

Pesquisa Censitária e Amostral

Matemática – 8º Ano

População, amostra, representatividade, amostragem e vieses estatísticos

Objetivos de aprendizagem

Ao final deste curso, o estudante deverá ser capaz de:

- 1) compreender o conceito de pesquisa estatística;
- 2) identificar a população de interesse;
- 3) definir uma unidade estatística;
- 4) distinguir população e amostra;
- 5) distinguir pesquisa censitária e pesquisa amostral;
- 6) reconhecer vantagens e limitações de um censo;
- 7) reconhecer vantagens e limitações de uma pesquisa amostral;
- 8) compreender o conceito de representatividade;
- 9) compreender por que uma amostra grande não é necessariamente representativa;
- 10) reconhecer a importância de um processo de seleção adequado;
- 11) compreender a amostragem aleatória simples;
- 12) compreender a amostragem sistemática;
- 13) compreender a amostragem estratificada;
- 14) distinguir métodos probabilísticos de seleção de métodos por conveniência;
- 15) reconhecer viés de seleção;
- 16) reconhecer viés de resposta;
- 17) reconhecer problemas de não resposta;
- 18) compreender a importância da formulação neutra de perguntas;
- 19) interpretar resultados amostrais com cautela;
- 20) compreender qualitativamente o conceito de erro amostral;
- 21) justificar a escolha entre censo e amostra em diferentes contextos;
- 22) planejar uma pesquisa amostral simples e coerente.

Sumário

1	Introdução	1
2	Pesquisa estatística	1
3	População	2
4	População não significa necessariamente pessoas	2
5	Unidade estatística	2
6	Amostra	3
7	Exemplo	3
8	Notação	4
9	Pesquisa censitária	4
10	Exemplo de censo	4
11	Pesquisa amostral	5
12	Exemplo de pesquisa amostral	5
13	Censo versus amostra	5
14	Vantagens de um censo	5
15	Limitações de um censo	6
16	Censo não significa ausência de erro	6
17	Vantagens de uma pesquisa amostral	7
18	Exemplo: teste destrutivo	7
19	Representatividade	7
20	Exemplo de falta de representatividade	7
21	Tamanho não garante representatividade	8

22 Amostra grande versus amostra bem selecionada	8
23 Amostragem	8
24 Amostragem aleatória simples	9
25 Exemplo de amostragem aleatória simples	9
26 Por que utilizar aleatoriedade?	9
27 Amostragem sistemática	10
28 Intervalo sistemático	10
29 Exemplo sistemático	10
30 Importância do ponto inicial aleatório	11
31 Risco de periodicidade	11
32 Amostragem estratificada	11
33 Por que estratificar?	12
34 Amostragem estratificada proporcional	12
35 Fórmula proporcional	13
36 Exemplo estratificado	13
37 Seleção dentro dos estratos	14
38 Comparação dos métodos	14
39 Amostragem por conveniência	14
40 Problema da conveniência	15
41 Amostra voluntária	15
42 Viés	15
43 Viés de seleção	15

44 Exemplo de viés de seleção	15
45 Viés de não resposta	16
46 Por que a não resposta pode ser problemática?	16
47 Viés de resposta	16
48 Pergunta tendenciosa	16
49 Perguntas duplas	17
50 Perguntas ambíguas	17
51 Ordem das perguntas	17
52 Quadro de amostragem	18
53 Problema de cobertura	18
54 Erro amostral	18
55 Erro amostral não significa erro do pesquisador	19
56 Exemplo intuitivo	19
57 Tamanho da amostra e variabilidade	19
58 Exemplo crucial	19
59 Parâmetro e estatística – aprofundamento	20
60 Exemplo	20
61 Inferência estatística – ideia	20
62 Exemplo completo: pesquisa escolar	21
63 Estratificação proporcional	21
64 Seleção dentro de cada série	22
65 Pesquisa censitária ou amostral?	22

66 Caso 1: turma pequena	22
67 Caso 2: população nacional	23
68 Caso 3: teste destrutivo	23
69 Caso 4: atualização cadastral individual	23
70 Planejamento de uma pesquisa	23
71 Fluxograma do planejamento	24
72 Pergunta estatística	24
73 Exemplo de pergunta não estatística	25
74 Variável da pesquisa	25
75 População-alvo	25
76 População acessível	26
77 Generalização indevida	26
78 Proporção amostral	26
79 Resultado da amostra não é resultado exato da população	27
80 Viés versus erro amostral	27
81 Comparação	28
82 Exemplo comparativo	28
83 Amostragem estratificada e proporcionalidade	28
84 Cálculo por estrato	29
85 Arredondamentos em estratificação	29
86 Exemplo de arredondamento	30
87 Escolha do método	30

88	Não existe método universal	31
89	Boas práticas para questionários	31
90	Exemplo de reformulação	31
91	Categorias exaustivas	32
92	Categorias mutuamente exclusivas	32
93	Pesquisa piloto	33
94	Ética e privacidade	33
95	Dados pessoais desnecessários	33
96	Estratégia geral	34
97	Erros comuns	34
98	Mapa conceitual	36
99	Resumo teórico	36
100	Exercícios	37
100.1	Parte A – População e amostra	37
100.2	Parte B – Censo ou amostra	37
100.3	Parte C – Escolha entre censo e amostra	38
100.4	Parte D – Aleatória simples	38
100.5	Parte E – Sistemática	38
100.6	Parte F – Estratificada	39
100.7	Parte G – Mais estratificação	39
100.8	Parte H – Representatividade	39
100.9	Parte I – Viés de seleção	39
100.10	Parte J – Perguntas tendenciosas	40
100.11	Parte K – Não resposta	40
100.12	Parte L – Tamanho versus qualidade	40
100.13	Parte M – Proporção na amostra	41
100.14	Parte N – Método adequado	41
100.15	Parte O – Desafio	41

101	Gabarito comentado	42
101.1	Parte A	42
101.2	Parte B	43
101.3	Parte C	43
101.4	Parte D	43
101.5	Parte E	44
101.6	Parte F	44
101.7	Parte G	45
101.8	Parte H	45
101.9	Parte I	45
101.10	Parte J	46
101.11	Parte K	47
101.12	Parte L	47
101.13	Parte M	47
101.14	Parte N	48
101.15	Parte O	48
102	Avaliação final	49
102.1	Gabarito da avaliação final	50
103	Conclusão	52

1 Introdução

Imagine que queremos responder à pergunta:

Qual é o meio de transporte mais utilizado pelos estudantes de uma escola?

Uma possibilidade seria entrevistar:

todos os estudantes.

Outra possibilidade seria entrevistar apenas uma parte deles.

Essas duas estratégias correspondem, respectivamente, a:

pesquisa censitária

e:

pesquisa amostral.

