

Visualização de Dados

Prof. Gustavo Willam Pereira



INSTITUTO FEDERAL
Sudeste de Minas Gerais

Como escolher a visualização certa para seus dados

- Ao visualizar os dados, é crucial usar o gráfico que melhor transmite sua mensagem de forma clara e precisa.
- Embora seus dados possam funcionar com várias opções de gráfico, você deve selecionar aquela que garante que o significado pretendido seja transmitido.
- Os dados só têm valor se você souber como exibi-los e contextualizá-los visualmente.
- Apresentaremos uma visão geral das diferentes categorias de gráficos e explicaremos como determinar o ajuste correto.

Como escolher a visualização certa para seus dados

- Primeiro, entenda a história que seus dados contam. Gráficos, mapas e infográficos ajudam o público a compreender números complexos, descobrir padrões, identificar tendências e receber informações. Considere o que você deseja enfatizar com sua visualização.
- Também é importante seguir as práticas recomendadas para gráficos. Seu visual deve ter números somados corretamente e dimensionados adequadamente.
- O que você pretende mostrar? Examinar cuidadosamente seus dados e objetivos o ajudará a escolher o tipo de visualização que melhor atende às suas necessidades.

Como escolher a visualização certa para seus dados

- Geralmente, existem quatro tipos principais de gráficos a serem considerados:
 - **Comparação**
 - **Composição**
 - **Relação**
 - **Distribuição**

1. Visualizações de comparação

- **Eles permitem examinar diferenças entre diferentes entidades.**
- **Gráficos de barras e gráficos de linhas** são visualizações de comparação comuns, pois mostram proporções e diferenças entre categorias.
- **Quando usar gráficos de barras?**
 - Comparando valores para diferentes categorias
 - Rastreando mudanças ao longo do tempo
 - quando as mudanças são grandes

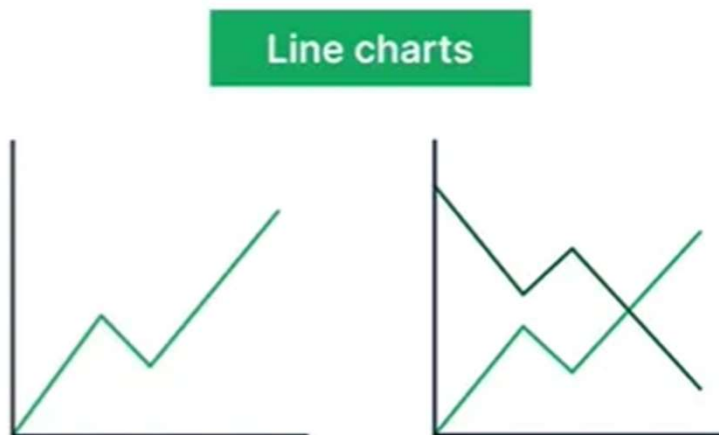
1. Visualizações de comparação



- Sua escolha entre um gráfico de barras horizontal e vertical pode ser orientada pela quantidade de texto que você precisa exibir na parte inferior de cada barra.
- Se você tiver mais texto, uma exibição horizontal costuma ser mais eficaz.

1. Visualizações de comparação

- Quando usar gráficos de linhas?
 - Rastreando mudanças ao longo do tempo
 - quando as mudanças são pequenas
 - Além disso, você pode comparar as alterações ao longo do tempo para mais de um grupo

