

IAA005 - ESTATÍSTICA APLICADA I

Parte 1

Prof. Arno P. Schmitz

UFPR – Universidade Federal do Paraná

APRESENTAÇÃO

ARNO P. SCHMITZ

Bacharelado em Ciências Econômicas – UFPR

Mestrado em Economia – UFBA

Doutorado em Desenvolvimento Econômico - UFPR

EMENTA

Introdução à Estatística: descritiva e inferencial. Estudo de populações e amostras. Introdução às séries estatísticas. Análise de gráficos. Estudo das medidas de posição e de dispersão. Análise de correlação e regressão linear. Introdução aos testes de hipótese com uma amostra.

BIBLIOGRAFIA

COSTA, H. V. V. da. **Introdução ao R. Recife:** UFPE. 2017.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica.** 5 ed. Porto Alegre: AMGH. 2011.

KAZMIER. L. J. **Estatística aplicada à economia e administração.** São Paulo: Pearson Makron Books. 1982.

NAVIDI, W. **Probabilidade e estatística para ciências exatas.** Porto Alegre: AMGH. 2012.

NEDER, H. D. **Amostragem em pesquisas socioeconômicas.** Campinas: Alínea. 2008.

SILVA, B. F. da; DINIZ, J.; BORTOLUZZI, M. A. **Minicurso de estatística básica: introdução ao software R.** Santa Maria: UFSM. 2009.

SOUZA, E. F. M. de; PETERNELLI, L. A.; MELLO, M. P. de. **Software livre R: aplicação estatística.** João Pessoa: UFPR. 2019.

CONTEÚDOS

1. Conceitos básicos
2. Apresentação tabular e gráfica
3. Medidas de posição e dispersão
4. Séries Estatísticas
5. Correlação e regressão

ANÁLISE EXPLORATÓRIA DE DADOS

6. População e Amostra
7. Modelos probabilísticos
8. Modelos estatísticos de inferência

INFERÊNCIA ESTATÍSTICA

CONCEITOS BÁSICOS

O QUE É ESTATÍSTICA

Estatística é uma ciência que estuda e aplica métodos e técnicas que sistematicamente objetivam:

- a) Organizar;
- b) Descrever;
- c) Analisar e;
- d) Interpretar dados.

Estes dados são geralmente oriundos de estudos e pesquisas ou de experimentos realizados em quaisquer áreas do conhecimento.

IMPORTÂNCIA DA ESTATÍSTICA

- Interpretação de dados

Exemplos:

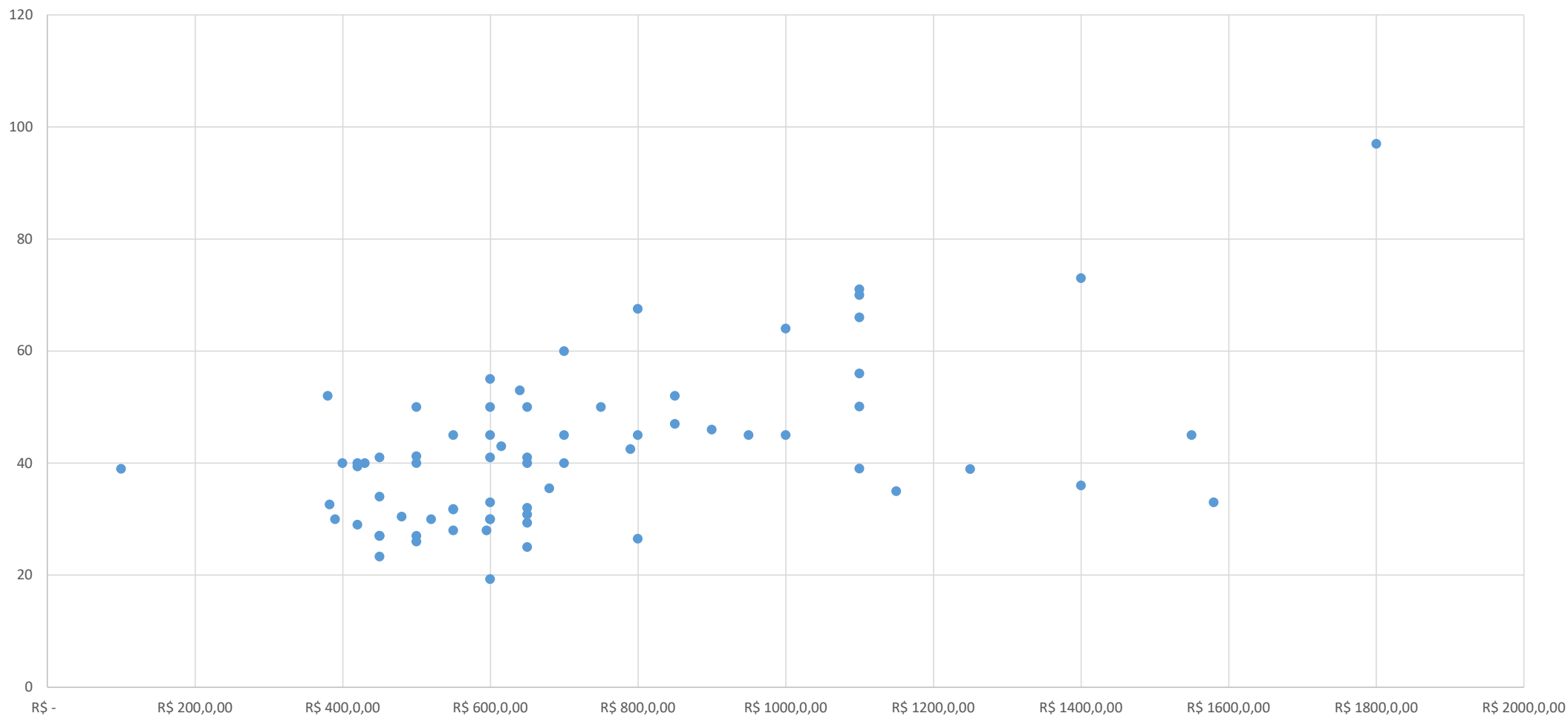
- ❖ Questões socioeconômicas e demográficas (renda, consumo, migração, etc.);
 - ❖ Questões de saúde e epidemiologia (doenças, tratamentos, etc.);
 - ❖ Questões climáticas (chuvas, matas, calor, etc.);
 - ❖ Agropecuária (produção, perdas, etc.)
 - ❖ Setor público (análise de impactos, projetos, etc.)
 - ❖ Política (pesquisas eleitorais e de opinião pública, etc.)
 - ❖ Propaganda (pesquisa de mercado, novos produtos, etc.)
-
- Vasta literatura sobre métodos e aplicações empíricas;
 - Aplicações científicas.

ECONOMIA - EXEMPLO

- Como é a variação espacial da renda média das famílias?
 - A renda média das famílias está distribuída aleatoriamente ou existe um padrão geográfico?
 - Como a renda média está relacionada com as características da residência (tamanho, número de moradores, etc.)?
 - Como as variações da renda média das famílias altera o padrão de consumo?

