ANÁLISE DE DADOS
TÉCNICAS ESTATÍSTICAS, MINERAÇÃO DE DADOS E VISUALIZAÇÃO

Versão 1.0 · Abril/2026

Autor
Prof. Dr. Rodrigo Almeida Tavares
Doutor em Estatística Aplicada
Instituto Nacional de Ciências Quantitativas

Revisor
Prof. Dr. Eduardo Sampaio Vieira
Doutor em Ciência de Dados
Instituto Nacional de Ciências Quantitativas

Publicação
Instituto Nacional de Ciências Quantitativas
Rua Dr. Sampaio Vieira, 1500 – Sala 12
Vila Nova – Cidade Alta – CEP 60.000-000
Fortaleza – Ceará – Brasil
www.inciquant.com.br

Contato
Prof. Dr. Rodrigo Almeida Tavares
rodrigo.tavares@inciquant.com.br
+55 (85) 3234-5678

Versão
1.0 – Abril/2026

SUMÁRIO EXECUTIVO

Análise de Dados: Técnicas Estatísticas, Mineração de Dados e Visualização é um manual técnico-científico que apresenta, de forma sistemática e rigorosa, os fundamentos teóricos e as técnicas práticas da análise de dados, da estatística descritiva à mineração de dados e à visualização de informações. Ao longo de 12 capítulos, o leitor percorre o caminho completo que separa um conjunto de dados brutos de uma decisão fundamentada: organização e limpeza dos dados, medidas de tendência central e de dispersão, distribuições de probabilidade, correlação e regressão linear, testes de hipótese, intervalos de confiança, análise de variância, regressão múltipla, classificação e agrupamento, e princípios de visualização de dados.

O livro foi escrito para profissionais que precisam tomar decisões com dados — analistas, gestores, pesquisadores e estudantes de graduação e pós-graduação — e pressupõe apenas o domínio de álgebra e noções básicas de cálculo. Cada capítulo combina exposição conceitual, exemplos numerados e exercícios propostos, de modo que o leitor possa acompanhar a teoria e verificar, por conta própria, o que aprendeu.

O que o leitor vai encontrar neste livro

- Os fundamentos da estatística descritiva e da inferência, com ênfase no que cada medida e cada teste realmente significam.
- As técnicas de mineração de dados mais usadas em contextos reais, com critérios para escolher entre elas.
- Princípios de visualização de dados que ajudam a comunicar resultados sem distorcer a informação.
- Exemplos numerados e exercícios propostos ao final de cada capítulo.

Como este livro foi produzido

Este livro foi produzido com o apoio de ferramentas de inteligência artificial, sob a direção e a revisão dos autores. A IA foi usada como auxiliar na redação e na organização do texto; os autores são responsáveis pelo conteúdo, pelas fórmulas e pelas conclusões apresentadas. Os exemplos numéricos foram conferidos manualmente e os exercícios foram revisados por pares.

Como citar este livro

Tavares, R. A. (2026). *Análise de Dados: Técnicas Estatísticas, Mineração de Dados e Visualização*. Versão 1.0. Fortaleza: Instituto Nacional de Ciências Quantitativas.

SUMÁRIO

Capítulo 1 – O que é análise de dados
Capítulo 2 – Tipos de dados e escalas de medição
Capítulo 3 – Organização e limpeza de dados
Capítulo 4 – Medidas de tendência central
Capítulo 5 – Medidas de dispersão e de posição
Capítulo 6 – Distribuições de probabilidade: o que o dado esperado diz sobre o que vai acontecer
Capítulo 7 – Correlação e regressão linear: quando duas variáveis andam juntas
Capítulo 8 – Testes de hipótese: como decidir com dados
Capítulo 9 – Intervalos de confiança: quanto confiar na estimativa
Capítulo 10 – Análise de variância: comparando mais de dois grupos
Capítulo 11 – Regressão múltipla: quando mais de uma variável explica o resultado
Capítulo 12 – Classificação e agrupamento: como organizar dados sem saber o gabarito
Apêndice A – Tabela de distribuição normal padrão
Apêndice B – Tabela de distribuição t de Student
Apêndice C – Tabela de distribuição qui-quadrado
Apêndice D – Tabela de distribuição F de Fisher-Snedecor
Apêndice E – Tabela de números aleatórios
Apêndice F – Glossário de termos e símbolos
Apêndice G – Fórmulas de referência rápida
Apêndice H – Sugestões de leitura e bibliografia comentada

PREFÁCIO

Este livro nasceu de uma pergunta simples: por que tantos profissionais que precisam tomar decisões com dados não conseguem fazer isso com segurança? A resposta, quase sempre, não está na falta de ferramentas — está na falta de um caminho claro entre o dado bruto e a decisão fundamentada. Este livro foi escrito para ocupar esse caminho.

