# Predicting Whiskey Ratings

Julia Nguyen

Value of High Ratings

- Consumers often research and rely on expert opinions
  - Overwhelming number of options for whiskey
  - Bottles are often expensive
- Liquor stores often provide 'expert' ratings
- Generally higher ratings correspond to higher quality whiskey

Methodology

- Data was scraped from WhiskeyAdvocate.com

  - 5883 different whiskey rated by experts

  - Scores from 60 to 97

- Variables: Age, price, alcohol by volume, type of whiskey, and descriptors, such as vanilla, smoke, spice, fruit, etc.

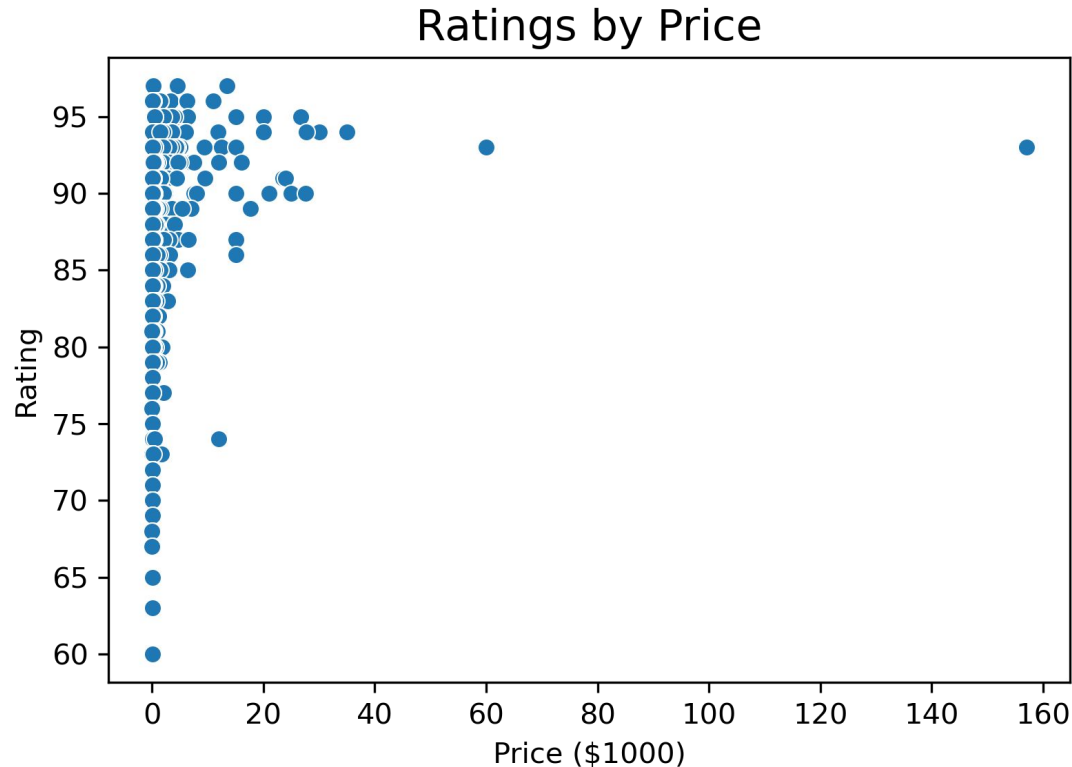- Target: Score

- Model: Linear Regression

Preliminary Data Analysis: Categories

Ratings by Whiskey Categories

Preliminary Data Analysis: Categories

Ratings by Whiskey Categories
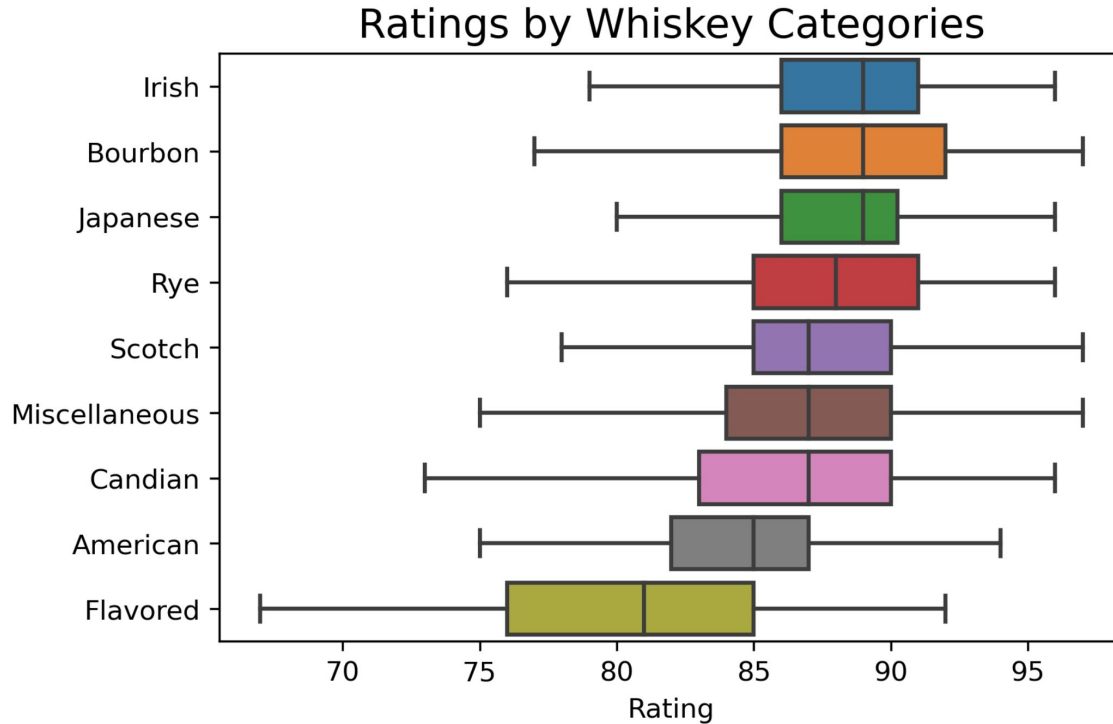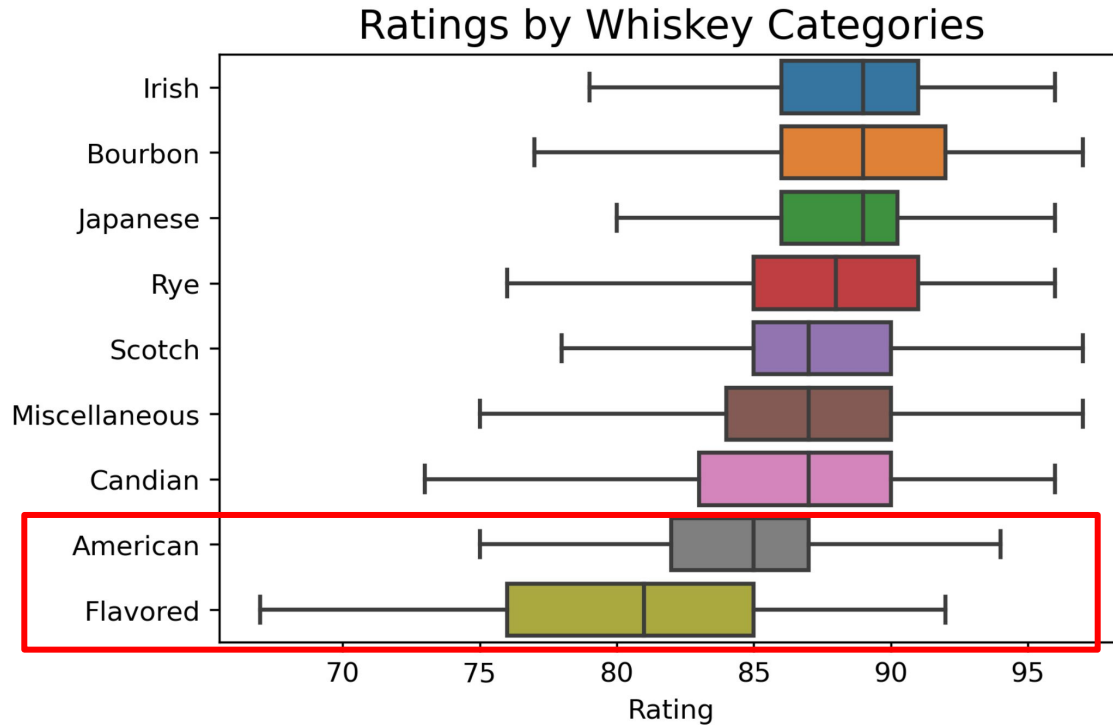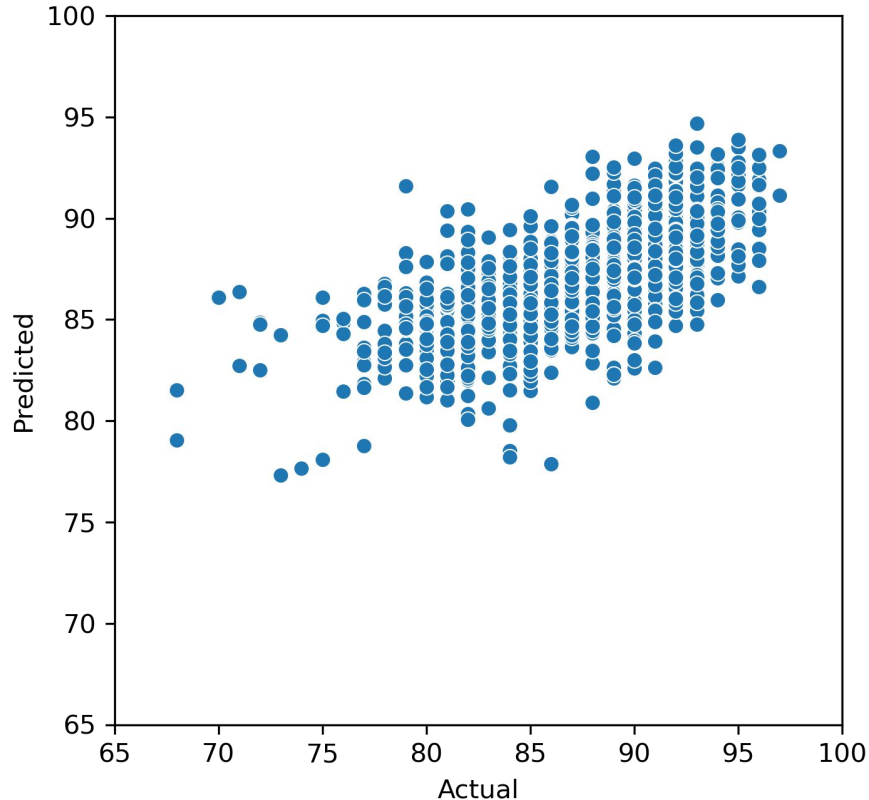
Preliminary Data Analysis: Price

