Project 4 - Project Proposal

Project Title: Wine Quality Analysis - Machine Learning Model

Team Members: AJ Domingo, Samuel Heaton, Alice Johnson,
Jananee Arjunan, Mia Tsivitse, William Julius

Problem Statement:

Our target variable will be points (rating score). We believe this analysis could be used to determine wine recommendations for a consumer analyzing a combination of factors such as top ratings, origin location and price. Using machine learning, our goal is to create a model that could predict the final rating of the wine as an indicator of its quality (‘classic’ or ‘good’).

The wines in our dataset were rated on a scale of 0-100 by wine enthusiasts. However, the only reviews shared were for wines that scored >= 80.

Scope of Project:

We are looking to analyze wine data based on factors such as:

- U.S. State
- Wine description
- WineEnthusiast Point Ratings
- Wine Price
- Wine Vineyard
- Wine Province
- Wine Region

Some example questions we are looking to answer include:

- Which state of origin has the highest average rating?
- What state of origin has the lowest average rating?
- What is the highest average rating by price?
- Which state of origin produces the most wine?
- What is the average wine price by country of origin?
-

Top classic wine comes from napa valley california is of variety Pinot Noir

- What are the highest producing Vineyards by country?

Data Used:

- Wine Reviews | Kaggle
- Google Maps Platform

- Visual Crossing Weather Data API

Format:
Python Pandas
Python Matplotlib
Amazon AWS RDS
PostgreSQL
Tableau
Tensor Flow & Python Script