

Winos



The Team

How is Climate Change affecting wine?

Kyle Johnson

Pepper

First Segment: Square
Second Segment: Circle
Third Segment: Triangle
Fourth Segment: X

Marisa Shideler

Allie

First Segment: Triangle
Second Segment: Triangle
Third Segment: X
Fourth Segment: Triangle

Brenya Skaggs

Milo & Olive

First Segment: Circle
Second Segment: X
Third Segment: Circle
Fourth Segment: Circle

Zackary Gheen

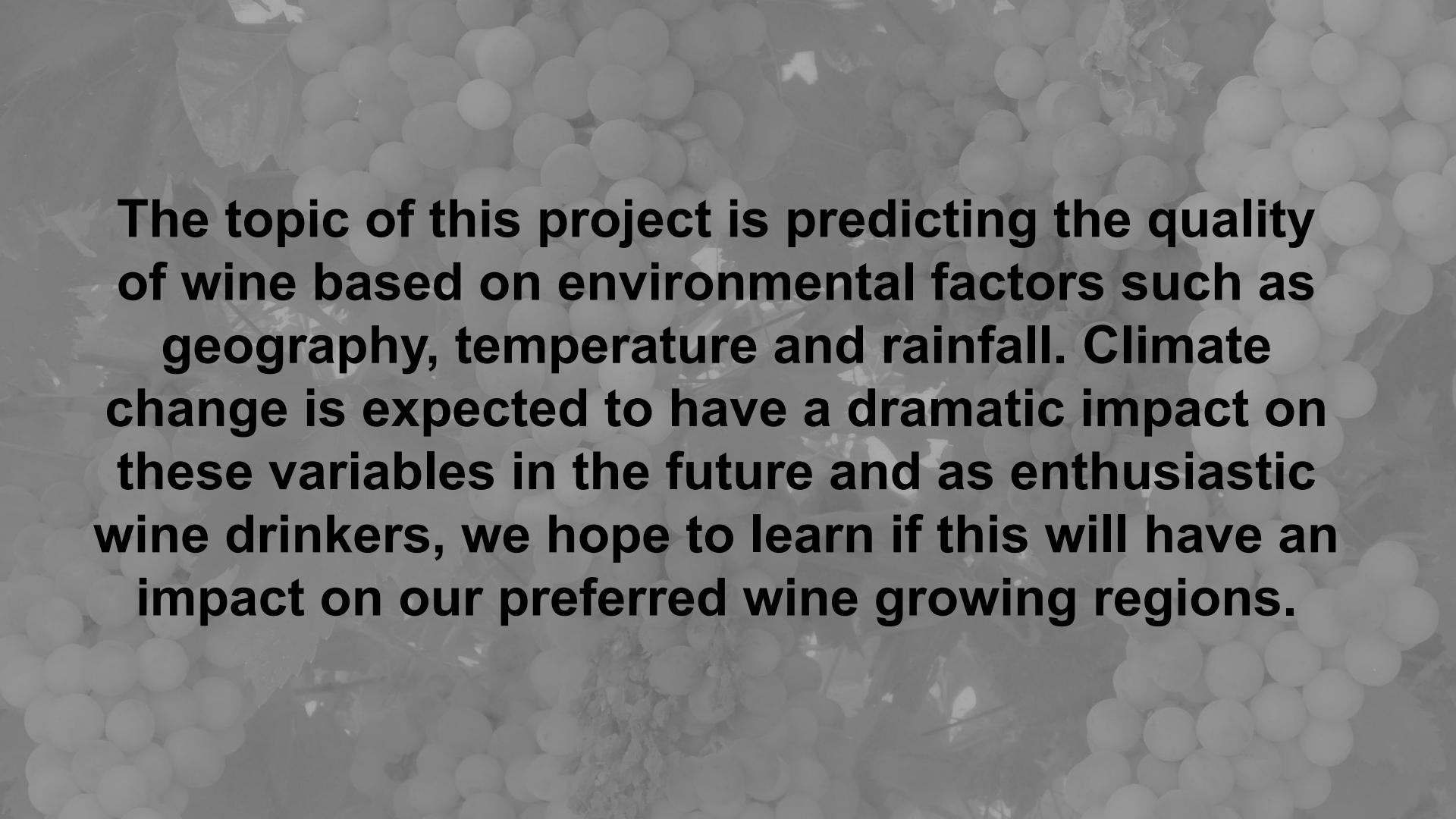
Bailey

First Segment: X
Second Segment: Square
Third Segment: Square
Fourth Segment: Square



