

Wine Score Prediction

CAPSTONE PROJECT



VIVINO

Problem

- New wines without any reviews have fewer buyers
- A wine without any scores is not trustworthy

Solution

- Regression model for wine score prediction
- Recommendation system based on keywords or similar wine

Impact on: Consumer satisfaction, wine producers, distributors, and retailers



The Dataset

→ Red wines from France

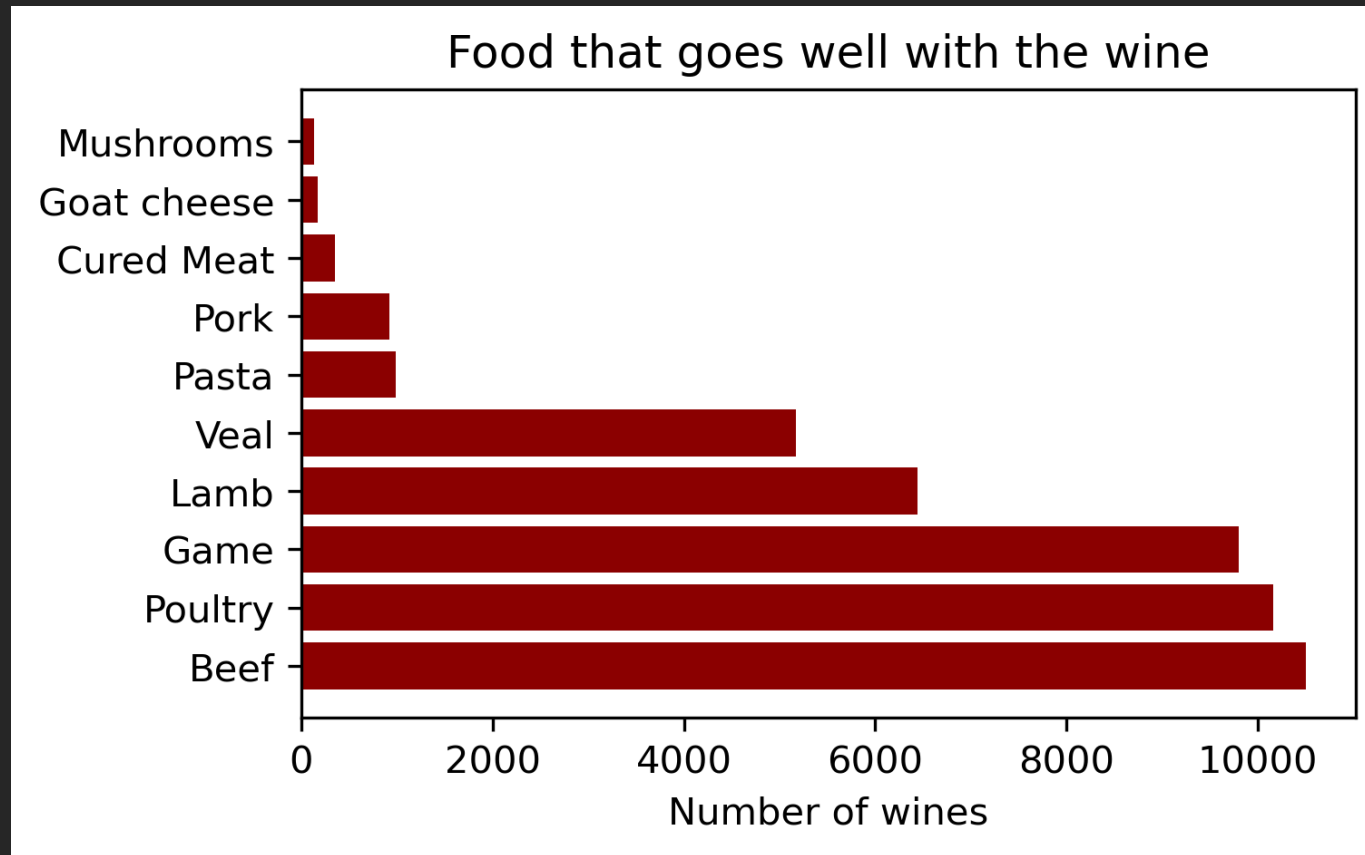
- Metadata
- Geographic Data
- Monetary Data
- Rating Data
- Taste Data
- Attribute Data

Data Cleaning

- Formatting: year, attribute data
- Duplicates: 0
- Missing Values: taste data, winery, description



