Wine Ratings vs Location Observable Trends

# Q 1: How does wine rating differ by country?

A screenshot of a cell phone

Description generated with high confidence

* USA has most recorded ratings accounting for 42% of total count of ratings followed by France and Italy.

# Q 2: Which provinces produce most wine (top 30)?

A screenshot of a cell phone

Description generated with high confidence

* USA tops the list with most provinces that produce wine. The top five provinces are California, Washington, Oregon, New York and Virginia which account for about 41% of total production.
* French provinces of Bordeaux, Burgundy, Alsace, Loire Valley, Champagne, Beaujolais, Southwest France, Provence and Rhone Valley account for 16% while the Italian provinces of Tuscany, Piedmont, Veneto, Northeastern Italy, Sicily & Sardinia, Southern Italy and Central Italy comprise about 14.5% of wine production worldwide.

# A screenshot of a cell phone Description generated with high confidenceQ 3: What is the average rating for most wines?

A screenshot of a cell phone

Description generated with high confidence

* For this dataset, average wine rating for most wine varietals is between 85-89 (68,497 ratings accounting for 52%. Wines from major producers like USA, France, Italy, Spain and Portugal fall in this category.
* Caveat: There are some outliers in the 95-100 rating range (2417 ratings accounting for 1.9%) suggesting that these wine varietals are most coveted and are among the most expensive.

# Q 4: Does having higher number of wineries/vineyards translate to higher wine ratings?

A screenshot of a cell phone

Description generated with high confidence

* Even though there are significantly higher number of wineries in USA, the average wine rating is lower when compared to some other countries. However, the quality of wine produced here is attested to by the higher number of ratings/ upvotes recorded. The same applies to the French and Italian wines.
* Caveat: Despite having significantly low number of ratings in comparison, England still maintains the highest average rating which could suggest production of decent quality wine. But, more datapoints are required to make this conclusion.

# Some Shortcomings with dataset

1. It’s not clear how many years of data has been recorded in this dataset. More number of unique datapoints are required to glean concrete insights and predict trends.
2. No information on how other factors (climate, soil, air quality, inclement conditions like fires or drought etc.) may have played a role in quality of grapes produced to make wine.
3. There isn’t a lot of diversity in wine tasters i.e. a lot of repetitive votes have been recorded for most wine varieties, strongly suggesting bias.
4. It isn’t clear if all the ratings were recorded in a year or two or over period of time.
5. Many vineyards/ wineries produce two or more wine varietals in their premises and have been listed multiple times. As a result, we cannot determine the relation between wine rating for every unique vineyard/ winery.

# for the future

1. Since we were cut for time and due to limitations in the dataset, we couldn’t answer some questions adequately. However, if given more time, we would use Google API and Weather API to be able to substantiate our conclusions and glean insights into trends better.