BAILEY





The topic of this project is predicting the quality of wine based on environmental factors such as geography, temperature and rainfall. Climate change is expected to have a dramatic impact on these variables in the future and as enthusiastic wine drinkers, we hope to learn if this will have an impact on our preferred wine growing regions.



Questions to Answer

- Do higher temperatures/rainfall correlate with higher or lower quality wine?
- What effect will future changes in rainfall and temperatures have on wine quality from various regions?
- Are new regions poised to emerge as premier locations for growing grapes and producing wine?

Technologies, Languages & Tools

Python Libraries

- Pandas
- Sklearn
- TensorFlow
- Train_test_split
- LinearRegression

AWS

- RDS
- S3

Jupyter Notebook

- VSCode
- Google Colaboratory
- Google Slides

- HTML
- CSS

- JavaScript
- PostgreSQL
- PGAdmin

Machine Learning

- Multiple Linear Regression

- Heroku
- Excel
- GitHub



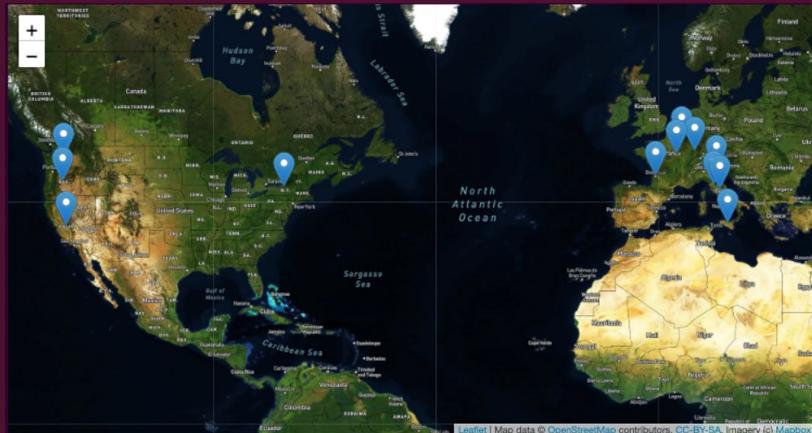
Data Sources

Environment Dataset

- The world bank provides observed rainfall and temperature data by year for regions within individual countries from 1901-present.
- Future predictions of the weather with the same structure are provided from 2020-2100.
- [Home | Climate Change Knowledge Portal \(worldbank.org\)](https://www.worldbank.org/en/topic/climate-change)

Wine Dataset

- This dataset includes 130,000 records of wine reviews from 2000-2017.
- [Wine Reviews | Kaggle](https://www.kaggle.com/datasets/wine-reviews/wine-reviews)



Dataset Analysis

Weather Data

Raw Data

Historical and projected precipitation and temperature were extracted from the country data for the following provinces/states:

- Alsace, France
- Aquitaine, France
- Burgundy, France
- Champagne Ardenne, France
- Piemonte, Italy
- Sardinia, Italy
- Sicily, Italy
- Tuscany, Italy
- Veneto, Italy
- Cantabria, Spain
- California, US
- New York, US
- Oregon, US
- Washington, US

Unnamed: 0	Spain	Andalucía	Aragón	Principado de Asturias	Illes Balears	Canarias	Cantabria	Castilla y León	Castilla-La Mancha	...	
0	1901	660.97	731.37	629.56	813.21	636.41	623.12	953.56	584.90	557.29	...
1	1902	660.49	653.91	633.91	837.45	642.47	292.17	960.72	636.91	552.15	...
2	1903	556.15	452.08	481.78	903.15	627.99	539.16	909.01	588.50	410.55	...
3	1904	582.28	614.13	485.96	767.75	562.99	534.67	834.29	549.63	526.19	...
4	1905	530.70	471.13	522.26	761.79	519.78	390.46	798.84	510.66	417.18	...

