MBA em Ciência de Dados
ICMC/USP - São Carlos

Estatística para Iniciantes
Aula 01 – Teoria

Francisco Louzada Neto
ICMC/USP
louzada@icmc.usp.br

Sobre do Curso

Objetivo Geral

O propósito central desse curso é expor os fundamentos introdutórios do **pensamento estatístico** e a importância que esse tipo de abordagem tem sobre o entendimento da nossa realidade.

Uma questão fundamental que move o pensamento estatístico é o entendimento de que a natureza se expressa de maneira aleatória, de modo que os modelos baseados nessa perspectiva se caracterizam pela consideração da **presença de padrões de aleatoriedade** cuja teoria é descrita pela matemática e a aplicação comumente é feita computacionalmente.

Sobre do Curso

Objetivos Específicos

Explicitaremos essas questões apresentando a seguinte linha de raciocínio:

- **Conceitos Estatísticos básicos (partes I e II)**
  Onde enfatizaremos os métodos descritivos fundamentais, que nos permitem expressar textualmente a informação observada em um conjunto de dados.
- **Técnica da Segmentação de Dados**
  Onde sistematizaremos o poder de exploração de um conjunto de dados, considerando tanto a perspectiva (e os maiores interesses) da área de estudo quanto o uso de métodos quantitativos automatizados.
- **Noções sobre Estimação para Comparações Estatísticas**
  Onde veremos como a incerteza caracteriza o modo como a estatística faz comparações.
- **Introdução à Regressão Linear Simples**
  Onde apresentaremos o ponto de vista estatístico sobre essa ferramenta que também é muito utilizada na matemática e na computação.

ESTRUTURAS DESCRITIVAS (PARTE I)

Resumo do Tópico

Assunto Abordado

Nestes dois primeiros tópicos, voltaremos nossa atenção às **Estruturas Estatísticas de Sumarização**.

Daremos uma introdução à forma de como a **Estatística** nos permite sintetizar massas de informação em **quantidades** ou **representações-chave**, no intuito de expressar características fundamentais e interpretáveis por trás de todo o conjunto de registros disponíveis.

Objetivos

Ao final deste tópico, o aluno será capaz de fazer uma **análise exploratória sistemática** sobre um conjunto de informações com características quantitativas e qualitativas de uma amostra ou população de interesse.

Ideias Básicas, Motivações e Aplicações

Se existe uma ciência que, desde o seu nascimento, se dedicou a sistematizar todas as etapas do processo de análise de dados, desde a *coleta* até a *interpretação*, essa ciência é a **Estatística**!

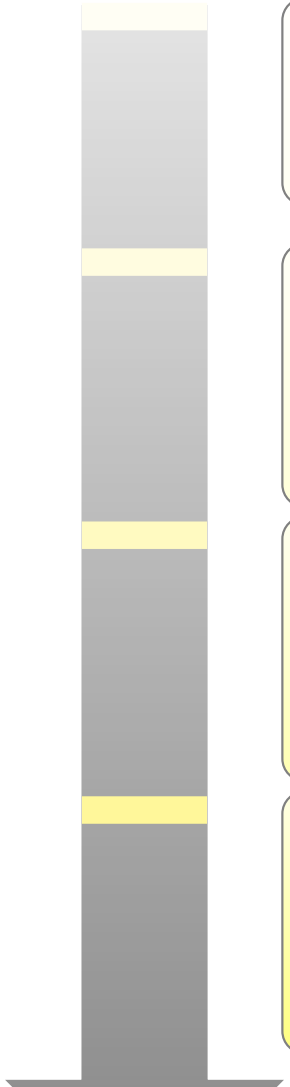
A **Estatística** tem se destacado nesses tempos em que a **Ciência de Dados** se mostra cada vez mais relevante.

O objetivo principal da **Estatística** é desenvolver metodologias que são matemática e computacionalmente embasadas para a coleta, organização, descrição, análise e interpretação de dados, levando em consideração a incerteza que governa a natureza.

NOTA

Por questões didáticas, a **Estatística** pode ser dividida em dois grandes ramos, **Descritiva** e **Indutiva** (ou **Inferencial**).

Estatística Descritiva



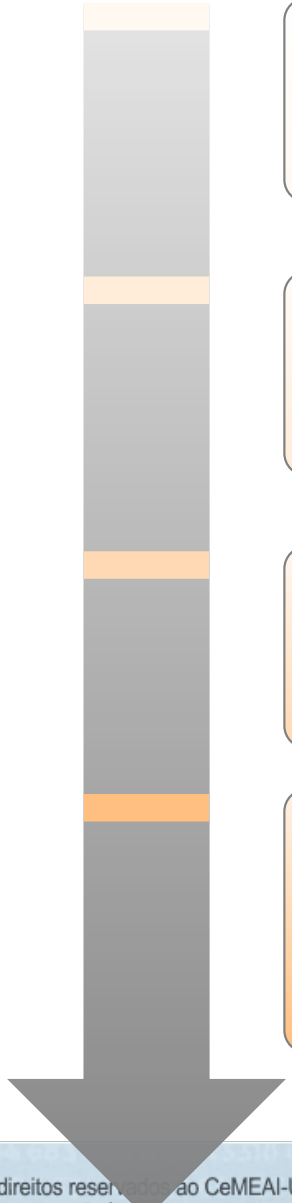
É a primeira etapa de uma boa análise.

A Estatística Descritiva é uma coleção de técnicas cujo propósito é **descrever e resumir** as informações contidas em um conjunto de dados.

Os resumos, em geral, correspondem à obtenção de **quantidades-chave** ou **representações-chave**.

Os resumos promovem interpretações práticas que permitem o entendimento do comportamento geral de uma massa de dados.

Estatística Indutiva (ou Inferência Estatística)



Etapa **posterior** à descritiva. Utilizada para **confirmar ou contrariar percepções** levantadas.

É uma coleção de técnicas que, **a partir de dados amostrais**, conclui algo da **população**.

Tem como pré-requisito a **validação** da técnica proposta para extrapolar os resultados obtidos.

