



ESTATÍSTICA ORIENTADA À CIÊNCIA DE DADOS



Gráficos

Por que utilizar gráficos?

São ferramentas poderosas para **explorar, comunicar e interpretar dados**;

A **escolha pelo gráfico adequado** para representar os **dados** pode transmitir informações de maneira clara e eficaz, auxiliando na análise e na tomada de decisões.

Exploração Visual de Padrões e Tendências nos Dados:

- Pode-se **identificar padrões, tendências e anomalias** nos dados que **podem não ser evidentes apenas olhando para os números;**
- Visualizar os dados nos permite **detectar relações, agrupamentos e variações** que podem fornecer *insights* valiosos sobre o fenômeno estudado.

Comunicação Eficaz dos Resultados:

- **Gráficos e visualizações tornam os resultados mais acessíveis e compreensíveis para um público amplo, incluindo aqueles sem conhecimento técnico avançado em estatísticas;**
- **As informações complexas podem ser comunicadas de maneira clara e concisa, facilitando a disseminação dos resultados e a colaboração entre diferentes partes interessadas.**

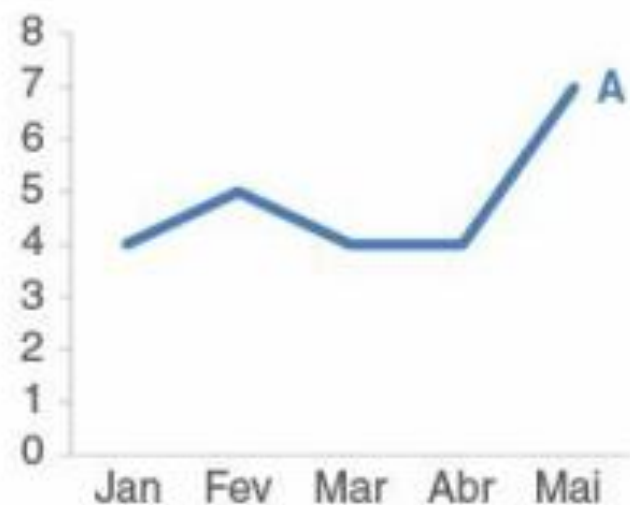
Comunicação Eficaz dos Resultados:

- **Gráficos e visualizações tornam os resultados mais acessíveis e compreensíveis para um público amplo, incluindo aqueles sem conhecimento técnico avançado em estatísticas;**
- **As informações complexas podem ser comunicadas de maneira clara e concisa, facilitando a disseminação dos resultados e a colaboração entre diferentes partes interessadas.**

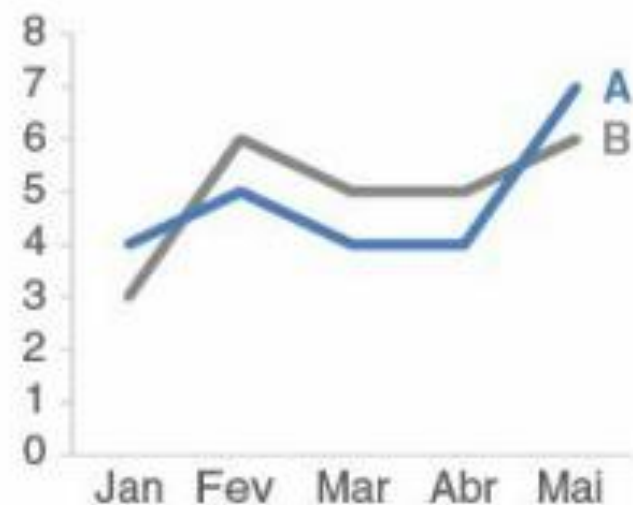
Gráficos “comuns”

