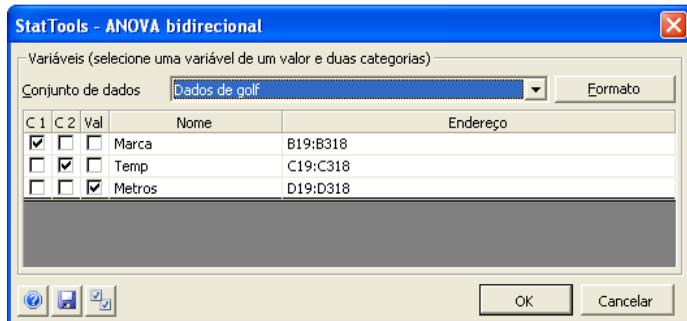
Executa uma ANOVA bidirecional sobre variáveis

O comando ANOVA bidirecional executa uma análise de variância bidirecional. Isto é normalmente feito no contexto de um projeto experimental, em que há “fatores”, cada qual definido em diversos níveis de “tratamento”. Por exemplo, em um estudo de desempenho de bola de golfe, os dois fatores poderiam ser *Marca* e *Temperatura externa*. Os níveis de tratamento para *Marca* seriam de “A” até “E”, e os níveis de tratamento para *Temperatura* seriam “Frio”, “Ameno” e “Quente”. A variável de valor seria *Lançamento em metros*, e as observações para esta variável seriam coletadas de um número de bolas de cada combinação de *Marca/Temperatura*. A finalidade do estudo seria verificar se há diferenças de média significativas entre as várias combinações de nível de tratamento.

Os dados da ANOVA bidirecional *devem* estar no formato empilhado. Ou seja, deve haver duas variáveis “de categoria” (para *Marca* e *Temperatura externa* no exemplo acima), e deve haver uma variável de “valor” (para *Lançamento em metros* acima). Além disso, os dados definidos devem estar “balanceados”, ou seja, deve haver um número igual de observações em cada combinação de nível de tratamento. Certamente é possível analisar um design não balanceado, mas este design é mais bem analisado com regressão (com variáveis fictícias).

Esta análise é definida com a caixa de diálogo **ANOVA bidirecional**:

Caixa de diálogo ANOVA



Duas variáveis de categoria (C1 e C2) e uma variável de valor (Val) precisam ser selecionadas para esta análise. O conjunto de dados selecionado deve estar no formato empilhado.

Relatório ANOVA bidirecional

Relatório do StatTools

Análise: ANOVA bidirecional
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Temporal

Tamanhos amostrais de ANOVA				
	Amena	Fria	Quente	Totais
A	20	20	20	60
B	20	20	20	60
C	20	20	20	60
D	20	20	20	60
E	20	20	20	60
Totais	100	100	100	

Equilibrado VERDADEIRO

Médias amostrais de ANOVA				
	Amena	Fria	Quente	Totais
A	236,45	218,82	258,44	237,90
B	245,13	224,15	258,27	242,52
C	242,72	228,00	263,04	244,58
D	237,62	215,00	256,11	236,24
E	255,75	224,79	270,94	250,49
Totais	243,53	222,15	261,36	

Desvio padrão amostral de ANOVA				
	Amena	Fria	Quente	Totais
A	8,83	10,90	11,01	19,22
B	9,80	11,70	8,93	17,36
C	14,25	10,85	7,08	18,15
D	10,18	13,64	12,13	20,69
E	10,96	10,67	9,05	21,84
Totais	12,78	12,28	10,98	

	Soma dos quadrados	Graus de liberdade	Quadrados médios	Razão F	Valor P
Marca	7702,44	4	1925,61	16,47	< 0,0001
Temp	77086,00	2	38543,00	329,58	< 0,0001
Interação	1999,97	8	250,00	2,14	0,0325
Erro	33329,13	285	116,94		
Total	120117,53	299			

Os três itens superiores são medidas de resumo (tamanhos de amostra, médias de amostra e desvios padrão de amostra) das várias combinações de nível de tratamento. A parte inferior do output mostra a tabela ANOVA. Há três valores p importantes nesta tabela: dois para “efeitos principais” e um para “interações”. Os principais efeitos indicam se há diferenças de médias significativas nos níveis de um dos fatores, em relação à média dos níveis do outro fator. Por exemplo, o efeito principal Temperatura indica se os valores nas células B17 a D17 são significativamente diferentes. (Eles são, como indica o valor p muito baixo para *Temperatura* na tabela ANOVA.)

Dados ausentes e conectar aos dados

- **Dados ausentes** – em razão da necessidade de um design balanceado, não deve haver dados ausentes.
- **Conectar aos dados** – todas as fórmulas da ANOVA são vinculadas a dados. Se os dados forem alterados, os resultados mudam automaticamente. O balanço do experimento é verificado quando a análise é executada. Alterações nos dados podem afetar o balanço e gerar resultados inválidos.

Comando Teste de independência de qui-quadrado

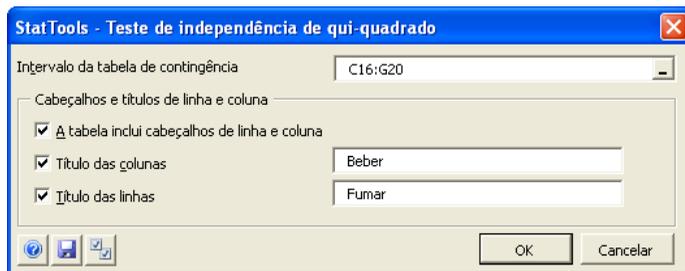
Testa a independência entre os atributos de linha e coluna de uma tabela de contingência

O comando Teste de independência de qui-quadrado testa a independência entre os atributos de linha e coluna de uma tabela de contingência. Por exemplo, se uma tabela de contingência lista contagens de pessoas em diferentes categorias de cigarros e bebidas, o procedimento testa se o hábito de fumar é independente do de beber. A tabela de contingência (também chamada “cross-tabs”) poderia ser uma tabela dinâmica do Excel.

Este procedimento é um pouco diferente de todos os procedimentos do StatTools. Para esta análise, basta uma tabela de contingência retangular. Cada célula nesta tabela deve ser uma contagem de observações em uma combinação específica de linha/coluna (abstêmios e fumantes inveterados, por exemplo). A tabela pode ter rótulos de linha e coluna (cabeçalhos) e/ou totais de linha e coluna, mas estes não são necessários e são utilizados apenas para aumentar a clareza dos relatórios do StatTools.

Caixa de diálogo Teste de independência de qui-quadrado

Esta análise é definida com a caixa de diálogo **Teste de independência de qui-quadrado**:



As opções da caixa de diálogo Teste de independência de qui-quadrado incluem:

- **Cabeçalhos e títulos de linha e coluna** – determina uma das seguintes opções: **A tabela inclui cabeçalhos de linha e coluna**, ou cabeçalhos na coluna mais à esquerda e linha superior da tabela; **Título das colunas**, ou o título que você usa para representar as colunas na tabela; **Título das linhas**, ou o título que você quer usar para representar as linhas da tabela.

Relatório de Teste de independência de qui-quadrado

Relatório do StatTools

Análise: Teste de independência de qui-quadrado
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Tempo real

Linhas: Fumar / Colunas: Beber					
Contagens originais	Nunca	Pouco	Médio	Muito	Total
Nunca	20	15	10	10	55
Às vezes	10	15	12	15	52
Médio	8	12	15	20	55
Muito	6	10	15	25	56
Total	44	52	52	70	218

Linhas: Fumar / Colunas: Beber					
Porcentagem de linhas	Nunca	Pouco	Médio	Muito	Total
Nunca	36,36%	27,27%	18,18%	18,18%	100,00%
Às vezes	19,23%	28,85%	23,08%	28,85%	100,00%
Médio	14,55%	21,82%	27,27%	36,36%	100,00%
Muito	10,71%	17,86%	26,79%	44,64%	100,00%

Linhas: Fumar / Colunas: Beber					
Porcentagem de colunas	Nunca	Pouco	Médio	Muito	Total
Nunca	45,45%	28,85%	19,23%	14,29%	100,00%
Às vezes	22,73%	28,85%	23,08%	21,43%	100,00%
Médio	18,18%	23,08%	28,85%	28,57%	100,00%
Muito	13,64%	19,23%	28,85%	35,71%	100,00%

Linhas: Fumar / Colunas: Beber					
Contagens esperadas	Nunca	Pouco	Médio	Muito	Total
Nunca	11,1009	13,1193	13,1193	17,6606	55
Às vezes	10,4954	12,4037	12,4037	16,6972	52
Médio	11,1009	13,1193	13,1193	17,6606	55
Muito	11,3028	13,3578	13,3578	17,9817	56

Linhas: Fumar / Colunas: Beber					
Distância do esperado	Nunca	Pouco	Médio	Muito	Total
Nunca	7,1340	0,2696	0,7416	3,3229	11,4681
Às vezes	0,0234	0,5435	0,0131	0,1725	0,7525
Médio	0,8662	0,0955	0,2696	0,3099	1,5412
Muito	2,4878	0,8441	0,2019	2,7393	6,2729

Estatística de qui-quadrado

Qui-quadrado	20,0349
Valor P	0,0177

O relatório acima mostra o resultado básico do teste, um valor p. Se este valor p for baixo (como neste caso), podemos concluir que os atributos de linha e coluna *não* são independentes. Podemos estudar os números desta planilha para compreender melhor como o fumo e a bebida estão relacionados.

Dados ausentes e conectar aos dados

- **Dados ausentes** – Não deve haver dados ausentes nas células da tabela de contingência.
- **Conectar aos dados** – as fórmulas do relatório de Teste de independência de qui-quadrado são vinculadas aos dados. Assim, se a contagem da tabela de contingência original mudar, os outputs dessa planilha também mudarão.

Menu Testes de normalidade

Pelo fato de tantos procedimentos estatísticos pressuporem que um conjunto de dados é normalmente distribuído, convém ter métodos para confirmar esta pressuposição. O StatTools oferece três verificações comumente utilizadas, descritos nesta seção.

Comando Teste de normalidade de qui-quadrado

Testa se os dados observados de uma variável está normalmente distribuído

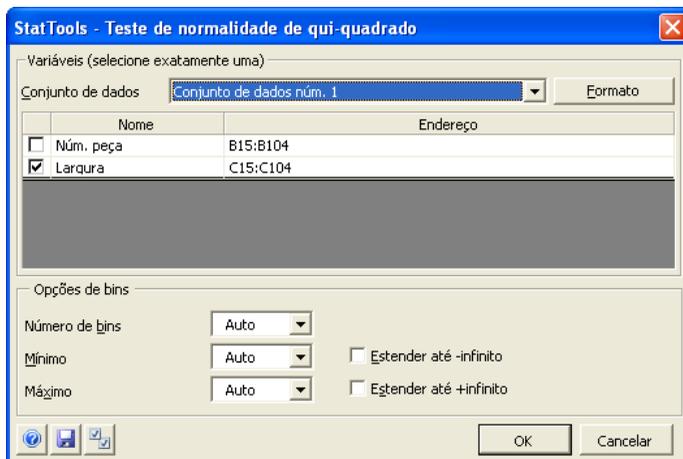
O procedimento de Teste de normalidade de qui-quadrado usa um teste de adequação de ajuste de qui-quadrado para verificar se os dados observados de uma variável específica poderiam ser provenientes de uma distribuição normal. Para fazer isso, ele cria um histograma desta variável, usando as categorias que você especificar, e sobrepõe um histograma de distribuição normal ao histograma dos dados. Se os dois histogramas tiverem basicamente a mesma forma, não podemos rejeitar a hipótese nula de um ajuste (*fit*) normal.

O teste formal é executado comparando-se as contagens observadas nas várias categorias com as contagens esperadas que são baseadas na pressuposição de normalidade. Na verdade, o procedimento permite que você teste a normalidade de diversas variáveis (separadamente). Um histograma é criado para cada variável selecionada, e o teste de qui-quadrado é executado em cada um deles.

A única exigência do teste de normalidade de qui-quadrado é que deve haver pelo menos uma variável numérica. Além disso, a maioria dos analistas sugere que deve haver pelo menos 100 observações – quanto mais, melhor.

**Caixa de diálogo
Teste de
normalidade de
qui-quadrado**

Esta análise é definida com a caixa de diálogo **Teste de normalidade de qui-quadrado**:

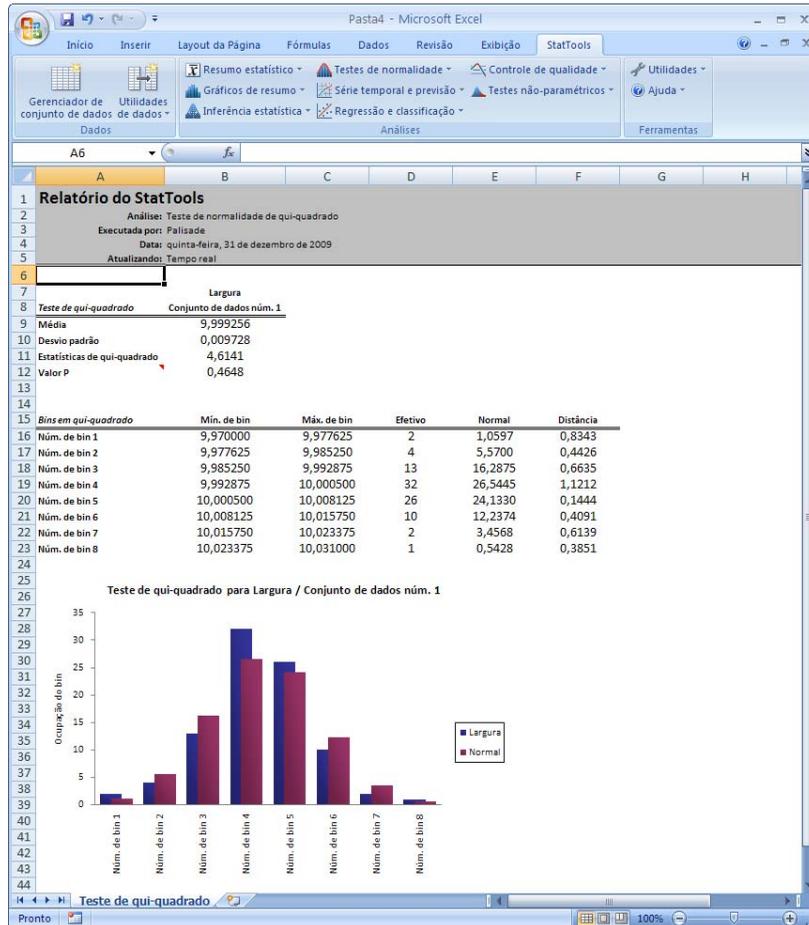


Uma variável pode ser selecionada para teste. O conjunto de dados selecionado pode ser de dados empilhados ou desempilhados.

As opções da caixa de diálogo **Teste de normalidade de qui-quadrado** incluem:

- **Número de bins** - especifica um número fixo de bins ou que o número de bins será automaticamente calculado para você.
- **Mínimo e máximo - Auto** especifica o mínimo e o máximo do conjunto de dados que serão utilizados para calcular o mínimo e o máximo de bins de intervalos iguais. O primeiro e o último bins, contudo, podem ser adicionados usando-se as opções **Estender até -infinito** e **Estender até +infinito**. Se **Auto** não for selecionado, você pode especificar um valor **Mínimo** e um **Máximo** para o início e o término dos bins. Dessa forma é possível especificar um intervalo específico para a execução dos bins independentemente dos valores mínimo e máximo do conjunto de dados.
- **Estender até -infinito** indica que o primeiro bin usado se estenderá desde o mínimo especificado até -infinito. Todos os outros bins terão o mesmo comprimento. Em determinadas circunstâncias, isso melhorará os teste de conjuntos de dados como limites inferiores desconhecidos.
- **Estender até +infinito** indica que o último bin usado se estenderá desde o máximo especificado até +infinito. Todos os outros bins terão o mesmo comprimento. Em determinadas circunstâncias, isso melhorará os teste de conjuntos de dados como limites superiores desconhecidos.

Relatório de teste de normalidade de qui-quadrado



Os resultados do teste são mostrados no relatório acima. O valor p de 0,4648 é um bom indício de que os valores estão distribuídos normalmente. Mais indícios deste efeito aparece nos histogramas da Figura 4 e dados de frequência. Contudo, atenção a dois fatores. Primeiro, se houver muito poucas observações (digamos, muito menos que 100), o teste de qui-quadrado não é eficiente para distinguir normalidade de não normalidade. O efeito é que o valor p normalmente não é baixo o bastante para rejeitar a hipótese de normalidade. Em essência, quase *tudo* tende a parecer normalizado com conjuntos de dados pequenos. Por outro lado, se o conjunto de dados for realmente grande (digamos, várias centenas de observações), o valor p normalmente será pequeno, indicando a não normalidade. O motivo é que, no caso de grandes conjuntos de dados, toda “protuberância” na curva tem grandes chances de criar um valor p baixo. Neste caso, o teste real é *prático*: os histogramas realmente diferem tanto para finalidades práticas?

***Dados ausentes
e conectar aos
dados***

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – o histograma e todas as fórmulas do teste são vinculados aos dados originais. Assim, se os dados mudarem, o histograma e os resultados de teste mudam automaticamente.

Comando Teste de Lilliefors

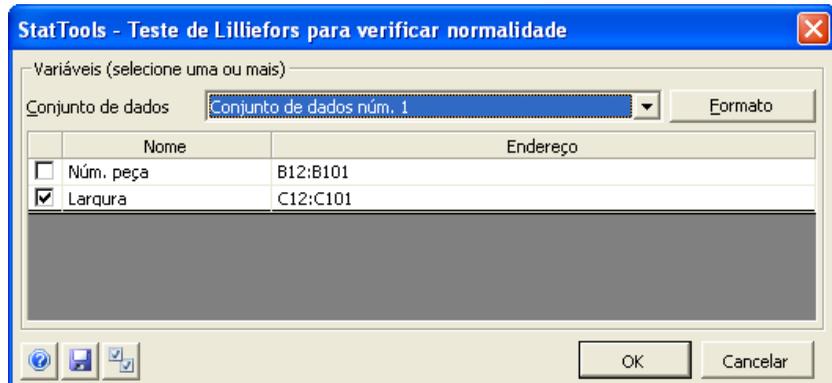
Testa se os dados observados de uma variável está normalmente distribuído

O procedimento do Teste de Lilliefors oferece um teste mais poderoso de normalidade que o teste de adequação de ajuste qui-quadrado. (Mais poderoso significa que há maior probabilidade de detectar não normalidade, se houver.) Baseia-se em uma comparação da “*fdc* empírica” e uma *fdc* normal, em que “*fdc*” é a sigla de função de distribuição cumulativa, mostrando a probabilidade de ser menor ou igual que algum valor específico.

A *fdc* empírica é baseada em dados. Por exemplo, se houver 100 observações e a 13ª mais baixa for 137, então a *fdc* empírica, avaliada em 137, será 0,13. O teste de Lilliefors encontra a distância vertical máxima entre a *fdc* empírica e a *fdc* normal, e compara este máximo com valores tabulados (com base no tamanho da amostra). Se a distância vertical máxima observada for suficientemente grande, não teremos evidências de que os dados *não* provêm de uma distribuição normal.

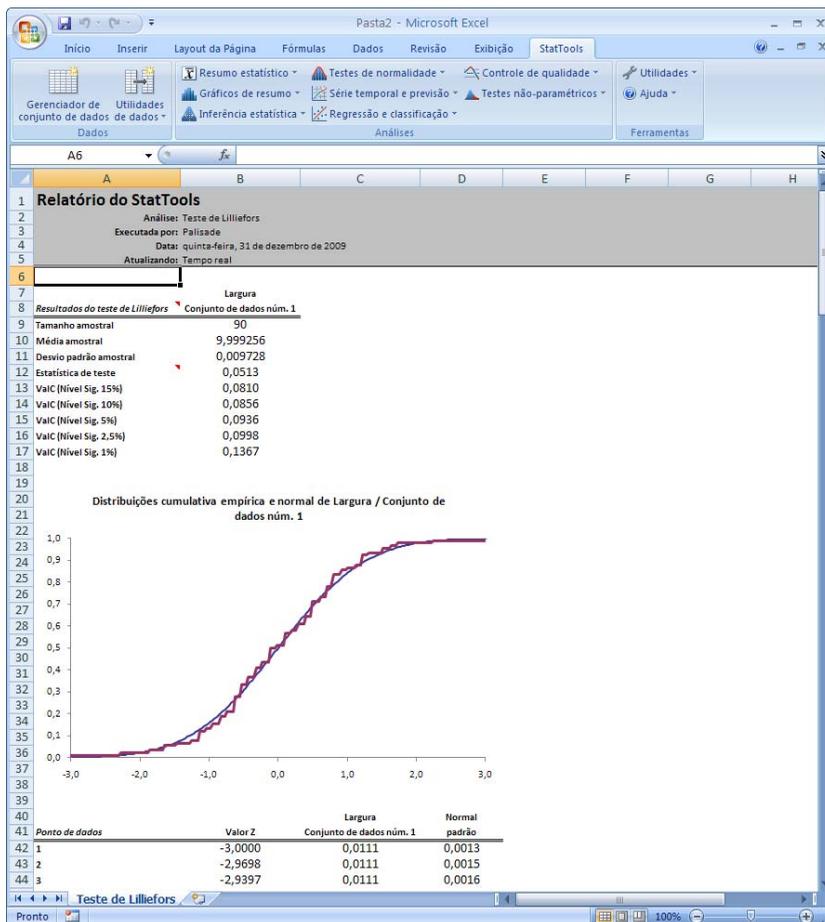
Esta análise é configurada na caixa de diálogo **Teste de Lilliefors para verificar normalidade**:

**Caixa de diálogo
Teste de Lilliefors
para verificar
normalidade**



Uma ou mais variáveis podem ser selecionadas para os testes. O conjunto de dados selecionado deve estar no formato desempilhado. As variáveis podem pertencer a conjuntos de dados diferentes.

Relatório de Teste de Lilliefors



Os resultados do teste são mostrados no relatório acima. Não há valor p (como na maioria dos testes de hipótese), mas podemos perceber pela declaração que a distância vertical máxima é suficientemente grande para colocar em dúvida a pressuposição de normalidade. Mais evidências deste efeito aparecem nas fdc's do gráfico incluído. Na verdade, o ajuste entre as duas curvas parece ser "bastante bom" e pode ser bom o suficiente para finalidades práticas. Em outras palavras, podemos concluir que estes dados estão "próximos o bastante" para serem normalmente distribuídos de acordo com as nossas finalidades.

***Dados ausentes
e conectar aos
dados***

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – as FDCs e todas as fórmulas do teste são vinculadas aos dados originais. Assim, se os dados mudarem, o gráfico e os resultados de teste mudam automaticamente.

