
Wine
Für dich, Für uns, Für alle
Bluberry Winery
Code Analytics
Arun & Jay
26.02.2021

AGENDA

01. **Introduction**

What, Why our story needs to be heard

02. **Data Collection**

What we found/have

03. **Methods/Analysis Plan**

What we going to do

04. **Data Analysis**

Exploratory analysis
Machine Learning Methods
Model Validation

05. **Price Prediction**

Price prediction
Price Suggestion

01. Introduction

- Background - Why our story needs to be heard
- Goal - What we want to tell

Background

- **Blueberry Winery** is one of the start-up Wine Manufacturing company in Portugal.
- **Blueberry Winery** is trying to enter the business with a good amount of analytics & research on domain knowledge

Goal

- **Blueberry Winery** wants to achieve the best 'Customer Satisfaction' towards the Quality and Price
- To Analyse and find the composition of factors that contribute to the quality wine

02.
Data Collection

- What we get

Wine Types

White Wine

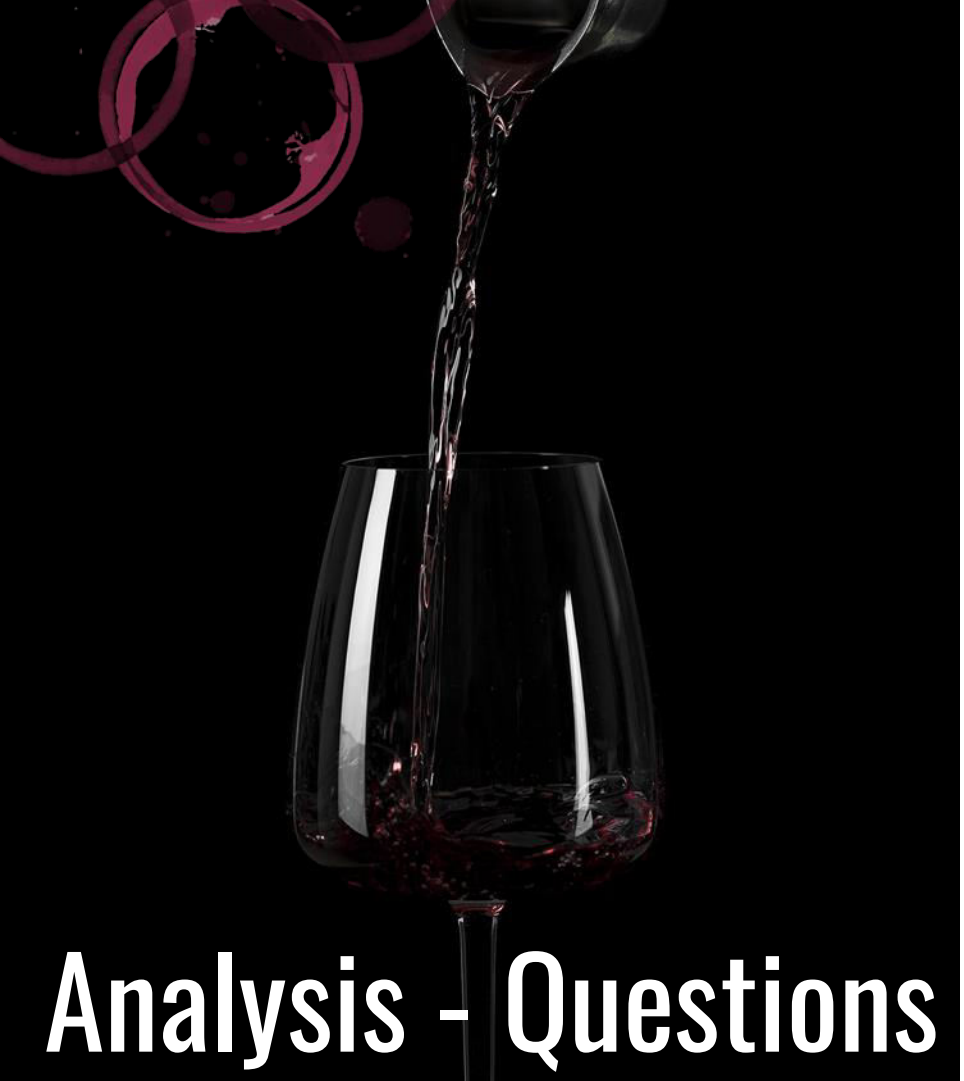
DataAttributes / Ingredients

- Fixed Acidity
- Volatile Acidity
- Citric Acid
- Residual Sugar
- Chlorides
- Free sulfur dioxide
- Total sulfur dioxide
- Density
- pH
- Sulphates
- Alcohol
- Quality

Red Wine

03. Methods/Analysis Plan

- How we want to find the valuable insights from the data



Analysis - Questions

Question 1

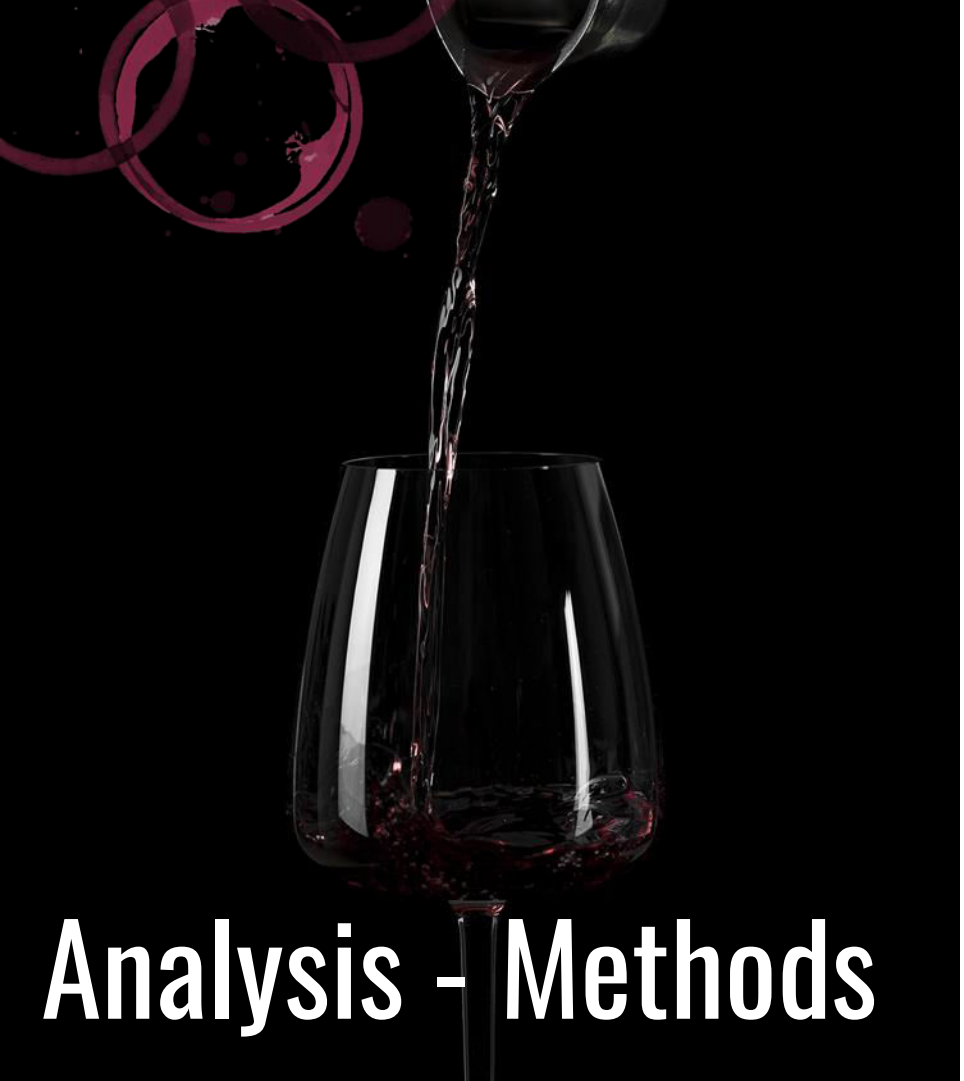
Which factor or combination of factors affect the quality of red & white wines?