Gráfico de linhas

Uma série



Duas séries



Várias séries

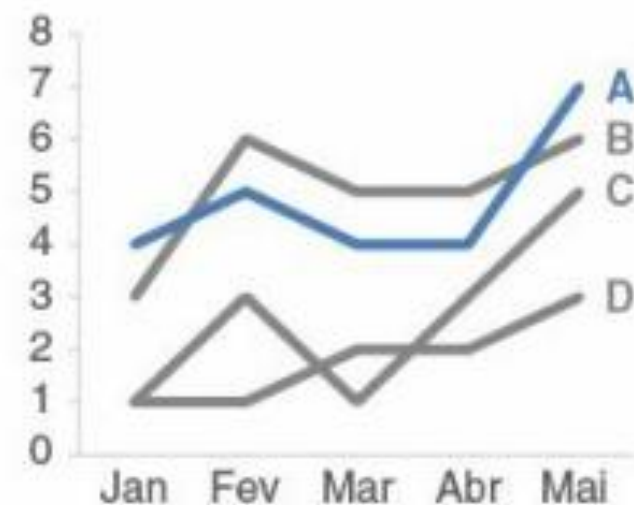
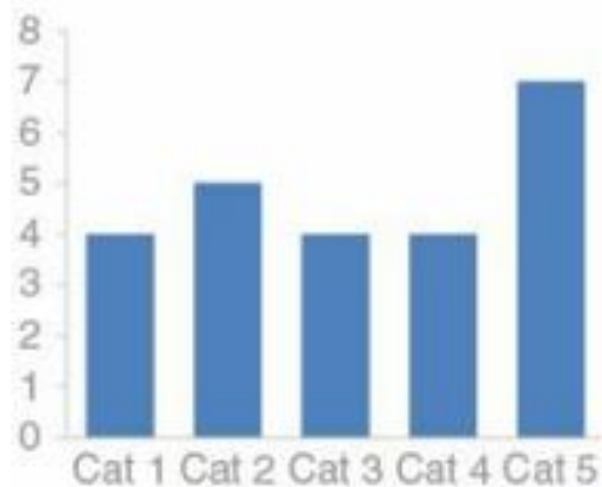


Gráfico de linhas

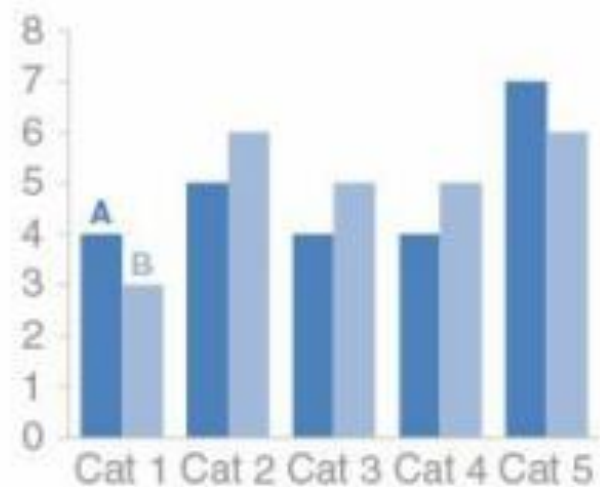
1. Os gráficos de linhas são mais usados para representar tendências ao longo do tempo ou relação entre variáveis;
2. Como os pontos são fisicamente conectados por meio da linha, isso indica uma conexão entre eles que pode não fazer sentido para dados categóricos;
3. Exemplo de dados adequados: **dias, meses, trimestres ou anos.**

Gráfico de barras

Uma série



Duas séries



Várias séries

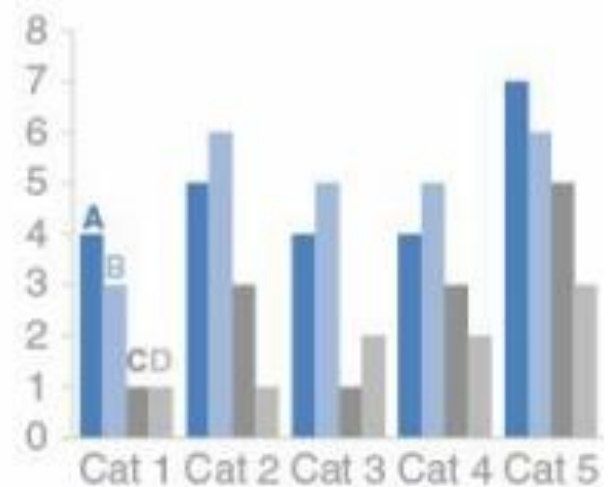


Gráfico de barras

1. Usado para **representar a frequência ou proporção de diferentes categorias;**
2. Excelente opção para **representar dados categóricos;**
3. Cada categoria é **representada por uma barra, cuja altura reflete a frequência.**

