

Controladoria e Contabilidade

Análise Multivariada Aplicada à Contabilidade

Prof. Dr. Marcelo Botelho da Costa Moraes

www.marcelobotelho.com

mbotelho@usp.br

Turma: 2º / 2016

Agenda – Aula 1/15

- Apresentação da Disciplina e Critérios de Avaliação;
- Tratamento de Dados e Métodos de Análise de Dados;
- Tipos de Variáveis e Escalas de Mensuração;
- Conceitos de Estatística e Probabilidade: Estatística Descritiva.

Ementa da Disciplina

- Tratamento de Dados e Métodos de Análise de Dados;
- Tipos de Variáveis e Escalas de Mensuração;
- Conceitos de Estatística e Probabilidade:
 - Estatística Descritiva;
 - Inferência Estatística: - Testes de Hipóteses;
 - Testes de Médias;
 - Testes Não Paramétricos;
- Análise Multivariada:
 - Conceitos Básicos e Tipos de Técnicas Multivariadas;
 - Definições de Exogeneidade/Causalidade;
 - Regressão Linear Simples;
 - Regressão Linear Múltipla;
 - Análise Discriminante Múltipla;
 - Regressão Logística;
 - Análise Fatorial;
 - Análise de Conglomerados.
- **Dados em Painel???**

Conhecendo a Disciplina

- Ver Programa da Disciplina
- <http://disciplinas.stoa.usp.br/>
- Análise Multivariada Aplicada à Contabilidade

Métodos Quantitativos

Quanto aos Métodos

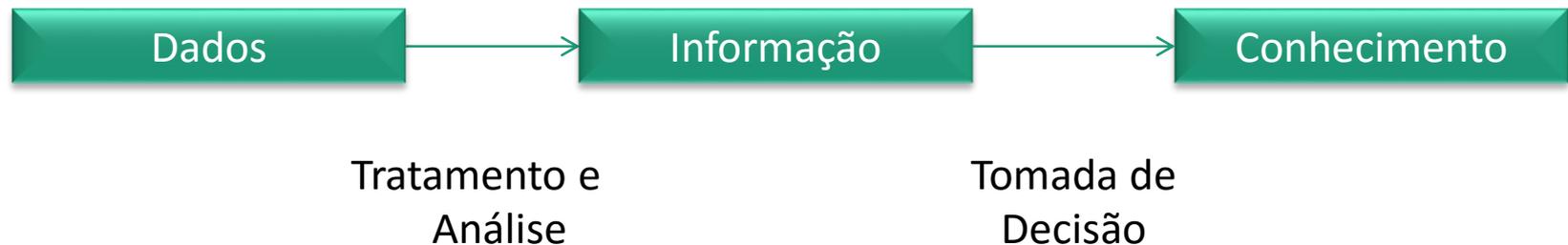
- Métodos Estatísticos
 - Inferência Clássica
 - Inferência Bayesiana
 - Econometria

- Métodos Computacionais
 - Simbólicos
 - Baseados em Memorização
 - Genéticos
 - Conexionistas

Tratamento de Dados e Métodos de Análise de Dados

Análise Multivariada

- Avaliação de comportamentos e tendências para a tomada de decisão



Processo de geração de conhecimento de um fenômeno (FÁVERO et al, 2009)

Técnicas Multivariadas de Análise de Dados

- Análise univariada: comportamento e distribuição de uma variável isolada
- Análise bivariada: associação, correlação e análise de variância relativas a duas variáveis simultaneamente
- Análise multivariada: todas as variáveis são aleatórias e inter-relacionadas, de modo que seus diferentes efeitos não podem ser interpretados separadamente

Técnicas Multivariadas de Análise de Dados



Tipos de análise de dados em função do número de variáveis (FÁVERO et al, 2009)

Objetivos da análise multivariada

- Redução dos dados ou simplificação estrutural
- Seleção ou agrupamento
- Investigação da existência de dependência entre variáveis
- Previsão
- Construção e verificação de hipóteses

Esquema de Banco de Dados

Variáveis

	Var 1	Var 2	Var 3	...	Var m
Obs 1					
Obs 2					
Obs 3					
...					
Obs n					

Observações

Observações

Matriz $n \times m$

Preparação da Análise de Dados

- Número de variáveis a serem analisadas ao mesmo tempo
- Nível de mensuração das variáveis de interesse
- Interesse descritivo ou inferencial da análise
- Interesse na inter-relação ou na relação de dependência das variáveis em estudo (para duas ou mais variáveis)

Dados Qualitativos e Quantitativos

- **Variável qualitativa:** análise estatística limitada. Síntese em número de observações ou proporções em cada categoria (dados categóricos) – sem operações aritméticas
- **Variável quantitativa:** dados quantitativos proporcionam maior facilidade e profundidade em análise e síntese

Dados de Seção Transversal

- Dados coletados no mesmo intervalo de tempo (ou aproximadamente) de diferentes elementos
- *Cross-Sectional*

Dados de Seção Transversal

Variáveis

	Var 1	Var 2	Var 3	...	Var m
Obs 1					
Obs 2					
Obs 3					
...					
Obs n					

Observações

Observações

Dados de Série Histórica

- Dados coletados ao longo de diversos períodos de um mesmo elemento
- *Time-Series* ou Série Temporal

Dados de Série Histórica

Variáveis

	Var 1	Var 2	Var 3	...	Var <i>m</i>
<i>T</i> ₁					
<i>T</i> ₂					
<i>T</i> ₃					
...					
<i>T</i> _{<i>n</i>}					

Tempo

Tempo

Dados em Painel

- Dados coletados de diferentes elementos coletados ao longo do tempo
- Seção Transversal + Série Histórica
- *Panel Data* ou Dados Longitudinais

Dados em Painel – CS Empilhada

	T	Var 1	Var 2	...	Var m
Obs 1	1				
Obs 2	1				
...	...				
Obs n	1				
Obs 1	2				
Obs 2	2				
...	...				
Obs n	2				
...	...				
Obs 1	<i>t</i>				
Obs 2	<i>t</i>				
...	...				
Obs n	<i>t</i>				

Dados em Painel – ST Empilhada

	T	Var 1	Var 2	...	Var m
Obs 1	1				
Obs 1	2				
...	...				
Obs 1	t				
Obs 2	1				
Obs 2	2				
...	...				
Obs 2	t				
...	...				
Obs n	1				
Obs n	2				
...	...				
Obs n	t				

Técnicas de Interpedendência

- **Análise de Conglomerados:** objetivo de agrupar indivíduos em grupos homogêneos em função da similaridade dos valores de suas variáveis
- **Análise Fatorial:** objetivo de condensar a informação contida em cada grupo de variáveis originais em um conjunto menor de variáveis (fatores), sem que haja perda considerável de informação

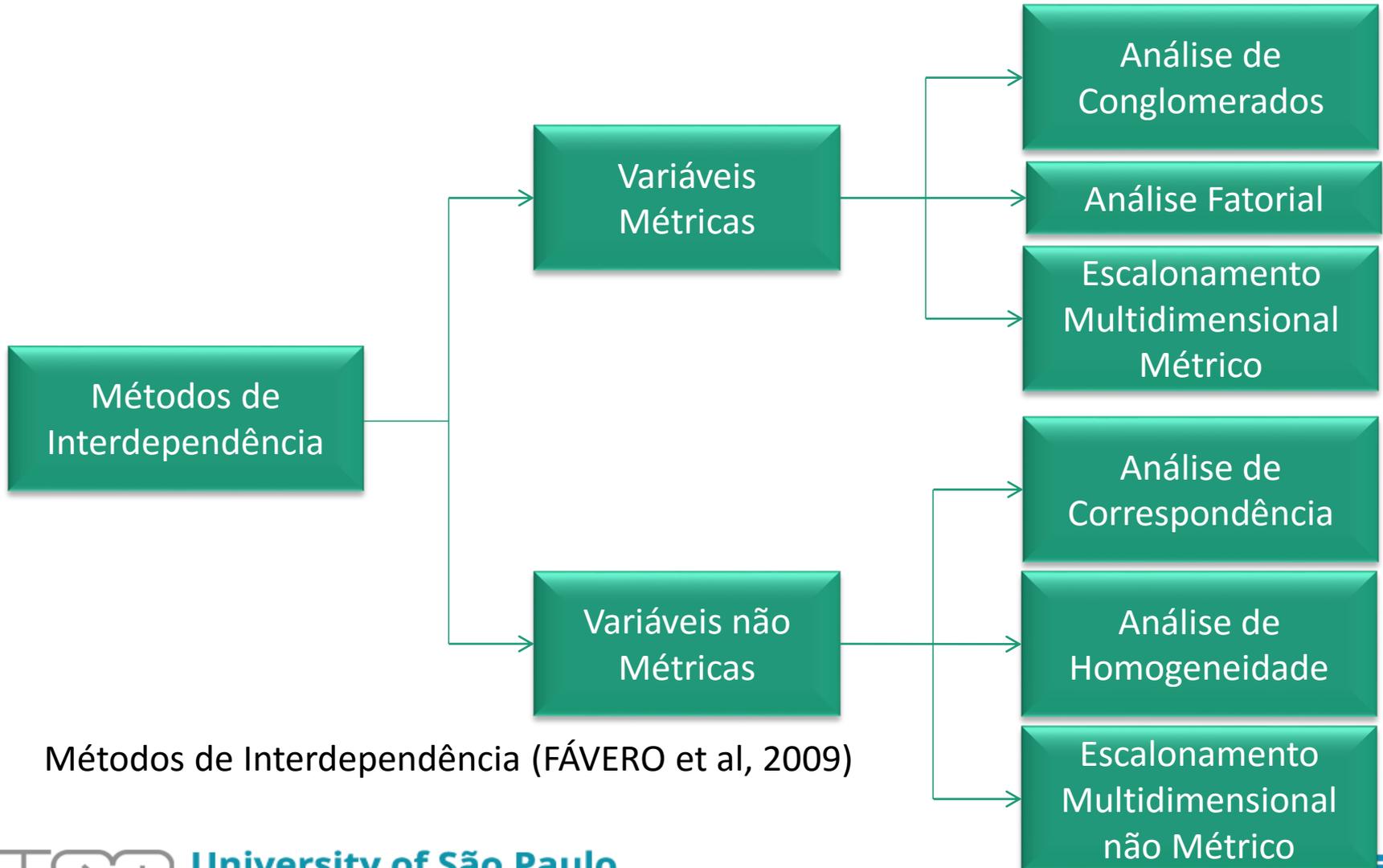
Técnicas de Interpedendência

- **Análise de Correspondência:** objetivo de estudar a associação das categorias presentes e duas variáveis não métricas e, ao mesmo tempo, criar mapas perceptuais que facilitem a visualização dessa associação
- **Análise de Homogeneidade:** objetivo de analisar dados categóricos associando variáveis qualitativas em função de seus níveis e do grau de complexidade