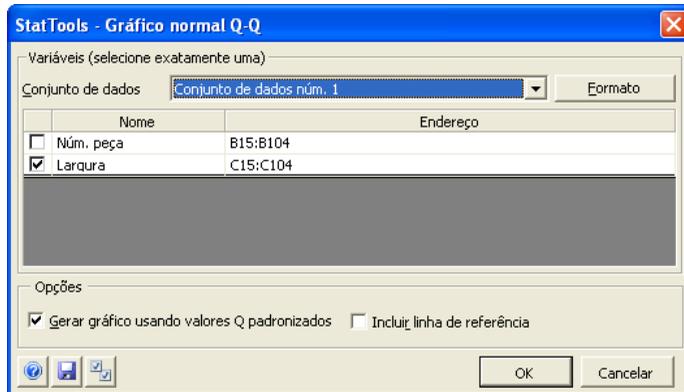
Comando Gráfico normal Q-Q

Testa se os dados observados de uma variável está normalmente distribuído

O comando Gráfico normal Q-Q cria um gráfico quantil-quantil (Q-Q) para uma variável simples. Proporciona um teste informal de normalidade. Embora os detalhes sejam um tanto complexos, o objetivo é bem simples: comparar os quantis (ou percentis) dos dados com os quantis de uma distribuição normal. Se os dados forme normal em essência, os pontos no gráfico Q-Q devem estar alinhados a aproximadamente 45 graus de inclinação. Contudo, uma curvatura nítida no gráfico é indicação de alguma forma de não normalidade (assimetria, por exemplo).

Caixa de diálogo Gráfico normal Q-Q

Esta análise é configurada na caixa de diálogo **Gráfico normal Q-Q**:

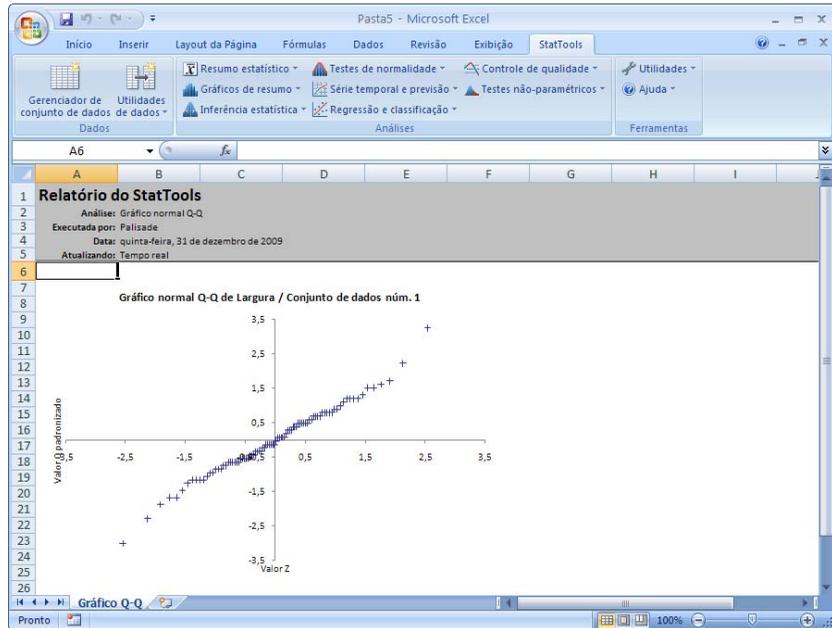


Uma variável pode ser selecionada para a criação do gráfico. O conjunto de dados selecionado deve estar no formato desempilhado.

As opções na caixa de diálogo **Gráfico normal Q-Q** incluem:

- **Gerar gráfico usando valores Q padronizados** – especifica o uso de um valor Q padronizado, em vez de dados Q-Q, no eixo Y do gráfico. Isso leva a comparações dos valores do eixo Y entre os gráficos normais Q-Q possíveis.

Relatório Gráfico normal Q-Q



Como foi afirmado antes, trata-se de um teste *informal* de normalidade. É difícil dizer “quão próximo” de uma linha de inclinação de 45 graus o gráfico deve ser para se aceitar a pressuposição de normalidade. Em geral, buscamos uma curvatura óbvia no gráfico, e não há nenhuma neste gráfico.

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – os gráficos e todas as fórmulas são vinculados aos dados originais. Assim, se os dados forem alterados, o gráfico muda automaticamente.

Menu Série temporal e previsão

Os procedimentos do menu Série temporal e previsão tratam da análise de dados coletados ao longo do tempo com aplicações de previsão e controle de qualidade. Os métodos de previsão fornecidos incluem dados de séries temporais usando método de médias móveis, suavização exponencial simples, método de suavização exponencial de Holt para a captura de tendências, e método de suavização exponencial de Winters para a captura da sazonalidade.

Comando Gráfico de série temporal

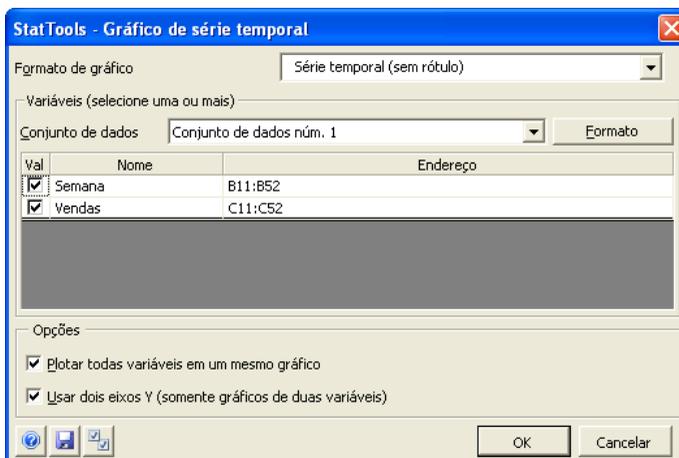
Cria um gráfico de série temporal para variáveis

O comando Gráfico de série temporal inclui uma ou mais variáveis de série temporal no mesmo gráfico. Se duas variáveis forem selecionadas, você terá a opção de usar escalas iguais ou diferentes no eixo Y das duas variáveis. A segunda opção é útil quando os intervalos de valores das duas variáveis forem consideravelmente diferentes. Contudo, se mais de duas variáveis forem incluídas no gráfico, elas devem compartilhar da mesma escala vertical.

Deve haver ao menos uma variável numérica no conjunto de dados. Também pode haver uma variável de “dados”, mas se ela tiver de ser usada para identificar o eixo horizontal do gráfico, deve ser selecionada como variável de “rótulo”.

Caixa de diálogo Gráfico de série temporal

Este tipo de gráfico é configurado na caixa de diálogo **Gráfico de série temporal**:

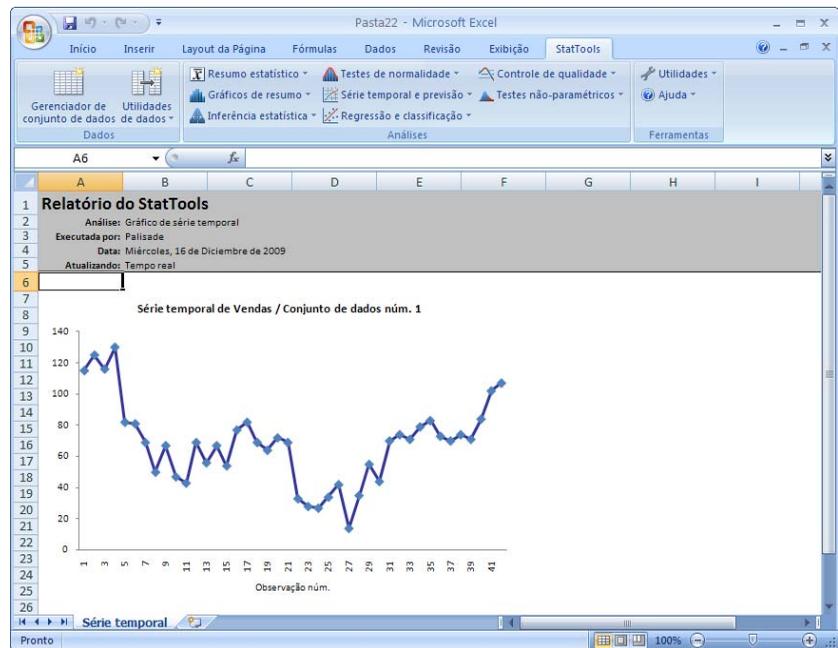


Uma ou mais variáveis podem ser selecionadas para a geração do gráfico. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes. A variável de rótulo (caixa de seleção Rótulo) aparece no eixo X.

As opções da caixa de diálogo **Gráfico de série temporal** incluem:

- Plotar todas variáveis em um mesmo gráfico. Seleciona a criação do gráfico com todas as variáveis.
- **Usar dois eixos Y.** Exibe um eixo Y separado para cada variável em um gráfico de duas variáveis. Unidades e valores de cada variável podem, assim, ser exibidos no gráfico.

Gráfico de série temporal de duas variáveis



Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – os gráficos são vinculados aos dados, assim, se os dados mudares o gráfico é automaticamente atualizado.

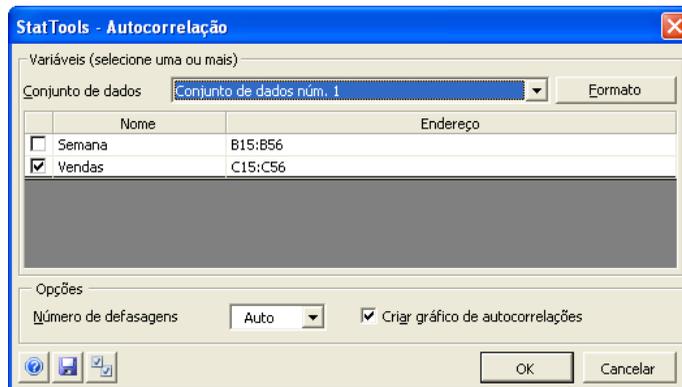
Comando Autocorrelação

Calcula as autocorrelações de variáveis

O comando Autocorrelação calcula as autocorrelações de qualquer variável numérica selecionada. Em geral, esta variável será de série temporal, embora o StatTools faz os cálculos de qualquer variável. Você pode selecionar o número de autocorrelações que desejar (ou seja, o número de defasagens). Também é possível gerar um gráfico das autocorrelações, que se chama correlograma. O output mostra quais autocorrelações, se houver, são significativamente diferentes de 0.

Caixa de diálogo Autocorrelação

Este tipo de gráfico é configurado na caixa de diálogo **Autocorrelação**:

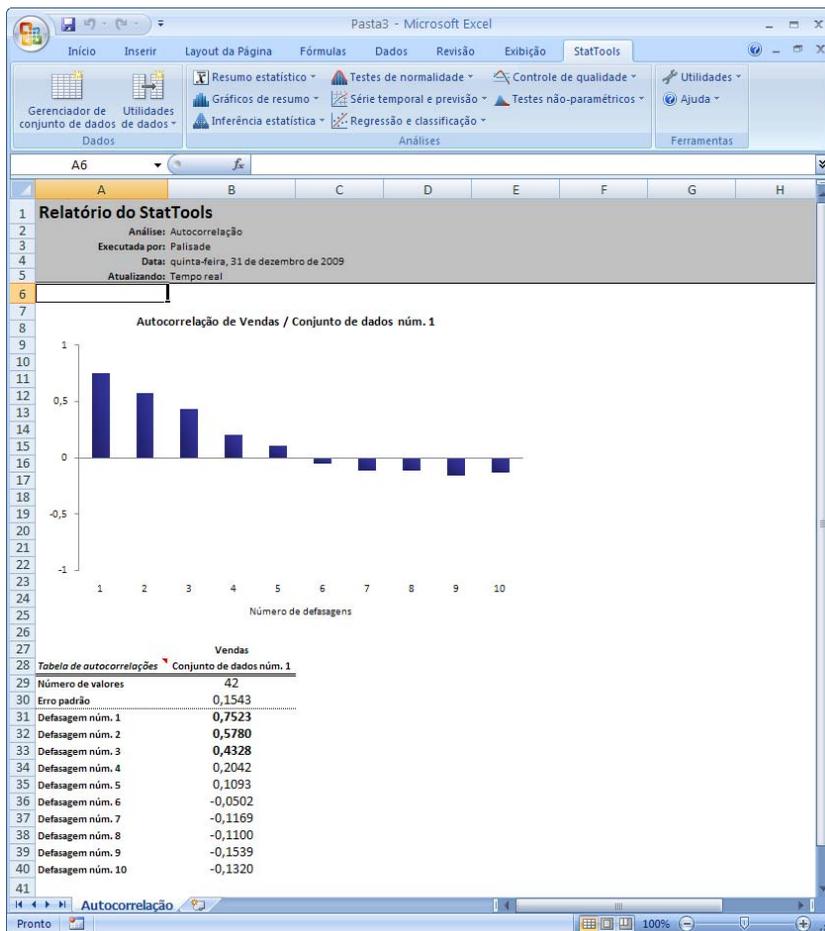


Uma ou mais variáveis podem ser selecionadas para análise. Os dados do conjunto selecionado não devem estar empilhados.

As opções da caixa de diálogo **Autocorrelação** incluem:

- **Número de defasagens.** O número de períodos para a defasagem no cálculo das autocorrelações. Se **Auto** for selecionado, o StatTools determina o número apropriado de defasagens a testar. Se você especificar um número específico de defasagens, o número máximo de defasagens possível é de 25% do número de observações na série. Por exemplo, se você tiver 80 valores mensais, poderá solicitar 20 defasagens.
- **Criar gráfico de autocorrelações.** Cria um gráfico de barras com a altura de cada barra igual à autocorrelação correspondente.

Relatório de autocorrelação



O relatório de autocorrelação está acima. Para cada defasagem, a autocorrelação correspondente aparece, junto com um erro padrão *aproximado*.

Dados ausentes e conectar aos dados

- **Dados ausentes** – este procedimento permite dados ausentes no início da série temporal, mas nenhum no meio ou no final da série.
- **Conectar aos dados** – o StatTools vincula o output aos dados. Portanto, se os dados forem alterados, as autocorrelações (e o correlograma) serão automaticamente atualizadas.

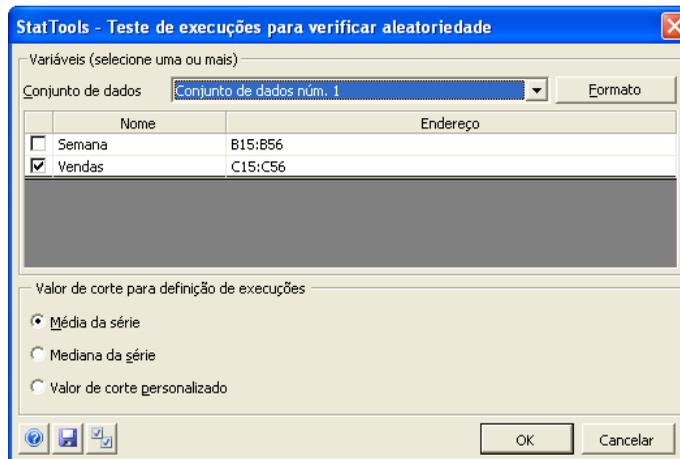
Comando Teste de execuções para verificar aleatoriedade

Executa um teste de execuções para verificar se uma variável é aleatória

O comando Teste de execuções para verificar aleatoriedade permite a verificação da “aleatoriedade” de uma sequência de valores em uma variável, normalmente uma variável de série temporal. Indica quantas “execuções” há em uma sequência, quando uma execução é um número consecutivo de valores em um lado ou outro de algum ponto de corte (com a média ou a mediana da sequência). Para uma sequência aleatória, espera-se que não haja poucas execuções nem muitas execuções. O teste de execuções conta o número de execuções e em seguida informa um valor p para o teste. Se o valor p for baixo, concluímos que provavelmente a sequência não é aleatória, ou seja, há muitas execuções ou muito poucas execuções.

Caixa de diálogo Teste de execuções para verificar aleatoriedade

Esta análise é definida com a caixa de diálogo **Teste de execuções para verificar aleatoriedade**:

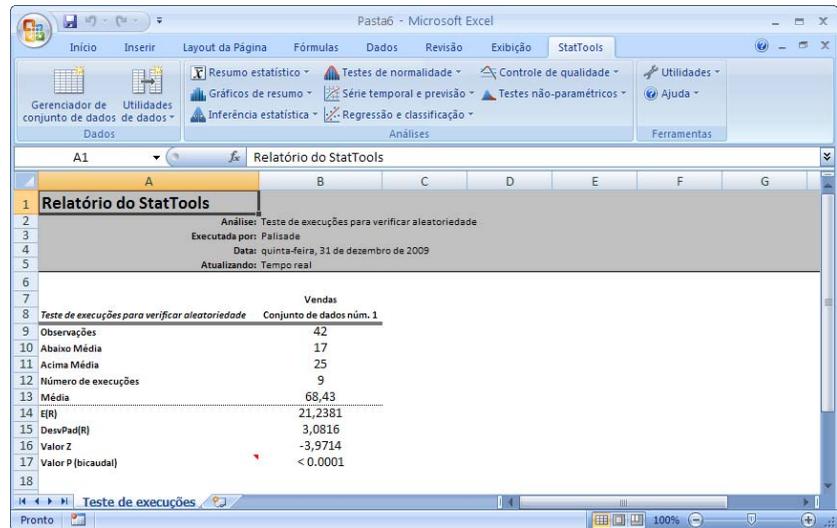


Uma ou mais variáveis podem ser selecionadas para análise. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Teste de execuções para verificar aleatoriedade** incluem:

- **Valor de corte para definição de execuções** – Um teste de execuções sempre é baseado nas execuções acima ou abaixo de um valor de corte. Este valor pode ser **Média da série**, **Mediana da série** ou qualquer outro **Valor de corte personalizado** que você queira especificar.

Relatório Teste de execuções para verificar aleatoriedade



Teste de execuções para verificar aleatoriedade	Vendas
Observações	42
Abaixo Média	17
Acima Média	25
Número de execuções	9
Média	68,43
E(R)	21,2381
DesvPad(R)	3,0816
Valor Z	-3,9714
Valor P (bicaudal)	< 0.0001

O relatório acima mostra o número de execuções e o número esperado de execuções sob aleatoriedade E(R). Como 20 é significativamente abaixo de 24,8333, esta série não é integralmente aleatória; ou seja, as vendas não têm um comportamento de “zigzague” como uma série aleatória.

Dados ausentes e conectar aos dados

- **Dados ausentes** – este procedimento permite dados ausentes no início e no final da série temporal, mas nenhum no meio da série.
- **Conectar aos dados** – o StatTools vincula o output aos dados. Portanto, se os dados forem alterados, os relatórios mudam automaticamente.

Comando Previsão

Gera previsões para variáveis de série temporal

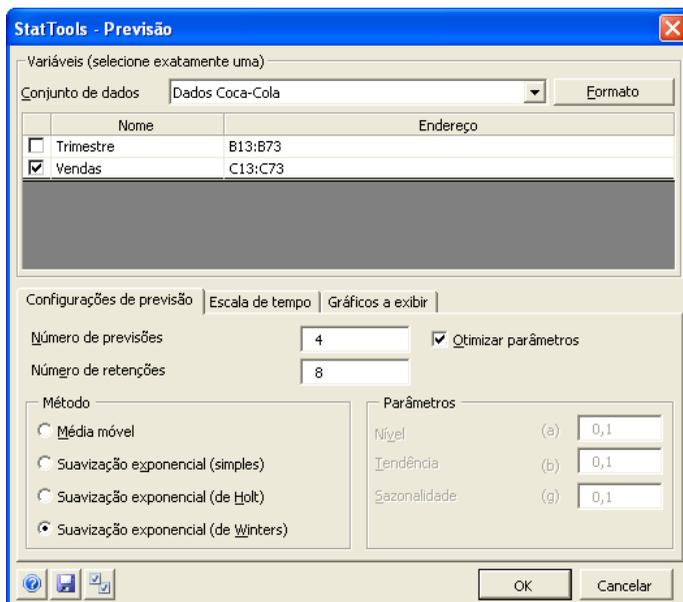
O comando Previsão proporciona vários métodos para prever uma variável de série temporal. Esses métodos incluem dados de séries temporais usando método de médias móveis, suavização exponencial simples, método de suavização exponencial de Holt para a captura de tendências, e método de suavização exponencial de Winters para a captura da sazonalidade. O comando Previsão também permite a eliminação inicial da sazonalidade dos dados, usando o método razão-médias móveis e um modelo de sazonalidade multiplicativa. Em seguida, você poderá usar qualquer um dos métodos de previsão (outros além que não o método de Winters) para prever os dados sem sazonalidade, e, por fim, “sazonalizar” as previsões de volta às unidades originais.

Os relatórios de previsão incluem um conjunto de colunas para a exibição dos vários cálculos (por exemplo, tendências e níveis suavizados do método de Holt, fatores sazonais do método razão-médias móveis e assim por diante), as previsões e os erros de previsão. Medidas de resumo também são incluídas (MAE, RMSE e MAPE) para o rastreamento do ajuste do modelo em relação aos dados observados. (Quando você usa os métodos de suavização exponencial, tem a opção de usar a otimização para encontrar a(s) constante(s) de suavização que minimizam o RMSE.)

Por fim, diversos gráficos de série temporal estão disponíveis, incluindo um gráfico da série original, um gráfico da série com previsões sobrepostas e um gráfico dos erros de previsão. No caso da eliminação da sazonalidade, os gráficos estão disponíveis para a série original e a série sem sazonalidade.

Caixa de diálogo Previsão

As previsões são definidas na caixa de diálogo **Previsão**:



Uma variável pode ser selecionada para a análise. Os dados do conjunto selecionado não devem estar empilhados.

Caixa de diálogo Previsão - Configurações de previsão

As opções de Configurações de previsão especificam o método de previsão usado e as definições do método selecionado. As opções são:

- **Número de previsões.** Especifica o número de períodos no *futuro* aos quais serão feitas previsões.
- **Número de retenções.** Especifica o número de observações a serem “retidas”, ou não utilizadas, no modelo de previsão. Você pode optar por usar todas as observações para estimar o modelo de previsão (0 retenção), ou reter algumas para validação. Em seguida, o modelo é estimado a partir das observações não retidas, e é utilizado para prever as observações retidas.
- **Otimizar parâmetros** (somente métodos de suavização exponencial). Localiza a constante de suavização que minimiza o RMSE (para o período de não retenção). A otimização exige que os parâmetros mostrados na caixa de diálogo estejam entre 0 e 1. Se você editar valores de parâmetro diretamente em um relatório de previsão, assegure-se de inserir valores dentro deste intervalo.

- **Dessazonalizar.** Determina a dessazonalização de dados antes da previsão. Para dados sazonais, ou seja, dados suspeitos de apresentar um padrão sazonal, há duas opções. Você pode usar o método de Winters, que lida diretamente com a sazonalidade, ou selecionar esta opção para dessazonalizar os dados antes, usando o método de razão-média móvel para a dessazonalização. Então *qualquer* método poderá ser usado para prever a série dessazonalizada.
- **Método** – seleciona o método de previsão a ser usado: **Média móvel** ou métodos de suavização exponencial **simples**, **Holt** ou **Winters**.
- **Parâmetros** – especifica os parâmetros a serem usados para o método de previsão selecionado:
 - **Intervalo** (apenas para o método de médias móveis), ou o número de observações consecutivas usado em cada média móvel.
 - **Nível** (todos os métodos de suavização exponencial), um parâmetro de suavização que pode assumir qualquer valor entre 0 e 1 (o valor padrão é 0,1).
 - **Tendência** (métodos de suavização exponencial de Holt e Winter), um segundo parâmetro de suavização que pode assumir qualquer valor entre 0 e 1 (o valor padrão é 0,1).
 - **Sazonalidade** (apenas o método de suavização exponencial de Winter), um terceiro parâmetro de suavização que pode assumir qualquer valor entre 0 e 1 (o valor padrão é 0,1).