A escolha entre essas estratégias depende de fatores como:

- tamanho da população;
- custo;
- tempo disponível;
- precisão desejada;
- facilidade de acesso aos indivíduos;
- possibilidade de selecionar uma amostra representativa.

2 Pesquisa estatística

Definição 2.1: Pesquisa estatística

Uma **pesquisa estatística** é um processo organizado de coleta, organização, análise e interpretação de dados com o objetivo de responder a uma questão.

Uma pesquisa pode envolver:

- 1) definição do problema;
- 2) identificação da população;
- 3) escolha das variáveis;
- 4) elaboração de instrumento de coleta;

- 5) seleção dos participantes;
- 6) coleta dos dados;
- 7) organização em tabelas e gráficos;
- 8) análise;
- 9) conclusão.

3 População

Definição 3.1: População estatística

A **população estatística** é o conjunto completo de indivíduos, objetos ou unidades sobre os quais desejamos obter informações.

Por exemplo:

- todos os estudantes de uma escola;
- todos os moradores de uma cidade;
- todas as lâmpadas produzidas por uma fábrica em um dia;
- todos os veículos registrados em determinado município.

4 População não significa necessariamente pessoas

Observação

Na Estatística, a palavra população pode representar qualquer conjunto de unidades de interesse.

Por exemplo, em uma pesquisa sobre a duração de lâmpadas, a população pode ser formada por:

lâmpadas,

e não por pessoas.

5 Unidade estatística

Definição 5.1: Unidade estatística

A **unidade estatística** é cada elemento individual pertencente à população estudada.

Se a população é formada pelos estudantes de uma escola, então cada estudante é uma unidade estatística.

Se a população é formada por peças produzidas em uma fábrica, então cada peça é uma unidade estatística.

6 Amostra

Definição 6.1: Amostra

Uma **amostra** é um subconjunto da população selecionado para fornecer informações sobre a população inteira.

Em símbolos:

$$A \subseteq P,$$

onde:

$$P$$

representa a população e:

$$A$$

a amostra.

7 Exemplo

Uma escola possui:

$$1200$$

estudantes.

Uma pesquisa entrevista:

$$150$$

deles.

Então:

$$N = 1200$$

é o tamanho da população e:

$$n = 150$$

é o tamanho da amostra.

8 Notação

Notação

É comum utilizar:

$$N$$

para o tamanho da população e:

$$n$$

para o tamanho da amostra.

Em geral:

$$n \leq N.$$

9 Pesquisa censitária

Definição 9.1: Pesquisa censitária

Uma pesquisa é **censitária** quando os dados são coletados em todas as unidades da população de interesse.

Assim:

$$\text{censo} \implies \text{toda a população é investigada.}$$

10 Exemplo de censo

Uma turma possui:

$$30$$

estudantes.

O professor pergunta a todos os 30 estudantes:

“Quantos livros você leu neste semestre?”

Como todos os membros da população foram consultados:

$$\text{a pesquisa é censitária.}$$

11 Pesquisa amostral

Definição 11.1: Pesquisa amostral

Uma pesquisa é **amostral** quando apenas uma parte da população é investigada.

O objetivo é utilizar a amostra para obter informações sobre a população.

12 Exemplo de pesquisa amostral

Uma cidade possui:

200 000

habitantes.

Uma pesquisa entrevista:

1200

moradores.

Como apenas parte da população foi investigada:

a pesquisa é amostral.

13 Censo versus amostra

	Censo	Pesquisa amostral
Unidades estudadas	todas	parte da população
Custo	geralmente maior	geralmente menor
Tempo	geralmente maior	geralmente menor
Erro amostral	não	pode existir
Viés de coleta	pode existir	pode existir
Aplicação	população acessível	população grande ou difícil

14 Vantagens de um censo

Um censo pode ser adequado quando:

- a população é pequena;
- todas as unidades são facilmente acessíveis;
- é necessário conhecer cada unidade;
- decisões individuais dependem dos dados coletados.

15 Limitações de um censo

Um censo pode apresentar:

- alto custo;
- grande necessidade de pessoal;
- maior tempo de coleta;
- dificuldade logística;
- erros de registro;
- ausência de respostas;
- desatualização se a coleta demorar muito.

16 Censo não significa ausência de erro

Atenção

Investigar toda a população elimina o **erro amostral**, pois não há amostragem. Entretanto, um censo ainda pode conter erros por:

- respostas incorretas;
- perguntas mal formuladas;
- registros duplicados;
- indivíduos não localizados;
- falhas na medição.

Portanto:

censo não significa dados automaticamente perfeitos.

17 Vantagens de uma pesquisa amostral

Uma pesquisa amostral pode:

- reduzir custos;
- reduzir tempo;
- permitir investigações mais detalhadas;
- ser realizada quando a população é muito grande;
- ser necessária quando a medição destrói ou inutiliza a unidade testada.

18 Exemplo: teste destrutivo

Uma fábrica deseja medir a resistência máxima de peças metálicas.

O teste exige levar cada peça até a ruptura.

Testar todas as peças destruiria toda a produção.

Nesse caso, uma pesquisa por amostragem é muito mais apropriada.

19 Representatividade

Definição 19.1: Amostra representativa

Uma amostra é considerada **representativa** quando reproduz adequadamente características relevantes da população para o objetivo da pesquisa.

Uma amostra representativa não precisa ser uma cópia perfeita da população.

Ela deve evitar distorções sistemáticas importantes.

20 Exemplo de falta de representatividade

Uma escola possui estudantes de:

6°, 7°, 8°, 9°

anos.

Queremos estudar a opinião de toda a escola.

Se entrevistarmos apenas estudantes do 9° ano, a amostra pode não representar adequadamente os demais anos.

21 Tamanho não garante representatividade

Teorema 21.1: Princípio qualitativo da representatividade

Uma amostra numerosa pode produzir resultados ruins se o processo de seleção for sistematicamente enviesado.

Por exemplo, entrevistar:

1000

pessoas de um único bairro pode ser pior para representar uma cidade inteira do que uma amostra menor, porém bem distribuída.

22 Amostra grande versus amostra bem selecionada

Atenção

Não devemos confundir:

tamanho da amostra

com:

qualidade da amostragem.

Ambos são importantes.

Uma amostra grande selecionada de forma inadequada pode permanecer enviesada.

23 Amostragem

Definição 23.1: Amostragem

Amostragem é o processo utilizado para selecionar as unidades que formarão uma amostra.

Neste curso estudaremos principalmente:

amostragem aleatória simples,

amostragem sistemática

e:

amostragem estratificada.

24 Amostragem aleatória simples

Definição 24.1: Amostragem aleatória simples

Na **amostragem aleatória simples**, as unidades são selecionadas por um procedimento aleatório que dá condições equivalentes de seleção às unidades da população, segundo o plano estabelecido.