Técnicas de Interpedendência

- **Escalonamento Multidimensional:** objetivo de mapear as distâncias entre pontos em determinada representação gráfica espacial e, portanto, ajuda o pesquisador na identificação de dimensões-chave inerentes a avaliações feitas por correspondentes quanto a determinados objetos

Técnicas de Interdependência



Métodos de Interdependência (FÁVERO et al, 2009)

Técnicas de Dependência

- **Análise de Regressão:** utilizada quando o problema apresenta uma única variável dependente métrica relacionada com uma ou mais variáveis explicativas (independentes ou preditoras) métricas ou não métricas (binárias, variáveis *dummy*)
- **Análise Discriminante:** utilizada quando uma variável dependente é dicotômica ou multicotômica (não métrica), com variáveis explicativas métricas ou binárias

Técnicas de Dependência

- **Regressão Logística e Regressão Logística Multinomial:** utilizada para a criação de modelos em que diversas variáveis métricas ou dicotômicas são utilizadas para a previsão de uma variável dependente binária
- **Análise Multivariada de Variância ou Análise Multivariada de Variância com Medidas Repetidas:** permite testar diferenças de médias para diversas variáveis dependentes simultaneamente

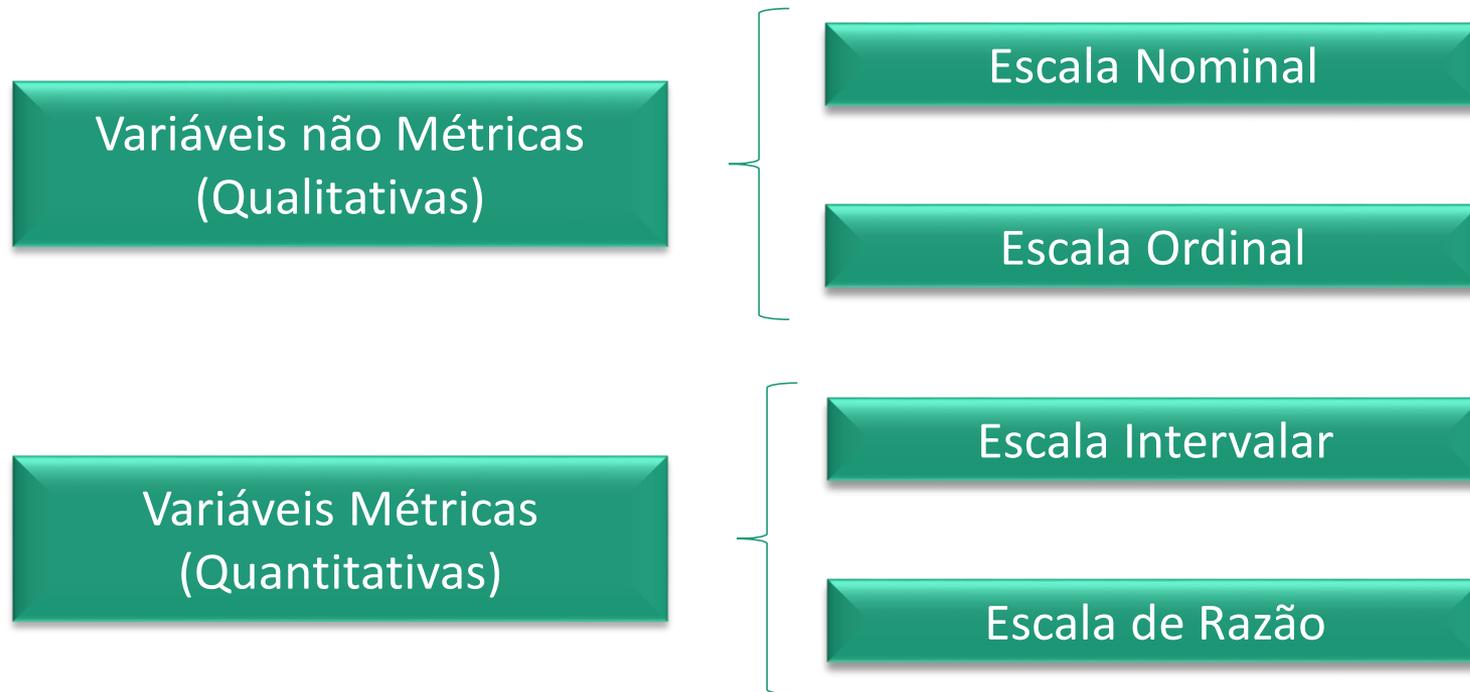
Técnicas de Dependência

- **Correlação Canônica:** tem como objetivo quantificar a força da relação existente entre dois vetores de variáveis, representados pelas dependentes métricas e pelas explicativas também métricas, e desenvolver uma combinação linear de cada conjunto de variáveis com o intuito de maximizar a correlação entre eles

Técnicas de Dependência	Variáveis Dependentes	Variáveis Explicativas
Análise de Regressão Simples	Y – métrica	X – métrica ou não métrica (<i>dummy</i>)
Análise de Regressão Múltipla	Y – métrica	X1, X2, ..., Xn – métricas ou não métricas (<i>dummies</i>)
Análise Discriminante	Y – não métrica	X1, X2, ..., Xn – métricas ou não métricas (<i>dummies</i>)
Regressão Logística	Y – não métrica (<i>dummy</i>)	X1, X2, ..., Xn – métricas ou não métricas (<i>dummies</i>)
Análise Multivariada de Variância	Y1, Y2, ..., Yn – métricas ou não métricas (<i>dummies</i>)	X1, X2, ..., Xn – não métricas
Correlação Canônica	Y1, Y2, ..., Yn – métricas ou não métricas (<i>dummies</i>)	X1, X2, ..., Xn – métricas ou não métricas (<i>dummies</i>)

Tipos de Variáveis e Escalas de Mensuração

Escalas de Mensuração



Tipos de variáveis e escalas de mensuração (FÁVERO et al, 2009)

Escalas de Mensuração

Variáveis Qualitativas (Não Métricas)

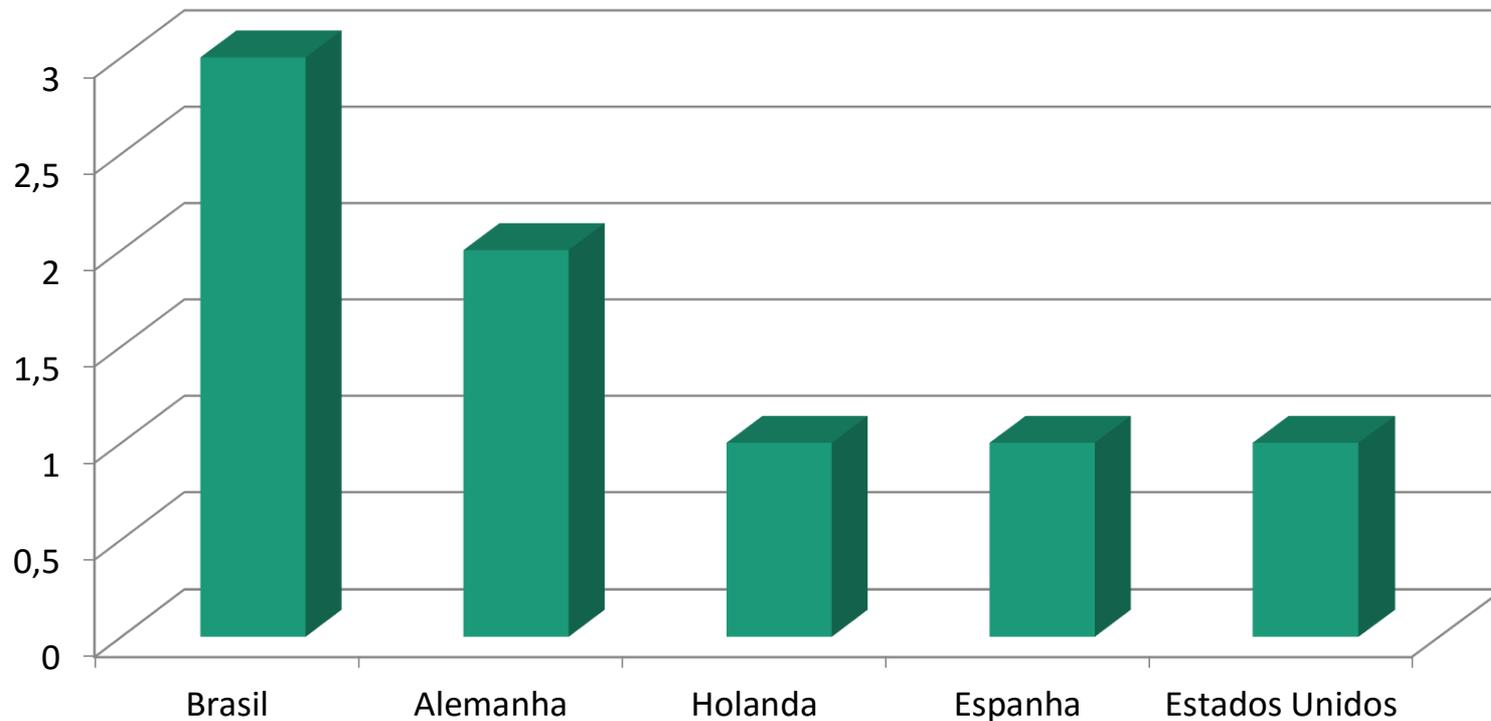
- **Escala nominal:** dados que consistem em rótulos ou nomes usados para identificar um atributo do elemento
- **Escala ordinal:** os dados apresentam as propriedades dos dados nominais e se a ordem ou classificação dos dados for significativa