Nota: se Otimizar parâmetros for selecionado, os parâmetros Nível, Tendência e Sazonalidade não poderão ser definidos, uma vez que seus valores não podem ser otimizados.

**Caixa de diálogo
Previsão - Escala
de tempo**

As opções de escala de tempo especificam os rótulos de escala de tempo e tempos da variável analisada. As opções são:

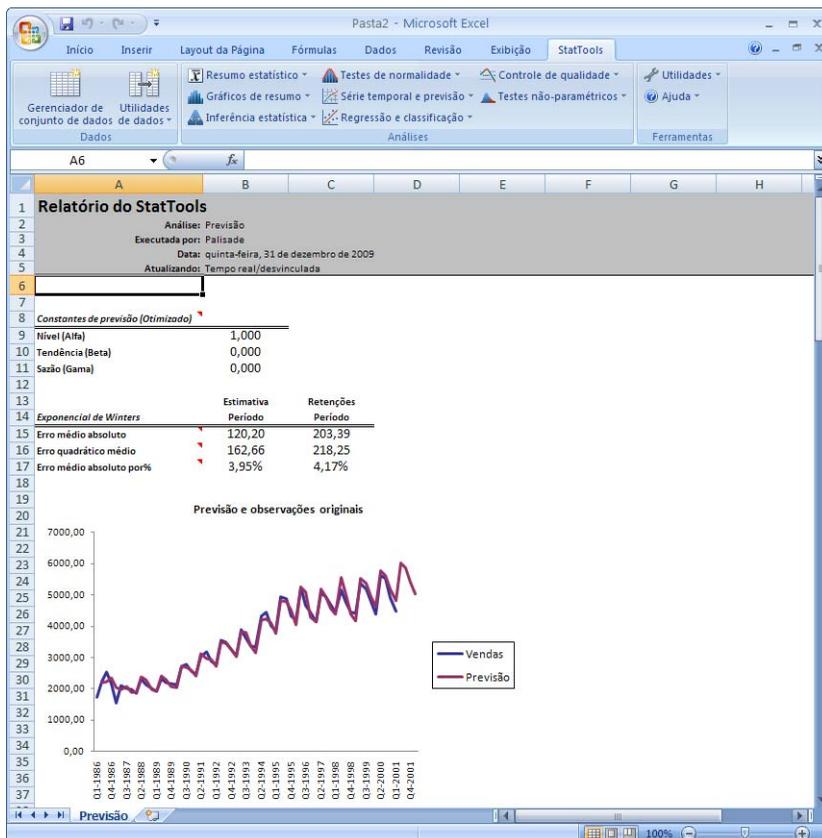
- **Período sazonal.** Especifica o tipo de dados de série temporal: Anual, Trimestral, Mensal, Semanal, Diária ou Nenhum. Usado para a sazonalização de dados e para rotulagem.
- **Estilo do rótulo.** Especifica como a escala de tempo será rotulada em todos os gráficos gerados.
- **Rótulo inicial.** Especifica a entrada do primeiro rótulo de escala de tempo no gráfico.

Caixa de diálogo Previsão - Gráficos a exibir

As opções de gráfico especificam os gráficos de previsão que podem ser gerados. Os gráficos disponíveis incluem:

- 1) **Sobreposição de previsão**, ou o gráfico de série temporal dos valores de dados gerados pela previsão
- 2) **Série original**, ou o gráfico de série temporal dos dados reais
- 3) **Erros de previsão**, ou o erro entre dados de previsão e reais
- 4) **Sobreposição de previsão dessazonalizado**, ou o gráfico de série temporal dos valores de dados gerados pela previsão após a dessazonalização dos dados originais
- 5) **Série original dessazonalizada**, ou o gráfico de série temporal dos dados reais após a dessazonalização dos dados originais
- 6) **Erros de previsão dessazonalizados**, ou o erro entre dados de previsão e reais após a dessazonalização dos dados originais

Exemplo de relatório de previsão



***Dados ausentes
e conectar aos
dados***

- **Dados ausentes** – dados ausentes são permitidos no *início* da série temporal, mas não no meio ou no final da série.
- **Conectar aos dados** – devido aos demorados cálculos exigidos, as previsões não são vinculadas aos dados. Se forem feitas alterações nos dados originais, o procedimento deverá ser re-executado.

Menu Regressão e classificação

Os comandos do menu “Regressão e classificação” executam análises de regressão e classificação. As análises de regressão disponíveis incluem múltiplo simples, passo a passo, prospectiva, retrospectiva e bloco. Outras análises do menu de regressão e classificação incluem análise discriminante e regressão logística.

Comando Regressão

Executa análise de regressão em um conjunto de variáveis

O comando de regressão realiza várias análises de regressão incluindo múltiplo simples, passo a passo, prospectiva, retrospectiva e bloco. Os relatórios de cada análise incluem medidas de resumo de cada execução da equação de regressão, uma tabela ANOVA para cada regressão, e uma tabela de coeficientes de regressão estimados, seus erros-padrão, seus valores t, valores p e seus intervalos de 95% de confiança de cada regressão.

Além disso, você tem a opção de criar duas novas variáveis, os valores ajustados e os residuais, criando vários gráficos de dispersão diagnósticos.

Tipos de regressão disponíveis

Os tipos de regressão disponíveis incluem múltiplo simples, passo a passo, prospectiva, retrospectiva e bloco. O procedimento de regressão múltipla simples constrói uma equação no ato, usando as variáveis explanatórias selecionadas. Os outros procedimentos permitem variáveis (ou blocos de variáveis) para entrar ou sair da equação de modo sequencial. Especificamente, o procedimento passo a passo permite que as variáveis sejam inseridas uma de cada vez. A próxima variável a incluir é aquela mais altamente correlacionada com a parte inexplicada da variável de resposta. Entretanto, a opção passo a passo também permite a saída das variáveis (desde que tenham sido inseridas), caso já não contribuam de modo significativo. O procedimento prospectivo é idêntico ao passo a passo, mas as variáveis não podem sair depois que tenham sido inseridas. O procedimento retrospectivo tem início com todas as possíveis variáveis explanatórias da equação e, depois, ele as exclui, uma por vez, caso já não contribuam de modo significativo. Por fim, o procedimento de bloco permite que blocos de variáveis explanatórias sejam inseridos (ou não) *em bloco*, na ordem especificada. Se um bloco não for significativo e não entrar, blocos futuros não serão considerados para inserção.

Caixa de diálogo Regressão

Essas análises são configuradas na caixa de diálogo **Regressão**:

StatTools - Regressão

Tipo de regressão: **Múltiplos**

Variáveis (selecione uma ou mais independentes e uma dependente)

Conjunto de dados: **Conjunto de dados núm. 1** Formato

I	D	Nome	Endereço
<input type="checkbox"/>	<input type="checkbox"/>	Mês	B20:B55
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Horas máq.	C20:C55
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Ciclos de prod.	D20:D55
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Despesas gerais	E20:E55

Parâmetros

Usar valores P Usar valores F

Valor P a inserir:

Valor P a retirar:

Definir constante como zero (origem)

Gráficos

Valores ajustados comparados aos valores Y efetivos

Valores ajustados comparados aos valores X

Valores residuais comparados aos valores ajustados

Valores residuais comparados aos valores X

Opções avançadas

Incluir informações detalhadas dos passos

Incluir previsão para o conjunto de dados: **Conjunto de dados núm. 1**

Nível de confiança: **95%**

OK Cancelar

Uma variável dependente (D) e uma ou mais variáveis independentes (I) precisam ser selecionadas para análises salvo no caso de uma regressão em bloco. Com a regressão em bloco, precisam ser selecionados uma variável dependente (D) e de um a sete blocos (B1 a B7). Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Regressão** incluem:

- **Tipo de regressão.** Selecione o tipo de regressão para executar: múltiplo simples, passo a passo, prospectiva, retrospectiva e bloco.

Os **Parâmetros** de regressão mudam de acordo com o Tipo de análise selecionada. As opções são:

- **Usar valores P:** afeta como as variáveis são incluídas ou removidas da equação de regressão. Depois de selecionada, você pode especificar um **Valor P a inserir** e/ou um **Valor P a retirar**, dependendo do método de regressão em uso. Quanto menor o valor de *p*, mais significativa deve ser a variável para entrar ou sair da equação de regressão. Os valores-padrão exibidos são normalmente aceitáveis. Lembre-se de duas coisas apenas. Primeiro: o valor de *p* para entrar não pode ser maior do que o valor de *p* para sair. Segundo: para facilitar a entrada das variáveis (e dificultar a saída), use valores de *p maiores*. Para dificultar a entrada das variáveis (e facilitar a saída), use valores de *p menores*. Os valores de *p* típicos ficam no intervalo de 0,01 a 0,1.
- **Usar valores F:** assim como com os valores de *p*, afeta o modo como as variáveis são incluídas ou removidas da equação de regressão. Permite que você especifique um **Valor F a inserir** e/ou um **Valor F a retirar**, dependendo do método de regressão em uso. Os valores típicos ficam no intervalo de 2,5 a 4.

O StatTools pode criar vários gráficos de dispersão opcionais, como exibidos nas opções **Gráficos** selecionados. Os gráficos são:

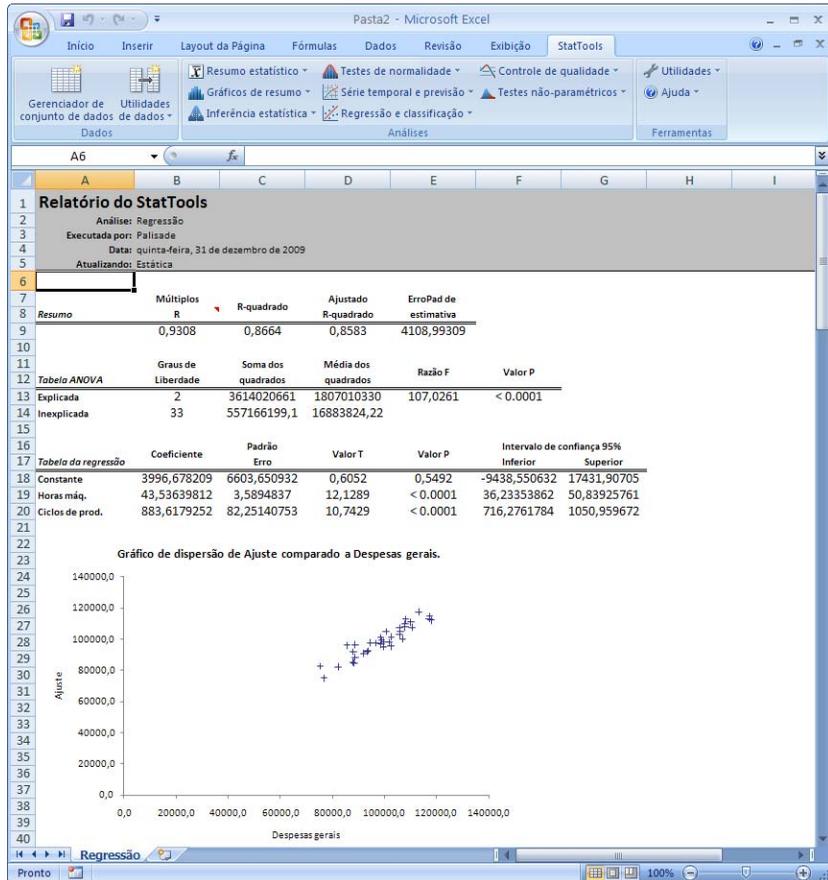
- **Valores ajustados comparados aos valores Y efetivos**
- **Valores ajustados comparados aos valores X**
- **Valores residuais comparados aos valores ajustados**
- **Valores residuais comparados aos valores X**

Esses gráficos são usados tipicamente em “análises residuais” para verificar se as pressuposições de regressão estão satisfeitas. Provavelmente, o gráfico mais útil será o dos resíduos (no eixo vertical) comparado aos valores ajustados (ou previstos) da variável de resposta.

As **Opções avançadas** para uma análise de regressão são:

- **Incluir informações detalhadas dos passos**, informando a r -quadrada e estatística de erro padrão da variável independente em cada passo intermediário da regressão.
- **Incluir previsão para o conjunto de dados**, na qual os valores previstos da variável dependente são gerados para os valores da variável independente de um segundo conjunto de dados. Esse conjunto de dados de previsão deve ter os mesmos nomes de variáveis do conjunto de dados original analisado pela regressão. Em geral, no conjunto de dados de previsão haverá conjuntos de valores das variáveis independentes para os quais deseja fazer a previsão do valor da variável dependente. A equação de regressão calculada a partir do primeiro conjunto de dados é usada para fazer as previsões. Os valores previstos para a variável dependente serão inseridos diretamente no conjunto de dados de previsão, preenchendo a coluna (ou linha) da variável dependente com os valores previstos. O **Nível de confiança** especifica os limites mínimo e máximo que serão gerados para os valores previstos.

Relatório da regressão passo a passo



Dados ausentes e conectar aos dados

- **Dados ausentes** – Se houver valores ausentes, qualquer linha com valores ausentes de *qualquer* das variáveis selecionadas será ignorada.
- **Conectar aos dados** – Não há vínculo com os dados originais. Se houver alteração nos dados, será preciso executar a análise novamente.

Comando Regressão logística

Executa regressão logística em um conjunto de variáveis

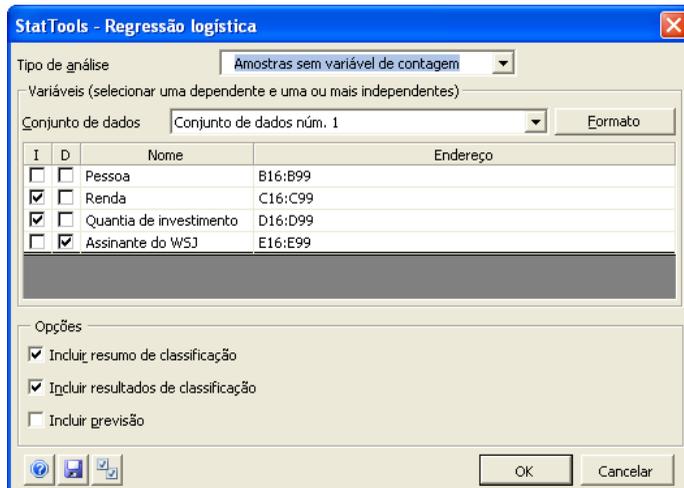
O comando Regressão logística executa uma análise de regressão logística em um conjunto de variáveis. Trata-se basicamente de um tipo não linear de análise de regressão em que a variável de resposta é binária: 0 ou 1. Deve haver uma variável de resposta 0-1 que especifique se cada observação é um “sucesso” ou “falha”, mais uma ou mais variáveis explicativas que podem ser usadas para estimar a probabilidade de sucesso.

A segunda opção da regressão logística é ter uma variável de “contagem” que especifique o número de tentativas observadas em cada combinação de variáveis explicativas. Assim, a variável de resposta deve indicar o número de tentativas que resultarão em “sucesso”. O resultado da regressão logística é uma equação de regressão similar a uma equação de regressão múltipla normal. Entretanto, ela deve ser interpretada de outra forma, como explicado a seguir.

O procedimento de regressão logística do StatTools baseia-se na otimização para localizar a equação de regressão. Esta otimização deve usar um algoritmo não linear complexo e, por isso, o procedimento pode demorar um pouco, dependendo da velocidade do PC.

Esta análise é configurada na caixa de diálogo **Regressão logística**:

Caixa de diálogo Regressão logística



Uma variável dependente ou de resposta (D) e uma ou mais variáveis independentes (I) precisam ser selecionadas para a análise. Os dados:

- 1) Precisam estar no formato “empilhado”, e ter também uma variável de resposta 0-1 que especifique se cada observação é um “sucesso” ou “falha”, aqui denominado **Amostras sem variável de contagem**.
- 2) Precisam ter uma variável de “contagem” e uma variável de resposta inteira, aqui denominada **Resumo das amostras (com variável de contagem)**.

As variáveis podem pertencer a conjuntos de dados diferentes. Se o Tipo de análise estiver configurado como **Resumo das amostras (com variável de contagem)**, será preciso selecionar outra variável de contagem.

As opções da caixa de diálogo **Regressão logística** são:

- **Tipo de análise**. Seleciona o tipo de regressão logística para executar **Amostras sem variável de contagem** ou **Resumo das amostras (com variável de contagem)**.

As **Amostras sem variável de contagem** têm uma variável de resposta 0-1 que especifica se cada observação é um “sucesso” ou “falha”, mais uma ou mais variáveis explicativas que podem ser usadas para estimar a probabilidade de sucesso.

A segunda opção da regressão logística é ter uma variável de “contagem”, que especifique o número de tentativas observadas em cada combinação de variáveis explicativas. Assim, a variável de resposta deve indicar o número de tentativas que resultarão em “sucesso”. Para este tipo de análise o sistema exibe no Seletor de variáveis uma coluna C, separada, que permite a seleção da variável de contagem.

- **Incluir resumo de classificação** com o relatório de regressão.
- **Incluir resultados de classificação** com o relatório de regressão.

- **Incluir previsão**, na qual os valores previstos da variável dependente são gerados para os valores da variável independente de um segundo conjunto de dados. Esse **conjunto de dados de previsão** deve ter os mesmos nomes de variáveis do conjunto de dados original analisado pela regressão. Em geral, no conjunto de dados de previsão haverá conjuntos de valores das variáveis independentes para os quais deseja fazer a **previsão** do valor da variável dependente. A equação de regressão calculada a partir do primeiro conjunto de dados é usada para fazer as previsões. Os valores previstos para a variável dependente serão inseridos diretamente no conjunto de dados de previsão, preenchendo a coluna (ou linha) da variável dependente com os valores previstos.

Relatório de regressão logística

Relatório do StatTools

Análise: Regressão logística
 Executada por: Falsade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Estática

Resumo de medidas

Desvio nulo	105,4941869
Desvio do modelo	26,83613587
Melhoria	78,65805103
Valor P	< 0,0001

Coefficientes de regressão

	Coefficiente	Erro padrão	Valor Wald	Valor P	Limite inferior	Limite superior	Exp(Coeff)
Constante	-7,930438797	3,10005134	-2,558163697	0,0105	-14,00653942	-1,854338171	0,000359629
Renda	-9,19596E-05	5,30141E-05	-1,734625165	0,0828	-0,000195867	1,1948E-05	0,999908045
Quantia de investimento	0,000351829	9,60677E-05	3,662306752	0,0002	0,000163537	0,000540122	1,000351891

Matriz de classificação

	1	0	Percentual Correta
1	23	4	85,19%
0	3	54	94,74%

Resumo de classificação

Correta	91,67%
Base	67,86%
Melhoria	74,07%

Probabilidades e classificações

	Renda	Quantia de investimento	Probabilidade	Classe de análise	Classe original
1	66400	26900	1,02%	0	0
2	68000	7100	0,00%	0	0
3	54900	21500	0,44%	0	0
4	50600	19300	0,30%	0	0
5	54100	16700	0,09%	0	0
6	78200	31900	1,99%	0	0

O relatório acima inclui os dados originais mais os dados usados para classificação. As classificações previstas, na coluna E, dependem de se as probabilidades estimadas de “sucesso”, na coluna D, estarem acima ou abaixo do valor de corte de 0,5 ou 50%. O relatório lista as estatísticas de resumo da regressão (algo similar à R-quadrada da regressão múltipla), informações detalhadas sobre a equação de regressão e os resultados de resumo do procedimento de classificação. (Existem muitos comentários nas células que ajudam o usuário a interpretar os resultados.) Neste exemplo, 90,5% das observações estão classificadas de modo correto. Sem dúvida, esperamos ter a maior porcentagem possível. Em geral, a única forma de melhorar os resultados é usar mais (ou melhores) variáveis explicativas. Os valores da coluna H, **Exp (Coef)**, são geralmente usados para interpretar a equação de regressão. Eles indicam a provável alteração nas chances de “sucesso”, se qualquer variável explicativa aumentar 1 unidade.

**Dados ausentes
e conectar aos
dados**

- **Dados ausentes** – Se houver valores ausentes, qualquer linha com valores ausentes de *qualquer* das variáveis selecionadas será ignorada.
- **Conectar aos dados** – Não há vínculo com os dados originais. Se houver alteração nos dados, será preciso executar a análise novamente.

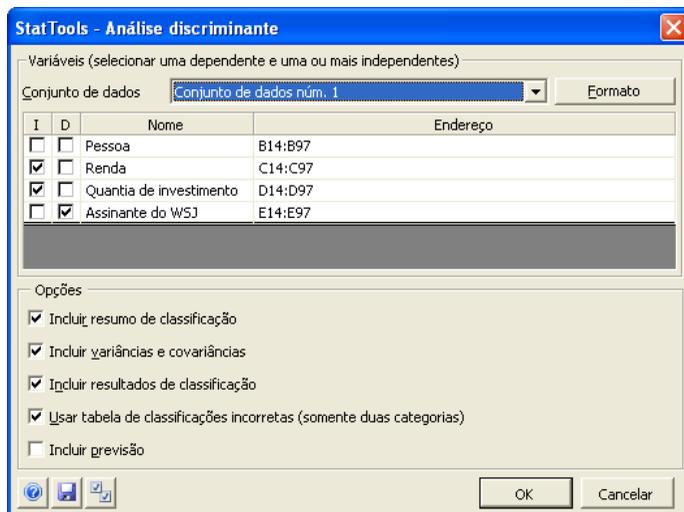
Comando Análise discriminante

Executa uma análise discriminante em um conjunto de variáveis

Este comando executa uma análise discriminante em um conjunto de dados. Nesta análise existe uma variável de “categoria” que especifica em quais dois ou mais grupos cada observação está, mais uma ou mais variáveis explicativas que podem ser usadas para prever a associação ao grupo. Existem duas formas de prever a associação ao grupo. O modo mais comum, válido para qualquer número de grupos, é calcular a “distância estatística” de cada observação à média de cada grupo, e classificar a observação de acordo com a menor distância estatística. O segundo modo, usado no caso de dois grupos, é calcular uma função discriminante (expressão linear das variáveis explicativas), e classificar cada observação dependendo de se o valor discriminante é inferior ou superior a algum valor de corte. Este segundo método também permite especificar probabilidades anteriores de associação ao grupo, assim como custos de classificação incorreta. Assim, o procedimento de classificação equivale a minimizar o custo esperado da classificação incorreta.