Question 2

Do the different types of wines have different factors affecting quality?

Question 3

Is there any other interesting trends that exist in other columns besides quality?



Analysis - Methods

Method 1

Using Exploratory Data Analysis to find the distribution of factors of each Variable

Method 2

Using Machine Learning Methods to find the Best Composition factors to contribute the quality wine

Method 3

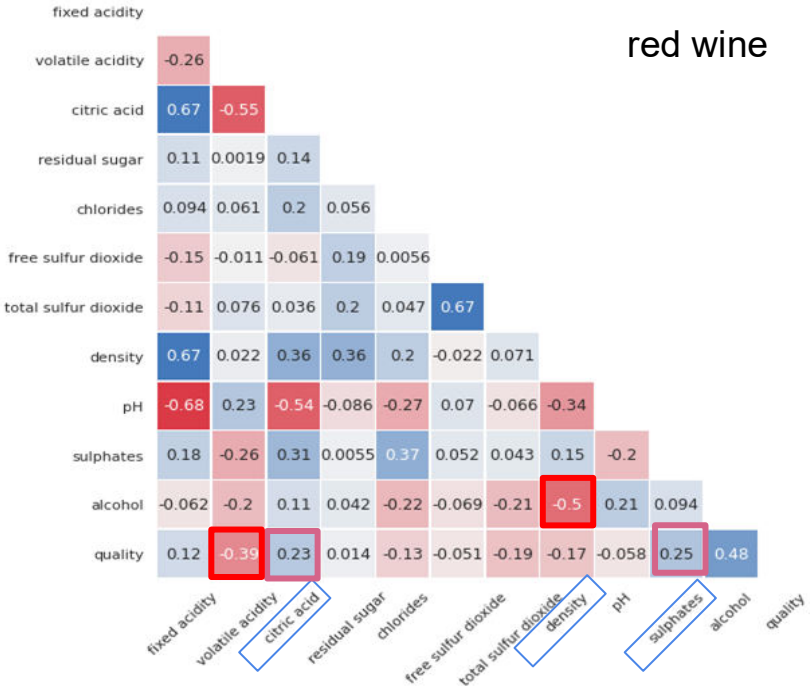
Explore Sales dataset and predict the price for the wine quality