Apresentação Variáveis Qualitativas

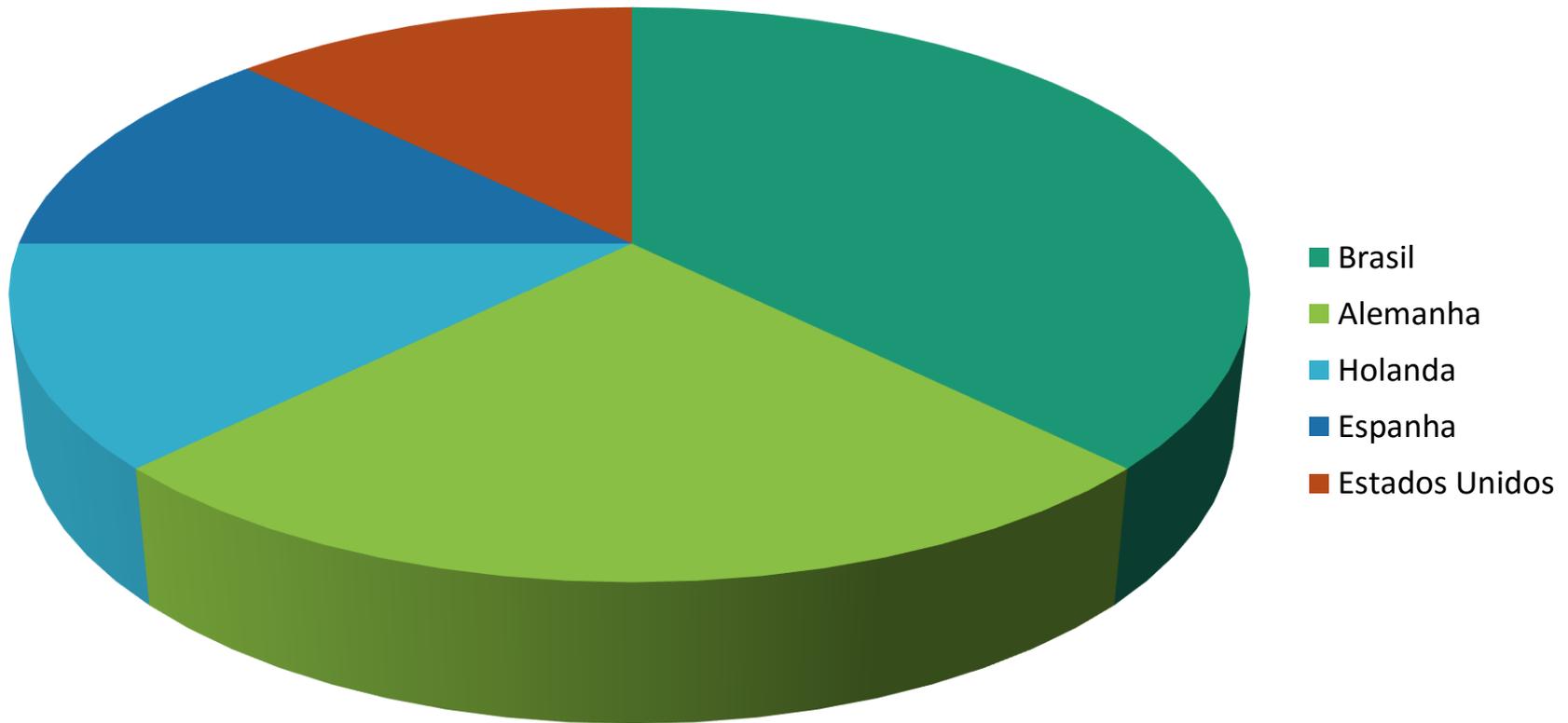
Empresa	País de Origem
Gerda	1
Citigroup	5
Santander	4
Volkswagen	2
Wal-Mart	5
Petrobras	1
Vale	1
Philips	3
Boeing	5
Siemens	2

- 1 – Brasil
- 2 – Alemanha
- 3 – Holanda
- 4 – Espanha
- 5 – Estados Unidos

Apresentação Variáveis Qualitativas



Apresentação Variáveis Qualitativas



Tipos de Questões – Escalas Qualitativas

- **Dicotômicas:** questões com duas respostas possíveis

Exemplo:

“Eu gostei do curso de Estatística Aplicada”

Sim []

Não []

Tipos de Questões – Escalas Qualitativas

- **Múltipla Escolha:** questões com três ou mais respostas possíveis

Exemplo:

“Com quem você pretende fazer este curso”

Ninguém []

Colegas de Trabalho []

Familiares / amigos []

Tipos de Questões – Escalas Qualitativas

- **Escala Likert:** afirmação na qual o entrevistado mostra o quanto concorda ou discorda

Exemplo:

“Este curso ajudou a aprimorar meus conhecimentos de estatística e me tornou mais apto a utilizar ferramentas para a tomada de decisão.”

Discordo totalmente []	Discordo parcialmente []	Nem concordo, nem discordo []	Concordo parcialmente []	Concordo totalmente []
----------------------------	------------------------------	-----------------------------------	------------------------------	----------------------------

Tipos de Questões – Escalas Qualitativas

- **Diferencial Semântico:** escala inserida entre duas palavras antônimas e o entrevistado seleciona o ponto que representa a direção e a intensidade de seus sentimentos

Exemplo:

“O professor, em sua opinião, é:”

Ágil ___:___:___:___:___: Moroso

Experiente ___:___:___:___:___: Inexperiente

Cativante ___:___:___:___:___: Não Cativante

Exemplo – Escalas Qualitativas

- A higiene e a limpeza do recinto são adequadas

Discordo totalmente	Discordo parcialmente	Nem concordo, nem discordo	Concordo parcialmente	Concordo totalmente
12%	7%	14%	42%	25%

Erro comum: adotar pesos para cada categoria e calcular a média!!!

Tipos de Questões Abertas – Escalas Qualitativas

- **Completamente desestruturada:** questão que o entrevistado pode responder de várias formas

Exemplo:

“Qual é sua opinião sobre o curso de Estatística?”

Tipos de Questões Abertas – Escalas Qualitativas

- **Associação de palavras:** apresenta-se uma palavra de cada vez, e o entrevistado deve mencionar a primeira palavra que lhe vier à cabeça

Exemplo:

“Qual é a primeira palavra que vem à sua cabeça quando ouve o seguinte:”

Dados _____

Estatística _____

Tomada de Decisão _____

Tipos de Questões Abertas – Escalas Qualitativas

- **Completar a sentença:** apresenta-se uma frase incompleta de cada vez, e o entrevistado deve completá-la

Exemplo:

“Quando escolho determinado curso, o mais importante para mim é:”

Tipos de Questões Abertas – Escalas Qualitativas

- **Completar uma história:** apresenta-se uma história incompleta e o entrevistado deve completá-la

Exemplo:

“Eu me matriculei em um curso de Estatística e, no primeiro dia de aula, percebi que os demais alunos eram brilhantes. Isso despertou em mim os seguintes pensamentos: ...” Agora você completa a história.

Tipos de Questões Abertas – Escalas Qualitativas

- **Completar um quadro:** apresenta-se um quadro com personagens, sendo que um deles faz uma declaração. O entrevistado deve preencher o balão vazio como se fosse o personagem

Exemplo:

Preencha o balão vazio



Tipos de Questões Abertas – Escalas Qualitativas

- **Testes de percepção temática (TAT – *Thematic Apperception Tests*)**: apresenta-se um quadro, e o entrevistado deve dizer o que acha que está ocorrendo ou o que pode ocorrer na situação apresentada

Exemplo:

Crie uma história sobre esta cena



Escalas de Mensuração

Variáveis Quantitativas (Métricas)

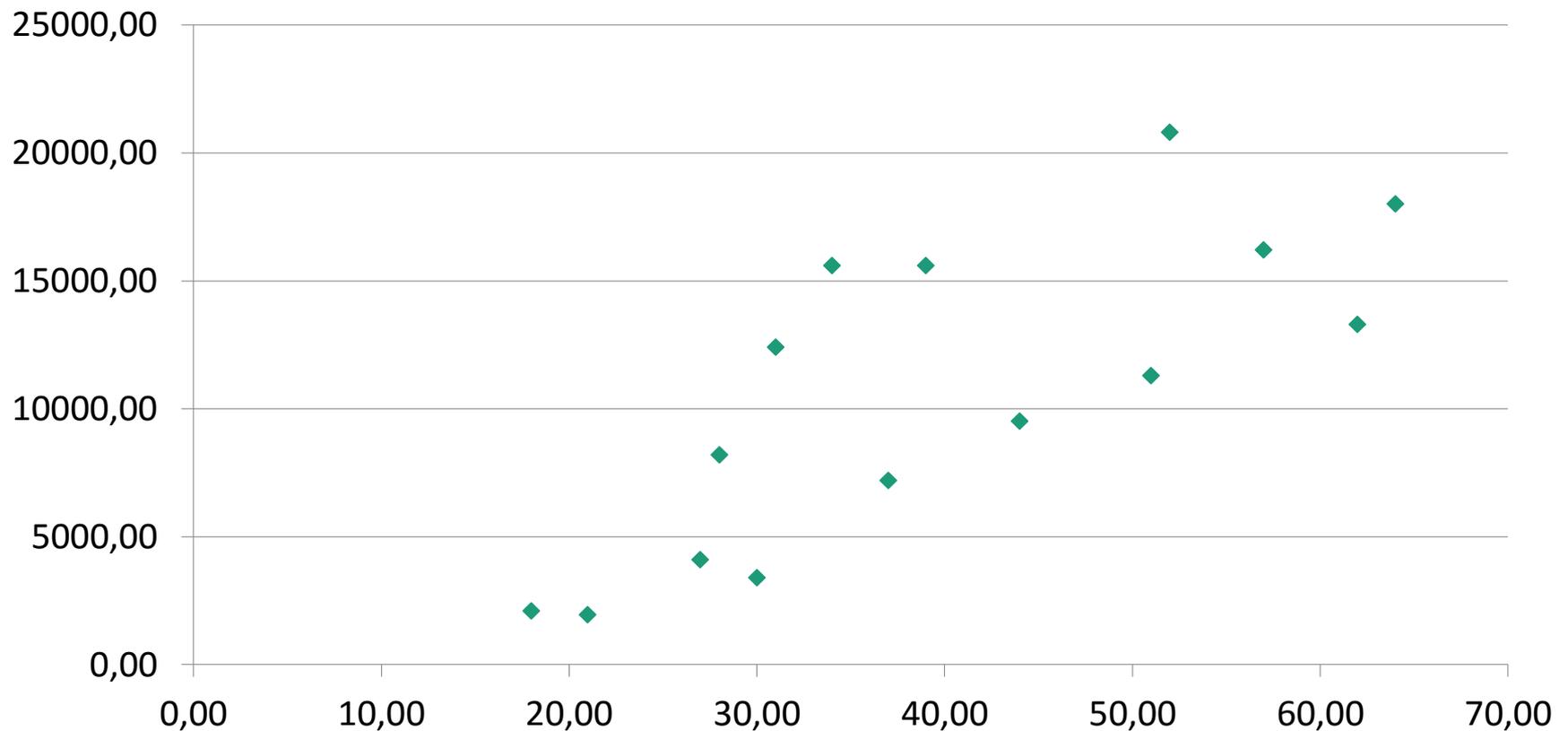
- **Escala intervalar:** os dados apresentam as propriedades dos dados ordinais e o intervalo entre os valores for expresso em termos de uma unidade de medida fixa
- **Escala de razão:** os dados apresentam as propriedades dos dados de intervalo e a proporção de dois valores for significativa, a partir da origem da escala

Apresentação Variáveis Quantitativas

nome	idade	renda_mensal
Patrícia	31,00	12400,00
Leonor	64,00	18000,00
Ovídio	57,00	16200,00
Mariana	37,00	7200,00
Roberto	52,00	20800,00
Letícia	18,00	2100,00
Dalila	21,00	1950,00
Luiz	34,00	15600,00
Paulo	30,00	3400,00
...

