Technical Task

● Imagine you are working as a Data Scientist for an Online Wine Shop named “The Wine Land” (Dream come true??).

● As the name suggests, the online store specializes in selling different varieties of wines.

● Wines and its allied activities are niches. Having said that, the online store receives a decent amount of traffic and reviews from its users.

● Your boss is hoping to leverage the “reviews” data and draw actionable insights from it.

What is Expected?

● The directives from your boss are -

○ To derive the top 5 actionable Insights from the Data.

○ To build a predictive model for predicting the wine “variety”.

○ To build an API for serving predictions. Feel free to define the I/O contract.

Outcome

● Submit a one-page word/pdf document highlighting the actionable insights from the analysis. Feel free to add code snippets wherever necessary.

● Clearly state the assumptions made (if any).

● Submit the source code used for building models in a zip or share the link to the GitHub repository.

About the Dataset -

● The dataset can be downloaded from - https://drive.google.com/file/d/1ra9lwNjK9G8Ns0bAfzipD0u3Xwii5hc0/view

● The Data Description is as follows -

○ user_name - user_name of the reviewer

○ country -The country that the wine is from.

○ review_title - The title of the wine review, which often contains the vintage.

○ review_description - A verbose review of the wine.

○ designation - The vineyard within the winery where the grapes that made the wine are from.

○ points - ratings given by the user. The ratings are between 0 -100.

○ price - The cost for a bottle of the wine

○ province - The province or state that the wine is from.

○ region_1 - The wine-growing area in a province or state (ie Napa).

○ region_2 - Sometimes there are more specific regions specified within a wine-growing area (ie Rutherford inside the Napa Valley), but this value can sometimes be blank.

○ winery - The winery that made the wine

○ variety - The type of grapes used to make the wine. Dependent variable for task 2 of the assignment.

Deadline

Submit the assignment on or before 14th May 2023 (Sunday) EOD.

Incase of any queries contact: s348ankit@gmail.com