WINE QUALITY PREDICTION

Team Members:

1. Tanmay Ambegaokar
2. Bhumika Mallikarjun Horapet
3. Kundana Chowdary Cherukuri
4. Jagadeeshwar Kalyanapu

Project Proposal:

Two of the most consumed varieties of wine are red and white, each with distinctive flavors and qualities. Cortez et al. created the two data sets that were used in this investigation. The sets include the physicochemical characteristics of the white and red Vinho Verdes wines as well as the sensory evaluations of the wines made by wine specialists. Both sets were merged into a single data frame for handling convenience. The columns include a quality score derived from professional tastings as well as details on the physicochemical characteristics of the wines.

The dataset contains 13 columns and 6497 rows which tell us about the different aspects such as pH level, alcohol content, residual sugar level, acidity, and chloride levels which influence the quality of the wine. We will perform a variety of statistical tests and explore the dataset using various plots to examine the influence of each attribute that affects the quality of the wine. A better chemical understanding of the wine quality will facilitate better wine production.

SMART QUESTIONS:

1. Which chemical attribute has the most significant influence on wine quality?
2. How does the distribution of sugar, sulphate, chloride, pH, and alcohol content change over different quality categories?
3. How these factors vary between red and white wine?

DATASET LINK:

http://www3.dsi.uminho.pt/pcortez/wine/

GITHUB REPOSITORY:

https://github.com/Jagadeeshwarkalyanapu/Team_Bees