Apresentação Variáveis Quantitativas

Gráfico de Dispersão



Apresentação Variáveis Quantitativas

Variável	Conceito Original	Faixa (grupo)
Idade	Menos de 29 anos	1
	De 29 a 37 anos	2
	De 37 a 52 anos	3
	Mais de 52 anos	4
Renda	Menos de R\$ 5.600,00	1
	De R\$ 5.600,00 a R\$ 11.300,00	2
	De R\$ 11.300,00 a R\$ 15.600,00	3
	Mais de R\$ 15.600,00	4

Regras de Mapeamento

- **Classificação:** podem ser utilizados números para agrupar ou designar as respostas, porém sem que haja uma ordem
- **Ordem:** podem ser utilizados números de acordo com uma ordenação
- **Distância:** as diferenças entre os números são ordenadas, ou seja, a diferença entre qualquer par de números é maior que, menor que ou igual à diferença entre outro par de números
- **Origem:** a série de números tem uma origem única indicada por zero

Regras de Mapeamento

Tipo de Dado	Característica dos dados	Exemplo
Nominal	Classificação sem ordem, distância ou origem	Sexo (masculino ou feminino)
Ordinal	Classificação e ordem, mas sem distância ou origem única	Preparo de carnes (mal passada, ao ponto, bem passada)
Intervalar	Classificação, ordem e distância, mas sem origem única	Temperatura em graus
De razão	Classificação, ordem, distância e origem única	Idade em anos

Fonte: Cooper e Schindler (2003)

Variáveis *Dummy*

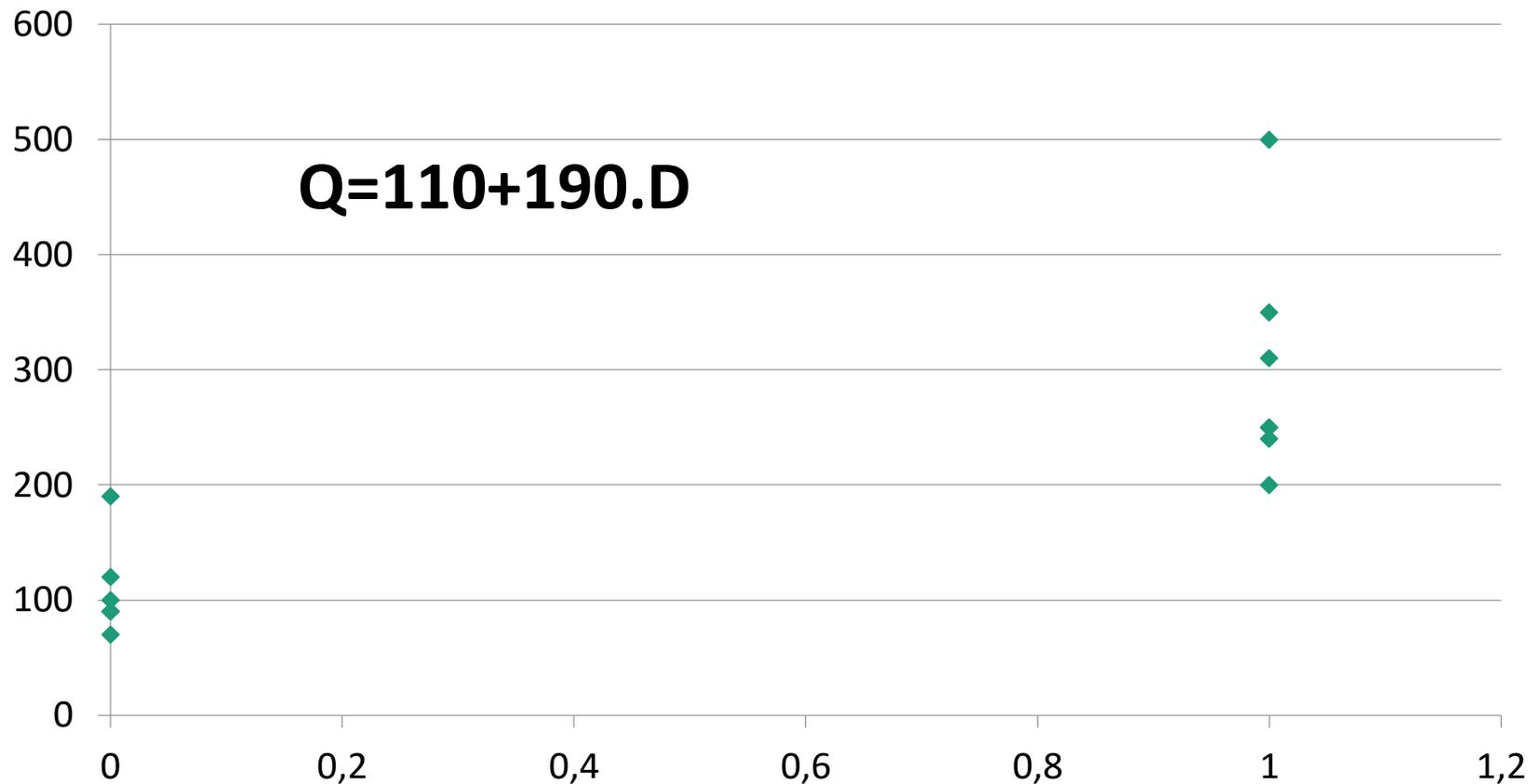
- Variável nominal dicotômica

Consumo de chocolate	Sexo
250	1
240	1
190	0
90	0
70	0
90	0
310	1
350	1
500	1
120	0
100	0
200	1
250	1

0 – Masculino
1 – Feminino

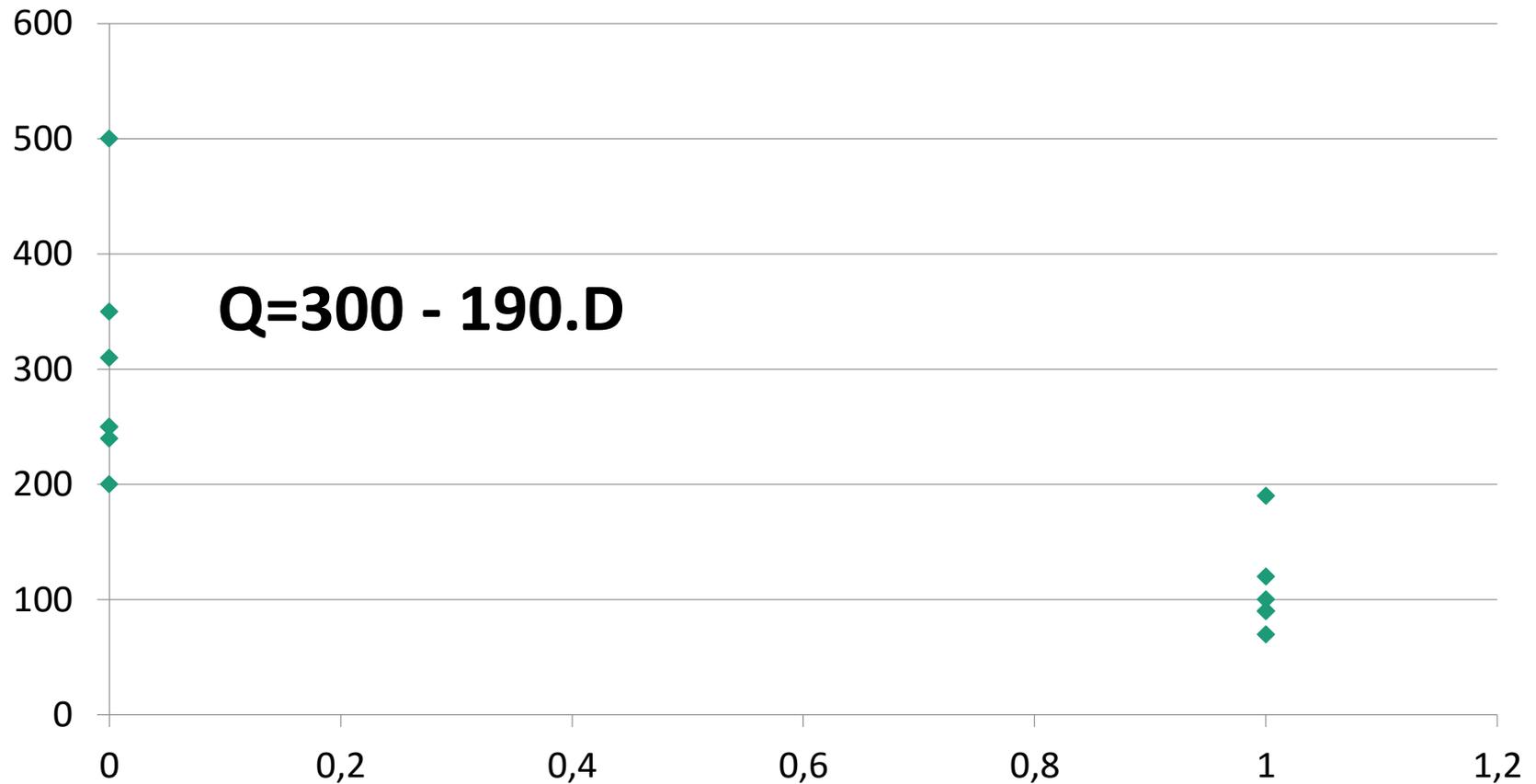
Variáveis *Dummy*

Consumo de chocolate



Variáveis *Dummy* (invertendo)

