DATASCI 203 Lab 2 Proposal

Team members: Theresa Azinge, Wesley Chang, Spandan Garg, Dakota Potere-Ramos
Github repository: https://github.com/mids-w203/andamooka/

Research Question: How does the chemical composition of white wine affect its relative qualitative rankings?

Actor: Acme Beverage Distribution Corporation (ABDC) is interested in expanding its wine offerings to incorporate more specialty and artisanal wines. Specifically, our team has been asked to look at the market feasibility of Portuguese "Vinho Verde" white wines based on quality rankings. Our analysis will give insight into the possible relationship between chemical composition and quality scoring of "Vinho Verde" white wines. Our results will help ABDC's product team assess potential new "Vinho Verde" wine offerings before any independent wine evaluation has been conducted. ABDC's Head Sommelier and Beverage Management Team are the actors who can change the chemical composition of the wines via selection of different wine samples.

Audience: Our intended audience is ABDC's Head Sommelier and Beverage Management Team who are interested in predicting "Vinho Verde" white wine quality score based on chemical composition. The Head Sommelier has also expressed interest in building a data-driven model that may later be generalized to other wines.

Data Source: https://archive.ics.uci.edu/ml/datasets/Wine+Quality
Data File: winequality-white.csv
X variable(s):

- fixed acidity (g/dm^3)
- volatile acidity (g/dm^3)
- citric acid (g/dm^3)
- residual sugar (g/dm^3)
- chlorides (g/dm^3)
- free sulfur dioxide (mg/dm^3)
- total sulfur dioxide (mg/dm^3)
- density (g/cm^3)
- pH
- sulphates (g/dm^3)
- alcohol (vol.%).

Y variable: quality score (0-10)

The X variables will be used to operationalize the chemical composition of a wine. The Y variable will be used to operationalize wine quality.

Unit of observation: Each entry in the data set represents the results of laboratory and sensory analysis on a unique sample of "Vinho Verde" white wine. The laboratory analysis provides eleven features related to the chemical composition of the wine. Nine features represent concentrations of various compounds while the other two represent pH and alcohol percentage by volume. The sensory analysis provides a single feature related to the quality score of the wine on a 0-10 scale. Each wine sample was evaluated by a minimum of three sensory assessors, or sommeliers, using blind taste testing. The final quality score value is given by the median score of all evaluations.