04.
Data Analysis

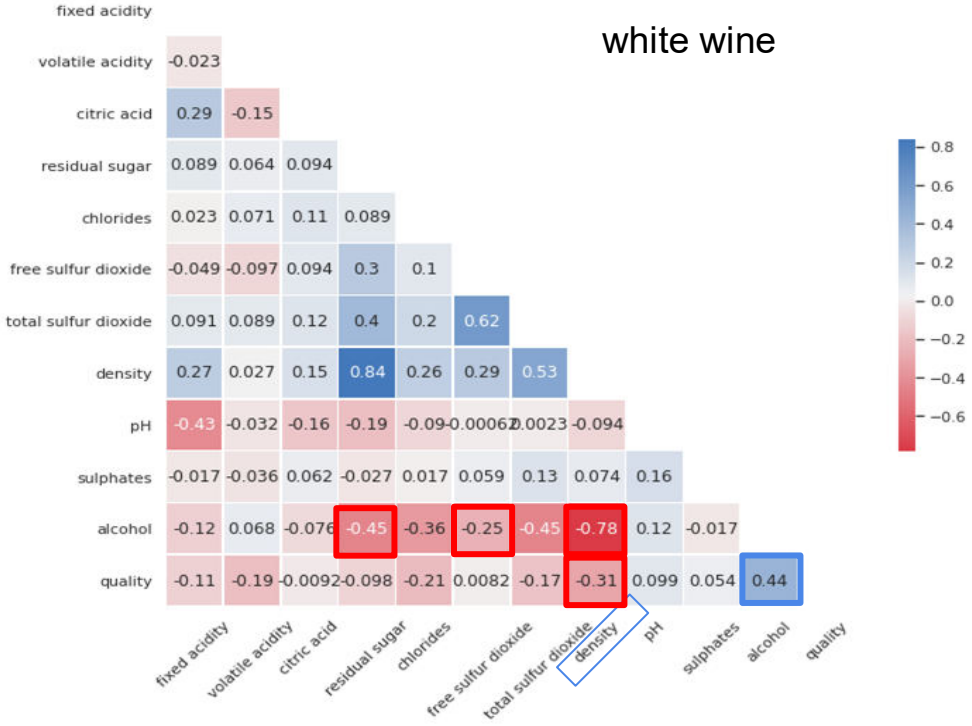
- Results - Exploratory analysis
- Feature Selection
- Machine Learning Methods