Consumo de chocolate



Variáveis *Dummy*

Consumo de chocolate	Profissão
250	Contador
240	Engenheiro
190	Engenheiro
90	Médico
70	Médico
90	Engenheiro
310	Contador
350	Contador
500	Engenheiro
120	Contador
100	Médico
200	Médico
250	Engenheiro

Variáveis *Dummy*

Profissão	<i>Dummies</i>	
	D ₁	D ₂
Contador	0	0
Engenheiro	1	0
Médico	0	1

$$Q = \alpha + \beta_1 \cdot D_1 + \beta_2 \cdot D_2$$

α : constante que representa o consumo médio de chocolate dos contadores ($D_1 = 0$; $D_2 = 0$);

β_1 : diferença entre o consumo médio de chocolate de engenheiros e contadores ($D_1 = 1$; $D_2 = 0$);

$\alpha + \beta_1$: consumo médio de chocolate de engenheiros;

β_2 : diferença entre o consumo médio de chocolate de médicos e contadores ($D_1 = 0$; $D_2 = 1$);

$\alpha + \beta_2$: consumo médio de chocolate de médicos;

$\beta_1 - \beta_2$: diferença entre o consumo médio de chocolates de médicos e engenheiros

Conceitos de Estatística e Probabilidade: Estatística Descritiva

Medidas de Estatística Descritiva

- Medidas de Posição
 - Medidas de Tendência Central
 - Medidas de Tendência Não Central
- Medidas de Dispersão ou Variabilidade
- Medidas de Assimetria e Curtose

Medidas de Tendência Central

- **Média Aritmética**, ou apenas média, representa a soma do total de valores de uma variável dividida pelo número total de observações

$$\bar{x} = \frac{\sum x_i}{n}$$

Medidas de Tendência Central

- **Mediana** (Md), valor que ocupa a posição central de uma série de observações ordenadas de forma crescente (divide a amostra em 50%)

$$Md(x) = \begin{cases} \frac{x_{\frac{n}{2}} + x_{(\frac{n}{2})+1}}{2}, & \text{se } n \text{ for par} \\ \frac{x_{(n+1)}}{2}, & \text{se } n \text{ for ímpar} \end{cases}$$

Medidas de Tendência Central

- **Moda** (M_o), valor mais frequente do conjunto de observações de determinada variável (única que pode ser calculada para variáveis qualitativas – frequência)

Medidas de Tendência Central

Mês	Produção
Jan.	210
Fev.	180
Mar.	205
Abr.	195
Mai.	205
Jun.	220
Jul.	185
Ago.	190
Set.	200
Out.	180
Nov.	205
Dez.	198

Medidas de Tendência Central

- **Média**

$$\bar{x} = \frac{210 + 180 + 205 + 195 + 205 + 220 + 185 + 190 + 200 + 180 + 205 + 198}{12}$$

$$\bar{x} = 197,75$$

Medidas de Tendência Central

180 < 180 < 185 < 190 < 195 < 198 < 200 < 205 < 205 < 205 < 210 < 220

1 2 3 4 5 6 7 8 9 10 11 12

$$Md(x) = \frac{198 + 200}{2} = 199$$

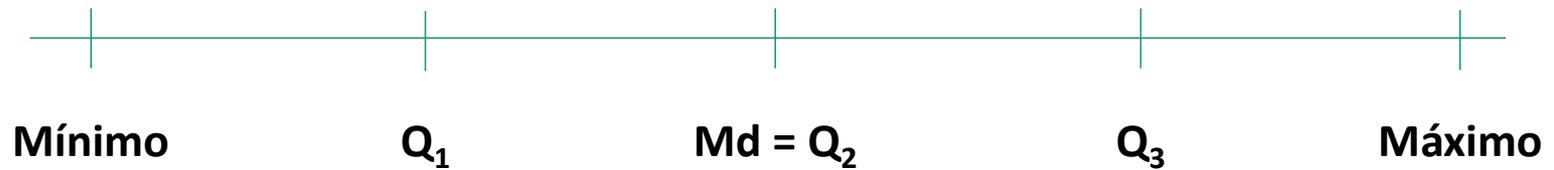
Mediana = 2º quartil, (Q_2), 50º percentil ou 5º decil

Medidas de Tendência Central

- Moda = 205 (repete três vezes)

Medidas de Tendência não Central

- **1º e 3º quartis** (Q_1 e Q_3), dividem um conjunto de dados, dispostos em ordem crescente, em quatro partes de dimensões iguais



Medidas de Tendência não Central

- **Decis**, dividem um conjunto de dados, dispostos em ordem crescente, em dez partes de dimensões iguais
- **Percentis**, dividem um conjunto de dados, dispostos em ordem crescente, em cem partes de dimensões iguais

Medidas de Tendência não Central

A partir do exemplo anterior, vamos calcular o p_{10} (1º decil), p_{25} (1º quartil) e o p_{75} (3º quartil)

- Posição do 10º percentil = $[(n - 1) \cdot (p/100)] + 1 = [(12 - 1) \cdot (10/100)] + 1 = 2,1$

$$p_{10} = (0,9 \times 180) + (0,1 \times 185) = 180,5$$

Medidas de Tendência não Central

A partir do exemplo anterior, vamos calcular o p_{10} (1º decil), p_{25} (1º quartil) e o p_{75} (3º quartil)

- Posição do 25º percentil = $[(n - 1) \cdot (p/100)] + 1 = [(12 - 1) \cdot (25/100)] + 1 = 3,75$

$$p_{25} = (0,75 \times 190) + (0,25 \times 185) = 188,75$$

Medidas de Tendência não Central

A partir do exemplo anterior, vamos calcular o p_{10} (1º decil), p_{25} (1º quartil) e o p_{75} (3º quartil)

- Posição do 75º percentil = $[(n - 1) \cdot (p/100)] + 1 = [(12 - 1) \cdot (75/100)] + 1 = 9,25$

$$p_{75} = (0,75 \times 205) + (0,25 \times 205) = 205$$

Existência de *Outliers*

- O valor x^* é considerado um possível outlier quando:

$$X^* < Q_1 - 1,5.(Q_3 - Q_1)$$

$$X^* > Q_3 + 1,5.(Q_3 - Q_1)$$

- $(Q_3 - Q_1) \rightarrow$ intervalo interquartil

Existência de *Outliers*

- O valor x^* é considerado um provável outlier quando:

$$X^* < Q_1 - 3.(Q_3 - Q_1)$$

$$X^* > Q_3 + 3.(Q_3 - Q_1)$$

- Uso do Box-Plot

Medidas de Dispersão ou Variabilidade

- **Amplitude**, diferença entre o maior e o menor valor do conjunto de observações
- **Variância**, dispersão dos dados em torno da média

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Para Amostra

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

Para População

Medidas de Dispersão ou Variabilidade

- **Desvio padrão**, raiz quadrada da variância. Pode ser entendido como a média das distâncias das observações em relação à média geral da variável

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

Para Amostra

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

Para População

Medidas de Dispersão ou Variabilidade

- **Coeficiente de variação (CV)**, mede a homogeneidade dos dados em relação à média, como medida de risco relativo

$$\left(\frac{S}{\bar{x}} \times 100 \right) \%$$

Para Amostra

$$\left(\frac{\sigma}{\mu} \times 100 \right) \%$$

Para População

Se for acima de 30%, o conjunto de dados poderá ser considerado heterogêneo

Medidas de Dispersão ou Variabilidade

- A partir do exemplo anterior, vamos calcular a amplitude, variância, desvio padrão e coeficientes de variação

- Amplitude = $220 - 180 = 40$

Medidas de Dispersão ou Variabilidade

- A partir do exemplo anterior, vamos calcular a amplitude, variância, desvio padrão e coeficientes de variação

- Variância = $s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} = \frac{(210-197,75)^2 + \dots + (198-197,75)^2}{11} = 151,66$

Medidas de Dispersão ou Variabilidade

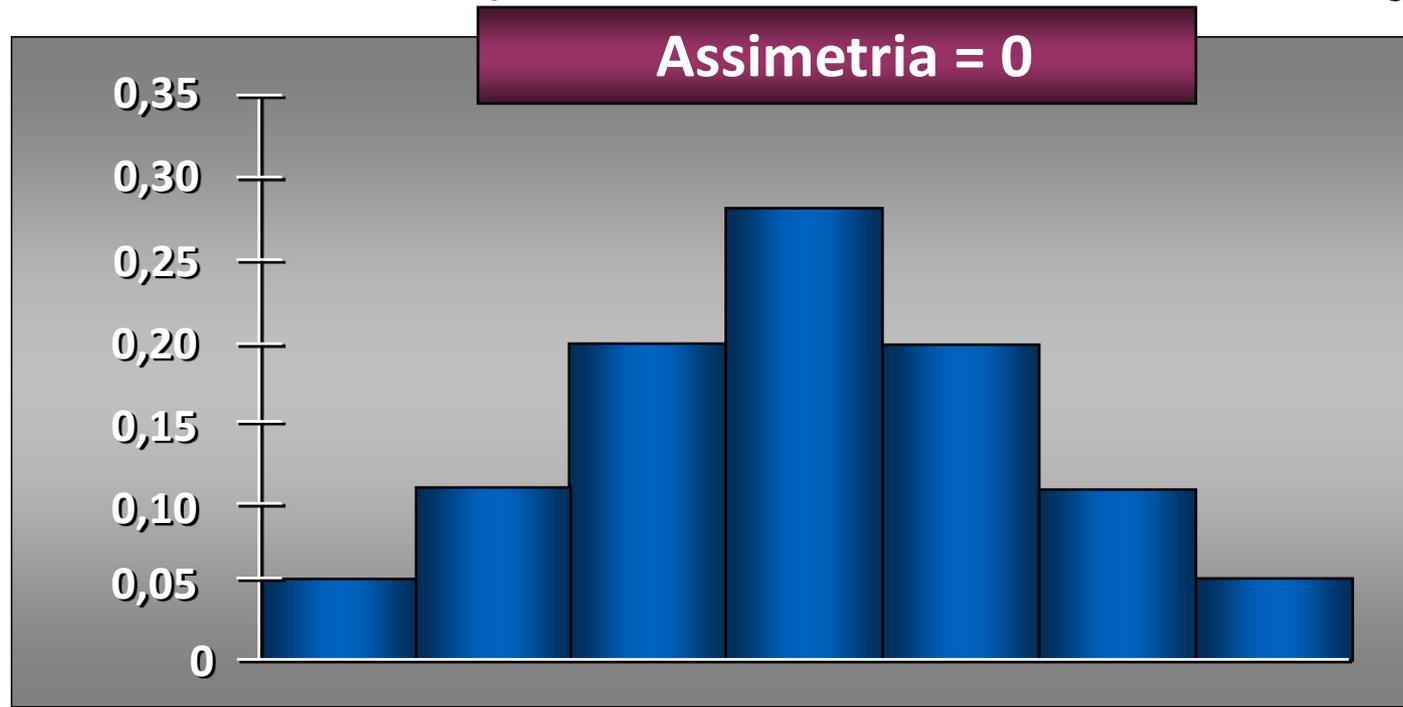
- A partir do exemplo anterior, vamos calcular a amplitude, variância, desvio padrão e coeficientes de variação
- Desvio padrão = $s = \sqrt{s^2} = \sqrt{151,66} = 12,31$

