

Alcoolismo e a Vida Acadêmica



Universidade do Vale do Itajaí
Mestrado em Computação Aplicada — Kobrasol
Disciplina: Análise Exploratória de Dados
Professor: Wemerson Delcio Parreira

Aluno: Khalil G. Q. de Santana — khalil.santana@edu.univali.br

Problema

Dentre os maiores desafios enfrentados pelos estudantes está o *stress*^{[3][4]}, desta maneira estes acadêmicos buscam formas de lazer como passeios ou vídeo-games.

Dentre estas está o **consumo de bebidas alcoólicas** como cervejas e vinhos, contudo, seria tal atividade danosa para o desempenho acadêmico dos indivíduos?



Hipóteses

H_0 : “Não há uma correlação negativa entre o consumo alcoólico no dia-a-dia e a notas finais obtidas pelos indivíduos”

H_1 : “Há uma correlação negativa significativa entre o consumo alcoólico diário e notas finais obtidas.”



Materiais e Metodos

- Grau de significância (α) = 0,05
- Python 3.11
 - Bibliotecas:
 - Pandas
 - Numpy
 - Scipy
 - Matplotlib
 - Seaborn
- Google Collab



Banco de dados selecionado

O *dataset*^[1] selecionado foi obtido por meio da plataforma Kaggle, e o mesmo teve como origem um artigo^[2] publicado na Conferência de Tecnologia de Negócios do Futuro (FUBUTEC) em 2008.


- Questionário de 37 perguntas
- Aplicado a 788 alunos
- Em duas escolas de Portugal
- Durante 2005 e 2006.



Banco de dados selecionado

São ao todo ~110KiB de dados distribuídos em 33 colunas, detalhando diversos aspectos de cada estudante, como:

Sexo	Idade	Endereço	Tamanho da Família	Acesso à Internet
M / F	Decimal	Rural ou Urbano	≥ 3 ou < 3	Sim / Não

- 1044 linhas
 - Formato CSV
 - Anonimizado
- 

Análise Exploratória de Dados

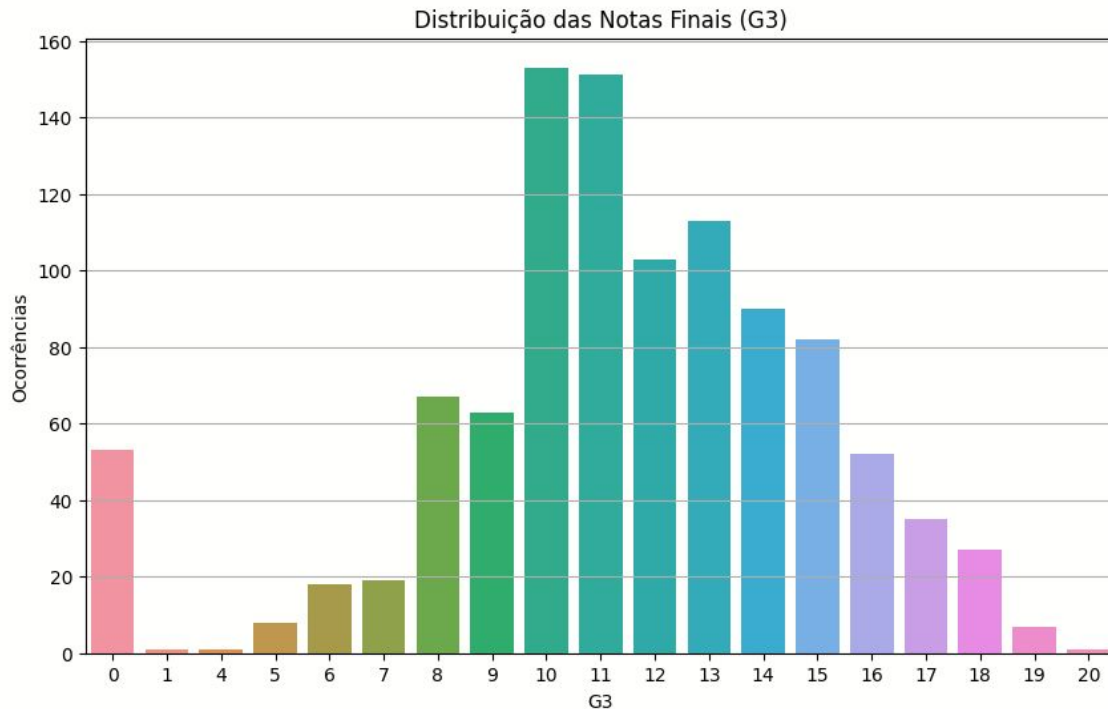
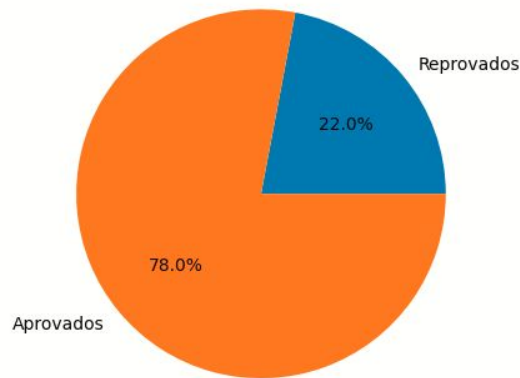
- Dataset já 'limpo'
 - Livre de linhas nulas
 - Dados consistentes
- Dividido entre curso de matemática e português
 - Unificado para os fins deste estudo



