

# Imbibing in NLP: Clustering of Wine Reviews

---

Michelle L. Gill

Metis NYC, DS8

August 19, 2016

# Objective

---

- Supervised machine learning has predicted price fluctuations in fine wine
- Occurrence of vintage years predicted by weather
- Improved recommendation systems and forecasting of economic factors can increase profits

"Machine Learning in Fine Wine Price Prediction", *Journal of Wine Economics*, 2015, <http://bit.ly/2bFiN1M>

"Quants and quaffs", *The Economist*, 08/08/2015, <http://econ.st/1MeIQqP>

# Objective

---

- Supervised machine learning has predicted price fluctuations in fine wine
- Occurrence of vintage years predicted by weather
- Improved recommendation systems and forecasting of economic factors can increase profits

Goal: Understand how language of expert wine reviews relates to wine variety and quality

# Data and Tools

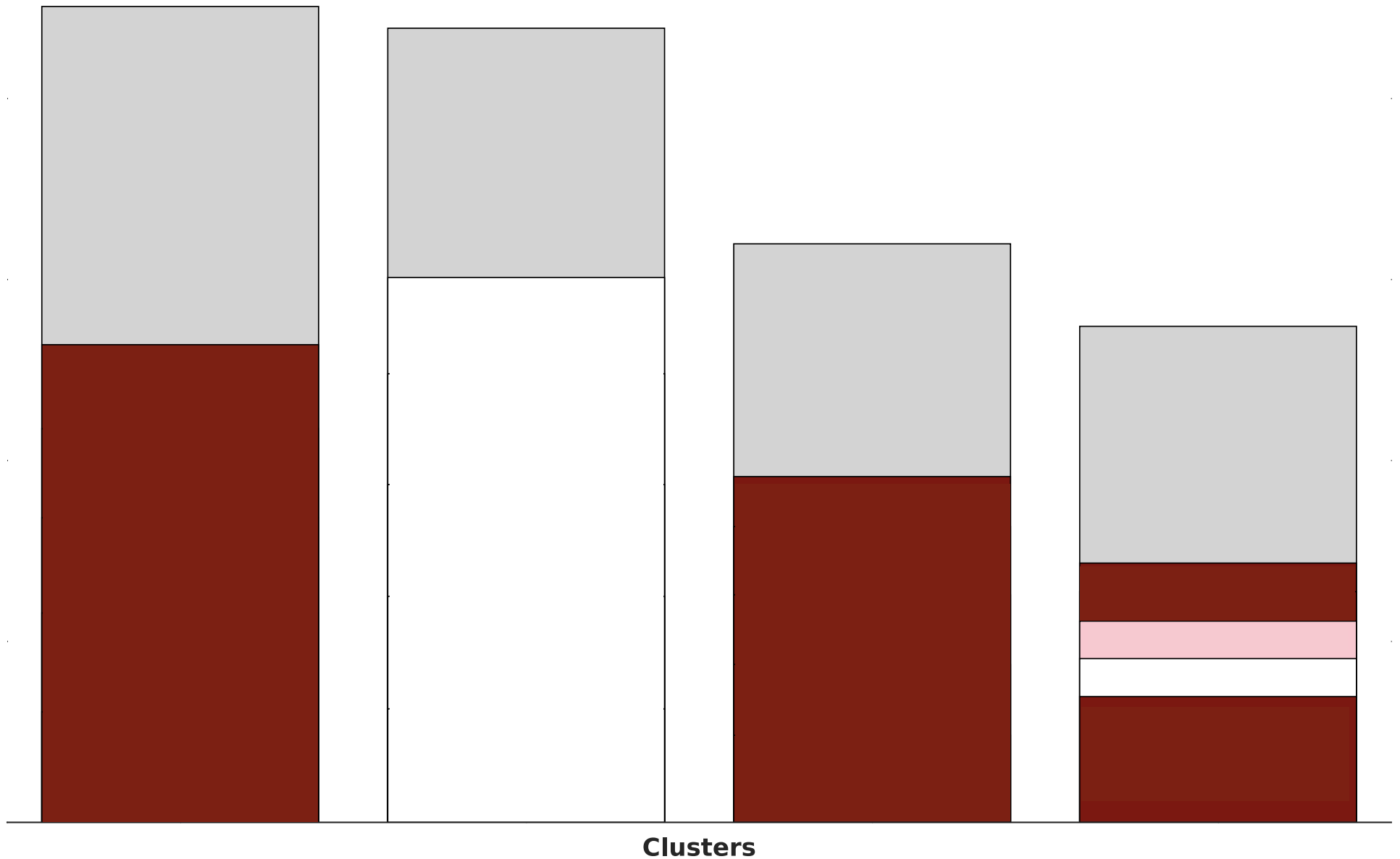
---

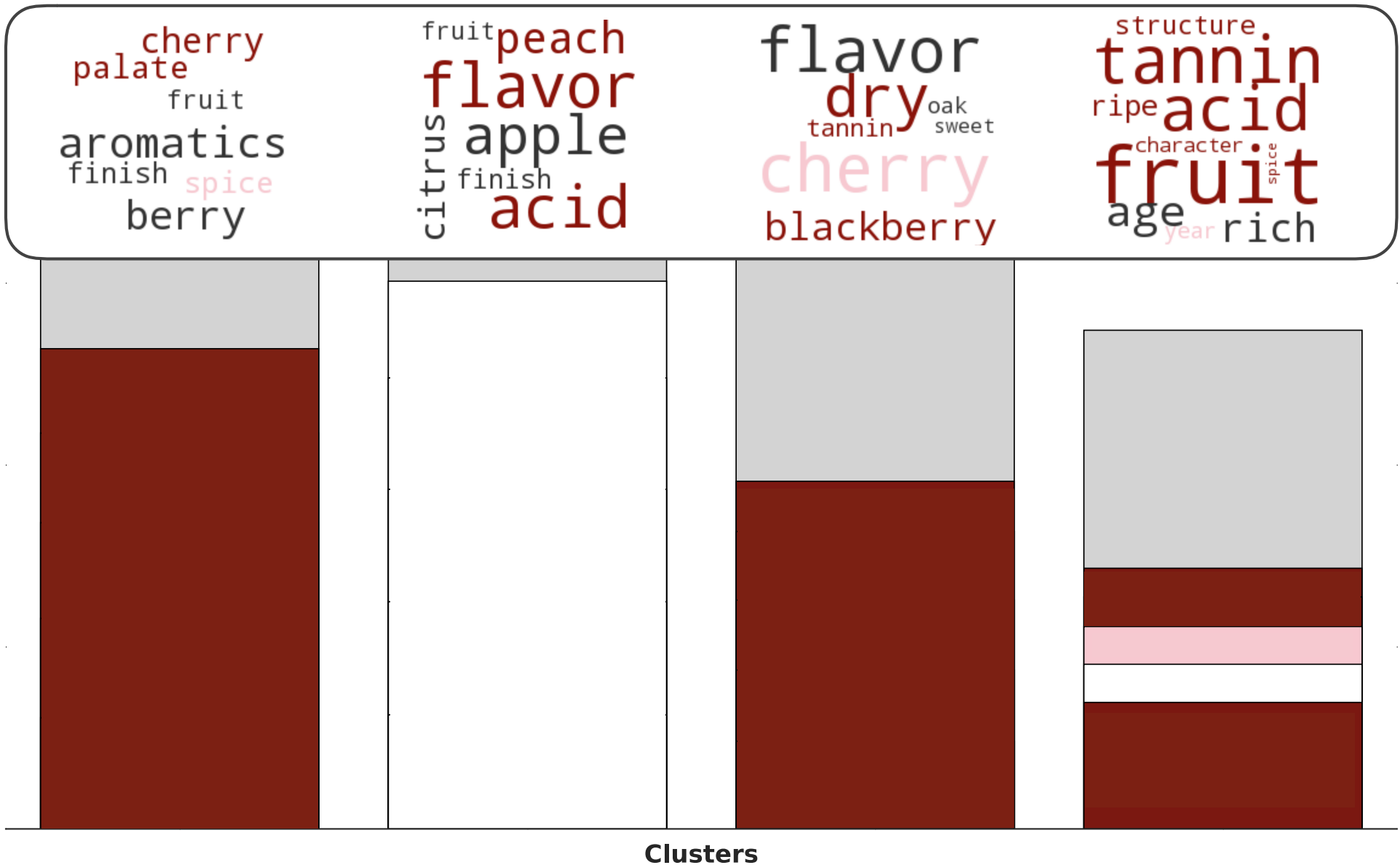
- Wine reviews (190,000) scraped from Wine Enthusiast
- Undersampled to 30,000 reviews
- Tokenization with lemmatization
- Linkage analysis, latent semantic analysis (LSA), latent dirichlet allocation (LDA), and sentiment analysis used