OUTROS EXEMPLOS

Preço do Aluguel (kitnetes) X Área do Imóvel (m²) – Centro de Curitiba/PR - 2016



Fatores que explicam o consumo de água

TABELA 10 - RESULTADOS DOS MODELOS DE REGRESSÃO PARA O CONSUMO DE ÁGUA

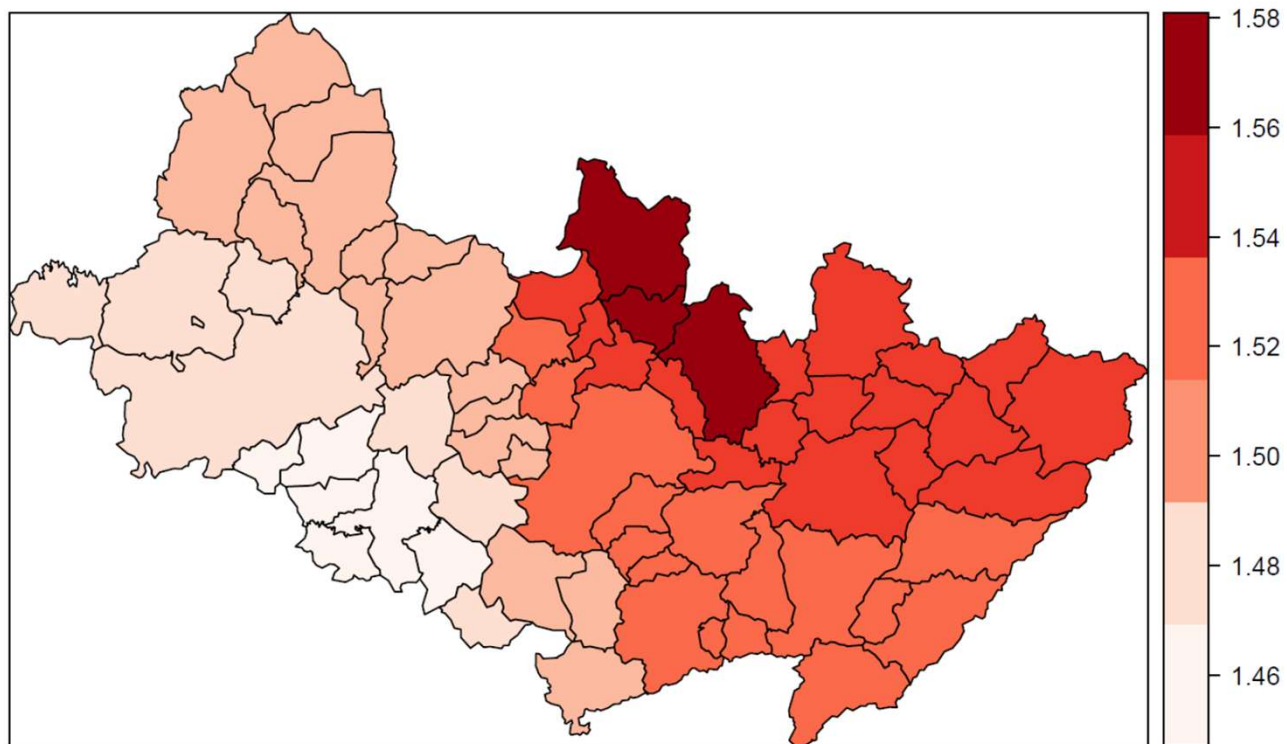
Variáveis	<i>Pooled</i>	<i>Pooled_ano</i>	EF	EF_ano	EF <i>Driscoll-Kraay</i>
Constante	-3,461***	-4,028***	6,122***	5,282***	5,282***
<i>lnpopurb</i>	1,004***	1,008***	0,097***	0,117***	0,117**
<i>lntrf</i>	-0,204***	-0,197***	-0,229***	-0,223***	-0,223***
<i>lnpib</i>	0,014***	0,008**	0,135***	0,026***	0,026***
<i>lndens</i>	0,007*	0,012***	0,758***	0,261***	0,261***
<i>lntemp</i>	0,623***	0,752***	-0,218**	0,020	0,020
<i>lnprecip</i>	-0,173***	-0,140***	-0,099***	0,044**	0,044**
<i>D</i>	0,004	-0,013	0,032***	0,004	0,004
<i>2006</i>	-	-0,105***	-	-0,138***	-0,138***
<i>2007</i>	-	-0,099***	-	-0,108***	-0,108***
<i>2008</i>	-	-0,086***	-	-0,114***	-0,114***
<i>2009</i>	-	-0,047***	-	-0,104***	-0,104***
<i>2010</i>	-	-0,006	-	-0,042***	-0,042***
<i>2011</i>	-	0,032**	-	-0,028***	-0,028***
<i>2012</i>	-	0,039***	-	0,011	0,011**
<i>2013</i>	-	0,035**	-	0,028***	0,028***
<i>2014</i>	-	-0,011	-	0,042***	0,042***
<i>2015</i>	-	-0,026	-	-0,026***	-0,026***
<i>R</i> ²	98,33	98,42	27,38	39,10	39,10

FONTE: Elaborada pela autora a partir dos *outputs* do *Stata* 15.

NOTA: Todas as variáveis do modelo foram logaritmizadas (*ln*), sendo o volume de água consumido (*lnvolag*, em 1.000 m³/s) a variável dependente. As siglas representam: *lnpopurb*=população urbana atendida pelo abastecimento de água, *lntrf*=tarifa média da água (R\$), *lnpib*=PIB municipal (R\$ 1.000,00 constantes), *lndens*=densidade demográfica (100 hab/km²), *lntemp*=temperatura média (°C), *lnprecip*=precipitação média anual (mm), *D*=cobrança (0=não há, 1=há). O asterisco indica o nível de significância dos parâmetros: * = p<10%, ** = p<5%, *** = p<1%.

Estatística Espacial

Desvio Padrão Local - Disponibilidade Hídrica (metros cúbicos/segundo)



Áreas da Estatística

- Estatística descritiva (análise exploratória de dados);
- Probabilidade;
- Inferência estatística.

Definições

- Dados:
 - ✓ Observações de um evento (pesquisa);
- Informação:
 - ✓ Alguma(s) conclusão(ões) a partir dos dados;
- População (universo):
 - ✓ Conjunto global de pesquisa ao qual se deseja extrair conclusões;
- Amostra:
 - ✓ Subconjunto finito de uma população (sobre a qual são feitas análises e inferências para a população)

Estatística Descritiva

- Objetivo:
 - Levantar hipóteses e conclusões sobre certas características da população ou amostra
- Contato com os dados e produção de informações básicas;
- Técnicas para descrever e apresentar informações inerentes às bases de dados:
 - Tabelas
 - Gráficos
 - Medidas-resumo informativas

Probabilidade

Estudo da Incerteza

- Fenômenos aleatórios:
 - Acontecimentos com resultados sem previsão certa (certeza)
 - Eventos independentes/dependentes
 - Probabilidade condicional
 - Funções de distribuição de probabilidade (discretas e contínuas)
 - a) Distribuição binomial;
 - b) Distribuição hipergeométrica;
 - c) Distribuição de Poisson;
 - d) Distribuição normal (Z , t);
 - e) Distribuição F ;
 - f) Distribuição Qui-Quadrado.