Um procedimento escolar típico é:

- 1) numerar os indivíduos;
- 2) utilizar um sorteio;
- 3) selecionar os números sorteados;
- 4) evitar substituições arbitrárias.

25 Exemplo de amostragem aleatória simples

Uma turma possui:

40

estudantes.

Queremos selecionar:

8

estudantes.

Numeramos:

$1, 2, \dots, 40$

e sorteamos 8 números diferentes.

Os respectivos estudantes formarão a amostra.

26 Por que utilizar aleatoriedade?

A aleatoriedade procura impedir que o pesquisador escolha deliberadamente indivíduos com determinadas características.

Assim, ela ajuda a reduzir:

viés de seleção.

27 Amostragem sistemática

Definição 27.1: Amostragem sistemática

Na **amostragem sistemática**, os elementos da população são organizados em uma lista e selecionados segundo um intervalo regular após a escolha de um ponto inicial.

28 Intervalo sistemático

Se uma população possui aproximadamente:

$$N$$

elementos e desejamos uma amostra de tamanho:

$$n,$$

podemos considerar o intervalo:

$$k \approx \frac{N}{n}.$$

29 Exemplo sistemático

Uma lista possui:

$$1000$$

nomes.

Desejamos selecionar:

$$100.$$

Então:

$$k = \frac{1000}{100} = 10.$$

Podemos escolher aleatoriamente um número inicial entre:

$$1$$

e:

10.

Suponha que o início seja:

7.

Selecionamos:

7, 17, 27, 37, ...

até completar a amostra.

30 Importância do ponto inicial aleatório

Atenção

Escolher sempre:

10, 20, 30, ...

sem aleatorizar o ponto inicial pode introduzir padrões indesejados.

Na amostragem sistemática, é recomendável que o primeiro elemento seja selecionado aleatoriamente dentro do primeiro intervalo.

31 Risco de periodicidade

Considere uma lista organizada em um padrão repetitivo.

Se o intervalo sistemático coincidir com esse padrão, a amostra pode ficar distorcida.

Por exemplo, se uma lista alterna sistematicamente tipos de indivíduos e o intervalo de seleção acompanha exatamente essa periodicidade, certos grupos podem ser super ou sub-representados.

32 Amostragem estratificada

Definição 32.1: Amostragem estratificada

Na **amostragem estratificada**, a população é dividida em grupos relativamente homogêneos, chamados **estratos**, e uma amostra é selecionada dentro de cada estrato.

Exemplos de estratos:

- séries escolares;
- regiões;

- turnos;
- setores de uma empresa.

33 Por que estratificar?

Quando a população possui grupos relevantes, selecionar indivíduos de cada grupo pode melhorar a representatividade.

Por exemplo, em uma escola com:

600

estudantes do ensino fundamental e:

400

do ensino médio, uma amostra que ignore completamente um dos grupos não representará a escola inteira.

34 Amostragem estratificada proporcional

Uma possibilidade é manter aproximadamente as proporções da população.

Se a população é composta por:

60%

do Grupo A e:

40%

do Grupo B,

uma amostra de:

100

pessoas pode conter:

60

do Grupo A e:

40

do Grupo B.

35 Fórmula proporcional

Se o estrato i possui:

$$N_i$$

indivíduos em uma população de tamanho:

$$N,$$

e queremos uma amostra de tamanho:

$$n,$$

uma alocação proporcional pode utilizar:

$$n_i \approx n \frac{N_i}{N}.$$

36 Exemplo estratificado

Uma escola possui:

$$300$$

estudantes pela manhã e:

$$200$$

à tarde.

Total:

$$N = 500.$$

Desejamos uma amostra de:

$$n = 50.$$

Turno da manhã:

$$n_1 = 50 \frac{300}{500}.$$

$$n_1 = 30.$$

Turno da tarde:

$$n_2 = 50 \frac{200}{500}.$$

$$n_2 = 20.$$

Portanto:

30 estudantes da manhã e 20 da tarde.

37 Seleção dentro dos estratos

Atenção

Dividir a população em estratos não basta.
É preciso também selecionar adequadamente os indivíduos dentro de cada estrato.
Por exemplo, pode-se realizar um sorteio aleatório simples dentro de cada grupo.

38 Comparação dos métodos

Método	Ideia principal
Aleatória simples	selecionar unidades por sorteio aleatório
Sistemática	selecionar a cada k posições após início aleatório
Estratificada	dividir em grupos relevantes e amostrar dentro de cada grupo

39 Amostragem por conveniência

Definição 39.1: Amostragem por conveniência

Na **amostragem por conveniência**, são escolhidos os indivíduos que são mais fáceis de acessar.

Exemplo:

entrevistar apenas as pessoas que passam perto do pesquisador.

40 Problema da conveniência

Esse método pode produzir forte viés.

Por exemplo, pesquisar hábitos de toda uma escola entrevistando apenas os estudantes que chegam uma hora antes das aulas pode favorecer um perfil específico.

41 Amostra voluntária

Em uma amostra voluntária, as próprias pessoas decidem participar.

Exemplo:

uma enquete aberta na internet.

Pessoas com opiniões mais fortes podem ter maior motivação para responder.

Portanto, o resultado pode não representar adequadamente toda a população.

42 Viés

Definição 42.1: Viés estatístico

Um **viés** é uma tendência sistemática que favorece determinados resultados ou produz distorção na coleta, seleção ou interpretação dos dados.

Um viés não é simplesmente uma pequena diferença aleatória.

Ele possui caráter:

sistemático.

43 Viés de seleção

Definição 43.1: Viés de seleção

Ocorre quando o processo de escolha da amostra favorece ou exclui sistematicamente certos grupos da população.

Exemplo:

Para estudar a opinião de todos os estudantes sobre a merenda, entrevistar apenas aqueles que compram alimentos fora da escola.

44 Exemplo de viés de seleção

Uma pesquisa pergunta:

“Você gosta de praticar esportes?”

e é realizada apenas na entrada de uma academia.

A amostra provavelmente superrepresentará pessoas interessadas em atividade física.

Logo, não é adequada para estimar a opinião de toda a cidade.

45 Viés de não resposta

Definição 45.1: Não resposta

Ocorre quando indivíduos selecionados para a pesquisa não respondem, e os não respondentes diferem sistematicamente dos respondentes em aspectos relevantes.

46 Por que a não resposta pode ser problemática?

Suponha que:

1000

pessoas sejam selecionadas.

Apenas:

150

respondem.

Mesmo que a seleção inicial tenha sido boa, os 150 respondentes podem formar um grupo diferente dos 850 que não responderam.

47 Viés de resposta

Definição 47.1: Viés de resposta

O viés de resposta ocorre quando as respostas fornecidas não refletem adequadamente a informação desejada, devido à forma da pergunta, constrangimento, memória, pressão ou outras influências.