5 rows × 21 columns

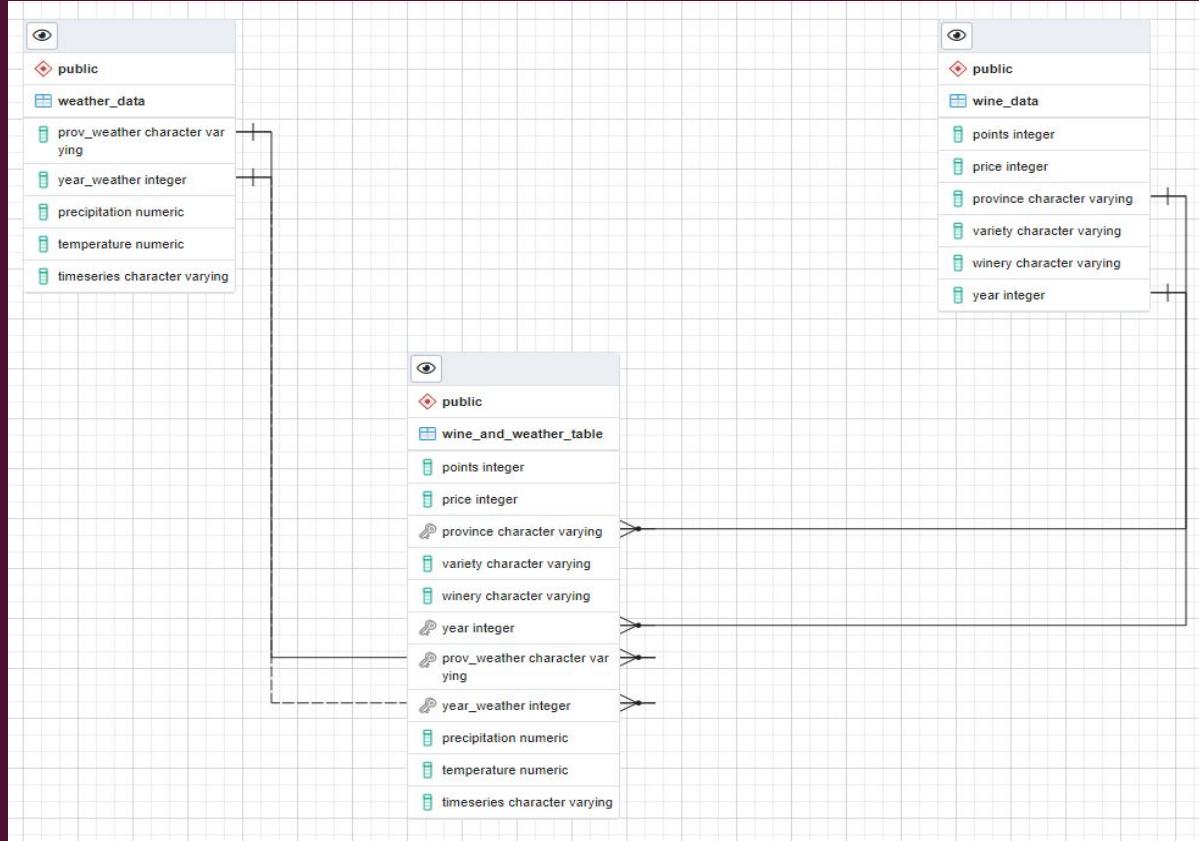
Wine Data

Raw Data

The raw wine data, which contain over 80,000 wine reviews, came from Kaggle. This dataset contains varieties of wine along with other information such as where and when they came from, what wineries were used, how much they cost and what the review rating is.