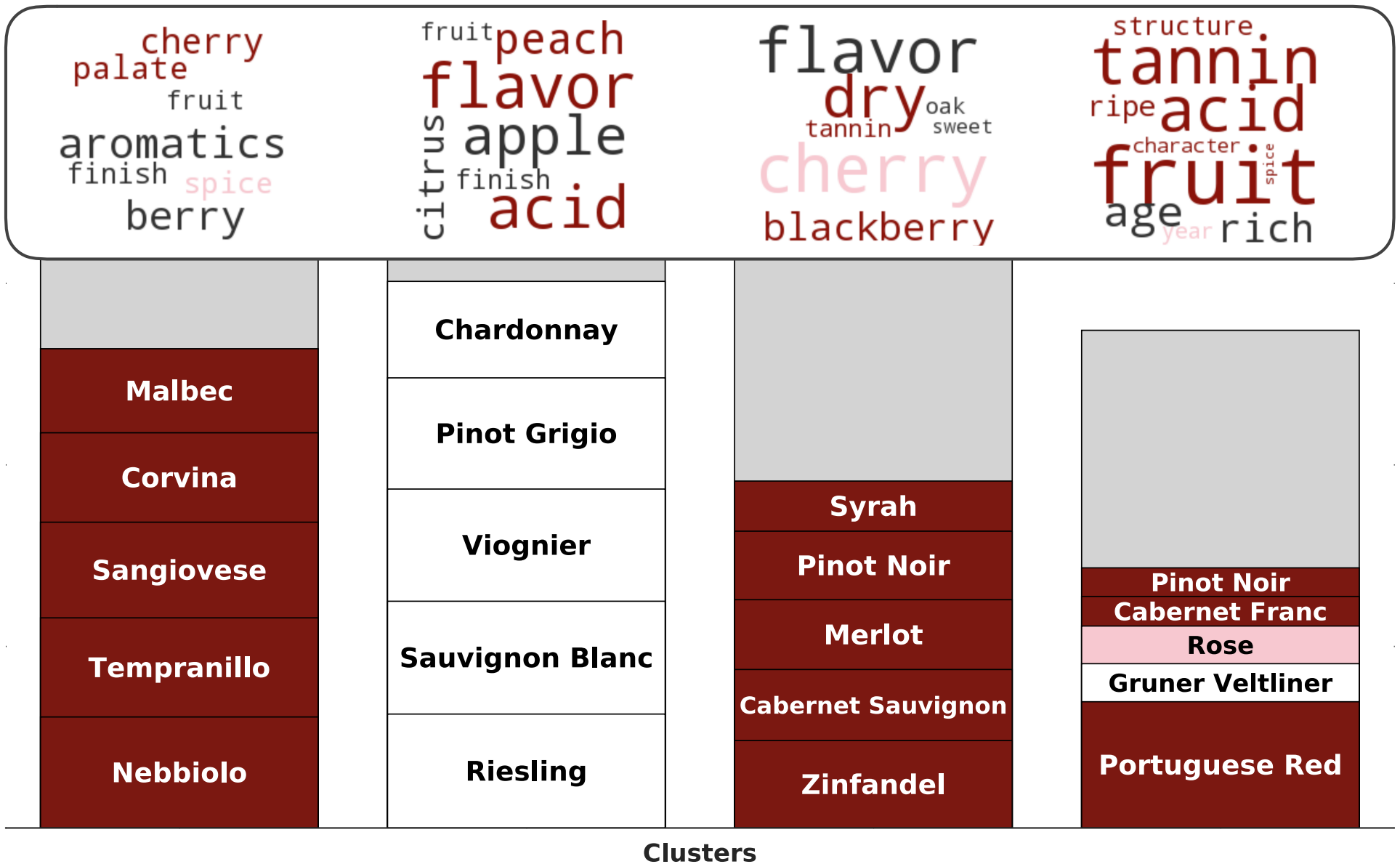
# Varietal Clustering of Reviews

---



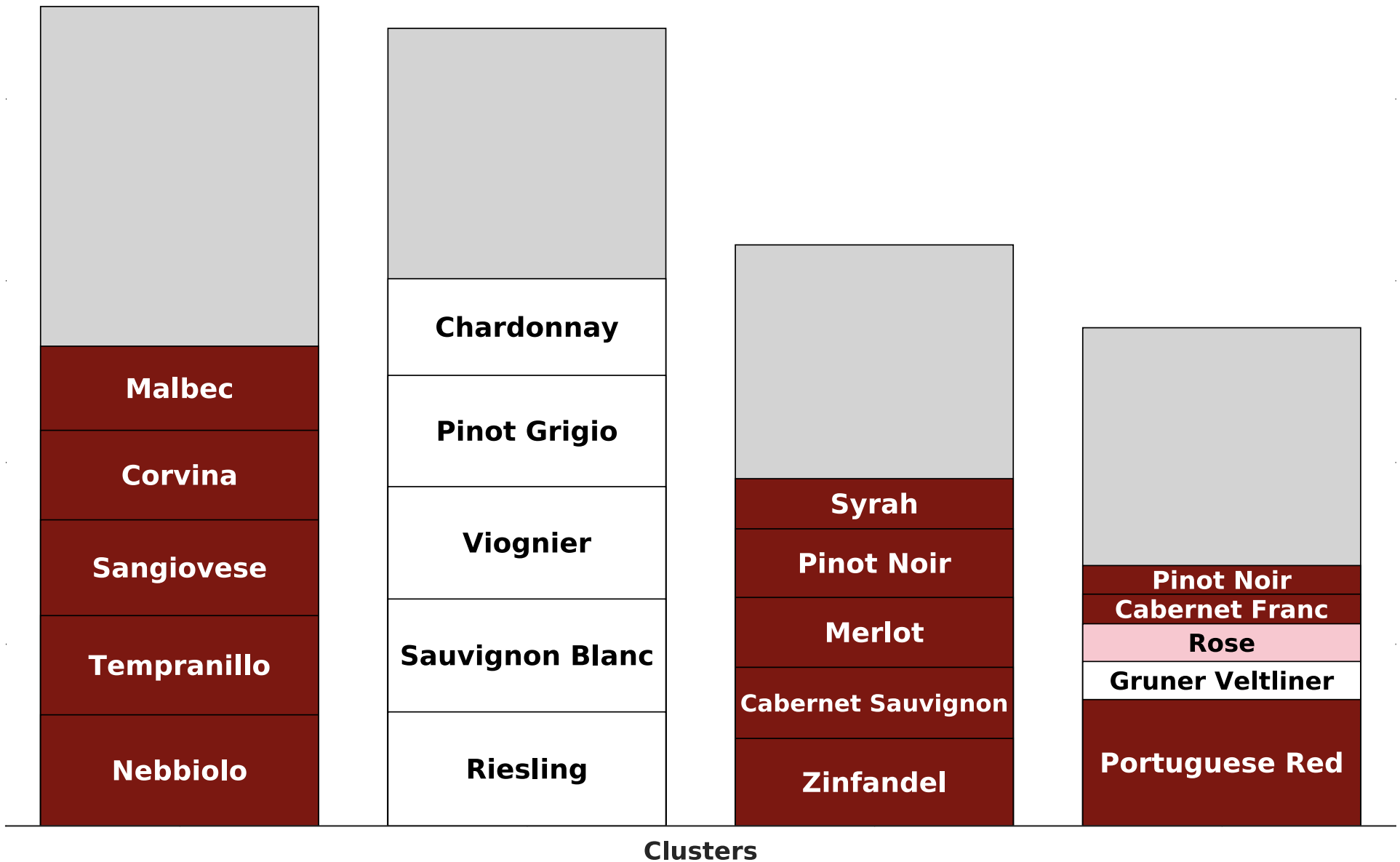


# Varietal Clustering of Reviews



# Varietal Clustering of Reviews

---