Distribuição das Notas Finais

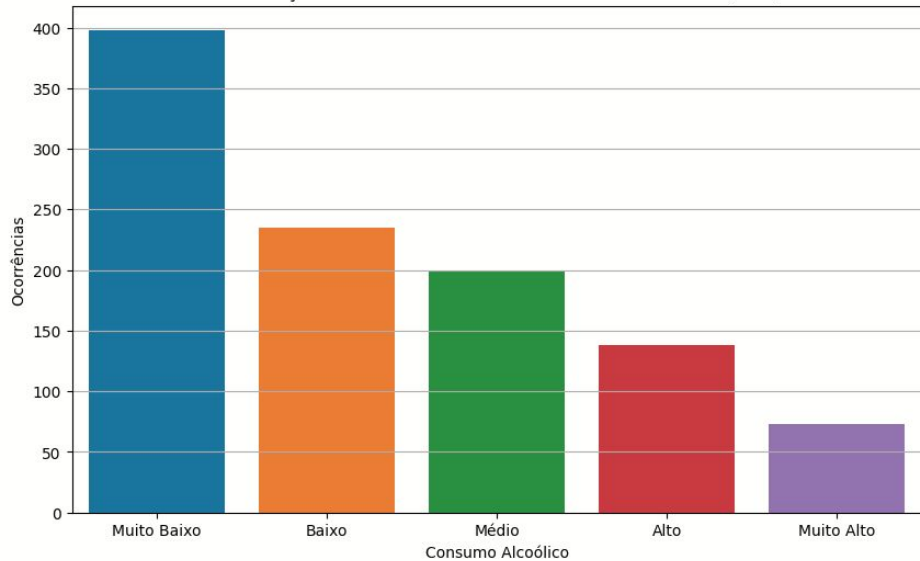
Segundo [5], notas ≥ 10 são aprovações em Portugal