Caixa de diálogo Análise discriminante

Esta análise é configurada na caixa de diálogo **Análise discriminante**:



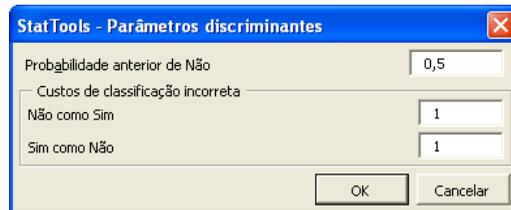
Uma variável dependente (D) e uma ou mais variáveis independentes (I) precisam ser selecionadas para a análise. Os dados precisam estar no formato “desempilhados”. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Análise discriminante** são:

- **Incluir resumo de classificação** com o relatório de regressão.
- **Incluir variâncias e covariâncias** com o relatório de regressão.
- **Incluir resultados de classificação** com o relatório de regressão.
- **Usar tabela de classificações incorretas**, selecionada quando for preciso alterar probabilidades anteriores ou custos de classificação incorreta.
- **Incluir previsão**, na qual os valores previstos da variável dependente são gerados para os valores da variável independente de um segundo conjunto de dados. Esse **conjunto de dados de previsão** deve ter os mesmos nomes de variáveis do conjunto de dados original analisado pela regressão. Em geral, no conjunto de dados de previsão haverá conjuntos de valores das variáveis independentes para os quais deseja fazer a **previsão** do valor da variável dependente. A equação de regressão calculada a partir do primeiro conjunto de dados é usada para fazer as previsões. Os valores previstos para a variável dependente serão inseridos diretamente no conjunto de dados de previsão, preenchendo a coluna (ou linha) da variável dependente com os valores previstos.

**Caixa de diálogo
Custos de
classificação
incorreta**

Se houver exatamente dois grupos possíveis para a categoria de variável dependente (como neste exemplo) e a opção **Usar tabela de classificações incorretas** estiver marcada, o sistema exibe uma caixa de diálogo que permite especificar probabilidades anteriores e/ou custos de classificação incorreta. As configurações padrão são: cada grupo tem a mesma probabilidade e os custos de classificação incorreta são iguais, mas essas configurações podem ser ignoradas.



StatTools - Parâmetros discriminantes

Probabilidade anterior de Não	0,5
Custos de classificação incorreta	
Não como Sim	1
Sim como Não	1

OK Cancelar

Relatório de análise discriminante

Pasta3 - Microsoft Excel

Resumo estatístico - Testes de normalidade - Controle de qualidade - Utilidades -
 Gráficos de resumo - Série temporal e previsão - Testes não-paramétricos - Ajuda -
 Inferência estatística - Regressão e classificação - Análises Ferramentas

A6

Relatório do StatTools

Análise: Discriminante
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Estática

Variáveis de configuração	Anterior	Custos de classificação incorreta	
		Probabilidade	Não
Não	0,5	0	1
Sim	0,5	1	0

Resumo amostral	Amostra	Média	
		Renda	Quantia de investimento
Não	57	66042,10526	24952,63158
Sim	27	80485,18519	53000

Função discriminante	Coefficiente
Renda	6,586E-05
Quantia de investimento	-0,000352735

Matriz de classificação	Não	Sim	Correto
Não	52	5	91,2%
Sim	2	25	92,6%

Resumo de classificação	
Correto	91,7%
Base	67,9%
Melhoria	74,1%

Matriz de var. e covars.		Renda	Quantia de investimento
Não			
Renda	209308552,6	79303458,65	
Quantia de investimento	79303458,65	99485394,74	
Sim			
Renda	118875156,7	64698846,15	
Quantia de investimento	64698846,15	80470769,23	
Combinada			
Renda	180634549	74672727,86	
Quantia de investimento	74672727,86	93456367,14	

Distância estatística	Renda	Quantia de investimento	Distância		Discrim	Original Classe	Análise Classe
			Não	Sim			
1	66400	26900	0,229005023	2,769086119	-5,115462701	Não	Não
2	68000	7100	2,363424601	5,232611294	1,974062536	Não	Não
3	54900	21500	0,84174592	3,258632671	-3,968084673	Não	Não

Discriminante

O relatório de Análise discriminante inclui os dados originais mais os dados usados para classificação. As classificações previstas dependem de se os valores discriminantes apresentados estão abaixo ou acima de um valor de corte. Se os campos probabilidades anteriores e custos de classificação incorreta ficarem com os valores padrão, este procedimento de classificação será equivalente à classificação básica na menor das duas distâncias estatísticas. Na verdade, se houvesse mais de dois grupos, os valores discriminantes não apareceriam e a classificação estaria baseada na menor das distâncias estatísticas. O relatório também apresenta estatísticas descritivas dos grupos e coeficientes da função discriminante (isso ocorre apenas quando existem dois grupos), as probabilidades anteriores, os custos de classificação incorreta e o valor de corte da classificação incorreta (novamente, se houver apenas dois grupos) e os resultados de resumo do procedimento de classificação (e os comentários da célula para ajudar na interpretação dos resultados). Neste exemplo, 89% das observações estão classificadas de modo correto. Sem dúvida, esperamos ter a maior porcentagem possível. Em geral, a única forma de melhorar os resultados é usar mais (ou melhores) variáveis explicativas.

**Dados ausentes
e conectar aos
dados**

- **Dados ausentes** – se houver valores ausentes, qualquer linha com valores ausentes de *qualquer* das variáveis selecionadas será ignorada.
- **Conectar aos dados** – não há vínculo com os dados originais. Se houver alteração nos dados, será preciso executar a análise novamente.

Menu Controle de qualidade

Os procedimentos do menu Controle de qualidade tratam da análise de dados coletados ao longo do tempo com aplicações de controle de qualidade.

O gráfico de Pareto exibe a importância relativa dos dados categorizados.

Os quatro tipos de gráficos de controle produzem gráficos de série temporal, que permitem observar se um processo está em controle estatístico. É possível ver se os dados ficam dentro dos limites de controle do gráfico, e também é possível verificar outro comportamento não randômico como execuções longas acima ou abaixo da linha central.

Comando Gráfico de Pareto

Comando gráfico de Pareto para a variável categorizada

Os gráficos de Pareto são úteis para determinar os itens mais significativos em um grupo de dados categorizados, além de apresentarem uma rápida representação visual de sua importância relativa. Em geral, os gráficos de Pareto são usados na área de Garantia da qualidade para determinar os poucos fatores de maior significado (Regra de Pareto 80/20).

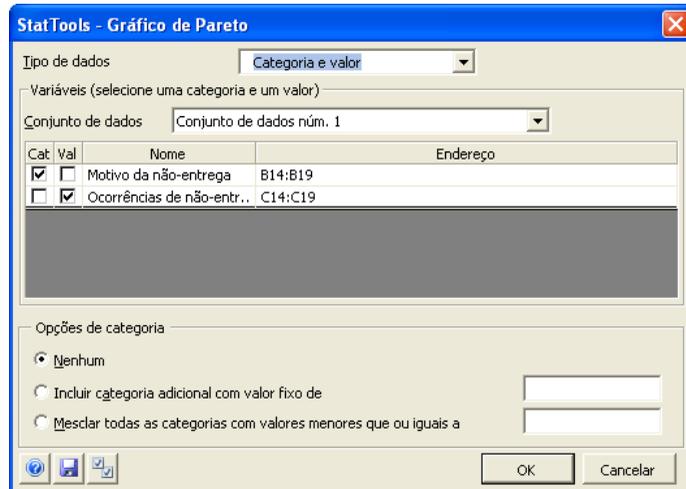
Por exemplo, um fabricante de peças para motor decidiu investigar o motivo pelo qual os clientes estavam rejeitando determinado produto. Quando cada lote era devolvido, informava-se o motivo (“tamanho errado”, “acabamento incorreto da superfície”, etc.) Após vários meses de coleta de dados, cria-se um gráfico de Pareto. Providências são tomadas para resolver as fontes dos problemas mais importantes.

O StatTools permite criar gráficos de Pareto com base em dados de um de dois formatos: “Somente categoria” ou “Categoria e valor”. A variável Somente categoria conterà, de modo geral, uma entrada de cada leitura. No exemplo acima, cada célula corresponderia ao motivo de um lote de peças devolvidas. O valor da célula poderia ser “acabamento incorreto da superfície” e, provavelmente haveria muitas células duplicadas. O StatTools conterà o número de vezes que cada entrada aparece na variável, e cria o respectivo gráfico de Pareto. Se o campo Categoria e valor for selecionado, as variáveis especificadas são as categorias e cada respectiva contagem.

Os eixos do gráfico de Pareto são construídos desta forma:

- As categorias são colocadas ao longo do eixo horizontal.
- A frequência (ou contagem) é colocada ao longo do eixo vertical esquerdo.
- A porcentagem cumulativa é colocada ao longo do eixo vertical direito.

Caixa de diálogo Gráfico de Pareto

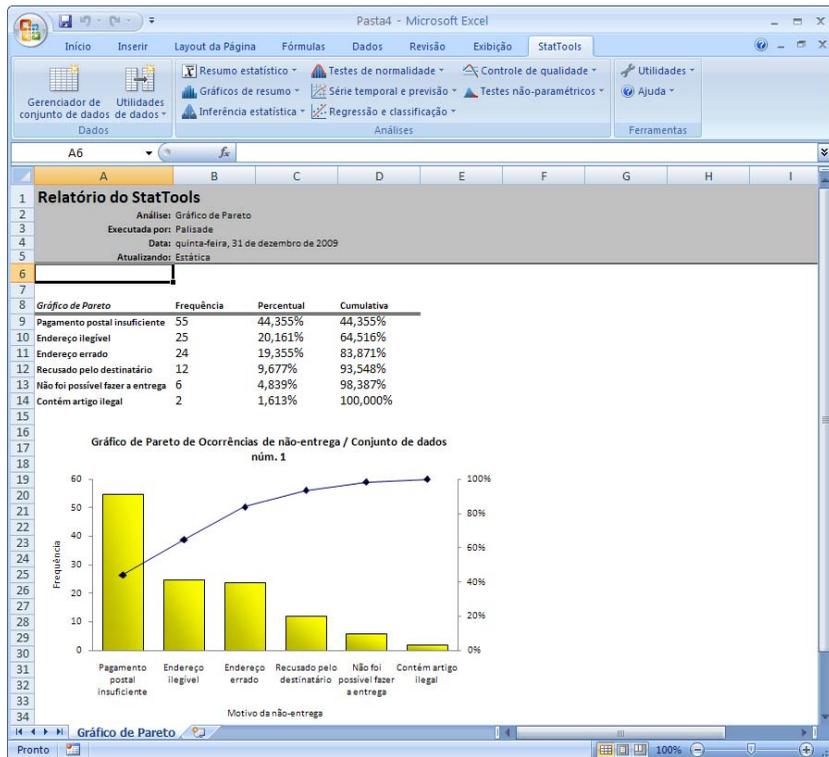


As opções da caixa de diálogo **Gráfico de Pareto** são:

- **Tipo de dados.** Seleciona o tipo de dados usado para construção do gráfico de Pareto: Categoria e Valor ou Somente categoria.

As **Opções de categoria** da caixa de diálogo **Gráfico de Pareto** são:

- **Nenhum** – cada uma das categorias será representada por uma barra no gráfico de Pareto.
- **Incluir categoria adicional com valor fixo de** – uma barra chamada “Div” será incluída na extrema direita do gráfico de Pareto, com uma frequência igual ao valor especificado.
- **Mesclar todas as categorias com valores menores que ou iguais a** – todas as categorias cuja frequência é menor ou igual ao valor especificado serão combinadas em uma categoria chamada “Div” e colocadas na extremidade direita do gráfico de Pareto.



Dados ausentes e conectar aos dados

- **Dados ausentes** – se houver valores ausentes, qualquer linha com valores ausentes de *qualquer* das variáveis selecionadas será ignorada.
- **Conectar aos dados** – não há vínculo com os dados originais. Se houver alteração nos dados, será preciso executar a análise novamente.

Comando Gráficos X/R

Cria gráficos de controle X e R para variáveis de série temporal

Esta análise produz gráficos de barras X/R para dados de série temporal. Ela pressupõe que os dados foram coletados em pequenas subamostras ao longo do tempo. Por exemplo, um operador pode coletar medições sobre as amplitudes das quatro peças selecionadas aleatoriamente a cada meia hora. Assim, o tamanho da subamostra é 4. Se os dados são coletados em 50 períodos de meia hora, os dados devem ser organizados em quatro colunas adjacentes e 50 linhas adjacentes, com cabeçalhos variáveis como SubAmostra1 a SubAmostra4 acima da primeira linha de dados.

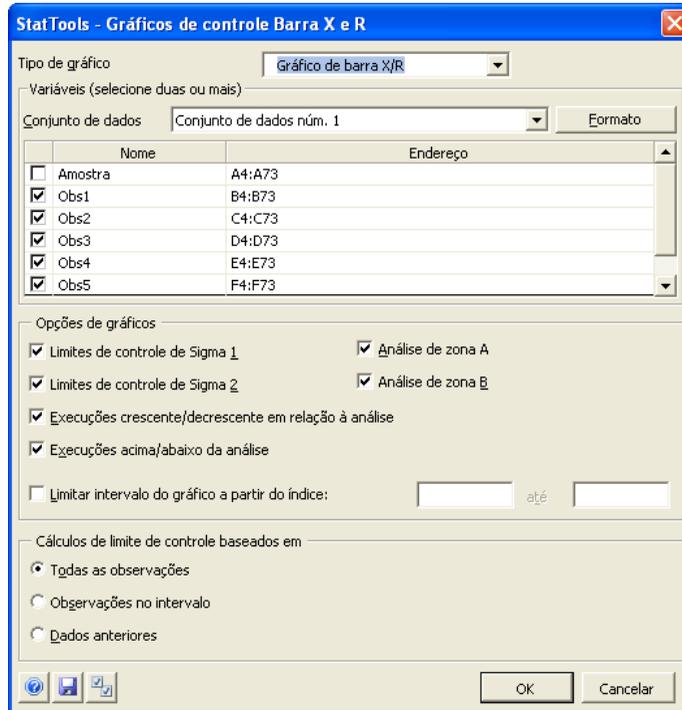
A finalidade do procedimento é verificar se o processo que gera os dados está em controle estatístico. Para tanto, o procedimento primeiramente calcula um gráfico de barras X e outro R, para cada linha do conjunto de dados. O gráfico de barras X será a média das observações daquela linha; R é o intervalo (máximo menos mínimo) das observação daquela linha.

As barras X e R são criadas em gráficos de série temporal separados em torno da linha central. A linha central do gráfico de barras X é a média das barras X (algumas vezes denominadas barras-duplas-X), e a linha central do gráfico R é a barra R, média das barras R. Um modo fácil de verificar se o processo está em controle é checar se qualquer das barras X ou R fica fora dos respectivos limites superior e inferior de controle (UCL e LCL), aproximadamente mais ou menos 3 desvios-padrão das linhas centrais. Os gráficos apresentam esses limites de controle e, assim, é fácil identificar qualquer valor extremo.

O procedimento também permite verificar outros possíveis comportamentos fora de controle, inclusive oito ou mais pontos em uma linha acima/abaixo da linha central, oito ou mais pontos em uma linha na direção ascendente/descendente, 4 de 5 pontos em uma linha, no mínimo, além de um desvio-padrão da linha central e no mínimo 2 de 3 pontos em uma linha além de dois desvios-padrão da linha central.

**Caixa de diálogo
Gráficos de
controle
Barra X e R**

Esses gráficos são configurados na caixa de diálogo **Gráficos de controle Barra X e R**:



Duas ou mais variáveis podem ser selecionadas para a análise. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

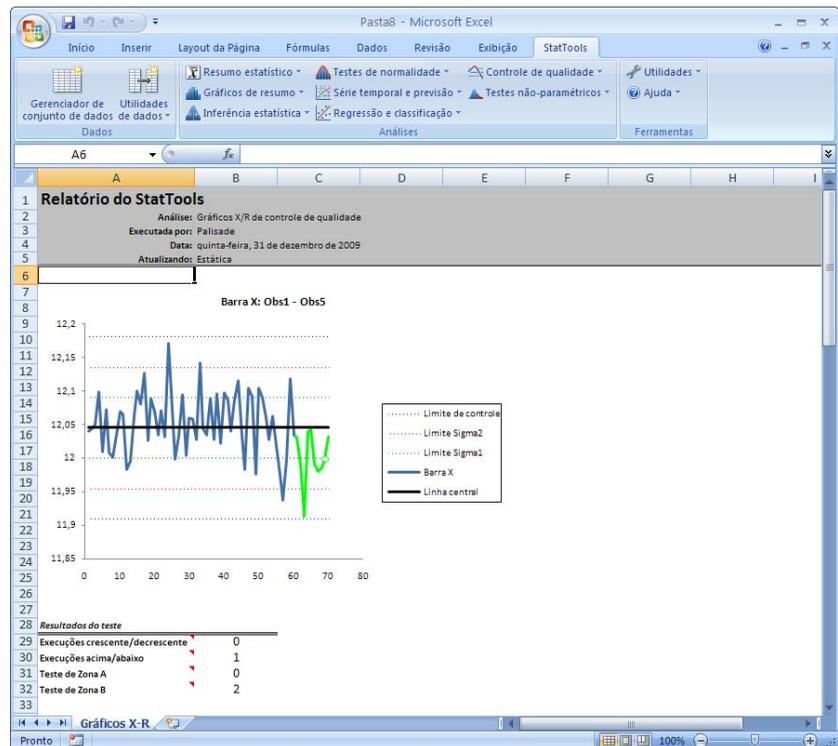
As **Opções de gráficos** na caixa de diálogo Gráficos de controle Barra X e R são:

- **Limites de controle de Sigma 1 e Sigma 2** – acrescenta linhas de limite de controle em um e/ou dois sigmas a partir da linha média. Essas linhas extras permitem verificar outros tipos de comportamento fora do controle (as chamadas regras da “zona”).
- **Análise de zona A e zona B** – número de pontos além da zona A (Sigma 2) e zona B (Sigma 1)
- **Execuções crescente/decrescente e acima/abaixo da análise** – movimentos sequenciais para crescentes/decrescentes de comprimento 8 ou superior.
- **Limitar intervalo do gráfico a partir do índice** – limita os pontos do gráfico em um intervalo de pontos de dados de uma variável (ou seja, intervalo entre os índices inicial e final)

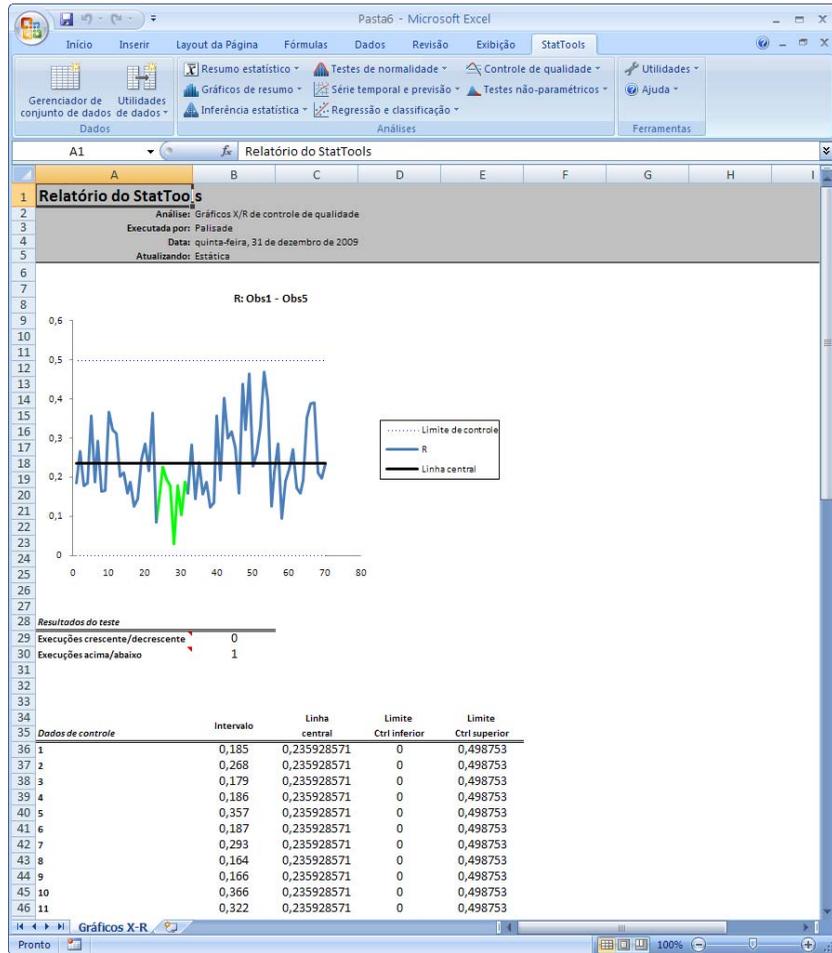
As opções de **Cálculos de limite de controle baseados em** na caixa de diálogo Gráficos de controle Barra X e R determinam os dados nos quais serão baseados os cálculos do limite de controle, incluindo:

- **Todas as observações** – usa todos os dados disponíveis nos cálculos de limite de controle
- **Observações no intervalo** – usa dados entre **Índice de início** e **Índice de parada** nos cálculos de limite de controle
- **Dados anteriores** – cria limites de controle de dados observados anteriormente. Basta inserir os valores de **Tamanho da subamostra**, **Média R** e **Média Barra X** calculados a partir dos dados anteriores.

Exemplo de gráfico Barra X



Exemplo de gráfico R



Dados ausentes e conectar aos dados

- **Dados ausentes** – não são permitidos.
- **Conectar aos dados** – os gráficos não se conectam aos dados.

Comando Gráfico P

Cria gráficos P para variáveis de série temporal

Os gráficos P destinam-se aos dados de “atributo”. Com dados de atributo, cada observação indica o número (ou fração) dos itens não conformes com as especificações a partir de uma amostra de itens. Por exemplo, um processo pode produzir determinado número de itens a cada meia hora, alguns dos quais não conformes. Assim, um gráfico P seria produzido a cada fração de meia hora com os itens não conformes. Como sempre, a finalidade é constatar se o processo está sob controle.

Este procedimento exige um conjunto de dados com uma das seguintes variáveis, no mínimo: uma variável que contenha o *número* de itens não conformes em cada amostra ou uma *fração* de itens não conformes em cada amostra. Opcionalmente, pode haver uma variável que contenha os tamanhos das amostras. Se não houver variável de tamanho de amostra, será preciso inserir um tamanho de amostra, pressupostamente uma constante de todas as amostras. Contudo, se houver uma variável de tamanho de amostra, os tamanhos das amostras não precisam ser iguais.

**Caixa de diálogo
Gráficos de
controle P**

Este gráfico é configurado na caixa de diálogo **Gráficos de controle P**:

Val	Nome	Endereço
<input type="checkbox"/>	Horas	A3:A27
<input type="checkbox"/>	Tam.amostra	B3:B27
<input checked="" type="checkbox"/>	Não-conforme	C3:C27

Para a análise são selecionadas uma variável de valor e, opcionalmente, uma variável de tamanho. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções de **Dados de input** na caixa de diálogo **Gráficos de controle P** são:

- **Números de itens não conformes** – especifica se a variável de valor apresenta o número real de itens não conformes do total da amostra
- **Frações de itens não conformes** – especifica se a variável de valor apresenta a fração de itens não conformes do total da amostra

As opções de **Tamanho amostral** na caixa de diálogo Gráficos de controle P são:

- **Usar variável de tamanho** – especifica se uma variável de tamanho é usada para apresentar o tamanho total de cada amostra
- **Usar tamanho comum** – especifica se nenhuma variável de tamanho é usada, uma vez que cada amostra é do tamanho informado.