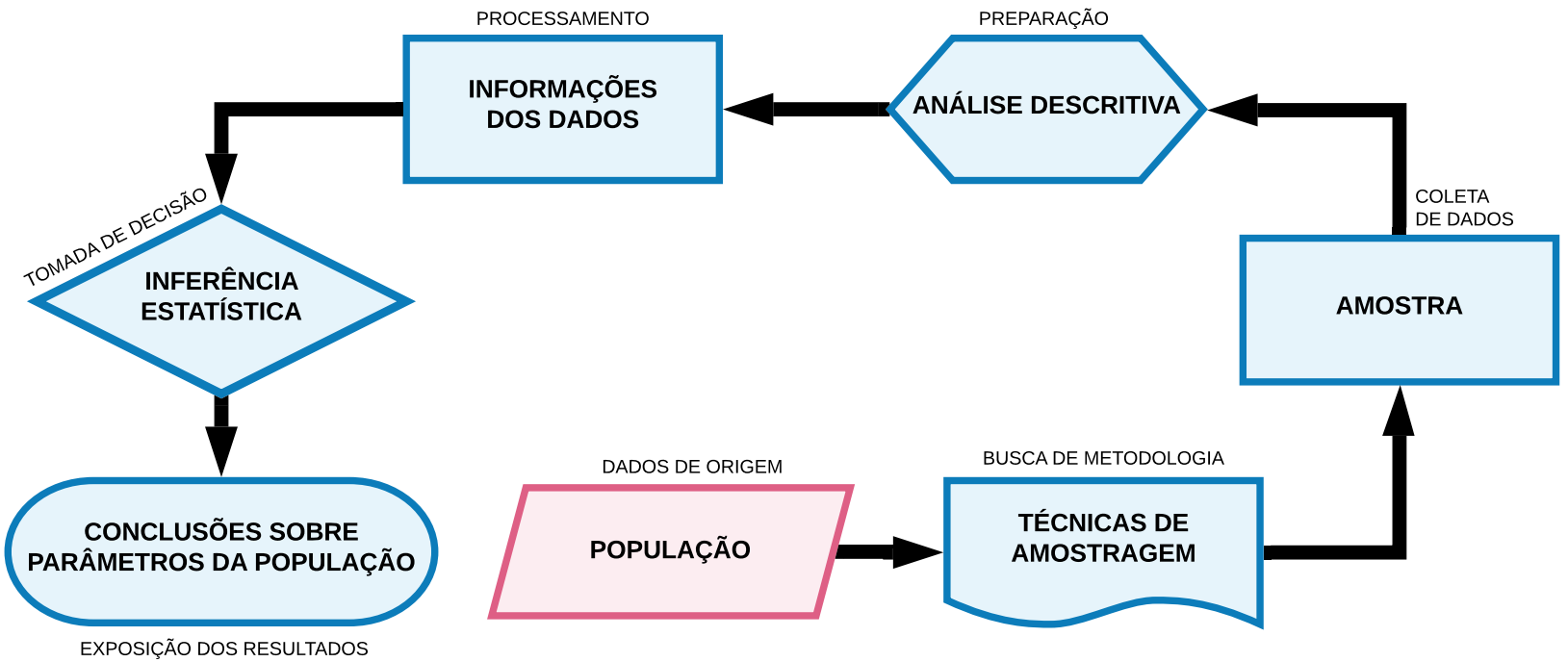
Infere sobre a população **quantificando e entendendo o padrão aleatório dos erros envolvidos** no processo de estimação.

Fluxo de Modelagem

Os métodos estatísticos se aplicam em uma infinidade de contextos.

Qualquer área que lide com a coleta, registro e análise de informações pode utilizar de metodologias estatísticas.

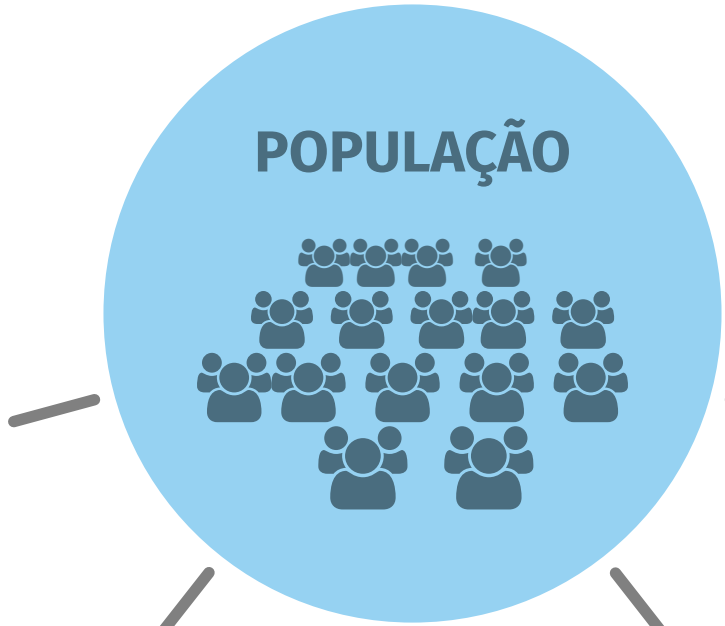
FLUXO



Conceitos Fundamentais

População

Denota o conjunto de todos os possíveis valores de uma característica observável, associada a uma coleção de indivíduos, animais ou objetos.



Todas as rendas familiares de todos os clientes de uma agência seguradora.

Todas as ocorrências de sinistro nos anos de 2000 a 2019.

Todas as franquias de seguros fechadas no ano de 2019.

Todos os lucros trimestrais de uma seguradora sobre a venda de seguros de automóveis.

Conceitos Fundamentais

Censo

Quando, em um estudo, o interesse reside em avaliar todos (ou *quase* a totalidade) os elementos pertencentes à população. É um recorte exato da população em um momento específico no tempo.

