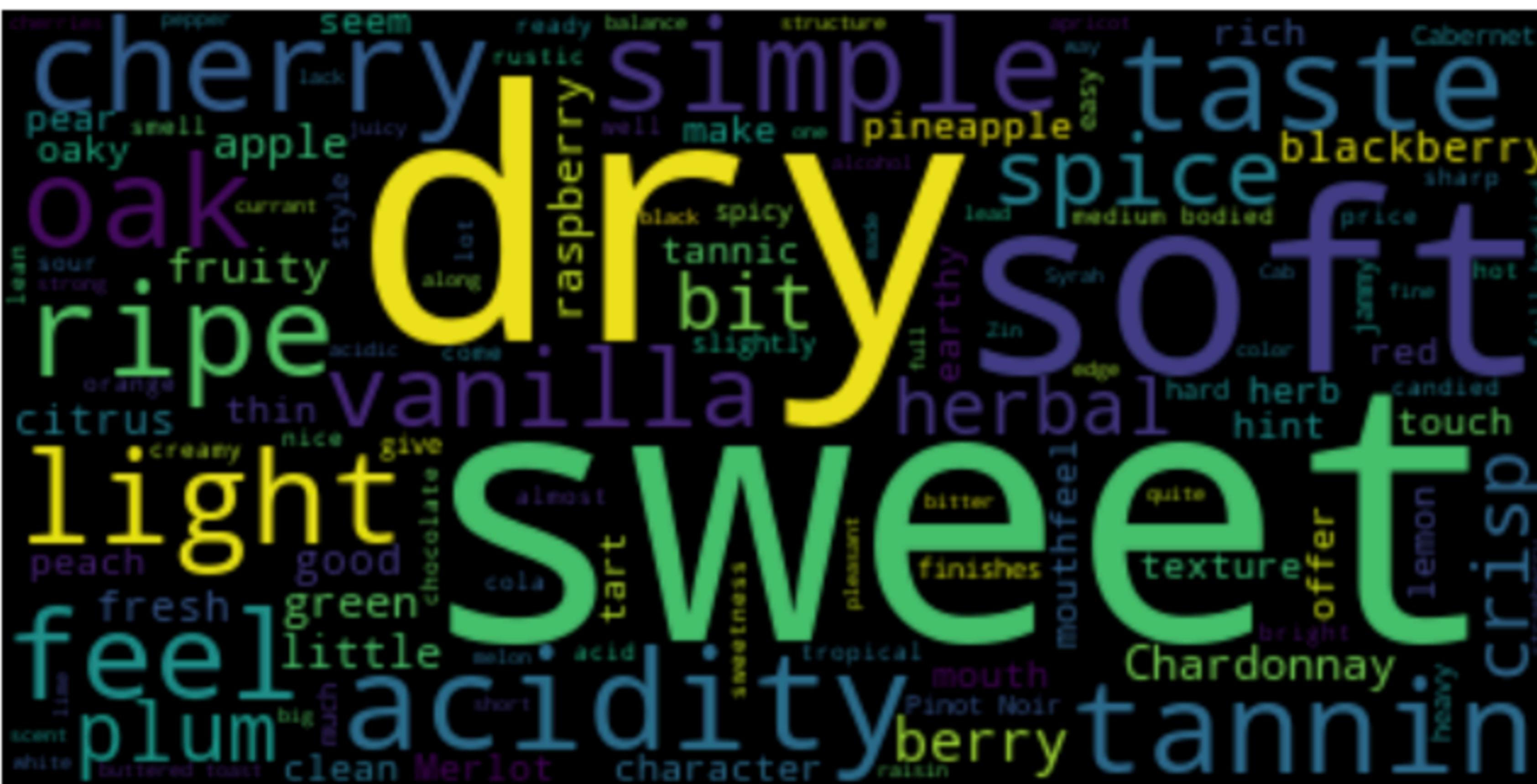


Viz 1: All wine reviews were aggregated into bins based off of their respective wine scores.
This WordCloud visualizes the frequency of descriptor words in reviews.

