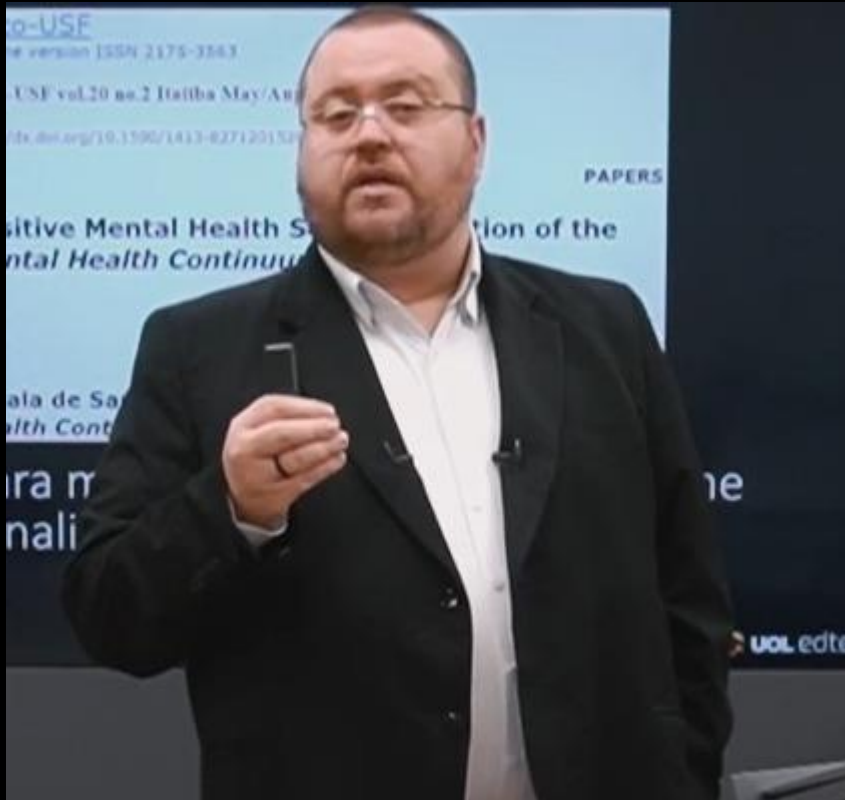


Pesquisa quantitativa

Prof. Dr. Wagner de Lara
Machado





Wagner de Lara Machado

<http://lattes.cnpq.br/5116682938541081>

<https://scholar.google.com/citations?user=fH6qCDoAAAAJ&hl=endo>

Pesquisa quantitativa

- Utiliza números para organizar a informação
 - Estatística
- Abordagens exploratórias e confirmatórias
 - Abdução
 - Hipotético-dedutivo



Desenhos de pesquisa

- Experimentos e quasi-experimentos
- Estudos transversais ou correlacionais
- Estudos longitudinais (séries temporais, painéis)
- Estudos de caso ($n=1$)
- ...

Dois tipos de estatística

- Estatística descritiva
 - Organizar, resumir e comunicar informação
- Estatística inferencial
 - Fazer inferências sobre uma população a partir de uma amostra



Depression Anxiety and Stress Scale (DASS-21) – Short Form:
Adaptação e Validação para Adolescentes Brasileiros

Naiana Dapieve Patias – Universidade Federal do Rio Grande do Sul e Faculdade Meridional (IMED), Passo Fundo, Brasil

Wagner De Lara Machado – Pontifícia Universidade Católica de Campinas, Campinas, Brasil

Denise Ruschel Bandeira – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Débora Dalbosco Dell’Aglio – Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil

Tabela 2

Médias e Desvios Padrão no EDAE-A por Sexo e por Faixa Etária

			Média	Dp	Teste <i>t</i>	<i>gl</i>	<i>p</i>
Depressão	Sexo	Meninos (n=162)	3,16	3,64	-4,38	402	<0,001
		Meninas (n=264)	4,95	4,73			
	Faixa etária	12 a 15 anos (n=256)	4,31	4,69	0,26	424	0,79
		16 a 18 anos (n=170)	4,20	4,01			
Ansiedade	Sexo	Meninos (n=162)	2,29	2,62	-3,63	415	<0,001
		Meninas (n=264)	3,40	3,72			
	Faixa etária	12 a 15 anos (n=256)	2,93	3,43	-0,37	424	0,71
		16 a 18 anos (n=170)	3,06	3,33			
Estresse	Sexo	Meninos (n=162)	4,32	3,53	-5,22	412	<0,001
		Meninas (n=264)	6,46	4,89			
	Faixa etária	12 a 15 anos (n=256)	5,60	4,64	-0,25	424	0,89
		16 a 18 anos (n=170)	5,71	4,39			
EDAE-A	Sexo	Meninos (n=162)	9,77	7,84	-5,37	419,7	<0,001
		Meninas (n=264)	14,82	11,54			
	Faixa etária	12 a 15 anos (n=256)	12,85	11,19	-0,12	424	0,90
		16 a 18 anos (n=170)	12,97	9,58			

Nota. EDAE-A: Escala de Depressão, Ansiedade e Estresse para adolescentes

Tabela 2

Médias e Desvios Padrão no EDAE-A por Sexo e por Faixa Etária

			Média	Dp	Teste <i>t</i>	<i>gl</i>	<i>p</i>
Depressão	Sexo	Meninos (n=162)	3,16	3,64	-4,38	402	<0,001
		Meninas (n=264)	4,95	4,73			
	Faixa etária	12 a 15 anos (n=256)	4,31	4,69	0,26	424	0,79
		16 a 18 anos (n=170)	4,20	4,01			
Ansiedade	Sexo	Meninos (n=162)	2,29	2,62	-3,63	415	<0,001
		Meninas (n=264)	3,40	3,72			
	Faixa etária	12 a 15 anos (n=256)	2,93	3,43	-0,37	424	0,71
		16 a 18 anos (n=170)	3,06	3,33			
Estresse	Sexo	Meninos (n=162)	4,32	3,53	-5,22	412	<0,001
		Meninas (n=264)	6,46	4,89			
	Faixa etária	12 a 15 anos (n=256)	5,60	4,64	-0,25	424	0,89
		16 a 18 anos (n=170)	5,71	4,39			
EDAE-A	Sexo	Meninos (n=162)	9,77	7,84	-5,37	419,7	<0,001
		Meninas (n=264)	14,82	11,54			
	Faixa etária	12 a 15 anos (n=256)	12,85	11,19	-0,12	424	0,90
		16 a 18 anos (n=170)	12,97	9,58			