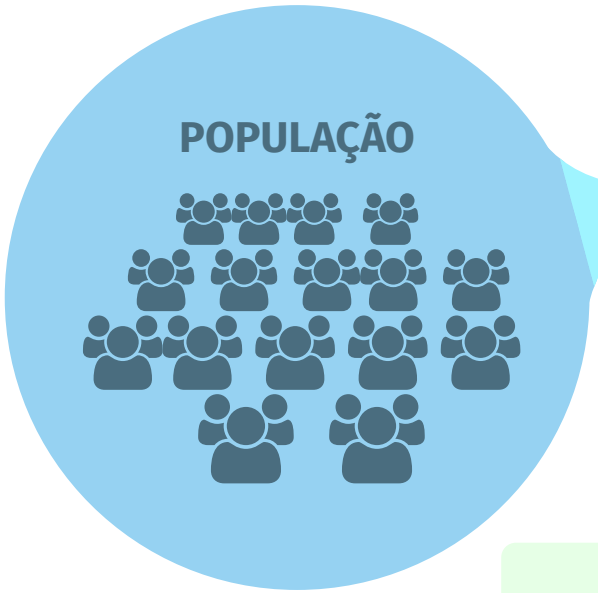
EXEMPLOS

- Censo Demográfico - IBGE;
- Determinação do perfil socio-demográfico dos profissionais de uma empresa específica, avaliando-se a totalidade dos funcionários.

Conceitos Fundamentais

Amostra

Qualquer subconjunto finito de elementos extraídos da população.



Rendas familiares de **alguns** dos clientes de uma agência seguradora.

Algumas ocorrências de sinistro nos anos de 2000 a 2019.

Uma fração representativa das franquias de seguros fechadas no ano de 2019.

Parte dos registros dos lucros trimestrais de uma seguradora sobre a venda de seguros de automóveis.

Conceitos Fundamentais

Variável

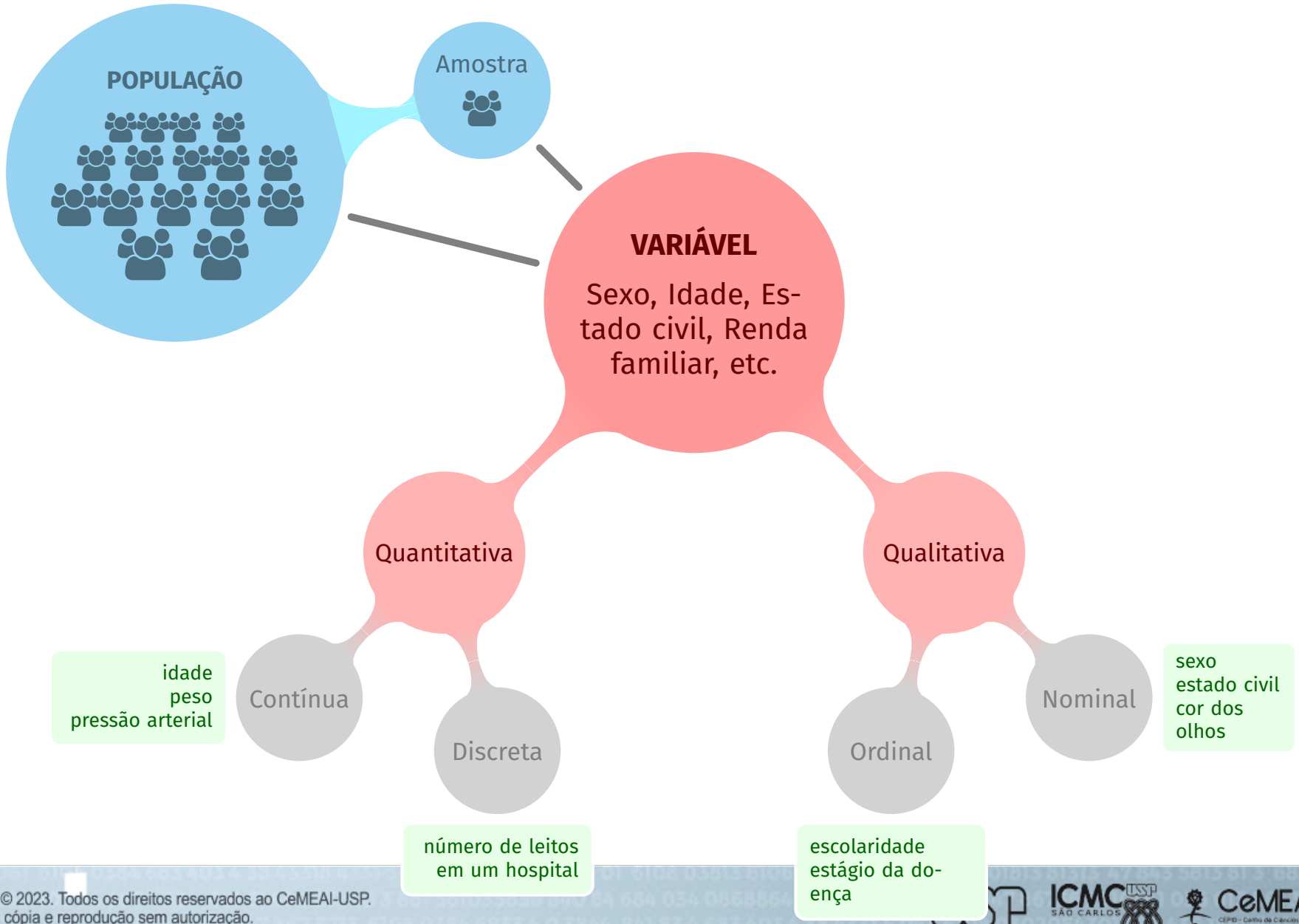
Uma variável (*feature*) representa qualquer característica ou atributo de interesse associado a uma população.
É sobre essas características que a estatística vai trabalhar.

EXEMPLOS

Qualquer característica sócio-demográfica do cliente de uma agência seguradora, como:

- Sexo;
- Idade;
- Estado civil;
- Renda familiar;
- etc.

Conceitos Fundamentais



Exemplos que serão abordados

Durante esse desenvolvimento, vamos utilizar, sempre que for necessário, dois possíveis exemplos para ilustrar o que estudaremos.

Ex 01

O primeiro deles, considera um conjunto de dados de pequeno porte e, portanto, de mais simples tratamento. Trata-se de um conjunto com de medidas das alturas de 11 pessoas.

Ex 02

O segundo deles, se refere aos valores pagos por 3142 pessoas, por seguros de automóveis. Maiores detalhes serão expostos a seguir.

