

Buy Better Wine

**Predicting Wine Ratings with
Linear Regression**

Matt Segall

Why predict wine ratings?



Hypotheses



Price



Vintage



Geography



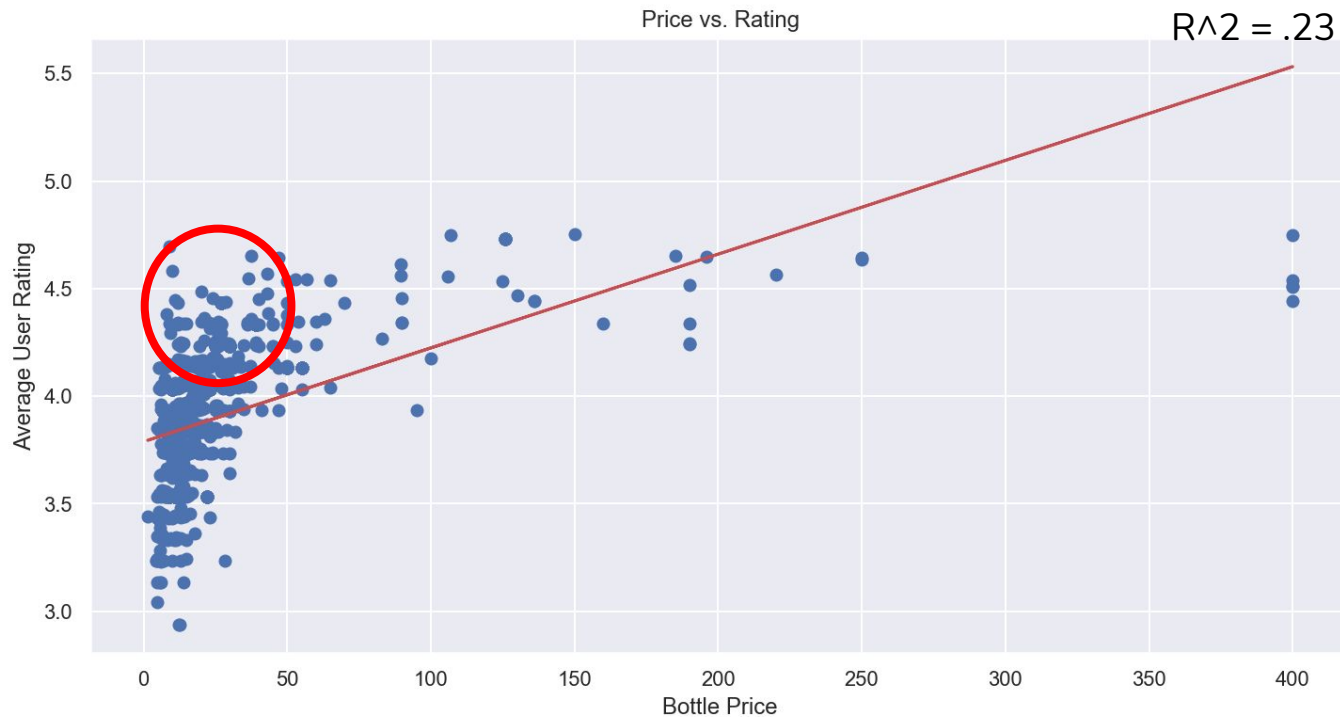
Style

Data

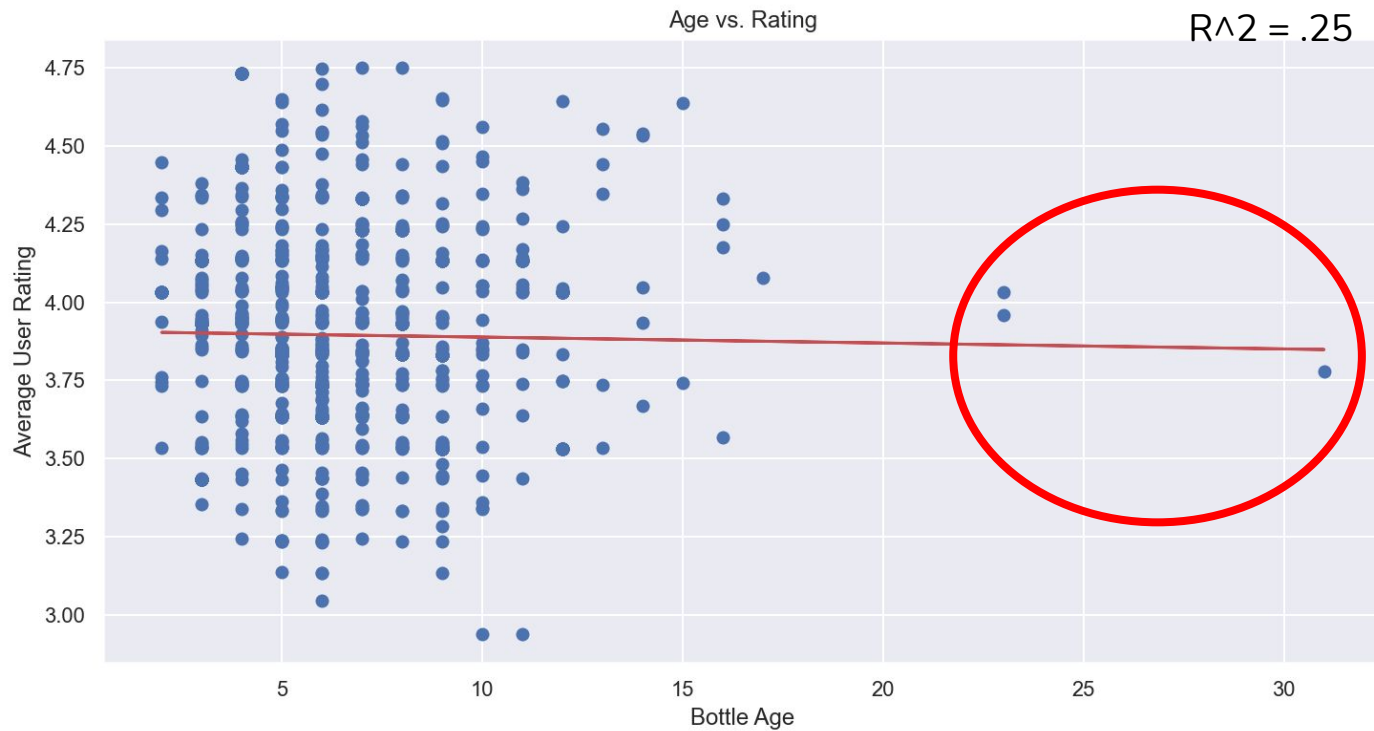
- ~4,700 wines with 9 attributes
- Target: average user rating
- Features
 - Winery
 - Wine Name
 - Vintage
 - Type (Red, White, Rose, Sparkling, etc)
 - Region
 - Country
 - Total # of Reviews
 - Price



Price



Vintage



Further Refining

- Individual transformations
- Interaction terms
- LASSO



Top Regional Specialties

1. **White Wines from US** ($\beta = .016$)
2. **White Wines from Italy** ($\beta = .015$)
3. **Sparkling Wines from France** ($\beta = .013$)
4. **Whites Wines from Germany** ($\beta = .012$)
5. **Rosés from Portugal** ($\beta = .010$)



Looking Forward



- Further refine current model
- Other models could be better fit
- New data could bring exciting insights
 - Climate/soil
 - Growing methods
 - Size of operation
 - Natural wine trends
 - Tasting notes

Questions? Cheers!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**.