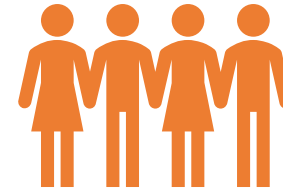
Nota. EDAE-A: Escala de Depressão, Ansiedade e Estresse para adolescentes

Populações e amostras



População

Todos os casos de um determinado grupo com uma ou mais características específicas



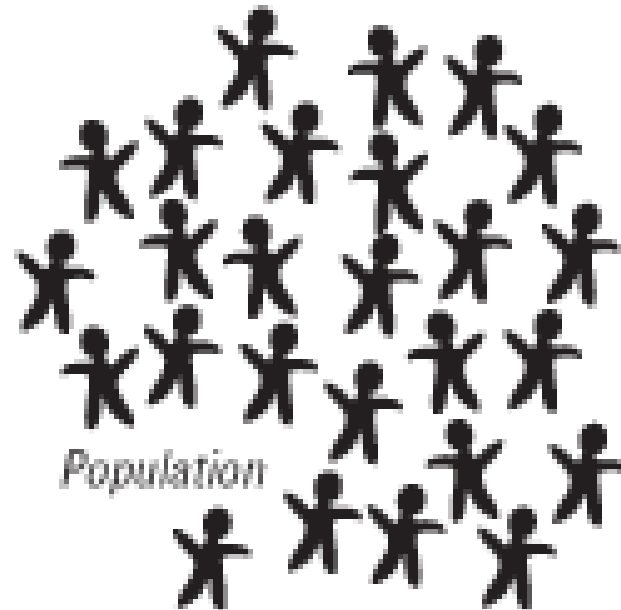
Amostra

Um conjunto de observações, uma porção da população

Podem ser probabilísticas ou não

Precisam ser grandes o suficiente para detectar um efeito

We want to know about these



Parameter

μ

(Population mean)

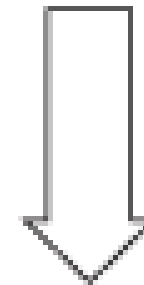
Random
selection



We have these to work with



Inference



\bar{X}

Statistic

(Sample mean)

Amostras e populações

- Por que usar amostras?
 - Custo
 - Eficiência
 - Precisão

Estatísticas



NÚMEROS



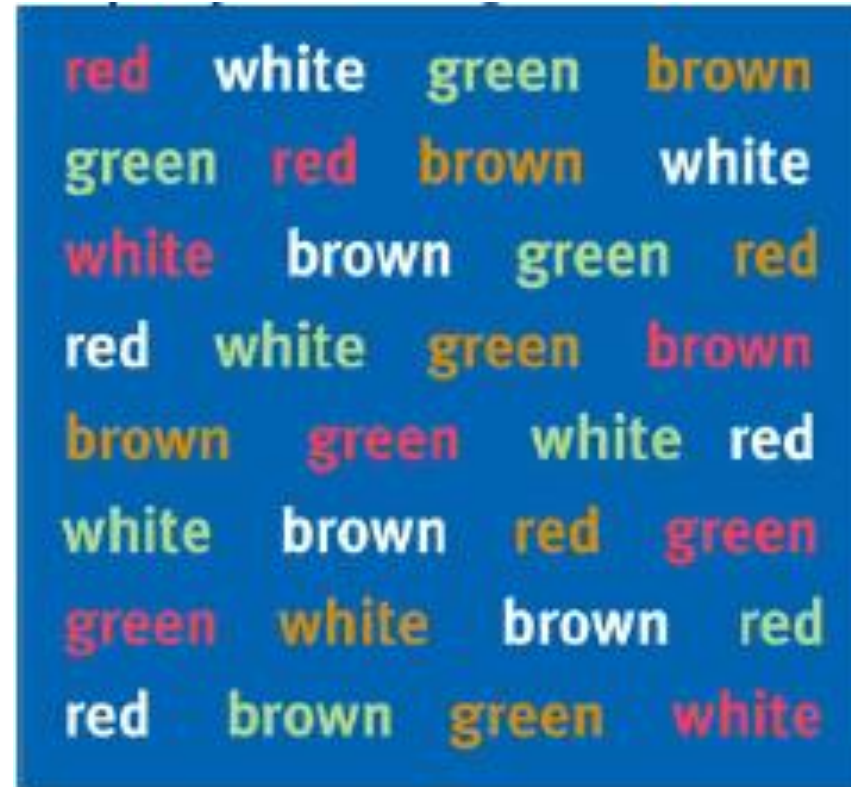
VARIÁVEIS (DIFERENTE
DE UMA CONSTANTE)



MEDIDAS
(PSICOMETRIA)

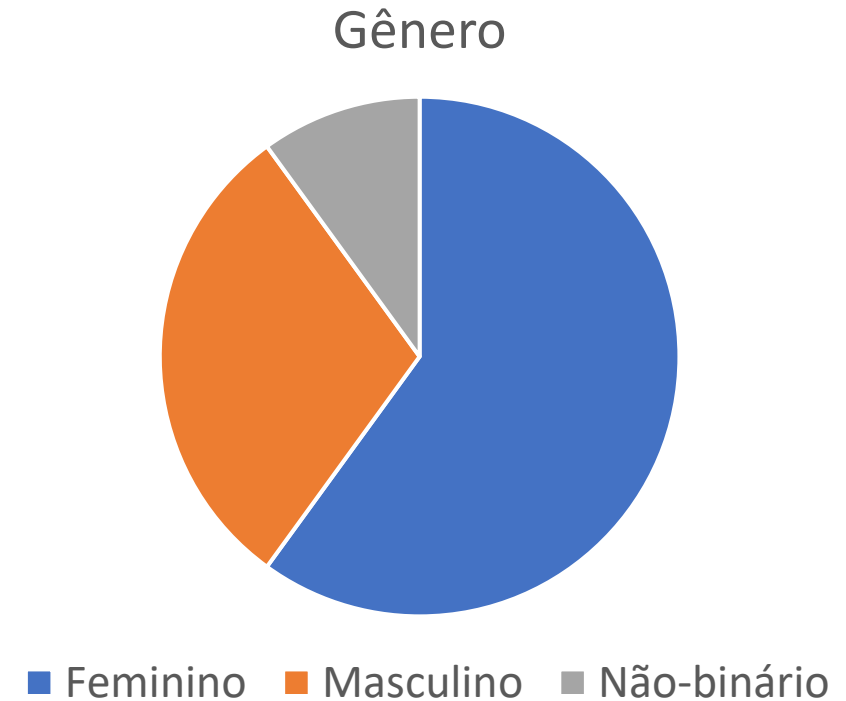
Variáveis

- Algo que muda de valor, de observação em observação, em uma amostra (população)
- Exemplo:



Variáveis discretas

- Identidade e ordenação
- Variáveis nominais
- Variáveis ordinais



Variáveis contínuas ou escalares

- A distância entre dois valores é conhecida
- Variáveis intervalares
- Variáveis de razão



- Número de erros em uma tarefa
- Número de itens recordados em uma tarefa de memória

?

Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	2	3	4	5

Estatística inferencial

Teste de hipótese



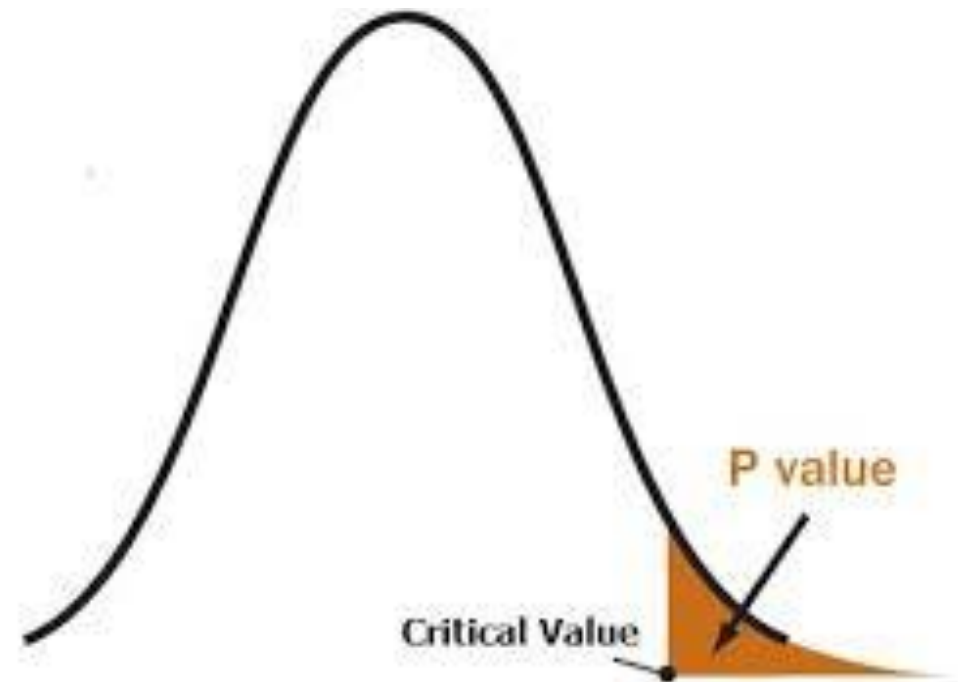
- Hipóteses são conclusões sobre o relacionamento entre variáveis
- Mas não só isso, elas também descrevem o processo generativo dos dados

Acessando variáveis

- Definição operacional de variáveis
 - A satisfação com o trabalho dos times de inovação será medida em uma escala tipo Likert
 - 1 (totalmente insatisfeito) a 5 (totalmente satisfeito)
 - A performance dos times de inovação será medida pelo número de soluções/produtos registrados ou pré-registrados
 - Geralmente requer uma medida válida, fidedigna, padronizada e normatizada

Estatística e probabilidade

- Teorias e distribuições de probabilidades para:
 - Comparação de resultados
 - Tomada de decisão

