

O Trabalho

Produção de um artigo a partir da análise dos microdados dos desempenhos dos estudantes das Escolas de São Paulo no ENEM, utilizando técnicas de análise de regressão estudadas durante o curso.

- Uso do software R para ajuste de modelos.
- Redação em editor de texto.
- Trabalho em duplas.
- Entrega pelo “Google Sala de Aula” em 26/07/2024.

O objetivo do trabalho é propor um modelo de regressão linear múltipla para explicar o desempenho escolar dos alunos das escolas de São Paulo no ENEM 2015.

- Utilização dos microdados dos alunos das escolas de São Paulo:
Microdados Alunos SP ENEM 2015.RData.
- Análise focada na nota média do ENEM (`nota.media.enem`).

Sugestão de Roteiro

- ❶ Familiarize-se com a base de dados.
- ❷ Identifique e argumente sobre as variáveis relevantes e irrelevantes para explicar o desempenho escolar ^{1,2}.
- ❸ Realize uma análise descritiva das variáveis selecionadas.
- ❹ Reflita sobre os impactos de omitir variáveis relevantes e de incluir variáveis irrelevantes na escolha das variáveis.

1. Recomenda-se a leitura do Trabalho de TCC do Rony de Souza para compreensão da nova geração de indicadores educacionais, contidos na base.
2. Além do dicionário contido no arquivo `.RData`, também foi disponibilizado o dicionário da base original de microdados do Enem (*Dicionário Original Microdados Enem 2015.xlsx*).

Sugestão de Roteiro

- ① Tente ajustar vários modelos para explicar o desempenho escolar.
- ② Faça uma comparação entre os diferentes modelos ajustados. Analise:
 - ▶ Grau de magnitude dos coeficientes estimados.
 - ▶ Significância estatística.
 - ▶ Significância prática.
- ③ Faça uma interpretação dos coeficientes.
- ④ Comente sobre a qualidade do ajuste.
- ⑤ Faça uma análise de resíduos do **melhor modelo**.

Interpretação dos Coeficientes

Essa é a parte mais importante do trabalho, focada em obter boas estimativas dos efeitos de cada variável explicativa no desempenho escolar:

- Analise os coeficientes cuidadosamente. Eles têm o sinal esperado?
- Qual o grau de magnitude do efeito? Conheça a escala de cada variável numérica para entender o impacto dos coeficientes.
- Observe a forma funcional do modelo. Por exemplo, variáveis em forma logarítmica demandam interpretação específica.
- Reflita sobre o significado desses efeitos. Quais recomendações práticas podem ser derivadas a partir do modelo?

A base é grande (65 mil alunos), contexto em que é comum que todos os coeficientes tenham significância estatística. Nesse caso, é mais importante refletir sobre o grau de magnitude dos coeficientes – o efeito refletido pelo coeficiente é “grande” no sentido prático?

Escreva um texto em forma de artigo seguindo a estrutura sugerida:

- 1 Introdução.
- 2 Análise descritiva.
- 3 Aspectos teóricos do modelo utilizado.
- 4 Análise e interpretação do modelo ajustado.
- 5 Conclusões.
- 6 Referências bibliográficas.
- 7 Anexo com análise de resíduos do modelo escolhido.

Critérios de Avaliação

O trabalho será avaliado de acordo com os seguintes aspectos:

- Utilização do maior número de aspectos teóricos apresentados no curso.
- Aplicação adequada das técnicas e sua correta interpretação.
- Redação e organização geral do artigo.
- Domínio da linguagem técnica e dos significados dos termos estatísticos.
- Interpretação e conclusão dos resultados.

Não serão aceitos trabalhos após a data final estabelecida.

- É necessário entregar tanto o arquivo de texto quanto o arquivo de script do R contendo a sequência de comandos usados para gerar as análises e os modelos.
- Utilize tabelas e gráficos para enriquecer a apresentação dos dados:
 - ▶ Todas as tabelas e gráficos inseridos devem ser analisados e mencionados no texto do artigo.