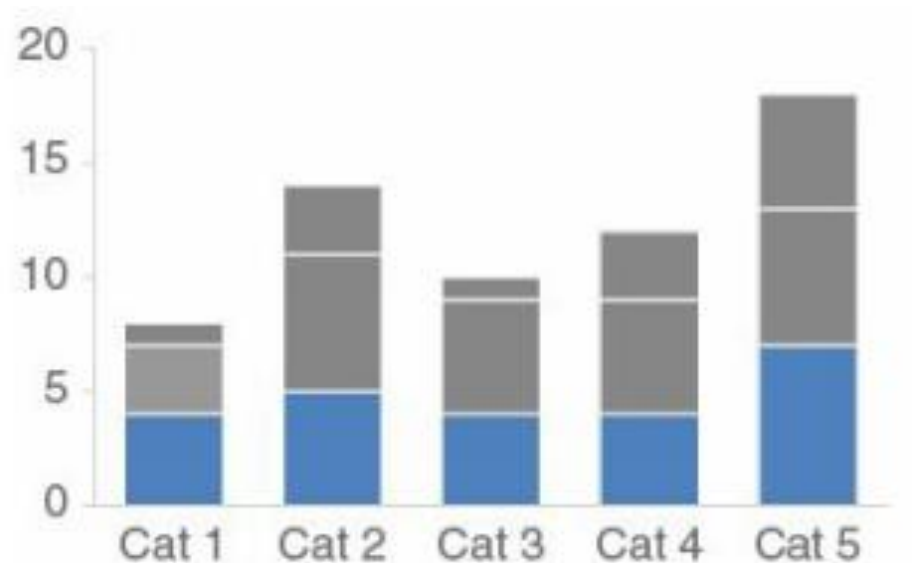


Gráfico de barras verticais empilhadas

Uma série



Gráfico de barras horizontais

Barras verticais empilhadas	Barras horizontais
Permite a comparação de totais entre categorias	Particularmente útil se seus nomes de categoria são longos
Possibilita a visualização das partes subcomponentes dentro de determinada categoria	Facilita a leitura, pois a estrutura é tal que nossos olhos atingem os nomes de categoria antes dos dados.
Pode ser estruturado como números absolutos; ou com cada coluna somando 100% (em que você representa a porcentagem do total de cada segmento vertical	Mostra os totais em diferentes categorias, mas também permite uma ideia das partes subcomponentes

Gráfico de área

Decomposição da entrevista



Gráfico de área

1. Mostra a **mudança na distribuição cumulativa ao longo do tempo ou entre diferentes categorias;**
2. **Difícil de ser compreendido**, pois os olhos humanos não conseguem atribuir bem o valor quantitativo no espaço bidimensional;
3. Recomendado quando é necessário **visualizar números de magnitudes amplamente diferentes.**

Gráfico de setores (pizza)