Unnamed: 0	country	description	designation	points	price	province	region_1	region_2	title	variety	winery
0	0	Italy Aromas include tropical fruit, broom, brimston...	Vulkà Bianco	87	NaN	Sicily & Sardinia	Etna	NaN	Nicosia 2013 Vulkà Bianco (Etna)	White Blend	Nicosia
1	1	Portugal This is ripe and fruity, a wine that is smooth...	Avidagos	87	15.0	Douro	NaN	NaN	Quinta dos Avidagos 2011 Avidagos Red (Douro)	Portuguese Red	Quinta dos Avidagos
2	2	US Tart and snappy, the flavors of lime flesh and...	NaN	87	14.0	Oregon	Willamette Valley	Willamette Valley	Rainstorm 2013 Pinot Gris (Willamette Valley)	Pinot Gris	Rainstorm
3	3	US Pineapple rind, lemon pith and orange blossom ...	Reserve Late Harvest	87	13.0	Michigan	Lake Michigan Shore	NaN	St. Julian 2013 Reserve Late Harvest Riesling ...	Riesling	St. Julian
4	4	US Much like the regular bottling from 2012, this...	Vintner's Reserve Wild Child Block	87	65.0	Oregon	Willamette Valley	Willamette Valley	Sweet Cheeks 2012 Vintner's Reserve Wild Child...	Pinot Noir	Sweet Cheeks

Weather & Wine Data



Combined Weather and Wine Data

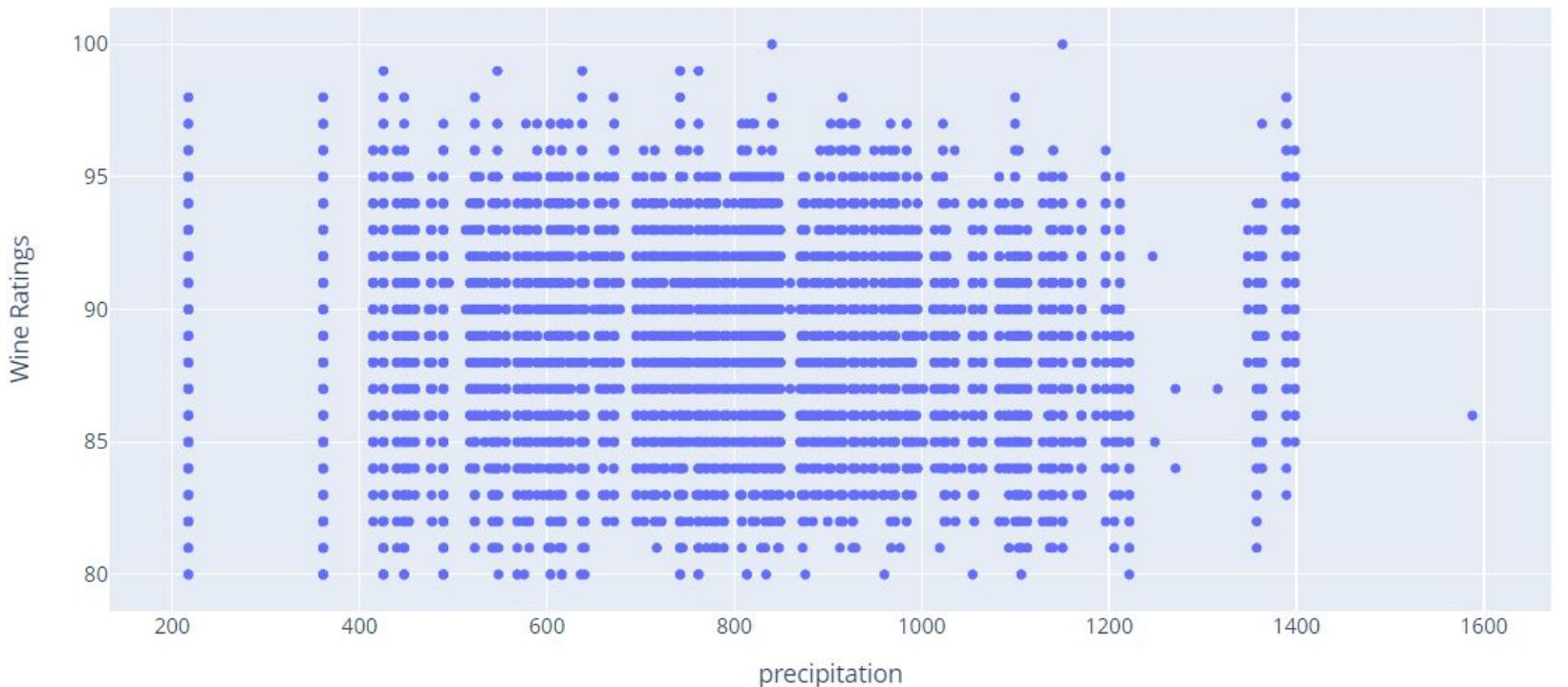
Google Colaboratory was used to write the weather and wine data to an Amazon RDS instance. Using SQL we were able to create a new table with the joined data. The following table will be used for our Machine Learning Model.

```
CREATE TABLE wine_weather_table
AS (SELECT * FROM wine_data
    LEFT JOIN weather_data ON wine_data.province=weather_data.Prov_Weather AND wine_data.year=weather_data.Year_Weather
    WHERE weather_data.Timeseries = 'Historical');
```

