



What Makes a Good Wine?



Data Set

User zackthoutt on Kaggle collected this data by scraping data from WineEnthusiast.com.

Description of the columns of the data set:

- `X1`: Number corresponding to the observation. Listed in ascending order from 0
- `country`: The country that the wine is from
- `description`: A few sentences from a sommelier describing the wine's taste, smell, look, feel, etc.
- `designation`: The vineyard within the winery where the grapes that made the wine are from
- `points`: The number of points WineEnthusiast rated the wine on a scale of 1-100 (though they say they only post reviews for wines that score greater than or equal to 80)
- `price`: The cost for a bottle of the wine
- `province`: The province or state that the wine is from
- `region_1`: The wine growing area in a province or state (ie Napa)
- `region_2`: Sometimes there are more specific regions specified within a wine growing area (ie Rutherford inside the Napa Valley), but this value can sometimes be blank
- `variety`: The type of grapes used to make the wine (ie Pinot Noir)
- `winery`: The winery that made the wine

Data Set

Glimpse of data

```
## Rows: 150,930
## Columns: 11
## $ X1          <dbl> 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 1...
## $ country     <chr> "US", "Spain", "US", "US", "France", "Spain", "Spain", ...
## $ description <chr> "This tremendous 100% varietal wine hails from Oakville...
## $ designation <chr> "Martha's Vineyard", "Carodorum Selección Especial Rese...
## $ points      <dbl> 96, 96, 96, 96, 95, 95, 95, 95, 95, 95, 95, 95, 95,...
## $ price       <dbl> 235, 110, 90, 65, 66, 73, 65, 110, 65, 60, 80, 48, 48, ...
## $ province    <chr> "California", "Northern Spain", "California", "Oregon",...
## $ region_1    <chr> "Napa Valley", "Toro", "Knights Valley", "Willamette Va...
## $ region_2    <chr> "Napa", NA, "Sonoma", "Willamette Valley", NA, NA, NA, ...
## $ variety     <chr> "Cabernet Sauvignon", "Tinta de Toro", "Sauvignon Blanc...
## $ winery      <chr> "Heitz", "Bodega Carmen Rodríguez", "Macauley", "Ponzi"...
```

Research Question

Based on our interests, we selected this data set to explore trends that will help to make informed decisions when buying wine. Therefore, our research question is:

How does one select the best wine?

Top 5s

Countries with highest rated wines on average

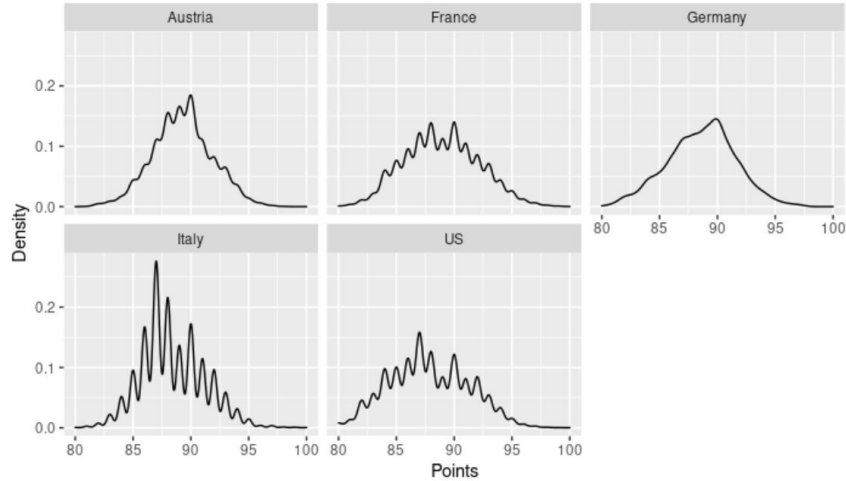
1. Austria
2. Germany
3. France
4. Italy
5. Canada

