



WINE RECOMMENDATIONS


RACHEL GOODRIDGE




PURPOSE

To create a wine recommendation system for wine enthusiasts of all experience levels that allows users to explore new options and effortlessly make decisions about their next wine purchase.









Data Features

-  Country
-  Description
-  Designation
-  Points
-  Price
-  Province
-  Region
-  Taster Name
-  Variety
-  Winery

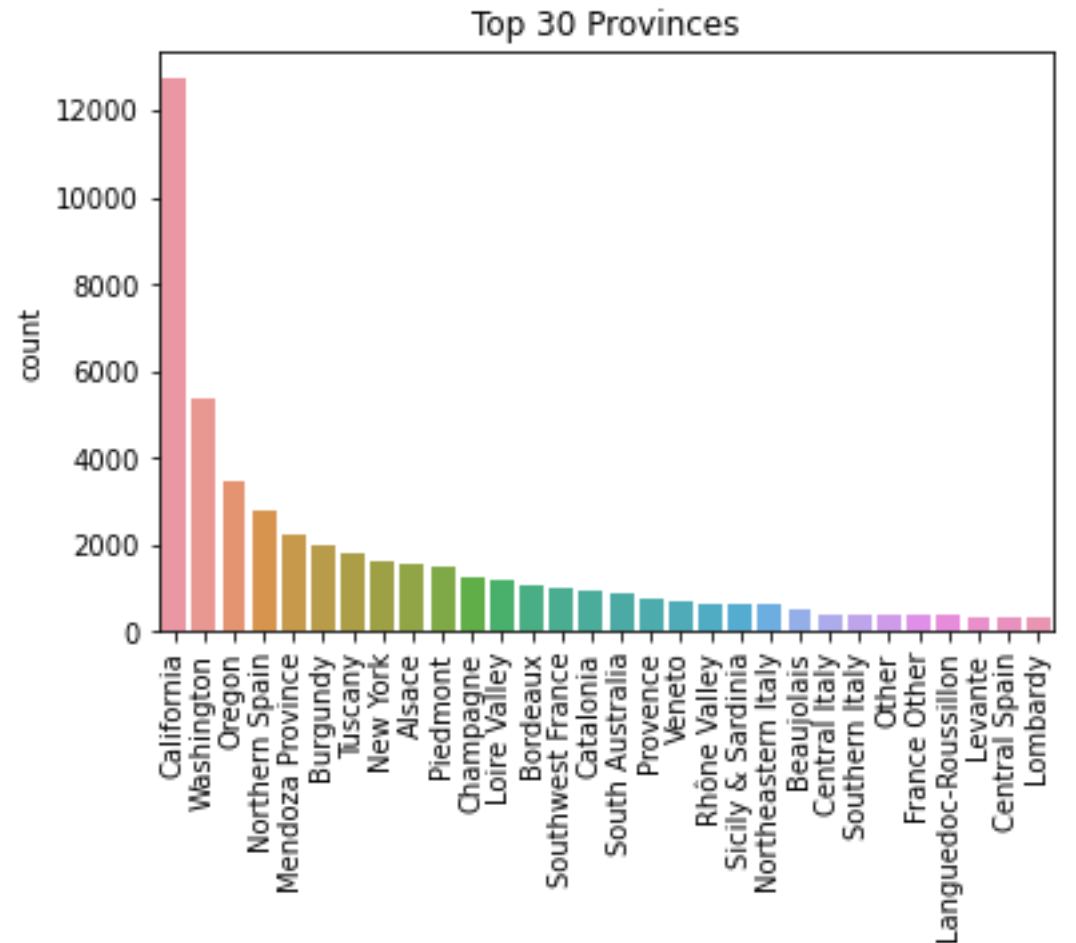
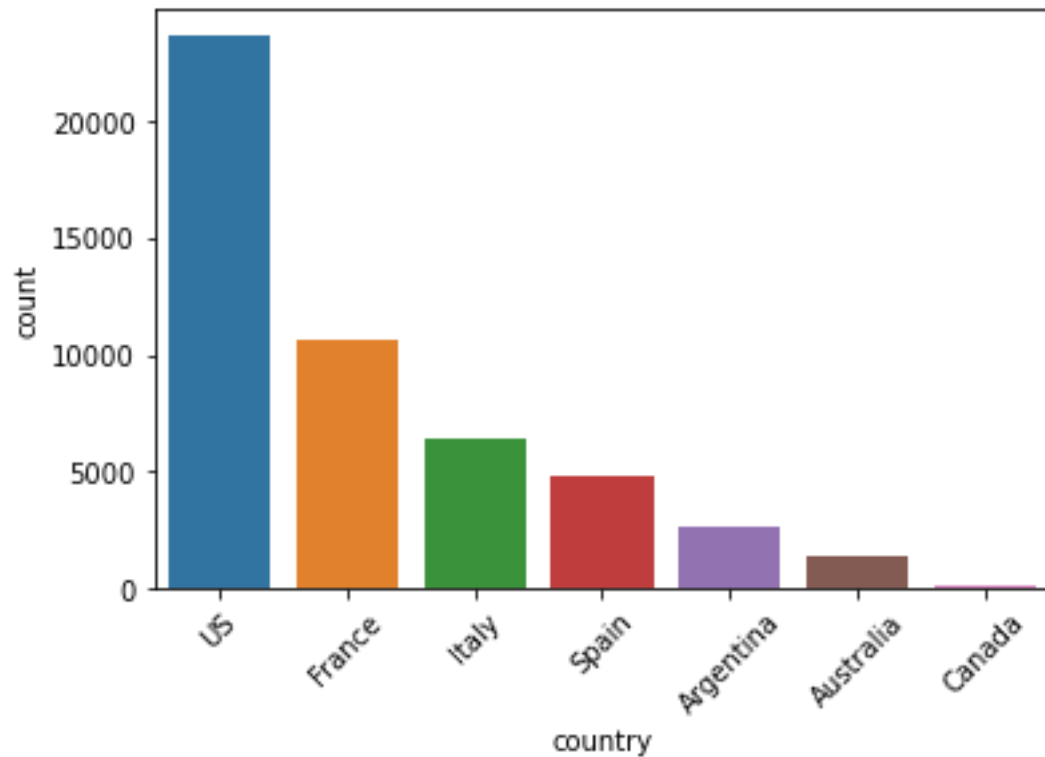
Data Features

