What Kind of Whiskey Gets High User Ratings?

An analysis of specific Total Wine spirits

Will Nobles

Introduction

Motivation:

- assist distillers in creating a marketable product

Objectives:

- investigate user rating predictability

Goals:

- show a correlation between features and user rating

Methodology

Data:

- https://www.totalwine.com/
- 1198 rows, 11 columns

Metrics:

- whiskey name
- critic rating, rating source, user rating, number of reviews, price, brand, country, spirit type, taste notes

Methodology (continued)

Models:

- Linear Regression through regularization, feature engineering

Tools:

- BeautifulSoup
- Selenium
- scikit-learn

Methodology (continued)

Features

- whiskey name
- critic rating
- rating source
- **user rating**
- **number of reviews**

- **price**
- **brand**
- **country**
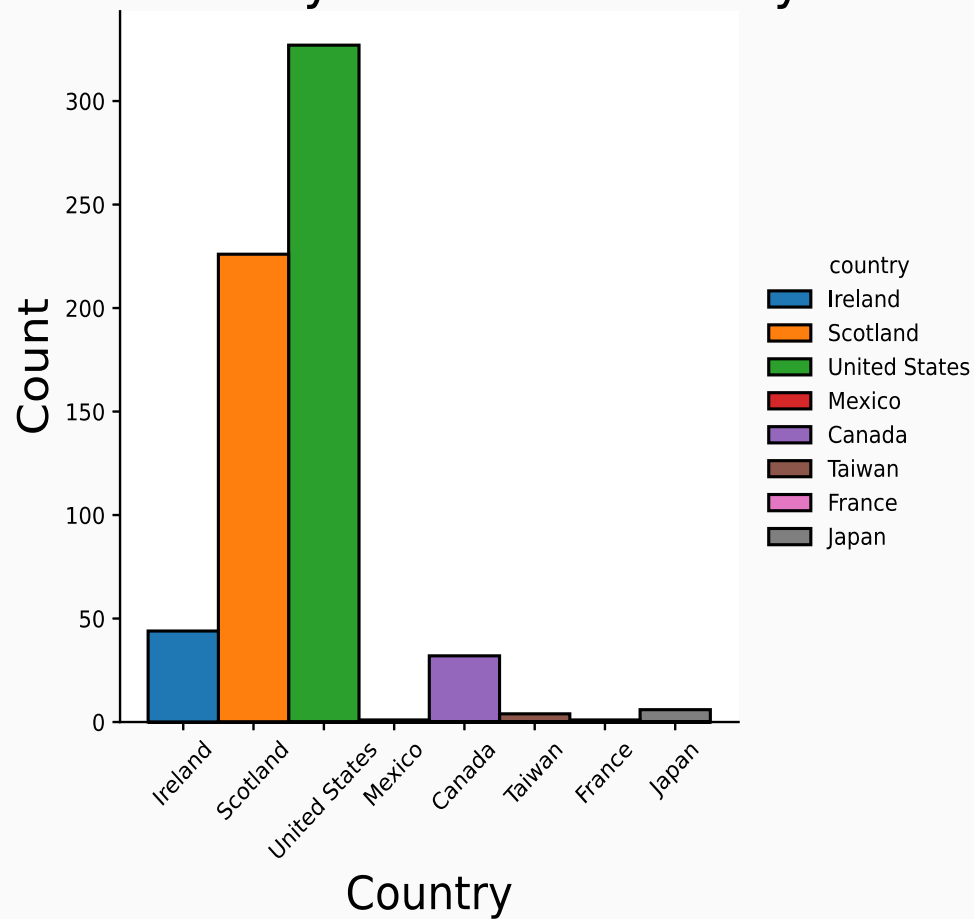- **spirit type**
- taste notes