Varieties with highest rated wines on average

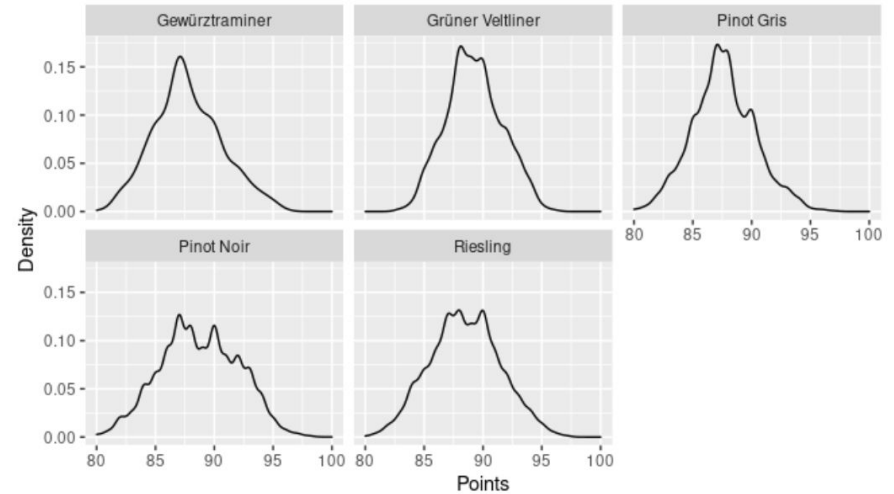
1. Gewürztraminer
2. Grüner Veltliner
3. Pinot Gris
4. Pinot Noir
5. Riesling