-  Country
-  Description
-  Designation
-  Points
-  Price
-  Province
-  Region
-  Taster Name
-  Variety
-  Winery

Tools

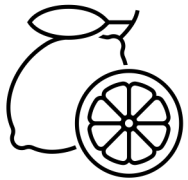
-  Data manipulation - pandas & numpy
-  Visualizations - matplotlib & seaborn
-  Text Preprocessing - re, string, & nltk
-  Dimensionality Reduction -
CountVectorizer from sklearn
-  Topic Modeling - corextopic
-  Recommender Systems
 -  Content-based filtering
 -  Collaborative filtering with LightFM

WINE LOCATION DISTRIBUTION



TOPIC MODELING WITH COREX

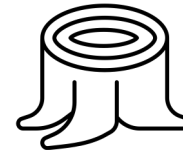
Citrus



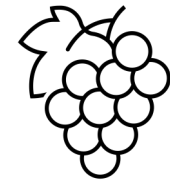
Berry



Woody



Vinification



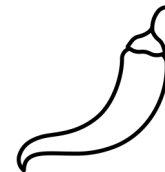
Varietal



Body



Spicy



Miscellaneous



CONTENT-BASED FILTERING

Wine the User Likes

Description:

Blackberry and raspberry **aromas** show a typical Navarran whiff of green **herbs** and, in this case, horseradish. In the mouth, this is fairly **full bodied**, with tomatoey acidity. Spicy, **herbal flavors** complement **dark plum** fruit, while the finish is **fresh** but grabby.

CONTENT-BASED FILTERING

Wine the User Likes

Description:




Blackberry and raspberry **aromas** show a typical Navarran whiff of green **herbs** and, in this case, horseradish. In the mouth, this is fairly **full bodied**, with tomatoey acidity. Spicy, **herbal flavors** complement **dark plum** fruit, while the finish is **fresh** but grabby.