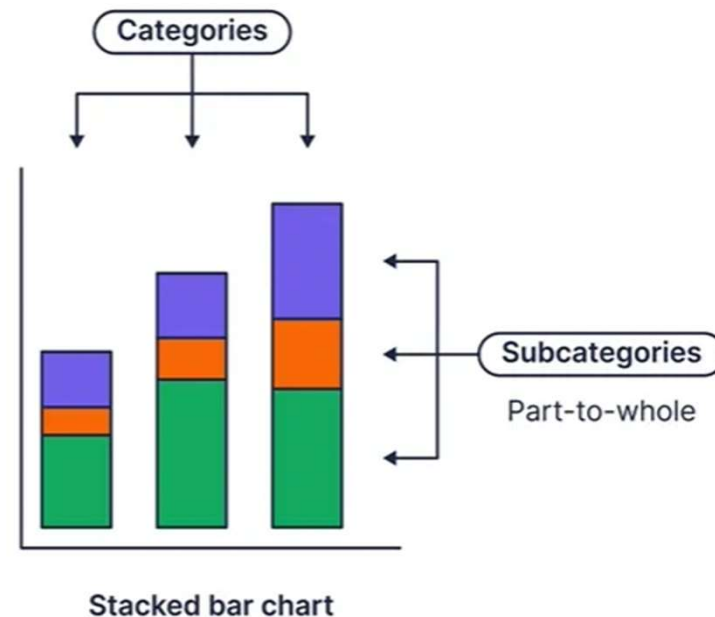


2. Visualizações de composição

- As visualizações de composição fornecem informações sobre como os componentes individuais se combinam para formar um sistema ou conjunto de dados completo. Esses recursos visuais são ideais para compreender a composição proporcional ou a estrutura hierárquica dos dados categorizados.
- Os gráficos de composição comumente usados são **gráficos de pizza, mapas de árvore, gráficos de barras, empilhadas e gráficos de áreas empilhadas**.

2. Visualizações de composição

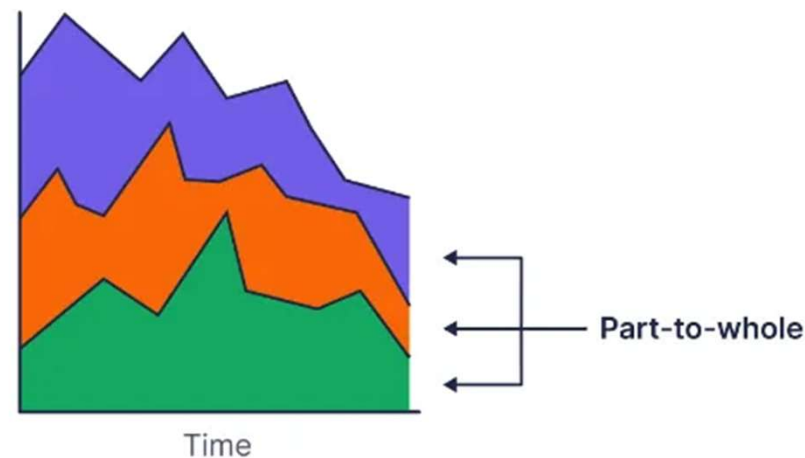
- Quando usar gráficos de barras empilhadas?
 - Como um gráfico de barras, ele compara diferentes categorias, cada barra ou categoria pode ser dividida em subcategorias.



- É útil comparar os totais entre categorias e também mostrar como cada subcategoria faz parte do todo.

2. Visualizações de composição

- Quando usar o gráfico de áreas empilhadas ?
 - O gráfico de áreas empilhadas é uma extensão do gráfico de linhas, enquanto o gráfico de linhas é usado para comparar valores de várias categorias ao longo do tempo, o gráfico de áreas empilhadas pode fazer o mesmo, ao mesmo tempo que mostra a composição de cada categoria como um todo.



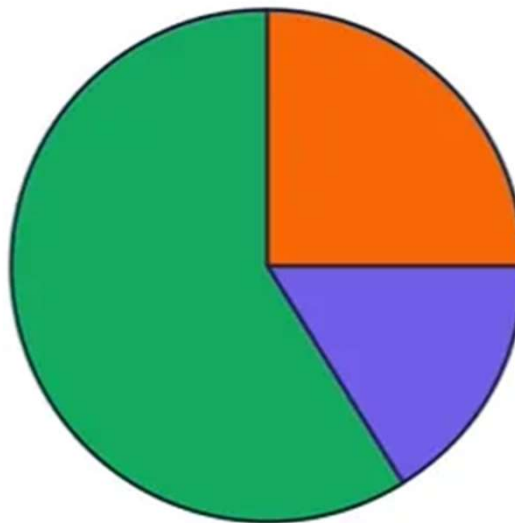
Stacked area chart

- A relação parte-todo será claramente indicada preenchendo as seções entre as linhas e o eixo x com cores diferentes. A soma de todas as áreas é igual a 100%.