Inferência Estatística

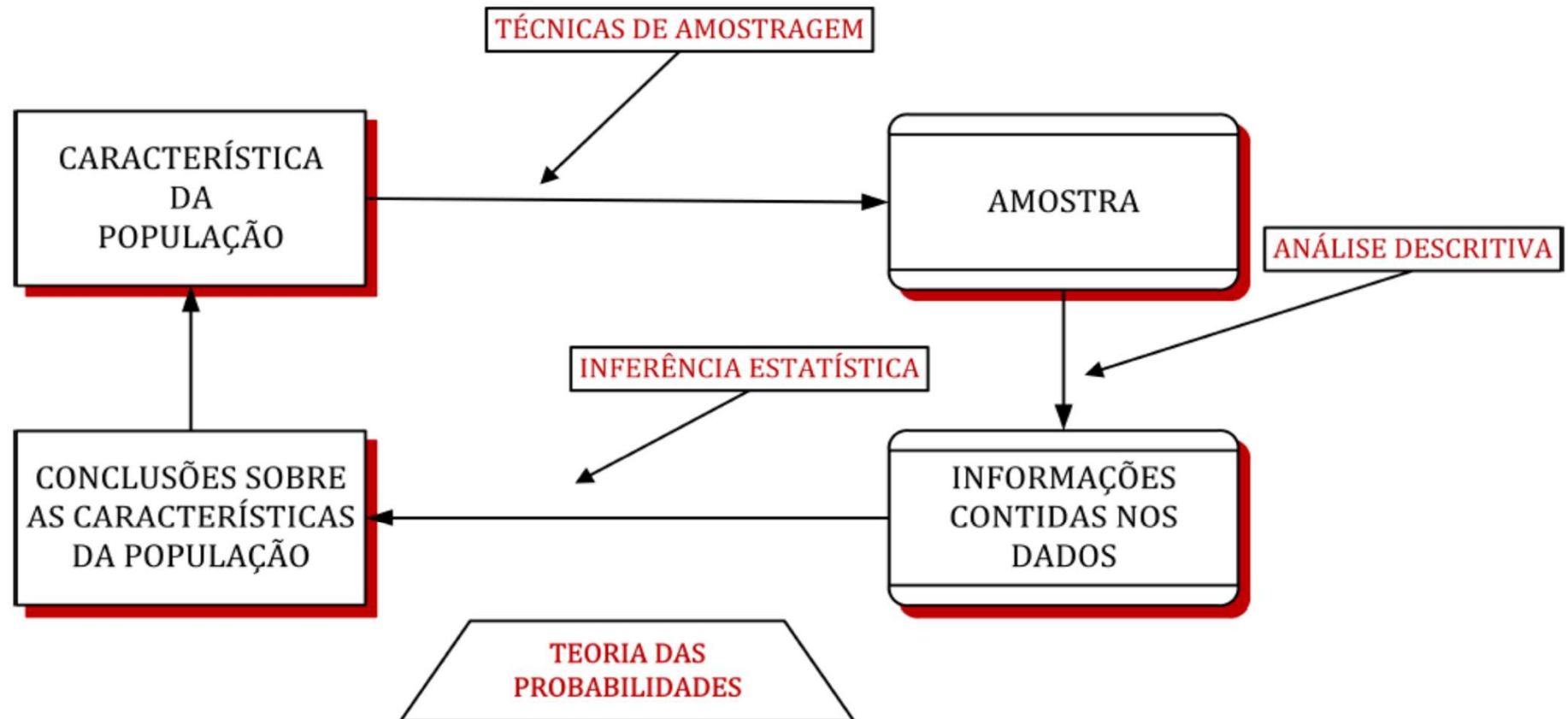
- Técnicas que permitem extrair (inferir) conclusões a partir de uma amostra;

Técnicas:

- Regressão linear simples e múltipla por Mínimos Quadrados Ordinários (cross-section ou séries temporais);
- Regressão com dados longitudinais (espaço e tempo);
- Regressão com variáveis dummy;
- Regressão com equações simultâneas;
- Regressão de modelos com resposta qualitativa (logit, probit, tobit, etc.);
- Modelos de previsão;
- Outros modelos.

População X Amostra

Etapas do Método de Análise Estatística



Fonte: adaptado de [Cancho(2010)]

Problema Fundamental da Estatística

A partir de uma amostra, “como” tirar “CONCLUSÕES” sobre a “População” ?

Resposta: MODELOS DE REGRESSÃO e INFERÊNCIA ESTATÍSTICA

População X Amostra

- População → Censo
- Amostra → Amostragem

Censo X Amostragem

- a) Para avaliar toda a população, considerando todos os elementos, realiza-se um censo. P. ex., a cada 10 anos o IBGE realiza o censo demográfico da população brasileira.
- b) Caso não haja recursos ou tempo disponíveis para execução do censo, deve-se realizar uma pesquisa por amostragem na população de interesse.

Amostragem – Why?

- ✓ Parâmetros populacionais desconhecidos;
- ✓ Impossibilidade de realização de um censo;
- ✓ Amostragem é mais barata e mais rápida;

➔ Não existe técnica estatística capaz de ajustar uma amostra mal coletada

➔ Plano Amostral (Teoria da Amostragem)

Em geral, uma amostra deve ser um subconjunto representativo da população (de alguma forma).

➤ Existem diversas maneiras para se retirar uma amostra de uma população.

Amostras Probabilísticas X Não-probabilísticas

- A amostra é obtida a partir de uma população, por meio de processos definidos pelo pesquisador.

Subdivide-se em dois grupos:

- **Amostragem Probabilística** = Cada elemento da população possui a mesma probabilidade de ser selecionado para compor a amostra → mecanismos de seleção aleatórios;

P. ex., se em uma população de 1.000 peças produzidas existem 20% de peças defeituosas, uma amostra de 100 dessas peças também deve possuir essa mesma proporção de peças defeituosas, ou seja, 20 peças.

- **Amostragem Não Probabilística** = A seleção da amostra depende do julgamento do pesquisador. Há uma escolha deliberada dos elementos para compor a amostra → mecanismos não aleatórios de seleção.

Amostragem com reposição X sem reposição

Quando a população é finita, a amostragem pode ser feita de duas formas:

- **Com reposição** - o elemento amostrado volta a fazer parte da população após ser analisado.
 - O espaço amostral permanece inalterado.

Exemplo: amostrar a quantidade de veículos que passam em determinada avenida entre as 17:30 e 19:30hs na primeira semana do mês, o veículo amostrado volta para a população.

- **Sem reposição** - o elemento amostrado não volta a compor a população após ser analisado.
 - O espaço amostral é alterado.

Exemplo: amostrar pessoas para avaliar hábitos de consumo, pessoas amostradas não voltam à população.

População Finita e Infinita

- Em populações infinitas, escolher entre amostragem com reposição ou sem reposição é irrelevante.
- Geralmente, costuma-se empregar o seguinte critério para saber se uma população pode ser considerada finita em relação à amostra:

Se $\frac{n}{N} \geq 0,05$, ou seja 5%, a população é considerada finita. Em outras palavras, quando a amostra for menor ou igual a 5% da população, ela deve ser considerada infinita.

- Sendo N e n os tamanhos populacional e amostral, respectivamente.

Plano Amostral

- 1) Definição da população alvo** (ex. famílias residentes na região metropolitana de Curitiba, ano 2020 nos meses de fevereiro e março);
- 2) Determinação do quadro amostral** (lista, localização ou disponibilidade da população; ex: questionário aplicado em residências sorteadas nos diversos bairros, ponderadas pelo tamanho populacional);
- 3) Seleção da técnica de amostragem** (Aleatória simples; Estratificada; Conglomerados; etc. – com uso de questionários, formulário eletrônico, etc.);
- 4) Determinar o tamanho da amostra** (Se a população é conhecida e a variável de interesse existe; caso contrário faz-se a determinação do tamanho da amostra baseada em experimento);
- 5) Executar o processo de amostragem** (no caso de experimento o processo de amostragem é feito em conjunto com a determinação do tamanho da amostra.

Amostras mais precisas são geralmente amostras maiores (reduz erro amostral)!!