Wine Recommendation

Description:

Concentrated **blackberry** and black currant **aromas** are ripe and come with a balsamic note. It feels juicy, **full**, lively and **fresh**, with oaky **dark-plum** flavors accented by rooty, **herbal notes**. Dry tannins mark the finish.

COLLABORATIVE FILTERING

-  Unsuccessful due to the nature of the data
-  No overlap between tasters
-  Average prediction error was very high



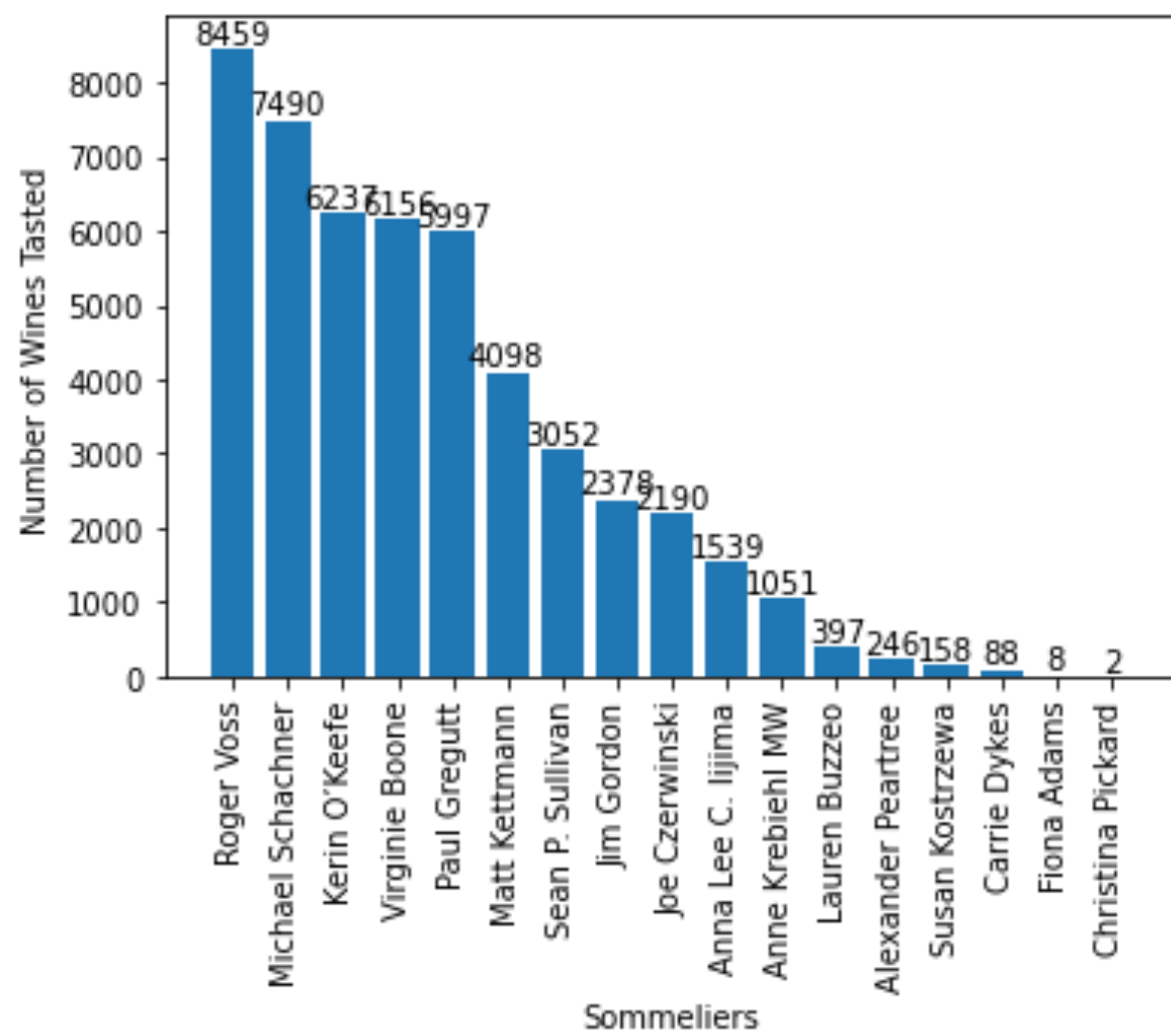
CONCLUSION

Using CorEx for topic modeling and a content-based filtering system, I created a functional wine recommendation system that can cater to everyone.