Medidas de Dispersão ou Variabilidade

- A partir do exemplo anterior, vamos calcular a amplitude, variância, desvio padrão e coeficientes de variação
- Coeficiente de Variação = $\frac{12,31}{197,75} \times 100 = 6,23\%$

Medidas de Assimetria e Curtose

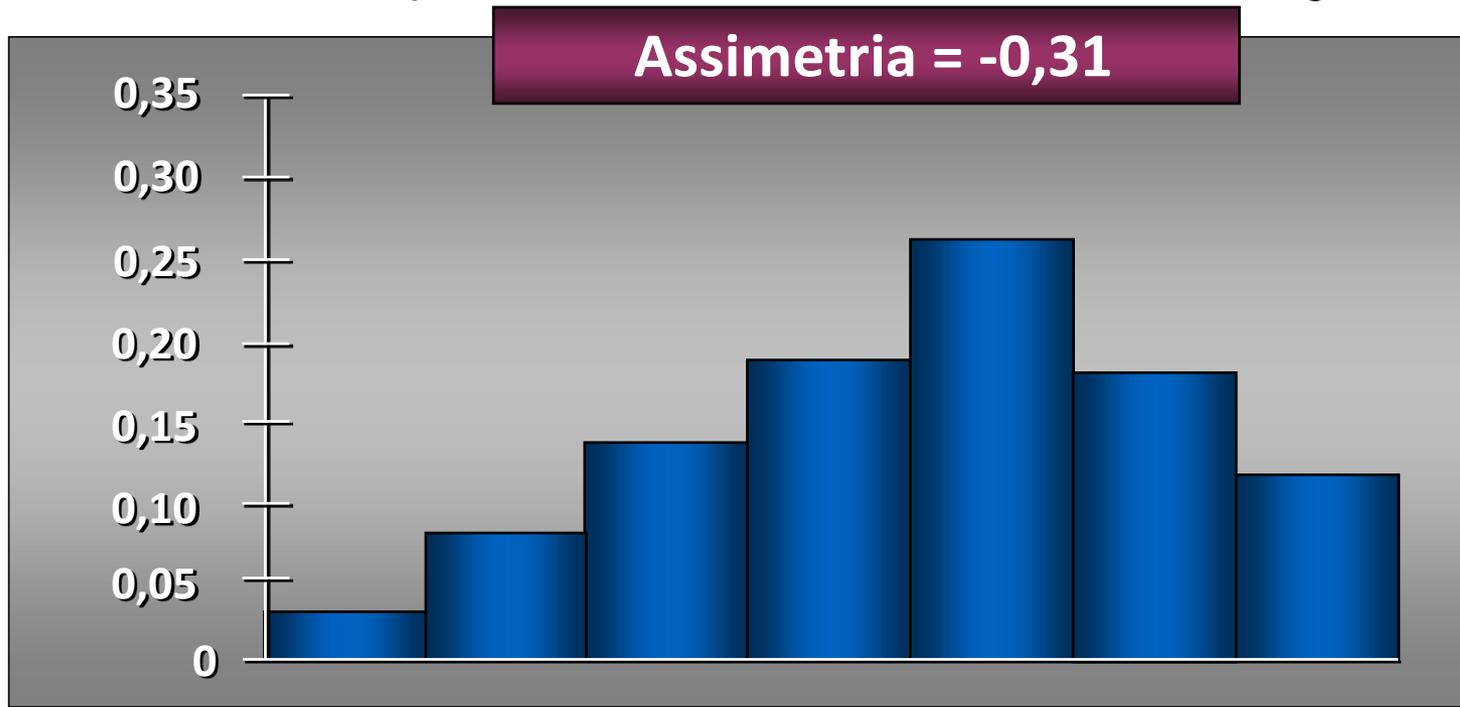
- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição



Média e mediana são iguais

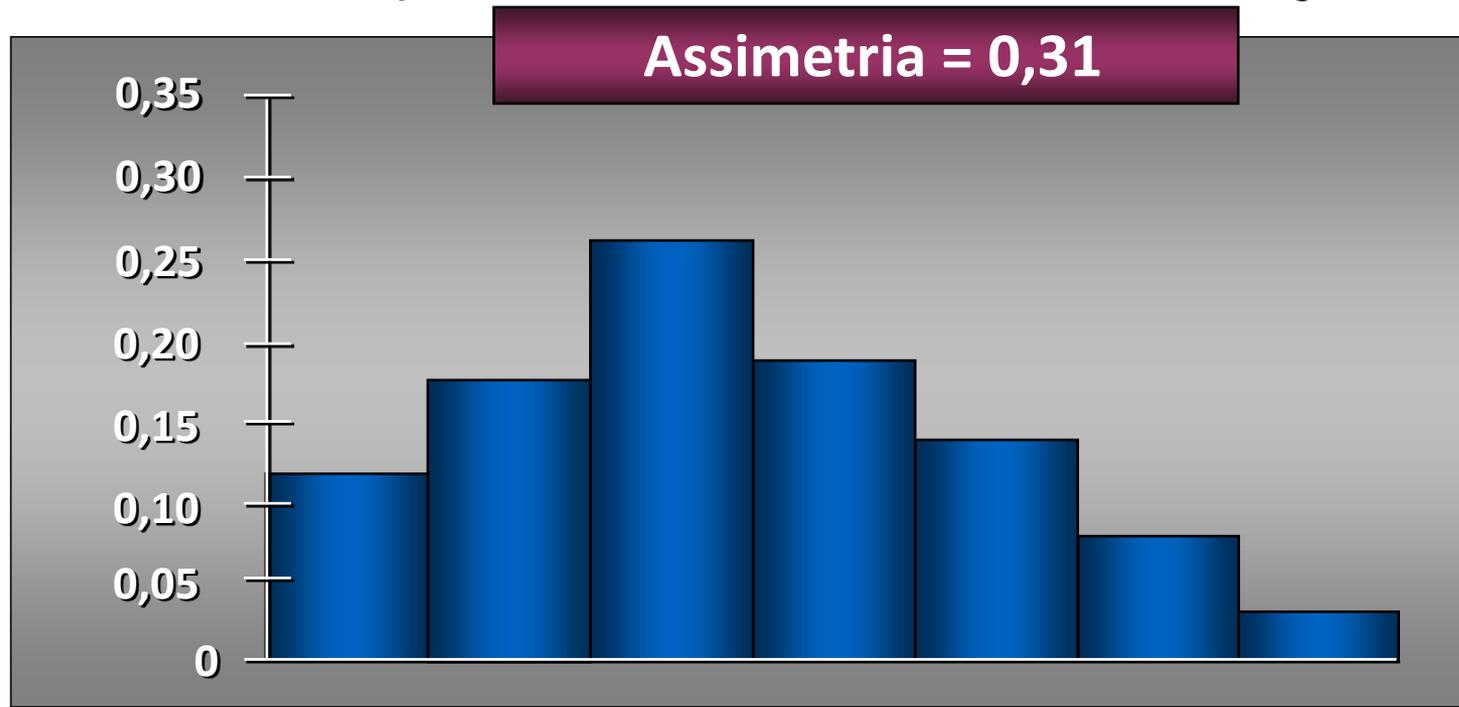
Medidas de Assimetria e Curtose

- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição



Medidas de Assimetria e Curtose

- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição



Medidas de Assimetria e Curtose

- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição
- Diferença entre a média e a moda, ponderada por uma medida de dispersão (desvio padrão), conhecida como 1º coeficiente de Pearson

$$\textit{Assimetria} = \frac{\bar{x} - Mo}{S}$$

Medidas de Assimetria e Curtose

- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição
- Para evitar o uso da moda, adota-se a relação empírica entre mediana e moda: $\bar{x} - Mo = 3(\bar{x} - Md)$, que corresponde ao 2º coeficiente de Pearson

$$Assimetria = \frac{3(\bar{x} - Md)}{S}$$

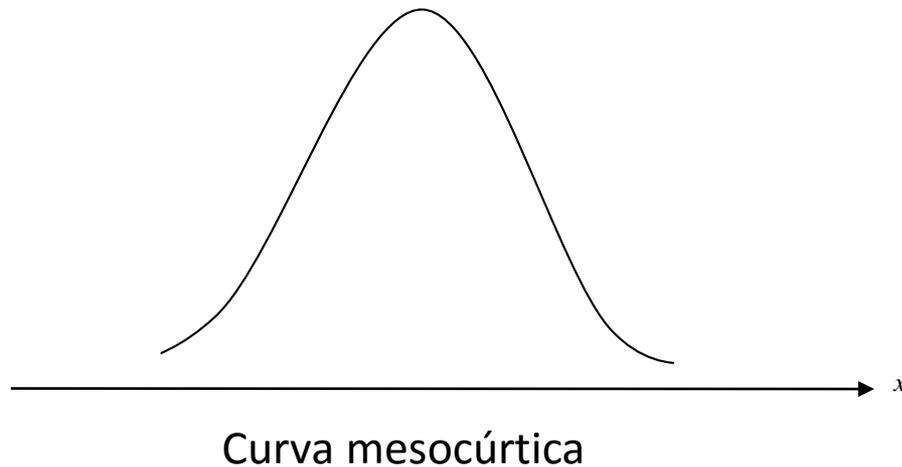
Medidas de Assimetria e Curtose

- **Assimetria**, representa o grau de desvio (ou achatamento) da simetria de uma distribuição
- Outra medida de assimetria é o coeficiente de assimetria de Fisher (g_1), calculado a partir do terceiro momento em torno da média (M_3)

$$g_1 = \frac{n^2 M_3}{(n-1)(n-2)S^3}, \text{ em que } M_3 = \frac{\sum (x_i - \bar{x})^3}{n}$$

