Dsa210 Project Report

This study analyzes the data from 1,044 Portuguese secondary‐school students. This examines whether alcohol consumption affects academic performance or not. After cleaning and merging the two datasets, I did an alcohol parameter (alcohol\_avg) and did exploratory analysis showing that most students report low weekday drinking but more moderate weekend consumption, that higher the drinking levels arethe more absences and lower median final grades (G3) are. Also the greater study time is associated with higher and better grades and also less absences. A two‐sample t‐test comparing low (≤2) and high (>2) drinkers shows a difference in final grade means, confirming the hypothesis of a negative alcohol effect. In predictive modeling, a linear regression using alcohol\_avg, studytime, and absences is made and in conclusion 3.1 % of grade variance (R² = 0.031) with coefficients of –0.288 for alcohol, +0.674 for study time, and –0.024 for absences, whereas a KNN regressor (k = 5) performed poorly (R² = –0.171). And lastly, k-means clustering (k = 3) on three features showed groups—one high-drink/low-grade cluster, one low-drink/high-grade cluster, and a larger group—suggesting targeted intervention potential. Future extensions may include adding the datasets of other relevant cases such as parent data or teacher data of the students, and adding more graphs and models will be possible.

ChatGpt Prompts

