2. Visualizações de composição

- Quando usar gráficos de pizza?

- O gráfico de pizza é ótimo para exibir a composição básica, mas é limitado pelo número de partes que você pode exibir.
- Quando as categorias têm tamanhos próximos, é difícil dizer qual é maior e torna-se difícil de ler.
- Um gráfico de pizza não deve ser selecionado se for dividido em mais de cinco partes.

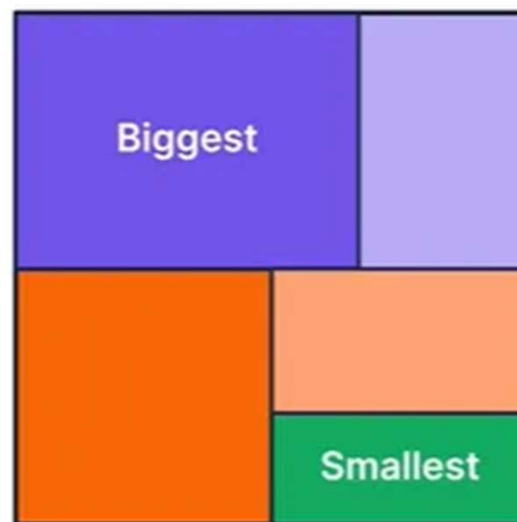


Pie chart

2. Visualizações de composição

- Quando usar Treemaps?

- Quando o todo consiste em cinco ou mais partes, um Treemap é uma boa escolha.
- Em vez de um círculo é usado um retângulo, o que facilita imediatamente a interpretação, pois não precisamos comparar ângulos, mas sim retângulos de tamanhos diferentes.
- Os retângulos serão ordenados em tamanho decrescente, do canto superior esquerdo para o canto inferior direito.



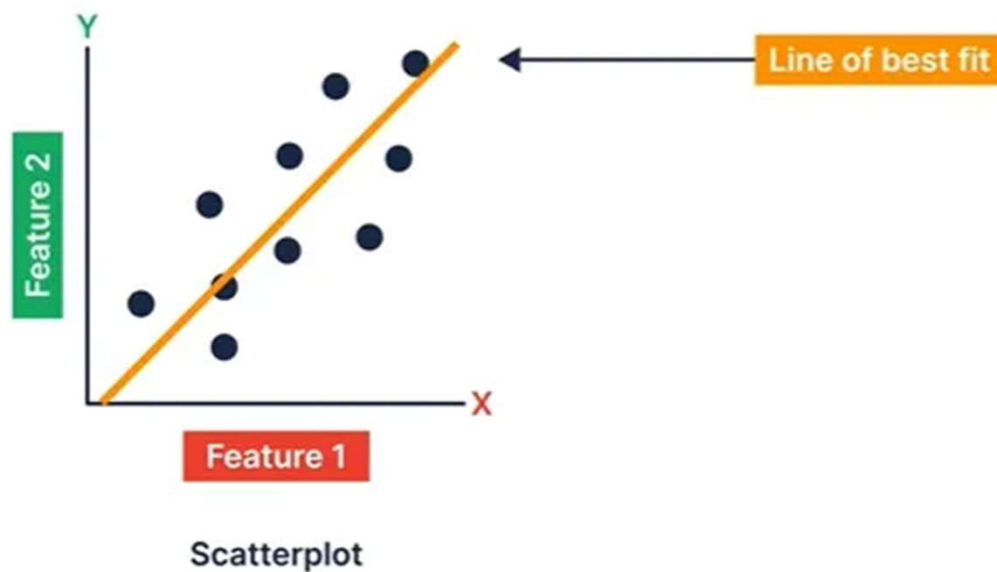
Treemap

3. Visualizações de relacionamento

- As visualizações de relacionamento revelam relacionamentos entre variáveis.
- **Gráficos de dispersão, gráficos de bolhas e mapas de calor** são bons para examinar correlações e tendências entre dois ou mais recursos numéricos . Consulte-os quando quiser entender como as mudanças em uma variável impactam outra.

