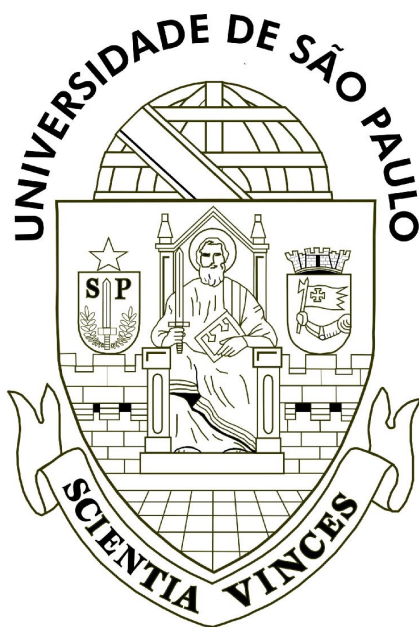


Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
SME0620 - Estatística I



Trabalho de Estatística

Alunos:

Lucas Lima Romero (13676325)

Marco Antonio Gaspar Garcia (11833581)

Professor: Vicente Garibay Cancho

Maio
2024

Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
SME0620 - Estatística I

Trabalho de Estatística

Trabalho de Pesquisa Estatística Descritiva

Alunos:

Lucas Lima Romero (13676325)

Marco Antonio Gaspar Garcia (11833581)

Professor: Vicente Garibay Cancho

*“As coisas que não têm
nome são mais pronunciadas
por crianças.”
(Manoel de Barros)*

Maio
2024

Sumário

1	Introdução	1
1.1	Motivação e Importância	1
1.2	Problema de Pesquisa	1
1.3	Objetivos	1
2	Metodologia	2
2.1	Coleta de Dados	2
2.2	Ferramentas Utilizadas	2
2.3	Procedimento de Análise	2
2.4	Categorização das Variáveis	2
3	Resultados	3
4	Conclusão	4
	Bibliografia	4

1 Introdução

1.1 Motivação e Importância

A motivação para o trabalho foi analisar possíveis fatores que influenciam na performance acadêmica dos estudantes. O desempenho acadêmico dos estudantes é um indicador crucial do sucesso educacional e pode influenciar significativamente as oportunidades futuras de um indivíduo. Entender os fatores que afetam esse desempenho é essencial para educadores, formuladores de políticas e a sociedade em geral.

Fatores como o nível educacional dos pais, o status socioeconômico representado pelo tipo de almoço que os alunos recebem, e a preparação para testes foram estudados para sua relação com o sucesso acadêmico. Além disso, as diferenças de desempenho baseadas em gênero e etnia também são de grande interesse, pois refletem as desigualdades subjacentes que podem existir dentro do sistema educacional.

1.2 Problema de Pesquisa

A pergunta que precisa ser respondida é como os fatores socioeconômicos e demográficos, como gênero, etnia, nível de educação dos pais e tipo de almoço, afetam o desempenho acadêmico dos estudantes em matemática, leitura e escrita?

1.3 Objetivos

O objetivo principal desta pesquisa é analisar o impacto de variáveis socioeconômicas e demográficas no desempenho acadêmico de estudantes em matemática, leitura e escrita. Especificamente, a pesquisa visa:

- Avaliar a relação entre o nível educacional dos pais e as pontuações acadêmicas dos estudantes.
- Estimar o efeito do tipo de almoço no desempenho acadêmico, como um indicador do status socioeconômico.
- Analisar o impacto da preparação para testes no desempenho em matemática, leitura e escrita.
- Examinar as diferenças de desempenho entre diferentes grupos étnicos e gêneros.

2 Metodologia

2.1 Coleta de Dados

Os dados utilizados neste estudo foram extraídos de um dataset encontrado no site *Kaggle*, divulgado pelo docente, contendo informações sobre o desempenho acadêmico de estudantes, incluindo suas pontuações em matemática, leitura e escrita, além de dados demográficos e socioeconômicos. Este dataset é composto por 1000 entradas e 8 variáveis.