Erros Amostrais

- **Erros amostrais:** Trata-se da diferença entre o que foi apurado na amostra e o verdadeiro valor da população.
 - ✓ Ocorrem pois as amostras são aleatórias
- **Erros não amostrais:** Ocorrem quando os dados amostrais são coletados incorretamente, devido a(o):
 - i) uma amostra tendenciosa;
 - ii) instrumento de medida defeituoso;
 - iii) anotações erradas no questionário;
 - iv) Etc.
- Observação: Os erros não amostrais devem ser minimizados

Dimensionamento do Tamanho da Amostra

- Para populações “infinitas” (amostras pequenas)

$$n = \left(\frac{Z_{\alpha} \cdot \sigma_x}{d} \right)^2$$

Em que:

n = tamanho da amostra;

Z_{α} = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\alpha}{2}$, ou seja, o valor tabelado de $Z = 1,96$;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

d = erro amostral absoluto admitido;

Dimensionamento do Tamanho da Amostra

- Quando o desvio padrão populacional da variável “x” é desconhecido, opta-se por realizar sua estimativa por experimento, durante o processo de amostragem.
 - Este experimento inicia-se com os valores das primeiras 50 observações (existem divergências) amostrais selecionadas aleatoriamente (para pesquisas socioeconômicas), eliminados os outliers.
 - Após efetua-se o cálculo do desvio-padrão amostral da média dos valores de x, $(Sx^- = s/\sqrt{n})$ e se utiliza este valor em substituição ao valor de “ σ_x ”. Este método é repetido com o crescimento da amostra para verificar a variação do valor do desvio padrão e a suficiência do tamanho da amostra.

Dimensionamento do Tamanho da Amostra

Para estimar o desvio padrão da amostra aplica-se a seguinte expressão:

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

Em que:

s = desvio padrão da amostra;

X = cada valor da variável X da amostra;

\bar{X} = média dos valores de X da amostra;

n = tamanho da amostra.

Dimensionamento do Tamanho da Amostra

O erro amostral absoluto admitido é dado por:

$$d = Z_{\alpha} \cdot \frac{\sigma_x}{\sqrt{n}}$$

Em que:

d = erro amostral absoluto admitido;

Z_{α} = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\alpha}{2}$, ou seja, o valor tabelado de $Z = 1,96$;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

n = tamanho da amostra.

Dimensionamento do Tamanho da Amostra

- Para populações “finitas” (grandes amostras)

$$n = \frac{1}{\frac{d^2}{Z_{\alpha}^2 \cdot \sigma_x^2} + \frac{1}{N}}$$

Em que:

n = tamanho da amostra;

d = erro amostral absoluto admitido;

Z_{α} = valor da variável aleatória normal padrão Z que deixa uma área de cauda a direita com valor $\frac{\alpha}{2}$, ou seja, o valor tabelado de $Z = 1,96$;

σ_x = desvio padrão populacional da variável x (ou sua estimativa);

N = Tamanho da população.

Escolha do Método de Amostragem

- Para a escolha do método deve-se levar em conta:
 - ✓ Tipo de pesquisa que será feita;
 - ✓ Acessibilidade e disponibilidade dos elementos da população;
 - ✓ Disponibilidade de tempo;
 - ✓ Disponibilidade de recursos financeiros e humanos;
 - ✓ Outros fatores específicos da pesquisa.

Amostragem Probabilística

- Os métodos mais usuais para se realizar uma amostragem probabilística são: amostragem aleatória simples, amostragem sistemática, amostragem estratificada e amostragem por conglomerados.

Amostragem aleatória simples

- É necessário o conhecimento do tamanho e características da população;
- A amostra é escolhida por sorteio (aleatoriamente) ou pela tabela de números aleatórios;
- Outras formas de acesso aleatório aos elementos da população são admitidas.

Amostragem Aleatória Simples

- **Para uma população infinita:**

- Pode-se utilizar faixas de valores.

- Porém, é necessário que se garanta que as probabilidades de intervalos de valores que sejam incluídos na amostra sejam iguais às porcentagens da população.

- Por exemplo, em uma pesquisa sobre massa corporal, se em uma população 40% dos elementos possuem massa corporal entre 50 e 70Kg, na amostra a porcentagem deve se manter: aproximadamente 40% dos elementos amostrados também devem ter massa corporal entre 50 e 70Kg.

Amostragem Sistemática

- Utilizada quando os elementos estão dispostos de maneira organizada (ex.: fila, lista) e aleatória.
- Escolhe-se um ponto de partida (por sorteio) e seleciona-se cada k-ésimo elemento da população (ex.: o 20º elemento; 40º elemento, etc.)
- Exemplo: Em uma fábrica de parafusos, a cada 1000 peças produzidas, uma é retirada para teste.

Amostragem Estratificada

- Indicada quando a população está dividida em grupos distintos, denominados estratos.
- Dentro de cada estrato é realizada uma amostragem aleatória simples.
- O tamanho da amostra pode ou não ser proporcional ao tamanho do estrato (quando é proporcional diz-se que trata-se de uma **Amostragem Estratificada Proporcional**).

Exemplo: Uma comunidade universitária com 8000 indivíduos está estratificada da seguinte forma:

Estrato	População	Amostra
Professores	800	80
Funcionários	1200	120
Estudantes	6000	600

Amostragem por Conglomerados

- A área da população é dividida em seções (ou conglomerados, ex.: bairros, quarteirões).
- Os conglomerados são selecionados aleatoriamente;
- Dentro de um conglomerado, todos os elementos são amostrados;
- Eventualmente pode-se ter Amostragem por conglomerados em que coleta-se uma amostra e utiliza-se pesos para expandir os resultados para a população;

Exemplo: POF do IBGE - conglomerados (municípios + setores censitários + domicílios = AAS) com estratificação (a partir do censo demográfico).

Amostragem Não Probabilística

- Na amostragem não probabilística, os elementos da população não tem a mesma probabilidade de serem selecionados, portanto, não há garantias da representatividade da população.

Amostragem por conveniência

Os elementos são selecionados por serem imediatamente disponíveis.
Exemplo: Uma repórter entrevistando pessoas na rua.

Amostragem por julgamento

Uma pessoa experiente no assunto escolhe intencionalmente os elementos a serem amostrados. Exemplo: Um novo produto a ser “testado” por pessoas.

Amostragem Não Probabilística

Amostragem por quotas

- A amostragem por quotas consiste em um refinamento da amostragem por conveniência. Nela, os elementos selecionados devem obedecer as proporções de características da população.

Exemplo: se em uma população, existem 20% de indivíduos da classe econômica A, 50% da classe B e 30% da classe C, a amostra por conveniência deve respeitar essas proporções (amostragem por quotas), selecionando 20% de indivíduos da classe econômica A, 50% da classe B e 30% da classe C.

Dados de Pesquisas e Variáveis

Dados

- São dados obtidos diretamente em pesquisas
 - Ainda sem qualquer tratamento, processo de síntese ou análise
- Podem estar incluídos em tabelas
 - Porém, não tratados em publicações

Variáveis

- Característica que pode variar em termos de valor entre elementos em uma amostra ou população
 - Elementos diferentes podem ter valores diferentes de uma variável (por exemplo: Renda do indivíduo)

Variáveis

Quantitativa:

- Escala de medida tem valores numéricos (12; 18; 1.253; 1.832,21; 1.326,18; etc.);
- Os valores podem representar magnitudes diferentes para cada variável (R\$; kg; metros; etc.).

Qualitativas:

- Escala de medida é um conjunto de categorias (M; F; etc.);
- Categorias distintas diferem em qualidade, não em magnitude numérica.

Escala de Medida

- Valores que a variável expressa

Exemplo:

- Gênero: masculino, feminino, outro;
 - Número de moradores do domicílio: 0, 1, 2, ...
 - Renda: R\$1.200,00; 1.800,00; etc.
 - Altura: 180cm; 1,76m; etc.
-
- Os métodos estatísticos consideram a escala das variáveis.