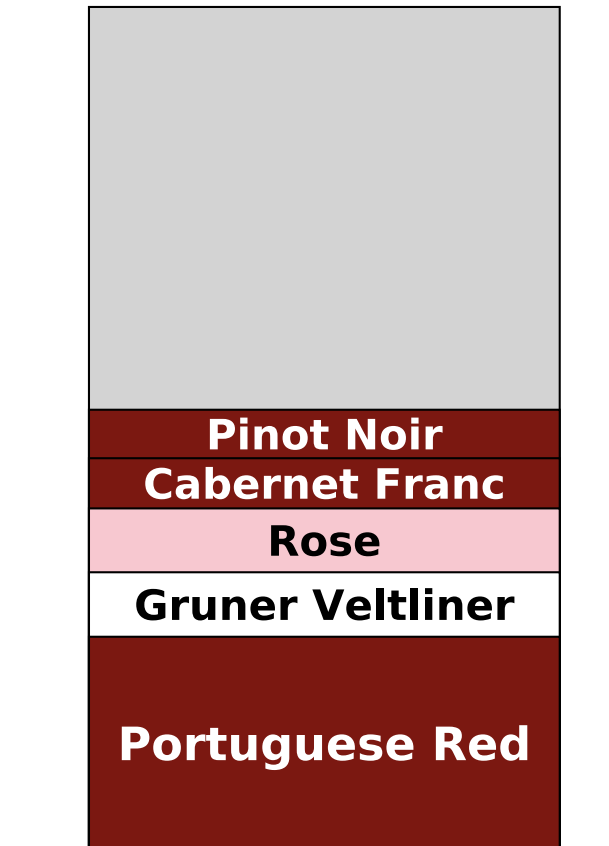




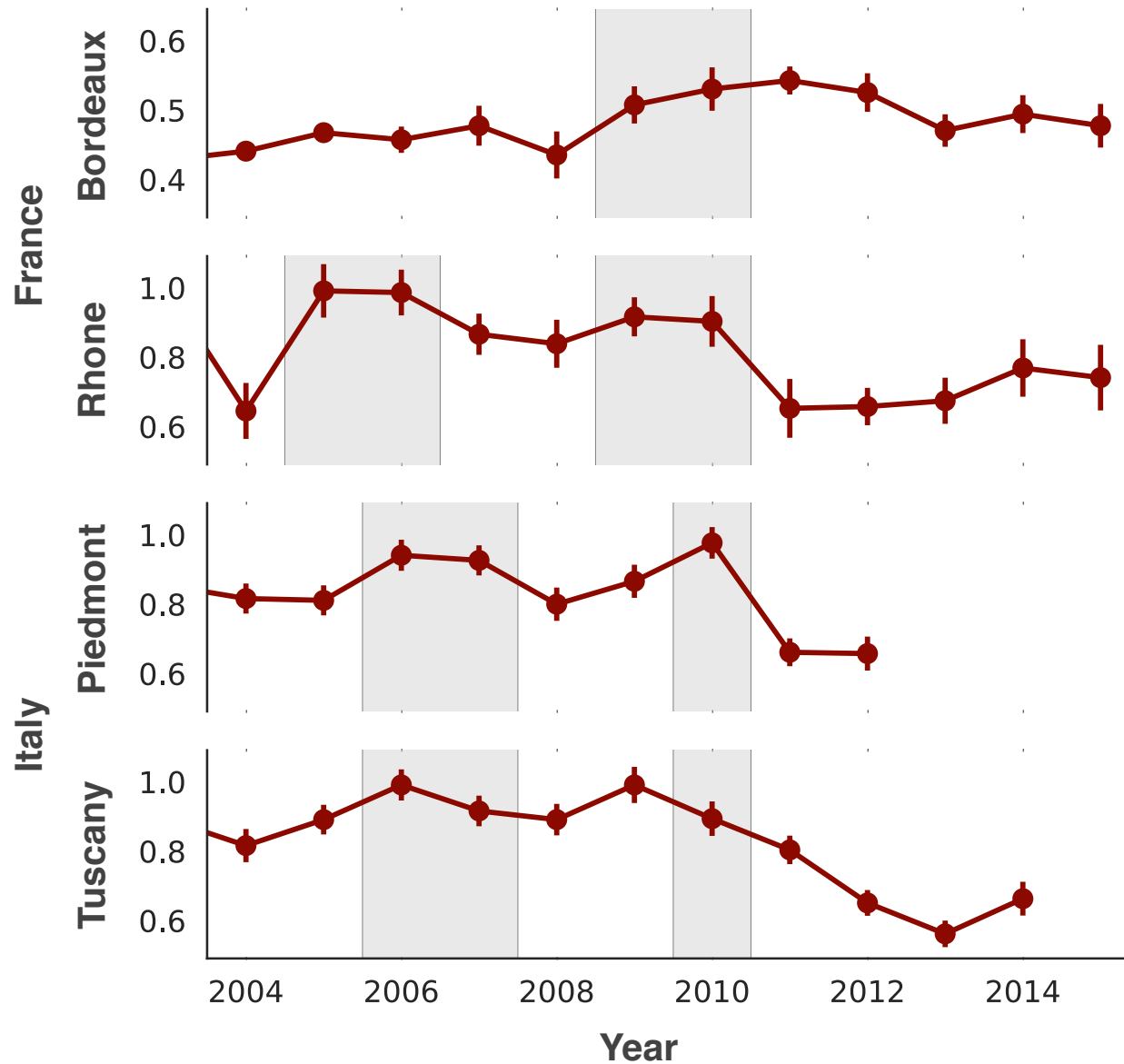
# Stalwart Portuguese Reds

---

- Portuguese reds cluster into single category