3. Visualizações de relacionamento

- Quando usar o gráfico de dispersão?
- Se houver uma relação linear, uma linha de melhor ajuste poderá ser adicionada ao gráfico.



3. Visualizações de relacionamento

- Quando usar os mapas de calor ?
 - Se você tiver mais de três recursos numéricos em um conjunto de dados, um mapa de calor será ideal para nos ajudar a determinar quais dois recursos estão relacionados entre si de maneira eficiente em termos de tempo.
 - Um mapa de calor mostra a correlação entre cada dois recursos em um conjunto de dados.
 - Correlação é uma estatística que mede o quão forte ou fracamente duas características estão relacionadas e é medida de -1 negativo a 1 positivo, com zero representando nenhuma relação.

3. Visualizações de relacionamento

- Quando usar os mapas de calor ?



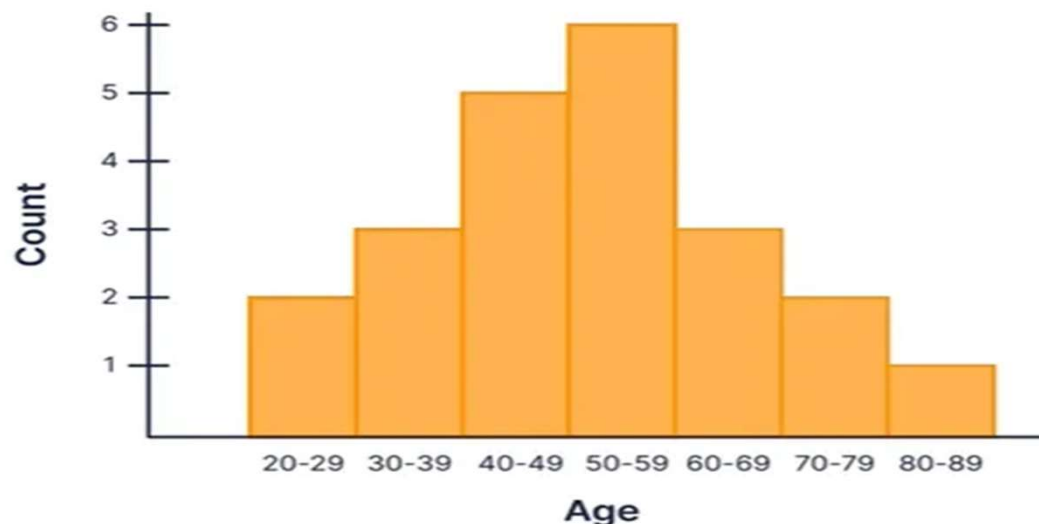
4. Visualizações de distribuição

- As visualizações de distribuição retratam a distribuição de frequência dos valores em um conjunto de dados. **Histogramas** , **gráficos de caixa** e **gráficos de densidade** representam efetivamente padrões na dispersão e concentração de valores.
- Isso ajuda a analisar **como os valores se agrupam e a identificar valores discrepantes** .

4. Visualizações de distribuição

- Quando usar histogramas?

