Context

- Imagine you are working as a Data Scientist for an Online Wine Shop named “The Wine Land”
- As the name suggests, the online store specializes in selling different varieties of wines.
- The online store receives a decent amount of traffic and reviews from its users.
- Leverage the “reviews” data and draw actionable insights from it.

What is Expected?

- Build a predictive model for predicting the wine “variety”. Provide the output along with all features to a CSV file. Both Training & test data is provided here
- Submit the source code used for building models in a zip or share the link to the GitHub repository.
- Also submit a short summary: Model used, features extracted, Model accuracy in train. Along with some visualization of data and top 5 actionable Insights from the Data.
- .

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