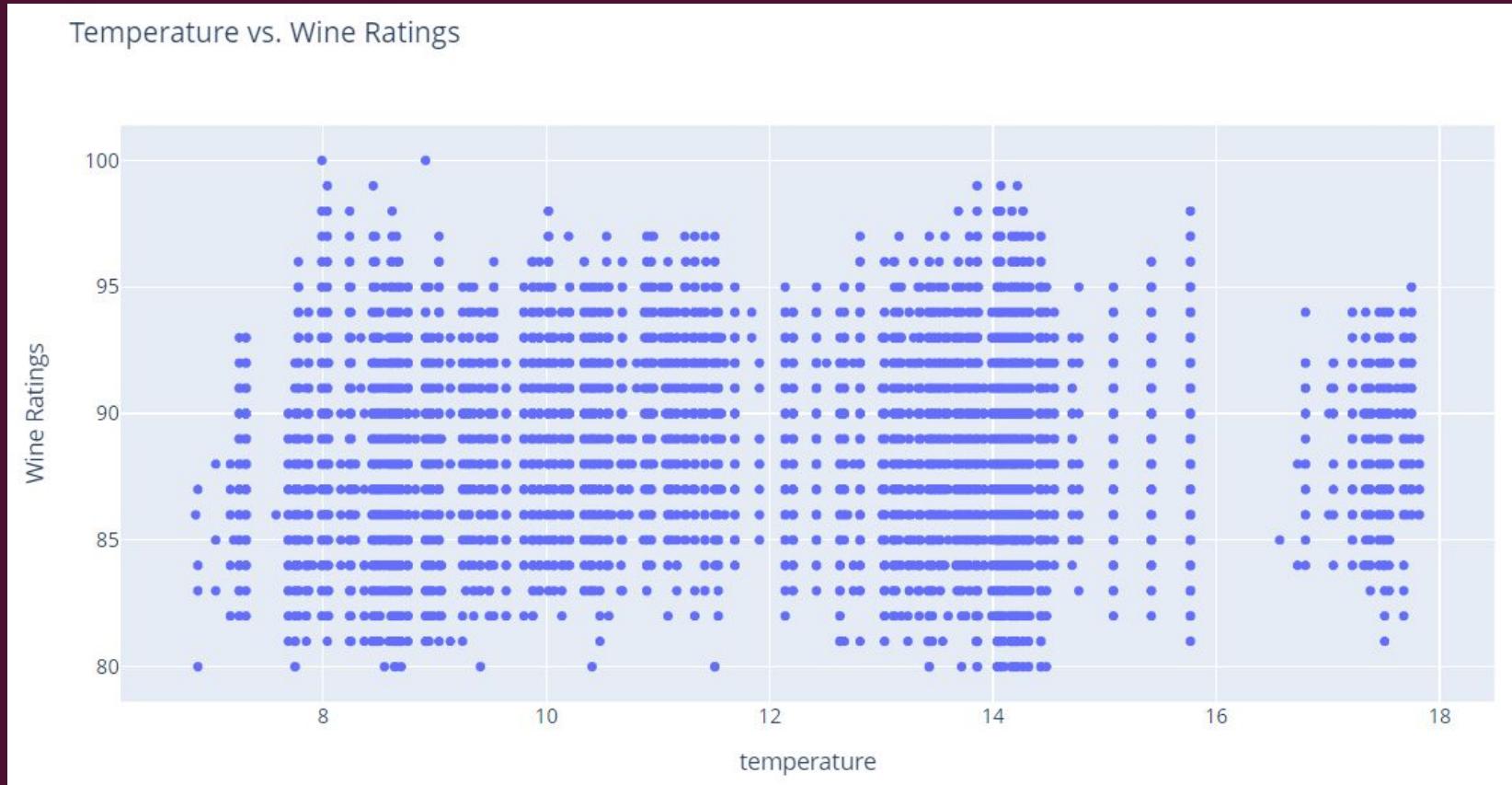
	points	price	province	variety		winery	year	prov_weather	year_weather	precipitation	temperature	timeseries
0	87	14	Oregon	Pinot Gris		Rainstorm	2013	Oregon	2013	440.10	8.68	Historical
1	87	65	Oregon	Pinot Noir		Sweet Cheeks	2012	Oregon	2012	808.24	9.04	Historical
2	87	24	Alsace	Gewürztraminer		Trimbach	2012	Alsace	2012	938.63	10.39	Historical
3	87	27	Alsace	Pinot Gris		Jean-Baptiste Adam	2012	Alsace	2012	938.63	10.39	Historical
4	87	19	California	Cabernet Sauvignon		Kirkland Signature	2011	California	2011	489.67	13.43	Historical
...
58846	90	28	Alsace	Pinot Gris	Domaine Riefländermann	2013	Alsace	2013	949.40	9.94	Historical	
58847	90	75	Oregon	Pinot Noir		Citation	2004	Oregon	2004	596.22	9.28	Historical
58848	90	30	Alsace	Gewürztraminer		Domaine Gresser	2013	Alsace	2013	949.40	9.94	Historical
58849	90	32	Alsace	Pinot Gris		Domaine Marcel Deiss	2012	Alsace	2012	938.63	10.39	Historical
58850	90	21	Alsace	Gewürztraminer		Domaine Schoffit	2012	Alsace	2012	938.63	10.39	Historical