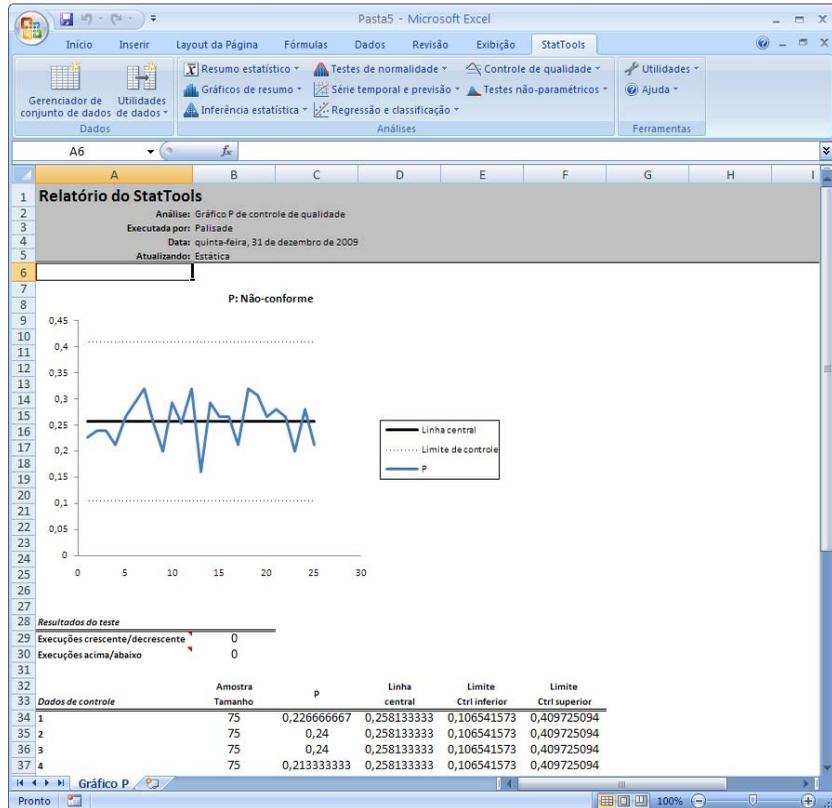
As **Opções de gráficos** na caixa de diálogo Gráficos de controle P são:

- **Limites de controle de Sigma 1 e Sigma 2** – acrescenta linhas de limite de controle em um e/ou dois sigmas a partir da linha média. Essas linhas extras permitem verificar outros tipos de comportamento fora do controle (as chamadas regras da “zona”).
- **Análise de zona A e zona B** – número de pontos além da zona A (Sigma 2) e zona B (Sigma 1)
- **Execuções crescente/decrescente e acima/abaixo da análise** – movimentos sequenciais para crescentes/decrescentes de comprimento 8 ou superior.
- **Limitar intervalo do gráfico a partir do índice** – limita os pontos do gráfico em um intervalo de pontos de dados de uma variável (ou seja, intervalo entre os índices inicial e final).

As opções de **Cálculos de limite de controle baseados em** na caixa de diálogo Gráficos de controle P determinam os dados nos quais serão baseados os cálculos do limite de controle, incluindo:

- **Todas as observações** – usa todos os dados disponíveis nos cálculos de limite de controle
- **Observações no intervalo** – usa dados entre **Índice de início** e **Índice de parada** nos cálculos de limite de controle
- **Dados anteriores** – cria limites de controle de dados observados anteriormente. Basta inserir os valores de **Tamanho da subamostra** e **Média P** calculados a partir dos dados anteriores.

Exemplo de gráfico P



Dados ausentes e conectar aos dados

- Dados ausentes – não são permitidos.
- Conectar aos dados – os gráficos não se conectam aos dados.

Comando Gráfico C

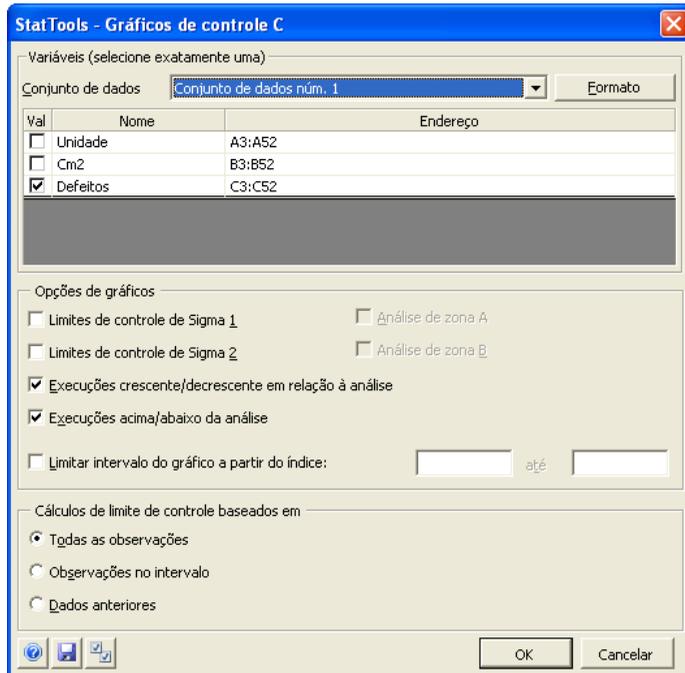
Cria gráficos C para variáveis de série temporal

Os gráficos C são usados para registrar o número de defeitos dos itens de um tamanho constante. Por exemplo, pressuponha que as portas dos automóveis sejam produzidas em lotes de 50 unidades. Em cada lote poderíamos contar o número de defeitos (mancha na pintura ou um canto áspero, por exemplo). Essas contagem são registradas em um gráfico. Como sempre, a finalidade é verificar se o processo está sob controle.

O conjunto de dados de um gráfico C deve incluir uma variável que contenha a contagem dos defeitos de cada item. Pressupõe-se que o tamanho do item seja igual em cada observação. Por exemplo, se um “item” for um lote de portas de automóveis, a nossa premissa seria de que cada lote tem o mesmo número de portas.

Este gráfico é configurado na caixa de diálogo **Gráficos de controle C**:

Caixa de diálogo
Gráficos de
controle C



Uma ou mais variáveis são selecionadas para análise. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

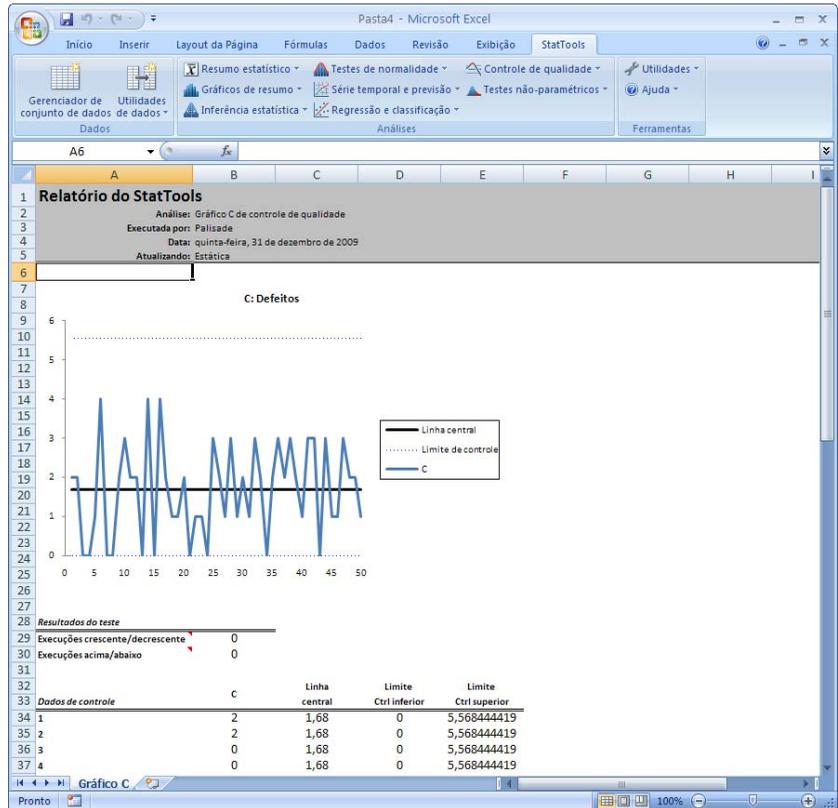
As **Opções de gráficos** na caixa de diálogo Gráficos de controle C são:

- **Limites de controle de Sigma 1 e Sigma 2** – acrescenta linhas de limite de controle em um e/ou dois sigmas a partir da linha média. Essas linhas extras permitem verificar outros tipos de comportamento fora do controle (as chamadas regras da “zona”).
- **Análise de zona A e zona B** – número de pontos além da zona A (Sigma 2) e zona B (Sigma 1)
- **Execuções crescente/decrescente e acima/abaixo da análise** – movimentos sequenciais para crescentes/ decrescentes de comprimento 8 ou superior.
- **Limitar intervalo do gráfico a partir do índice** – limita os pontos do gráfico em um intervalo de pontos de dados de uma variável (ou seja, intervalo entre os índices inicial e final)

As opções de **Cálculos de limite de controle baseados em** na caixa de diálogo Gráficos de controle C determinam os dados nos quais serão baseados os cálculos do limite de controle, incluindo:

- **Todas as observações** – usa todos os dados disponíveis nos cálculos de limite de controle
- **Observações no intervalo** – usa dados entre **Índice de início** e **Índice de parada** nos cálculos de limite de controle
- **Dados anteriores** – cria limites de controle de dados observados anteriormente. Basta inserir o valor de **Média C** calculado a partir dos dados anteriores.

Exemplo de gráfico C



Dados ausentes e conectar aos dados

- Dados ausentes – não são permitidos.
- Conectar aos dados – Os gráficos não se conectam aos dados.

Comando Gráfico U

Cria gráficos U para variáveis de série temporal

Os gráficos U são semelhantes aos gráficos C, mas agora registramos a taxa de defeitos. Usando o mesmo exemplo das portas de automóveis, vamos pressupor que os tamanhos do lote não sejam necessariamente iguais, ou seja, lotes diferentes têm quantidades diferentes de portas. Assim, em um gráfico U, registraríamos a taxa de defeitos por porta de automóvel, ou seja, o número de defeitos em um lote dividido pelo número de portas do lote. Como sempre, a finalidade é verificar se o processo está sob controle.

Este procedimento exige um conjunto de dados com uma das seguintes variáveis, no mínimo: uma variável que contenha o *número* de defeitos de cada observação ou uma variável que contenha a *taxa* de defeitos em cada observação. Opcionalmente, pode haver uma variável que contenha os tamanhos dos itens. Se não houver variável de tamanho, será preciso inserir um tamanho de item, pressupostamente o tamanho constante do item de todas as observações. Contudo, se houver uma variável de tamanho, os tamanhos dos itens não precisam ser iguais.

Caixa de diálogo Gráficos de controle U

Este gráfico é configurado na caixa de diálogo **Gráficos de controle U**:

T	Val	Nome	Endereço
<input type="checkbox"/>	<input type="checkbox"/>	Unidade	A3:A52
<input checked="" type="checkbox"/>	<input type="checkbox"/>	Cm2	B3:B52
<input type="checkbox"/>	<input checked="" type="checkbox"/>	Defeitos	C3:C52
<input type="checkbox"/>	<input type="checkbox"/>	Taxa	D3:D52

Para a análise são selecionadas uma variável de valor e, opcionalmente, uma variável de tamanho. Os dados do conjunto selecionado não devem estar empilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções de **Dados de input** na caixa de diálogo **Gráficos de controle U** são:

- **Números de defeitos** – especifica se a variável de valor apresenta o número real de itens não conformes do total da amostra.
- **Taxas de defeitos** – especifica se a variável de valor apresenta a fração de itens não conformes da amostra

As opções de **Tamanho amostral** na caixa de diálogo Gráficos de controle U são:

- **Usar variável de tamanho** – especifica se uma variável de tamanho é usada para apresentar o tamanho total de cada amostra
- **Usar tamanho comum** – especifica se nenhuma variável de tamanho é usada, uma vez que cada amostra é do tamanho informado.

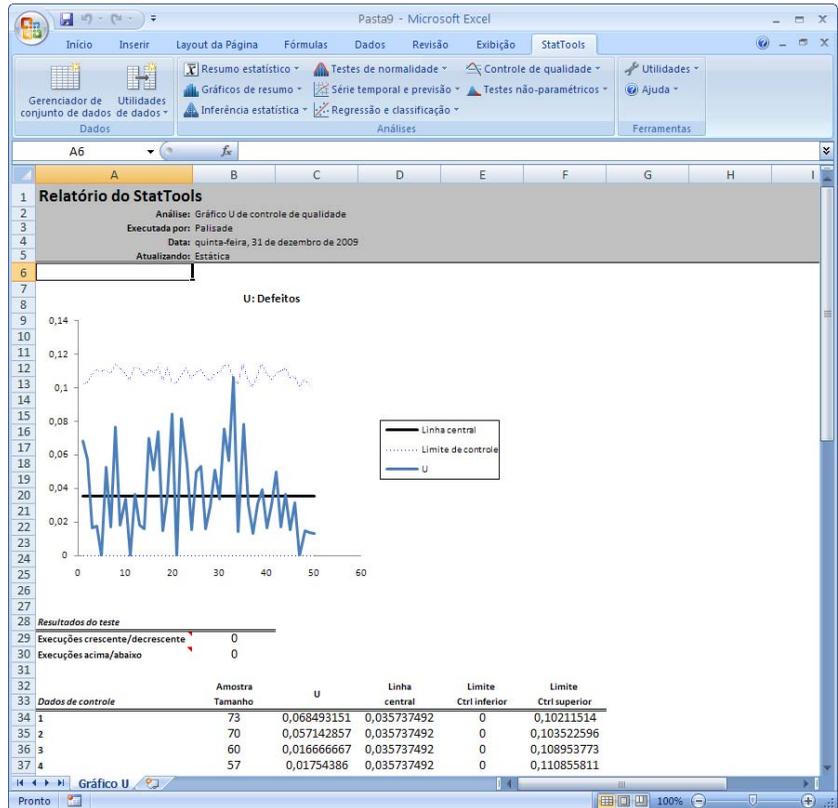
As **Opções de gráficos** na caixa de diálogo Gráficos de controle U são:

- **Limites de controle de Sigma 1 e Sigma 2** – acrescenta linhas de limite de controle em um e/ou dois sigmas a partir da linha média. Essas linhas extras permitem verificar outros tipos de comportamento fora do controle (as chamadas regras da “zona”).
- **Análise de zona A e zona B** – número de pontos além da zona A (Sigma 2) e zona B (Sigma 1)
- **Execuções crescente/decrecente e acima/abaixo da análise** – movimentos sequenciais para crescentes/decrecentes de comprimento 8 ou superior.
- **Limitar intervalo do gráfico a partir do índice** – limita os pontos do gráfico em um intervalo de pontos de dados de uma variável (ou seja, intervalo entre os índices inicial e final)

As opções de **Cálculos de limite de controle baseados em** na caixa de diálogo Gráficos de controle U determinam os dados nos quais serão baseados os cálculos do limite de controle, incluindo:

- **Todas as observações** – usa todos os dados disponíveis nos cálculos de limite de controle
- **Observações no intervalo** – usa dados entre **Índice de início** e **Índice de parada** nos cálculos de limite de controle
- **Dados anteriores** – cria limites de controle de dados observados anteriormente. Basta inserir os valores de **Tamanho da subamostra** e **Média U** calculados a partir dos dados anteriores.

Exemplo de gráfico U



Dados ausentes e conectar aos dados

- Dados ausentes – não são permitidos.
- Conectar aos dados – os gráficos não se conectam aos dados.

Menu Testes não-paramétricos

Testes “não-paramétricos” são procedimentos estatísticos aplicados a amostras de dados para testar hipóteses sobre distribuições de probabilidades associadas. Testes de hipótese “paramétrica” são mais conhecidos e muito usados; entretanto, a alternativa não paramétrica oferece vantagens que fazem com que seja a escolha mais adequada em muitas situações.

Os testes de hipótese paramétrica consideram premissas sobre o tipo de distribuição associada (em geral, isso é o normal), e estimam os parâmetros desse tipo de distribuição (em geral, a média e o desvio-padrão). Em muitas aplicações a premissa normalidade estaria incorreta. Por exemplo, os números de chamadas por hora de uma central de atendimento ao consumidor e o tempo de espera em um caixa de supermercado não são distribuídos normalmente. Os testes não-paramétricos não exigem pressuposições sobre o tipo de distribuição associada. Alguns deles fazem determinadas pressuposições gerais sobre a forma da distribuição: nesse pacote, o teste dos postos sinalizados Wilcoxon pressupõe que a distribuição seja simétrica. Os outros dois testes do pacote (Teste de sinal e Teste Mann-Whitney) não pressupõem nada sobre a forma da distribuição.

Com amostras de tamanho pequeno, a alternativa não paramétrica é sempre mais apropriada. Se a amostra for grande, pode-se aplicar o teste de normalidade. Se a pressuposição de que a distribuição é normal justificar-se, pode-se usar um teste paramétrico. Entretanto, para amostras de tamanho pequeno os testes de normalidade têm pouco poder para diferenciar a distribuição normal das outras. Os testes não-paramétricos podem solucionar o dilema.

Dados ordinais

Para determinados tipos de dados, não é possível usar os testes paramétricos mas, sim, alguns não-paramétricos. **Dados ordinais** é um desses casos, em que as observações são descritas em termos de números que expressam classificações; entretanto, a diferença entre dois desses números não é significativa. Por exemplo, níveis de escolaridade podem ser codificados como 0 (primeiro grau), 1 (secundário incompleto), 2 (secundário), 3 (universitário incompleto), 4 (universitário) e 5 (pós-graduação). Quando essa escala é usada, não há implicação de que a diferença no nível de escolaridade entre “secundário” e “secundário incompleto” seja equivalente à diferença entre estar nas categorias “pós-graduação” e “universitário incompleto”, ainda que nos dois casos a diferença entre as classificações seja igual a 2. Este pacote inclui testes que podem ser aplicados a esses dados: o Teste de sinal (tipo de análise uniamostrai) e Teste Mann-Whitney.

Resumo dos usos

Em resumo, os testes não-paramétricos são aplicáveis nas seguintes situações, nas quais não se aplicam os testes paramétricos:

- quando há pouca informação sobre a distribuição de probabilidade associada,
- quando a amostra é de tamanho pequeno para testar a pressuposição normalidade de modo confiável,
- quando os dados são ordinais.

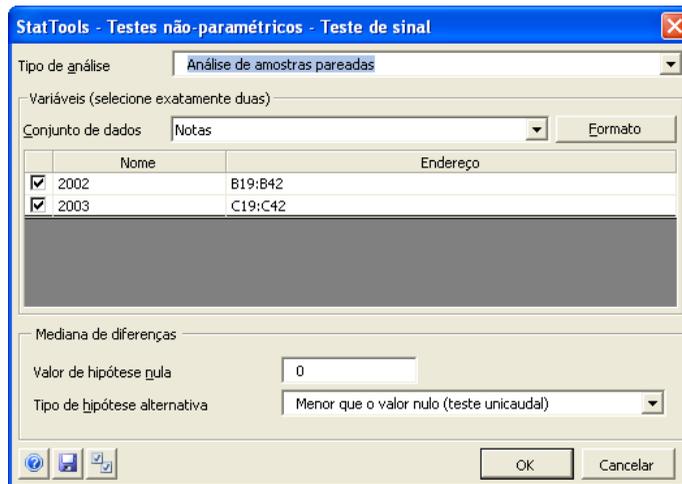
Comando Teste de sinal

Executa teste de sinal nas variáveis

O comando Teste de sinal executa testes de hipótese da mediana de uma variável simples (ou de uma mediana de diferenças de um par de variáveis (**Análise uniamostrai**) ou da mediana das diferenças de um par de variáveis (**Análise de amostras pareadas**). O teste não faz pressuposições sobre a forma de distribuição (e, especificamente, não pressupõe que seja normal). A análise uniamostrai pode ser usada com **dados ordinais** como descrito no capítulo Visão geral.

Caixa de diálogo Teste de sinal

Esta análise é configurada na caixa de diálogo **Teste de sinal**:



O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise uniamostrai exige uma ou mais variáveis, enquanto uma análise de amostras pareadas exige duas variáveis. Para uma análise uniamostrai, as variáveis selecionadas podem ter dados empilhados ou desempilhados; se for análise de amostras pareadas, os dados precisam estar desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Teste de sinal** são:

- **Tipo de análise.** Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostr.** Executa testes de hipótese para uma variável numérica simples.
 - **Análise de amostras pareadas.** Esta análise é adequada quando duas variáveis formam um par, naturalmente. Equivale a uma análise uniamostr. das diferenças entre pares.
- **Mediana (ou Mediana de diferenças).**
 - **Valor de hipótese nula,** ou o valor do parâmetro de população sob hipótese nula.
 - **Tipo de hipótese alternativa,** ou a alternativa ao **Valor de hipótese nula** a ser avaliado durante a análise. A hipótese alternativa pode ser **“monocaudal”** (ou seja, maior ou menor que a hipótese nula) ou **“bicaudal”** (ou seja, não igual à hipótese nula).

Relatório de teste de sinal

Teste de sinal (amostras pareadas)		2002 - 2003
Estatísticas amostrais		
Tamanho amostral		24
Média amostral		-6,358333333
Desvio padrão amostral		13,11745988
Mediana amostral		-5,75
Hipóteses		
Mediana hipotética (MH)		0
Hipótese alternativa		< 0
Ajuste de tamanho amostral		
Núm. de valores = MH		0
Núm. de valores < ou > MH (tamanho ajustado)		24
Contagens de valor		
Núm. de valores < MH		17
Núm. de valores > MH (estatística de teste)		7
Níveis de significância e valor P		
Valor P		0,0320
Hipótese nula com significância de 10%		Rejeitar
Hipótese nula com significância de 5%		Rejeitar
Hipótese nula com significância de 1%		Não rejeitar

***Dados ausentes
e conectar aos
dados***

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

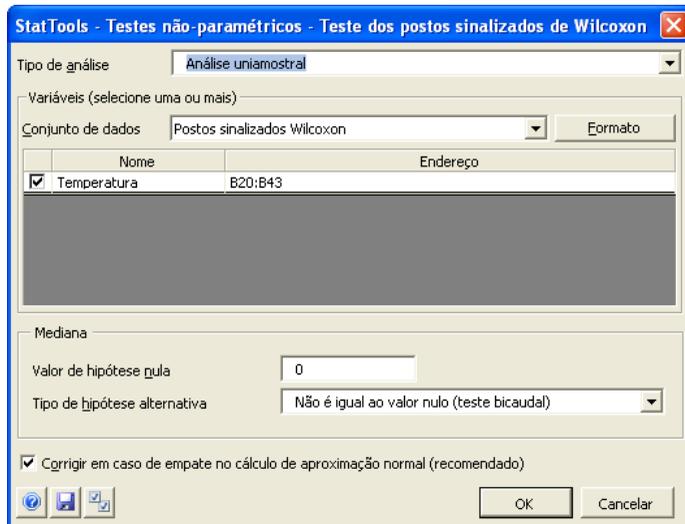
Comando Teste dos postos sinalizados de Wilcoxon

Executa testes dos postos sinalizados de Wilcoxon nas variáveis

O comando Teste dos postos sinalizados de Wilcoxon executa testes de hipótese da mediana de uma variável simples (**Análise uniamostr**) ou da mediana das diferenças de um par de variáveis (**Análise de amostras pareadas**). O teste pressupõe que a distribuição de probabilidade seja simétrica (mas não pressupões que isso seja normal).

