**Wine Variety Prediction:**

**Summary:**

The first part of this project consists of data visualizations for features ‘country’,’winery’ with the price and rating.

The second part of this project consists of insights from review descriptions

The third part is the model for predicting the variety of the wine.

The features used for the same are:

**['price','rating','one\_word\_review','one\_word\_review\_2','one\_word\_review\_3','country','province',**

**'region\_1','region\_2','points']**

The other features were created from review insights and domain Knowledge

The model used for training was ‘Random forest classifier’.

While the other models like kNN classifier was used as well, the **Random forest Classifier** worked out to be best with arounr 40 percent accuracy.

**Note: “The insights and visualizations are explained in the ipynb files itself”**