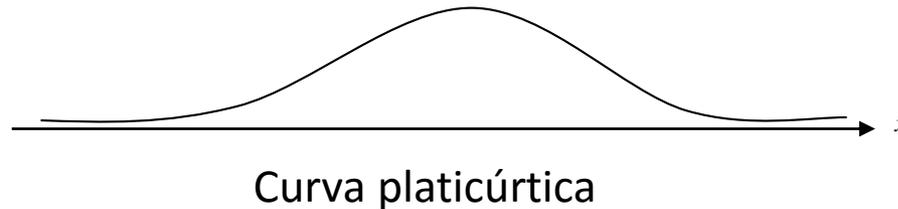
Medidas de Assimetria e Curtose

- **Curtose**, representa a altura do ponto máximo da curva de distribuição



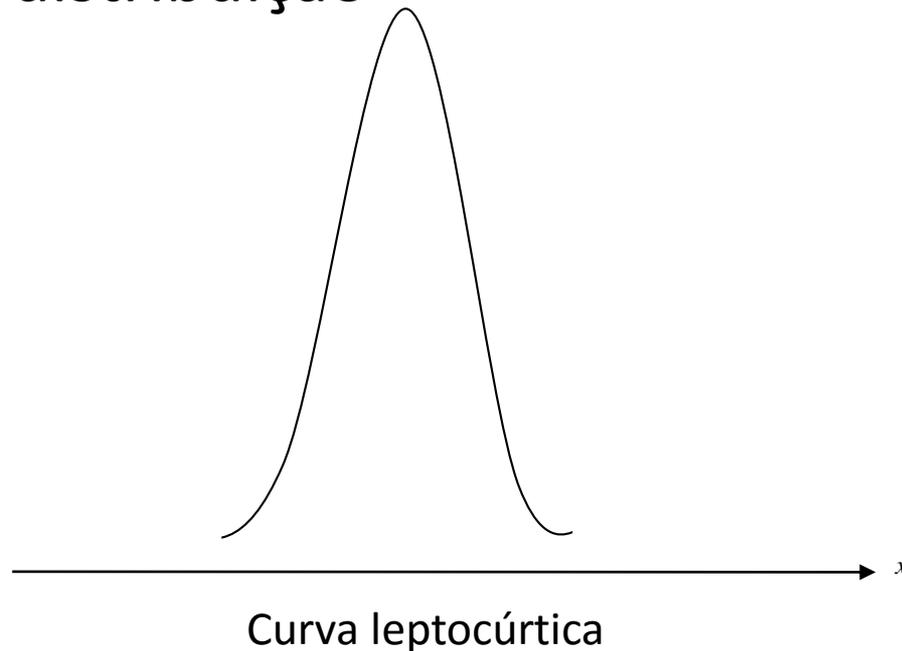
Medidas de Assimetria e Curtose

- **Curtose**, representa a altura do ponto máximo da curva de distribuição



Medidas de Assimetria e Curtose

- **Curtose**, representa a altura do ponto máximo da curva de distribuição



Medidas de Assimetria e Curtose

- **Curtose**, representa a altura do ponto máximo da curva de distribuição
- Coeficiente percentílico de achatamento ou curtose

$$k = \frac{(Q_3 - Q_1)}{2 \cdot (p_{90} - p_{10})}$$

Se $k = 0,263$, diz-se que a curva é mesocúrtica. Para $k > 0,263$ → platicúrtica. Para $k < 0,263$ → leptocúrtica

Medidas de Assimetria e Curtose

- **Curtose**, representa a altura do ponto máximo da curva de distribuição
- Coeficiente de achatamento de Fisher (g_2), calculado a partir do terceiro momento em torno da média (M_4)

$$g_2 = \frac{n^2(n+1)M_4}{(n-1)(n-2)(n-3)S^4} - 3 \times \frac{(n-1)^2}{(n-2)(n-3)}$$

$$\text{em que } M_4 = \frac{\sum(x_i - \bar{x})^4}{n}$$

Para a curva com distribuição normal $g_2 = 0$. Negativa \rightarrow achatada; Positiva \rightarrow

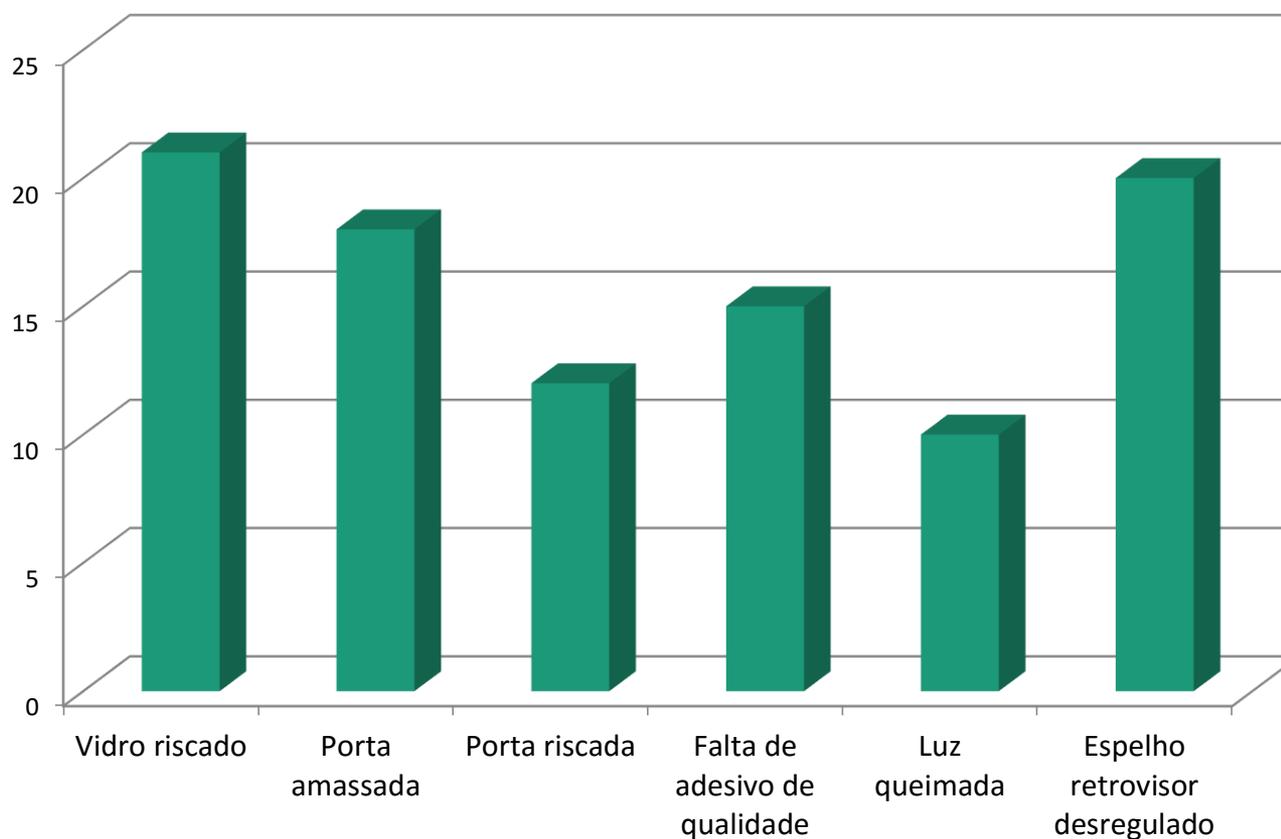
Representação Gráfica de Resultados

- Gráfico de Barras

Não-conformidade	Frequência	Percentual
Vidro riscado	21	21,88%
Porta amassada	18	18,75%
Porta riscada	12	12,50%
Falta de adesivo de qualidade	15	15,53%
Luz queimada	10	10,42%
Espelho retrovisor desregulado	20	20,83%
Total	96	100%