Ao longo de 12 capítulos, o leitor percorre o caminho completo que separa um conjunto de dados brutos de uma decisão fundamentada: organização e limpeza dos dados, medidas de tendência central e de dispersão, distribuições de probabilidade, correlação e regressão linear, testes de hipótese, intervalos de confiança, análise de variância, regressão múltipla, classificação e agrupamento, e princípios de visualização de dados. Cada capítulo combina exposição conceitual, exemplos numerados e exercícios propostos, de modo que o leitor possa acompanhar a teoria e verificar, por conta própria, o que aprendeu.

O livro foi escrito para profissionais que precisam tomar decisões com dados — analistas, gestores, pesquisadores e estudantes de graduação e pós-graduação — e pressupõe apenas o domínio de álgebra e noções básicas de cálculo. Quando um conceito exige mais do que isso, o texto avisa e oferece o caminho para quem quer ir mais longe.

Este livro foi produzido com o apoio de ferramentas de inteligência artificial, sob a direção e a revisão dos autores. A IA foi usada como auxiliar na redação e na organização do texto; os autores são responsáveis pelo conteúdo, pelas fórmulas e pelas conclusões apresentadas. Os exemplos numéricos foram conferidos manualmente e os exercícios foram revisados por pares.

Fortaleza, abril de 2026.

Prof. Dr. Rodrigo Almeida Tavares
Prof. Dr. Eduardo Sampaio Vieira

CAPÍTULO 1 – O QUE É ANÁLISE DE DADOS

Toda decisão que alguém toma hoje pode ser apoiada por dados — e, quase sempre, é melhor que ela seja. Mas dados brutos não decidem nada sozinhos. Entre o número que chega à tela e a decisão que muda o rumo de uma empresa, de um hospital ou de uma escola existe um caminho inteiro a ser percorrido. Este capítulo mapeia esse caminho.

1.1 Dados, informação e decisão

Um dado é um registro isolado: o preço de uma passagem, a temperatura de um dia, a idade de um paciente. Informação começa quando alguém organiza esses registros de modo que eles respondam a uma pergunta. Decisão começa quando alguém usa essa informação para escolher entre duas ou mais alternativas.

A análise de dados é o caminho que liga os três. Ela pega os dados brutos, organiza-os, resume-os, testa hipóteses e devolve ao tomador de decisão uma informação que ele consegue usar. Não substitui o julgamento — alimenta o julgamento.

1.2 Estatística descritiva e inferência

Existem dois grandes caminhos dentro da análise de dados. A estatística descritiva resume o que se observou: média, mediana, desvio-padrão, gráfico de barras, tabela de frequências. A inferência usa o que se observou para falar do que não se observou: estimar um parâmetro populacional, testar uma hipótese, construir um intervalo de confiança.

Os dois caminhos são complementares. Quase todo estudo sério começa pela descritiva e só depois avança para a inferência.

1.3 Mineração de dados e visualização

Mineração de dados é o conjunto de técnicas para encontrar padrões em bases grandes demais para serem analisadas à mão: regras de associação, agrupamentos, classificações, modelos preditivos. Visualização de dados é a arte de transformar tabelas em gráficos que comunicam sem distorcer.

1.4 Quando a análise de dados ajuda e quando não ajuda

A análise de dados ajuda quando há dados suficientes e relevantes para a pergunta feita, quando o tomador de decisão consegue agir sobre o resultado e quando o custo da análise é menor que o custo da decisão errada. Não ajuda quando a pergunta não está bem definida, quando os dados são tendenciosos demais para qualquer correção, quando a decisão já foi tomada por outro caminho e quando o resultado não pode ser medido.

1.5 Exercícios propostos

1. Cite três decisões do seu dia a dia que poderiam ser apoiadas por dados. Para cada uma, diga que dado bastaria e que pergunta ele responderia.
2. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
3. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?
4. Dê um exemplo em que a análise de dados ajuda e um em que ela não ajuda, e justifique cada um.

CAPÍTULO 2 – TIPOS DE DADOS E ESCALAS DE MEDIÇÃO

Antes de calcular qualquer coisa, é preciso saber o que se está medindo. Este capítulo apresenta os tipos de dados e as escalas de medição que ditam, em grande parte, quais técnicas podem ser usadas e quais não podem.

2.1 Variáveis quantitativas e qualitativas

Variáveis quantitativas são aquelas que assumem valores numéricos e podem ser somadas, comparadas e medidas. Variáveis qualitativas são aquelas que descrevem categorias, sem ordem ou com ordem apenas parcial.

2.2 Escala nominal

Na escala nominal, os valores são apenas rótulos. Não há ordem nem distância entre eles. Exemplo: cor do olho, sexo, cidade de nascimento.

2.3 Escala ordinal

Na escala ordinal, há ordem entre os valores, mas a distância entre eles não é medida. Exemplo: grau de satisfação (ruim, regular, bom, ótimo), posição em uma competição.

2.4 Escala intervalar

Na escala intervalar, há ordem e a distância entre os valores é medida, mas o zero é arbitrário. Exemplo: temperatura em graus Celsius, ano de nascimento.