Participação no mercado de fornecedores

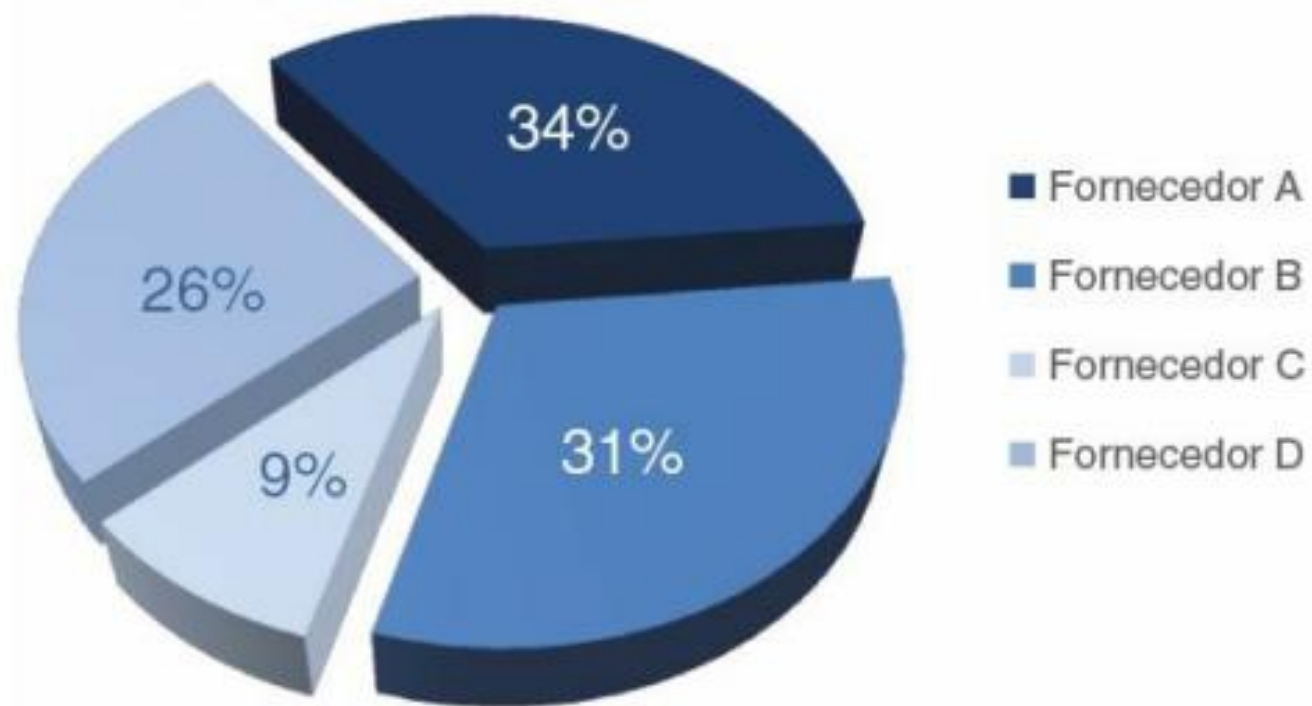
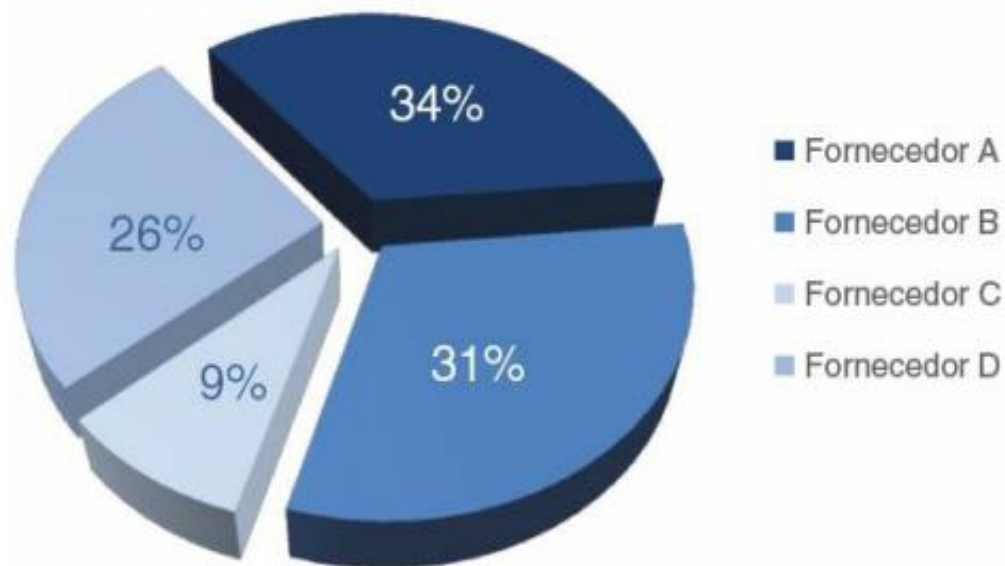


Gráfico de setores (pizza)