Ex 01 (Alturas): O Conjunto de Dados

Considere o conjunto de dados com 11 alturas, cujo os valores são apresentados a seguir:

1.85 1.85 1.76 1.78 1.80 1.71
1.73 1.76 1.80 1.83 1.80

Ex 02 (Seguros): O Conjunto de Dados

Considere alguns registros sobre clientes de uma determinada seguradora. Os dados podem ser acessados aqui. O conjunto de dados conta com 3142 linhas, uma para cada contrato fechado pela seguradora. Para cada contrato estão disponíveis as seguintes informações:

- **PUBLICO**: O sexo do cliente;
- **VEICULO**: O modelo do veículo segurado (duas categorias);
- **ESTADO**: O estado em que o seguro foi vendido (SP e RJ);
- **MES**: O mês em que o seguro foi vendido;
- **VALOR_FIPE**: O valor do veículo na tabela FIPE;
- **VALOR_SEGURO_VENDIDO**: O valor pelo qual o seguro foi vendido para aquele cliente.

PUBLICO	VEICULO	ESTADO	MES	VALOR_FIPE	VALOR_SEGURO_VENDIDO
Feminino	NOVO KA S 1.0 TICVT FLEX 4P (2020/2020)	SP	1	44244	1298.67
Feminino	NOVO KA S 1.0 TICVT FLEX 4P (2020/2020)	SP	1	44244	1211.33
Feminino	NOVO KA S 1.0 TICVT FLEX 4P (2020/2020)	SP	1	44244	1038.20
⋮	⋮	⋮	⋮	⋮	⋮

Medidas de Posição

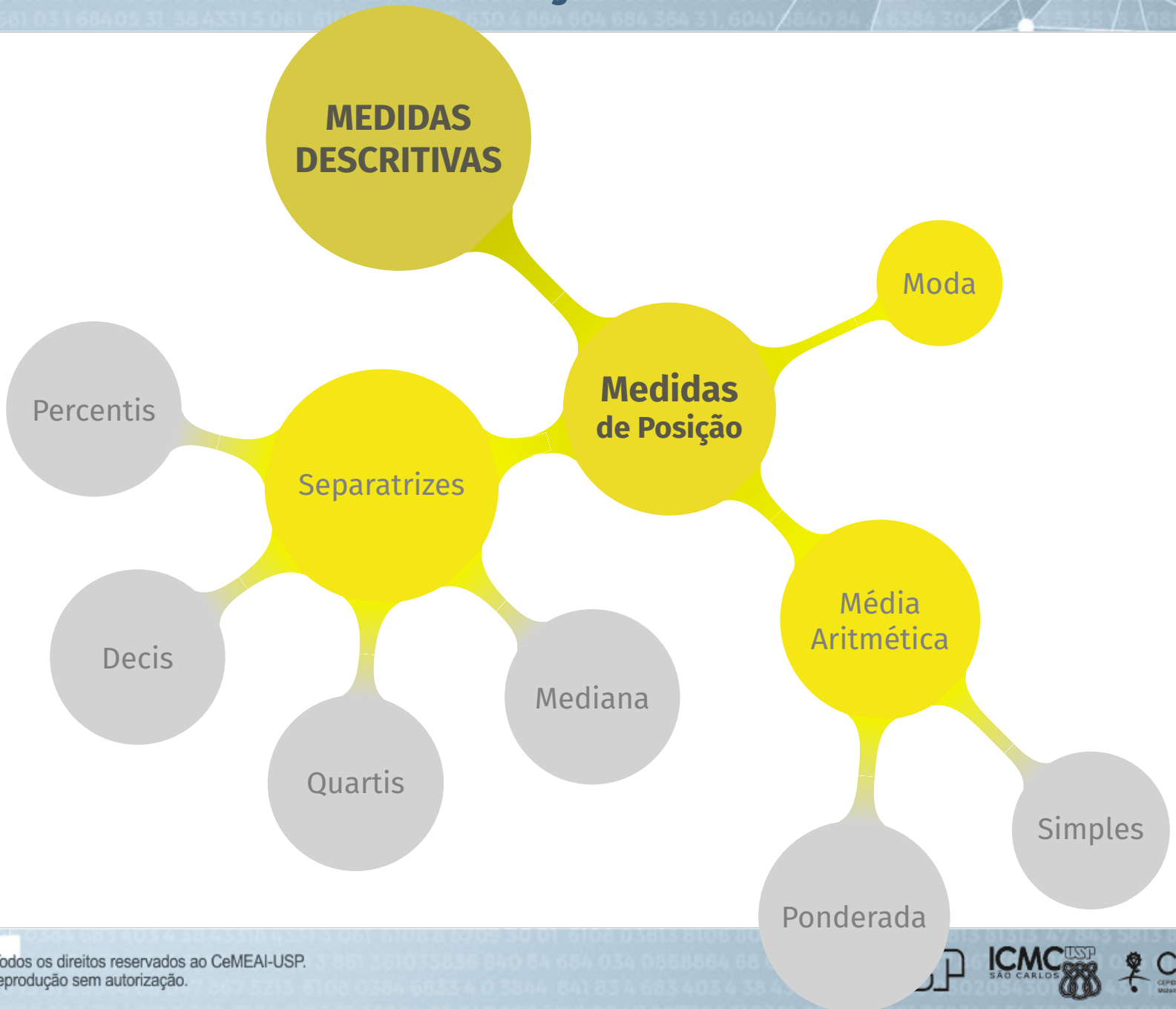
As **medidas de posição** (ou **medidas de tendência central**) resumem informações ao retornarem um valor sobre o qual todos os outros se distribuem com maior frequência.

São pontos de concentração.

As **medidas de posição** nos permitem ter uma **primeira percepção** dos dados sobre qualquer quantidade de registros.

Deixamos de tatear, por exemplo, uma grande tabela numérica às cegas, e entendemos sobre qual valor os registros se distribuem com maior ou menor frequência.

Estruturas de Sumarização



Média Aritmética Simples

A **média aritmética simples** (ou apenas **média**) é a medida de posição mais utilizada.
Possui fácil interpretação e não distingue os valores entre si.
Todas as observações são igualmente importantes para o cálculo dessa média.