Most Used Descriptors used per Score Range

Scores: 80-84



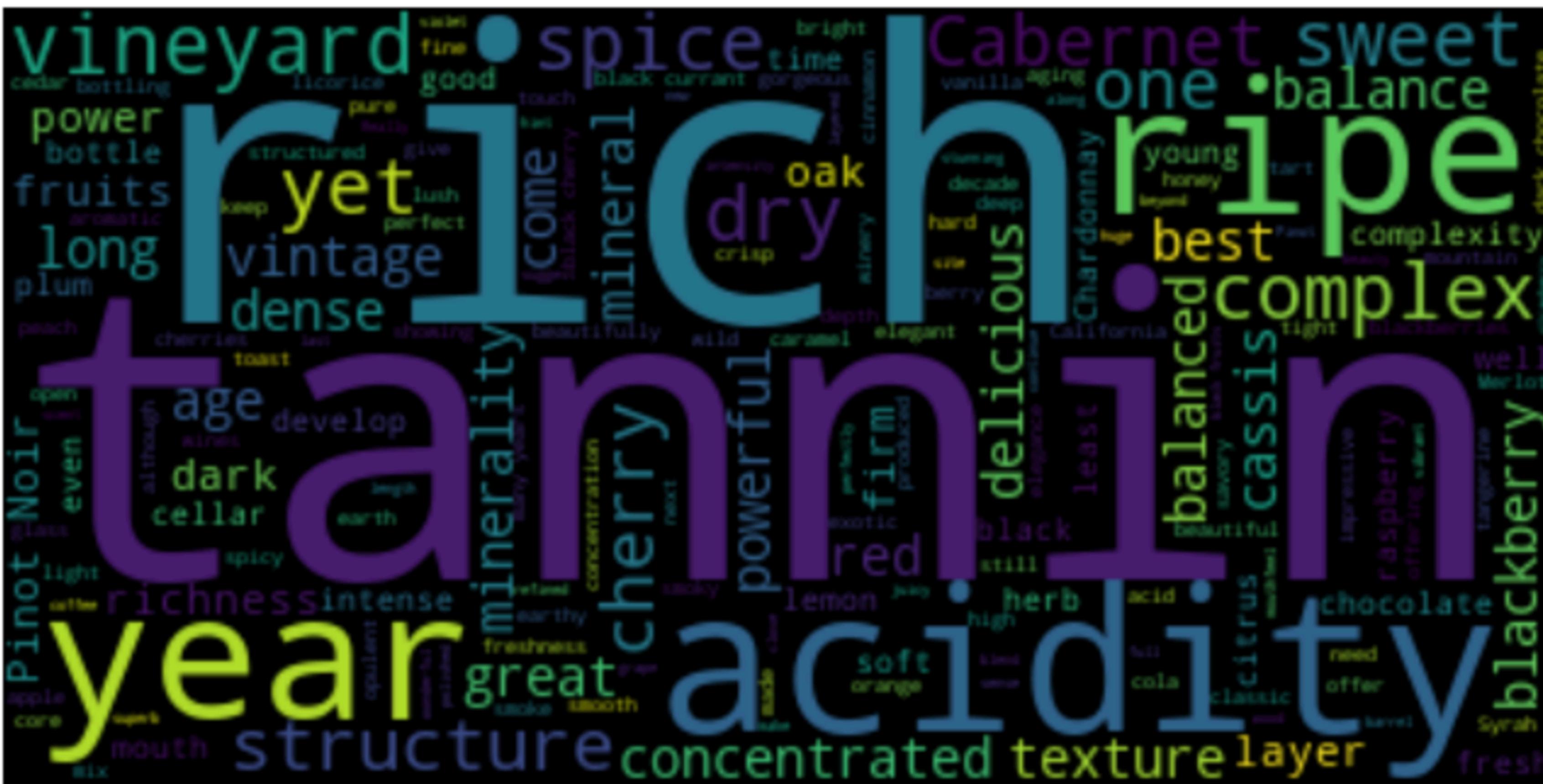
Scores: 85-89



Scores:90-94



Scores: 95-100



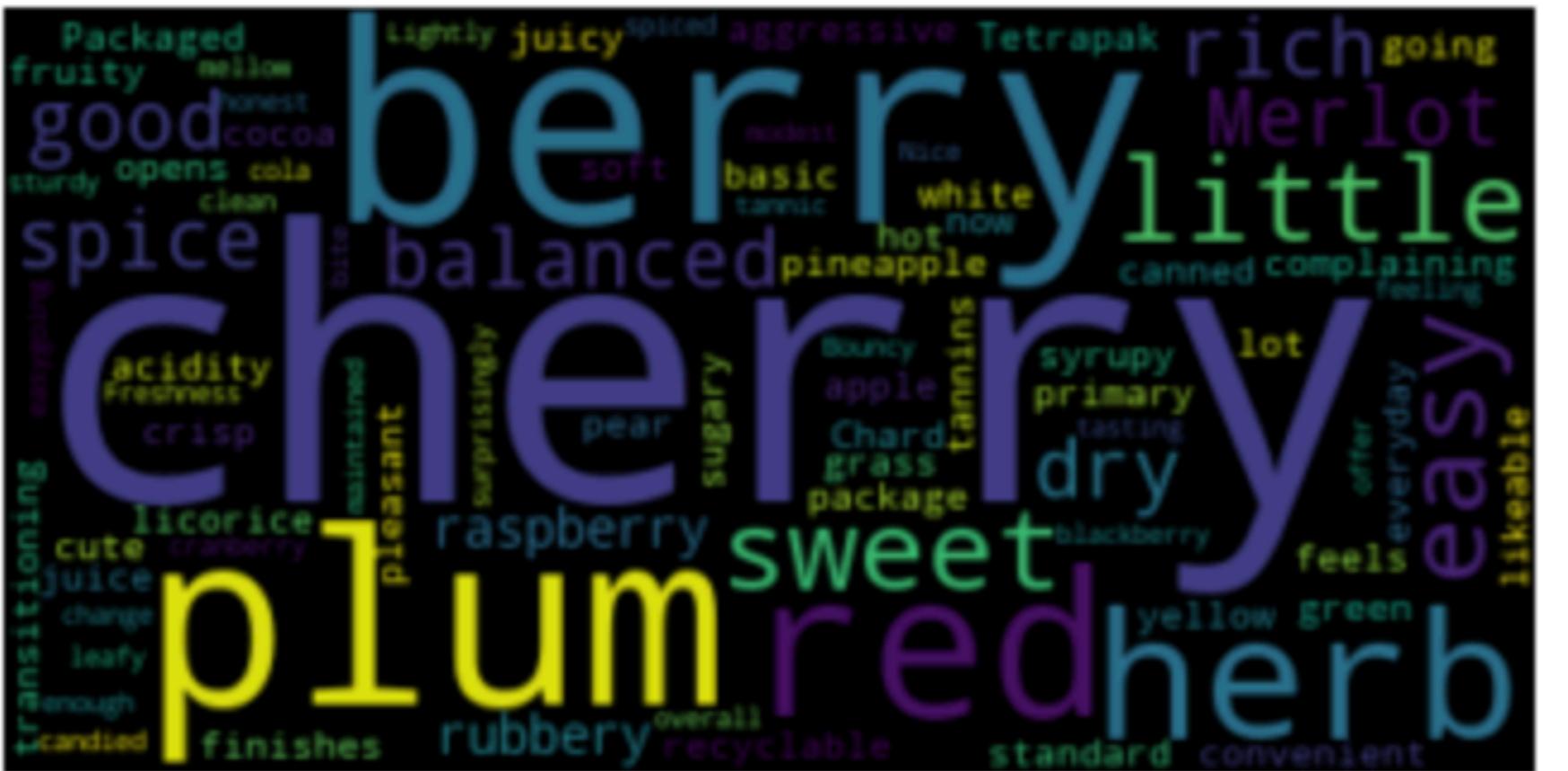
Insights

- The descriptors used for wines in the highest score range (95-100) emphasize 'tannin', 'rich', 'acidity' and 'ripe' and 'now'. These are descriptors that usually correspond with age and complexity, suggesting that older wines may net better scores. Future research is needed to test this hypothesis.
- In the lowest score range (80-84), the descriptors 'soft', 'sweet', 'acidity' and 'dry' are dominant. The descriptors in this range also reflect less dimensionality than higher-scored wines. For example, 'cherry', 'spice', 'oak', 'sweet', 'herbal', 'vanilla',etc. are descriptors that most likely were very obvious to judges. In contrast, higher scoring wines show much less emphasis on single-word taste descriptors -- in favor of descriptors that refer to the structure and balance of the wine.
- In the highest score range (95-100), terms that indicate a higher level of detail and uniqueness such as 'vineyard' and 'vintage' have higher frequencies in wine review descriptions than at lower score ranges.
- Balance, structure, and mouthfeel are emphasized in higher-scoring wines (90-100)
- Emphasis on fruit-forward descriptors and de-emphasis on mouthfeel and structure descriptors in lower scoring wines(80-89)