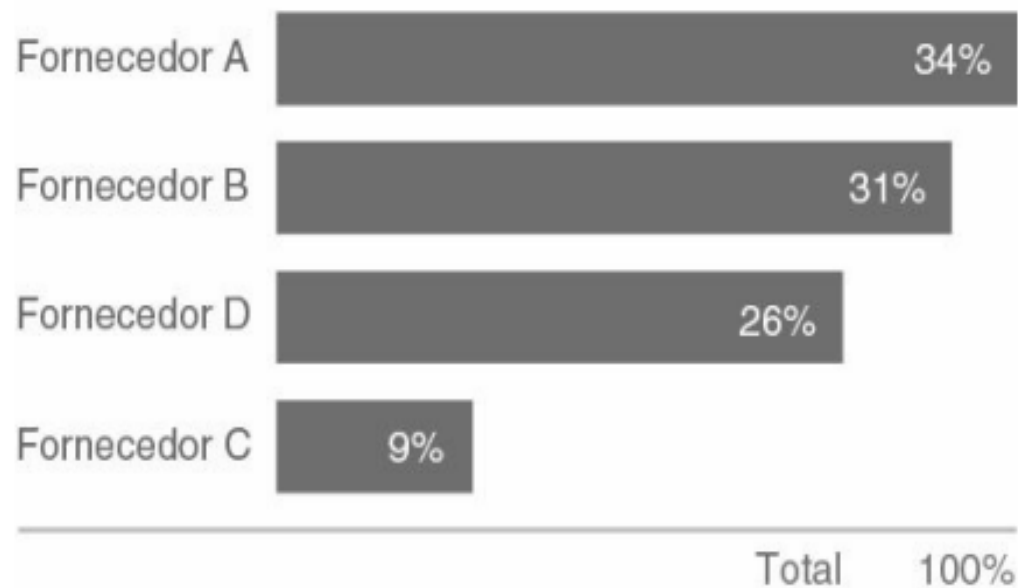
1. Mostra a **proporção de cada categoria em relação ao todo**. O círculo é dividido em setores, e o tamanho angular de cada setor representa a proporção da categoria;
2. É difícil ler os gráficos de pizza. Quando os segmentos têm tamanhos parecidos, é difícil (se não impossível) dizer qual é o maior;
3. Quando seus tamanhos não são parecidos, **tende-se a determinar que um é maior que outro, mas não se pode avaliar o quanto**.

Gráficos de setores podem ser alterados por gráficos de barras horizontais, por exemplo.

