Predicting Wine Grape Varietal Suitability Based on Climate and Precipitation

Viticulture at a Climate Crossroads

- Global warming have impacted the viticulture industry in a great amount.
  - It made certain traditional wine zones too hot to grow the wine grapes that they use to grow.
  - Causes places that use to be unsuitable to grow grapes to become more practical, such as England.
  - It causes overgrowth in areas that were traditionally used to grow cool climate wine styles, such as Germany.
  - It ruins the traditional flavor profiles of certain wine regions, such as Chabli.
- This project aims to create a machine learning model to estimate the suitability of wine grape varietals for specific locations based on climate data.
- It have the potential to help winemakers adapt to the challenges of climate change in viticulture, by offering a tool for winemakers to assess the feasibility of cultivating different grape varieties or exploring new regions.

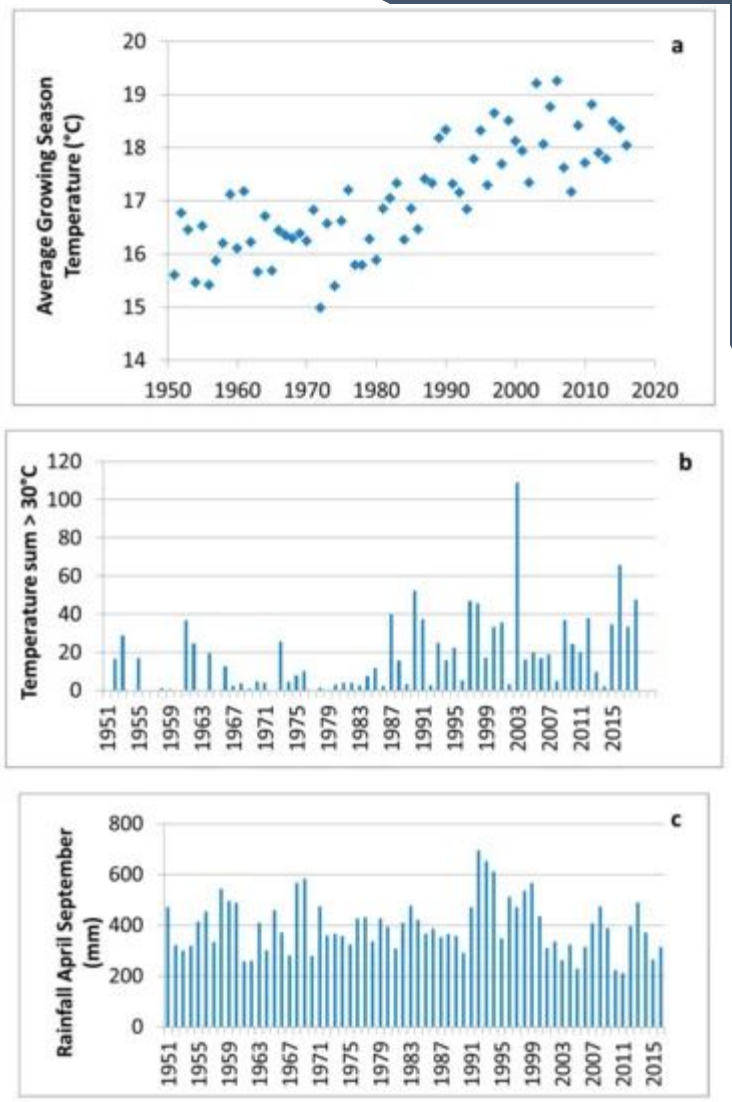


Figure 1. Climate data for Bordeaux (Bordeaux Mérignac weather station) from 1951 to 2018: (**a**) average growing season temperature, (**b**) temperature sum >30 °C during 45 days prior to harvest, (**c**) rainfall April–September. Source: (van Leeuwen et al., *An update on the impact of climate change in viticulture and potential adaptations* 2019)