48 Pergunta tendenciosa

Considere:

“Você concorda que a excelente nova biblioteca deve continuar recebendo investimentos?”

A palavra:

“excelente”

pode influenciar a resposta.

Uma formulação mais neutra seria:

“Você é favorável à continuidade dos investimentos na nova biblioteca?”

49 Perguntas duplas

Considere:

“Você está satisfeito com a merenda e com o horário das aulas?”

Uma pessoa pode gostar da merenda e não gostar do horário.

Logo, a pergunta mistura duas questões diferentes.

É melhor separá-las.

50 Perguntas ambíguas

Considere:

“Você pratica exercícios regularmente?”

O que significa:

regularmente?

Uma formulação mais precisa poderia ser:

“Em quantos dias da última semana você praticou pelo menos 30 minutos de atividade física?”

51 Ordem das perguntas

A ordem das questões também pode influenciar respostas.

Perguntas anteriores podem alterar a maneira como o participante interpreta uma questão posterior.

Assim, o planejamento do questionário deve considerar:

clareza, neutralidade e ordem.

52 Quadro de amostragem

Definição 52.1: Quadro de amostragem

O **quadro de amostragem** é a lista ou sistema utilizado para identificar as unidades que podem ser selecionadas.

Exemplo:

lista atualizada de estudantes matriculados.

53 Problema de cobertura

Se o quadro de amostragem deixa de fora parte importante da população, pode surgir:

viés de cobertura.

Exemplo:

Realizar uma pesquisa sobre todos os moradores utilizando uma lista que contém apenas pessoas com telefone fixo.

54 Erro amostral

Definição 54.1: Erro amostral

O **erro amostral** é a diferença que pode ocorrer entre uma medida calculada na amostra e o valor correspondente na população devido ao fato de apenas parte da população ter sido observada.

55 Erro amostral não significa erro do pesquisador

Observação

Mesmo uma amostra bem escolhida pode não reproduzir exatamente a população. Essa diferença natural é chamada erro amostral. Ela não é necessariamente uma falha.

56 Exemplo intuitivo

Suponha que exatamente:

50%

dos estudantes de uma escola prefiram a opção A.

Uma amostra aleatória de 20 estudantes pode produzir:

55%

ou:

45%.

Outra amostra poderia produzir valor diferente.

Essa variação decorre do acaso na seleção.

57 Tamanho da amostra e variabilidade

De modo geral, mantendo um bom plano de amostragem, amostras maiores tendem a fornecer estimativas mais estáveis.

Entretanto:

aumentar o tamanho não corrige automaticamente um viés.

58 Exemplo crucial

Imagine uma cidade com:

1 000 000

de habitantes.

Uma enquete pergunta a opinião de:

100 000

seguidores de uma única página temática na internet.

O tamanho da amostra é enorme.

Ainda assim, ela pode representar mal a cidade, porque:

a seleção não foi adequada.

59 Parâmetro e estatística – aprofundamento

Definição 59.1: Parâmetro populacional

Um **parâmetro** é uma medida numérica que descreve uma característica da população.

Definição 59.2: Estatística amostral

Uma **estatística amostral** é uma medida calculada a partir dos elementos da amostra.

60 Exemplo

A média de altura de:

todos os estudantes da escola

é um parâmetro populacional.

A média calculada em:

uma amostra de 100 estudantes

é uma estatística amostral.

Utilizamos a segunda para estimar a primeira.

61 Inferência estatística – ideia

O processo de utilizar informações de uma amostra para fazer afirmações sobre a população é chamado, em níveis mais avançados:

inferência estatística.

No 8º ano, o ponto essencial é compreender que essa passagem exige:

uma amostra adequadamente selecionada.

62 Exemplo completo: pesquisa escolar

Uma escola possui:

1000

estudantes distribuídos da seguinte forma:

6° ano : 250,

7° ano : 250,

8° ano : 300,

9° ano : 200.

Desejamos entrevistar:

100

estudantes.

63 Estratificação proporcional

Calculamos:

$$100 \cdot \frac{250}{1000} = 25.$$

Logo, para o 6° ano:

25.

Igualmente para o 7°:

25.

Para o 8°:

$$100 \cdot \frac{300}{1000} = 30.$$

Para o 9º:

$$100 \cdot \frac{200}{1000} = 20.$$

Assim:

$$25 + 25 + 30 + 20 = 100.$$

64 Seleção dentro de cada série

Depois de determinar o número por série, realizamos uma seleção aleatória dentro de cada grupo.

Isso combina:

estratificação

com:

aleatoriedade.

65 Pesquisa censitária ou amostral?

A decisão depende do contexto.

66 Caso 1: turma pequena

Queremos conhecer o lanche preferido de uma turma com:

25

estudantes.

Todos estão presentes.

É simples consultar os 25.

Logo, um:

censo

pode ser adequado.

67 Caso 2: população nacional

Queremos estimar a opinião de milhões de pessoas.

Consultar todas seria caro e demorado.

Uma:

pesquisa amostral

bem planejada tende a ser mais prática.

68 Caso 3: teste destrutivo

Queremos avaliar a resistência de um lote de peças, mas o teste destrói a peça.

Um censo destruiria todo o lote.

Logo:

amostragem.

69 Caso 4: atualização cadastral individual

Uma escola precisa verificar os dados cadastrais de cada estudante.

Como a informação de cada indivíduo é necessária:

censo.

70 Planejamento de uma pesquisa

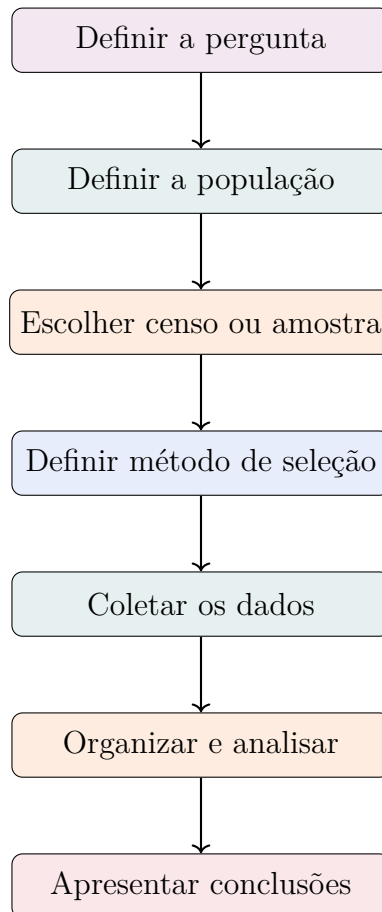
Uma pesquisa estatística deve começar pela pergunta, e não pelo gráfico.