Participação no mercado de fornecedores



Participação no mercado de fornecedores



Gráficos estadísticos

Gráfico de dispersão

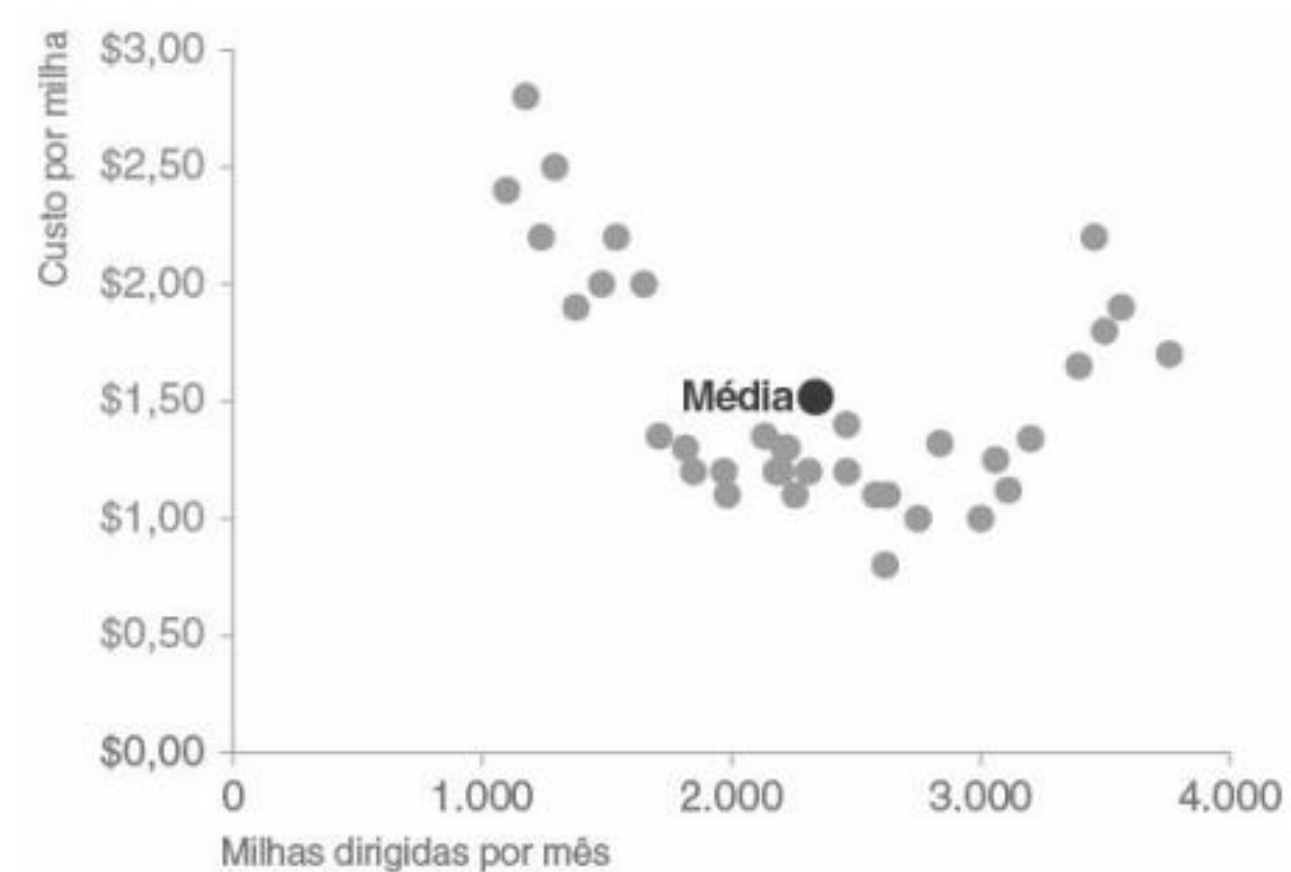
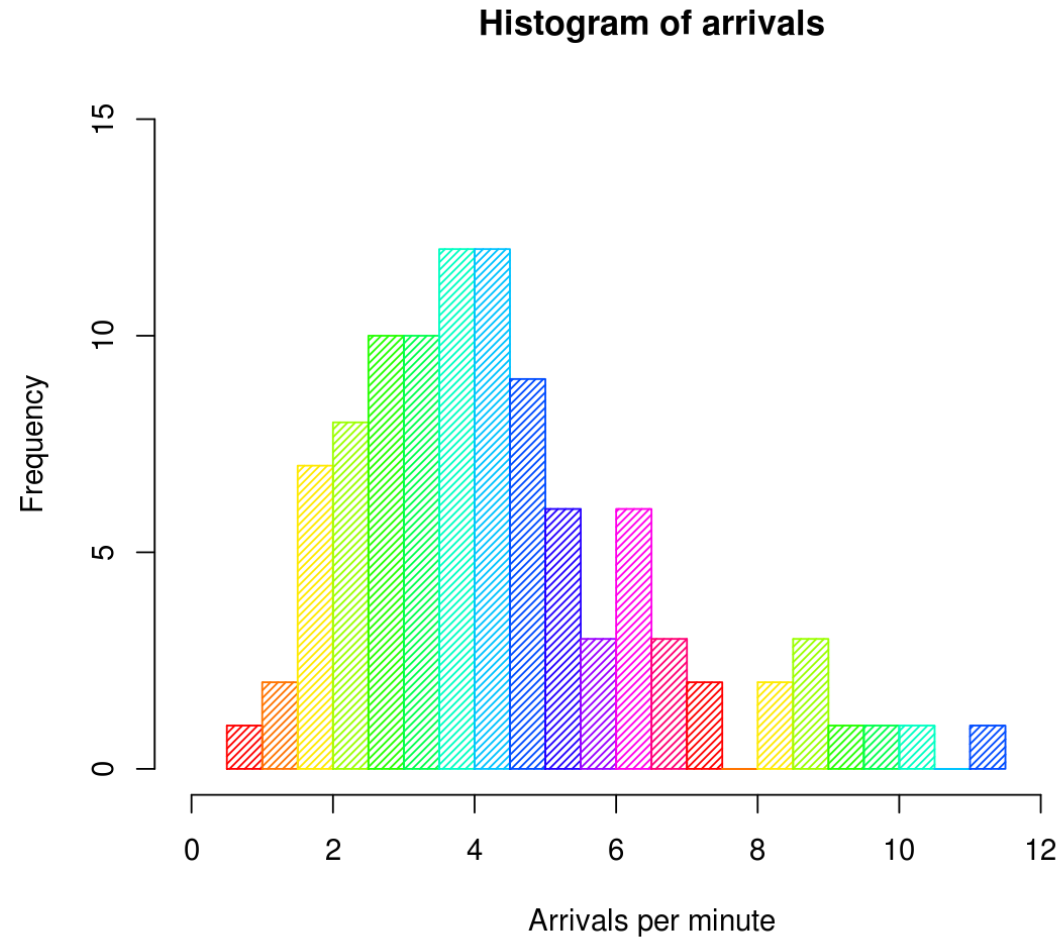


Gráfico de dispersão

1. Mostra a **relação entre duas variáveis numéricas**;
2. Cada **ponto representa uma observação**, e sua **posição no gráfico indica os valores das variáveis**;
3. Tendem a ser **mais usados nos campos científicos**.

