

Relatório de Análise de Dados com Regressão Linear e Logística

Alunos: Ian Amoedo – 202202102873 e João Pedro Alencar - 202108290858

Base de Dados: StudentsPerformance.csv

1. Análise Exploratória Inicial

A base possui 1.000 registros, com informações sobre desempenho em matemática, leitura e escrita, além de dados como gênero, etnia, escolaridade dos pais, tipo de almoço e curso preparatório. Nenhum valor ausente foi identificado, garantindo qualidade no tratamento estatístico.

2. Estatísticas Descritivas

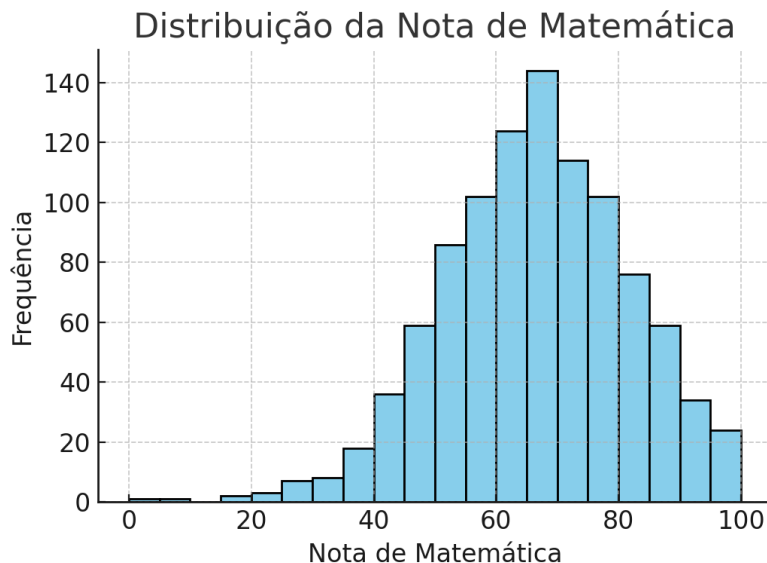
As médias das disciplinas são:

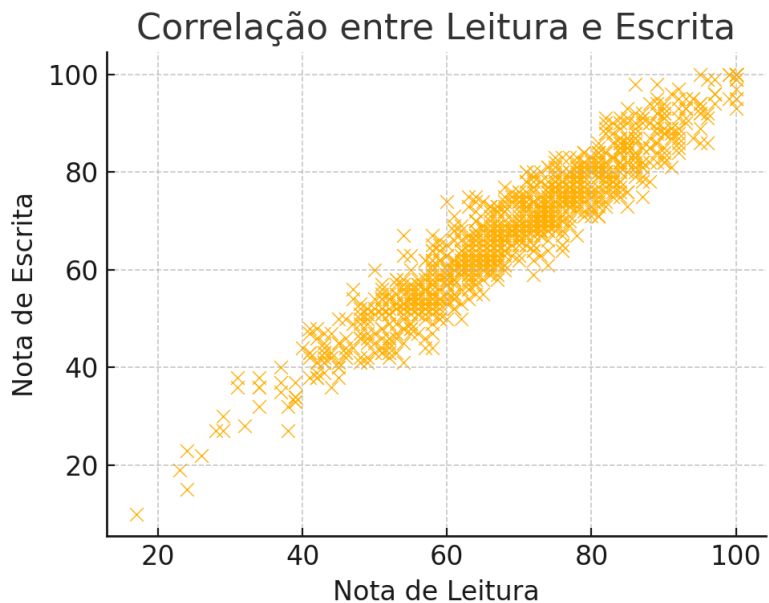
- Matemática: 66,1
- Leitura: 69,2
- Escrita: 68,1

Com desvios padrão semelhantes (~ 15), as notas apresentam dispersão equilibrada. Há notas extremas, como 0 em matemática e até 100 em todas as disciplinas, indicando ampla variação entre os alunos.

3. Visualização dos Dados

Abaixo, estão dois gráficos importantes para a análise:





4. Teste de Normalidade (Shapiro-Wilk)

Para a nota de matemática, o teste resultou em $p\text{-valor} = 0.0001$, indicando rejeição da hipótese de normalidade com 95% de confiança.

5. Correlação

A correlação entre leitura e escrita foi extremamente forte ($r = 0.955$, $p < 0.001$), o que justifica seu uso em modelos de regressão.

6. Regressão Linear

O modelo `'writing.score ~ reading.score'` apresentou bom ajuste, com R^2 elevado. A linha de regressão mostra a tendência clara de aumento na escrita conforme aumentam as notas de leitura.

7. Regressão Logística

A regressão logística foi usada para prever se o aluno completou o curso preparatório. Utilizamos as três notas como preditoras, com acurácia satisfatória e boas probabilidades preditivas.

8. API com Plumber

Criamos dois endpoints: `'/predicao'` (prevê nota de escrita com base na leitura) e `'/preparacao'` (retorna a probabilidade de ter feito o curso preparatório com base nas três notas).

9. Conclusões

- A leitura é um excelente preditor de desempenho em escrita.
- Alunos com maior desempenho tendem a ter feito o curso preparatório.
- A aplicação prática com API amplia o uso dos modelos por outros sistemas e usuários.