Uma sequência adequada é:

- 1) definir o problema;
- 2) definir a população;
- 3) definir as variáveis;
- 4) escolher censo ou amostra;
- 5) definir o método de seleção;
- 6) elaborar perguntas;
- 7) coletar;

- 8) organizar;
- 9) analisar;
- 10) concluir.

71 Fluxograma do planejamento



72 Pergunta estatística

Definição 72.1: Pergunta estatística

Uma pergunta estatística é uma pergunta para a qual se espera variabilidade nas respostas ou nos dados coletados.

Por exemplo:

Qual é o tempo de deslocamento dos estudantes até a escola?

Espera-se que diferentes estudantes apresentem tempos diferentes.

73 Exemplo de pergunta não estatística

A pergunta:

Quantos estudantes estão matriculados hoje nesta turma?

pode ter uma resposta única diretamente registrada.

Não envolve necessariamente investigar uma distribuição de valores.

74 Variável da pesquisa

Se perguntamos:

“Quantos minutos você leva para chegar à escola?”

a variável é:

tempo de deslocamento.

Se perguntamos:

“Qual meio de transporte você utiliza?”

a variável é:

meio de transporte.

75 População-alvo

Definição 75.1: População-alvo

A **população-alvo** é o conjunto sobre o qual desejamos efetivamente tirar conclusões.

É importante defini-la com precisão.

Compare:

estudantes

com:

estudantes do 8º ano da Escola X matriculados em 2026.

A segunda formulação é muito mais precisa.

76 População acessível

Em algumas pesquisas, a população-alvo é maior do que o conjunto efetivamente acessível ao pesquisador.

Essa diferença pode criar limitações.

Por exemplo, desejar estudar:

todos os adolescentes de uma cidade

mas entrevistar apenas:

adolescentes matriculados em uma determinada escola.

77 Generalização indevida

Atenção

As conclusões devem respeitar a população representada pelos dados.

Uma pesquisa realizada em uma única escola não permite automaticamente afirmar:

“todos os adolescentes da cidade pensam assim”.

Essa conclusão ultrapassaria o alcance da amostra.

78 Proporção amostral

Suponha uma amostra com:

$$n = 200.$$

Desses:

$$130$$

preferem a opção A.

A proporção observada é:

$$\frac{130}{200} = 0,65.$$

Em porcentagem:

$$65\%.$$

Esse número descreve diretamente:

a amostra.

Podemos utilizá-lo como estimativa para a população apenas com cautela e um bom desenho amostral.

79 Resultado da amostra não é resultado exato da população

Atenção

Se:

65%

dos entrevistados escolheram uma opção, não devemos concluir automaticamente:

exatamente 65% de toda a população.

O valor amostral é uma estimativa.

80 Viés versus erro amostral

É importante distinguir:

erro amostral

e:

viés.

O erro amostral decorre da variabilidade natural de selecionar apenas parte da população.

O viés é uma distorção sistemática.

81 Comparação

Conceito	Descrição
Erro amostral	diferença natural entre amostra e população
Viés de seleção	certos grupos têm seleção inadequadamente favorecida/excluída
Viés de resposta	respostas são influenciadas ou registradas inadequadamente
Não resposta	pessoas selecionadas não participam e podem diferir das demais
Cobertura	quadro de seleção não representa toda a população-alvo

82 Exemplo comparativo

Pesquisa 1:

500

pessoas escolhidas aleatoriamente.

Pesquisa 2:

5000

pessoas que decidiram voluntariamente responder a uma enquete de um perfil especializado.

Não podemos concluir que a Pesquisa 2 seja automaticamente melhor apenas porque:

$$5000 > 500.$$

O método de seleção pode ser mais importante para evitar viés.

83 Amostragem estratificada e proporcionalidade

Considere três bairros:

A : 5000 habitantes,

B : 3000,

$$C : 2000.$$

Total:

$$N = 10\,000.$$

Desejamos:

$$n = 500.$$

84 Cálculo por estrato

Bairro A:

$$500 \frac{5000}{10\,000} = 250.$$

Bairro B:

$$500 \frac{3000}{10\,000} = 150.$$

Bairro C:

$$500 \frac{2000}{10\,000} = 100.$$

Verificação:

$$250 + 150 + 100 = 500.$$

Logo:

250,	150,	100.
------	------	------

85 Arredondamentos em estratificação

Nem sempre:

$$n \frac{N_i}{N}$$

resulta em número inteiro.

Nesse caso, é necessário realizar arredondamentos de modo que:

$$\sum_i n_i = n.$$

O ajuste deve preservar, tanto quanto possível, as proporções da população.

86 Exemplo de arredondamento

Suponha três estratos com proporções:

33%, 33%, 34%.

Queremos uma amostra de:

10.

Os valores exatos seriam:

3,3, 3,3, 3,4.

Uma distribuição possível é:

3, 3, 4.

87 Escolha do método

Notação

Uma regra qualitativa útil:

- população homogênea e lista disponível:

aleatória simples;

- lista ordenada e grande:

sistemática;

- população com grupos relevantes:

estratificada.

88 Não existe método universal

Observação

A escolha da técnica depende da:

- população;
- disponibilidade de listas;
- custos;
- objetivo;
- existência de grupos importantes;
- logística de coleta.

Não existe um método único adequado para todas as pesquisas.

89 Boas práticas para questionários

Uma boa pergunta deve ser:

clara,

específica,

neutra

e:

compatível com a informação desejada.

90 Exemplo de reformulação

Pergunta inadequada:

“Você não acha que deveríamos melhorar a péssima cantina?”

Problemas:

- formulação negativa;
- julgamento prévio;

- indução de resposta.

Uma opção melhor:

“Como você avalia a cantina da escola?”

com alternativas equilibradas.

91 Categorias exaustivas

Em perguntas fechadas, as opções devem cobrir adequadamente as respostas possíveis.

Por exemplo:

Como você chega à escola?

Alternativas:

- a pé;
- bicicleta;
- transporte público;
- carro;
- transporte escolar;
- outro.

A categoria:

outro

pode evitar excluir respostas legítimas.

92 Categorias mutuamente exclusivas

Idealmente, cada respondente deve saber em qual alternativa se encaixar.

Por exemplo, classes de idade:

10–12,

12–14

são ambíguas para quem possui exatamente:

12.

É melhor utilizar convenção clara, como:

$[10, 12)$, $[12, 14)$.

93 Pesquisa piloto

Definição 93.1: Pesquisa piloto

Uma **pesquisa piloto** é uma aplicação preliminar do instrumento de pesquisa em um grupo pequeno para identificar problemas antes da coleta principal.

Ela pode revelar:

- perguntas ambíguas;
- opções ausentes;
- questionário excessivamente longo;
- dificuldades de interpretação.

94 Ética e privacidade

Pesquisas envolvendo pessoas devem tratar os dados de forma responsável.