Correlation between all variables for red & white

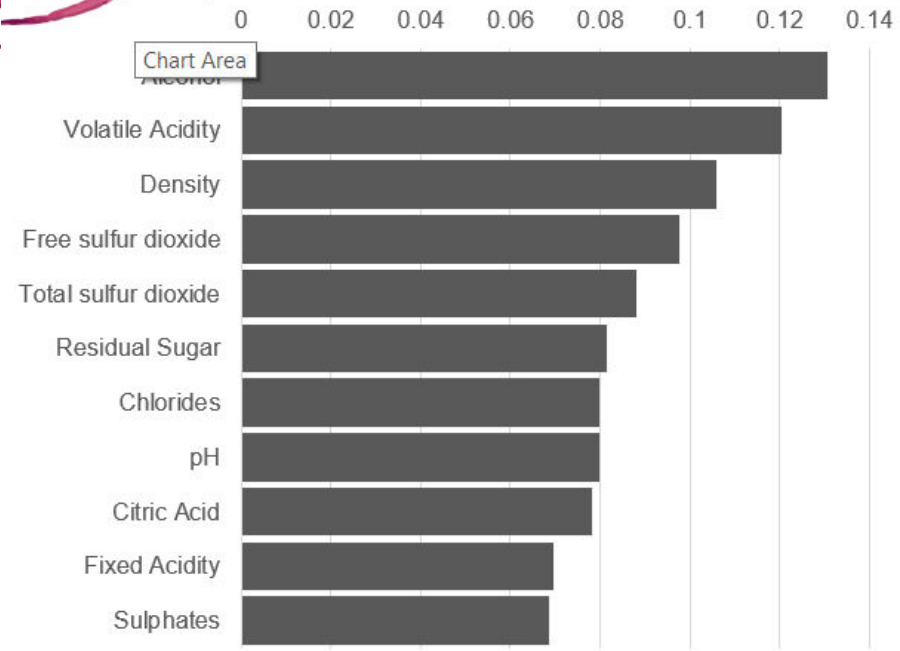
red wine



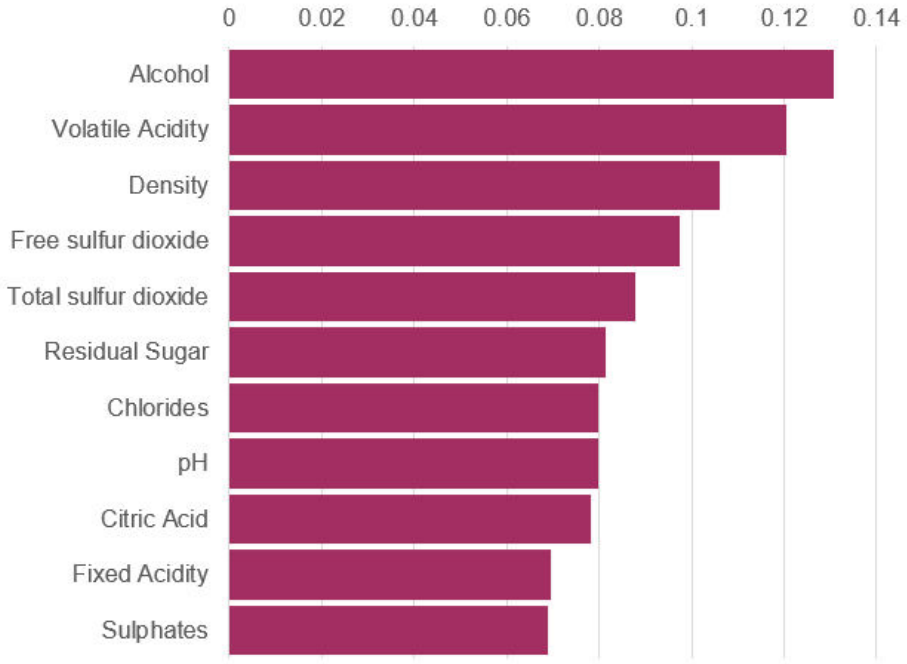
white wine



Important Factors towards predicting Wine Quality

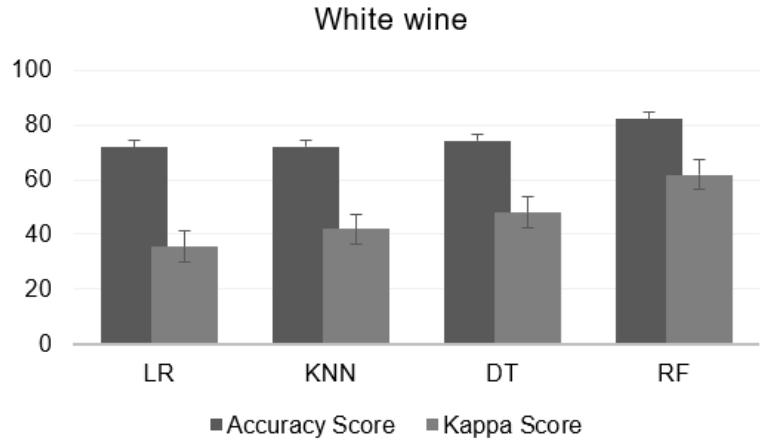


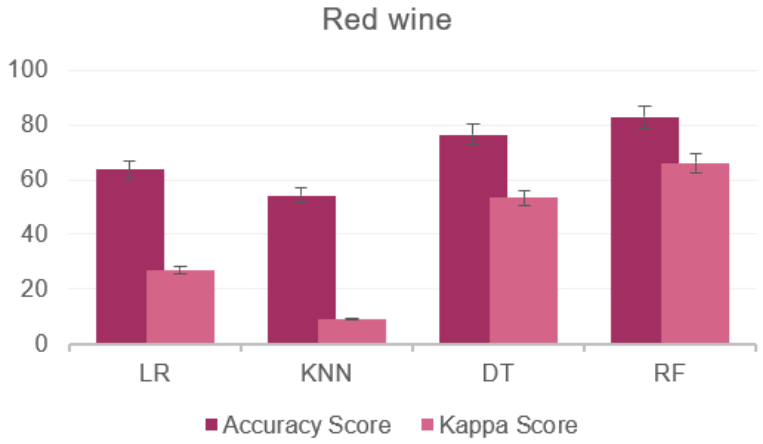