THANK YOU!

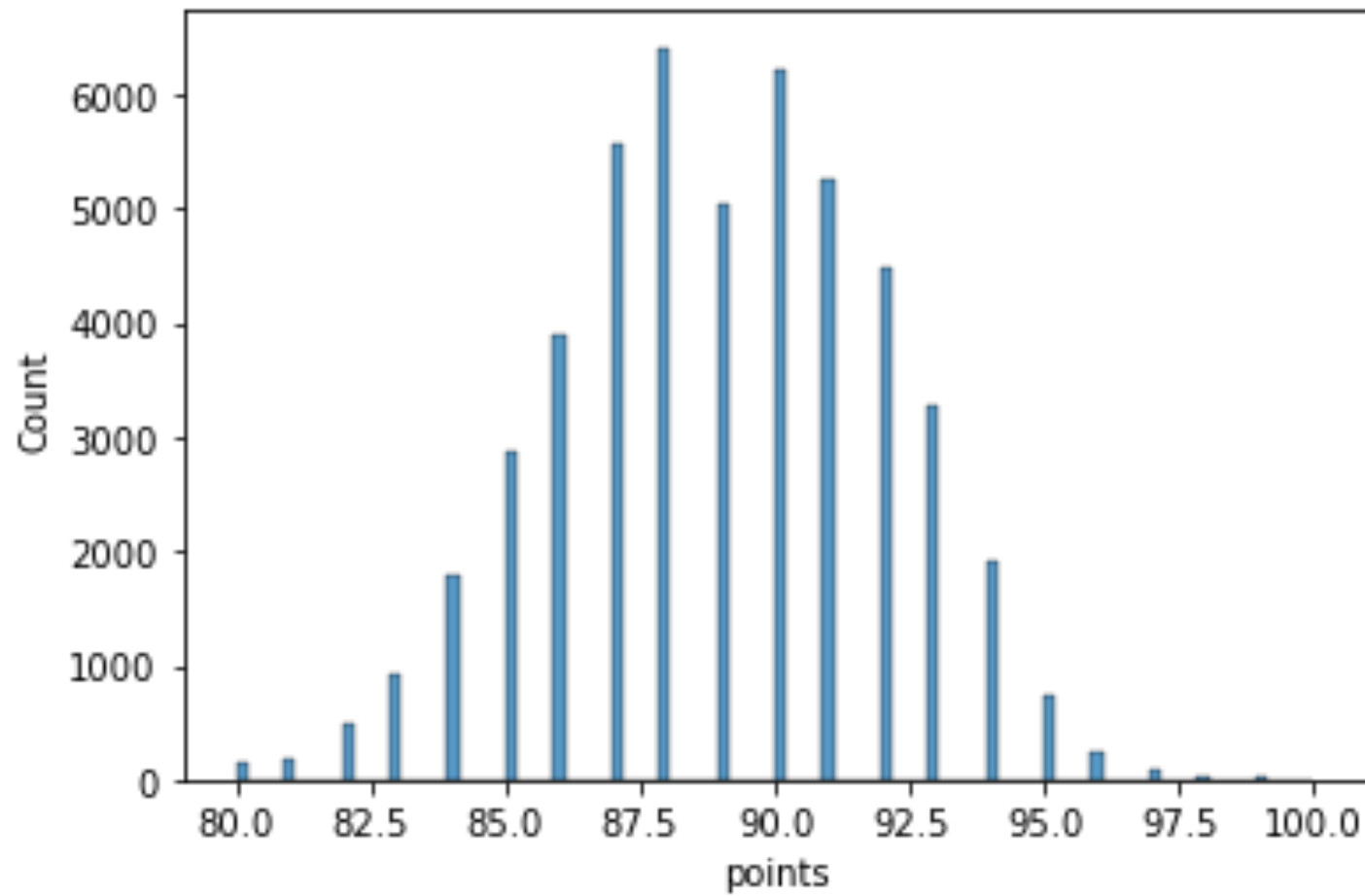
APPENDIX



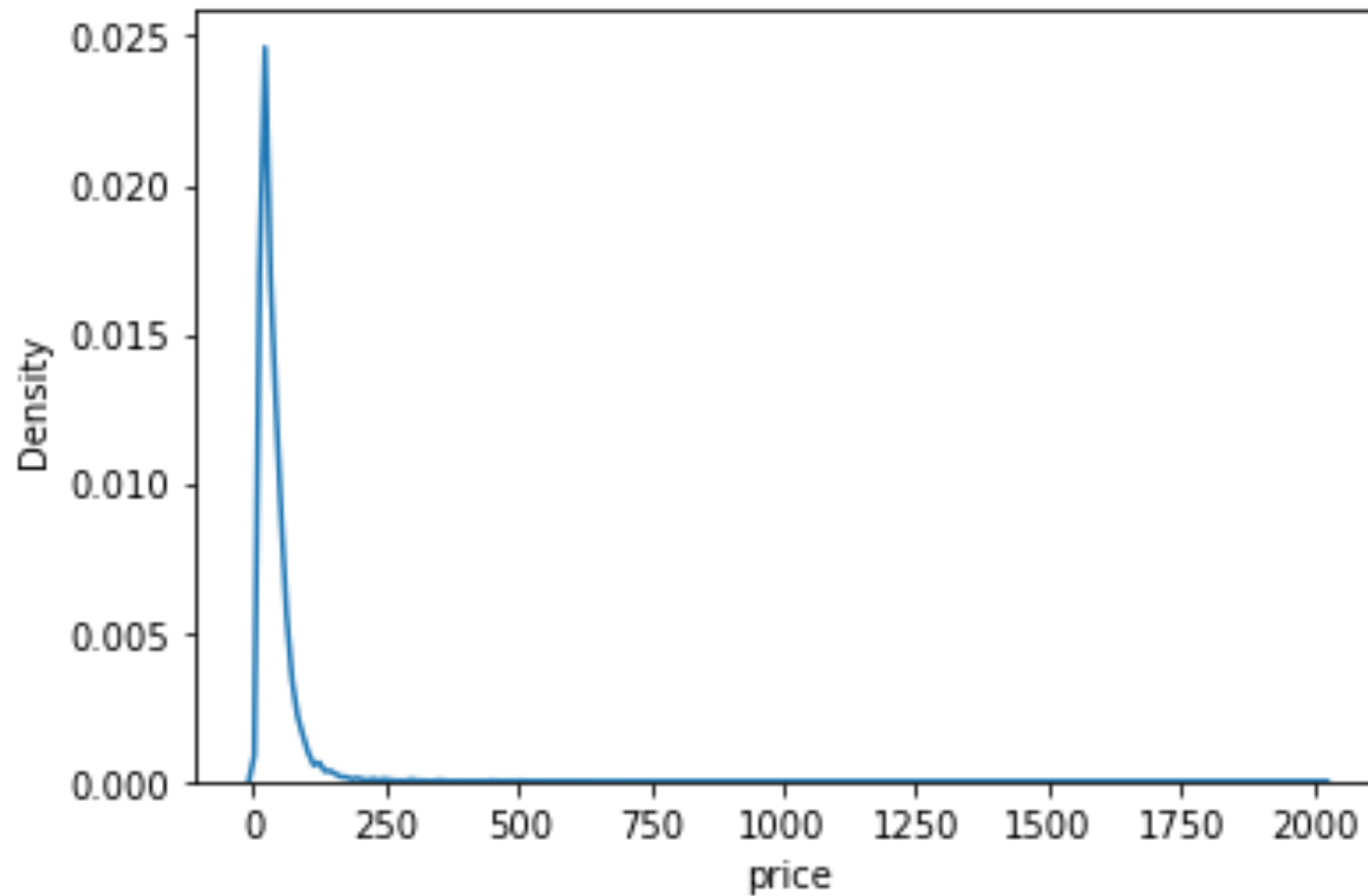
SOMMELIERS

WINE COUNTS

	wines	counts	tasters	taster_counts
40078	Segura Viudas NV Aria Estate Extra Dry Sparkli...	7	Michael Schachner	1
40083	Segura Viudas NV Extra Dry Sparkling (Cava)	7	Michael Schachner	1