Variáveis Quantitativas

- **Variável discreta:**

- Os valores possíveis formam um conjunto de números como: 0, 1, 2, 3, ... ; (Ex.: número de moradores, número de veículos, etc.)

- **Variável contínua:**

- Os valores são um “contínuo” de valores reais possíveis: 0,01 ; 0,02; 0,03; ...; 1,00; 1,01; ..; (Ex.: Renda do indivíduo; temperatura média; etc.)

Variáveis Qualitativas

- **Variável nominal ou categórica:**

- Os valores são apresentados em categorias mutuamente exclusivas (Ex.: Gênero, Região, etc.)

- **Variável ordinal:**

- Os valores são apresentados em categorias mutuamente exclusivas que possuem uma ordenação natural (Ex.: Grau de instrução; Classe socioeconômica; etc.)

Séries Estatísticas

- As séries estatísticas são a apresentação de informações (variáveis estatísticas) na forma de tabelas, cujo objetivo é sintetizar as informações estatísticas observadas e torná-las compreensíveis.
- Uma tabela ou um gráfico deve ser composto de: **título; corpo e; rodapé.**
- **O título** deve conter informações suficientes para compreender as informações expostas:
 - ✓ Fato ou variável de interesse – fenômeno descrito;
 - ✓ Espaço Geográfico – local a que se refere o fato;
 - ✓ Época – tempo em que o fato foi observado.
- **O corpo** deve conter o registro dos dados.
- **O rodapé** deve conter a identificação da fonte de dados.

Série Temporal

- Série temporal é uma série estatística em que os dados são apresentados segundo a época de ocorrência.
- ✓ O tempo varia, mas o fato - variável(is) de interesse - e o local são fixos.

Tabela 6 - Incentivos à exportação de manufaturados (como % do valor FOB das exportações de manufaturados) - Período 1969-85.

Ano	Isenções			Subsídios				TOTAL GERAL
	ICM	IPI	Draw-back	Crédito Prêmio	Isenção I.R.	Subs. Crédito	Sub-Total	
1969	20,5	7,4	4,0	6,7		4,1	10,8	42,7
1970	20,5	7,2	4,0	13,5		7,5	21,0	52,7
1971	19,8	7,1	4,0	13,2	1,3	7,8	22,3	53,1
1972	19,1	9,0	4,9	16,3	1,3	8,2	25,8	58,8
1973	18,3	8,9	7,2	16,2	1,3	6,5	23,9	58,3
1974	17,7	5,0	12,6	12,0	1,8	6,1	19,9	55,2
1975	17,0	5,4	8,3	12,1	1,7	11,5	25,3	56,0
1976	16,3	5,2	11,8	11,7	1,3	15,9	28,9	62,2
1977	16,3	5,4	12,6	12,4	1,5	19,6	33,5	67,9
1978	16,3	6,1	9,1	12,8	1,8	17,0	31,6	63,1
1979	16,3	6,5	10,5	12,8	2,1	13,9	28,8	62,1
1980	17,7	6,3	9,0	0,0	1,9	2,0	3,9	37,0
1981	18,3	6,8	9,4	6,5	1,8	18,7	27,0	61,6
1982	19,1	7,2	10,3	9,1	1,6	21,7	32,4	69,0
1983	19,1	7,2	8,6	7,8	1,6	9,3	18,7	53,6
1984	20,5	7,0	9,1	7,8	1,6	2,7	12,1	48,7
1985	20,5	7,2	9,1	1,4	1,6	3,6	6,6	43,4

Fonte: Bauman e Moreira (1987) in CLEMENTS, J. C. & McCLAIN, J. S. *The political economy of export promotion in Brazil*. In: GRAHAM, L. S. & WILSON, R. H. *The political economy of Brazil: public policies in an era of transition*. Institute of Latin American Studies Symposia on Latin America Series. Austin: The University of Texas Press, 1990.

Série Geográfica

- Série geográfica é uma série estatística na qual os dados são apresentados segundo a localidade de ocorrência.
- O local varia e o tempo e a variável(eis) são fixos

Tabela 1: Tarifa média máxima e aplicada; membros selecionados da OMC, 2013

País (PIB per capita)	Produtos	Média das tarifas	
		máximas	Média das tarifas aplicadas
Estados Unidos (US\$ 53000)	Todos	4	3
	Agricultura	5	5
	Vestuário	12	12
Brasil (US\$ 11208)	Todos	31	14
	Agricultura	35	10
	Vestuário	35	35
Índia (US\$ 1498)	Todos	49	14
	Agricultura	114	34
	Vestuário	38	13

Fonte: World Trade Organization Tariff Profiles e World Bank

Série Específica

- Série Específica – Os dados são agrupados segundo a modalidade de ocorrência. Fato – variável(eis) de interesse, tempo e local fixos.

**Taxa (%) de crescimento semestral
dos componentes da DA ¹**

2011.I

Setores	(%)
Consumo das Famílias	5,7
Consumo da Adm. Pública	2,3
Formação Bruta de Capital Fixo	7,3
Exportação	5,2
Importação	13,9

Fonte: Contas Nacionais Trimestrais - IBGE.

Nota: (1) Refere-se ao crescimento com respeito ao
mesmo período do ano anterior

Série Mista

- Série Mista é uma combinação de duas ou mais dos 3 tipos de séries (Temporal, Geográfica e Específica).

Tabela 1
Participação (%) das Economias Fluminense e Paulista por Atividade Econômica
em Valor Adicionado Bruto a Preço básico no Brasil
(1996 e 2008)

Atividade Econômica	1996		2008	
	RJ	SP	RJ	SP
Agropecuária	1,4	8,6	0,8	7,9
Indústria	8,6	42,9	12,7	33,9
Indústria Extrativa Mineral	18,7	5,0	53,5	1,2
Indústria de Transformação	6,3	46,8	6,7	43,7
Construção	13,9	36,5	10,6	27,6
SIUP*	8,2	45,4	6,0	25,6
Serviços	13,0	34,9	11,6	33,4
Comércio	8,4	41,3	8,3	31,2
Financeiro	10,7	49,9	8,8	51,1
APU**	14,7	20,9	12,7	19,0
Outros	14,2	35,7	12,9	37,7
Total	11,2	35,6	11,2	32,0

Fonte: Elaborado pelo Grupo de Economia da Inovação (IE/UFRJ) com base na Retropolação das Contas Nacionais e Regionais, IBGE.

Nota: Preço básico em R\$ de 1995.

* Produção e distribuição de eletricidade e gás, água, esgoto e limpeza urbana.