Harnessing Data Science for Grape Varietal Suitability

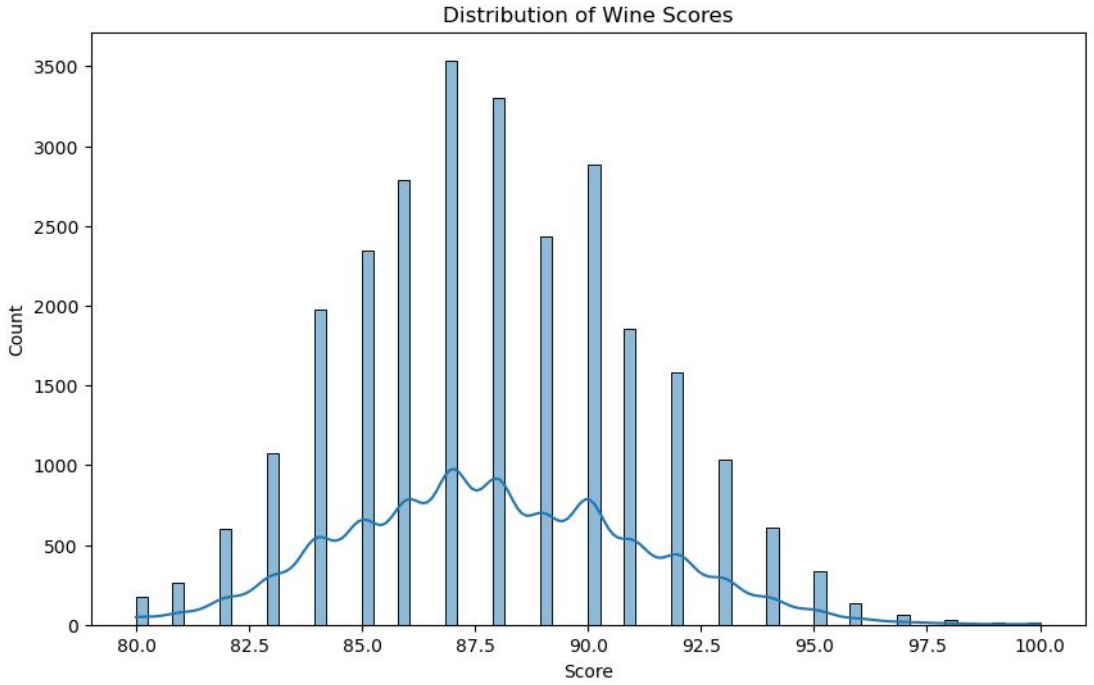
- **Predictive Insights:** Leveraging historical climate data to forecast future grape varietal success, empowering vintners with forward-looking guidance.
- **Data Integration:** Combining weather patterns, soil characteristics, and wine scores to create a multifaceted model of viticulture suitability.
- **Machine Learning:** Utilizing cutting-edge algorithms to parse complex datasets, revealing nuanced relationships between climate and grape viability.
- **User-Centric Design:** Tailoring the model interface for accessibility by vineyard owners and wine producers, providing actionable recommendations.
- **Dynamic Adaptation:** Ensuring the model adapts to ongoing climate trends, offering winemakers a tool to navigate long-term changes.
- **Collaborative Approach:** Engaging industry experts and data scientists in iterative model refinement to align with practical viticultural needs.

The Far-Reaching Impact of Predictive Modeling

- **Climate Resilience:** Equipping winemakers with data to adapt to unpredictable weather patterns, ensuring the resilience of vineyards to climate variability.
- **Innovative Cultivation:** Identifying new regions for viticulture, expanding the global wine map beyond traditional boundaries.
- **Economic Boost:** Driving regional development by introducing winemaking as a viable agricultural practice, stimulating job creation, and diversifying income sources.
- **Sustainability Focus:** Promoting sustainable viticulture practices by optimizing grape variety selection to suit local climatic conditions, reducing the need for environmental intervention.
- **Quality Consistency:** Enhancing wine quality by matching grape varietals with ideal growing conditions, leading to consistent, high-quality production.
- **Market Expansion:** Opening up new markets for winemakers by discovering untapped regions, offering a competitive edge in a global industry.

Data Insights: Wine Scores & Climate Correlations

- Data Sources:
  - Wine Data:
    - Kaggle’s wine review dataset scrapped by Zack Thoutt
    - Own dataset scrapped by my own custom scraper targeting Wine Enthusiast ratings.
  - Weather Data:
    - NOAA's historical climate and weather datasets, published by World Weather Records. This required a custom parsing script to transform it into workable dataset.
    - NOAA Weather API dataset to fill in missing weather data.
- Preliminary EDA:
  - Current data is skewed because not all data finish processing.
  - Some initial insights:
    - Optimal grape development for high-quality wines might require warmer temperatures during the early growth stages and a balance of warm temperatures with controlled water stress during the flowering and fruit set stage.
    - Varieties with broader climatic adaptability might show a wider range of scores due to variability in growing conditions across regions.
    - Higher quality wines are associated with a warmer and slightly wetter early growth stage, suggesting the importance of adequate water availability for initial grape development.
    - The warmth during the flowering, fruit set, and verasion stage is critical, with high scorers showing significantly higher temperatures. The slightly lower precipitation could suggest a need for controlled water stress to concentrate flavors in grapes.
    - Weather have no noticeable impact during late dormancy stages of the grape growth cycle.



Model Development Roadmap

- **Data Refinement:** Iterative cleansing and transformation to ensure robust model inputs.
  - Addressing data quality issues identified in preliminary EDA.
- **Feature Engineering:** Crafting predictive variables from raw data.
  - Calculating climate averages tied to grapevine phenological stages.
  - Create interface so users can input an address and get their varietal predictions.
- **Baseline Model Building:** Establishing a starting point for predictive accuracy.
  - Developing a suite of baseline models to set initial performance benchmarks.
  - Utilizing regression and classification techniques tailored to the dataset's characteristics.
- **Validation and Testing:** Ensuring model reliability and generalizability.
  - Cross-validation with diverse datasets to assess model stability.
  - Real-world testing with current year data for practical viability assessment.
- **Future Dataset Expansion:** Planning for model scalability.
  - Preparing for integration of a larger dataset once available.
- **Sustainability and Adaptation Metrics:** Measuring beyond accuracy.
  - Developing metrics to quantify economic and environmental impacts of varietal suitability.
  - Aligning model objectives with long-term sustainability goals in viticulture.