

Trabalho 1 — Estatística (2025.1)

Objetivo do Trabalho

Neste trabalho, vocês assumirão o papel de pesquisadores responsáveis por elaborar um texto (semelhante a um artigo científico) com base em dados reais.

O objetivo é aplicar os conceitos estudados até aqui na disciplina, com foco em análise exploratória de dados. Espera-se que vocês elaborem perguntas relevantes sobre o tema e as respondam ao longo do texto, com base na análise do conjunto de dados fornecido.

Organização e Grupos

1. Cada aluno deve colocar seu nome e e-mail no arquivo de seleção de temas.
2. Os grupos terão tamanho variável, dependendo do tamanho do conjunto de dados.
3. Após o prazo de manifestação de interesse, enviarei por e-mail a cada grupo um conjunto de dados referente ao tema escolhido e o método de amostragem.
4. Para fins didáticos, trataremos esse conjunto de dados como a população de estudo.

Avaliação

O trabalho será avaliado com nota de 0 a 10. As situações abaixo acarretarão automaticamente nota 0:

- Entrega fora do prazo.
- Trabalho desorganizado ou mal estruturado.
- Não utilizar a linguagem R.

- Não enviar todos os arquivos solicitados.
- Não seguir o método de amostragem informado no e-mail.
- Não preencher o arquivo de seleção de temas com nome e e-mail no prazo indicado.
- Plágio.

Datas e Prazos

- Manifestação do tema de interesse: 22:00h de 25/04 até 30/04 00:00h.
- Entrega do trabalho: até 06/06 00:00h.
- A entrega deve ser feita no e-mail: ismael@dme.ufrj.br usando a seguinte estrutura:
 - Assunto: Trabalho de Estatística.
 - O corpo do texto deve conter os nomes dos integrantes e o título do trabalho.
 - Em anexo devem colocar os arquivos requisitados.
 - Apenas **um** dos membros do grupo deve fazer o envio.

Itens obrigatórios para entrega:

- Arquivo contendo a amostra utilizada nas análises.
- Arquivo qmd e pdf com os resultados e análises da pesquisa (Parte 1 do trabalho).
- Arquivo qmd e pdf contendo as respostas (Com os gráficos) da Parte 2 do trabalho (Pode ser entregue no mesmo arquivo do anterior).

Parte 1 — Análise Exploratória de Dados

O trabalho deve seguir a seguinte estrutura:

Título

Relacionado à principal pergunta da pesquisa. Pode ser simples, mas deve representar o objetivo do trabalho.

Capítulo 1 — Introdução

Apresente o tema, os objetivos da pesquisa, definição da pergunta principal que deseja responder através da pesquisa e perguntas adicionais caso verifique necessidade. Além disso, a Introdução pode conter qualquer outro aspecto que julgar relevante (como o tamanho da população, contexto da pesquisa, etc.).

Capítulo 2 — Descrição dos Dados

Liste as variáveis do conjunto de dados e explique o que cada uma representa. Classifique-as como:

- Quantitativa (Discreta ou Contínua) ou
- Qualitativa (Nominal ou Ordinal)

No caso de variáveis qualitativas, descreva os possíveis valores que a variável assume e seus significados.

Capítulo 3 — Metodologia

Descreva detalhadamente o método de amostragem adotado, evidenciando o tamanho amostral selecionado. Explícite também as vantagens e desvantagens do método de amostragem adotado.

Observação: Não é necessário explicar o motivo da seleção do tamanho amostral, haja vista que isso será discutido mais para o fim do curso.

Capítulo 4 — Análise Exploratória

Utilize gráficos e tabelas de frequência para analisar **todas** as variáveis do seu conjunto de dados, excetuando-se apenas variáveis como nome ou id. Você pode (e deve) utilizar os gráficos discutidos em aula: dispersão, barras, histogramas e boxplots.

Interprete os resultados de forma crítica e os use para responder as perguntas iniciais ou gerar novos insights.

Nesta seção também é interessante você analisar relações entre a variável de interesse e as demais variáveis, fazendo uso do gráfico de dispersão e do coeficiente de correlação.

Capítulo 5 — Conclusão

Discuta os principais achados da sua análise. Responda a pergunta principal e comente limitações que afetaram a análise.

Importante: A estrutura apresentada acima é o mínimo exigido. A exploração de mais elementos de maneira criativa será positivamente avaliada.

Parte 2 — Tamanho Amostral e Propriedades da Média Amostral

Exercício 1

- a) Escolha uma variável quantitativa da sua população.
- b) Calcule a média populacional e a armazene na variável `media_populacional`.
- c) Crie dois conjuntos de dados (vetores) vazios chamados `media_amostrai` e `variancia_media_amostrai`.
- d) Considerando n como sendo cada uma das seguintes proporções do tamanho da população (% do tamanho da população):
2%, 4%, 6%, ..., 98%, realize:
 - Extraia uma amostra de tamanho n via Amostragem Aleatória Simples.
 - Calcule a média amostral da variável selecionada.
 - Calcule a variância das médias amostrais.
 - Armazene a média amostral e a variância das médias amostrais nos vetores `media_amostrai` e `variancia_media_amostrai`, respectivamente.
- e) Construa um gráfico constituído dos seguintes elementos:
 - Gráfico de dispersão da `media_amostrai` vs tamanhos amostrais.
 - Linha horizontal vermelha em $y = \text{media_populacional}$.
- f) Construa outro gráfico composto pelos seguintes elementos:
 - Gráfico de dispersão da `variancia_media_amostrai` vs tamanhos amostrais.
 - Linha horizontal roxa em $y = 0$.
 - Obs: Ajuste seu gráfico para que a linha acima seja visível.
- g) Escreva um texto interpretando os gráficos.

Exercício 2

- a) Escolha (de preferência a mesma do Exercício 1) uma variável quantitativa da sua população.
- b) Calcule a média populacional e a armazene na variável `media_populacional`.
- c) Crie um conjunto de dados (vetor) vazio chamado `media_amstras`.
- d) Extraia 1.000 amostras de tamanho 40% do tamanho da população utilizando o método de Amostragem Aleatória Simples e calcule a média de cada uma, armazenando-as no vetor `media_amstras`.
- e) Crie um vetor `media_das_medias`, onde cada posição i armazena a média dos i primeiros valores de `media_amstras`.
- f) Construa um gráfico contendo:
 - Um gráfico de dispersão de `media_das_medias` vs número de amostras (1, 2, ..., 1000).
 - Linha horizontal vermelha em $y = \text{media_populacional}$.
- g) Escreva um texto interpretando o gráfico.