- Reviews written by a single editor
- Reviews for wines from editor's other regions do not cluster

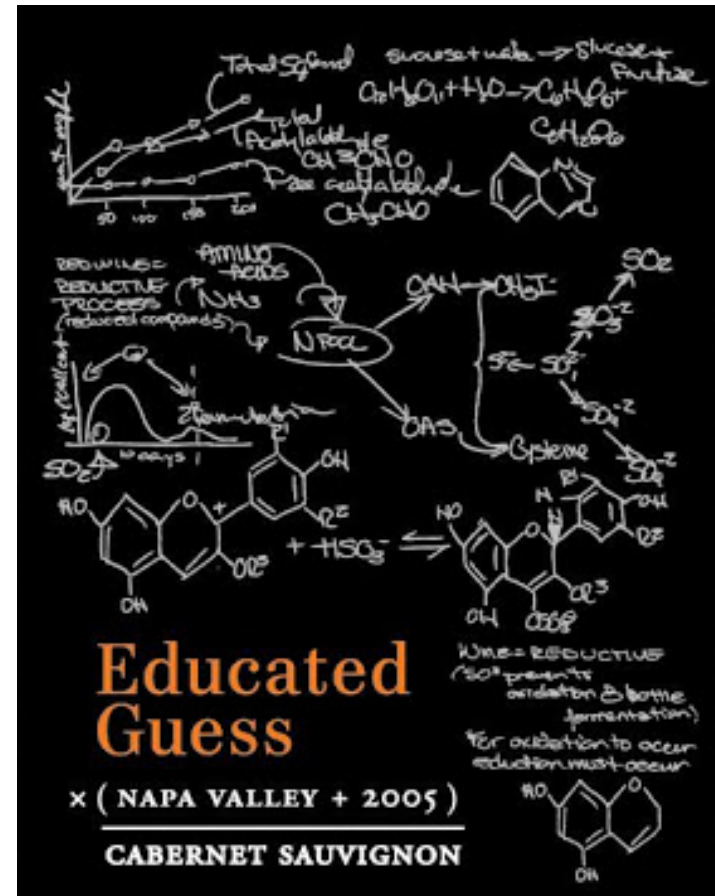


# Sentiment Analysis and Vintage



# Conclusions

- Wine reviews analyzed using NLP and unsupervised learning
- Clustering of reviews can reveal variety
- Sentiment analysis is consistent with vintage years



Thank You!