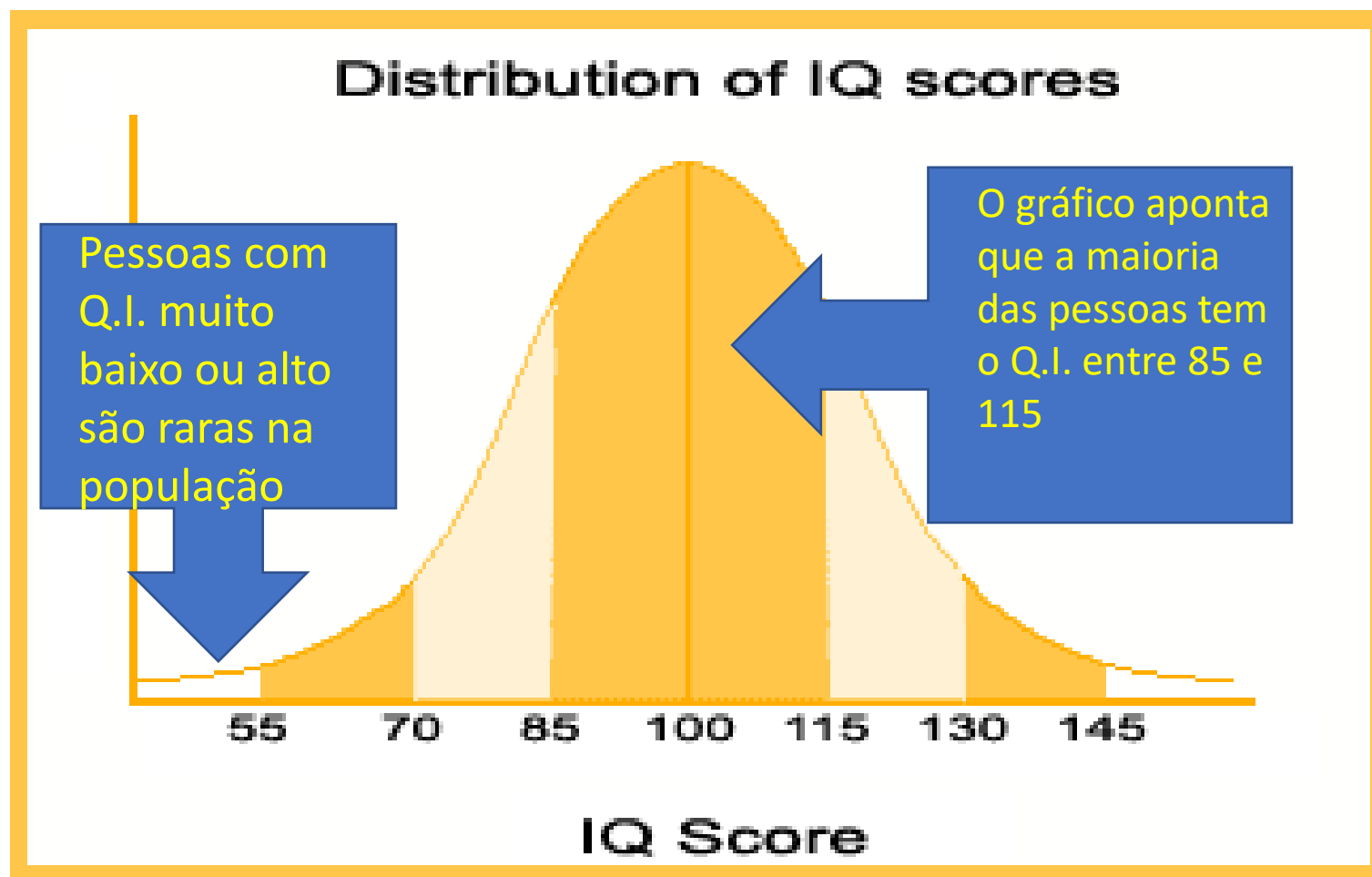


Probabilidade

- Probabilidade é a razão entre a ocorrência de um fenômeno e a possibilidades de ocorrência:
- Moeda de “cara ou coroa” (x):
 - 10 lançamentos
 - 6 caras
 - $p(x = \text{cara}) = 6/10 = 0,6$ ou 60%



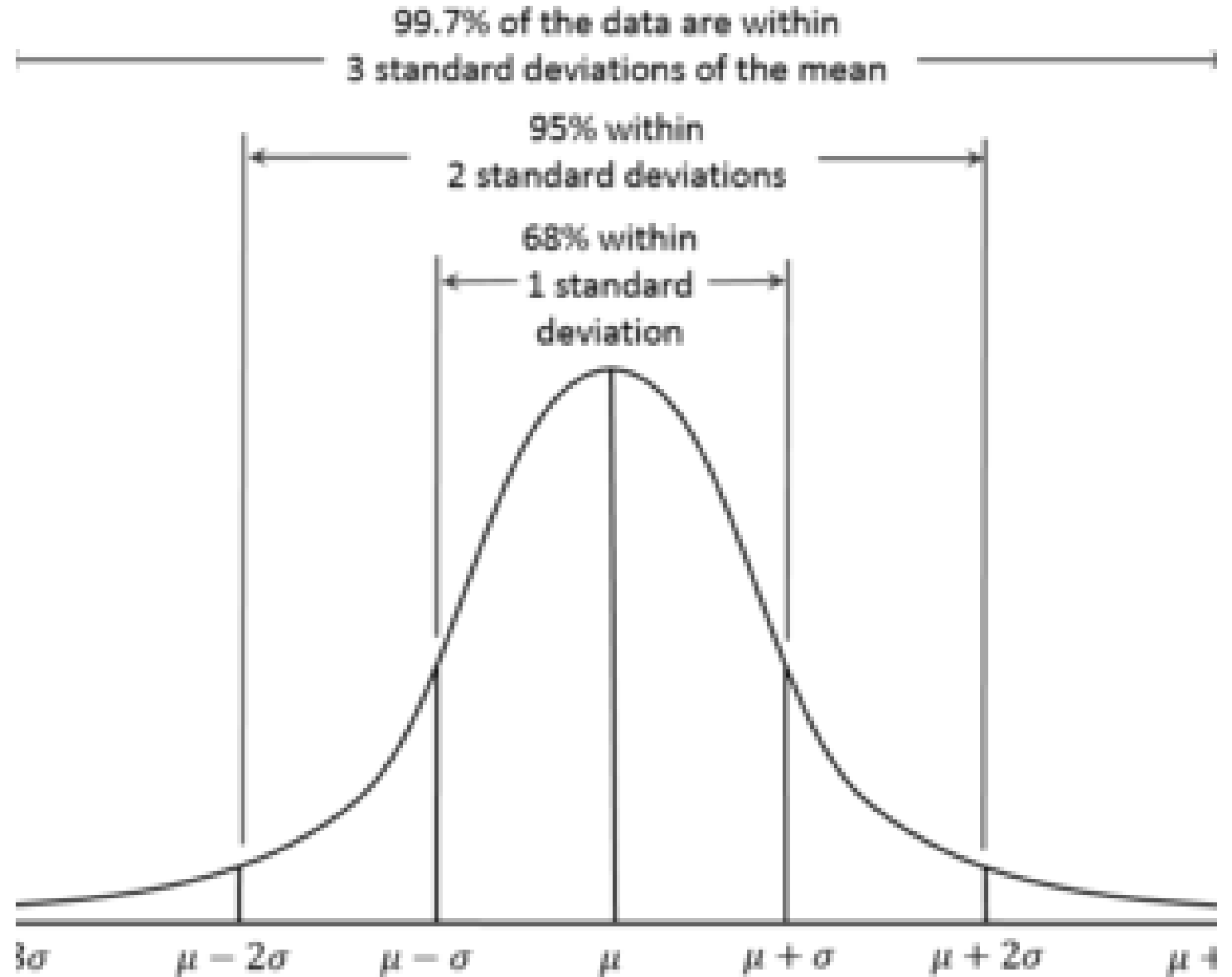
Exemplo de distribuição: Q.I.
 $\mu = 100$ e $\sigma = 15$



Probabilidades e a
distribuição normal
(Gauss ou
distribuição z)