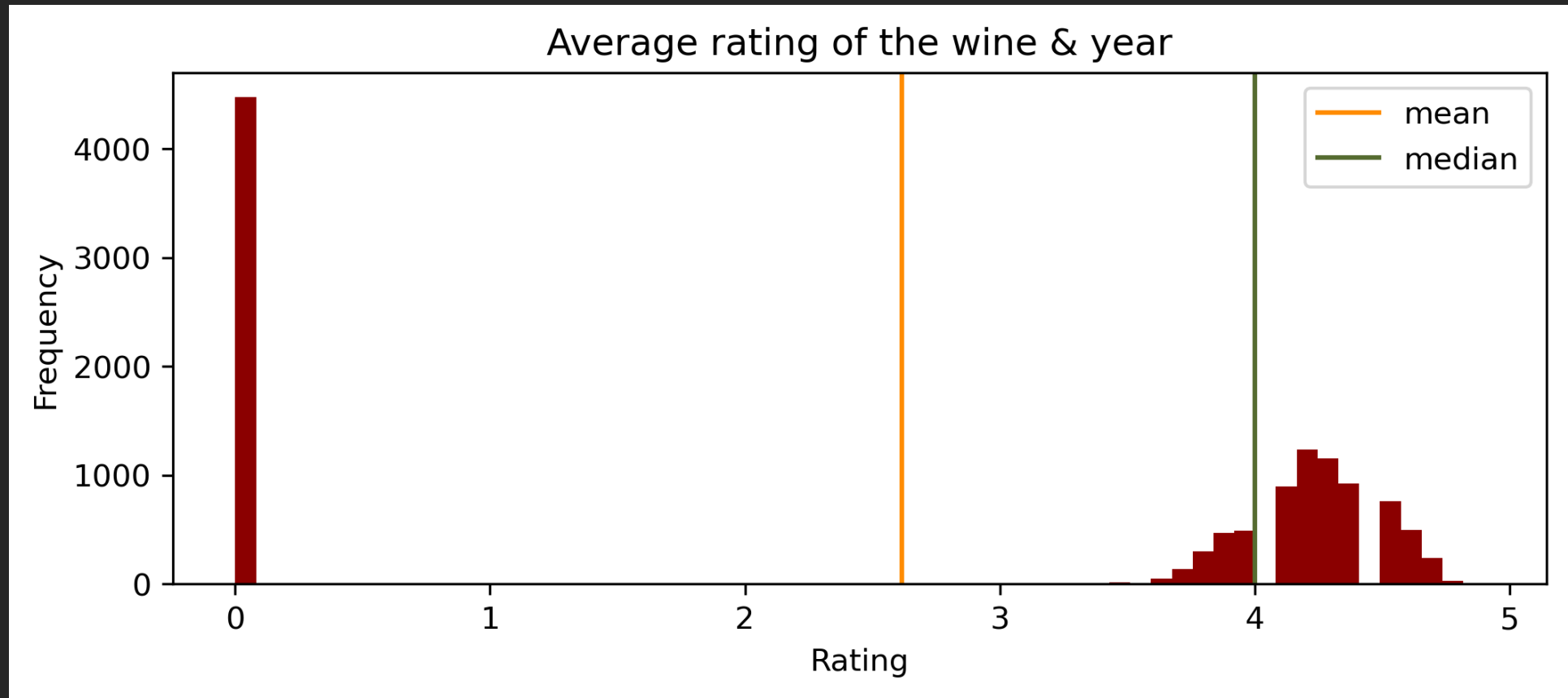
EDA – Categorical Columns



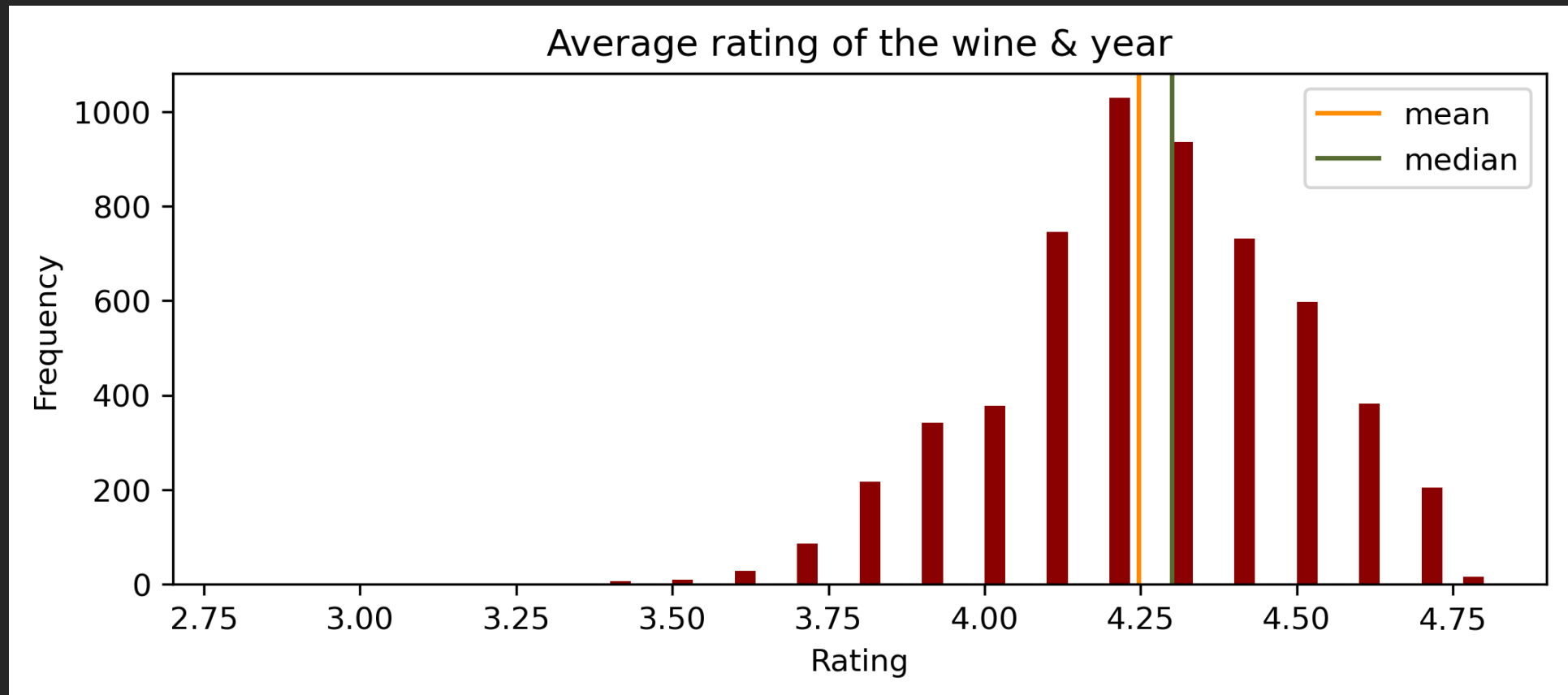
- Most French red wines in this dataset go well with beef, poultry and game