Ratings by Price

Preliminary Data Analysis: Price

Ratings by ln(Price)
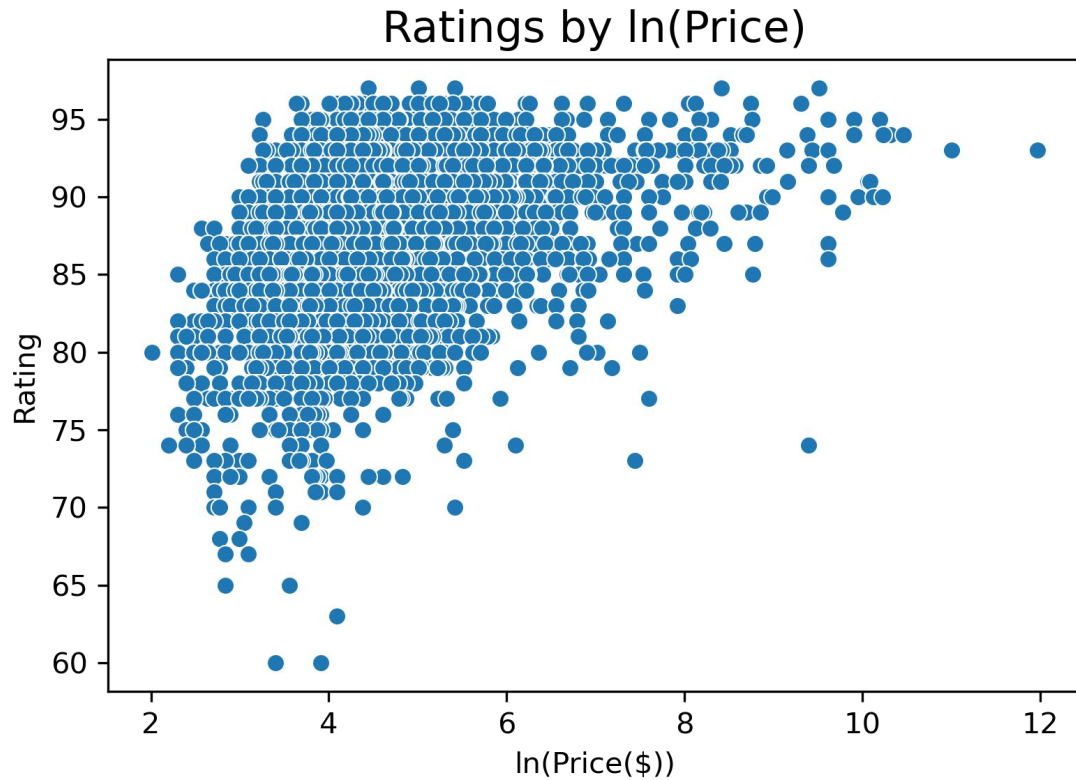
Linear Regression Model:

- Strong Correlation between Actual and Predicted

- R^2 = 0.44

- RMSE = 3.4

Linear Regression Model: Selected Coefficients

Variable Name	Coefficient
ABV	0.073
Age	0.152
Flavored Whiskey	-0.041
Irish Whiskey	3.624
Leather	0.687
Sweet	-0.013

Linear Regression Model: Selected Coefficients

Variable Name	Coefficient
ABV	0.073
Age	0.152
Flavored Whiskey	-0.041
Irish Whiskey	3.624
Leather	0.687
Sweet	-0.013

Linear Regression Model: Selected Coefficients

Variable Name	Coefficient
ABV	0.073
Age	0.152
Flavored Whiskey	-0.041
Irish Whiskey	3.624
Leather	0.687
Sweet	-0.013

Linear Regression Model: Selected Coefficients

Variable Name	Coefficient
ABV	0.073
Age	0.152
Flavored Whiskey	-0.041
Irish Whiskey	3.624
Leather	0.687
Sweet	-0.013

Recommendations

Target

- Older, higher ABV, higher price
- Categories: Japanese whisky, Irish whisky, scotch, bourbon
- Flavors: Cinnamon, chocolate, smoke, leather
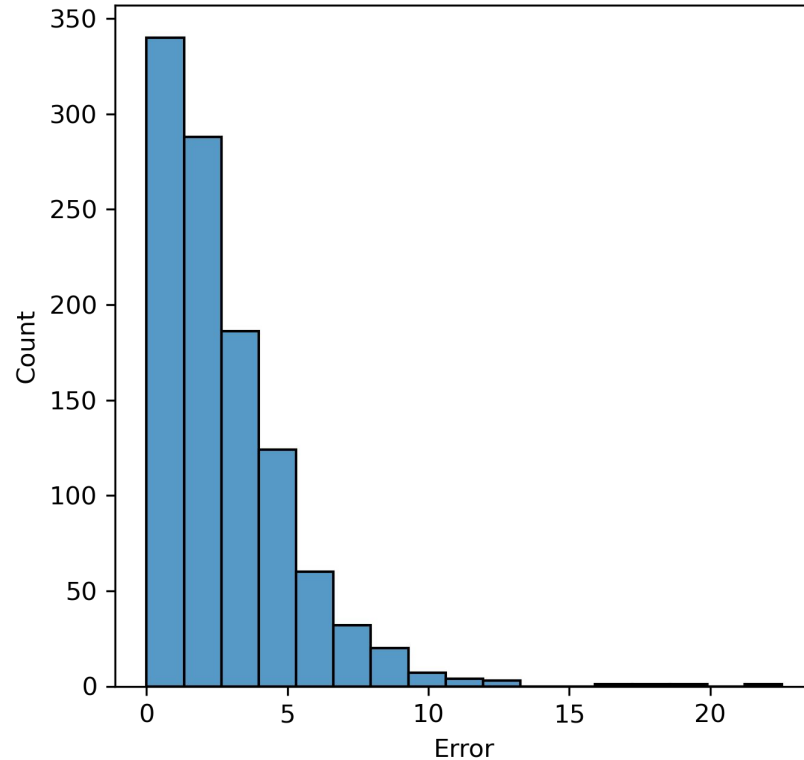- Complexity and richness

Avoid

- Categories: English whisky, Flavored whiskey, blended whiskeys
- Flavors: Sweet, floral, citrus, tropical, pepper

Future Work

- Combine with datasets that provide volume of whiskey sold and sales
- Analyze specific category of whiskey

Questions?

Appendix: Error Distribution

Appendix: Top 5 Most common Adjectives

Adjective	Count
Vanilla	2287
Sweet	2253
Fruit	1872
Chocolate	1680
Spice	1480

Appendix: Recommendation Coefficients

Variable	Coefficient
Ln(price)	269017.52
ABV	0.073
ln(Age)	0.153
Japanese Whisky	3.173
English Blended Whisky	-5.203
Complexity	1.357
Rich	1.134
Balance	0.759

Variable	Coefficient
Cinnamon	0.813
Chocolate	0.561
Smoke	0.588
Leather	0.687
Sweet	-0.012
Floral	0.258
Citrus	0.256
Pepper	0.227