# Future Work

---

- Clustering by review author and geography
- Effect of weather on sentiment analysis
- Comparison to similar analysis of amateur reviews

# Amateur (Hour) Wine Reviews

---

## RECOMMENDED!

Jeff W. says:

Chardonnay? More like ChardonYAYYY  
I'm drunk at work AND IT FEELS GREAT  
anyone wanna wrestle? NO SARAH  
I WILL NOT STEP INTO YOUR OFFICE

Pairs nicely with:  
Unemployment

## RECOMMENDED!

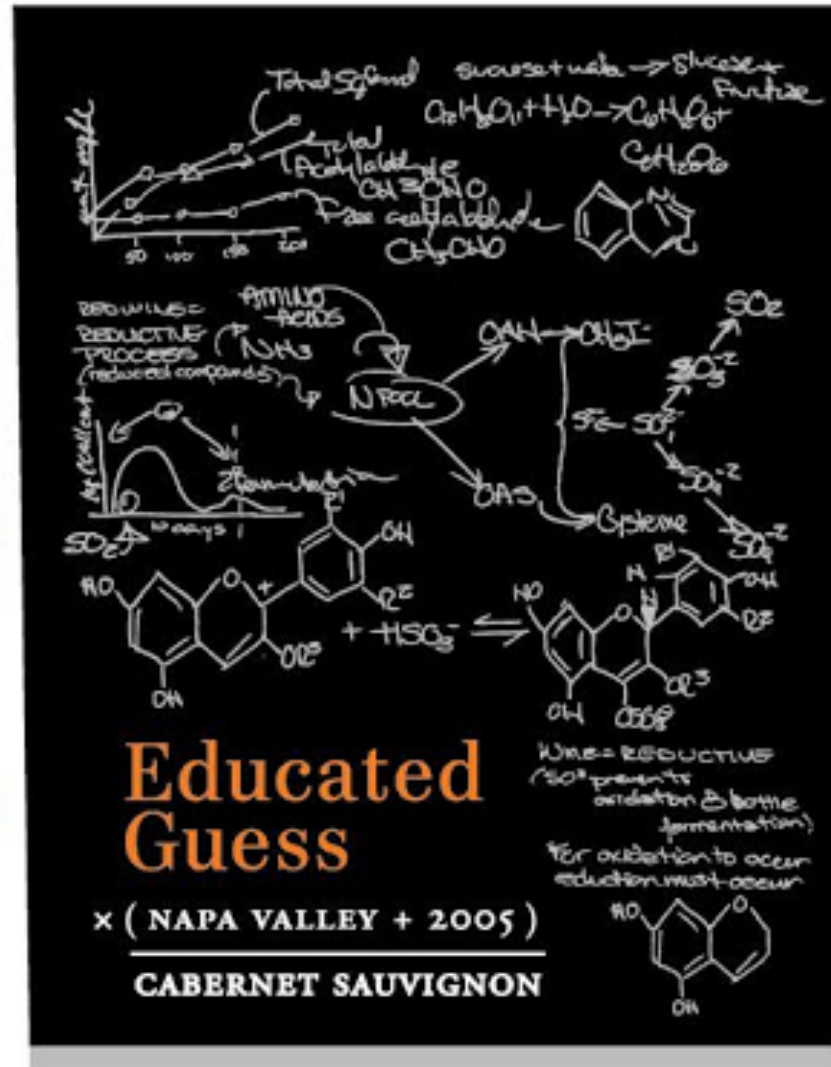
Jeff W. says:

Probably the only good thing to ever  
come out of New Jersey

Pairs nicely with:  
red meat and moving to New York

# Wine for Biochemists

Sulfur dioxide and enology: formation of acetylaldehyde in response to  $\text{SO}_2$  in Fermentation.  
(Cornell University)



General Chemistry:  
Sucrose conversion to  
glucose and fructose,  
with structures.  
(UC Davis)



Sulfide production  
during fermentation.  
(Virginia Tech)



Bleaching of Red Wine  
with excess sulphur  
dioxide, flavylum cation  
of anthocyanins.  
(Cornell University)

$\text{SO}_2$  and wine  
quality (reductive  
process graph) showing  
how  $\text{SO}_2$  kills bacteria.  
 $\text{SO}_2$  inhibits oxidation  
and bottle fermentation,  
and  $\text{SO}_2$  blocks  
polymerization.  
(Cornell University)