2.2 Ferramentas Utilizadas

Para a análise dos dados, utilizaremos a linguagem de programação Python, juntamente com bibliotecas especializadas como Pandas para manipulação de dados, NumPy para operações numéricas e Matplotlib e Seaborn para visualização de dados. Todas as análises serão realizadas no ambiente Jupyter Notebook, que permite uma integração eficiente de código e texto, facilitando a visualização dos passos analíticos e resultados. Além disso, os dados foram compilados num arquivo em Excel, conforme solicitado.

2.3 Procedimento de Análise

- **Carregamento dos Dados:** Os dados serão carregados usando a biblioteca de manipulação de dados Pandas.
- **Limpeza e Preparação dos Dados:** Verificação de valores ausentes ou atípicos, e adequação dos tipos de dados para análise.
- **Análise Descritiva:** Cálculo de estatísticas descritivas para variáveis quantitativas e criação de tabelas de frequência para variáveis qualitativas.
- **Visualização dos Dados:** Geração de gráficos para uma melhor compreensão das distribuições e relações entre as variáveis, com as bibliotecas Matplotlib e Seaborn.

2.4 Categorização das Variáveis

Qualitativas Nominais:

- *gender*: Gênero do estudante.
- *race_ethnicity*: Grupo étnico do estudante.
- *lunch*: Tipo de almoço.
- *test_preparation_course*: Participação em curso de preparação.

Qualitativas Ordinais:

- *parental_level_of_education*: Nível de educação dos pais.

Quantitativas Discretas:

- *math_score*: Pontuação em matemática.
- *reading_score*: Pontuação em leitura.
- *writing_score*: Pontuação em escrita.

3 Resultados

As plotagens dos gráficos e análises sobre os resultados se encontram no Jupyter Notebook em anexo.

4 Conclusão

A análise dos dados sobre o desempenho acadêmico de estudantes em matemática, leitura e escrita revela várias intersecções importantes entre o desempenho acadêmico e variáveis socioeconômicas e demográficas. Primeiramente, é perceptível que existem variações significativas nas pontuações com base no grupo étnico e no nível de educação dos pais, o que sugere que fatores culturais e o ambiente educacional em casa podem desempenhar papéis cruciais no sucesso acadêmico dos alunos.

Além disso, o tipo de almoço, que serve como um indicador para o status socioeconômico, parece influenciar as pontuações dos alunos. Aqueles que recebem almoço gratuito ou a preço reduzido, indicativo de menor status socioeconômico, tendem a ter pontuações mais baixas em comparação com aqueles que recebem almoço padrão. Isso destaca a importância do suporte socioeconômico como um fator influente no desempenho educacional.

O curso de preparação para o teste também mostra uma relação clara com as pontuações obtidas, onde os estudantes que completaram esses cursos geralmente apresentaram melhores resultados. Isso sugere que intervenções direcionadas, como programas de preparação para exames, podem ser estratégias eficazes para melhorar o desempenho acadêmico.

Por fim, ao observar as diferenças de gênero, pode-se notar que as pontuações variam entre masculino e feminino em diferentes disciplinas, o que poderia indicar a influência de fatores culturais ou de ensino nas preferências e competências acadêmicas entre os gêneros.

Essas intuições podem ser utilizadas para formular políticas educacionais mais eficazes, focadas no fortalecimento das oportunidades educacionais para grupos desfavorecidos e na implementação de programas de apoio que possam ajudar todos os estudantes a alcançar seu potencial pleno. As autoridades educacionais e as escolas devem considerar essas variáveis ao desenvolver estratégias para melhorar o desempenho estudantil e reduzir as disparidades educacionais.

Bibliografia

Bussab, WO; Morettin, PA. Estatística Básica. São Paulo: Editora Saraiva, 2017 (9ª Edição). Magalhães, MN; Lima, ACP.