Visualizations

Austria has the highest density of wines that scored 90 or above

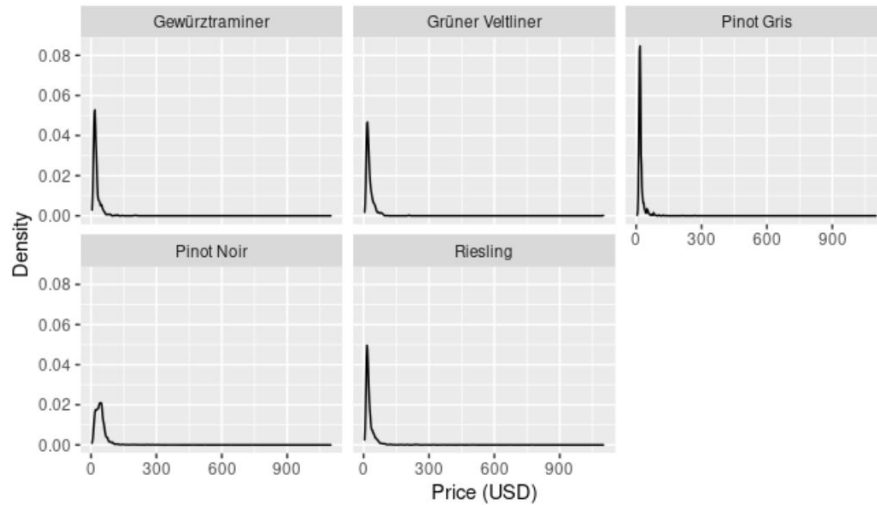


Grüner Veltliner has a concentrated density between 85 and 90 points

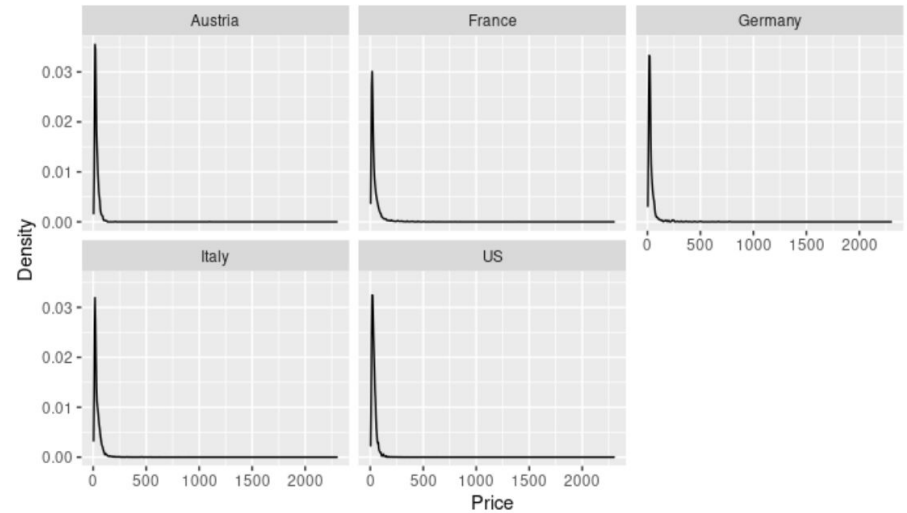


Visualizations

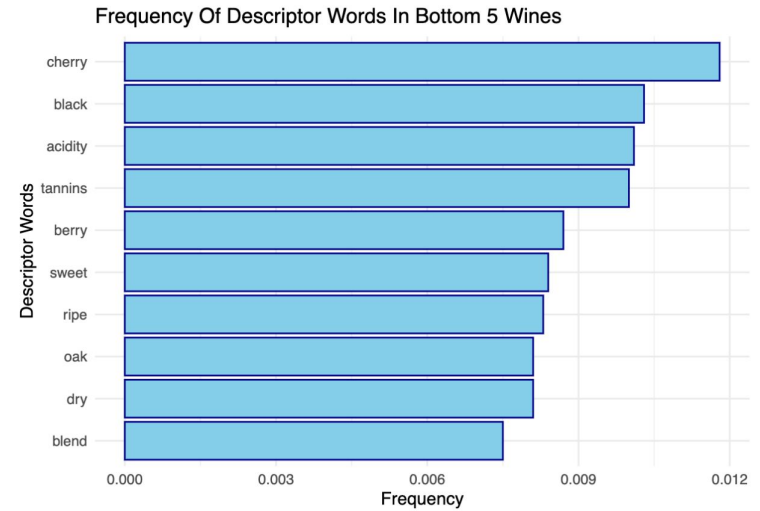
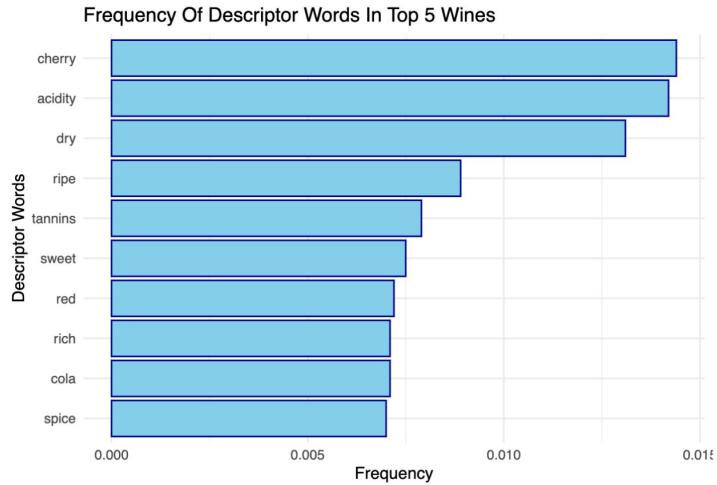
The majority of wines in all top varieties are cheap