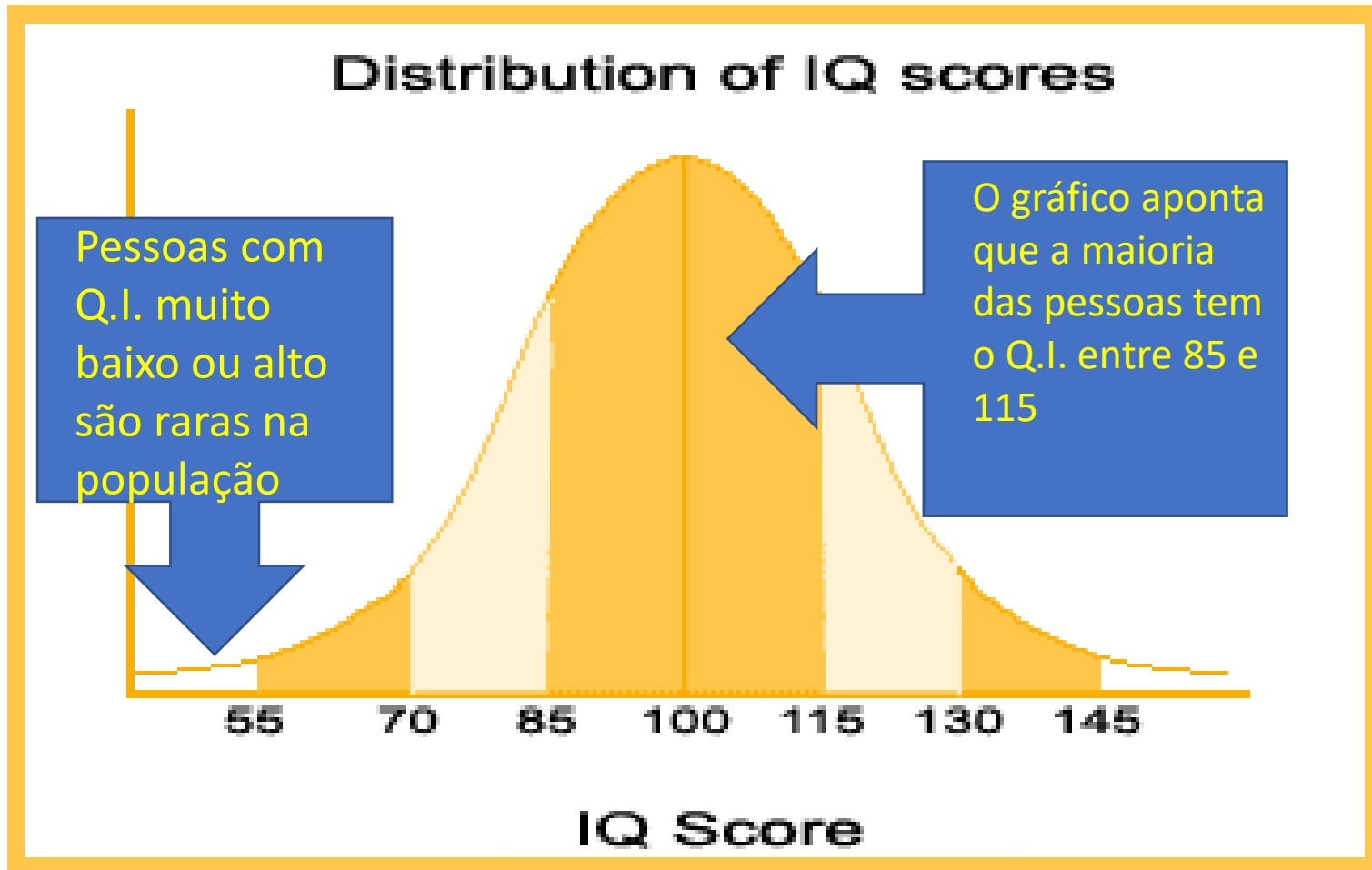
Escores
padronizados ou
“escores z”:

$$\frac{x - \mu}{\sigma}$$



Exemplo de distribuição: Q.I.

$\mu = 100$ e $\sigma = 15$



$$x_j = 120$$

$$z = \frac{120 - 100}{15}$$

$$z = 1,33$$

SE vivêssemos em um mundo justo...

- Não haveria diferenças de salários entre mulheres e homens
- Não haveria diferença na proporção de cargos superiores entre
 - Mulheres e homens
 - Negros/pardos e brancos

Teste de hipótese estatística

- Em estatística é comum as pessoas “inverterem” o entendimento sobre o teste de hipótese
- Quando realizamos um teste, em geral, gostaríamos de observar “efeitos”. Estes efeitos podem ser diferenças entre grupos, correlações, etc.
- Exemplo de hipóteses (efeitos):
 - Na pandemia, as empresas que aumentaram o número de anúncios digitais tiveram maiores vendas. (predição ou explicação)
 - Quem está com maiores demandas no trabalho durante a pandemia se sente mais ansioso, estressado e consome mais álcool. (associação)

Teste de hipótese estatística

- A questão é que em estatística nós testamos a “hipótese nula” ou H_0 , ao invés da hipótese alternativa ($H_1 \dots H_n$)
- Exemplo de hipóteses (efeitos):
 - H_1 - Na pandemia, as empresas que aumentaram o número de anúncios digitais tiveram maiores vendas. (predição ou explicação)
 - H_0 - Não há diferenças em vendas entre as empresas que aumentaram ou não o número de anúncios digitais.
 - H_1 - Quem está com maiores demandas no trabalho durante a pandemia se sente mais ansioso, estressado e consome mais álcool
 - H_0 - Não há associação entre o nível de demandas no trabalho, ansiedade, estresse e o consumo de álcool durante a pandemia

Teste de hipótese estatística

- Logo, não testamos se estamos “corretos” com a nossa hipótese alternativa, mas o quão errados estamos (qual probabilidade, “p valor”) em assumir a hipótese nula como verdadeira
- Assim, temos duas opções ou escolhas a fazer:
 - Não rejeitar a H_0
 - Rejeitar H_0
- Em geral, utiliza-se para essa decisão o ponto de corte de $\alpha = 5\%$, ou $p \leq 0,05$ ou ainda, probabilidade de 5 em cada 100