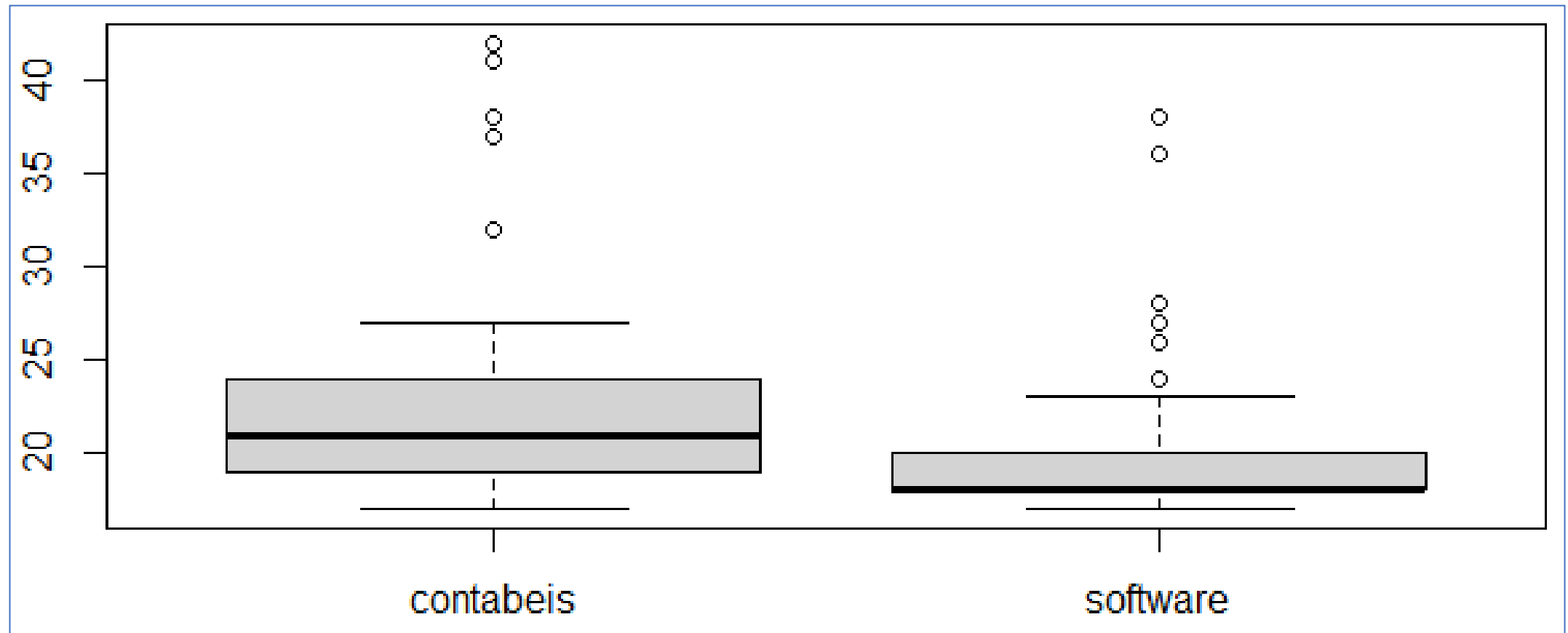
Histograma



Histograma

1. Representa a **distribuição de frequência de uma variável**;
2. **Divide os dados em intervalos (blocos)** e mostra a frequência com que os valores caem em cada intervalo;
3. Utilizado para **visualizar a forma da distribuição dos dados, identificar padrões, assimetrias e concentrar-se nas frequências em diferentes faixas.**