Results and Insights
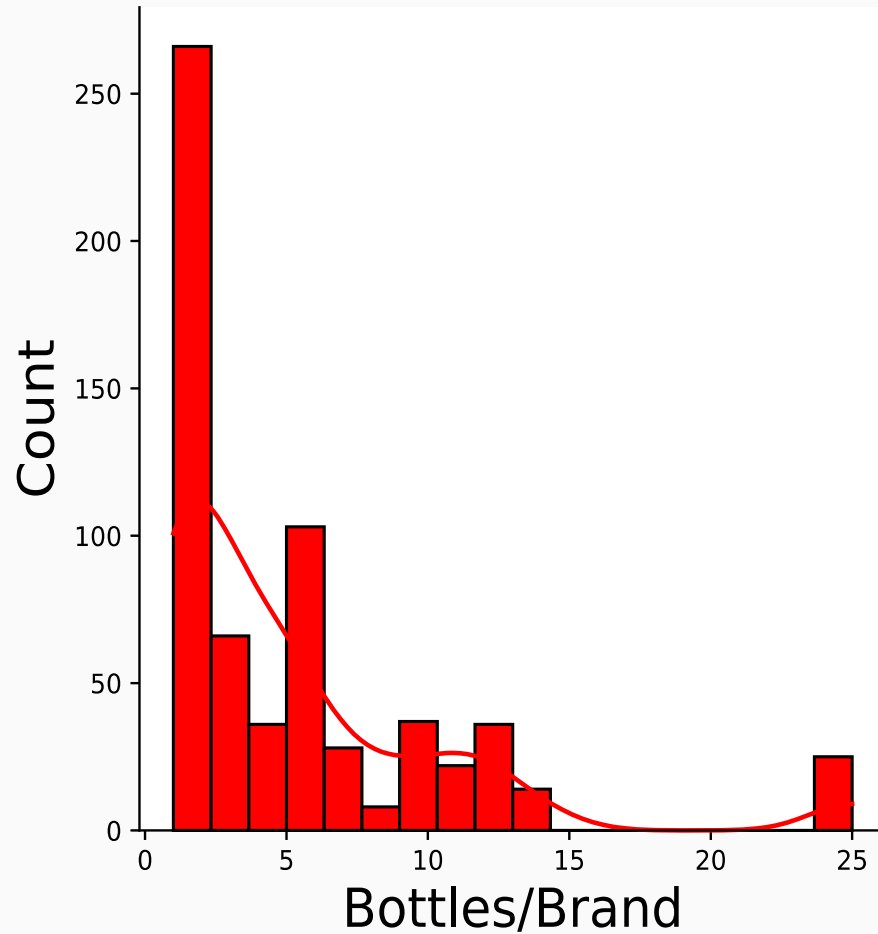
4.25

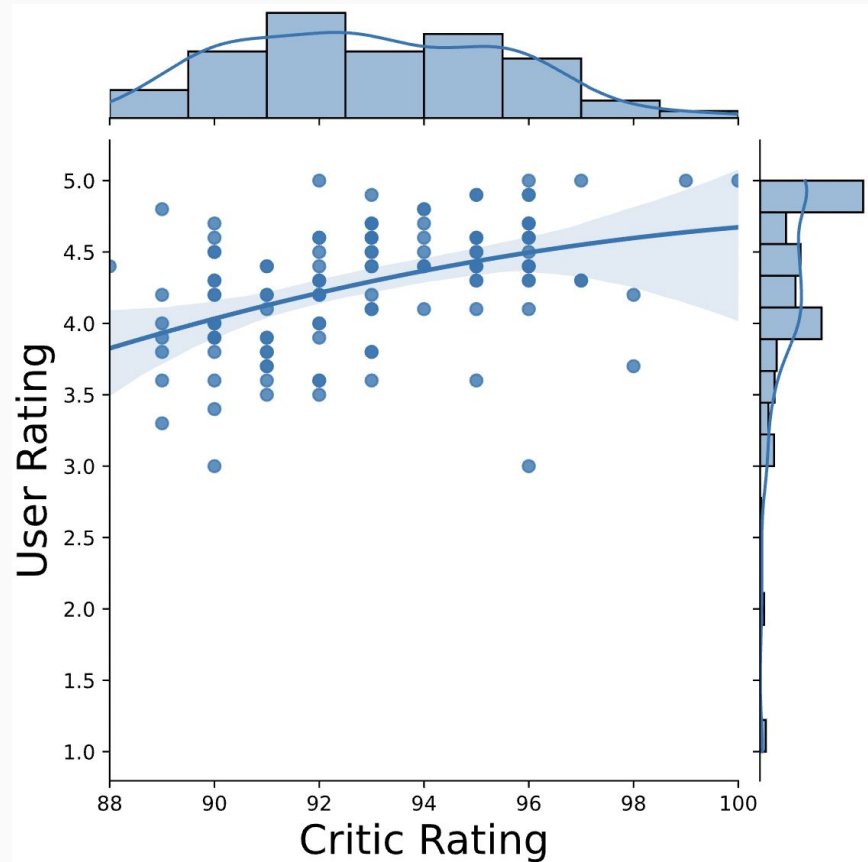Average User Rating

$6.99 - $19999.99

Bottle Price Range

Number of Whiskeys Manufactured by Country

Plot of Counts for Number of Bottles Produced Per Brand

Joint Plot Showing a Polynomial Relationship Between User Rating and Critic Rating

Train

Test

0.593 -6.196×10^{24}

Simple Linear Regression R^2 Values

Train
0.063

Test
0.051

LASSO Cross-Validation R^2 Values

Conclusions

Recommendations:
- Professional critic ratings are important

Interesting insights:
- Taste notes are similar
- Several brands account for most of the whiskey produced