The majority of wines in all top countries are cheap

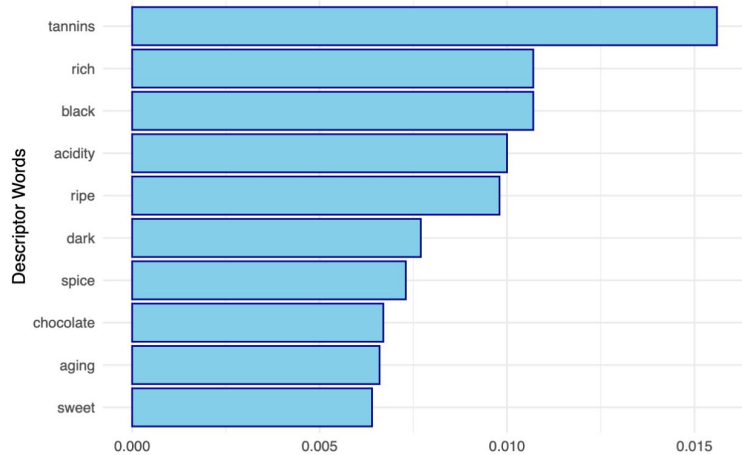


Visualizations

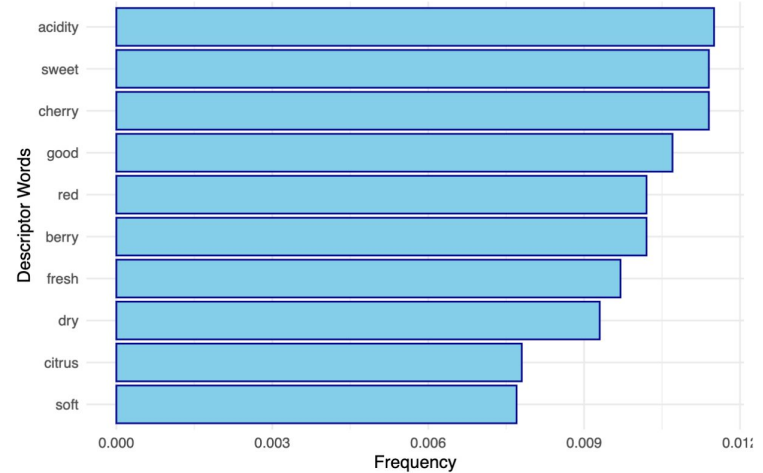


Visualizations

Frequency Of Descriptor Words In Expensive Wines

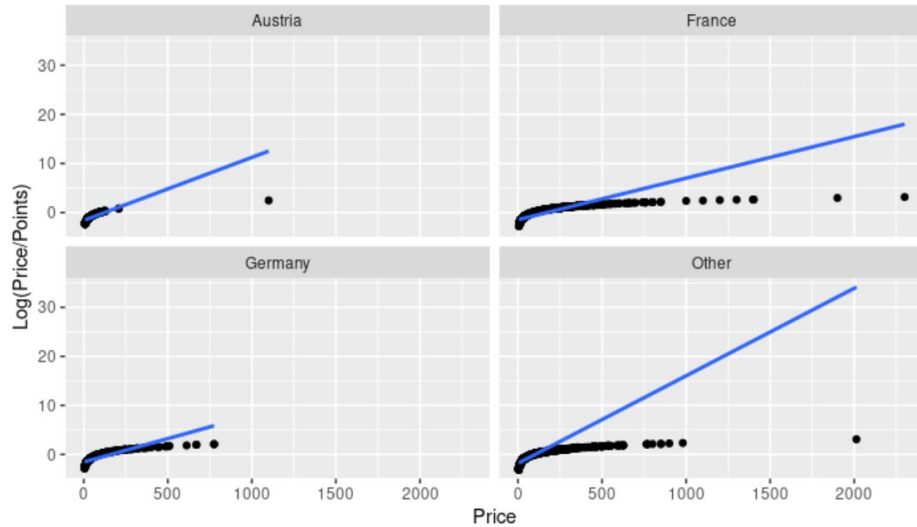


Frequency Of Descriptor Words In Cheap Wines

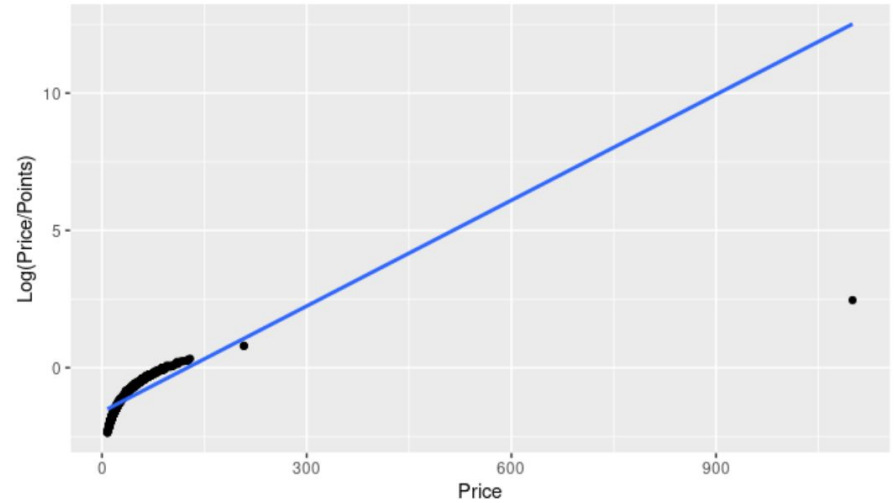


Visualizations

Positive Relationship Between Weighted Points Value and Price



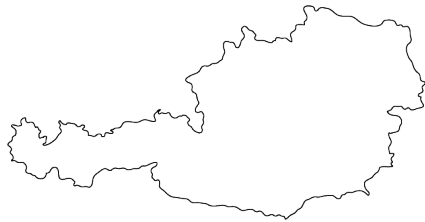
Austria Has Steepest Slope for Weighted Points and Price Model



Conclusions

Top Countries and Varieties

Wines from Austria had not only the highest average wine score, but also the highest density of wines that scored a 90 or above.



The Grüner Veltliner variety primarily has wines with scores of at least 90.

The Gewürztraminer wine variety had a more evenly distributed range of wine scores and prices, but a higher wine score on average.



Conclusions Points and Price

To maximize points...

Pick a wine from a top 5 country



Pick a wine from a top 5 variety



Price and points have a positive correlation.

Conclusions Text Analysis

Highly rated wines are likely to be dry reds with hints of cherry



Lowly rated wines are likely to be fruity whites or fruity dark reds (excluding nodes of cherry)



References

“Adorn Cabernet Sauvignon – at Drizly.com.” *Drizly*,

drizly.com/wine/red-wine/cabernet-sauvignon/adorn-cabernet-sauvignon/p82352.

“Blossom Hill White Wine – Blossom Hill.” *Home – Blossom Hill*, www.blossomhill.com/our-wines/white-wine.

