

Winos



The Team

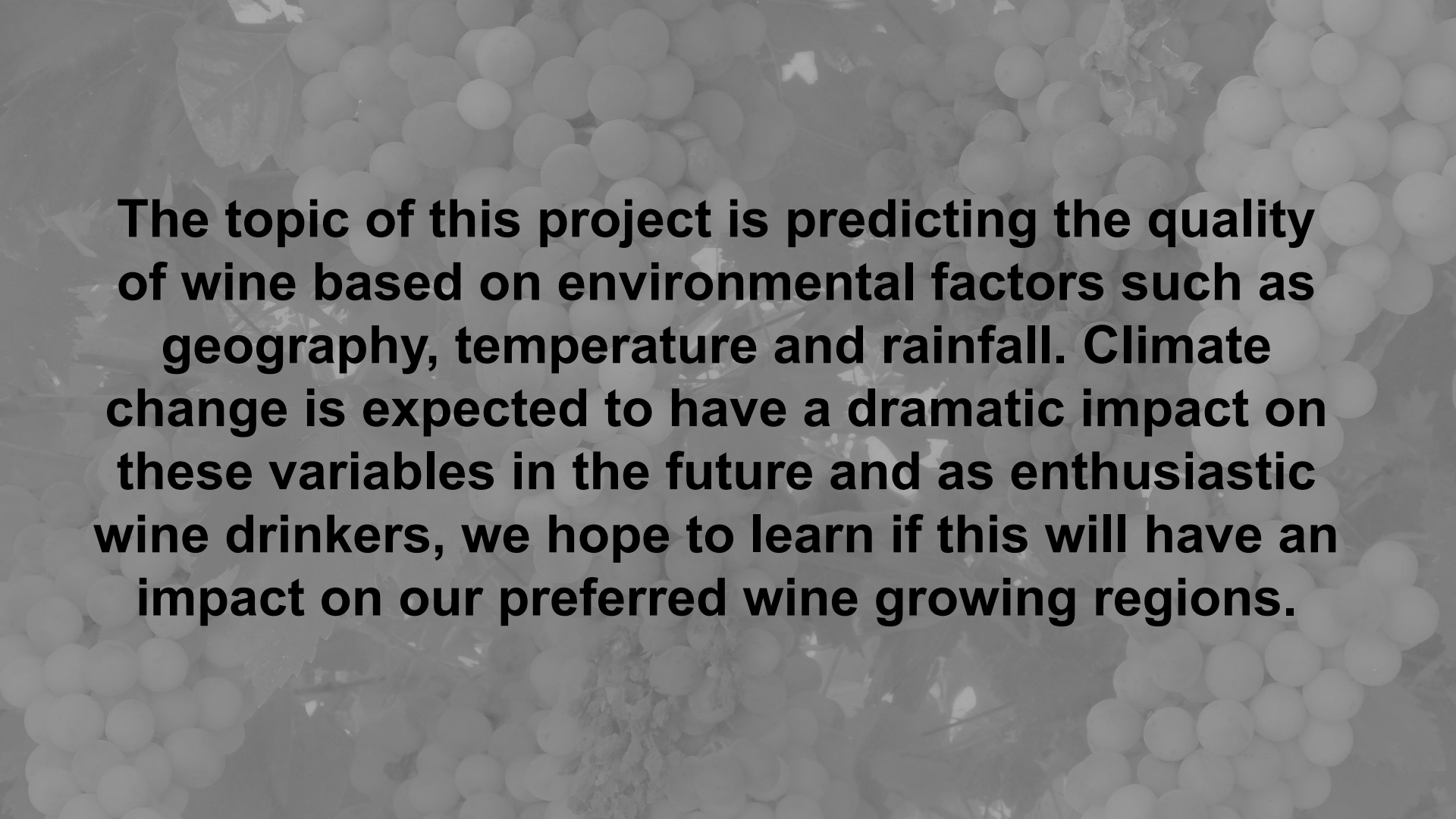
How is Climate Change affecting wine?

Kyle Johnson
Square


Marisa Shideler
Triangle

Brenya Skaggs
Circle

Zackary Gheen
X



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.

The background is a dark grey collage. It features a bar chart with an upward-trending arrow on the left, a donut chart with segments labeled 14%, 15%, and 10% in the upper right, and a bar chart with blue and grey bars at the bottom. In the foreground, a laptop is open, and a pair of hands is typing on the keyboard.

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 premiere locations for growing grapes and producing wine?

Group 2 Machine Learning Model Outline

Exploring the relationship between weather and wine

1) Datasets

- Wine reviews
- Historical mean temperature
- Historical precipitation

3) Types of Data Cleaning

Wine data:

- Remove row with excluded provinces
- Use regex to get year from Title field
- Perform feature selection (TBD)
- Drop rows with null values

Weather data requires no cleaning

5) Training and Evaluate Model

- Recommend Multiple Linear Regression
- Make regression
- Fit the model
- Predict wine quality

2) Features and Target

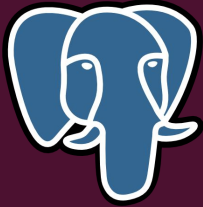
- Wine data features: Country, Description, Price, Province, Region_1, Region_2, Title, Variety, Winery
- Weather features: Year, Rainfall, Temperature, Timeseries
- Target: Points (wine rating)

4) Preprocessing

- All input data is tabular
- Merge wine data with weather data
- Drop rows with null values
- Split data into input (X) and output (y)
 - X – Features from 2
 - y – Target from 2
- Split X and y into training and test datasets

6) Reevaluate ML model as necessary

Technologies



Language



Tools



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://climateknowledgeportal.worldbank.org/)

Wine Dataset

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

Regions to be examined

- California, US
- Washington, US
- Bordeaux, France
- Tuscany, Italy
- Oregon, US
- Cantabria, Spain
- Piemonte, Italy
- Veneto, Italy
- New York, US
- Alsace, France
- Sicily, Italy
- Champagne, France

Dataset Analysis

Machine Learning

Results



Predictions

Recommendations