Future Work

- Categorize the taste column
- Filter based on the top brands

"There is no bad whiskey. There are only some whiskeys that aren't as good as others."
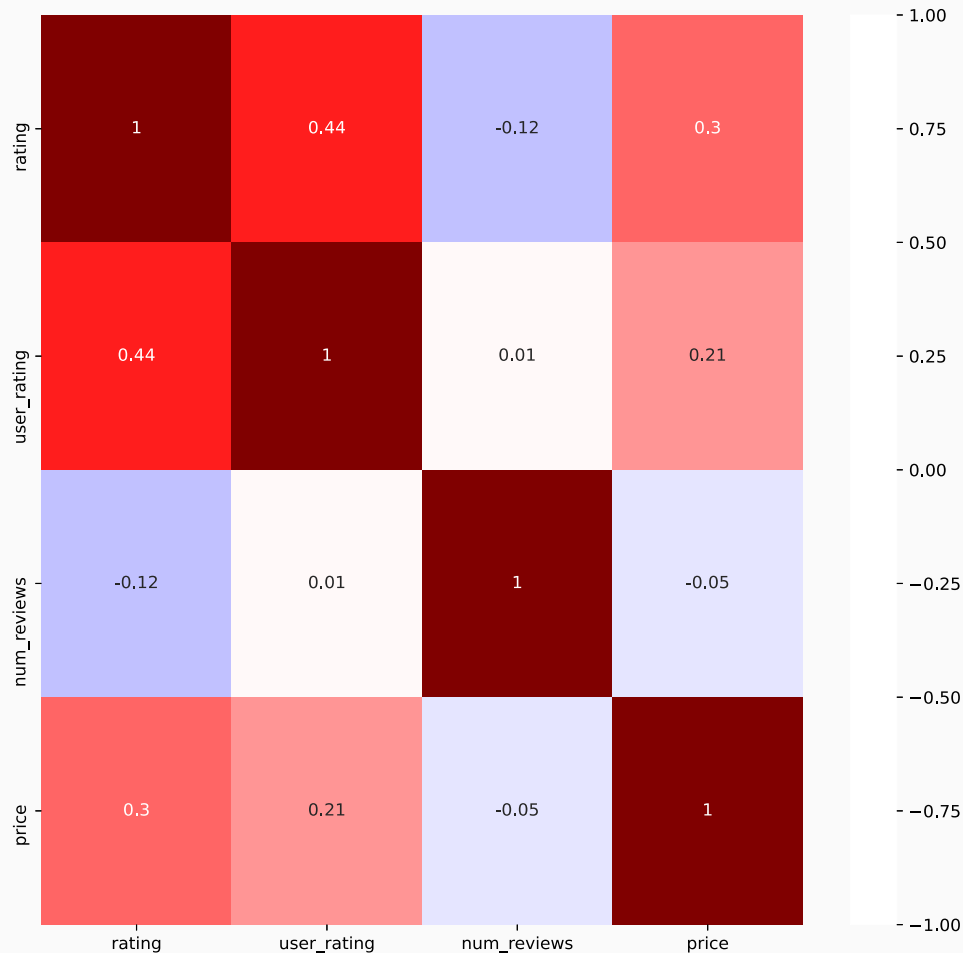