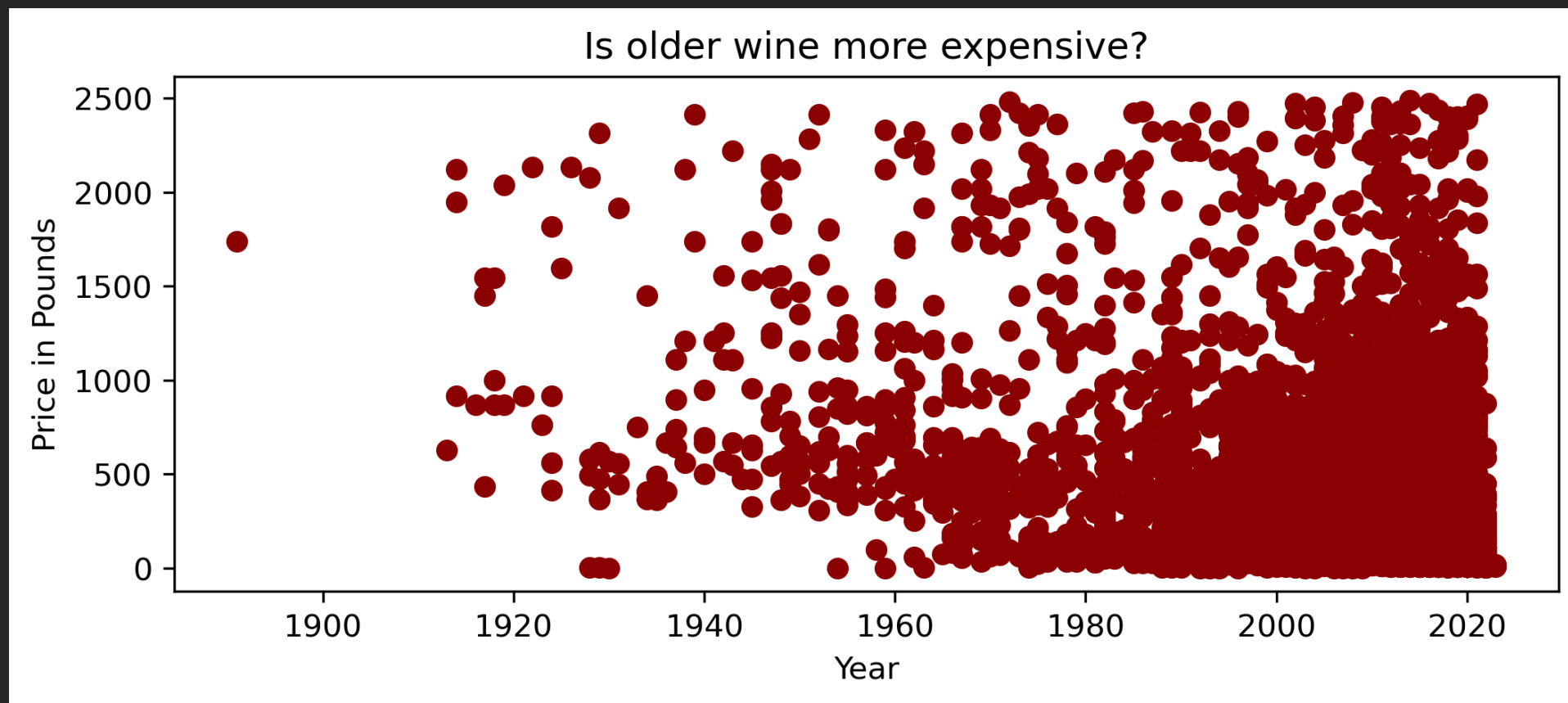
EDA – Numerical Columns



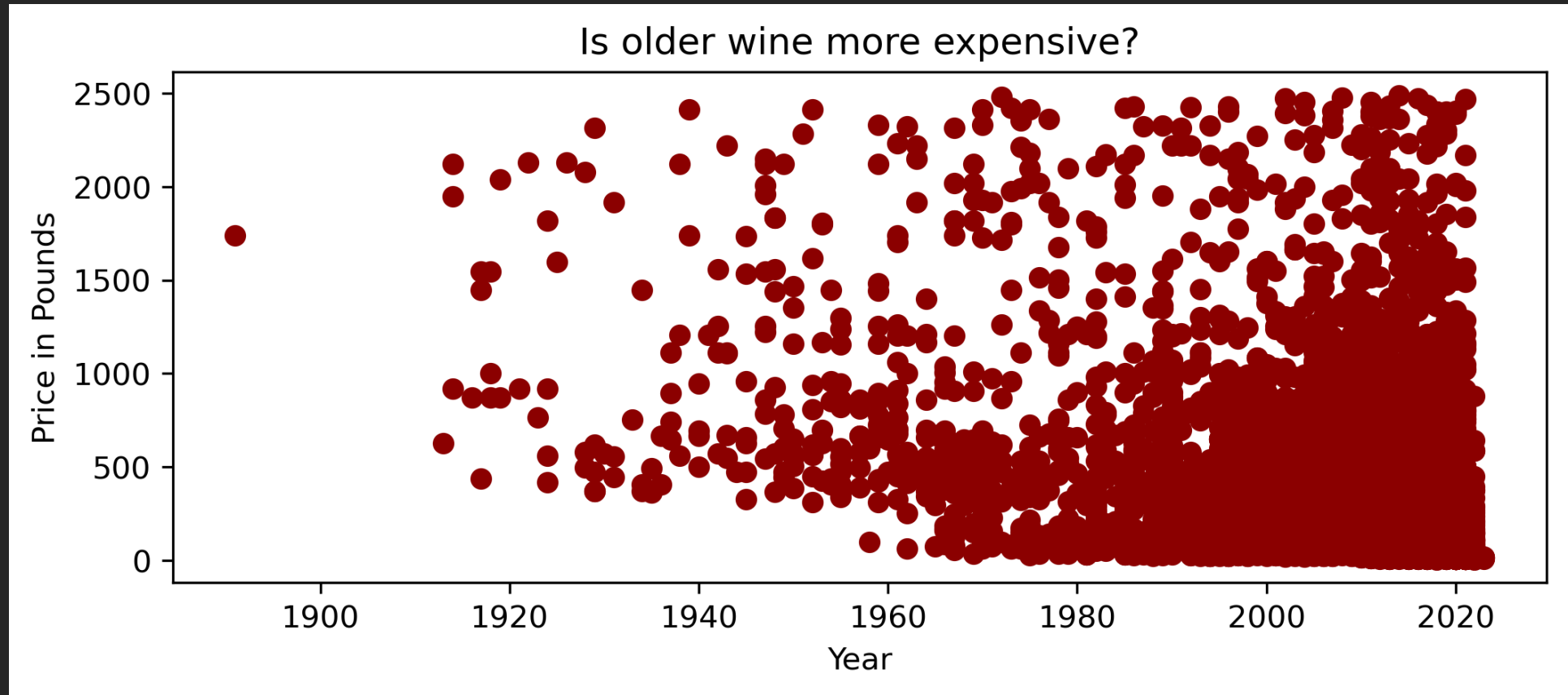
EDA – Numerical Columns



Advanced EDA



Advanced EDA



Next Steps

- EDA on specific questions and related to the target
ETA: 13.03.2024
- Preprocessing the data for modelling
ETA: 18.03.2024
- First Model: Linear Regression for predicting wine score
ETA: 22.03.2024