Esta análise é configurada na caixa de diálogo **Teste dos postos sinalizados de Wilcoxon**:

**Caixa de diálogo
Teste dos postos
sinalizados de
Wilcoxon**



O número de variáveis selecionadas depende do tipo de análise utilizada. Uma análise uniamostr exige uma ou mais variáveis, enquanto uma análise de amostras pareadas exige duas variáveis. Para uma análise uniamostr, as variáveis selecionadas podem ter dados empilhados ou desempilhados; se for análise de amostras pareadas, os dados precisam estar desempilhados. As variáveis podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Teste dos postos sinalizados de Wilcoxon** são:

- **Tipo de análise.** Seleciona o tipo de análise executada. As opções são:
 - **Análise uniamostr.** Executa testes de hipótese para uma variável numérica simples.
 - **Análise de amostras pareadas.** Este tipo de análise é adequado quando duas variáveis formam um par, naturalmente. Equivale a uma análise uniamostr. das diferenças entre pares.
- **Mediana (ou Mediana de diferenças).**
 - **Valor de hipótese nula,** ou o valor do parâmetro de população sob hipótese nula.
 - **Tipo de hipótese alternativa,** ou a alternativa ao **Valor de hipótese nula** a ser avaliado durante a análise. A hipótese alternativa pode ser “**monocaudal**” (ou seja, maior ou menor que a hipótese nula) ou “**bicaudal**” (ou seja, não igual à hipótese nula).
- **Correção de empate.** Seleção recomendada que corrige classificações empatadas no teste apenas quando for usada a aproximação normal. A correção envolve contar os números de elementos em grupos de classificações empatadas, reduzindo a variância proporcionalmente. A correção dos empates aumentará sempre o valor da estatística Z , se houver classificações empatadas. (Nota: a correção do empate não produzirá alterações na variância se não houver empates.)

Relatório de teste dos postos sinalizados de Wilcoxon

Relatório do StatTools	
Análise: Teste dos postos sinalizados de Wilcoxon	
Executada por: Pálissade	
Data: quinta-feira, 31 de dezembro de 2009	
Atualizando: Tempo real	
Temperatura	
Teste de postos sinalizados (uni amostral) Postos sinalizados Wilcoxon	
Estadísticas amostrais	
Tamanho amostral	24
Média amostral	-6,34
Desvio padrão amostral	13,14
Mediana amostral	-5,75
Hipóteses	
Mediana hipotética (MH)	0
Hipótese alternativa	<> 0
Ajuste de tamanho amostral	
Núm. de valores = MH	0
Núm. de valores < ou > MH (tamanho ajustado)	24
Informações de ranking	
Número de valores empatados	2
Soma de postos negativos	226,5
Soma de postos positivos (estatística de teste)	73,5
Cálculo de valor p	
Aproximação normal (AN) não usada	Sim
Empate presente mas não corrigido para	Não
Média para AN	150
Desvio padrão para AN com correção de empate	34,95711659
Estatística Z para AN com correção de empate	-2,1741
Valor P	0,0297
Níveis de significância	
Hipótese nula com significância de 10%	Rejeitar
Hipótese nula com significância de 5%	Rejeitar
Hipótese nula com significância de 1%	Não rejeitar

Nota: neste relatório, o valor p é calculado com aproximação normal quando o tamanho da amostra for maior que 15.

Dados ausentes e conectar aos dados

- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

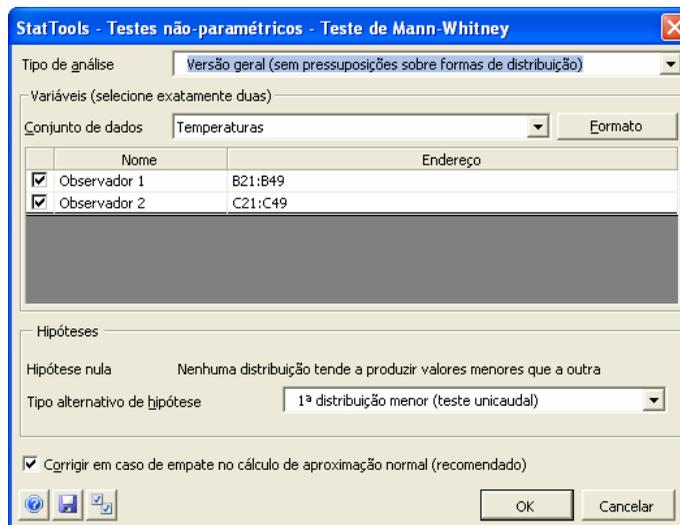
Comando Teste de Mann-Whitney

Executa teste de Mann-Whitney nas variáveis

O comando Teste de Mann-Whitney executa um teste de hipótese em duas amostras. Em uma versão do teste (versão mediana) a hipótese informa que as medianas das duas populações são idênticas. Nesta versão, pressupõe-se que as distribuições de probabilidade têm a mesma forma. Na outra versão (versão geral) esta pressuposição não é feita e a hipótese nega que qualquer distribuição de probabilidade tende a produzir valores menores que a outra (mais precisamente, informa que $P[X_1 > X_2] = P[X_2 > X_1]$, onde $P[X_1 > X_2]$ é a probabilidade de que uma observação da população 1 é maior que uma observação da população 2). Observe que o teste de Mann-Whitney pode ser usado para rejeitar a hipótese de que duas amostras são geradas pela mesma distribuição de probabilidade. O teste de Mann-Whitney é muitas vezes denominado o teste de soma de postos Wilcoxon.

Esta análise é configurada na caixa de diálogo **Teste de Mann-Whitney**:

Caixa de diálogo Teste de Mann-Whitney



A análise exige duas variáveis. Podem ser empilhadas ou desempilhadas, e podem pertencer a conjuntos de dados diferentes.

As opções da caixa de diálogo **Teste de Mann-Whitney** são:

- **Tipo de análise.** Seleciona a formulação da hipótese nula e as alternativas. As opções são:
 - **Versão geral.** Executa um teste de hipótese para ver se uma distribuição de probabilidade tende a produzir valores menores que a outra.
- **Hipóteses.**
 - **Hipótese nula:** Nega que nenhuma distribuição tenda a produzir valores menores que a outra. Mais precisamente, informa que $P[X_1 > X_2] = P[X_2 > X_1]$, onde $P[X_1 > X_2]$ é a probabilidade que uma observação da população 1 seja maior que uma observação da população 2, e $P[X_2 > X_1]$ tem interpretação análoga. Para distribuições contínuas, isto equivale a dizer que estas duas probabilidades são 0,5 ($P[X_1 > X_2] = P[X_2 > X_1] = 0,5$).
 - **Hipótese alternativa:** A hipótese alternativa pode ser “monocaudal” (ou seja, uma probabilidade é maior ou menor que a outra) ou “bicaudal” (ou seja, as duas probabilidades não são iguais).
 - **Versão mediana.** Executa um teste de hipótese para ver se a mediana de uma população é a mesma ou diferente da mediana da outra população. Pressupões que as duas distribuições têm a mesma forma.
- **Hipóteses.**
 - **Hipótese nula:** informa que as duas medianas são iguais.
 - **Hipótese alternativa:** A hipótese alternativa pode ser “monocaudal” (ou seja, a mediana da primeira população é maior ou menor que a segunda) ou “bicaudal” (ou seja, as duas medianas não são iguais).

Nota: os cálculos executados nas versões geral e mediana do teste são os mesmos; as versões diferem apenas quanto à presença das pressuposições das formas de distribuição iguais, e com relação à hipótese nula. O objetivo de ter essas duas versões é deixar claro que o teste de Mann-Whitney pode ser aplicado ainda que não se possa pressupor que as distribuições têm formas aproximadamente idênticas, desde que se considere uma hipótese nula apropriada. Especificamente, se o teste for executado quando as duas distribuições tiverem claramente formas diferentes, e o teste rejeitar a hipótese nula, pode ser que as medianas sejam diferentes, mas também pode ser que as variâncias sejam diferentes ou por outros motivos.

- **Correção de empate.** Seleção recomendada que corrige classificações empatadas no teste apenas quando for usada a aproximação normal. A correção envolve contar os números de elementos em grupos de classificações empatadas, reduzindo a variância proporcionalmente. A correção dos empates aumentará sempre o valor da estatística Z, se houver classificações empatadas. (Nota: a correção do empate não produzirá alterações na variância se não houver empates.)

Relatório do teste de Mann-Whitney

Pasta5 - Microsoft Excel

Relatório do StatTools

Análise: Teste de Mann-Whitney
 Executada por: Palisade
 Data: quinta-feira, 31 de dezembro de 2009
 Atualizando: Tempo real

	Observador 1	Observador 2
Estatísticas amostrais		
Temperaturas		
Tamanho amostral	29	24
Média amostral	29,292	32,017
Desvio padrão amostral	4,671	5,076
Mediana amostral	29,400	31,195

Teste de Mann-Whitney (Versão geral)

Hipóteses

Hipótese nula	Nenhuma das distâncias menor
Hipótese alternativa	Observador 1 Menor

Informações de ranking

Número de valores em ranking	53
Número de valores empatados	2
Soma de postos para Variável 1 (estatística de teste)	690,5
Soma de postos para Variável 2	740,5

Cálculo de valor P

Aproximação normal (AN) não usada	Sim
Empate presente mas não corrigido para	Não
Média para AN	783
Desvio padrão para AN com correção de empate	55,96314618
Estatística Z para AN com correção de empate	-1,6439
Valor P	0,0501

Níveis de significância

Hipótese nula com significância de 10%	Rejeitar
Hipótese nula com significância de 5%	Não rejeitar
Hipótese nula com significância de 1%	Não rejeitar

No relatório acima, o *valor p* é calculado usando a aproximação normal quando o tamanho de uma das duas amostras é maior que 10 (salvo quando um dos tamanhos for igual a 11 ou 12, enquanto a outra é igual a 3 ou 4).

Dados ausentes e conectar aos dados

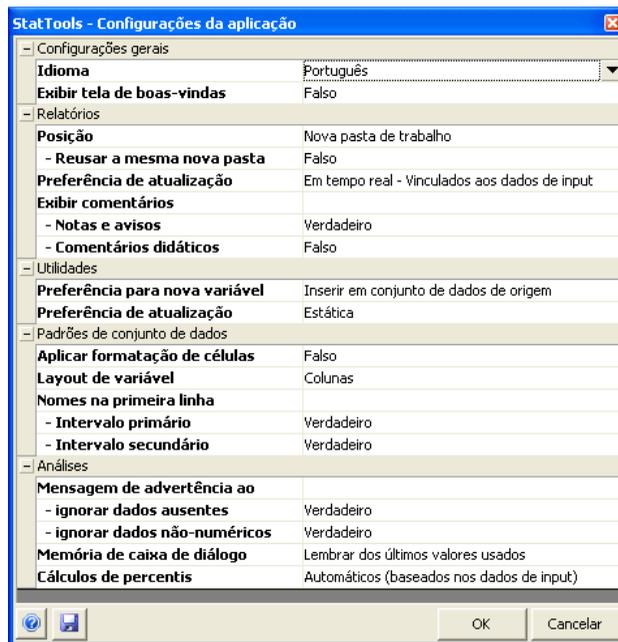
- **Dados ausentes** – são permitidos. Todas as linhas com dados ausentes nas variáveis selecionadas são ignorados.
- **Conectar aos dados** – todos os relatórios são calculados por fórmulas vinculadas com dados. Se os valores da variável selecionada mudarem, os outputs mudam automaticamente.

Menu Utilidades

Comando Configurações da aplicação

Especifica configurações para relatórios, gráficos, utilidades, conjuntos de dados e análises do StatTools

O comando **Configurações da aplicação** permite especificar configurações gerais para relatórios, gráficos, utilidades, conjuntos de dados e análises do StatTools. Essas configurações aplicam-se a todas as análises e conjuntos de dados. Outras configurações específicas para análise são definidas na respectiva caixa de diálogo.



As **Configurações de relatórios** especificam opções para relatórios e gráficos criados pelas análises do StatTools. As configurações definem:

- **Posição** – seleciona o local no Excel de novos relatórios e gráficos como:
 - **Pasta de trabalho ativa** – é criada uma nova planilha para cada relatório.
 - **Nova pasta de trabalho** – é criada uma nova pasta de trabalho de relatório do StatTools (se necessário) e cada relatório é colocado em uma planilha dessa pasta de trabalho.
 - **Depois da última coluna usada na planilha ativa** – cada relatório do StatTools é colocado na planilha ativa, à direita da última coluna utilizada.
 - **Perguntar qual é a célula inicial** – depois de executar uma análise, você seleciona a célula e o relatório ou gráfico é colocado no canto superior esquerdo da mesma.
 - **Reusar a mesma nova pasta de trabalho** – se for criada uma nova pasta de trabalho, ela será usada para todos os relatórios.
- **Preferência de atualização** – especifica como os resultados devem mudar quando são mudados os dados de variáveis. As opções de atualização de resultados são:
 - **Em tempo real – Vinculados aos dados de input** – os relatórios são automaticamente atualizados à medida que os dados de input mudam.
 - **Estática – Valores fixos** – os relatórios não se alteram com a mudança dos dados de input. As estatísticas são fixas, com base nos valores dos dados de input do momento em que o procedimento foi executado.

Os resultados são produzidos em tempo real no StatTools usando fórmulas do Excel e funções personalizadas do StatTools. Por exemplo, a fórmula:

```
=StatMean('Intervalo de confiança.xls'!Par)
```

calcula a média do *Par* de variáveis (que usa dados do nome de intervalo do Excel “Par” localizado na pasta de trabalho Confidence Interval.xls). Se houver mudanças nos dados do intervalo Par, o valor retornado pela função StatMean será atualizado.

Os relatórios e gráficos de todos os procedimentos do StatTools podem ser atualizados em tempo real, com as seguintes exceções:

- 1) Regressão
- 2) Análise discriminante
- 3) Regressão logística
- 4) Previsão

Esses procedimentos exigem recálculos demorados, que poderiam fazer com que o Excel tivesse reações mais lentas durante a atualização em tempo real.

- **Exibir comentários** – especifica quais categorias das mensagens do StatTools serão incluídas nos relatórios. Notas, avisos e notas didáticas podem ser exibidas como notas pop-up nas células em relatórios, como abaixo:

StatTools – Exemplo: Teste de Mann-Whitney (versão geral, dados desempilhados)

O teste de Mann-Whitney do StatTools é usado para verificar se um conjunto de dados amostrais demonstra se certas hipóteses devem ser aceitas ou rejeitadas. Em uma versão do teste (teste de mediana), a hipótese diz que as medianas de duas populações são idênticas. Na versão, despois-se que as distribuições de probabilidades têm a mesma forma. Em outra versão (teste geral) está pressuposição não é feita, e a hipótese nega que qualquer uma das distribuições de probabilidade tenha a gerar valores menores do que a outra (mais brevemente, a hipótese diz que $P(X_1 < X_2) = P(X_2 < X_1)$, onde $P(X_1 < X_2)$ é a probabilidade de que uma observação da população 1 seja maior do que uma observação da população 2).

Obs.: este procedimento permite usar dados empilhados ou desempilhados. Este exemplo usa dados desempilhados.

O conjunto de dados mostrado consiste nas medidas de temperaturas feitas por dois observadores em uma estação meteorológica local. Abaixo são mostrados os resultados de análise feita com o teste de Mann-Whitney (versão geral) para identificar se as observações da primeira pessoa tendem a produzir valores menores do que as da outra pessoa. Ao usar a alternativa monocaudal, podemos regular a hipótese numa significância de 10%, mas não na significância de 5% ou na de 1%.

Este exemplo foi criado pela Palisade Corporation.

Observador 1	Observador 2	Observador 1 Temperaturas	Observador 2 Temperaturas
29,24	29,74	29	24
25,28	37,88	29,292	32,017
34,11	32,59	4,671	5,076
36,81	33,54	29,400	31,195
26,54	29,08		
24,66	29,03		
25,33	29,64		
32,8	23,55		
21,24	30,47		
31,84	30,34		
19,75	35,69		
31,35	39,33		
27,74	29,24		
35,86	34,76		
38,62	29,27		
27,48	24,54		
32,44	35,16		

Estatísticas amostrais		Observador 1	Observador 2
Tamanho amostral		29	24
Média amostral		29,292	32,017
Desvio padrão amostral		4,671	5,076
Mediana amostral		29,400	31,195

Teste de Mann-Whitney (Versão geral)

Hipóteses	Resultado
Hipótese nula	Nenhuma das distâncias menor
Hipótese alternativa	Observador 1 Menor

Informações de ranking	
Número de valores em ranking	53
Número de valores empatados	2
Soma de postos para Variável 1 (estatística)	690,5
Soma de postos para Variável 2	740,5

Cálculo de valor P

Aproximação normal (A/N) não usada	Resultado
	Sim

Nota do StatTools:
A hipótese de versão geral nega que a qualquer distribuição tendam a produzir valores menores do que a outra ($P(X_1 < X_2) = P(X_2 < X_1)$, onde $P(X_1 < X_2)$ probabilidade de que a observação da população 1 > obs. da pop. 2).

As **Configurações de utilidades** especificam opções para novas variáveis criadas pelas utilidades do StatTools. Essas utilidades podem ser localizadas no menu Utilidades de dados. Eles criam novas variáveis transformando, combinando ou processando de qualquer outra forma as variáveis existentes.

- **Preferência para nova variável** – Especifica o local desejado das variáveis novas criadas pelas utilidades dos dados. As opções são:
 - **Inserir em conjunto de dados de origem** – cada nova variável criada é inserida no lado direito (ou inferior) do conjunto de dados que contém a variável original.
 - **Criar novo conjunto de dados**, onde novas variáveis serão colocadas em um novo conjunto de dados.

Todavia, haverá casos em que a Preferência para nova variável selecionada talvez não possa ser seguida. Essas situações incluem dados **empilhados e desempilhados** (neste caso as variáveis novas serão sempre colocadas em um novo conjunto de dados) e quando os **dados originais provêm de um conjunto de dados de múltiplos intervalos** (neste caso as variáveis novas só podem ser inseridas no conjunto de dados de origem).

- **Preferência de atualização** – especifica como os valores das variáveis criadas por uma utilidade mudarão quando mudam os dados da variável original. As opções para atualização são:
 - **Em tempo real – Vinculado aos dados de input** – os valores da variável nova são automaticamente atualizados à medida que os dados de input mudam.
 - **Estática – Valores fixos** – os valores da nova variável não se alteram com a mudança dos dados de input. Os valores da variável nova são fixos, com base nos valores dos dados de input do momento em que o procedimento foi executado.

Entretanto, existem casos em que o campo Preferência de Atualização não é seguido e as variáveis novas sempre usam a preferência **Estática – Valores fixos**. Esses casos incluem dados **Empilhados e desempilhados e Amostras aleatórias**. A atualização em tempo real não é aplicável a essas utilidades.

Configurações de conjuntos de dados

As **Configurações de conjuntos de dados** especificam opções dos novos conjuntos de dados criados usando o comando Gerenciador de conjunto de dados. Essas configurações são apenas padrões que aparecem na criação de um novo conjunto de dados. As configurações podem ser alteradas, se necessário, na caixa de diálogo Gerenciador de conjunto de dados.

- **Aplicar formatação de células** – Especifica se o conjunto de dados será formatado pelo StatTools.
- **Layout** – Seleciona o layout da variável padrão (por linha ou coluna).
- **Nomes na primeira coluna/linha (Intervalo primário)** – Especifica se os nomes são inseridos na primeira coluna ou linha do primeiro intervalo definido.
- **Nomes na primeira coluna/linha (Intervalo secundário)** – Especifica se os nomes são inseridos na primeira coluna ou linha do segundo intervalo definido e todos os intervalos subsequentes de um conjunto de dados de múltiplos intervalos.

Configurações de análises

As **Configurações de análises** especificam as entradas padrão exibidas nas caixas de diálogo usadas para configurar análises. Estas são apenas as entradas exibidas quando uma caixa de diálogo de uma análise for exibida pela primeira vez. As configurações podem ser alteradas, se necessário, nas respectivas caixas de diálogo.

- **Mensagens de advertência** – Seleciona se as mensagens de advertência serão exibidas durante a execução de uma análise se o StatTools detectar dados faltantes em uma variável ou se o StatTools detectar dados não numéricos.
- **Memória da caixa de diálogo** - Especifica as entradas padrão exibidas nas caixas de diálogo usadas para configurar análises. As opções são:
 - **Lembrar dos últimos valores usados (pela pasta de trabalho)**, em que uma caixa de diálogo exibida apresenta as entradas feitas nessa caixa na última vez em que foi exibida na pasta de trabalho ativa. Se uma caixa de diálogo for exibida pela primeira vez, apresentará as configurações padrão do sistema salvas para a análise.
 - **Usar sempre valores padrão do sistema**, em que uma caixa de diálogo apresenta as configurações padrão do sistema salvas para a análise.

- **Cálculos de percentis** – Selecciona o método a ser usado para calcular percentis. Dependendo da natureza dos seus dados, diferentes métodos podem produzir respostas melhores. Os métodos disponíveis (e o tipo de dados para os quais são adequados) são:
 - **Automáticos (baseados nos dados de input)**
 - **Interpolado com pontos terminais assimétricos (contínuo)**
 - **Interpolado com pontos terminais simétricos (contínuo)**
 - **Função de percentil do Excel (contínuo)**
 - **Observação mais próxima (discreta)**
 - **Função de dist. empírica (discreta)**
 - **Função de dist. empírica com cálculo de média (discreta)**

Comando Excluir conjuntos de dados

Exclui conjuntos de dados do StatTools na pasta de trabalho ativa

O comando **Excluir conjuntos de dados** exclui todos os conjuntos de dados definidos de uma pasta de trabalho ativa. Os dados reais não são excluídos do Excel, apenas a definição dos conjuntos de dados.

Comando Limpar memória da caixa de diálogo

Limpa toda a memória das entradas nas caixas de diálogo da análise

O comando Limpar memória da caixa de diálogo limpa toda a “memória” das entradas das caixas de diálogo da análise. As caixas de diálogo exibidas subsequentemente apresentarão, inicialmente, as configurações padrão do sistema salvas de cada análise.

Comando Descarregar o add-in StatTools

Descarrega o add-in StatTools

O comando Descarregar o add-in StatTools descarrega o StatTools e fecha todas as suas janelas.

Menu Ajuda

Ajuda do StatTools

Abre o arquivo de ajuda online do StatTools.

O comando Ajuda do StatTools do menu Ajuda abre o arquivo principal de ajuda do StatTools. Todas as funções e comandos do StatTools estão descritos neste arquivo.

Manual on-line

Abre o manual on-line do StatTools

O comando **Manual on-line do menu Ajuda** abre este manual no formato PDF. É necessário ter o Adobe Acrobat instalado para exibir o manual on-line.

Comando Ativação da licença

Exibe as informações de licenciamento do StatTools e permite o licenciamento de versões de teste.

O comando Ativação da licença do menu Ajuda exibe a caixa de diálogo Ativação da licença, listando as informações de versão e licenciamento da sua cópia do StatTools. Com esta caixa de diálogo você também pode converter uma versão de teste do StatTools em uma cópia licenciada.