** Administração, saúde e educação públicas e seguridade social.

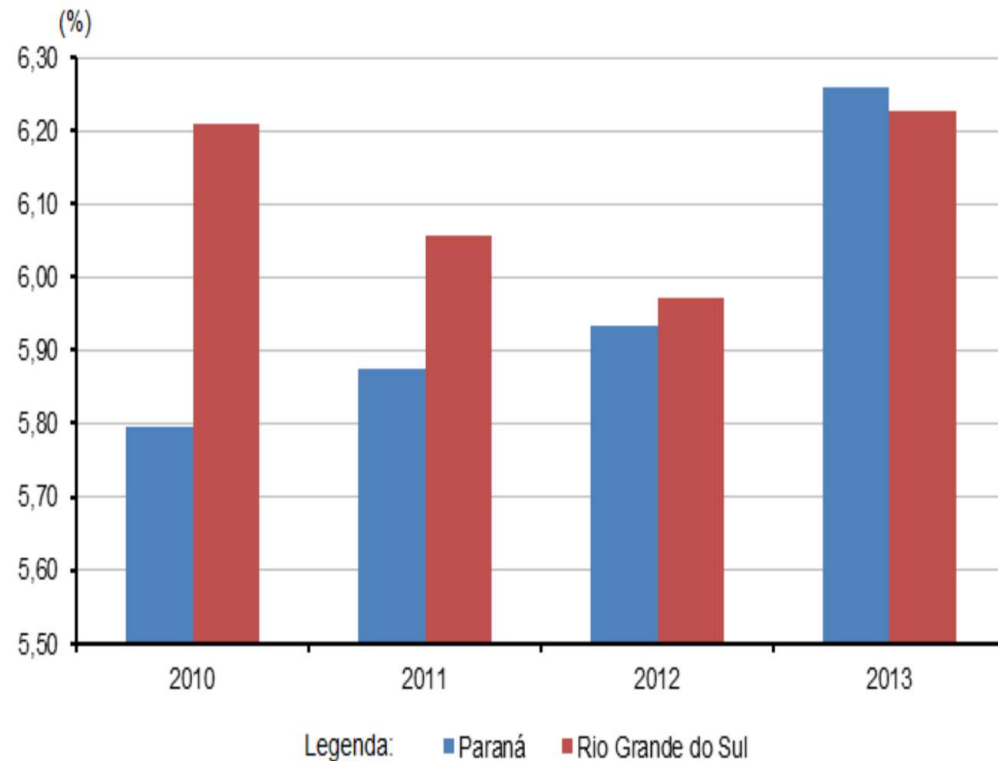
Gráficos

- A apresentação gráfica de dados estatísticos objetiva expor de forma concisa e simples as séries.
- Isso permite concluir sobre a evolução do fenômeno – variável - ou sobre o comportamento dos valores da série.
- Existem várias maneiras de se representar graficamente os dados estatísticos de acordo com os tipos de séries:
 - ✓ Colunas ou Barras;
 - ✓ Setores;
 - ✓ Linhas;
 - ✓ Dispersão;
 - ✓ Outros.

Gráfico de Colunas ou Barras

Gráfico 1

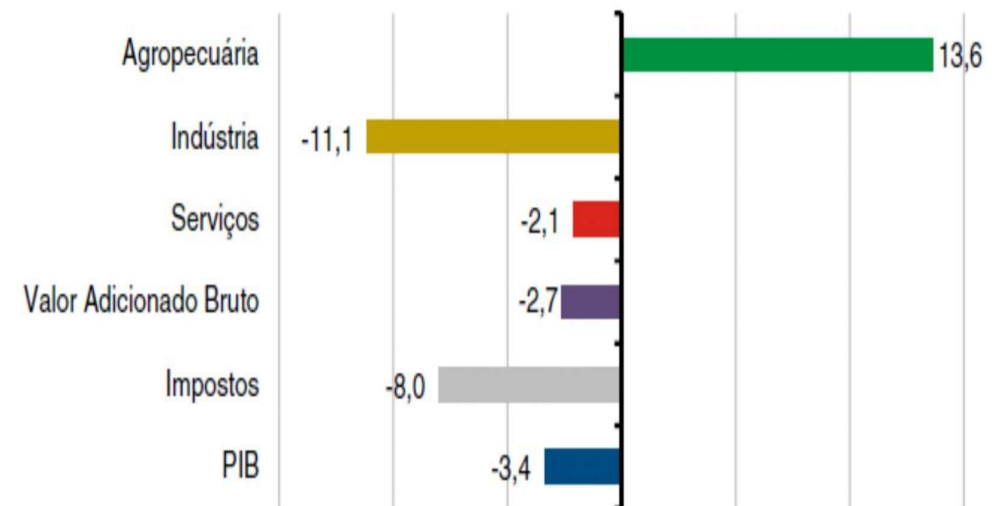
Participação dos Produtos Internos Brutos (PIB) do Rio Grande do Sul e do Paraná no do Brasil — 2010-13



Fonte: IBGE.
FEE.

Gráfico 1

Taxas percentuais de crescimento do Produto Interno Bruto (PIB), acumuladas ao longo do ano, do Rio Grande do Sul — 4.º trim./15

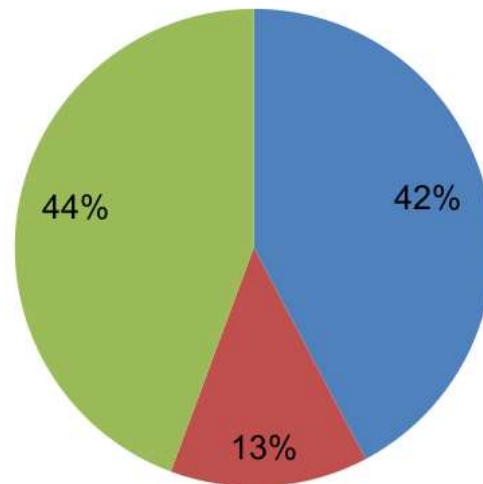


Fonte: FEE/CIES/NCR.

Gráfico de Setores

Gráfico 1

Distribuição dos componentes do Produto Interno Bruto, pela ótica da renda, no Rio Grande do Sul — 2015



- Remunerações
- Impostos, líquidos de subsídios, sobre a produção e a importação
- Excedente Operacional Bruto (EOB) e Rendimento Misto Bruto (RMB)

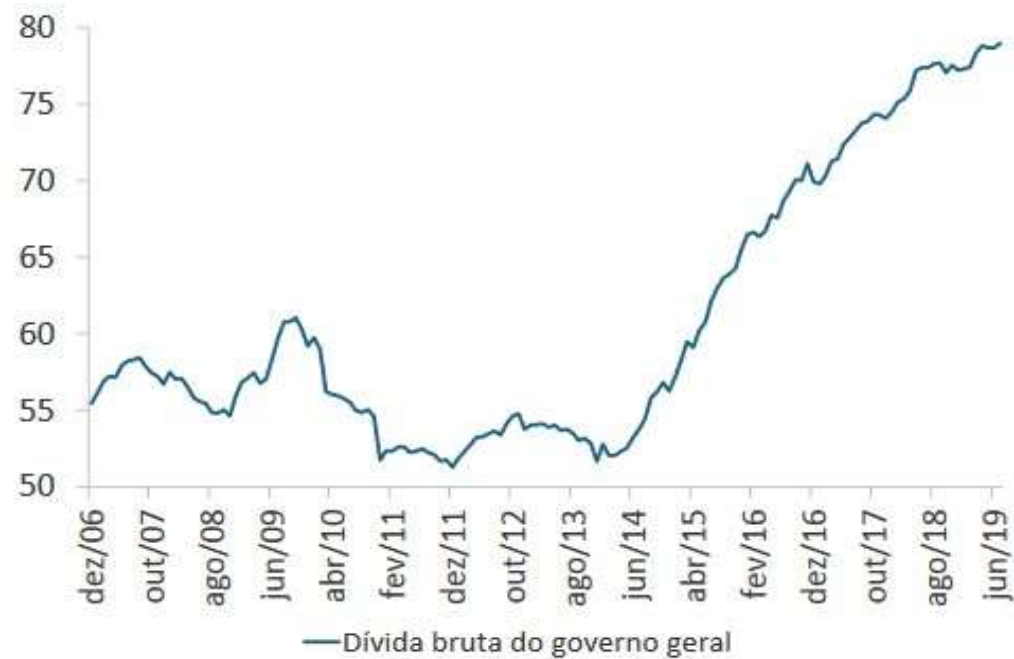
FONTE: IBGE.
FEE.