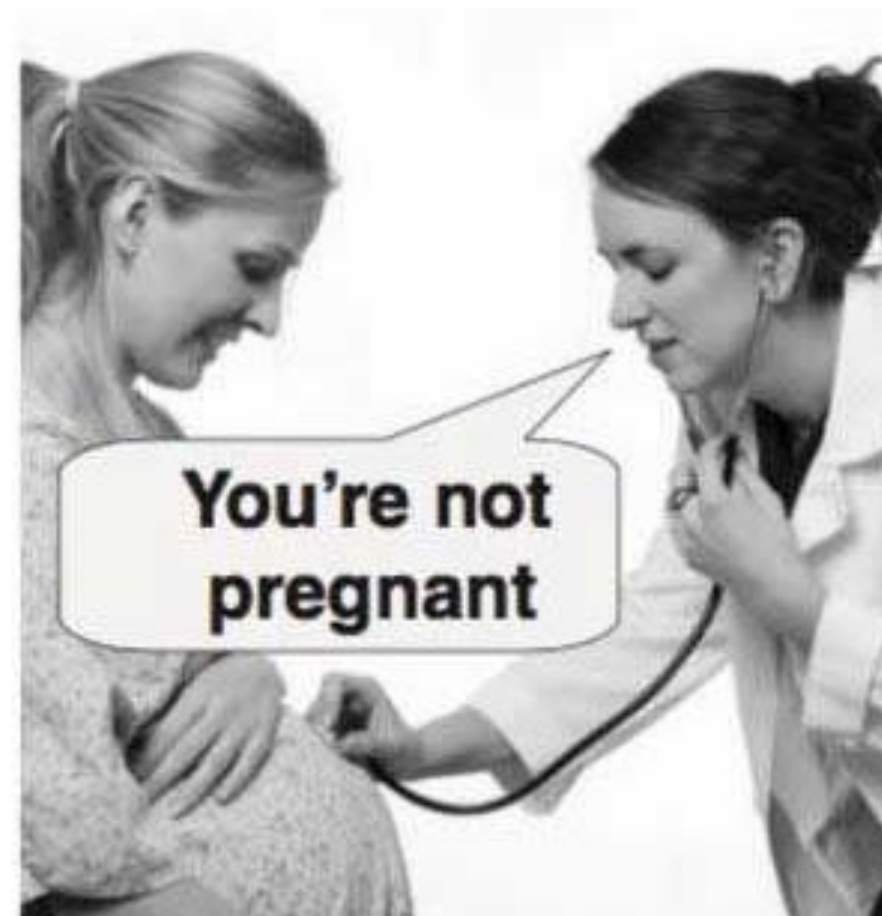
Teste de hipótese estatística

- Nossa probabilidade é condicional, isto é, qual a probabilidade de observar o efeito x caso a hipótese nula seja verdadeira
- Assim, tem-se que um valor de $p = 0.03$ indica a probabilidade de 3% em encontrar um efeito, sendo a hipótese nula verdadeira
- Com essa probabilidade, nós não rejeitamos ou rejeitamos a hipótese nula (H_0)?

Teste de hipótese estatística

	<i>Conclusão do teste</i>	
Verdade	Não se rejeita H0	Rejeita-se H0
H0 verdadeira	Decisão certa Probabilidade: $1 - \alpha$	Decisão errada Probabilidade: α ERRO TIPO I
H0 falsa	Decisão errada Probabilidade: β ERRO TIPO II	Decisão certa Probabilidade: $1 - \beta$ (Poder do teste)