Devemos considerar:

- necessidade da informação;
- privacidade;
- anonimização quando apropriada;
- clareza sobre a finalidade da coleta;
- respeito aos participantes.

95 Dados pessoais desnecessários

Atenção

Não devemos coletar informações pessoais apenas porque podem ser coletadas.
Uma boa pesquisa coleta:

apenas os dados necessários para responder à questão proposta.

96 Estratégia geral

Notação

Para analisar ou planejar uma pesquisa:

- 1) identifique a pergunta;
- 2) identifique a população-alvo;
- 3) identifique a unidade estatística;
- 4) verifique se será utilizado censo ou amostra;
- 5) se houver amostra, determine o método de seleção;
- 6) avalie a representatividade;
- 7) procure possíveis vieses;
- 8) avalie clareza e neutralidade das perguntas;
- 9) organize os dados;
- 10) apresente conclusões compatíveis com a população efetivamente estudada.

97 Erros comuns

Atenção

Erro 1: confundir população com amostra.

População:

conjunto total de interesse.

Amostra:

parte selecionada da população.

Atenção

Erro 2: achar que uma amostra grande é necessariamente boa.

Uma amostra de grande tamanho pode apresentar forte viés de seleção.

Atenção

Erro 3: chamar qualquer grupo entrevistado de amostra representativa.
Representatividade depende do processo de seleção e da composição da população.

Atenção

Erro 4: concluir que um censo não contém erros.
O censo elimina o erro amostral, mas não elimina:

erros de medição, resposta, cobertura ou registro.

Atenção

Erro 5: realizar sorteio apenas entre indivíduos facilmente acessíveis.
A aleatoriedade deve ocorrer dentro de um quadro que represente adequadamente a população-alvo.

Atenção

Erro 6: utilizar amostragem sistemática sem observar possíveis padrões na lista.
Uma periodicidade na organização dos indivíduos pode distorcer a amostra.

Atenção

Erro 7: estratificar, mas escolher os indivíduos de cada estrato por conveniência.
A estratificação não elimina automaticamente vieses internos aos grupos.

Atenção

Erro 8: generalizar além da população estudada.
Resultados obtidos em uma escola não representam automaticamente todas as escolas.

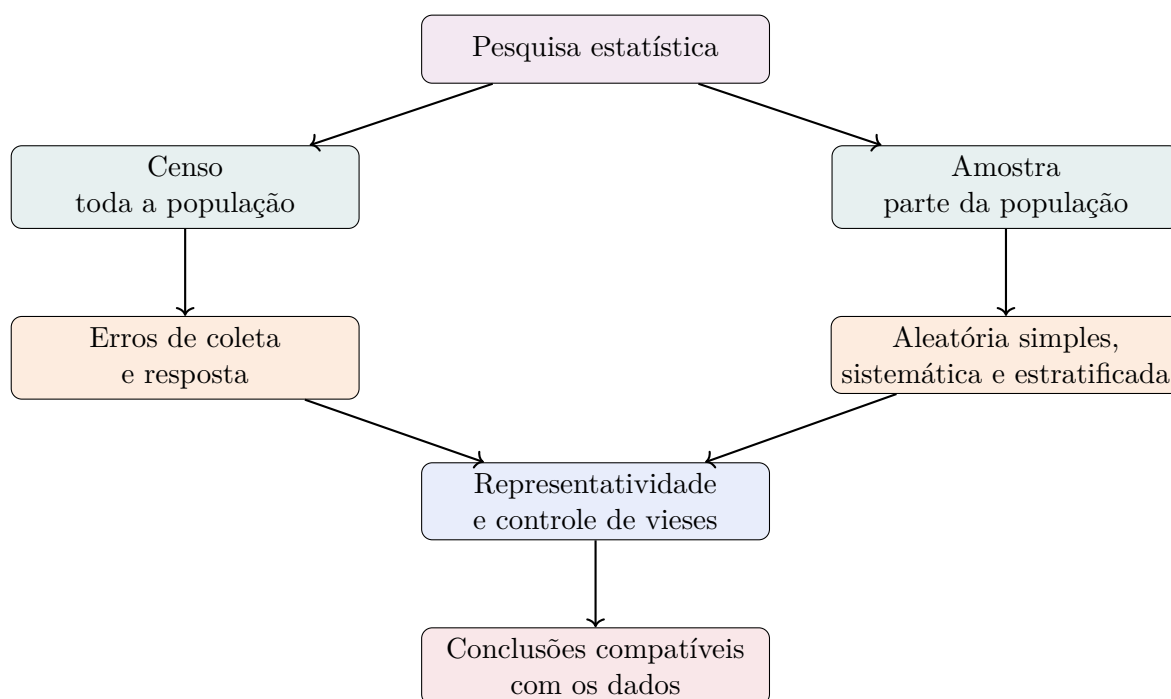
Atenção

Erro 9: formular perguntas tendenciosas.
A própria linguagem do questionário pode alterar as respostas.

Atenção

Erro 10: confundir variação aleatória com viés.
O erro amostral é uma variação natural.
O viés é uma distorção sistemática.

98 Mapa conceitual



99 Resumo teórico

Conceito	Definição resumida
População	conjunto total de interesse
Amostra	subconjunto da população
Censo	investigação de toda a população
Pesquisa amostral	investigação de parte da população
Aleatória simples	seleção por sorteio aleatório
Sistemática	seleção periódica após início aleatório
Estratificada	amostragem dentro de grupos relevantes
Viés	distorção sistemática
Erro amostral	diferença natural causada pela amostragem

Notação

Tamanhos:

N = tamanho da população

$$n = \text{tamanho da amostra.}$$

Na amostragem sistemática:

$$k \approx \frac{N}{n}.$$

Na estratificação proporcional:

$$n_i \approx n \frac{N_i}{N}.$$

Princípio fundamental:

uma boa amostra deve ser selecionada de forma compatível com a população que pretende represent.

100 Exercícios

100.1 Parte A – População e amostra

Exercício

Identifique a população e a amostra.

- a) Uma escola possui 900 estudantes e 120 são entrevistados sobre transporte.
- b) Uma fábrica produz 5000 peças e 100 são testadas.
- c) Uma cidade possui 80 000 moradores e 800 respondem a uma pesquisa selecionada aleatoriamente.

100.2 Parte B – Censo ou amostra

Exercício

Classifique cada situação como censitária ou amostral.

- a) Todos os 28 estudantes de uma turma respondem a uma pesquisa.
- b) 500 eleitores são entrevistados em uma cidade com 120 000 eleitores.
- c) Todas as peças de um pequeno lote de 30 unidades são inspecionadas.
- d) 200 produtos são analisados em uma produção de 20 000.