Gráfico de Linhas

GRÁFICO 5

Dívida bruta do governo geral (DBGG)

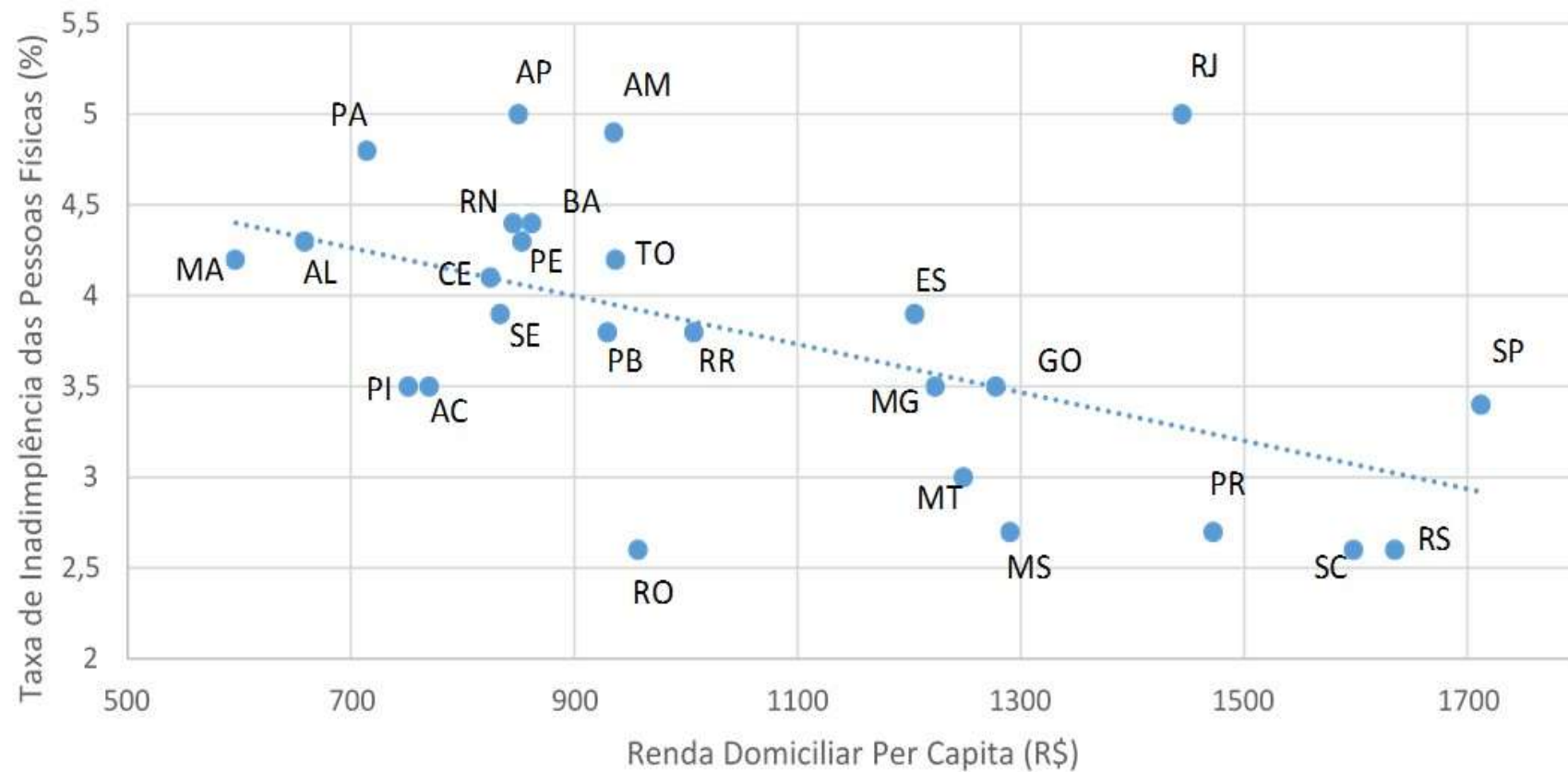
(Em % do PIB)



Fonte: Banco Central do Brasil (BCB).

Diagrama de Dispersão

Gráfico 1: Taxa de Inadimplência das Pessoas Físicas x Renda Domiciliar Per Capita - 2017. Fonte: BCB, IBGE. Elaboração: Boa Vista SCPC