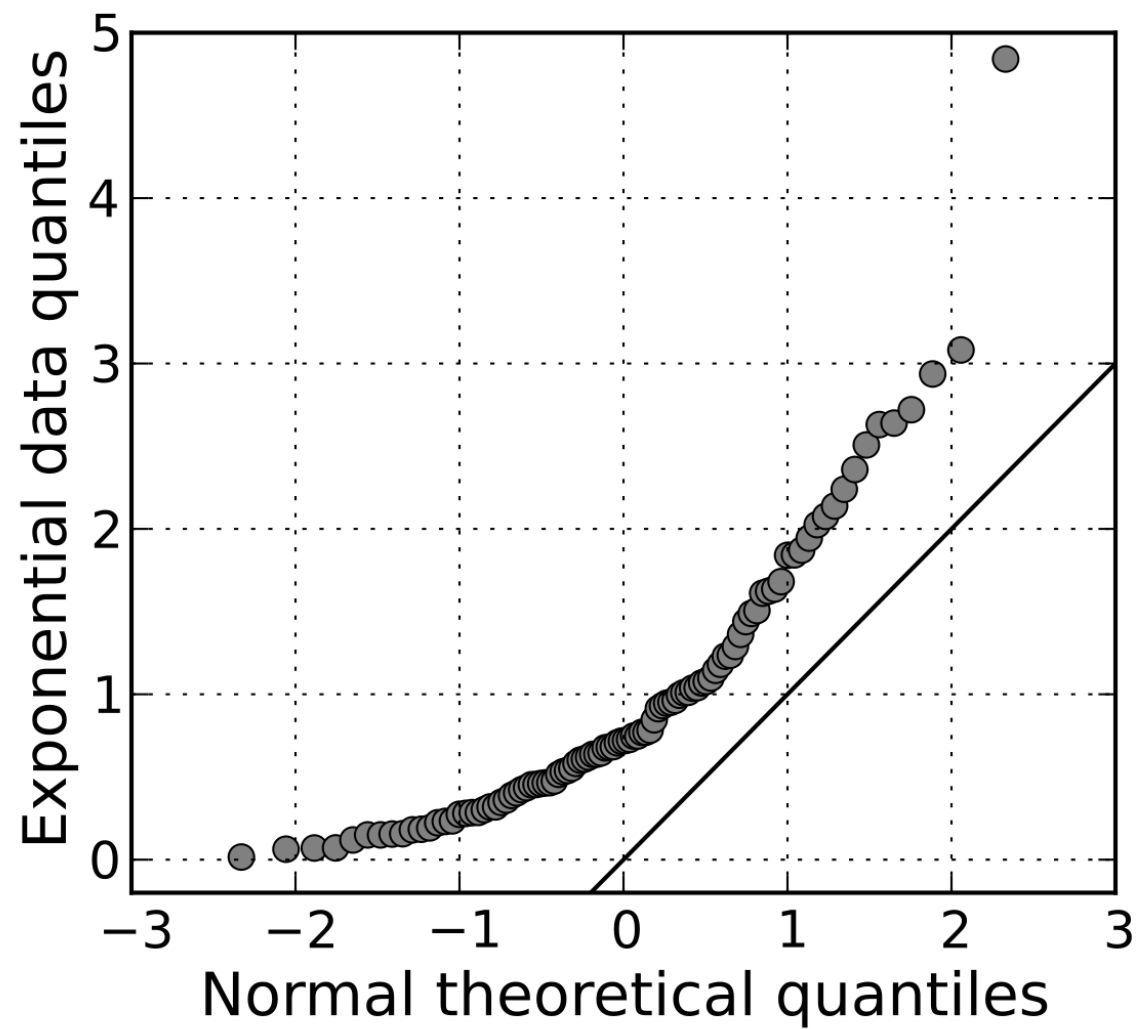
Boxplot (Diagrama de Caixa)



Boxplot (Diagrama de Caixa)

1. Exibe a **distribuição de um conjunto de dados, incluindo quartis, mediana e outliers.**
2. Ajuda a **identificar a dispersão e a presença de valores extremos;**
3. Apresenta a distribuição de um conjunto de dados através de cinco medidas principais: **valor mínimo, primeiro quartil (Q1), mediana, terceiro quartil (Q3) e valor máximo.**

QQ-plot



QQ-plot

1. Também conhecido como gráfico de probabilidade normal, é uma ferramenta gráfica **utilizada para avaliar a normalidade de um conjunto de dados em relação à distribuição normal teórica;**
2. É um gráfico de dispersão onde os **quantis (valores ordenados) dos dados observados são plotados em relação aos quantis esperados da distribuição normal.**



Prática