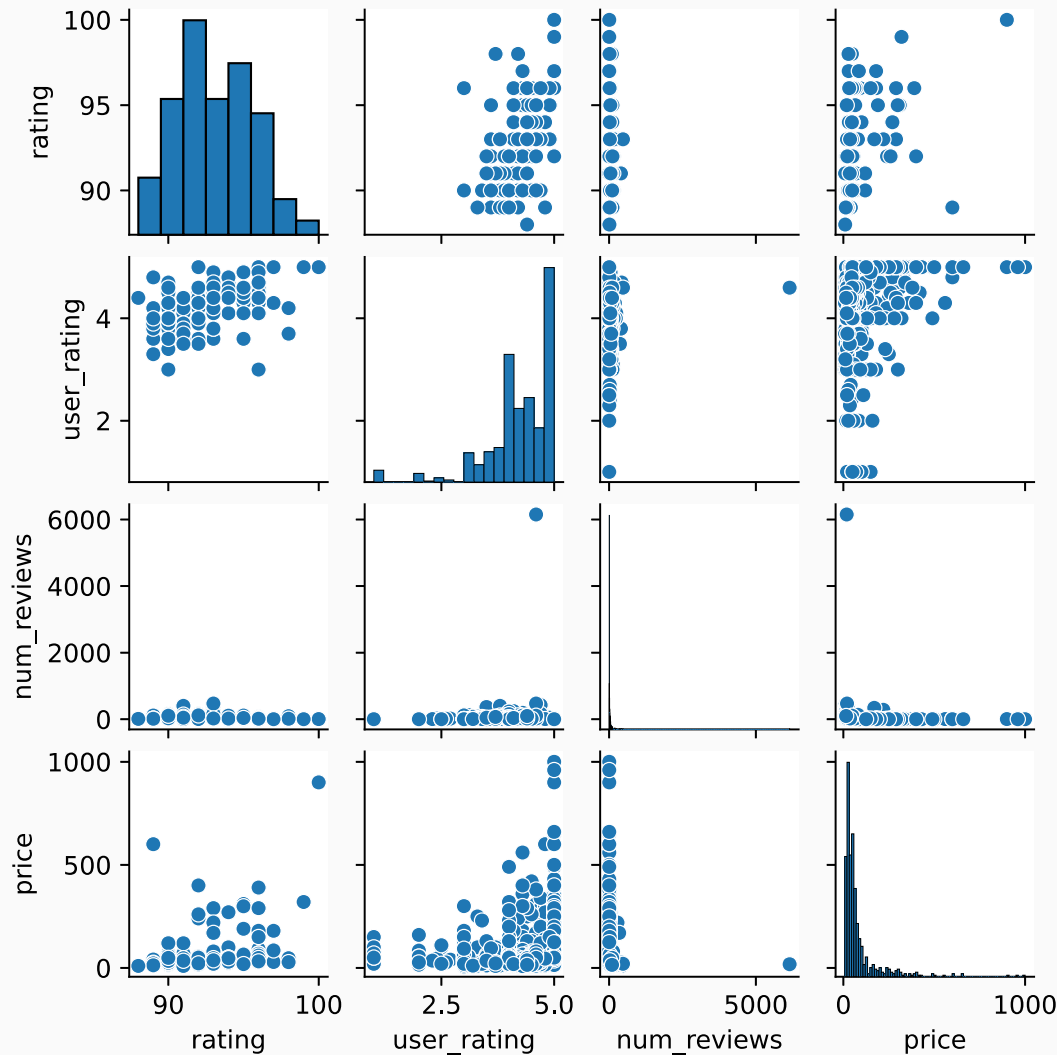
- Raymond Chandler

Appendix

Heat map showing correlations between features

Appendix

Pair plot showing correlations between features

Thank you!