Porcentagem dos alunos aprovados/reprovados

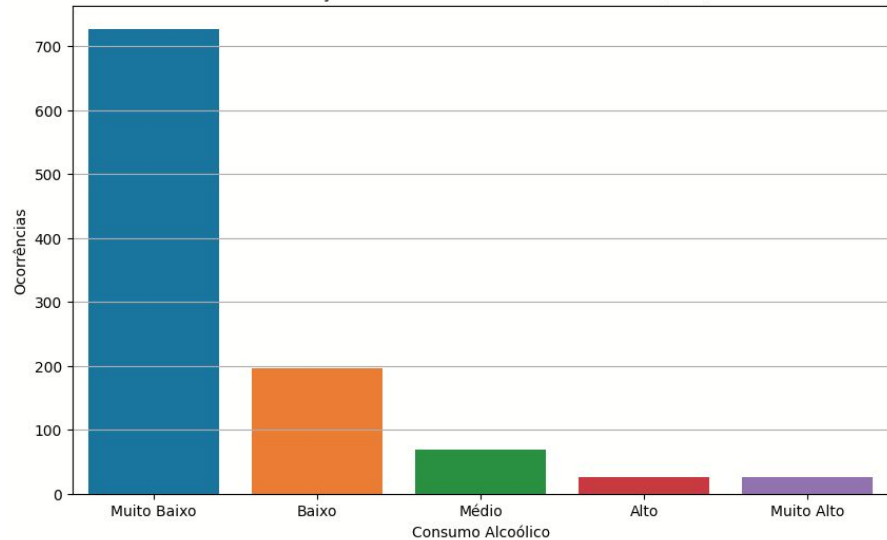


Distribuição do Consumo Alcoólico

Distribuição do consumo alcoólico nos finais de semana (Walc)



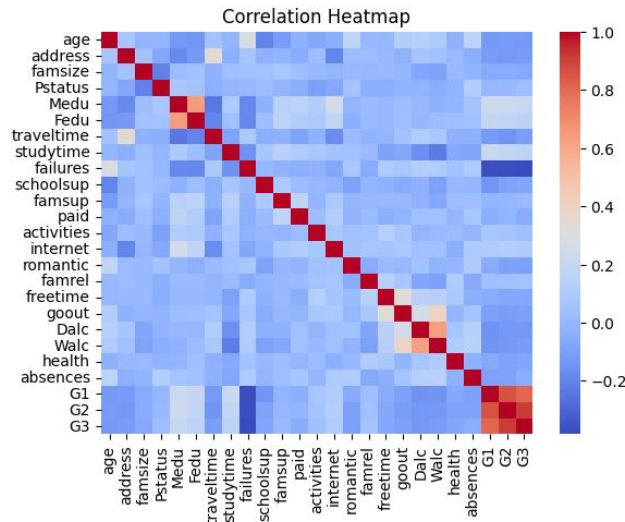
Distribuição do consumo alcoólico no dia-a-dia (Dalc)



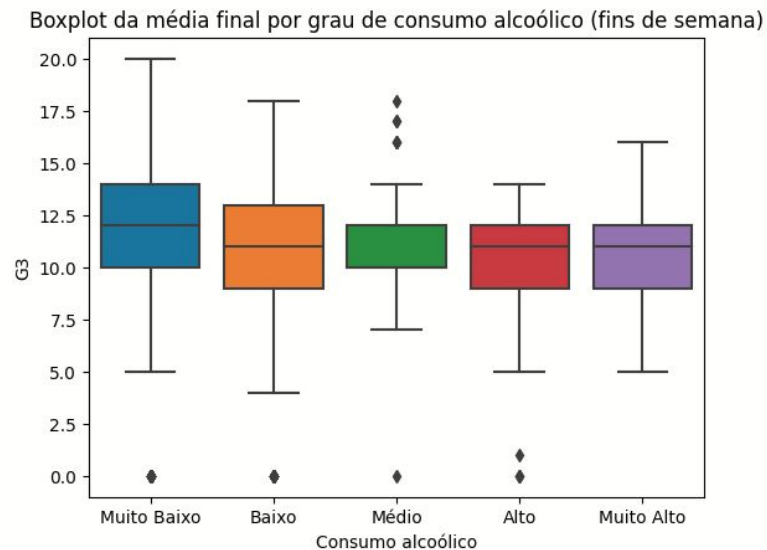
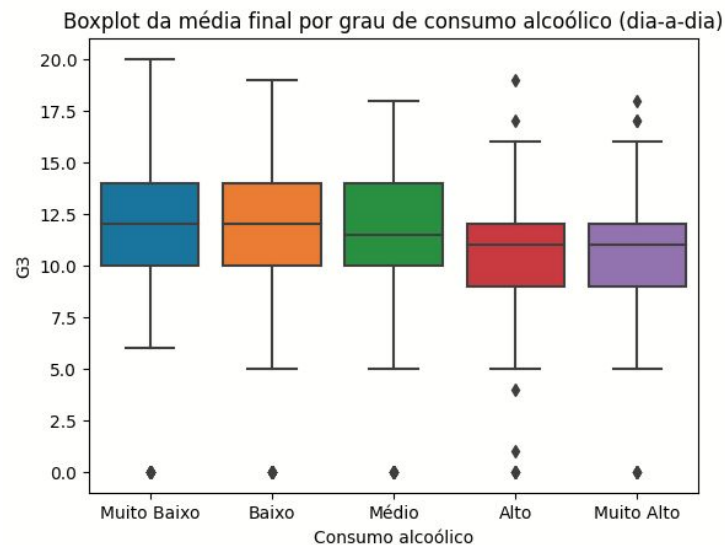
Consumo alcoólico varia conforme a semana