Definição

Sejam $x_1, x_2, \ldots, x_n$ valores observados de uma variável X. A **média aritmética simples** associada a esses valores é calculada como

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Ao trabalharmos com todos os N elementos da população, geralmente denotamos a média por μ (indicando a **média populacional**) e a calculamos por:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}.$$

Dificuldade

A média aritmética simples é **sensível a valores extremos**. Isso faz com que a ela deixe de ter a representatividade que se espera dela.

Indicação

Utilizá-la em conjunto com outras medidas mais robustas.

Ex 01 (Alturas): Média Aritmética Simples

Considerando o conjunto de dados com tamanho $n = 11$***. Os valores das alturas são dados por:***

1.85 1.85 1.76 1.78 1.80 1.71
1.73 1.76 1.80 1.83 1.80

$$\bar{X} = \sum_{i=1}^{11} \frac{x_i}{11}$$

$$= \frac{x_1 + x_2 + \ldots + x_{11}}{11}$$

$$= \frac{1.85 + 1.85 + 1.76 + 1.78 + 1.80 + 1.71 + 1.73 + 1.76 + 1.80 + 1.83 + 1.80}{11}$$

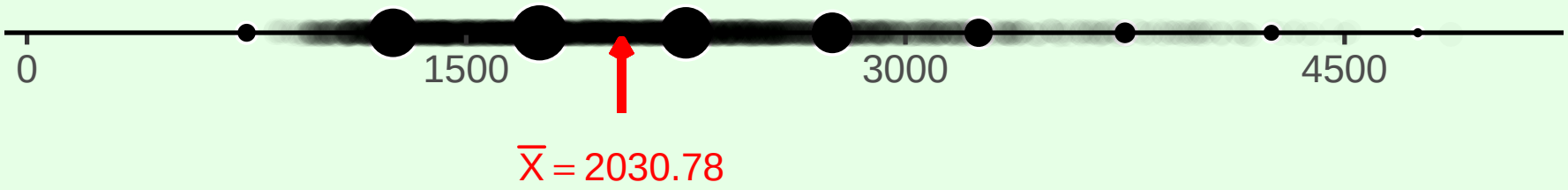
$$= 1.79$$

Ex 02 (Seguros): Média Aritmética Simples

Questão: Independente de qualquer detalhamento mais específico, qual é a média do valor cobrado pela seguradora, representado pela variável VALOR_SEGURO_VENDIDO, durante o primeiro semestre para os dois modelos, estados e públicos observados?

O valor médio do seguro vendido durante o primeiro semestre de 2020, para os dois tipos de veículos e estados observados é igual a R$ 2030.78.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{847{,}83 + \cdots + 4862{,}63}{3142} = 2030{,}78$$



Média Aritmética Ponderada

A **média aritmética ponderada** (ou apenas **média ponderada**) é uma medida de posição que faz distinção entre os valores, por considerar que alguns deles se expressam com maior frequência, ou tem maior peso sobre as demais observações. Além dos valores observados, essa média leva em consideração um conjunto de pesos associados.

Definição

Sejam $x_1, x_2, \ldots, x_n$ valores observados de uma variável X, e $w_1, w_2, \ldots, w_n$ os seus respectivos pesos. A **média aritmética ponderada** é calculada como:

$$\overline{X}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i}.$$

Ex 02 (Seguros): Média Aritmética Ponderada

Questão: Suponha que, para o segundo semestre de 2020, planeja-se que os valores dos seguros passem por um acréscimo de 15% para o público feminino e 25% para o público masculino, sendo mantidas as demais configurações de precificação. Nesse contexto, com base somente nos registros históricos disponibilizados, qual seria o valor médio do seguro vendido por essa seguradora, nessa nova configuração?

Novamente, independente de qualquer detalhamento, e considerando as ponderações definidas, qual é a média ponderada do valor do seguro vendido?

Ex 02 (Seguros): Média Aritmética Ponderada

O valor médio do seguro vendido, considerando as novas ponderações, durante o primeiro semestre de 2020, para os dois veículos e estados observados é igual a R$ 2040.65.

PUBLICO	PESO	VALOR_SEGURO_VENDIDO
Feminino	1.15	1298.67
Feminino	1.15	1211.33
⋮	⋮	⋮
Masculino	1.25	4515.94
Masculino	1.25	4862.63

$$\overline{X}_w = \frac{\sum_{i=1}^{n} x_i w_i}{\sum_{i=1}^{n} w_i} = \frac{1{,}15 \times 847{,}83 + \cdots + 1{,}25 \times 4862{,}63}{1{,}15 + \cdots + 1{,}25} = 2040{,}65$$

Moda

A **moda** é o valor (ou atributo) que ocorre com maior frequência em um conjunto de dados. Ela representa o valor que mais provavelmente será observado em um experimento.

Definição

Sejam $x_1, x_2, \ldots, x_n$ valores observados de uma variável X, a **moda** (ou **valor modal**) é dado por:

$$\text{Mod} = X_{\text{freq}},$$

em que X_{freq} é o valor mais frequente observado em $x_1, x_2, \ldots, x_n$.

OBSERVAÇÃO

Ex 01 (Alturas): Moda

Voltando ao conjunto de dados com tamanho $n = 11$. ***Os valores das alturas são dados por:***

1.85 1.85 1.76 1.78 1.80 1.71
1.73 1.76 1.80 1.83 1.80

Nesse caso, quando ordenamos e contabilizamos as repetições, identificamos um conjunto de dados unimodal.

				1.80		
		1.76		1.80		1.85
1.71	1.73	1.76	1.78	1.80	1.83	1.85

Simplificação do Ex 2 (Seguros): Moda

Considere um subconjunto de 11 valores dos seguros vendidos, utilizados nos exemplos anteriores:

1408.50 1630.07 1771.85 977.94 1139.64
1137.29 1408.50 2494.12 3617.60 3006.26 1534.65

Ordenando os valores e contabilizando suas repetições, identificamos o seguinte comportamento:

			1408.5						
977.94	1137.29	1139.64	1408.5	1534.65	1630.07	1771.85	2494.12	3006.26	3617.60

De onde concluímos que este é um conjunto unimodal.