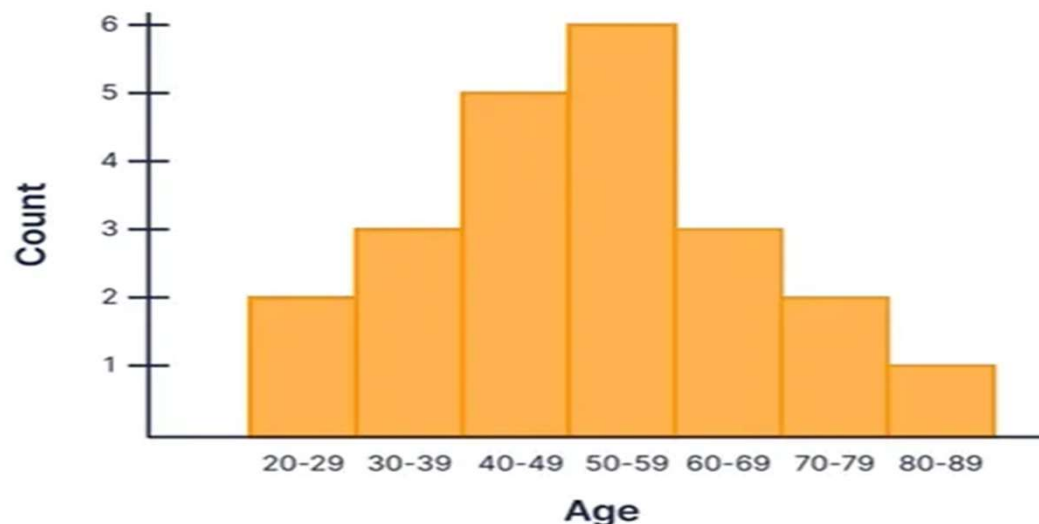
- Os histogramas agrupam dados contínuos (numéricos) em compartimentos de larguras fixas e contam o número de observações em cada compartimento.
- As caixas são exibidas como retângulos adjacentes plotados no eixo x, com altura indicando frequência.



4. Visualizações de distribuição

- Quando usar histogramas?

- Os histogramas agrupam dados contínuos (numéricos) em compartimentos de larguras fixas e contam o número de observações em cada compartimento.
- As caixas são exibidas como retângulos adjacentes plotados no eixo x, com altura indicando frequência.

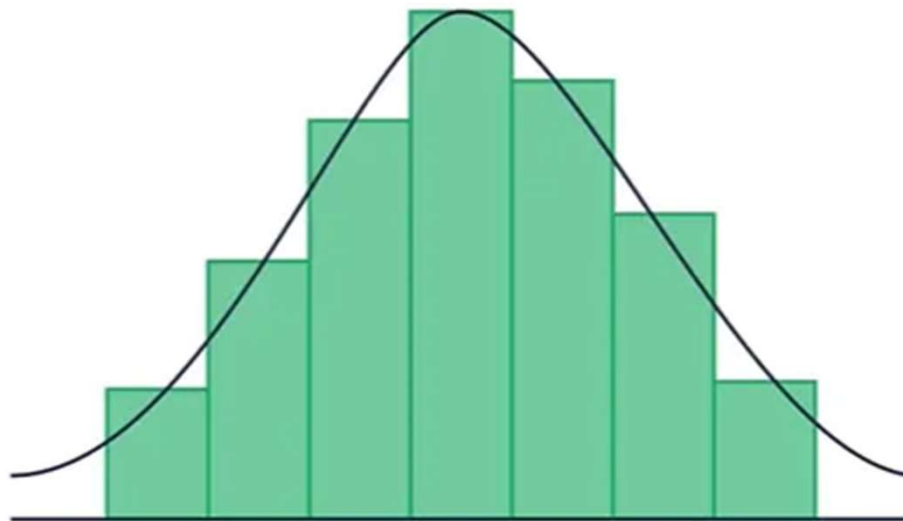


4. Visualizações de distribuição

- **Quando usar histogramas?**
 - Eles mostram efetivamente a forma geral e os clusters na distribuição de um conjunto de dados.
 - Os histogramas são ótimos se você deseja comunicar a distribuição dos dados de forma rápida e fácil a outras pessoas.