100.3 Parte C – Escolha entre censo e amostra

Exercício

Indique qual estratégia parece mais adequada e justifique.

- a) atualizar o telefone de cada funcionário de uma empresa com 40 empregados;
- b) estimar a opinião de milhões de consumidores;
- c) testar resistência de produtos quando o teste destrói o objeto;
- d) descobrir a preferência de todos os 20 alunos de uma pequena turma presente em sala.

100.4 Parte D – Aleatória simples

Exercício

Uma turma possui:

35

estudantes.

Deseja-se selecionar:

7.

Descreva um procedimento de amostragem aleatória simples.

100.5 Parte E – Sistemática

Exercício

Uma lista contém:

600

nomes.

Desejamos uma amostra de:

60.

- a) determine o intervalo k ;
- b) se o primeiro número sorteado for 4, escreva os cinco primeiros selecionados.

100.6 Parte F – Estratificada

Exercício

Uma escola possui:

400

estudantes pela manhã e:

600

à tarde.

Deseja-se uma amostra proporcional de:

100.

Quantos estudantes devem ser selecionados em cada turno?

100.7 Parte G – Mais estratificação

Exercício

Uma população possui:

$$A = 500, \quad B = 300, \quad C = 200.$$

Deseja-se uma amostra proporcional de:

200.

Determine o número de elementos de cada estrato.

100.8 Parte H – Representatividade

Exercício

Uma escola possui estudantes de quatro séries.

Uma pesquisa sobre toda a escola entrevista apenas alunos de uma das séries.

Explique o problema dessa amostra.

100.9 Parte I – Viés de seleção

Exercício

Identifique o possível viés.

- a) Uma pesquisa sobre prática esportiva da população é feita apenas em uma academia.
- b) Uma pesquisa sobre uso de transporte público é realizada apenas dentro de uma

estação de metrô.

- c) Uma enquete escolar é realizada somente entre estudantes que participam do grêmio.

100.10 Parte J – Perguntas tendenciosas

Exercício

Explique o problema e proponha uma formulação mais neutra.

- a) “Você concorda com a excelente proposta de ampliar a biblioteca?”
b) “Você não acha que a cantina cobra preços absurdos?”
c) “Você está satisfeito com o transporte e com o horário escolar?”

100.11 Parte K – Não resposta

Exercício

Uma pesquisa seleciona:

1000

pessoas, mas apenas:

180

respondem.

Explique por que isso pode comprometer a representatividade mesmo que a seleção inicial tenha sido aleatória.

100.12 Parte L – Tamanho versus qualidade

Exercício

Compare:

- Pesquisa A: 500 pessoas selecionadas aleatoriamente;
- Pesquisa B: 10 000 respostas voluntárias em uma página especializada.

É possível afirmar apenas pelo tamanho que a Pesquisa B representa melhor a população? Justifique.

100.13 Parte M – Proporção na amostra

Exercício

Uma amostra possui:

250

pessoas.

Dessas:

160

preferem a opção A.

Determine:

- a) a proporção amostral;
- b) a porcentagem;
- c) explique por que esse valor não deve ser automaticamente tratado como valor exato da população.

100.14 Parte N – Método adequado

Exercício

Associe cada situação a um método que pode ser adequado:

aleatória simples,

sistemática,

estratificada.

- a) lista completa de 200 pessoas sem grupos relevantes;
- b) população dividida entre regiões muito diferentes;
- c) lista de 5000 registros da qual queremos selecionar aproximadamente um em cada 50.

100.15 Parte O – Desafio

Exercício

Uma escola possui:

$6^\circ : 180,$

7° : 220,

8° : 260,

9° : 140.

Deseja-se uma amostra estratificada proporcional de:

80

estudantes.

Determine quantos estudantes devem ser selecionados de cada série.

101 Gabarito comentado

101.1 Parte A

Resolução

a)

População:

900 estudantes.

Amostra:

120 estudantes.

b)

População:

5000 peças.

Amostra:

100 peças.

c)

População:

80 000 moradores.

Amostra:

800 moradores.

101.2 Parte B

Resolução

- a) todos os 28: censitária ;
- b) 500 de 120 000: amostral ;
- c) todas as 30: censitária ;
- d) 200 de 20 000: amostral .

101.3 Parte C

Resolução

- a) Como os dados de cada funcionário são necessários e a população é pequena: censo .
- b) População muito grande: amostra bem planejada .
- c) Teste destrutivo: amostra .
- d) Turma pequena e totalmente acessível: censo .

101.4 Parte D

Resolução

Um procedimento possível:

- 1) numerar os estudantes de 1 a 35;
- 2) sortear 7 números distintos utilizando um mecanismo aleatório;
- 3) selecionar os estudantes correspondentes;
- 4) não substituir arbitrariamente um sorteado por outro.

Isso constitui uma:

amostragem aleatória simples.

101.5 Parte E

Resolução

$$k = \frac{600}{60}.$$

Logo:

$$k = 10.$$

Se o início é:

$$4,$$

os selecionados começam por:

$$4, \quad 14, \quad 24, \quad 34, \quad 44.$$

Logo:

$$4, \quad 14, \quad 24, \quad 34, \quad 44.$$

101.6 Parte F

Resolução

Total:

$$N = 400 + 600 = 1000.$$

Manhã:

$$100 \frac{400}{1000} = 40.$$

Tarde:

$$100 \frac{600}{1000} = 60.$$

Logo:

$$40 \text{ da manhã e } 60 \text{ da tarde.}$$

101.7 Parte G

Resolução

Total:

$$N = 500 + 300 + 200 = 1000.$$

Estrato A:

$$200 \frac{500}{1000} = 100.$$

Estrato B:

$$200 \frac{300}{1000} = 60.$$

Estrato C:

$$200 \frac{200}{1000} = 40.$$

Logo:

100, 60, 40.

101.8 Parte H

Resolução

A amostra exclui sistematicamente estudantes das outras séries.

Assim, ela pode não reproduzir as diferenças existentes entre os grupos.

Existe risco de:

viés de seleção

e falta de:

representatividade.

101.9 Parte I

Resolução

a)

Frequentadores de academia tendem a praticar mais atividades físicas que a população em geral.

viés de seleção.

b)

Pessoas presentes no metrô tendem a utilizar transporte público com maior frequência.

viés de seleção.

c)

Participantes do grêmio podem ter perfil de participação e opinião diferente dos demais estudantes.

viés de seleção.

101.10 Parte J

Resolução

a)

A palavra:

“excelente”

induz avaliação positiva.

Uma formulação possível:

“Você é favorável à ampliação da biblioteca?”

b)

A expressão:

“preços absurdos”

induz uma resposta crítica.

Uma opção:

“Como você avalia os preços praticados pela cantina?”

c)

A pergunta combina dois temas diferentes.