White Wine



Red Wine

Machine Learning Model Scores



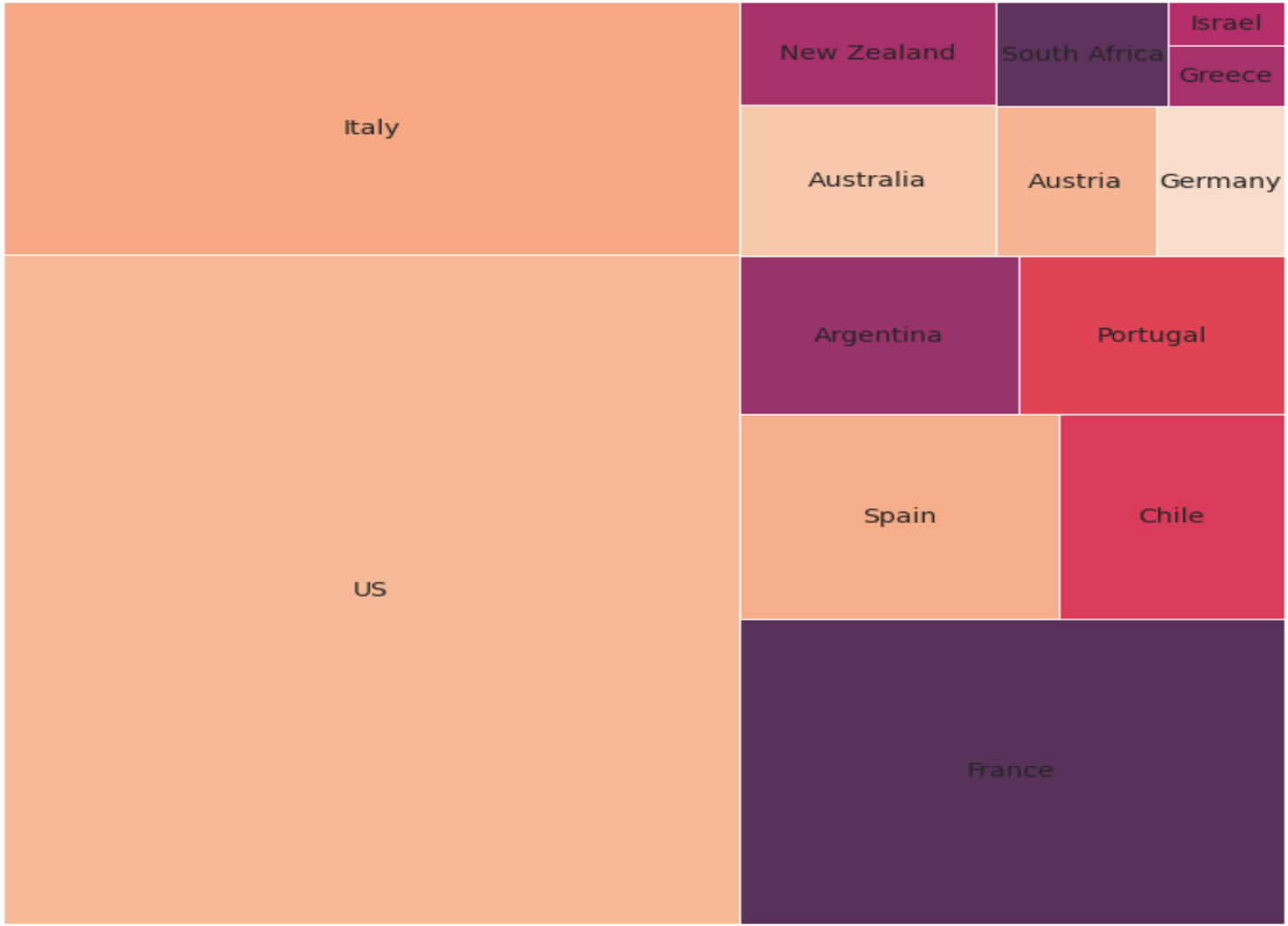


LR – Logistic Regression Algorithm
KNN – K Nearest Neighbor's Algorithm
DT – Decision Tree Classifier Algorithm
RF – Random Forrest Classifier Algorithm