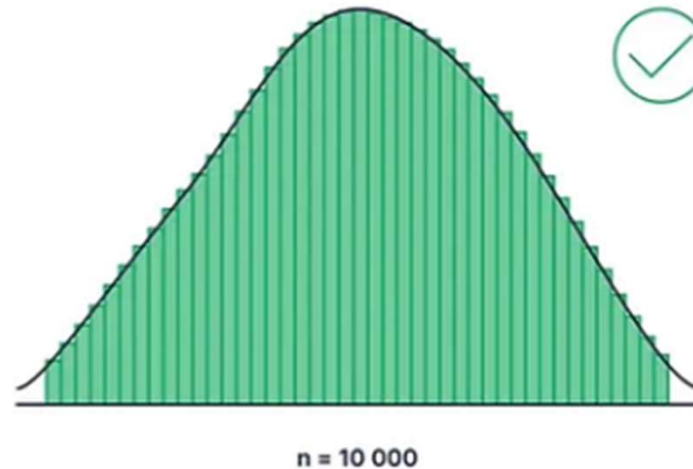
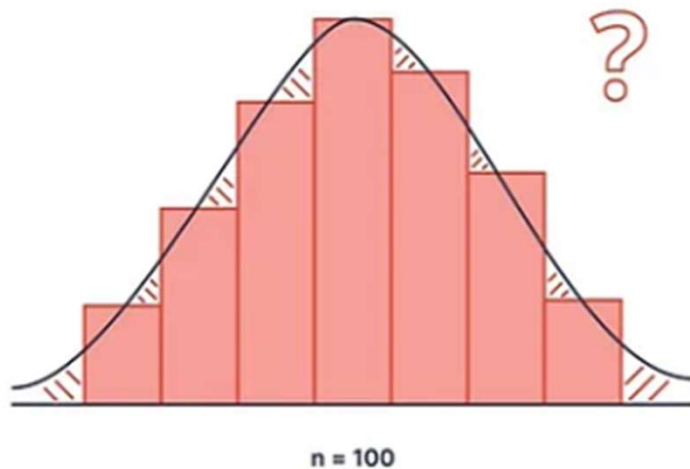
4. Visualizações de distribuição

- Quando usar gráficos de densidade?
- Os gráficos de densidade usam a estimativa de densidade do kernel (KDE) para suavizar as caixas do histograma em curvas, revelando padrões mais sutis.



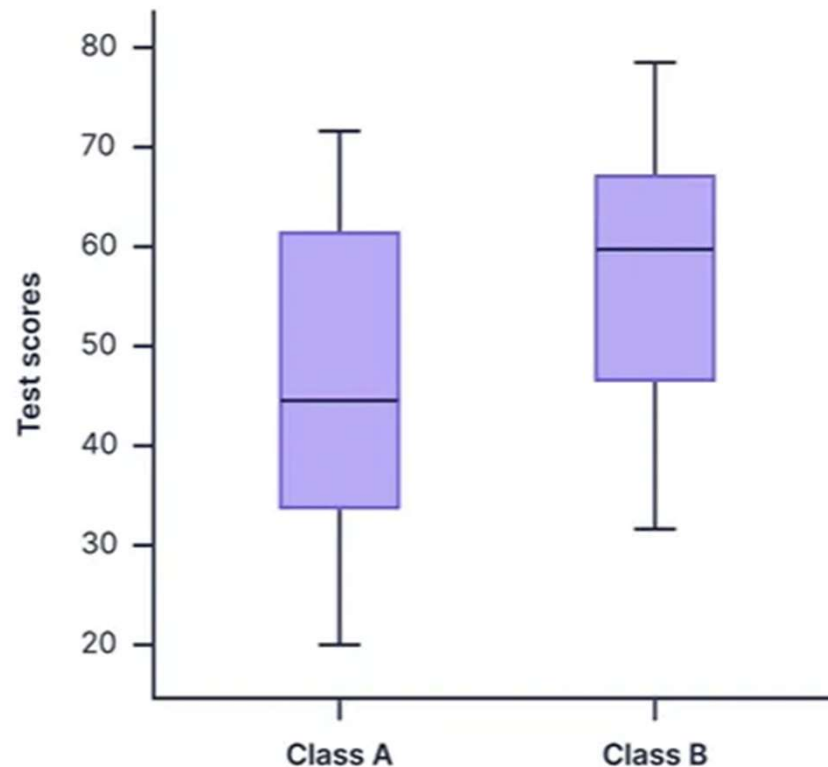
4. Visualizações de distribuição

- Quando usar gráficos de densidade?
- As curvas de densidade são muito úteis ao trabalhar com amostras grandes.
- Em amostras pequenas, as curvas de densidade podem ser imprecisas devido à falta de lacunas.



