

Linhas de tendência

Bem-vindo a este vídeo sobre linhas de tendência.

Você pode baixar a pasta de trabalho de exercícios para acompanhar os procedimentos usando sua própria cópia do Tableau.

Adicionar linhas de tendência

As tendências podem destacar informações importantes em uma análise. Por exemplo, responder à pergunta “qual é o aumento da potência de saída em relação à velocidade do vento”? Podemos ver a relação aqui: quando a velocidade do vento aumenta, a potência de saída também aumenta. Mas quanto, exatamente?

É muito fácil adicionar linhas de tendência à exibição.

- Clique no painel Análise e arraste Linha de tendência para o tipo de modelo desejado.
- Remover as linhas de tendência é tão fácil quanto adicioná-las, basta arrastá-las para fora da exibição. Como queremos manter nossa linha de tendência, vou desfazer sua remoção.

Opções da linha de tendência

Por padrão, as linhas de tendência são aplicadas por painel e por cor.

- Na guia Dados, se adicionarmos outra dimensão à exibição, como Local, teremos uma linha de tendência por painel ou por gráfico de dispersão.
- Do mesmo modo, se arrastarmos Local para Cor, nossa linha de tendência será dividida em três.
- Se quisermos ver nossos Locais por Cor, mas tivermos apenas uma tendência geral, podemos modificar a linha de tendência.

Para editar a linha de tendência, basta clicar com o botão direito do mouse nela, selecionar Linhas de tendência e Editar linhas de tendência.

- Desmarcar a opção “Permitir uma linha de tendência por cor” exibirá novamente apenas uma linha de tendência geral.
- Também podemos desmarcar a opção Mostrar faixas de confiança para simplificar a exibição

Há muitas outras coisas que podemos fazer com esta caixa de diálogo.

- Primeiro, o tipo de modelo.
 - Estas são as mesmas opções que temos quando adicionamos uma linha de tendência pelo painel Análise
 - Essas opções solicitam que o Tableau crie um modelo de regressão linear com base neste tipo de transformação em uma ou nas duas variáveis.
 - A linearidade se refere aos coeficientes, e não à relação das variáveis
 - Para obter mais informações sobre tipos de modelo e transformações, consulte o artigo [Tipos de modelo da linha de tendência](#) na Ajuda on-line.
- As faixas de confiança mostram o intervalo de confiança de 95% para o modelo.
- E também temos a opção de forçar a interceptação de y em 0.

Significância da linha de tendência

É importante avaliar se uma linha de tendência fornece informações valiosas.

- Posicionar o ponteiro do mouse sobre a linha de tendência exibe uma dica de ferramenta com a equação da linha de tendência, o valor p e o coeficiente de determinação
- Em estatística, o valor p é o número atribuído ao conceito de significância.
 - Se o valor p for menor do que um valor de corte, geralmente 0,05, os resultados serão interpretados como significativos.
 - Um valor p alto (em uma escala de 0 a 1) pode indicar que a tendência aparente nos dados pode ser casual, e não resultante dos fatores no modelo.
- Neste exemplo, a linha de tendência tem um valor p muito pequeno, o que é bom

No entanto, para avaliar corretamente se um modelo é adequado, precisamos saber mais do que apenas o valor p .

- Também temos o coeficiente de determinação, que nos informa como nosso modelo se adequa aos dados. Os valores do coeficiente de determinação podem variar entre 0 e 1, e os valores altos são considerados melhores.
 - Em nosso exemplo, vemos que o coeficiente de determinação é bastante alto (0,956)
 - Isso significa que nosso modelo se adequa muito bem aos dados. Um coeficiente de determinação igual a 1 indicaria uma adequação perfeita, mas fique atento se ele for extremamente alta, por exemplo 0,999. Isso pode indicar que seu modelo é enganoso. Um indicador comum de um coeficiente de determinação artificialmente alto são graus de liberdade baixos ou um excesso de observações.

Resíduos da linha de tendência

Para decidir se uma linha de tendência representa os dados com precisão, não é suficiente ter um valor p baixo ou um coeficiente de determinação alto.

Nem todos os nossos pontos estarão na linha de tendência prevista. A distância de um determinado ponto até seu valor previsto é o erro, ou o resíduo.

Em um modelo correto, esses resíduos devem seguir uma distribuição normal aleatória próxima ao zero quando representados em relação à variável explicativa. Se esse gráfico de resíduos não estiver distribuído normalmente, isso indica que há tendências em como os dados *não conseguem* se alinhar aos valores previstos, o que significa que este não é o melhor modelo.

Para obter os valores residuais de uma exibição com uma linha de tendência,

- vá para Planilha > Exportar > Dados
- Essa operação vai solicitar que salvemos um arquivo (a única opção de formato é um arquivo do Microsoft Access). Vamos renomear esse arquivo como Resíduos da linha de tendência e clicar em Salvar.
- Em seguida, vamos selecionar “Conectar após exportação”
- Esta fonte de dados contém os dados originais do nosso gráfico de dispersão, os valores previstos (da linha de tendência) e os resíduos.

A representação dos resíduos é criada com a variável explicativa no eixo horizontal (Velocidade do vento em Colunas) e os resíduos no eixo vertical (em Linhas). Então, arrastamos Moinho de vento para Nível de detalhe.

Lembre-se que um bom modelo geralmente apresenta uma dispersão distribuída em torno de zero.

Claramente, nosso modelo atual não é muito bom para prever nossa Potência com base nos valores de Velocidade do vento. Embora o valor p e o coeficiente de determinação para a linha de tendência sejam bons, o gráfico de resíduos é péssimo.

CONCLUSÃO

Esta foi uma visão geral das linhas de tendência no Tableau e de como interpretar estatísticas relevantes relacionadas a elas. Obrigado por assistir a este vídeo. Sugerimos que você continue assistindo aos vídeos de treinamento gratuitos para aprender mais.