Initial Analysis

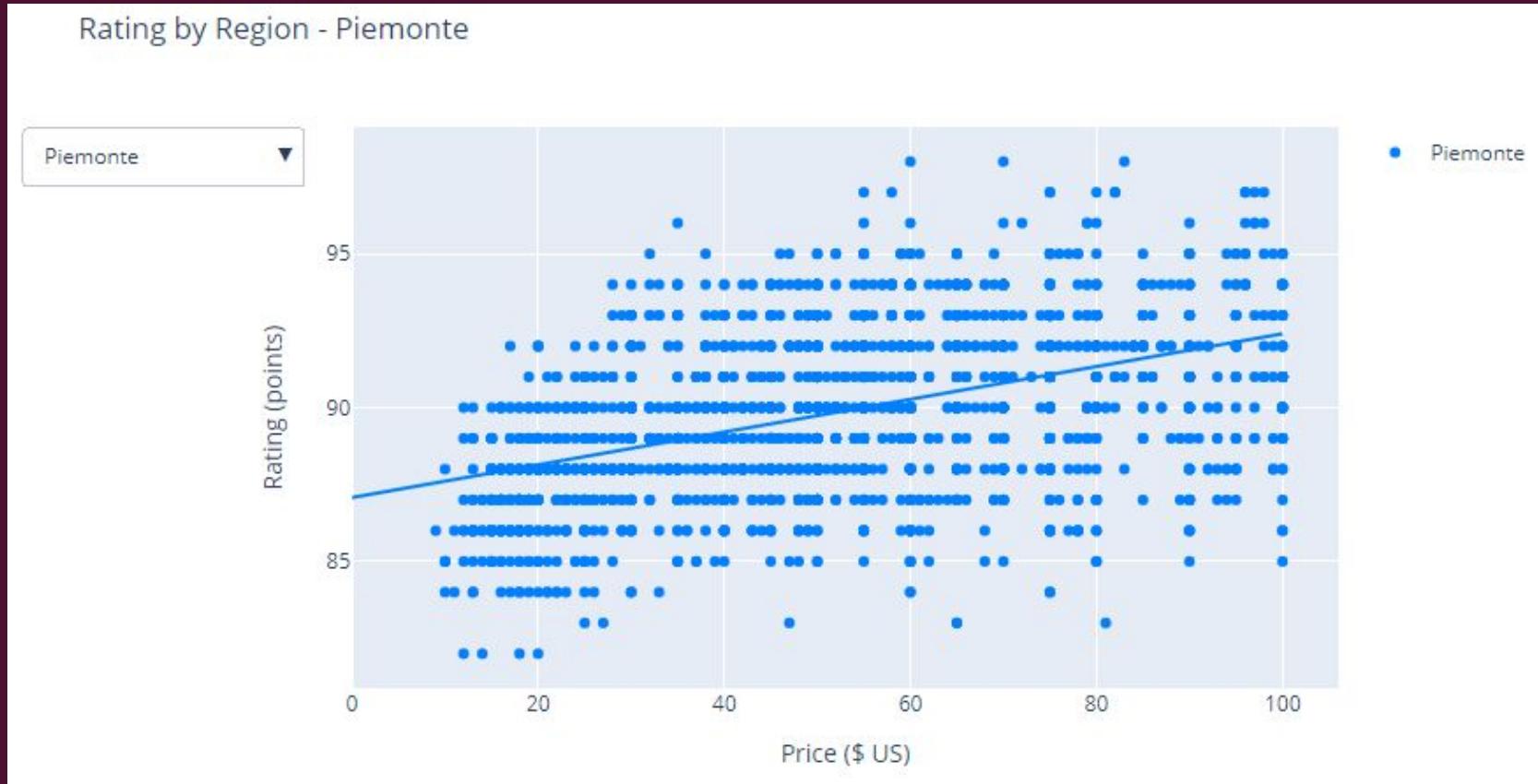
Precipitation vs. Wine Ratings



Initial Analysis



Initial Analysis



Machine Learning

Multiple Linear Regression

Output: Wine Ratings

	Inputs: Province, Variety, Price, Temperature and Precipitation	Inputs: Price, Temperature, Precipitation
MSE	5.73	6.17
R-Squared	33.2%	28.1%

Deep Learning

Deep Learning

```
[47]: # Define the model - deep neural net
number_input_features = 68
hidden_nodes_layer1 = 132
hidden_nodes_layer2 = 68

nn = tf.keras.models.Sequential()

# First hidden Layer
nn.add(tf.keras.layers.Dense(units=hidden_nodes_layer1, input_dim=number_input_features, activation="relu"))

# Second hidden Layer
nn.add(tf.keras.layers.Dense(units=hidden_nodes_layer2, activation="relu"))

# Output Layer
nn.add(tf.keras.layers.Dense(units=1, activation="relu"))

# Check the structure of the model
nn.summary()

Model: "sequential_1"
-----

| Layer (type)    | Output Shape | Param # |
|-----------------|--------------|---------|
| dense_3 (Dense) | (None, 132)  | 9108    |
| dense_4 (Dense) | (None, 68)   | 9044    |
