Capstone project 1

Proposal for Capstone Project

Title: predicting the quality of wine by using physicochemical properties

Problem:

What makes good wine good? I always have this question in my mind when having a glass with friends. Normally, I select my favorite bottle from beautiful label, deepness of the bottom and rating website in the internet. However, it would be much more legit and informative to investigate this problem through physicochemical properties which are the fundamental of flavor and taste of wine. After eliminating all the psychological biases such as price, popularity, year, branding, packaging etc., we can ask something like. What physicochemical properties will affect the quality of wine? Is there any magic mixture of acid and sugar that makes wine guru give a thump up?

Who care? :

By digging deeper, it can be beneficial to both supply side and demand side. For supply side, this is obvious that the wine makers want to brew the best wine for market. Getting known more about physicochemical properties will give them more resource to achieve their goal. In demand side, wine customers will learn more about physicochemical properties and how they affect wine quality. They will have more scientific indicators for selecting their own favorite bottle.

Data:

The physicochemical properties and rating data will be acquired from the UCI machine learning repository. This data contains about physicochemical properties and rating of the Portuguese "Vinho Verde" wine. On this data set, there is no information about name, year, and price of each bottle due to privacy and logistic issue.

Limitation:

First, our data set is not included wine from other country and different kind of wine. This can be critical drawback when we want to generalize our finding. Second, there is only one rating in our dataset. Adding more rating from different source can increase more credibility of the result. On the other word, the result heavily depends on just one rating system. Last, lack of price, name and year limits the application of the finding. It would be much more interesting, if we can investigate more in the relationship between physicochemical properties and price or year, so we might find answer to some myths in winery industry.