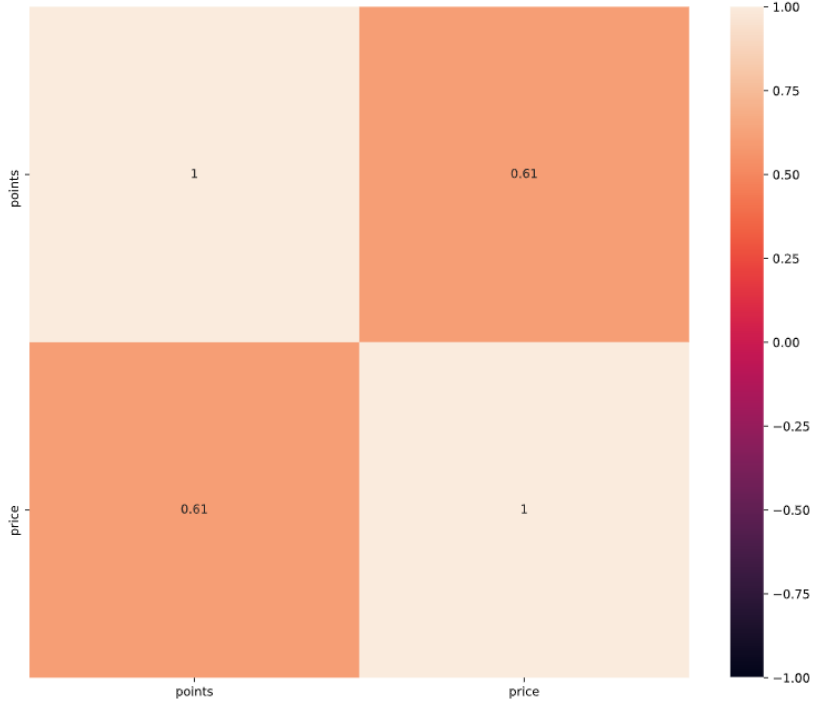
05.
Price Prediction

- Sales Dataset Exploration
- Price Prediction & Suggestion

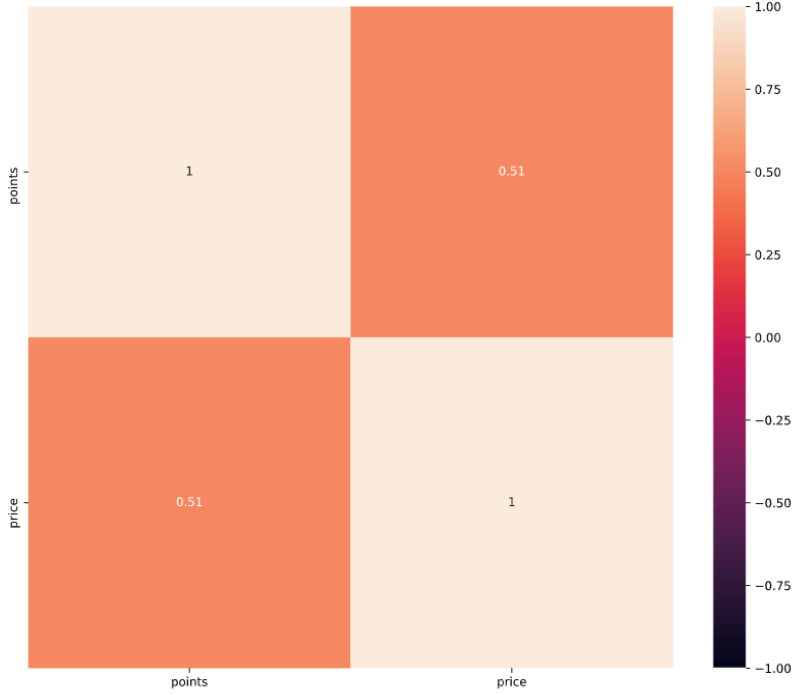
Sales data distribution



Points Vs Prize - Correlation



White Wine