Simplificação do Ex 2 (Seguros): Moda

Podemos olhar para um outro subconjunto com 12 valores dos seguros vendidos:

1408.50 1630.07 1771.85 977.94 1139.64
1137.29 1408.50 2494.12 3617.60 3006.26 1534.65

Nesse caso, quando ordenamos e contabilizamos as repetições, identificamos um conjunto de dados bimodal.

				1627.1					4078.9
1197.73	1285.61	1380.09	1394.53	1627.1	1969.95	2208.3	3531.26	3677.57	4078.9

Ex 02 (Seguros): Moda

Independente de qualquer detalhamento mais específico, qual é o valor que cobrado pela seguradora com maior recorrência, durante o primeiro semestre para os dois modelos, estados e públicos observados?

Observamos que os valores cobrados com maior frequência pela seguradora são:

966.92, 1097.23, 1157.77, 1171.84, 1340.59 1408.50, 1455.31, 1563.71, 1627.11, 1781.28 1788.43, 1922.19, 1959.21, 1972.73, 2132.01 2208.30, 2223.81, 2246.48, 2536.31, 2599.00 2916.46, 4078.93

*

Observação

Veja que determinar um valor para a moda em um conjunto amplo de observações, particularmente representando uma variável contínua (o valor pago por seguro) não parece ser razoável. Neste exemplo, identificamos 22 modas no conjunto de dados.

Será que essas “modas” encontradas realmente refletem os valores mais prováveis para os valores dos seguros?

Agora, vamos contabilizar quantas vezes esses valores ocorrem no conjunto de dados.

Ex 02 (Seguros): Moda

966.92	2
1097.23	2
1157.77	2
1171.84	2
1340.59	2
1408.50	2
1455.31	2
1563.71	2
1627.11	2
1781.28	2
1788.43	2
1922.19	2
1959.21	2
1972.73	2
2132.01	2
2208.30	2
2223.81	2
2246.48	2
2536.31	2
2599.00	2
2916.46	2
4078.93	2

(843.814, 1249.31]	296
(1249.31, 1650.79]	713
(1650.79, 2052.27]	753
(2052.27, 2453.75]	624
(2453.75, 2855.23]	411
(2855.23, 3256.71]	211
(3256.71, 3658.19]	68
(3658.19, 4059.67]	47
(4059.67, 4461.15]	16
(4461.15, 4862.63]	3

*

Interpretação
Ex 02 (Seguros): Moda

Note que o intervalo com maior acúmulo de observações é o intervalo **(1650.79, 2052.27]**, com **753** registros;

Muitos valores anteriormente indicados como "modas"do conjunto de dados estão em intervalos com um número de registros muito inferior a 753;

Observe o valor 4078.93 que, embora tenha sido observado tantas vezes quanto o valor 1788.43, por exemplo, está em um intervalo onde encontram-se somente **16** registros, **(4059.67, 4461.15]**;

Nesse contexto, é importante sabermos que, quando estamos lidando com variáveis cuja **possibilidade de valores é expressivamente grande**, ou mesmo **infinita**, como o caso de uma **variável contínua**, deve-se usar a definição anterior com cautela. A técnica de dividir o conjunto de dados em “classes de frequências” é a mais apropriada, e será abordada posteriormente.

Observação

Dificuldade

A Moda é pouco informativa quando muitos valores apresentam frequências similares.

Indicação

Utilizá-la em conjunto de dados agrupados (veremos com mais detalhes posteriormente).

Separatrizes

As **separatrizes** (ou **quantis**) são medidas que dividem um conjunto de dados ordenados em partes de igual proporção.

As principais medidas separatrizes são:

- Mediana
- Quartis
- Decis
- Percentis ou Centis

Mediana

A **mediana** é o valor que divide um conjunto de dados ordenados (de forma crescente ou decrescente) em duas partes iguais, de modo que 50% das observações se encontra abaixo e, 50% das observações se encontra acima dela.

Definição

Sejam $x_1, x_2, \ldots, x_n$ valores observados de uma variável X. Considere o conjunto $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ que denota os mesmos dados, ordenados de forma crescente ou decrescente. O cálculo da **mediana** leva em conta dois casos:

Se n for ímpar, a mediana é dada por:

$$\text{Med} = X_{(n+1)/2}.$$

Se n for par, a mediana é dada por:

$$\text{Med} = \frac{X_{n/2} + X_{(n+2)/2}}{2}.$$

Observação

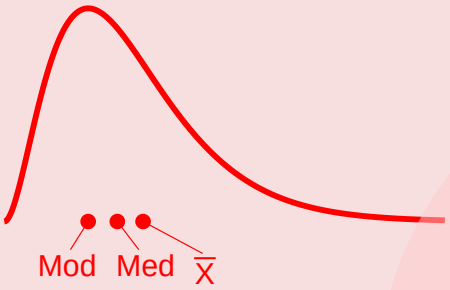
A mediana é uma medida de posição resistente, pois não é afetada por valores atípicos (**outliers**). Ao contrário da média, que é sensível à presença de valores extremos.

A mediana se relaciona empiricamente com **média** e a **moda**. A forma como essa relação acontece caracteriza a distribuição dos valores de uma variável.