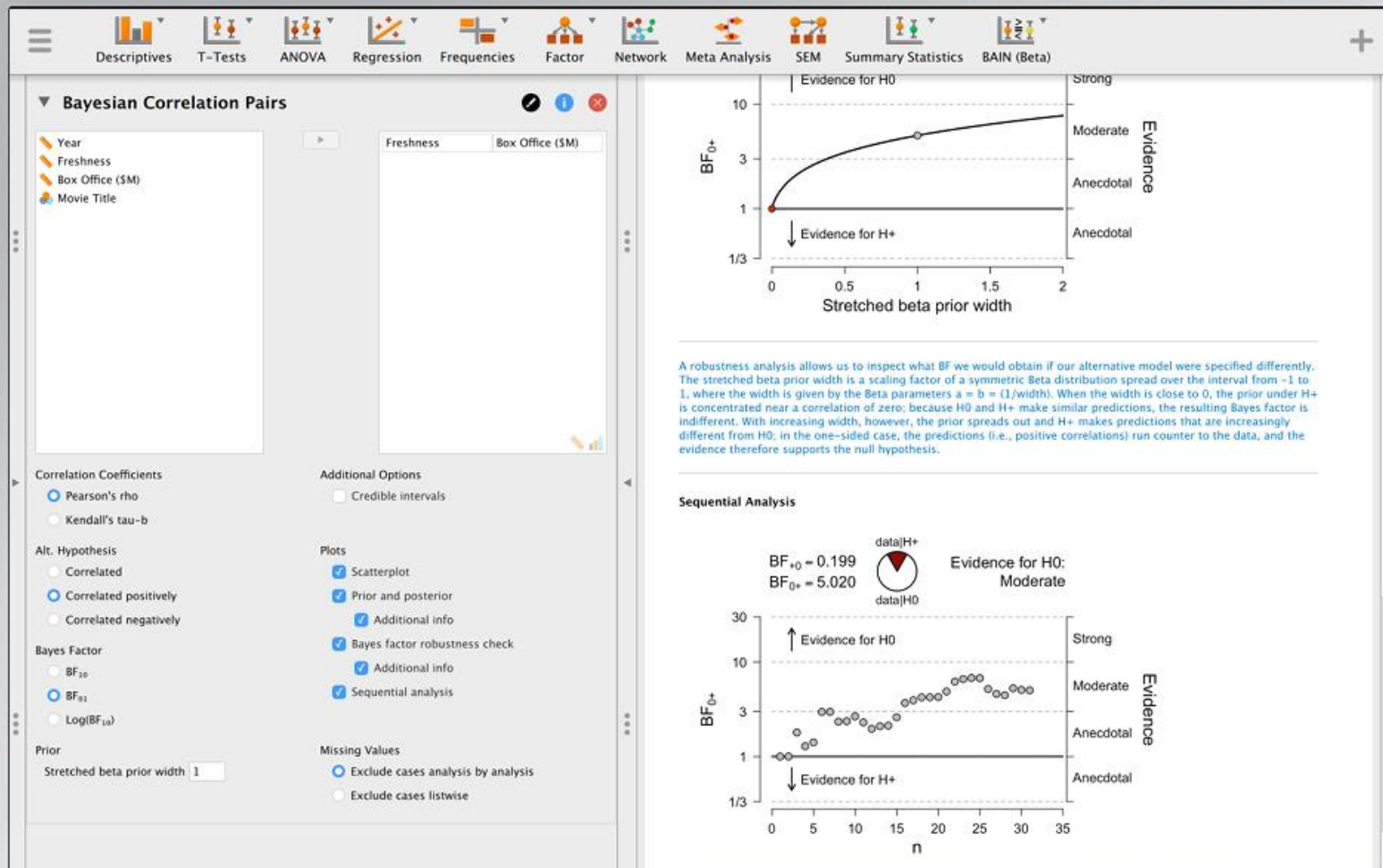
Erros do tipo I e II

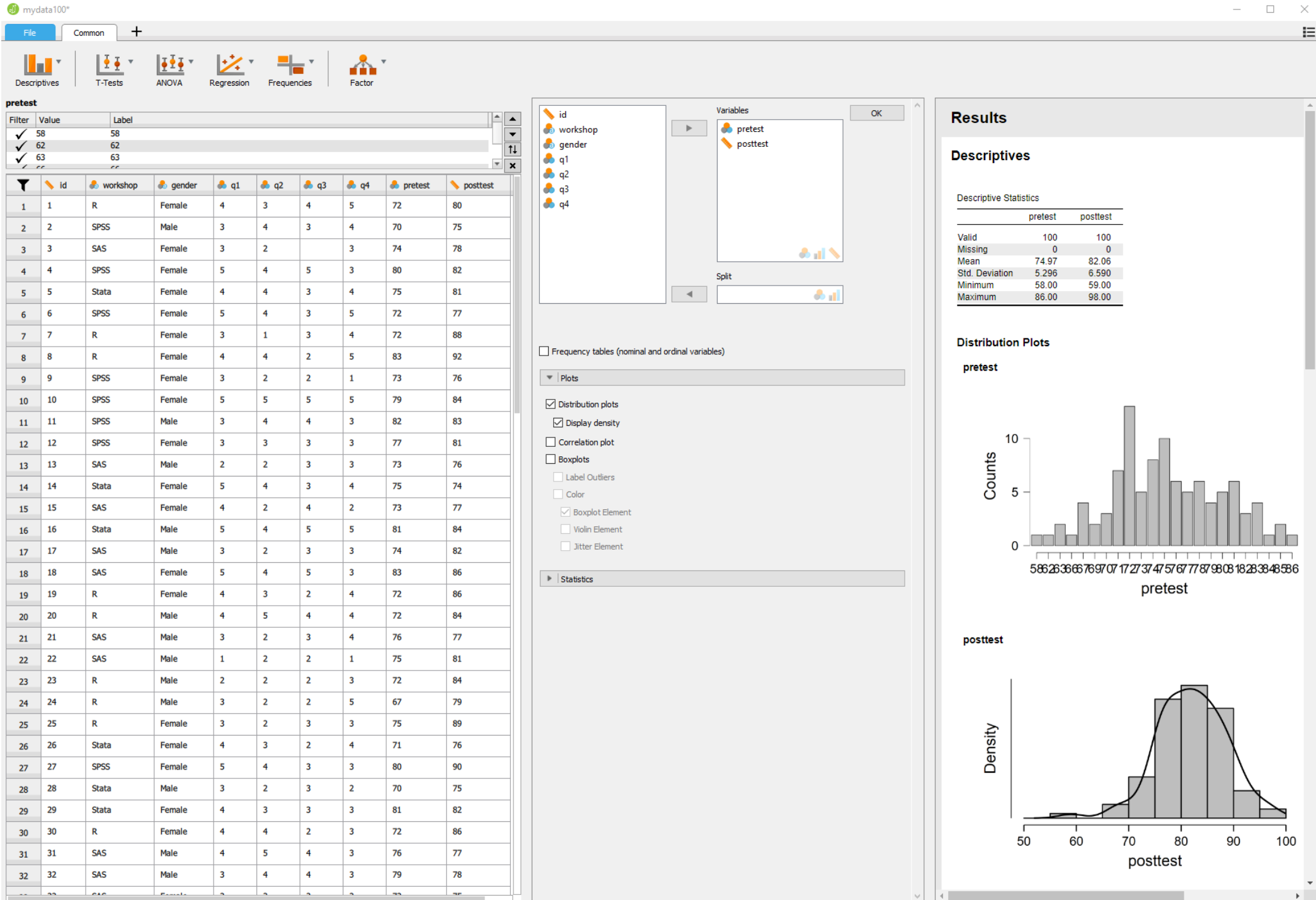


JASP!



- Free, Friendly and Flexible
- Links:
- <https://jasp-stats.org/>
- <https://www.youtube.com/watch?v=HxqB7CUA-XI>
- Download:
- <https://jasp-stats.org/download/>







Descriptives

T-Tests

ANOVA

Regression

Frequencies

Factor

Machine Learning



K-Means Clustering



Species



Variables

Sepal.Length
Sepal.Width
Petal.Length
Petal.Width

Tables

- ☒ Cluster means
- ☐ Cluster information
 - ☒ Within sum of squares
 - ☐ Silhouette score
 - ☐ Centroids
 - ☐ Between sum of squares
 - ☐ Total sum of squares
- ☐ Evaluation metrics

Plots

- ☐ Elbow method
- ☐ Cluster means
 - ☒ Display barplot
 - ☒ Group into one figure
- ☒ Cluster densities
- ☒ t-SNE cluster plot
 - ☒ Legend
 - ☐ Labels

Training Parameters

- ☐ Add predicted clusters to data

Results

K-Means Clustering

K-Means Clustering

Clusters	N	R ²	AIC	BIC	Silhouette
5	150	0.848655	130.200000	190.410000	0.340000

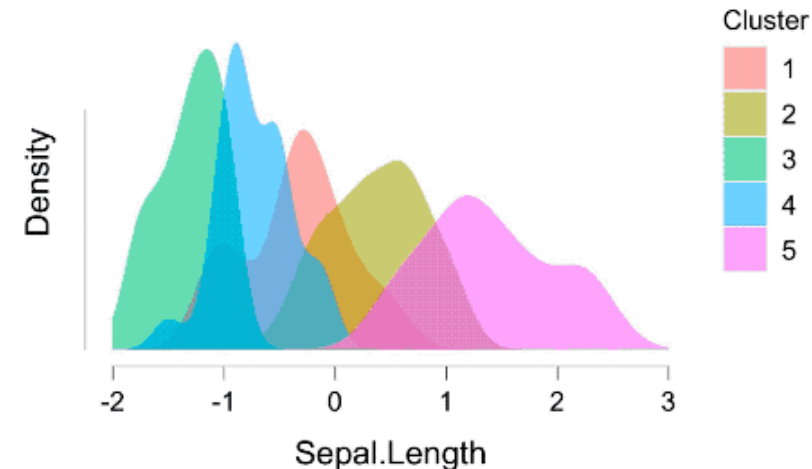
Note. The model is optimized with respect to the *BIC* value.

