

Wine Review Project Proposal

Choosing a Wine Based on Reviews

By Zoë Bakker

Intro

When deciding which wine to buy, one has to consider a number of factors:

- * Price
- * Wine type
- * Country of origin
- * Rating
- * And so on

Choosing a wine can be a complicated process, but, for most people, the goal is to choose the best wine at the best price.

Problem

Unless one is either a trained sommelier or has created some arbitrary rule for selecting wine

*I'll just always buy the **second-cheapest** bottle of Chardonnay*, trying to select the best value wine can be stressful.

I aim to make the process less stressful by creating a confidence-boosting program that can take away the guesswork in wine buying by accurately predicting wine value based on an analysis of a number of factors included in wine reviews.

Clients

- Wine buyers for restaurants, hotels, event spaces
- Wine clubs, wine subscription services
- Professionals who buy client gifts meant to impress
- Average consumers

The Dataset

Wine review will be pulled from:

- * Kaggle
- * Global Wine Database
- * Wine HQ

The Approach

1. Data Wrangling and Cleaning

- Formatting the databases in the same manner so they can be combined
- Dealing with NULL values
- Re-formatting characters that were corrupted on download

2. Exploratory Analysis

- Perform regression analysis to determine which variables have the greatest impact on wine value
- Identify patterns that connect different variables

3. Prediction

- Apply data science techniques to make predictions about wine based on a number of variables

Deliverables

- Presentation of findings
- Summary of findings
- R code