OBSERVAÇÃO

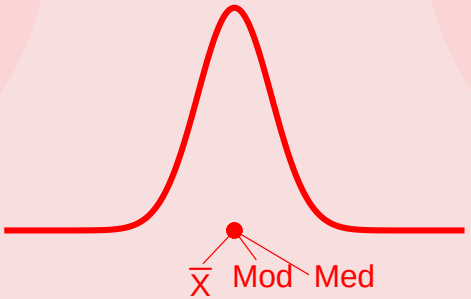
QUANTO A SIMETRIA UMA VARIÁVEL PODE SER

ASSIMÉTRICA À DIREITA



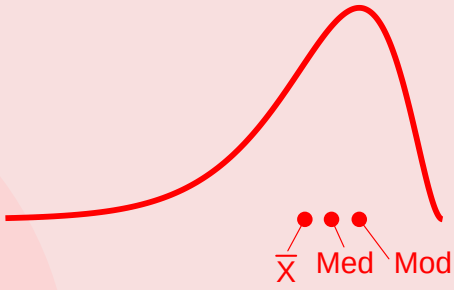
$Mod < Med < \overline{X}$

SIMÉTRICA



$\overline{X} \approx Med \approx Mod$

ASSIMÉTRICA À ESQUERDA



$\overline{X} < Med < Mod$

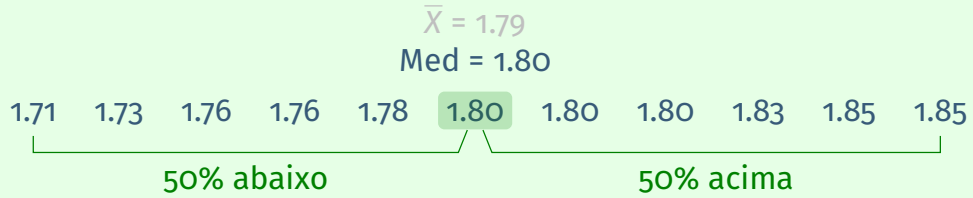
Ex 01 (Alturas): Mediana

Voltando ao conjunto de dados com tamanho $n = 11$. ***Os valores das alturas são dados por:***

1.85 1.85 1.76 1.78 1.80 1.71
1.73 1.76 1.80 1.83 1.80

Qual o valor da mediana associada a esse conjunto de dados?

Nesse caso, a mediana é exatamente o valor central do conjunto de dados.

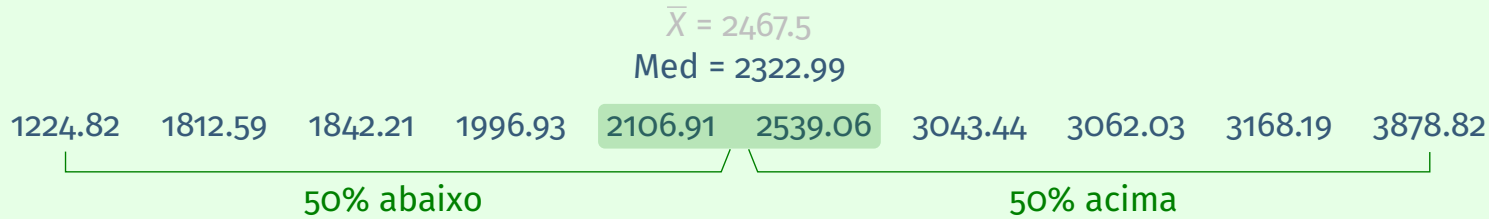


Simplificação do Ex 02 (Seguros): Mediana

Considere uma sub-amostra de 10 indivíduos, isto é, $n = 10$***, dos valores dos seguros vendidos, utilizados nos exemplos anteriores:***

2106.91 3062.03 3168.19 2539.06 1812.59
3043.44 3878.82 1842.21 1996.93 1224.82

A mediana é determinada considerando os dois valores centrais

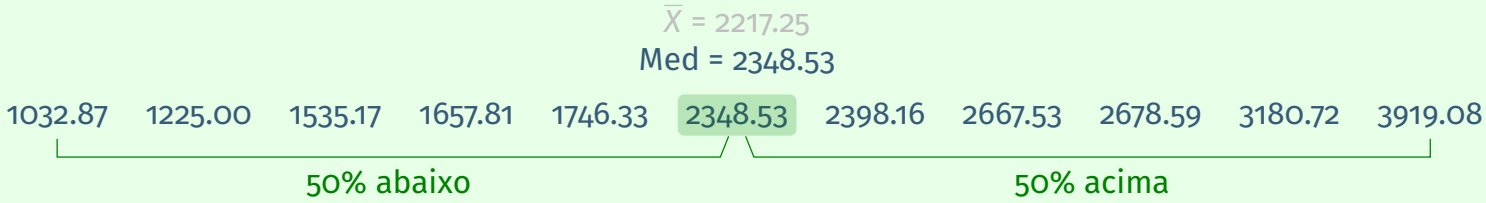


Simplificação do Ex 02 (Seguros): Mediana

Agora considere uma amostra com tamanho $n = 11$***. Os valores de seguros negociados são dados por:***

2678.59 1225.00 1657.81 1032.87 1746.33 2348.53
3180.72 1535.17 2398.16 2667.53 3919.08

Portanto, a mediana nesse caso é exatamente o valor central do conjunto de dados.

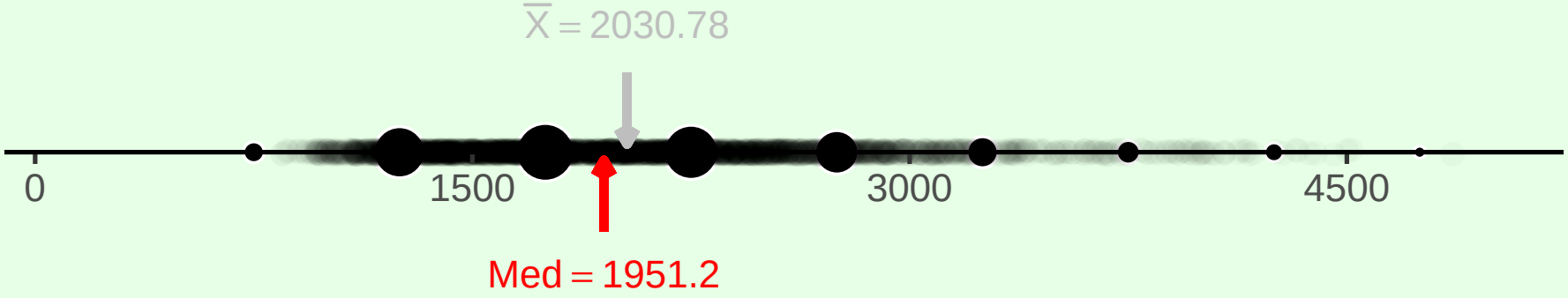


Ex 02 (Seguros): Mediana

Considere os valores dos seguros vendidos utilizados nos exemplos anteriores.