39083	Ruinart NV Brut Rosé (Champagne)	6	Joe Czerwinski, Roger Voss	2
2956	Bailly-Lapierre NV Brut (Crémant de Bourgogne)	6	Roger Voss	1
20805	Freixenet NV Cordon Negro Extra Dry Sparkling ...	5	Michael Schachner	1
40086	Segura Viudas NV Reserva Heredad Sparkling (Cava)	5	Michael Schachner	1
14796	Domaine Charles Baur NV Brut Sparkling (Créman...	4	Anne Krebiehl MW, Roger Voss	2
23047	Henri Abele NV Brut (Champagne)	4	Paul Gregutt, Roger Voss	2
24568	Jacquart NV Brut Mosaïque (Champagne)	4	Paul Gregutt, Roger Voss	2
36243	Pierre Sparr NV Brut Réserve Sparkling (Créman...	4	Anne Krebiehl MW, Joe Czerwinski	2

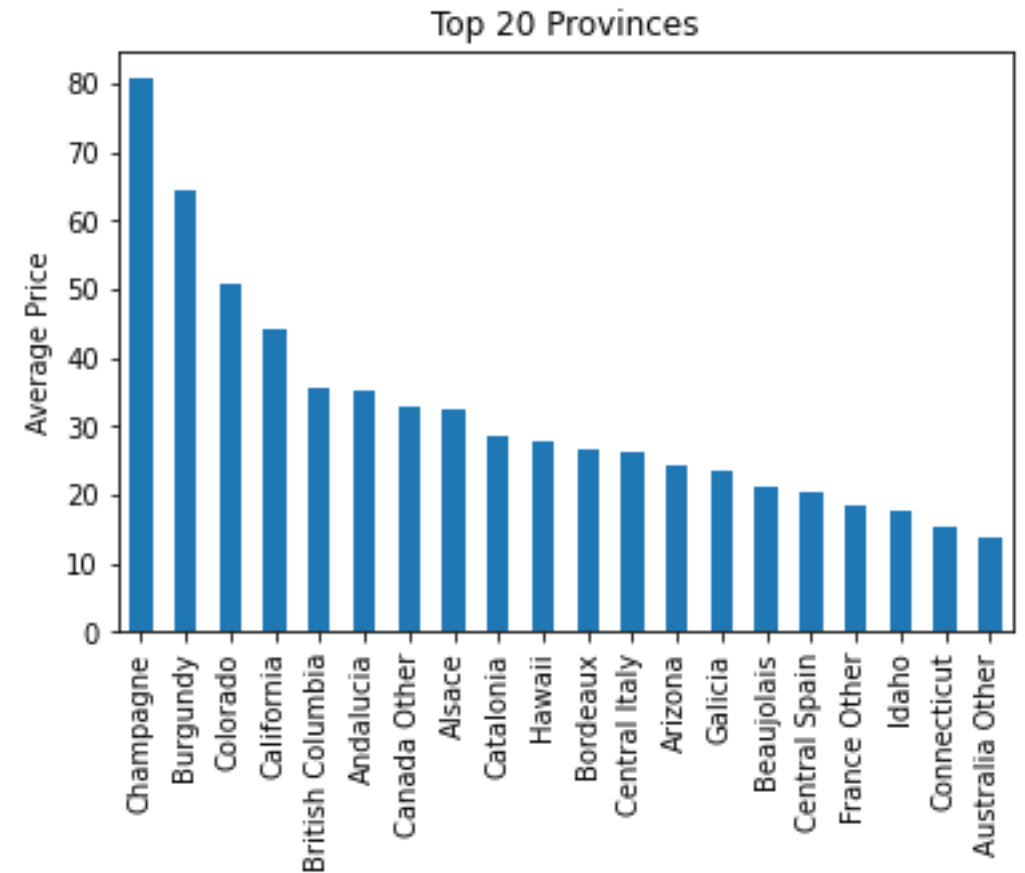
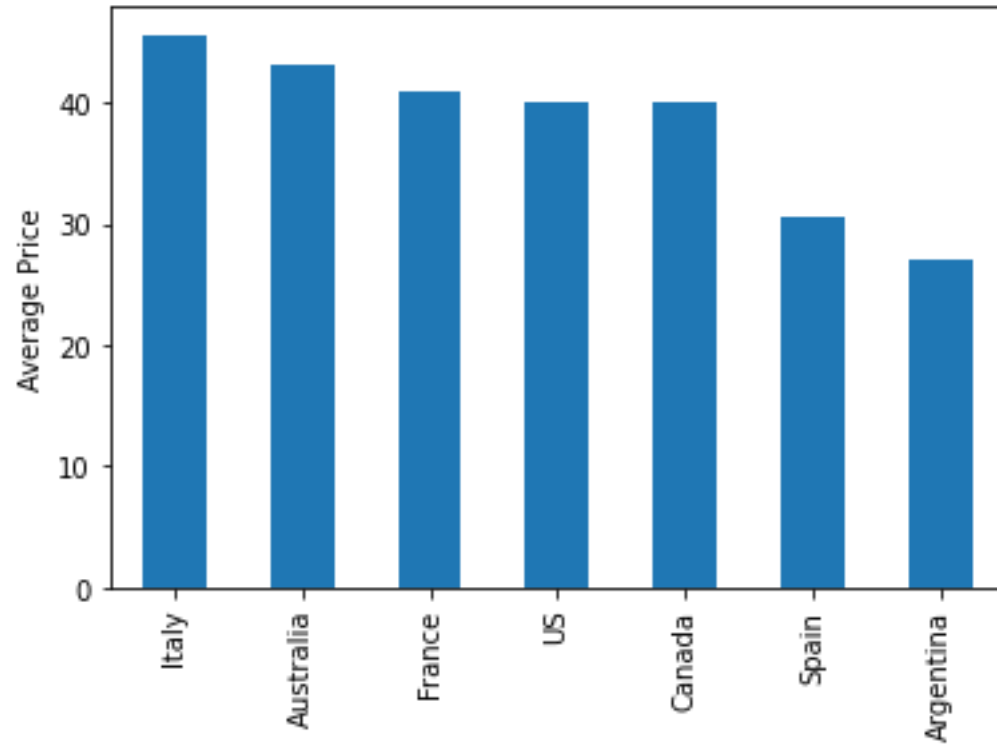