2.5 Escala razão

Na escala razão, há ordem, distância e o zero é real. Exemplo: idade, peso, altura, renda.

2.6 Exercícios propostos

1. Classifique cada variável abaixo em nominal, ordinal, intervalar ou razão: cor do olho, sexo, cidade de nascimento, grau de satisfação, temperatura em graus Celsius, ano de nascimento, idade, peso, altura, renda.
2. Por que a média de temperatura em graus Celsius faz sentido, mas a média de ano de nascimento não faz?
3. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
4. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.

CAPÍTULO 3 – ORGANIZAÇÃO E LIMPEZA DE DADOS

Antes de qualquer análise, os dados precisam estar organizados e limpos. Este capítulo apresenta o caminho que separa um conjunto de dados brutos de um conjunto de dados pronto para ser analisado.

3.1 Tabelas de frequências

Uma tabela de frequências organiza os dados de uma variável qualitativa ou quantitativa discretizada, mostrando quantas vezes cada valor aparece.

3.2 Dados faltantes

Dados faltantes são valores que deveriam estar na base e não estão. Eles podem surgir por erro de coleta, por recusa do respondente, por falha de sistema ou por simples esquecimento.

3.3 Outliers

Outliers são valores que se afastam muito do restante da amostra. Nem todo outlier é erro — alguns são o dado mais interessante da base.

3.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 4 – MEDIDAS DE TENDÊNCIA CENTRAL

Medidas de tendência central são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

4.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

4.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

4.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

4.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 5 – MEDIDAS DE DISPERSÃO E DE POSIÇÃO

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

5.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

5.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

5.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

5.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 6 – DISTRIBUIÇÕES DE PROBABILIDADE: O QUE O DADO ESPERADO DIZ SOBRE O QUE VAI ACONTECER

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

6.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

6.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

6.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

6.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 7 – CORRELAÇÃO E REGRESSÃO LINEAR: QUANDO DUAS VARIÁVEIS ANDAM JUNTAS

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

7.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

7.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

7.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

7.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 8 – TESTES DE HIPÓTESE: COMO DECIDIR COM DADOS

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

8.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

8.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

8.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

8.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 9 – INTERVALOS DE CONFIANÇA: QUANTO CONFIAR NA ESTIMATIVA

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

9.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

9.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

9.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

9.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 10 – ANÁLISE DE VARIÂNCIA: COMPARANDO MAIS DE DOIS GRUPOS

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

10.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

10.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

10.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

10.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 11 – REGRESSÃO MÚLTIPLA: QUANDO MAIS DE UMA VARIÁVEL EXPLICA O RESULTADO

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

11.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

11.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

11.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

11.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

CAPÍTULO 12 – CLASSIFICAÇÃO E AGRUPAMENTO: COMO ORGANIZAR DADOS SEM SABER O GABARITO

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

12.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

12.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

12.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

12.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE A – TABELA DE DISTRIBUIÇÃO NORMAL PADRÃO

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

A.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

A.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

A.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

A.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE B – TABELA DE DISTRIBUIÇÃO T DE STUDENT

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

B.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

B.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

B.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

B.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE C – TABELA DE DISTRIBUIÇÃO QUI-QUADRADO

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

C.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

C.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

C.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

C.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE D – TABELA DE DISTRIBUIÇÃO F DE FISHER-SNEDECOR

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

D.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

D.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

D.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

D.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE E – TABELA DE NÚMEROS ALEATÓRIOS

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

E.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

E.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

E.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

E.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE F – GLOSSÁRIO DE TERMOS E SÍMBOLOS

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

F.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

F.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

F.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

F.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE G – FÓRMULAS DE REFERÊNCIA RÁPIDA

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

G.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

G.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

G.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

G.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?

APÊNDICE H – SUGESTÕES DE LEITURA E BIBLIOGRAFIA COMENTADA

Medidas de dispersão e de posição são valores que tentam representar o centro de um conjunto de dados. Este capítulo apresenta a média, a mediana e a moda, e mostra quando cada uma delas é a mais adequada.

H.1 Média aritmética

A média aritmética é a soma dos valores dividida pela quantidade de valores. É a medida mais usada e a mais sensível a outliers.

H.2 Mediana

A mediana é o valor que divide a amostra ordenada em duas metades iguais. Ela não é afetada por outliers e é a medida mais robusta para dados assimétricos.

H.3 Moda

A moda é o valor mais frequente. Ela é a única medida de tendência central que faz sentido para variáveis qualitativas nominais.

H.4 Exercícios propostos

1. Um questionário pergunta "Você está satisfeito?" com as opções sim, não, talvez. Que escala essa variável pode ser? Justifique.
2. Dê um exemplo de variável qualitativa ordinal que não seja grau de satisfação.
3. Um hospital registra a temperatura de seus pacientes todos os dias. Que pergunta descritiva e que pergunta inferencial podem ser feitas com esses dados?
4. Por que a média de temperatura de um paciente isolado não é informação, mas a média de um grupo pode ser?