Para mais informações sobre como licenciar a sua cópia do StatTools, consulte o **Capítulo 1: Primeiros Passos**, deste manual.

Comando Sobre

Exibe informações de versão e direitos autorais do StatTools

O comando **Sobre** do menu **Ajuda** exibe a caixa de diálogo Sobre, listando informações de versão e direitos autorais da sua cópia do StatTools.

Referência: Funções do StatTools

Introdução

Funções personalizadas de planilha são usadas pelo StatTools para retornar estatística calculada para fórmulas do Excel. Essas funções permitem:

- 1) Cálculos estatísticos a serem embutidos nas fórmulas da planilha, assim como as funções padrão do Excel.
- 2) Estatísticas em “tempo real”, ou seja, os resultados mudam quando os dados originais mudam.

Se observar as fórmulas nas células de um relatório do StatTools, você verá as funções da planilha do StatTools. Todas as funções do StatTools começam com o prefixo “Stat”, como **StatMean()** ou **StatStdDev()**. Todas as funções do StatTools são exibidas na caixa de diálogo **Inserir função do Excel** para facilitar a entrada.

Funções do StatTools vs funções do Excel

Em alguns casos, o StatTools substitui as funções estatísticas incorporadas no Excel pelos próprios cálculos sólidos e rápidos. A exatidão dos cálculos estatísticos incorporados no Excel tem sido muitas vezes questionada, e o StatTools não usa nenhum deles! Até mesmo as funções estatísticas padrão do Excel – como DESVPAD() – são substituídas por versões novas e mais robustas do StatTools, como a StatSTDEV(). Os cálculos estatísticos do StatTools satisfazem os mais rigorosos testes de exatidão, e tem o desempenho otimizado pelo uso de .DLLs de C++, e não cálculos por macros.

As funções do StatTools, ao contrário das funções incorporadas no Excel, aceitam o uso de dados empilhados. A função **StatDestack** do StatTools automaticamente desempilha os dados de um conjunto de dados empilhados (para a categoria especificada). Depois, transfere esses dados para um função estatística do StatTools para análise.

As funções do StatTools também aceitam a análise dos dados residentes em planilhas diferentes. Os conjuntos de dados com várias planilhas aceitam mais de 65.535 pontos por variável. Esses pontos são inseridos clicando no botão **Múltiplos** da caixa de diálogo Gerenciador de conjunto de dados.

Funções de distribuição

O StatTools inclui um conjunto de funções de distribuição (como a **StatBinomial**) que substitui as funções de distribuição incorporadas no Excel (como **BinomDist**). Ao contrário das funções de distribuição do Excel, as funções de distribuição do StatTools podem retornar vários valores diferentes de uma distribuição de probabilidade. O valor retornado é definido pelo argumento **estatística** (o penúltimo argumento da função). Este argumento pode ser um valor de 1 a 12 ou um string que indica a estatística desejada para a distribuição inserida:

Lista de possíveis estatísticas retornadas

Valor ou string inserido	Estatística retornada
1 ou "média"	média
2 ou "stddev"	desvio padrão
3 ou "variance"	variância
4 ou "skewness"	assimetria
5 ou "kurtosis"	curtose
6 ou "mode"	modo
7 ou "discrete media"	média discreta (ou o valor mais próximo da média verdadeira que realmente poderia ocorrer)
8 ou "x to y"	x para y (a distribuição do valor y para um valor x inserido)
9 ou "x to p"	x para p (a distribuição do valor p para um valor x inserido)
10 ou "p to x"	p para x (a distribuição do valor x para um valor p inserido)
11 ou "x to q"	x para q (a distribuição do valor q para um valor x inserido)
12 ou "q to x"	q para x (a distribuição do valor x para um valor q inserido)

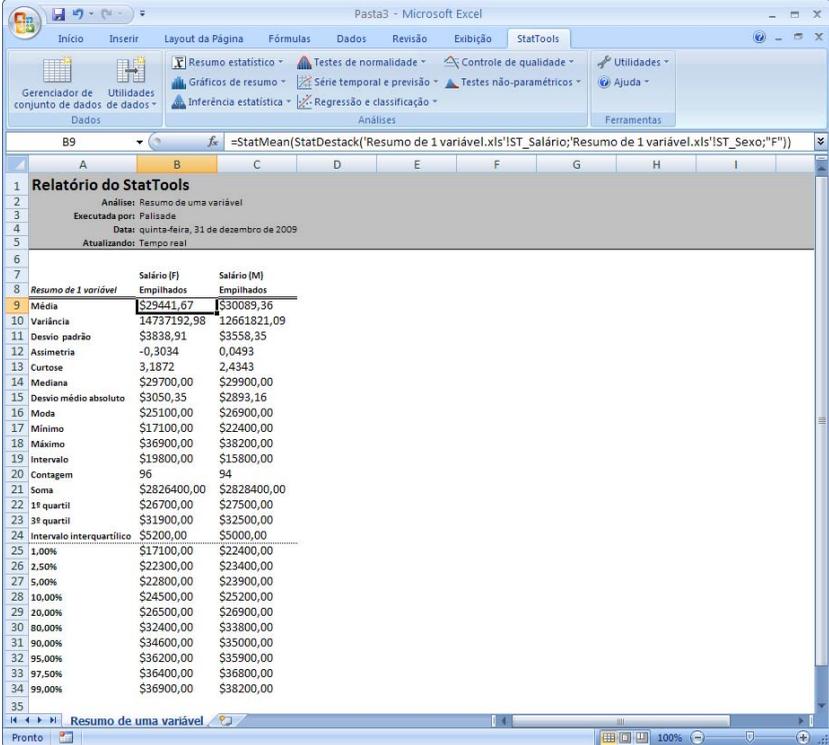
Por exemplo, a função de distribuição do StatTools:

StatNormal(10;1;"x to p";9,5)

Retorna o valor p associado ao valor x de 9.5 em uma distribuição normal com uma média de 10 e um desvio padrão de 1.

Relatórios “em tempo real”

O StatTools usa funções personalizadas para fazer com que os resultados sejam o mais “em tempo real” possível. Ou seja, sempre que possível, os relatórios têm fórmulas vinculadas aos dados originais. Por exemplo, suponha que haja uma variável *Peso* e você queira indicadores de resumo sobre *Peso*, como média e desvio padrão. O procedimento Resumo estatístico intitula a gama de pesos com *Peso*, e em seguida insere as fórmulas nas células de output: **=StatMean(Peso)** e **=StatStdDev(Peso)**. *StatMean* e *StatStdDev* são incorporados nas funções do StatTools para o cálculo da média e do desvio padrão. Elas substituem as funções incorporadas no Excel para as mesmas estatísticas. Devido a essas funções, quando os seus dados mudam, os resultados mudam automaticamente, assim não é necessário executar o procedimento mais uma vez.



The screenshot shows the Microsoft Excel interface with the StatTools ribbon active. The formula bar displays the formula `=StatMean(StatDestack('Resumo de 1 variável.xls'!ST_Salário;Resumo de 1 variável.xls'!ST_Sexo;'F'))`. The active cell (B9) contains the value `$29441,67`. The report titled "Relatório do StatTools" provides the following data:

Resumo de 1 variável		
	Salário (F)	Salário (M)
Média	\$29441,67	\$30089,36
Variancia	14737192,98	12661821,09
Desvio padrão	\$3838,91	\$3558,35
Assimetria	-0,3034	0,0493
Curtose	3,1872	2,4343
Mediana	\$29700,00	\$29900,00
Desvio médio absoluto	\$3050,35	\$2899,16
Moda	\$23100,00	\$26900,00
Mínimo	\$17100,00	\$22400,00
Máximo	\$36900,00	\$38200,00
Intervalo	\$19800,00	\$15800,00
Contagem	96	94
Soma	\$2826400,00	\$2828400,00
1º quartil	\$26700,00	\$27500,00
3º quartil	\$31900,00	\$32500,00
Intervalo interquartilico	\$5200,00	\$5000,00
1.00%	\$17100,00	\$22400,00
2.50%	\$22300,00	\$23400,00
5.00%	\$22800,00	\$23900,00
10.00%	\$24500,00	\$25200,00
20.00%	\$26500,00	\$26900,00
30.00%	\$32400,00	\$33800,00
40.00%	\$34600,00	\$35000,00
50.00%	\$36200,00	\$35900,00
60.00%	\$36400,00	\$36800,00
70.00%	\$36900,00	\$38200,00

Às vezes não é prático vincular os resultados aos dados. O exemplo por excelência é a regressão. O StatTools não fornece as fórmulas utilizadas para criar o output da regressão, mas apenas os resultados numéricos. Nesses casos, se os seus dados forem alterados, os procedimentos deverão ser executados novamente.

É possível usar a opção **Estático** do comando **Configurações de relatórios** para que os seus relatórios não fiquem vinculados aos dados. Isto é útil se o tempo recálculo do Excel se tornar um problema sempre que os dados forem alterados.

Referência: listagem das funções estatísticas

Tabela de funções disponíveis

Esta tabela lista as funções personalizadas adicionadas ao Excel pelo StatTools. Quando usadas, todas as funções são precedidas pelo prefixo **Stat**.

Função	Retornos
AUTOCORRELATION(<i>dados</i> ; <i>númDesfasagens</i>)	Calcula a autocorrelação para os valores em um conjunto de dados
AVEDEV(<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>)	Calcula o desvio médio absoluto dos dados com relação à sua média. Os argumentos podem ser números, vetores ou intervalos.
BINOMIAL(<i>n</i> ; <i>p</i> ; <i>estatística</i> ; <i>valor</i>)	Calcula a <i>estatística</i> para a distribuição binomial inserida.
CATEGORYINDICES(<i>intervalo</i> ; <i>nome_categoria</i>)	Obtém os índices de células para a categoria especificada.
CATEGORYNAMES(<i>intervalo</i>)	Obtém os nomes das categorias em um intervalo.
CATEGORYOCCURRENCECOUNT(<i>intervalo</i> ; <i>nome_categoria</i>)	Calcula o número de células no intervalo da categoria especificada.
CHISQ(<i>graus_liberdade</i> ; <i>estatística</i> ; <i>valor</i>)	Calcula a <i>estatística</i> para a distribuição qui-quadrado monocaudal.
CORRELATIONCOEFF(<i>dados1</i> ; <i>dados2</i>)	Calcula o coeficiente de correlação entre dois conjuntos de dados.
COUNT(<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>)	Calcula o número de elementos em seus argumentos, que podem ser números, vetores ou intervalos.
COUNTCATEGORIES(<i>intervalo</i>)	Conta o número de categorias do intervalo.
COUNTCELLSBYTYPE(<i>intervalo</i> ; <i>tipo</i>)	Calcula o número de células do tipo especificado no intervalo.
COUNTRANGE(<i>intervalo</i> ; <i>valorMin</i> ; <i>valorMax</i> ; <i>incluirMin</i> ; <i>incluirMax</i>)	Calcula o número de valores no <i>intervalo</i> que fica entre o <i>valorMin</i> e o <i>valorMax</i> .
COVARIANCE(<i>dados1</i> ; <i>dados2</i>)	Calcula a covariância da amostra entre 2 conjuntos de dados.
COVARIANCEP(<i>dados1</i> ; <i>dados2</i>)	Calcula a covariância da população entre 2 conjuntos de dados. Quaisquer valores omitidos retornam uma célula em branco.

Função	Retornos
DESTACK(<i>dados_intervalo; categorias_intervalo_1; categoria_1; categorias_intervalo_2; categoria_2</i>)	Extrai dados de uma categoria específica de dados empilhados.
DURBINWATSON(<i>dados</i>)	Calcula a estatística Durbin-Watson para os valores em um conjunto de dados.
F(<i>graus_liberdade1; graus_liberdade2; estatística; valor</i>)	Calcula uma <i>estatística</i> para a distribuição F para 2 conjuntos de dados.
GETCELLVALUES(<i>intervalo</i>)	Obtém os valores para todas as células de um tipo especificado no intervalo.
KURTOSIS(<i>dados1; dados2; ... dadosN</i>)	Calcula a curtose amostral de seus argumentos, que podem ser números, vetores ou intervalos.
KURTOSISP(<i>dados1; dados2; ... dadosN</i>)	Calcula a curtose da população de seus argumentos, que podem ser números, vetores ou intervalos.
LN(<i>x</i>)	Calcula o logaritmo natural do número real, positivo.
MAX(<i>dados1; dados2; ... dadosN</i>)	Calcula o máximo de seus argumentos, que podem ser números, vetores ou intervalos.
MEAN(<i>dados1; dados2; ... dadosN</i>)	Calcula a média aritmética(média) de seus argumentos, que podem ser números, vetores ou intervalos.
MEANABS(<i>dados1; dados2; ... dadosN</i>)	Calcula a média aritmética(média) dos valores absolutos de seus argumentos, que podem ser números, vetores ou intervalos.
MEDIAN(<i>dados; flag_discreta</i>)	Calcula a mediana do conjunto de dados.
MIN(<i>dados1; dados2; ... dadosN</i>)	Calcula o mínimo de seus argumentos, que podem ser números, vetores ou intervalos.
NORMAL(<i>média; desv_pad; estatística; valor</i>)	Calcula a <i>estatística</i> para a distribuição normal inserida(Gaussiana).
PAIRCOUNT(<i>dados1; dados2</i>)	Conta o número de pares de células para os quais cada uma das células do par é numérica.
PAIRMEAN(<i>dados1; dados2</i>)	Calcula a média das diferenças entre pares de células.

Função	Retornos
PAIRMEDIAN(<i>dados1; dados2</i>)	Calcula a mediana das diferenças entre pares de células.
PAIRSTDDEV(<i>dados1; dados2</i>)	Calcula o desvio padrão da amostral das diferenças entre pares de células.
PERCENTILE(<i>dados; p; flag_discreta</i>)	Calcula o percentil p para o conjunto de dados.
PRODUCT(<i>dados1; dados2; ... dadosN</i>)	Calcula o produto de seus argumentos, que podem ser números, vetores ou intervalos.
QUARTILE(<i>dados; q; flag_discreta</i>)	Calcula o quartil especificado do conjunto de dados.
RAND()	Retorna o número aleatório no intervalo 0 – 1.
RANGE(<i>dados1; dados2; ... dadosN</i>)	Calcula o intervalo(máximo - mínimo) de seus argumentos, que podem ser números, vetores ou intervalos.
RUNSTEST(<i>dados; valor_corte</i>)	Calcula a estatística “runs” para valores do conjunto de dados.
SKEWNESS(<i>dados1; dados2; ... dadosN</i>)	Calcula a assimetria amostral de seus argumentos, que podem ser números, vetores ou intervalos.
SKEWNESSP(<i>dados1; dados2; ... dadosN</i>)	Calcula a assimetria da população de seus argumentos, que podem ser números, vetores ou intervalos.
STANDARDIZE(<i>x; média; desv_pad</i>)	Calcula o valor normalizado da distribuição com a média e desvio padrão especificados.
STDDEV(<i>dados1; dados2; ... dadosN</i>)	Calcula o desvio padrão amostral de seus argumentos, que podem ser números, vetores ou intervalos.
STDDEVP(<i>dados1; dados2; ... dadosN</i>)	Calcula o desvio padrão da população de seus argumentos, que podem ser números, vetores ou intervalos.
SUM(<i>dados1; dados2; ... dadosN</i>)	Calcula a soma de seus argumentos, que podem ser números, vetores ou intervalos.
SUMDEVSQ(<i>dados1; dados2; ... dadosN</i>)	Calcula a soma dos quadrados dos desvios da média de seus argumentos, que podem ser números, vetores ou intervalos.

Função	Retornos
SUMSQ(<i>dados1; dados2; ... dadosN</i>)	Calcula a soma do quadrado de seus argumentos, que podem ser números, vetores ou intervalos.
STUDENT(<i>graus_liberdade; estatística; valor</i>)	Calcula a <i>estatística</i> para a distribuição t de Student.
VARIANCE(<i>dados1; dados2; ... dadosN</i>)	Calcula a variância amostral de seus argumentos, que podem ser números, vetores ou intervalos.
VARIANCEP(<i>dados1; dados2; ... dadosN</i>)	Calcula a variância da população de seus argumentos, que podem ser números, vetores ou intervalos.

Descrições detalhadas das funções

As funções estatísticas estão aqui mencionadas com os respectivos argumentos necessários.

AUTOCORRELATION

Descrição	AUTOCORRELATION (<i>dados</i> ; <i>númDesfasagens</i>) calcula a autocorrelação para os valores em um conjunto de dados usando o número de defasagens <i>numLags</i> . <i>dados</i> é a matriz ou intervalo de dados a que se destina o cálculo da autocorrelação. <i>numLags</i> é o número de defasagens a ser usado.
Exemplos	StatAutocorrelation (C1:C100; 1) retorna a autocorrelação para os dados dentro do intervalo C1:C100 com 1 defasagem.
Diretrizes	<i>númDesfasagens</i> precisa ser igual a ou maior que 1.

AVEDEV

Descrição	AVEDEV (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula o desvio médio absoluto de <i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i> com relação à sua média.
Exemplos	StatAveDev (1; 2; 5) calcula o desvio médio absoluto de 1, 2 e 5 com relação à sua média.
Diretrizes	Os argumentos <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

BINOMIAL

Descrição	BINOMIAL (<i>N</i> ; <i>P</i> ; <i>estatística</i> ; <i>valor</i>) calcula a <i>estatística</i> para a distribuição binomial com os valores <i>N</i> e <i>P</i> especificados.
Exemplos	StatBinomial (2; .3; "x to p"; 1) calcula o valor de p para um valor x de 1 na distribuição binomial com N=2 e P = 0; 3
Diretrizes	<i>N</i> é o número de tentativas ou eventos e precisa ser um número inteiro >0. <i>P</i> é a probabilidade e precisa ser >=0 e <=1. <i>estatística</i> é um valor inteiro entre 1 e 12 ou um vetor indicando a estatística a ser retornada. Para obter mais informações, consulte a seção Funções de distribuição no início deste capítulo. <i>valor</i> deve ser >= 0 (quando for inserido um valor para x) ou >=0 e <=1 (quando for inserido um valor para p).

CATEGORYINDICES

Descrição	CATEGORYINDICES (<i>intervalo</i> ; <i>categoria_name</i>) obtém os índices de células contendo o <i>categoria_name</i> especificado com o <i>intervalo</i> do Excel. Esta é uma função de vetor e os índices retornados são números de índice (entre 1 e # das células no <i>intervalo</i>) em função das posições das células contendo <i>categoria_name</i> dentro do <i>intervalo</i> .
Exemplos	StatCategoryIndices (C1:C100; "Masculino") retorna os índices (entre 1 e 100) das células com o vetor Masculino.
Diretrizes	<i>intervalo</i> é um intervalo válido do Excel <i>categoria_name</i> é um vetor, valor ou referência de célula que especifica a categoria a ser procurada.

CATEGORYNAMES

Descrição	CATEGORYNAMES (<i>intervalo</i>) obtém os nomes das categorias no <i>intervalo</i> especificado do Excel. Esta é uma função de vetor e o número de nomes retornados fica entre 1 e # das células do <i>intervalo</i> .
Exemplos	StatCategoryNames (C1:C100) retorna os nomes das categorias no intervalo C1:C100.
Diretrizes	<i>intervalo</i> é um intervalo válido do Excel

CATEGORYOCCURRENCECOUNT

Descrição	CATEGORYOCCURRENCECOUNT (<i>intervalo</i> ; <i>nome_categoria</i>) retorna o número de células contendo o <i>nome_categoria</i> especificado no <i>intervalo</i> do Excel.
Exemplos	StatCategoryOccurrenceCount (C1:C100; "Masculino") retorna o número de células no intervalo C1:C100 que contenha "Masculino".
Diretrizes	<i>intervalo</i> é um intervalo válido do Excel <i>categoria_name</i> é um vetor, valor ou referência de célula que especifica a categoria a ser procurada.

CHISQ

Descrição	CHISQ (<i>graus_liberdade</i> ; <i>estatística</i> ; <i>valor</i>) calcula a <i>estatística</i> para a distribuição qui-quadrado monocaudal usando os graus de liberdade especificados <i>graus_liberdade</i> .
Exemplos	StatChisq (2; "x to p"; 5) calcula a distribuição qui-quadrado monocaudal no valor 5 com 2 graus de liberdade.
Diretrizes	O número de graus de liberdade <i>graus_liberdade</i> devem estar no intervalo 1-32767. <i>estatística</i> é um valor inteiro entre 1 e 12 ou um vetor indicando a estatística a ser retornada. Para obter mais informações, consulte a seção Funções de distribuição no início deste capítulo. <i>valor</i> deve ser ≥ 0 (quando for inserido um valor para x) ou ≥ 0 e ≤ 1 (quando for inserido um valor para p).

CORRELATIONCOEFF

Descrição	CORRELATIONCOEFF (<i>dados1</i> ; <i>dados2</i>) calcula o coeficiente de correlação entre dois conjuntos de dados <i>dados1</i> e <i>dados2</i> .
Exemplos	StatCorrelationCoeff (A1:A100; B1:B100) calcula o coeficiente de correlação entre dois conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> devem ter o mesmo número de elementos.

COUNT

Descrição	COUNT (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula o número de elementos em <i>dados1</i> , <i>dados2</i> a <i>dadosN</i> , que podem ser números, vetores ou intervalos.
Exemplos	StatCount (A1:A100; B1:B100) calcula o número de elementos nos dois conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> são 1 a 30 argumentos, que podem ser números, vetores ou intervalos.

COUNTCATEGORIES

Descrição	COUNTCATEGORIES (<i>intervalo</i>) retorna o número de categorias no <i>intervalo</i> especificado do Excel.
Exemplos	StatCountCategories (C1:C100) retorna o número das categorias no intervalo C1:C100.
Diretrizes	<i>intervalo</i> é um intervalo válido do Excel

COUNTCELLSBYTYPE

Descrição	COUNTCELLSBYTYPE (<i>intervalo</i> ; <i>tipo</i>) calcula o número de elementos no <i>intervalo</i> inserido do <i>tipo</i> especificado.
Exemplos	StatCountCellsByType (A1:A100; 1) calcula o número de elementos numéricos no conjunto de dados localizado em A1:A100.
Diretrizes	<i>intervalo</i> é um intervalo válido do Excel. <i>tipo</i> é 1=numérico, 2=não vazio, 3=não vazio, não numérico, 4=vazio. Nota: StatTools considera vazia a célula que só contém espaços.

COUNTRANGE

Descrição	COUNTRANGE (<i>intervalo</i> ; <i>valorMin</i> ; <i>valorMax</i> ; <i>incluirMin</i> ; <i>incluirMax</i>) calcula o número de valores no <i>intervalo</i> que fica entre <i>valorMin</i> e <i>valorMax</i> . Valores que correspondam a <i>valorMin</i> e <i>valorMax</i> podem ser incluídos configurando <i>incluirMin</i> e/ou <i>incluirMax</i> como VERDADEIRO.
Exemplos	StatCountRange (A1:A100; 1; 10; VERDADEIRO; VERDADEIRO) calcula o número de valores no conjunto de dados localizado em A1:A100 que ficam no intervalo ≥ 1 e ≤ 10 .
Diretrizes	<i>intervalo</i> é o intervalo de células do qual se deseja contar valores. <i>valorMin</i> é o valor mínimo do intervalo. <i>valorMax</i> é o valor máximo do intervalo. <i>incluirMin</i> é um valor booleano indicando se o <i>valorMin</i> deve ser incluído na contagem. O padrão é VERDADEIRO. <i>incluirMax</i> é um valor booleano indicando se o <i>valorMax</i> deve ser incluído na contagem. O padrão é VERDADEIRO.