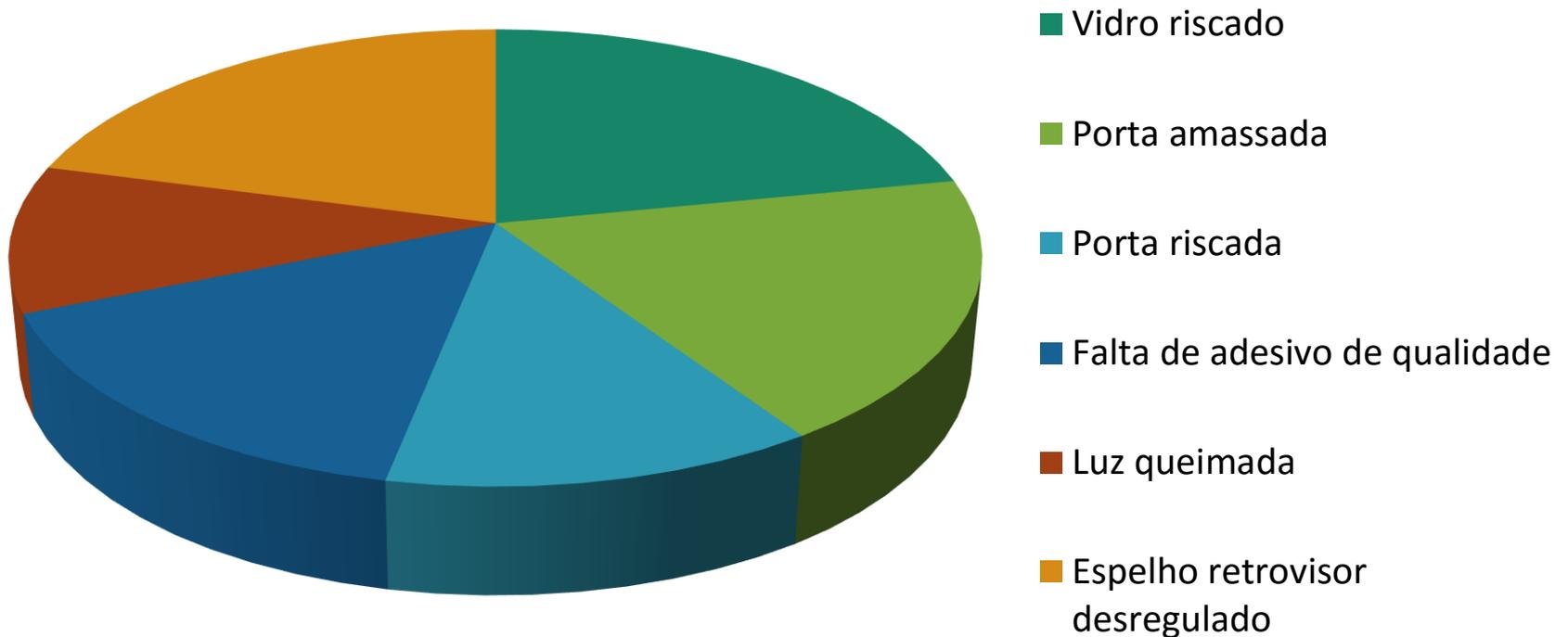
Representação Gráfica de Resultados

- Gráfico de Barras



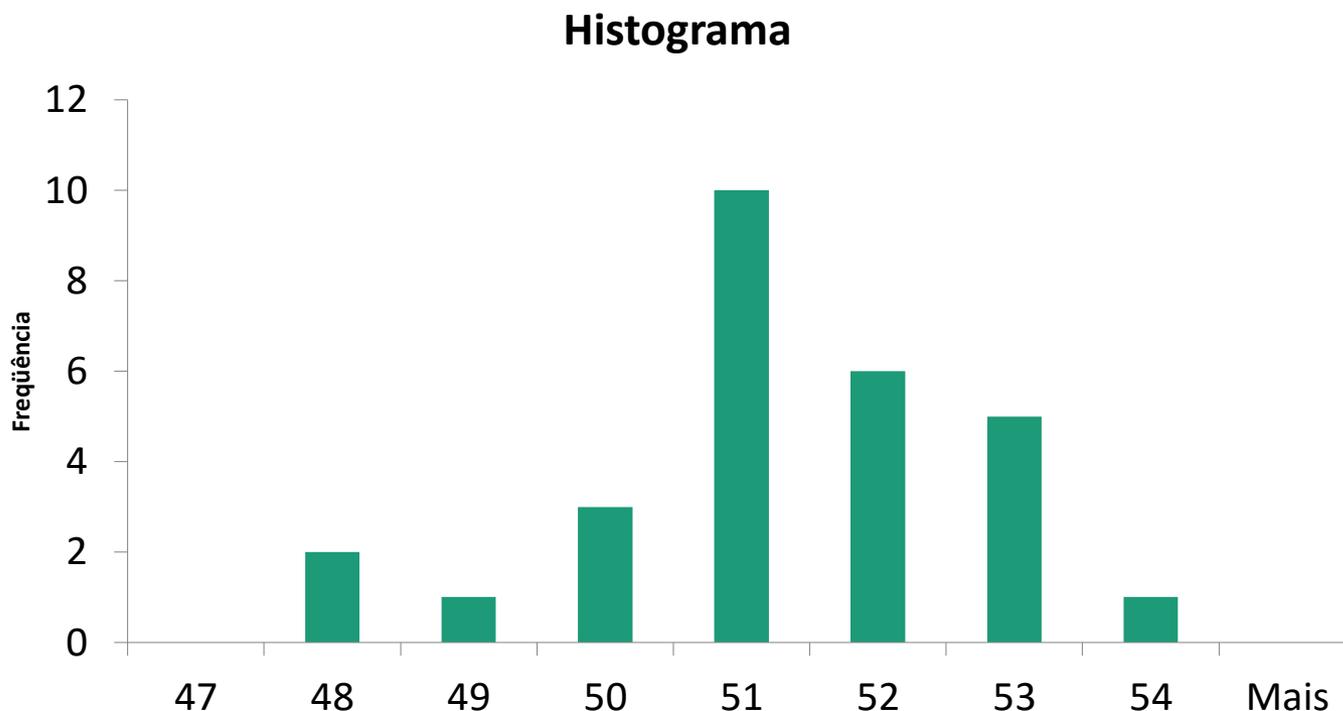
Representação Gráfica de Resultados

- Diagrama Circular (Gráfico de Pizza)



Representação Gráfica de Resultados

- Histograma



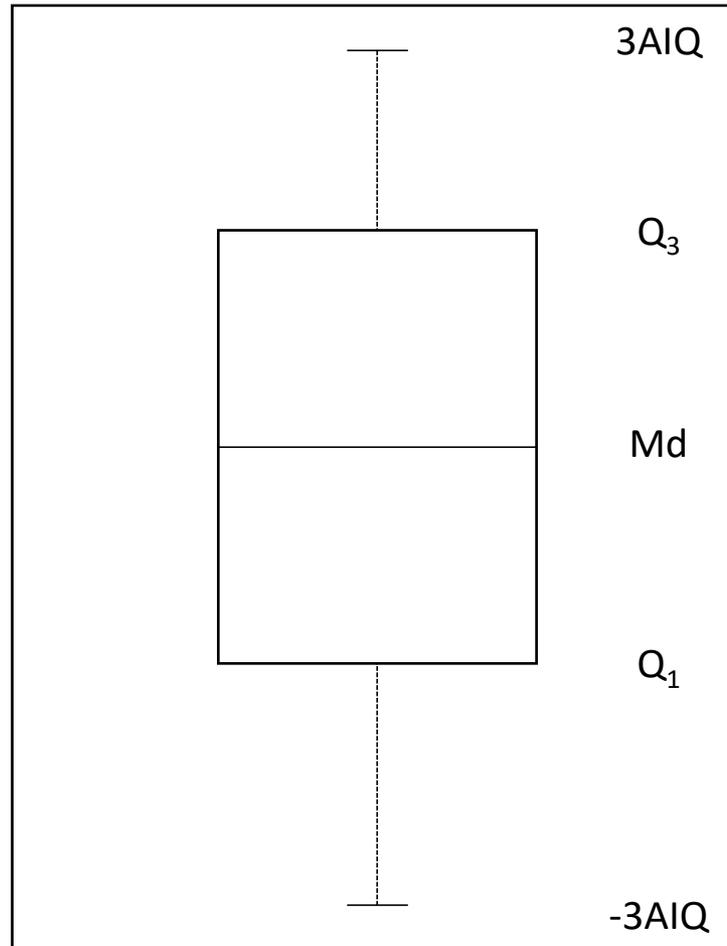
Representação Gráfica de Resultados

- Gráfico ramo-e-folhas

47	9									
48	0	5								
49	5	7								
50	0	3	4	5	5	5	6	6	7	
51	0	0	1	1	3	4	4	6		
52	1	2	4	5	8					
53	2									

Representação Gráfica de Resultados

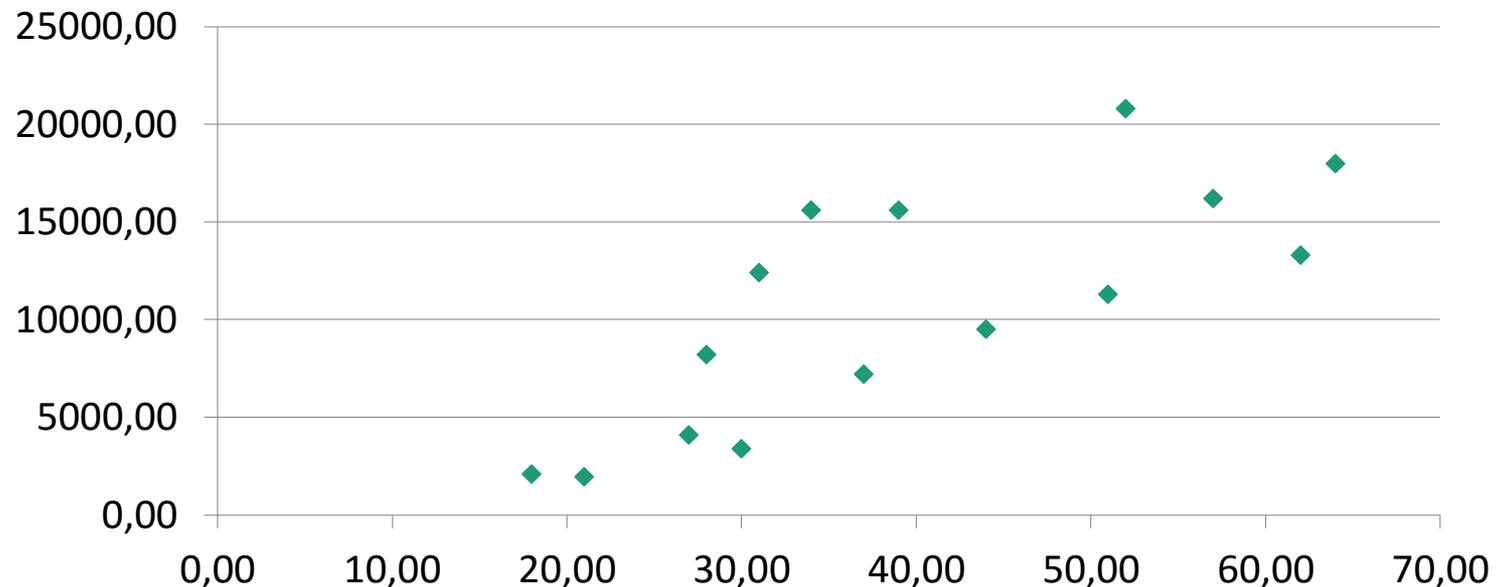
- Box-Plot



Representação Gráfica de Resultados

- Diagrama de dispersão (*scatterplot*)

Gráfico de Dispersão



Obrigado pela Atenção!!!

Até a próxima aula

mbotelho@usp.br

www.marcelobotelho.com