| dense_5 (Dense) | (None, 1)    | 69      |


-----  

Total params: 18,221  

Trainable params: 18,221  

Non-trainable params: 0
```

```
: # Train the model
fit_model = nn.fit(X_train,y_train,epochs=10)

Epoch 1/10
1380/1380 [=====] - 2s 902us/step - loss: -134513.7500 - accuracy: 0.0000e+00
Epoch 2/10
1380/1380 [=====] - 1s 1ms/step - loss: -1667468.7500 - accuracy: 0.0000e+00
Epoch 3/10
1380/1380 [=====] - 1s 858us/step - loss: -5906646.0000 - accuracy: 0.0000e+00
Epoch 4/10
1380/1380 [=====] - 1s 918us/step - loss: -13431134.0000 - accuracy: 0.0000e+00
Epoch 5/10
1380/1380 [=====] - 1s 978us/step - loss: -24701812.0000 - accuracy: 0.0000e+00
Epoch 6/10
1380/1380 [=====] - 1s 912us/step - loss: -40187644.0000 - accuracy: 0.0000e+00
Epoch 7/10
1380/1380 [=====] - 1s 912us/step - loss: -60381660.0000 - accuracy: 0.0000e+00
Epoch 8/10
1380/1380 [=====] - 1s 918us/step - loss: -85774152.0000 - accuracy: 0.0000e+00
Epoch 9/10
1380/1380 [=====] - 1s 913us/step - loss: -116855528.0000 - accuracy: 0.0000e+00
Epoch 10/10
1380/1380 [=====] - 1s 919us/step - loss: -154096944.0000 - accuracy: 0.0000e+00
```