Cluster Means

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Cluster 1	-0.351614	-1.328555	0.102606	0.012283
Cluster 2	0.380404	-0.389645	0.606791	0.563910
Cluster 3	-1.303439	0.198838	-1.304029	-1.284814
Cluster 4	-0.718944	1.501990	-1.297231	-1.216593
Cluster 5	1.392665	0.232382	1.156745	1.213276

Cluster Density Plots

Sepal.Length

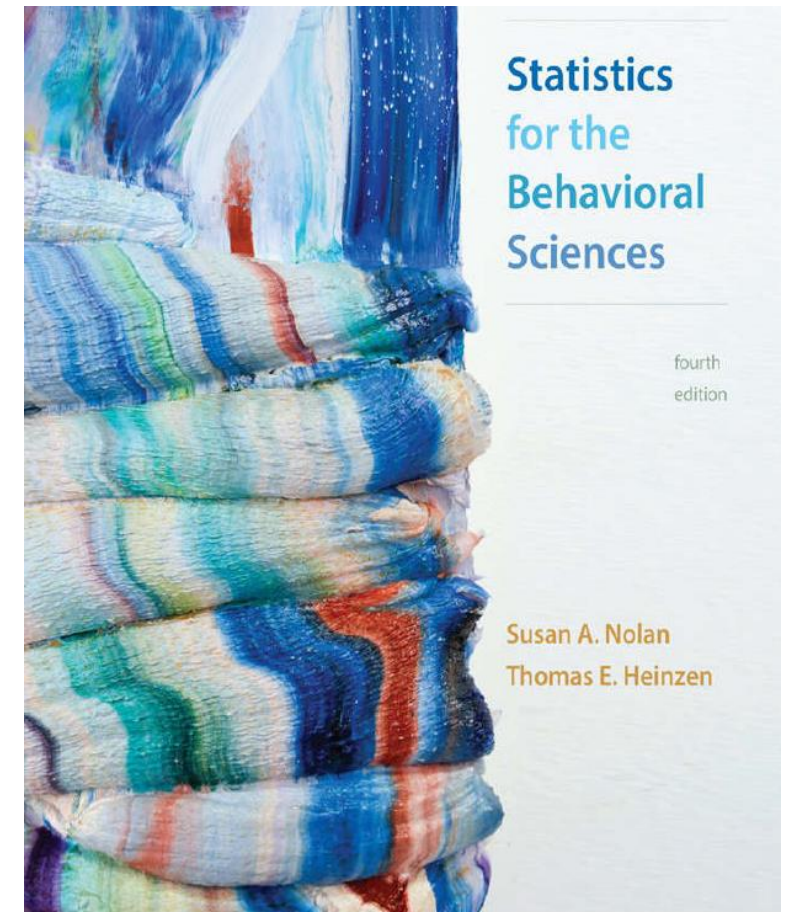
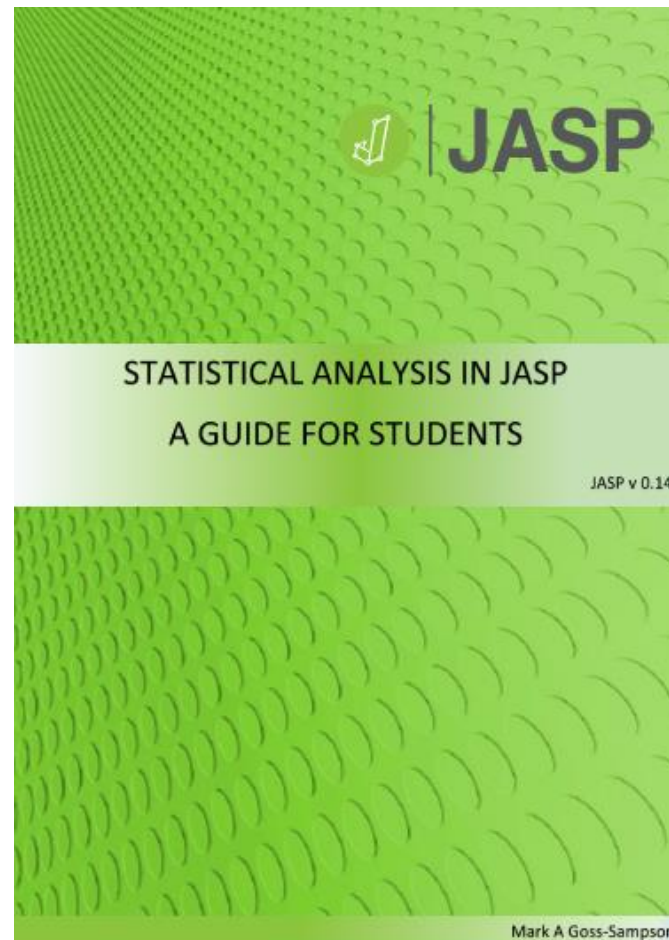
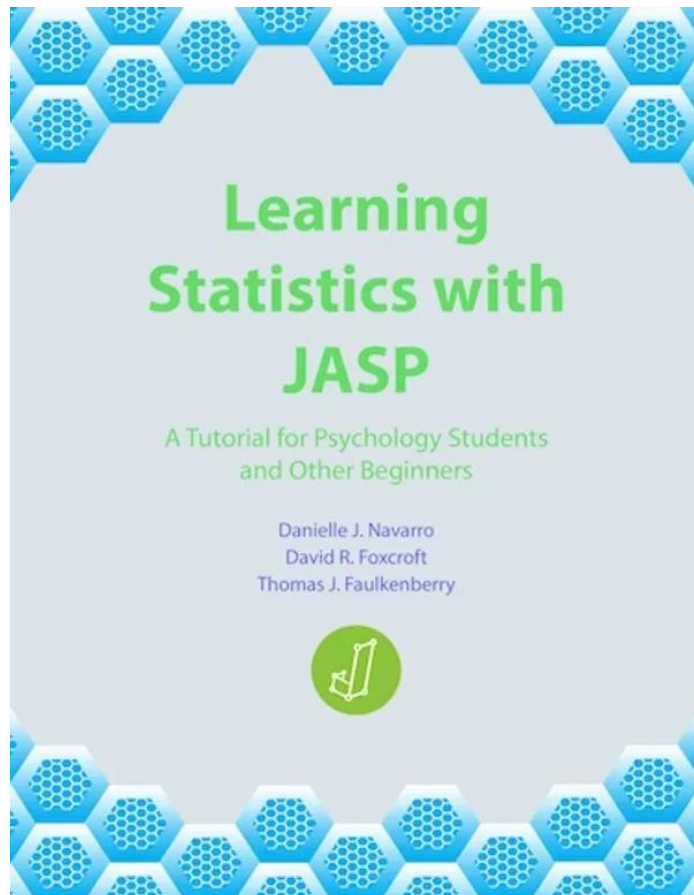


Sepal.Width

Livros da disciplina

Com JASP

<https://www.dropbox.com/sh/vfq0e8i1ereo8fm/AACZXVHmY62JX6wfiQwiGxLaa?dl=0>



Sugestão de leitura