Red Wine

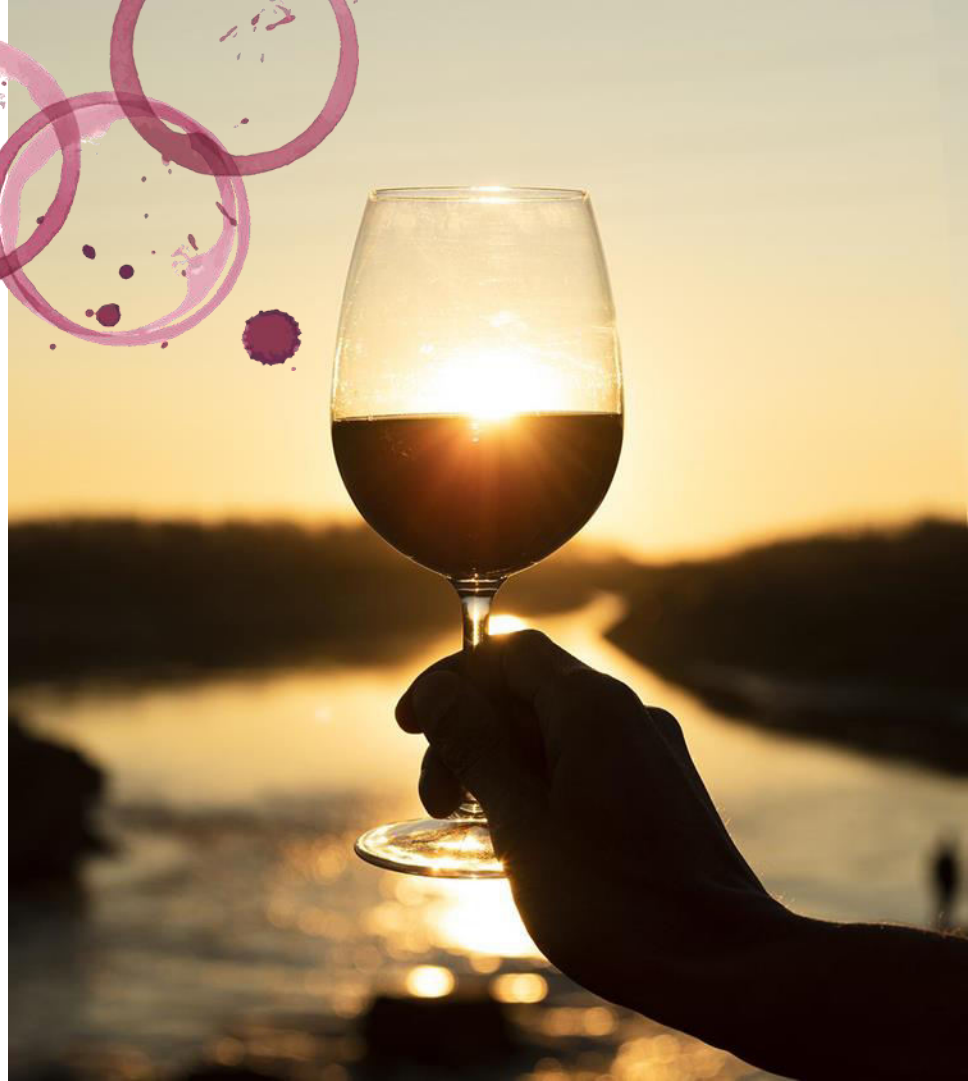
Comparison of estimated price for wines in dataset1 to the price in the dataset2

White Wine

Red Wine

THANKS

Do you have any questions?




Wine
Für dich, Für uns, Für alle