

STA 3000 Project Proposal

Data Analysis Option

Dataset: Red wine quality

Source: UCI machine learning repository

Link: <https://archive.ics.uci.edu/ml/datasets/wine+quality>

Topic: Discuss what attributes make a wine “good” and what make a wine “bad”

Intro: Simple and clean practice dataset for regression or classification modelling

The dataset has 12 variables and 1,599 observations. The predictor variables include fixed acidity, volatile acidity (amount of acetic acid - too much acetic acid can have an unpleasant taste), citric acid (can add a fresh taste), residual sugars (amount of sugar after fermentation stops), chlorides (amount of salt), free sulfur dioxide (prevents oxidation), total sulfur dioxide (can affect the taste), density (depending on alcohol and sugar content), pH (how acidic or basic), sulphates (contributes to SO₂ levels), and alcohol (the percentage of alcoholic content).