Histograma

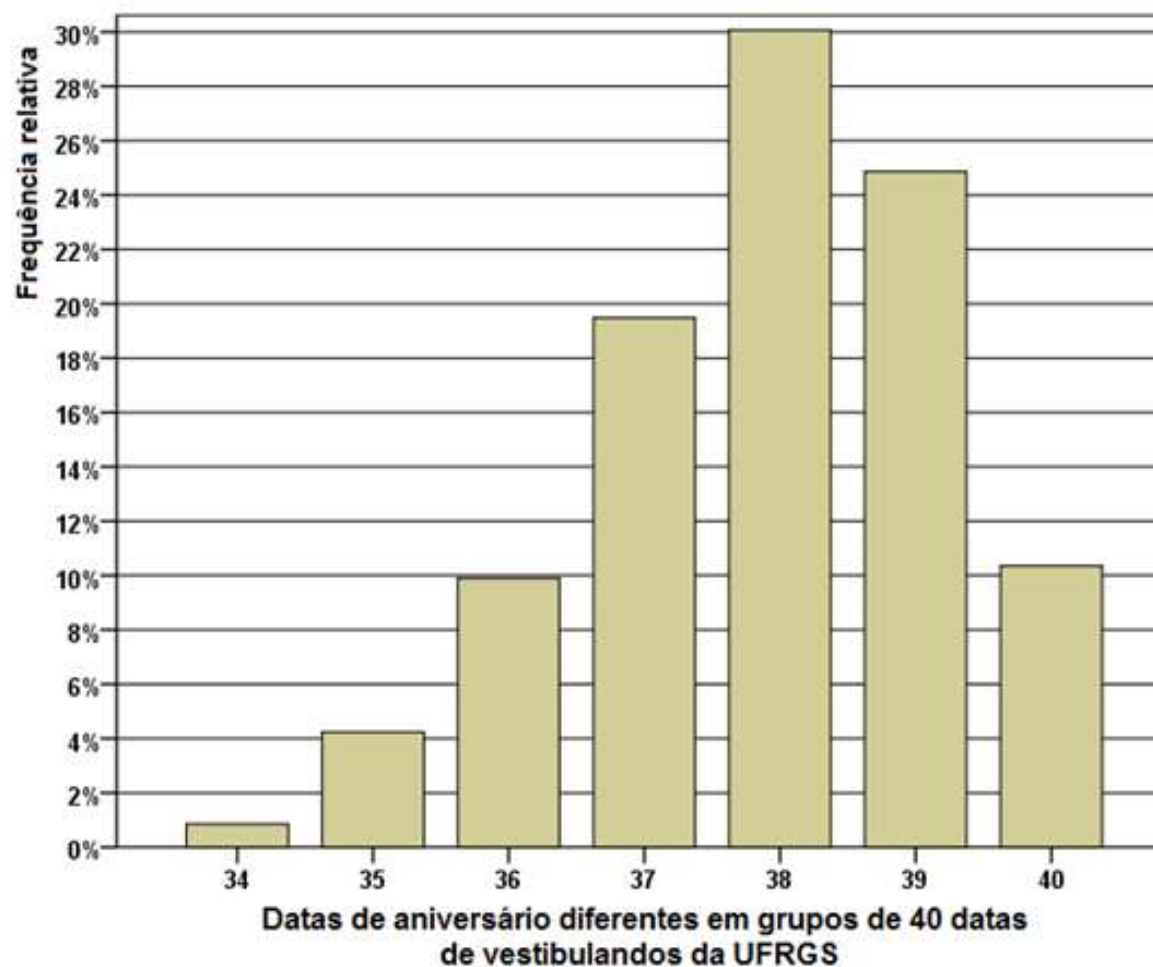
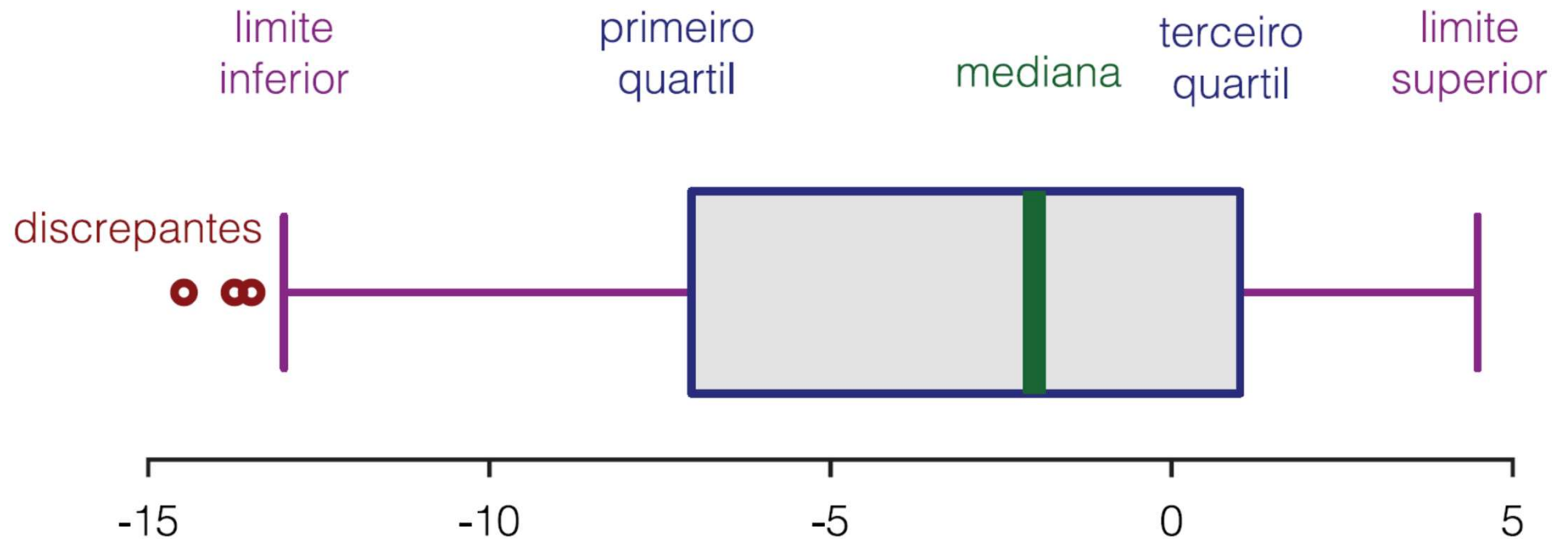


Figura 2 - Distribuição de frequências relativas das datas de aniversário diferentes em 2454 grupos de 40 vestibulandos da UFRGS.

Diagrama de Caixa – Box Plot



Distribuição de Frequência

Tabela 5 - Número de pessoas residentes nos domicílios das crianças e dos adolescentes internados no Hospital João XXIII

Distribuição	Frequência	
	ABS	%
De 2 a 4 pessoas	118	10,2
De 5 a 9 pessoas	65	5,6
Acima de 10 pessoas	4	,3
Sem Informação	965	83,8
Total	1152	100,0

Fonte: SAME-Hospital João XXIII

Distribuição de Frequência

Como construir?

- 1) Ter em mãos a tabela de dados (amostra ou população);
- 2) Determinar o número de classes (K)

Alternativas:

a) $K = 5$ se $n < 25$

$$K = \sqrt{n} \text{ se } n \geq 25;$$

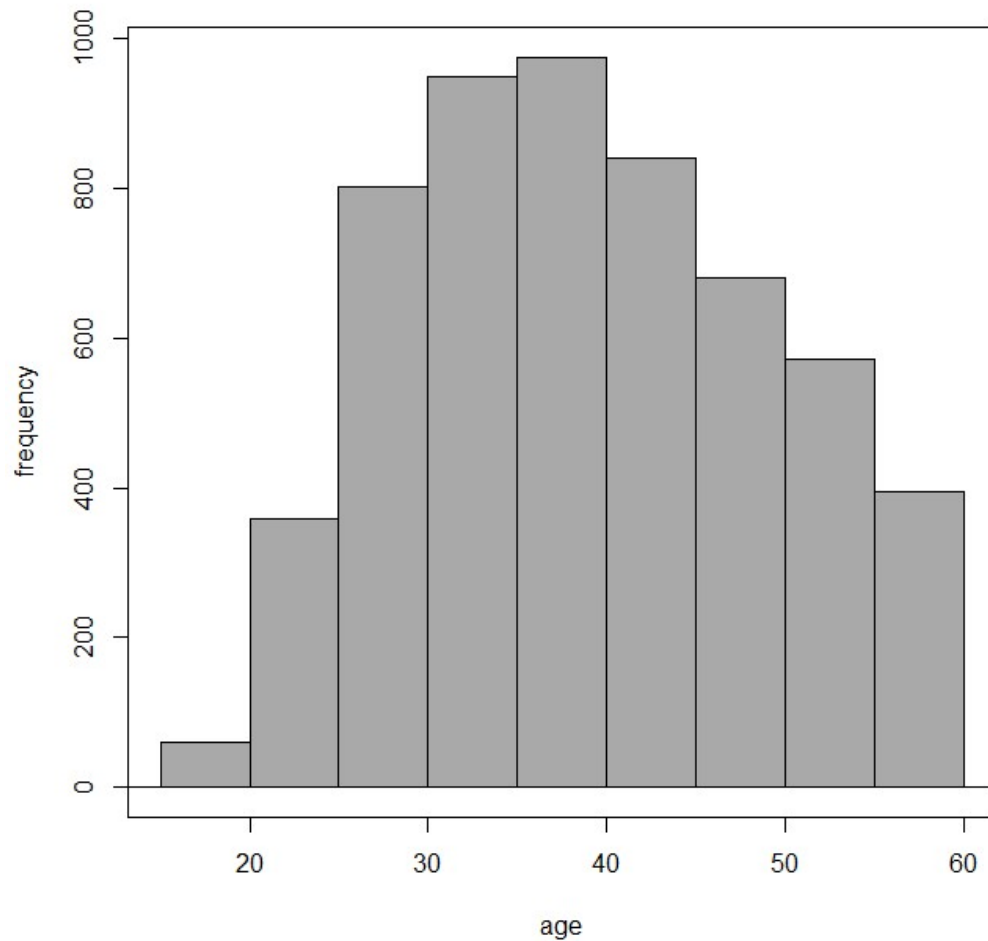
b) Fórmula de Sturges $\rightarrow K = 1 + 3,32 \cdot \log(n)$

em que n = tamanho da amostra ou população;

c) Determinação ad hoc (a gosto do pesquisador).

Distribuições de Probabilidade – 5634 pessoas (Idade)

Histograma



Curva de Densidade