Análise Exploratória de Dados - Correlações

- Maiores correlações **positivas** à G3:
 - Grau de escolaridade materno (0.20)
 - Tempo de estudo (0.16)
 - Grau de escolaridade paterno (0.15)
 - Acesso à Internet (0.10)
- Maiores correlações **negativas** à G3:
 - Uso de álcool no dia-a-dia (-0.12)
 - Idade (-0.12)
 - Endereço Urbano (-0.11)
 - Uso de álcool durante finais de semana (-0.11)
 - Tempo em trânsito (-0.10)



Boxplot de notas por consumo alcoólico



Comparação de notas por consumo alcoólico

	Muito Baixo	Baixo	Médio	Alto	Muito Alto
Média	11.74	11.47	11.29	10.53	10.39
Mediana	12.0	12.0	11.5	11.0	11.0

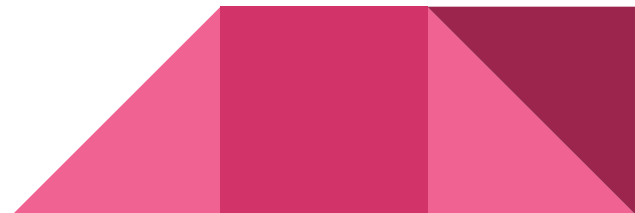


Verificação de Hipóteses

H_0 : “Não há uma correlação negativa entre o consumo alcoólico no dia-a-dia e a notas finais obtidas pelos indivíduos”

H_1 : “Há uma correlação negativa significativa entre o consumo alcoólico diário e notas finais obtidas.”

- Correlação (Pearson) **negativa fraca (-0.12)** entre consumo alcoólico e médias finais
 - P-value: $2,65 \cdot 10^{-5}$ → **suporta rejeitar H_0**



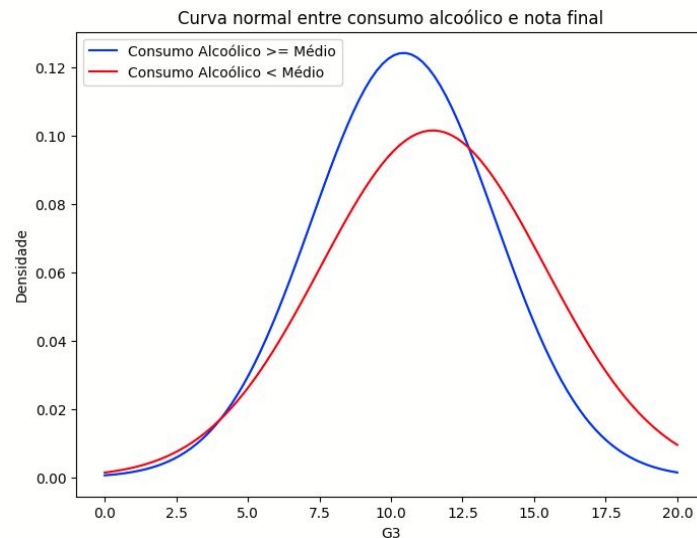
Verificação de Hipóteses

Teste ANOVA:

- Mede a variância entre grupos
 - F-statistic: 6,24
 - p-value: 5,79e-05 (menor que $\alpha=0,05$)

As diferenças entre os grupos são estatisticamente significantes!

H_0 Rejeitada.



Conclusões

- Consumo alcoólico está negativamente relacionado ao desempenho acadêmico
 - Não há evidências para suportar uma relação de causalidade
- Trabalhos futuros
 - Análise de outras correlações
 - Uso de um dataset mais atual



Referências Bibliográficas

- [1] UCI MACHINE LEARNING (org.). Student Alcohol Consumption. 2016. Disponível em: <https://www.kaggle.com/datasets/uciml/student-alcohol-consumption>. Acesso em: 24 jun. 2023.
- [2] CORTEZ, Paulo; SILVA, Alice Maria Gonçalves. Using data mining to predict secondary school student performance. 2008.
- [3] NYU. NYU Study Examines Top High School Students' Stress and Coping Mechanisms. 2015. Disponível em: <https://www.nyu.edu/about/news-publications/news/2015/august/nyu-study-examines-top-high-school-students-stress-and-coping-mechanisms.html>. Acesso em: 25 jun. 2023.
- [4] LEONARD, Noelle R. et al. A multi-method exploratory study of stress, coping, and substance use among high school youth in private schools. *Frontiers in psychology*, p. 1028, 2015.
- [5] WIKIPEDIA (org.). Grading systems by country. Disponível em: https://en.wikipedia.org/wiki/Grading_systems_by_country. Acesso em: 28 jun. 2023.

Referências Bibliográficas

- [6] PORTUGAL. Decreto nº 50/2013, de 2013. Novo Regime de Disponibilização, Venda e Consumo de Bebidas Alcoólicas em Locais Públicos e em Locais Abertos Ao Público. Portugal, 16 abr. 2013. Disponível em: <https://diariodarepublica.pt/dr/legislacao-consolidada/decreto-lei/2013-58361867>. Acesso em: 27 jun. 2023.
- [7] RITCHIE, Hannah; ROSER, Max. Alcohol Consumption. Elaborado por Our World in Data. Disponível em: <https://ourworldindata.org/alcohol-consumption>. Acesso em: 27 jun. 2023.
- [8] SUDHINARASET, May; WIGGLESWORTH, Christina; TAKEUCHI, David T. Social and cultural contexts of alcohol use: Influences in a social–ecological framework. Alcohol research: current reviews, v. 38, n. 1, p. 35, 2016.
- [9], Organização Mundial da Saúde. Fact Sheet: alcohol. Alcohol. 2022. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/alcohol>. Acesso em: 27 jun. 2023.
- [10] RODRIGUES, Leonardo. ANOVA: aprenda para o que serve, como calcular e em que momento utilizar essa variância!. Aprenda para o que serve, como calcular e em que momento utilizar essa variância!. 2019. Disponível em: <https://www.voitto.com.br/blog/artigo/anova>. Acesso em: 27 jun. 2023.
- [11] MUKAKA, Mavuto M. A guide to appropriate use of correlation coefficient in medical research. Malawi medical journal, v. 24, n. 3, p. 69-71, 2012.

Alcoolismo e a Vida Acadêmica



Universidade do Vale do Itajaí
Mestrado em Computação Aplicada — Kobrasol
Disciplina: Análise Exploratória de Dados
Professor: Wemerson Delcio Parreira

Aluno: Khalil G. Q. de Santana — khalil.santana@edu.univali.br