“Germany Outline Map: Germany Tattoo, Germany Map, Map Art.” *Pinterest*, www.pinterest.com/pin/303641199848244190/.

“Gewürztraminer Spl. Feinherb – Weingut Diehl.” *Blees*, 29 Oct. 2020,

weingut-diehl-blees.de/produkt/gewuerztraminer-spaetlese-feinherb/.

“Hugl Gruner Veltliner.” *Hugl Gruner Veltliner | Total Wine & More*,

www.totalwine.com/wine/white-wine/gruner-veltliner/hugl-gruner-veltliner/p/101175010.

“J Vineyards Pinot Gris – at Drizly.com.” *Drizly*, drizly.com/wine/white-wine/pinot-grigio/j-vineyard-pinot-gris/p5106.

References

“Low Hanging Fruit Sweet Red.” *Low Hanging Fruit Sweet Red | Total Wine & More*,

www.totalwine.com/wine/red-wine/red-blend/low-hanging-fruit-sweet-red/p/111734750.

Moen, John. “United States Outline Map.” *WorldAtlas*, WorldAtlas,

www.worldatlas.com/webimage/countrys/namerica/usstates/usashape.htm.

“Outline Map Of Italy Printable ~ Free Printable Coloring Page for Kids: Italy Map, Kids Travel Journal, Italy for Kids.” *Pinterest*,

www.pinterest.com/pin/283304632791078951/.

“Outline Map of Austria: Free Vector Maps: Austria Map, Map Vector, Vector Free.” *Pinterest*,

www.pinterest.com/pin/618330223817176966/.

“Pin on Printable Patterns at PatternUniverse.com.” *Pinterest*, www.pinterest.com/pin/325455510548716537/.

“Semi-Dry Riesling.” *Montezuma Winery*, 30 Oct. 2020, www.montezumawinery.com/product/semi-dry-riesling/.

“Woodbridge Pinot Noir by Robert Mondavi – at Drizly.com.” *Drizly*,

drizly.com/wine/red-wine/pinot-noir/woodbridge-pinot-noir-by-robert-mondavi/p2162.