Dashboard

Group 2 Winos

Predicting Wine Quality

Wine Rating

Province

Variety

Price (\$)

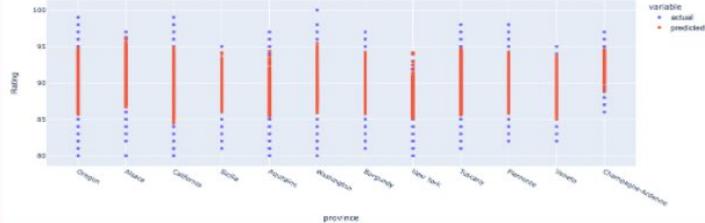
Temperature (C)

Precipitation (mm)

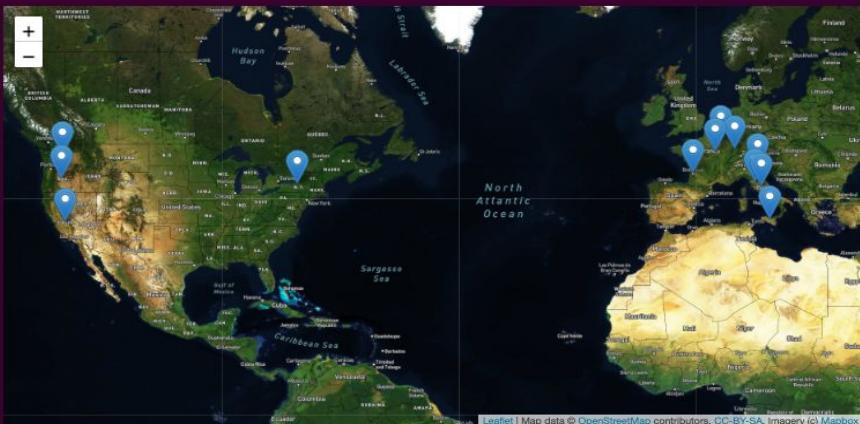
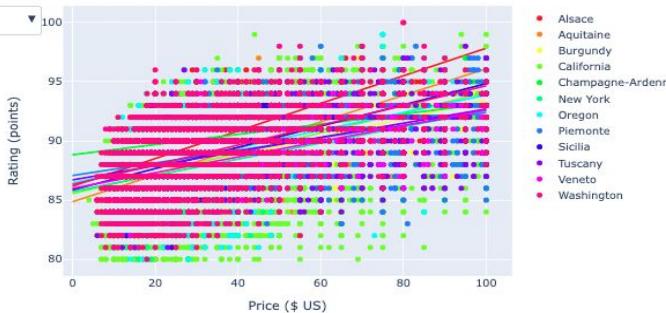
Predicted Wine Rating

Welcome to the Group 2 dashboard where we explored if there is a relationship between quality of wine and environmental variables such as geography, temperature and precipitation. Based on our analysis, there is no relationship. The R² value for the Multiple Linear Regression model we used is about 33%, which indicates the model isn't a good fit as can be seen in the charts for Predicted v Actual Ratings by Province and Wine Ratings by Region. The Wine Ratings by Region chart is interactive so that a single province can be viewed at one time instead of all provinces at once. The interactive map shows average temperature, precipitation and pricing for each province analyzed. Finally, the Wine Rating area is also interactive. Feel free to change province, variety of wine, price, temperature and precipitation and then click the "Predict" button to see what wine rating is produced using our model; however, considering the inaccuracy of our model, don't rely on it as valuable information. We feel the model would be improved with different data such as production levels, quality of grapes, types of grapes, and quality of soil.

Predicted v Actual Ratings by Province



Wine Ratings by Region



Conclusion

*Climate conditions in a wine growing region are
not by themselves accurate predictors of wine quality*