COVARIANCE

Descrição	COVARIANCE (<i>dados1</i> ; <i>dados2</i>) calcula a covariância amostral entre os dois conjuntos de dados <i>dados1</i> e <i>dados2</i> .
Exemplos	StatCovariance (A1:A100; B1:B100) calcula a covariância amostral entre dois conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> podem ser vetores ou intervalos. <i>dados1</i> e <i>dados2</i> devem ter o mesmo número de elementos.

COVARIANCEP

Descrição	COVARIANCEP (<i>dados1</i> ; <i>dados2</i>) calcula a covariância da população entre dois conjuntos de dados <i>dados1</i> e <i>dados2</i> .
Exemplos	StatCovarianceP (A1:A100; B1:B100) calcula a covariância da população entre dois conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> podem ser vetores ou intervalos.

DESTACK

Descrição	DESTACK (<i>dados_intervalo</i> ; <i>categorias_intervalo_1</i> ; <i>categoria_1</i> ; <i>categorias_intervalo_2</i> ; <i>categoria_2</i>) extrai dados de uma categoria específica <i>categoria_1</i> dos dados empilhados em <i>dados_intervalo</i> . Esta função permite outras funções estatísticas do StatTools tomarem dados empilhados como input. Assim, a função StatDestack será vista apenas como “inserida” em outras funções, como ilustrado no exemplo a seguir. Ela retorna um vetor de dados da categoria especificada, extraída de <i>dados_intervalo</i> .
Exemplos	StatMean(StatDestack (B1:B100; A1:A100; "Masculino")) calcula a média dos valores no intervalo B1:B100 em que o intervalo da categoria correspondente A1:A100 tem o valor “Masculino”.
Diretrizes	<i>dados_intervalo</i> é o intervalo com os valores empilhados. <i>categorias_intervalo_1</i> é o intervalo com o nome da primeira categoria. <i>categoria_1</i> é a primeira categoria para obter os dados. <i>categorias_intervalo_2</i> (<i>opcional</i>) é o intervalo com o nome da segunda categoria. <i>categoria_2</i> (<i>opcional</i>) é a segunda categoria para obter os dados.

DURBINWATSON

Descrição	DURBINWATSON (<i>dados</i>) calcula a estatística Durbin-Watson para os valores em um conjunto de dados <i>dados</i> .
Exemplos	StatDurbinWatson (A1:A100) calcula a estatística Durbin-Watson para o intervalo de dados localizado em A1:A100.
Diretrizes	<i>dados</i> pode ser um vetor ou intervalo de dados.

F

Descrição	F (<i>graus_liberdade1</i> ; <i>graus_liberdade2</i> ; <i>estatística</i> ; <i>valor</i>) calcula a <i>estatística</i> para a distribuição F usando o número de graus de liberdade do numerador <i>graus_liberdade1</i> e o número de graus de liberdade do denominador <i>graus_liberdade2</i> .
Exemplos	StatF (1; 1; "x to p"; 1,5) calcula o valor p da distribuição F para um valor de x de 1, 5 com um número de graus de liberdade do numerador =1 e um número de graus de liberdade do denominador =1.
Diretrizes	<i>graus_liberdade1</i> e <i>graus_liberdade2</i> deve ser um inteiro > 0. <i>estatística</i> é um valor inteiro entre 1 e 12 ou um vetor indicando a estatística a ser retornada. Para obter mais informações, consulte a seção Funções de distribuição no início deste capítulo. <i>valor</i> deve ser > 0.

GETCELLVALUES

Descrição	GETCELLVALUES (<i>intervalo</i> ; <i>typeOfCell</i>) obtém os valores para todas as células de um tipo especificado no intervalo.
Exemplos	StatGetCellValues (A1:A100; 2) obtém os valores para células não vazias no intervalo localizado em A1:A100.
Diretrizes	<i>intervalo</i> é o intervalo de células do qual se deseja contar valores. <i>typeOfCell</i> é 0 para todos, 1 para numérico, 2 para não vazio, 3 para não vazio e não numérico e 4 para células vazias.

KURTOSIS

Descrição	KURTOSIS (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula a curtose amostral dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> . Nota: StatKurtosis calculada sobre dados distribuídos normalmente retorna o valor 3.
Exemplos	StatKurtosis (A1:A100; {1;2;3;2,4}) calcula a curtose amostral do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

KURTOSISP

Descrição	KURTOSISP (<i>dados1; dados2; ... dadosN</i>) calcula a curtose da população dos dados especificados em <i>dados1, dados2, ... dadosN</i> . Nota: StatKurtosisP calculada sobre dados distribuídos normalmente retorna o valor 3.
Exemplos	StatKurtosisP (A1:A100; {1;2;3;2,4}) calcula a curtose da população do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

LN

Descrição	LN (<i>x</i>) calcula o logaritmo natural.
Exemplos	StatLN (4,5) calcula o logaritmo natural de 4, 5.
Diretrizes	<i>x</i> precisa ser um número real, positivo.

MAX

Descrição	MAX (<i>dados1; dados2; ... dadosN</i>) calcula o máximo dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatMax (A1:A100; {1;2;3;2,4}) calcula o valor máximo do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

MEAN

Descrição	MEAN (<i>dados1; dados2; ... dadosN</i>) calcula a média dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatMean (A1:A100; {1;2;3;2,4}) calcula o valor médio do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

MEANABS

Descrição	MEANABS (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula a média do valor absoluto dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> .
Exemplos	StatMeanAbs (A1:A100; {1;2;3;2,4}) calcula a média dos valores absolutos do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

MEDIAN

Descrição	MEDIAN (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula a mediana dos valores localizados em <i>dados</i> . Pode ser calculada usando qualquer um dos cinco métodos alternativos, como opcionalmente especificado por <i>calcFlag</i> .
Exemplos	StatMedian (A1:A100; 1) calcula o valor da mediana no conjunto de dados localizado em A1:A100. Os dados são contínuos.
Diretrizes	<i>dados</i> é um intervalo do Excel. <i>calcFlag</i> é um argumento opcional que pode ter um valor inteiro no intervalo -1 a 5. Este valor corresponde ao método desejado para cálculo do percentil. -1 ou não especificado) Automático (baseados nos dados de input) 0) Igual à função de percentil do Excel (contínuo) 1) Interpolado com pontos terminais assimétricos (contínuo) 2) Observação mais próxima (discreta) 3) Função de dist. empírica (discreta) 4) Interpolado com pontos terminais simétricos (contínuo) 5) Função de dist. empírica com cálculo de média (discreta)

MIN

Descrição	MIN (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula o mínimo dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> .
Exemplos	StatMin (A1:A100; {1;2;3;2,4}) calcula o valor mínimo do conjunto de dados localizado em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

MODE

Descrição	MODE (<i>dados; discreta</i>) calcula o modo de um conjunto de dados.
Exemplos	StatMode (A1:A100; FALSO) calcula o modo do conjunto de dados localizado em A1:A100. Os dados são contínuos.
Diretrizes	<i>dados</i> é a matriz ou intervalo de dados a que se destina o cálculo do modo. <i>discreta</i> é um argumento opcional que especifica se os dados devem ser tratados como discretos (verdadeiro) ou contínuos (falso). Se não estiver presente, será determinado automaticamente com base nos dados.

NORMAL

Descrição	NORMAL (<i>média; desv_pad; estatística; valor</i>) calcula a <i>estatística</i> da distribuição normal especificada pela <i>média</i> e por <i>desv_pad</i> .
Exemplos	StatNormal (2; 1; "x to p"; 3) calcula o valor de p para um valor x de 3 na distribuição normal com média=2 e desvio padrão=1
Diretrizes	<i>média</i> é a média aritmética da distribuição. <i>desv_pad</i> é o desvio padrão da distribuição. Deve ser > 0. <i>estatística</i> é um valor inteiro entre 1 e 12 ou um vetor indicando a estatística a ser retornada. Para obter mais informações, consulte a seção Funções de distribuição no início deste capítulo. <i>valor</i> deve ser >=0 e <=1 se for inserido um valor para p.

PAIRCOUNT

Descrição	PAIRCOUNT (<i>dados1; dados2</i>) conta o número de pares de células em <i>dados1</i> e <i>dados2</i> . Apenas pares de células numéricas são contados. Um valor ausente em qualquer dos intervalos não será contado.
Exemplos	StatPairCount (A1:A100, B1:B100) conta o número de pares de células numéricas nos conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> devem ser do mesmo tamanho dos intervalos do Excel. Pares de células de <i>dados1</i> e <i>dados2</i> são selecionados por linha, começando no canto superior esquerdo.

PAIRMEAN

Descrição	PAIRMEAN (<i>dados1</i> , <i>dados2</i>) calcula a média das diferenças entre os pares de células em <i>dados1</i> e <i>dados2</i> . Os cálculos são feitos apenas para pares de células numéricas. Um valor ausente em qualquer dos intervalos não será contado.
Exemplos	StatPairMean (A1:A100; B1:B100) calcula a média das diferenças entre os pares de células numéricas nos conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> devem ser do mesmo tamanho dos intervalos do Excel. Pares de células de <i>dados1</i> e <i>dados2</i> são selecionados por linha, começando no canto superior esquerdo.

PAIRMEDIAN

Descrição	PAIRMEDIAN (<i>dados1</i> ; <i>dados2</i> ; <i>flag_método</i>) calcula a mediana das diferenças entre os pares de células em <i>dados1</i> e <i>dados2</i> . Os cálculos são feitos apenas para pares de células numéricas. Um valor ausente em qualquer dos intervalos não será contado. A mediana pode ser calculada usando qualquer um dos cinco métodos alternativos, como opcionalmente especificado por <i>flag_método</i> .
Exemplos	StatPairMedian (A1:A10; B1:B100) calcula a mediana das diferenças entre os pares de células numéricas nos conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> devem ser do mesmo tamanho dos intervalos do Excel. Pares de células de <i>dados1</i> e <i>dados2</i> são selecionados por linha, começando no canto superior esquerdo. <i>flag_método</i> é um argumento opcional que pode ter um valor inteiro no intervalo -1 a 5. Este valor corresponde ao método desejado para cálculo da mediana. -1 ou não especificado) Automático (baseados nos dados de input) 0) Igual à função de percentil do Excel (contínuo) 1) Interpolado com pontos terminais assimétricos (contínuo) 2) Observação mais próxima (discreta) 3) Função de dist. empírica (discreta) 4) Interpolado com pontos terminais simétricos (contínuo) 5) Função de dist. empírica com cálculo de média (discreta)

PAIRSTDDEV

Descrição	PAIRSTDDEV (<i>dados1</i> ; <i>dados2</i>) calcula o desvio padrão amostral das diferenças entre os pares de células em <i>dados1</i> e <i>dados2</i> . Os cálculos são feitos apenas para pares de células numéricas. Um valor ausente em qualquer dos intervalos não será contado.
Exemplos	StatPairStdDev (A1:A100; B1:B100) calcula o desvio padrão amostral das diferenças entre os pares de células numéricas nos conjuntos de dados localizados em A1:A100 e B1:B100.
Diretrizes	<i>dados1</i> e <i>dados2</i> devem ser do mesmo tamanho dos intervalos do Excel. Pares de células de <i>dados1</i> e <i>dados2</i> são selecionados por linha, começando no canto superior esquerdo.

PERCENTILE

Descrição	PERCENTILE (<i>dados</i> ; <i>p</i> ; <i>flag_método</i>) calcula o percentil <i>p</i> de <i>dados</i> . A mediana pode ser calculada usando qualquer um dos cinco métodos alternativos, como opcionalmente especificado por <i>flag_método</i> .
Exemplos	StatPercentile (A1:A100; 0,15; 0) calcula o décimo quinto percentil dos dados localizados em A1:A100. Os dados são contínuos e usam o método de cálculo de percentil pela média ponderada.
Diretrizes	<i>p</i> deve ser 0-1, inclusive. <i>flag_método</i> é um argumento opcional que pode ter um valor inteiro no intervalo -1 a 5. Este valor corresponde ao método desejado para cálculo do percentil. -1 ou não especificado) Automático (baseados nos dados de input) 0) Igual à função de percentil do Excel (contínuo) 1) Interpolado com pontos terminais assimétricos (contínuo) 2) Observação mais próxima (discreta) 3) Função de dist. empírica (discreta) 4) Interpolado com pontos terminais simétricos (contínuo) 5) Função de dist. empírica com cálculo de média (discreta)

PRODUCT

Descrição	PRODUCT (<i>dados1; dados2; ... dadosN</i>) calcula o produto dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatProduct (A1:A10; {1;2;3;2,4}) calcula o produto de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

QUARTILE

Descrição	QUARTILE (<i>dados; q; flag_método</i>) calcula o quartil especificado de <i>dados</i> . Os quartis podem ser calculados usando qualquer um dos cinco métodos alternativos, como opcionalmente especificado por <i>flag_método</i> .
Exemplos	StatQuartile (A1:A100; 1; FALSO) calcula o primeiro quartil dos dados localizados em A1:A100. Os dados são contínuos.
Diretrizes	<i>dados</i> deve ser um intervalo do Excel. Q é o quartil; 0 = mínimo, 1 = 1o quartil, 2 = 2o quartil (mediana), 3 = 3o quartil, 4 = máximo. <i>flag_método</i> é um argumento opcional que pode ter um valor inteiro no intervalo -1 a 5. Este valor corresponde ao método desejado para cálculo do percentil. -1 ou não especificado) Automático (baseados nos dados de input) 0) Igual à função de percentil do Excel (contínuo) 1) Interpolado com pontos terminais assimétricos (contínuo) 2) Observação mais próxima (discreta) 3) Função de dist. empírica (discreta) 4) Interpolado com pontos terminais simétricos (contínuo) 5) Função de dist. empírica com cálculo de média (discreta)

RAND

Descrição	RAND () retorna um número aleatório no intervalo 0 a 1. Esta função usa o gerador de número aleatório do produto @RISK da Palisade e não o gerador de número aleatório que faz parte do Excel.
Exemplos	StatRand () retorna um número aleatório no intervalo 0 a 1.

RANGE

Descrição	RANGE (<i>dados1; dados2; ... dadosN</i>) calcula o intervalo (máximo - mínimo) dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatRange (A1:A100; {1;2;3;2,4}) calcula o intervalo (máximo - mínimo) dos dados localizados em A1:A100 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

RUNSTEST

Descrição	RUNSTEST (<i>dados; valor_corte</i>) calcula a estatística "runs" para os valores em <i>dados</i> usando valor de <i>corte</i> .
Exemplos	StatRunsTest (A1:A100; StatMean (A1:A100)) calcula a estatística do teste "run" nos dados em A1:A100 usando a média dos dados como valor de corte.
Diretrizes	<i>dados</i> deve ser um intervalo do Excel.

SKEWNESS

Descrição	SKEWNESS (<i>dados1; dados2; ... dadosN</i>) calcula a assimetria amostral dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatSkewness (A1:A10; {1;2;3;2,4}) calcula a assimetria amostral de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

SKEWNESSP

Descrição	SKEWNESSP (<i>dados1; dados2; ... dadosN</i>) calcula a assimetria da população dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatSkewnessP (A1:A10; {1;2;3;2,4}) calcula a assimetria da população de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

STANDARDIZE

Descrição	STANDARDIZE (<i>x</i> ; <i>média</i> ; <i>desv_pad</i>) calcula um valor normalizado da distribuição com a <i>média</i> especificada e o desvio padrão <i>desv_pad</i> . <i>x</i> é o valor a ser normalizado.
Exemplos	StatStandardize (2; 1; 3) calcula um valor normalizado no valor 2 da distribuição com média 1 e desvio padrão de 3.
Diretrizes	<i>x</i> é o valor a ser normalizado. <i>média</i> é a média aritmética da distribuição. <i>desv_pad</i> é o desvio padrão da distribuição. Deve ser > 0.

STDDEV

Descrição	STDDEV (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula o desvio padrão amostral dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> .
Exemplos	StatStdDev (A1:A10; {1;2;3;2,4}) calcula o desvio padrão amostral de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

STDDEVP

Descrição	STDDEVP (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula o desvio padrão da população dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> .
Exemplos	StatStdDevP (A1:A10; {1;2;3;2,4}) calcula o desvio padrão da população de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

SUM

Descrição	SUM (<i>dados1</i> ; <i>dados2</i> ; ... <i>dadosN</i>) calcula a soma dos dados especificados em <i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> .
Exemplos	StatSum (A1:A10; {1;2;3;2,4}) calcula a soma de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1</i> , <i>dados2</i> , ... <i>dadosN</i> podem ser números, vetores ou intervalos.

SUMDEVSQ

Descrição	SUMDEVSQ (<i>dados1; dados2; ... dadosN</i>) calcula a soma do quadrado do desvio da média de seus argumentos, que podem ser números, vetores ou intervalos. Quaisquer valores omitidos retornam uma célula em branco.
Exemplos	StatSumDevSq (A1:A10; {1;2;3;2,4}) calcula a soma do quadrado do desvio da média de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

SUMSQ

Descrição	SUMSQ (<i>dados1; dados2; ... dadosN</i>) calcula a soma do quadrado de seus argumentos, que podem ser números, vetores ou intervalos. Quaisquer valores omitidos retornam uma célula em branco.
Exemplos	StatSumSq (A1:A10; {1;2;3;2,4}) calcula a soma do quadrado de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

STUDENT

Descrição	STUDENT (<i>graus_liberdade; estatística; valor</i>) calcula a <i>estatística</i> da distribuição t inserida de Student.
Exemplos	StatStudent (5; 1; "x to p"; 2) calcula o valor p da distribuição t de Student, com 5 graus de liberdade no valor de x = 2.
Diretrizes	<i>graus_liberdade</i> é um inteiro indicando o número de graus de liberdade. Deve estar no intervalo 1-32767. <i>estatística</i> é um valor inteiro entre 1 e 12 ou um vetor indicando a estatística a ser retornada. Para obter mais informações, consulte a seção Funções de distribuição no início deste capítulo. <i>x</i> é o valor numérico no qual avaliar a distribuição. Deve ser >= 0.

VARIANCE

Descrição	VARIANCE (<i>dados1; dados2; ... dadosN</i>) calcula a variância amostral dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatVariance (A1:A10; {1;2;3;2,4}) calcula a variância amostral de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

VARIANCEP

Descrição	VARIANCE (<i>dados1; dados2; ... dadosN</i>) calcula a variância da população dos dados especificados em <i>dados1, dados2, ... dadosN</i> .
Exemplos	StatVarianceP (A1:A10; {1;2;3;2,4}) calcula a variância amostral de todos os valores do conjunto de dados localizado em A1:A10 e os valores 1, 2, 3 e 2,4.
Diretrizes	<i>dados1, dados2, ... dadosN</i> podem ser números, vetores ou intervalos.

Índice

- Amostra aleatória, comando, 62
- Análise discriminante, comando, 132
- ANOVA bidirecional, comando, 96
- ANOVA unidirecional, comando, 93
- Ativação da licença, comando, 175
- Autocorrelação, comando, 113
- Autorização, 175
- Barra de ferramentas
 - StatTools, 27
- Combinação, comando, 58
- Configurações da aplicação, comando, 167
- Correlações e covariância, comando, 68
- Defasagem, comando, 52
- Descarregar o add-in StatTools, comando, 173
- Desempilhar variáveis, comando, 49
- Desinstalação do StatTools, 7
- Diferença, comando, 54
- Empilhar variáveis, comando, 47
- Excluir conjuntos de dados, comando, 173
- Fictício, comando, 60
- Funções do StatTools, 177
- Funções StatTools
 - Lista alfabética das, 183
 - StatAutocorrelation, 187
 - StatAveDev, 187
 - StatBinomial, 187
 - StatCategoryIndices, 188
 - StatCategoryNames, 188
 - StatCategoryOccurrenceCount, 188
 - StatChiDist, 189
 - StatCorrelationCoeff, 189
 - StatCount, 189
 - StatCountCellByType, 190
 - StatCountRange, 190
 - StatCovariance, 190
 - StatCovarianceP, 191
 - StatDestack, 191
 - StatDurbinWatson, 191
 - StatF, 192
 - StatGetCellValues, 192
 - StatKurtosis, 192
 - StatKurtosisP, 193
 - StatLN, 193
 - StatMax, 193
 - StatMean, 193
 - StatMeanAbs, 194
 - StatMedian, 194
 - StatMin, 194
 - StatNormal, 195
 - StatPairCount, 195
 - StatPairMean, 196
 - StatPairMedian, 196
 - StatPairStdDev, 197
 - StatPercentile, 197
 - StatProduct, 198
 - StatQuartile, 198
 - StatRand, 198
 - StatRange, 199
 - StatRunsTest, 199
 - StatSkewness, 199
 - StatSkewnessP, 199
 - StatStandardize, 200
 - StatStdDev, 200
 - StatStdDevP, 200
 - StatStudent, 201
 - StatSum, 200
 - StatSumDevSq, 201
 - StatSumSq, 201
 - StatVariance, 202
 - StatVarianceP, 202
- Gerenciador de conjunto de dados, comando, 41

Gráfico Box-Whisker, comando, 76

Gráfico de dispersão, comando, 74

Gráfico de Pareto, comando, 137

Gráfico de série temporal, comando, 111

Gráfico normal Q-Q, comando, 108

Gráficos C, comando, 148

Gráficos P, comando, 144

Gráficos U, comando, 151

Gráficos X/R, comando, 140

Histograma, comando, 71

Ícones

- Área de trabalho, 8
- StatTools, 27

Instruções de instalação, 7

Interação, comando, 56

Intervalo de confiança –

- Proporção, comando, 82

Intervalo de confiança de média/desvio padrão, comando, 79

Limpar memória da caixa de diálogo, comando, 173

Menus

- Menu Ajuda (Janela de modelo), 175
- Menu da série temporal e previsão, 137, 155
- Menu Inferência estatística, 79
- Menu Regressão e classificação, 123
- Menu Resumo estatístico, 65
- Menu Série temporal e previsão, 111
- Menu Testes de normalidade, 101
- Menu Utilidades, 167
- Menu Utilidades de dados, 47

Palisade Corporation, 5

Previsão, comando, 117

Regressão logística, comando, 128

Regressão, comando, 123

Requisitos de sistema, 6

Resumo de 1 variável, comando, 65

Seleção de tamanho amostral, comando, 91

Sobre, comando, 175

Teste de execuções para verificar aleatoriedade, comando, 115

Teste de hipótese - média/desvio padrão, comando, 85

Teste de hipótese - Proporção, comando, 88

Teste de independência de qui-quadrado, comando, 98

Teste de Lilliefors, comando, 105

Teste de Mann-Whitney, comando, 163

Teste de normalidade de qui-quadrado, comando, 101

Teste de sinal, comando, 157

Teste dos postos sinalizados de Wilcoxon, comando, 160

Transformar variáveis, comando, 50