4. Visualizações de distribuição

- Quando usar gráficos de caixa?
- Os gráficos de caixa são usados para mostrar distribuições de valores de dados numéricos, especialmente quando você deseja compará-los entre vários grupos.

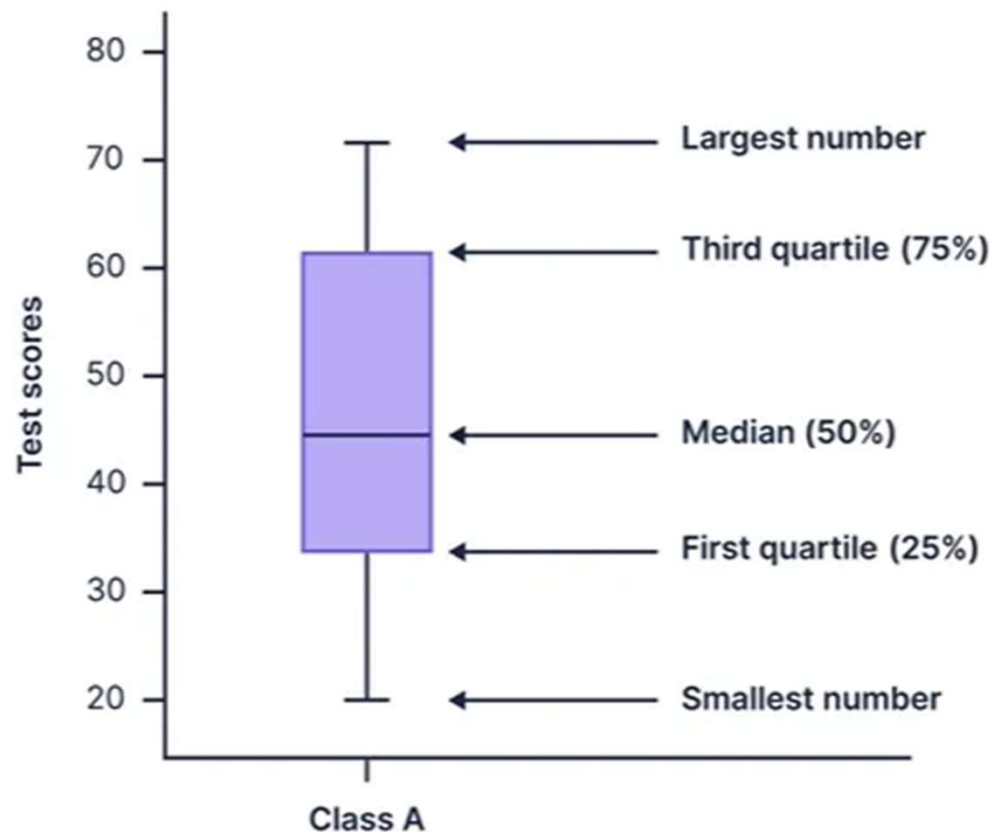


4. Visualizações de distribuição

- Em sua essência, o box plot retrata aspectos do “**resumo de cinco números**” – um conjunto de valores que descrevem a tendência central e a dispersão de um conjunto de dados.

- O resumo dos cinco números:

- Valor mínimo
- Primeiro quartil (Q1)
- Mediana
- Terceiro quartil (Q3)
- Valor máximo



Visualizações de dados

- A prática recomendada é explorar seus dados e considerar seus objetivos antes de escolher. Pergunte a si mesmo quais perguntas você mais deseja que sejam respondidas com base nos dados. Em seguida, selecione um tipo de visualização que corresponda ao insight necessário.
- Com a prática, você aprenderá a combinar intuitivamente a visualização correta com seus dados específicos e necessidades analíticas. A chave é encontrar representações visuais que tornem os padrões e insights fáceis de ver à primeira vista.



INSTITUTO FEDERAL
Sudeste de Minas Gerais