Independente de qualquer detalhamento, qual é a mediana do Valor do Seguro Vendido, representado pela variável VALOR_SEGURO_VENDIDO? Quão distinta da média ela é?

O valor mediano do seguro vendido durante o primeiro semestre de 2020, para os dois veículos e estados observados é igual a R$ 1951.2, por outro lado, o valor médio é igual a R$ 2030.78.



*

Interpretação
Ex 02 (Seguros): Mediana

Note a diferença entre a **média** e a **mediana**. Ela é próxima de R$ 80.00;

A **mediana** nos diz que 50% de todos os seguros vendidos, são vendidos por valores menores do que R$ 1951.2, enquanto o 50% restante se distribui até o valor máximo observado;

Visto que $\overline{X} > \text{Med}$, entendemos que os valores dos seguros vendidos se distribuem de modo assimétrico à direita, isto é um indicativo de que a média está sendo "puxada" pelos maiores valores.

A visualização gráfica desse tipo de comportamento será estudada posteriormente.

Outras Separatrizes

As demais **medidas separatrizes** se distinguem da mediana somente na quantidade de partes em que o conjunto de dados é dividido. As mais comuns entre elas, são

- **Quartis** (Q_1, Q_2 e Q_3)
  Dividem um conjunto em 4 partes iguais;
- **Decis** ($D_1, \ldots, D_9$)
  Dividem um conjunto em 10 partes iguais;
- **Percentis** ou **Centis** ($P_i, \ldots, P_{99}$)
  Dividem um conjunto em 100 partes iguais.

ALGUMAS RELAÇÕES ENTRE AS SEPARATRIZES

$Q_1 = P_{25}$;
$Q_2 = D_5 = P_{50}$ = Med;
$Q_3 = P_{75}$;
$D_1 = P_{10}, \ldots, D_9 = P_{90}$.

Ex 01 (Alturas): Quartis

Considerando o conjunto de dados com tamanho $n = 11$***. Os valores das alturas são dados por:***

1.85 1.85 1.76 1.78 1.80 1.71
1.73 1.76 1.80 1.83 1.80

Quais são os três quartis associados a esse conjunto de dados?

O valor de 25% das alturas está abaixo de 1.76.
O valor de 50% das alturas está abaixo de 1.80.
O valor de 75% das alturas está abaixo de 1.83.

A diferença entre Q2 e Q1 é: 4cm.
A diferença entre Q3 e Q2 é: 3cm.

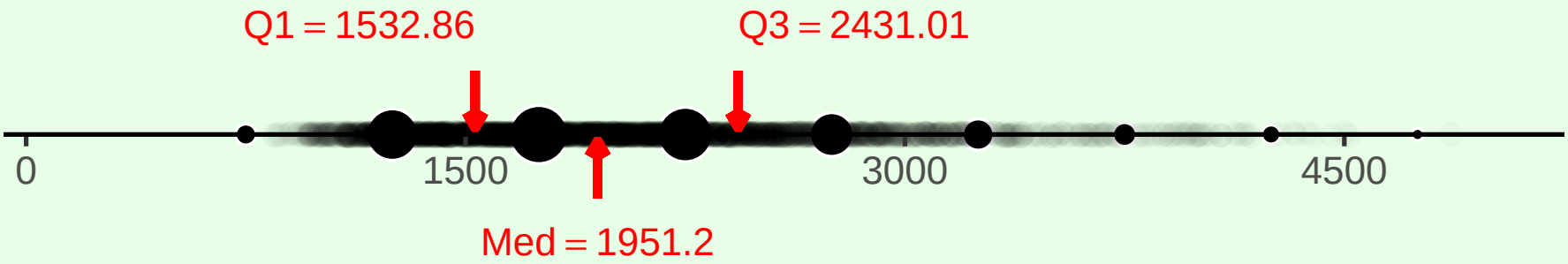
Ex 02 (Seguros): Quartis

Considere os valores dos seguros vendidos utilizados nos exemplos anteriores.

Independente de qualquer detalhamento, quais são os três quartis associados ao Valor do Seguro Vendido, representado pela variável VALOR_SEGURO_VENDIDO? Que informação eles podem nos passar?

O valor de 25% dos seguros vendidos está abaixo de R$ 1532.86.
O valor de 50% dos seguros vendidos está abaixo de R$ 1951.2.
O valor de 75% dos seguros vendidos está abaixo de R$ 2431.01.

A diferença entre Q2 e Q1 é: 418.34.
A diferença entre Q3 e Q2 é: 479.81.



Interpretação
Ex 02 (Seguros): Quartis

Note a magnitude das diferenças:

- $Q_2 - Q_1 = 418.34$;
- $Q_3 - Q_2 = 479.81$.

Os 25% dos valores entre Q_1 e Q_2, se concentram em uma diferença aproximadamente R$ 60.00 inferior aos 25% entre Q_2 e Q_3, o que indica a assimetria com que os valores observados se distribuem.

Observação

Embora as **medidas de posição** sejam muito populares na prática, **não recomendamos utilizá-las isoladamente**, pois, elas não carregam consigo a noção de um conceito muito importante na estatística, a **variabilidade**.

Por exemplo, considere os conjuntos

$$X_1 = \{5, 5, 5, 5, 5\} \quad \text{e} \quad X_2 = \{2.7, 5, 3, 8\},$$

ambos possuem médias e medianas iguais a cinco, isto é $\overline{X}_1 = \overline{X}_2 = 5$ e $\text{Med}_1 = \text{Med}_2 = 5$, no entanto, existe uma característica muito importante que os distingue, a **variabilidade**.

Resumo do tópico

- **IDEIAS BÁSICAS, MOTIVAÇÕES E APLICAÇÕES**
- **CONCEITOS FUNDAMENTAIS**
  - ***POPULAÇÃO***
  - ***CENSO***
  - ***AMOSTRA***
  - ***VARIÁVEL***
- **ESTRUTURAS DE SUMARIZAÇÃO**
  - ***MEDIDAS DE POSIÇÃO***

    Média Aritmética Simples e Ponderada, Moda, Separatrizes (Mediana, Quartis, Decis, Percentis);

Obrigado!

Continuamos no próximo encontro!