Melhor separar:

“Você está satisfeito com o transporte?”

e:

“Você está satisfeito com o horário escolar?”

101.11 Parte K

Resolução

Mesmo que os 1000 tenham sido selecionados corretamente, os 180 que responderam podem ter características diferentes dos 820 que não responderam.

Logo, pode ocorrer:

viés de não resposta.

101.12 Parte L

Resolução

Não.

O fato de:

$$10\,000 > 500$$

não garante melhor representatividade.

A Pesquisa B é voluntária e realizada em uma página especializada, podendo apresentar forte viés.

A Pesquisa A, embora menor, possui seleção aleatória.

Portanto:

o método de seleção é essencial, não apenas o tamanho.

101.13 Parte M

Resolução

Proporção:

$$\frac{160}{250}.$$

Simplificando:

$$\frac{16}{25} = 0,64.$$

Logo:

0,64.

Porcentagem:

64%.

Esse valor descreve diretamente a amostra.

Devido ao erro amostral e à possibilidade de outros problemas de seleção, não podemos afirmar automaticamente que:

exatamente 64% da população prefere A.

101.14 Parte N

Resolução

a) lista completa, sem estratos importantes:

aleatória simples;

b) regiões diferentes:

estratificada;

c) selecionar um em cada 50:

sistemática.

101.15 Parte O

Resolução

Total:

$$N = 180 + 220 + 260 + 140.$$

$$N = 800.$$

Amostra:

$$n = 80.$$

Para o 6°:

$$80 \frac{180}{800} = 18.$$

Para o 7°:

$$80 \frac{220}{800} = 22.$$

Para o 8°:

$$80 \frac{260}{800} = 26.$$

Para o 9°:

$$80 \frac{140}{800} = 14.$$

Verificação:

$$18 + 22 + 26 + 14 = 80.$$

Logo:

18,	22,	26,	14.
-----	-----	-----	-----

102 Avaliação final

Avaliação final – Pesquisa censitária e amostral

Resolva.

1. Defina:

- a) população;
- b) amostra;
- c) censo.

2. Uma universidade possui:

6000

estudantes.

Uma pesquisa entrevista:

450.

Identifique:

- a) população;
- b) amostra;
- c) N ;
- d) n .

3. Explique uma vantagem e uma limitação de uma pesquisa censitária.

4. Explique por que uma amostra muito grande pode continuar sendo inadequada.

5. Uma lista possui:

1200

pessoas.

Desejamos uma amostra sistemática de:

120.

Determine:

k .

Se o início aleatório for 6, escreva os quatro primeiros selecionados.

6. Uma escola possui:

700

estudantes no turno da manhã e:

300

à tarde.

Deseja-se uma amostra estratificada proporcional de:

200.

Determine quantos serão selecionados em cada turno.

7. Uma pesquisa sobre hábitos alimentares de toda uma cidade é feita apenas entre clientes de uma loja de produtos naturais.

Identifique e explique o problema.

8. Explique a diferença entre:

erro amostral

e:

viés.

9. Analise a pergunta:

“Você concorda com o maravilhoso novo projeto da escola?”

Explique o problema e reescreva-a de forma mais neutra.

10. Uma pesquisa selecionou aleatoriamente:

800

pessoas, mas apenas:

120

responderam.

Explique por que o resultado deve ser interpretado com cautela.

102.1 Gabarito da avaliação final

Resolução

1.
 - a) População é o conjunto total de unidades sobre o qual se deseja obter informações.
 - b) Amostra é um subconjunto selecionado da população.
 - c) Censo é uma pesquisa que investiga todas as unidades da população.
2. População:

os 6000 estudantes.

Amostra:

os 450 entrevistados.

Logo:

$$N = 6000$$

e:

$$n = 450.$$

3. Uma vantagem é obter dados de toda a população e eliminar o erro decorrente da seleção de uma amostra.

Uma limitação possível é:

maior custo, tempo ou dificuldade logística.

4. Porque representatividade depende do método de seleção. Uma amostra enorme pode excluir ou super-representar grupos sistematicamente.

5. :

$$k = \frac{1200}{120}.$$

$$k = 10.$$

Começando em:

6,

temos:

6, 16, 26, 36.

6. Total:

$$N = 1000.$$

Manhã:

$$200 \frac{700}{1000} = 140.$$

Tarde:

$$200 \frac{300}{1000} = 60.$$

Logo:

140 pela manhã e 60 à tarde.

7. Clientes de uma loja de produtos naturais podem ter hábitos alimentares diferentes da população da cidade.

Assim, existe:

viés de seleção.

A amostra pode não ser representativa.

8. Erro amostral é a diferença natural que ocorre porque observamos apenas parte da população.

Viés é uma distorção sistemática produzida pelo método de seleção, coleta, resposta ou cobertura.

9. A palavra:

“maravilhoso”

induz avaliação positiva.

Uma versão mais neutra seria:

“Você é favorável ao novo projeto da escola?”

10. Apesar de a seleção inicial ser aleatória, apenas uma pequena parte respondeu.

Se respondentes e não respondentes diferirem em aspectos relevantes, pode existir:

viés de não resposta.

Portanto, o resultado pode não representar adequadamente todos os selecionados nem a população.

103 Conclusão

Pesquisas censitárias e amostrais são ferramentas fundamentais para produzir informações estatísticas.

A escolha entre elas depende do objetivo, da população e das condições de coleta.

As principais ideias estudadas foram:

- 1) a população é o conjunto total de interesse;

- 2) uma amostra é uma parte da população;
- 3) um censo investiga toda a população;
- 4) uma pesquisa amostral investiga apenas uma parte;
- 5) amostras podem reduzir custo e tempo;
- 6) uma amostra deve ser selecionada de modo coerente com a população que pretende representar;
- 7) tamanho grande não garante representatividade;
- 8) a amostragem aleatória simples utiliza sorteio;
- 9) a amostragem sistemática utiliza intervalos regulares após um início aleatório;
- 10) a amostragem estratificada assegura representação de grupos relevantes;
- 11) amostras por conveniência podem produzir forte viés;
- 12) não resposta pode alterar a composição efetiva da amostra;
- 13) perguntas tendenciosas podem gerar viés de resposta;
- 14) erro amostral e viés são conceitos diferentes;
- 15) as conclusões devem respeitar a população efetivamente representada pelos dados.

Notação

Resumo fundamental:

N = tamanho da população

n = tamanho da amostra

censo $\implies$ toda a população

pesquisa amostral $\implies$ parte da população

Amostragem sistemática:

$$k \approx \frac{N}{n}$$

Estratificação proporcional:

$$n_i \approx n \frac{N_i}{N}.$$

Finalmente:

boa pergunta + boa seleção + boa coleta $\implies$ conclusões estatísticas mais confiáveis.

Fim do curso