POINTS DISTRIBUTION



PRICE DISTRIBUTION

AVERAGE PRICE BY LOCATION





WINE AROMA WHEEL

WORDS IN TOPICS

- Citrus Topic - white, citrus, lemon, appl, tannin, lime, peach, grapefruit, zest, plum
- Berry Topic - cherri, berri, blackberri, black cherri, pear, red berri, berri fruit, red cherri, cherri fruit, honey
- Woody Topic - oak, vanilla, toast, french oak, french, toast oak, barrel, toast oak, coconut, age French
- Vinification Topic - vineyard, bottl, singl vineyard, nose bottl, singl, bottl age, nose, sourc, site, vineyard design
- Varietal Topic - blend, cabernet, sauvignon, pinot, cabernet sauvignon, merlot, cabernet franc, franc, pinot noir, petit
- Body Topic - textur, structur, age, rich, ripe, fruiti, wood age, readi, aftertast, charact
- Spicy Topic - pepper, black, black pepper, white pepper, licor, underbrush, clove, crush, offer, deliv
- Misc Topic - bodi, medium, medium bodi, linger, long, complex, color, now, lush, sen

TOPIC DISTRIBUTION

- Percent Citrus Topic Documents : 0.2790336253178864
- Percent Berry Topic Documents : 0.4108101562184637
- Percent Woody Topic Documents : 0.23578089048560932
- Percent Vinification Topic Documents : 0.19930973236991886
- Percent Varietal Topic Documents : 0.21137932426432002
- Percent Body Topic Documents : 0.273402494651435
- Percent Spicy Topic Documents : 0.17452468413191782
- Percent Misc Topic Documents : 0.3580309207605054