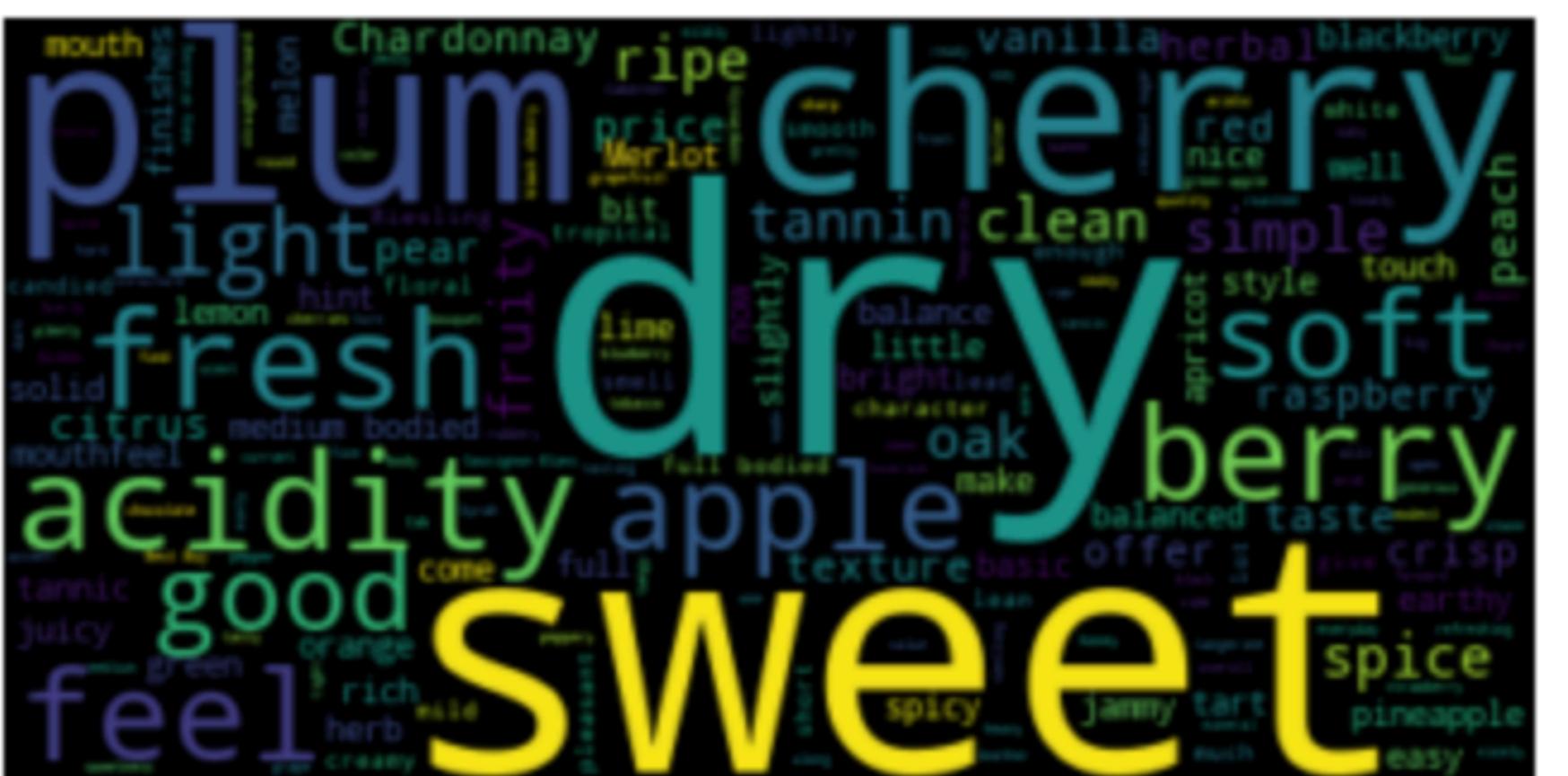
Viz 2: All wine reviews were aggregated into bins based off of their respective wine price segment. This WordCloud visualizes the frequency of descriptor words in reviews.

Most Used Descriptors used per Price Segment

Extreme Value, <\$4



Value, \$4-10



Popular Premium, \$10-15



Premium, \$15-20



Super Premium, \$20-30



Ultra Premium, \$30-50



Luxury, \$50-100



Super Luxury, \$100-200



Icon, >\$200

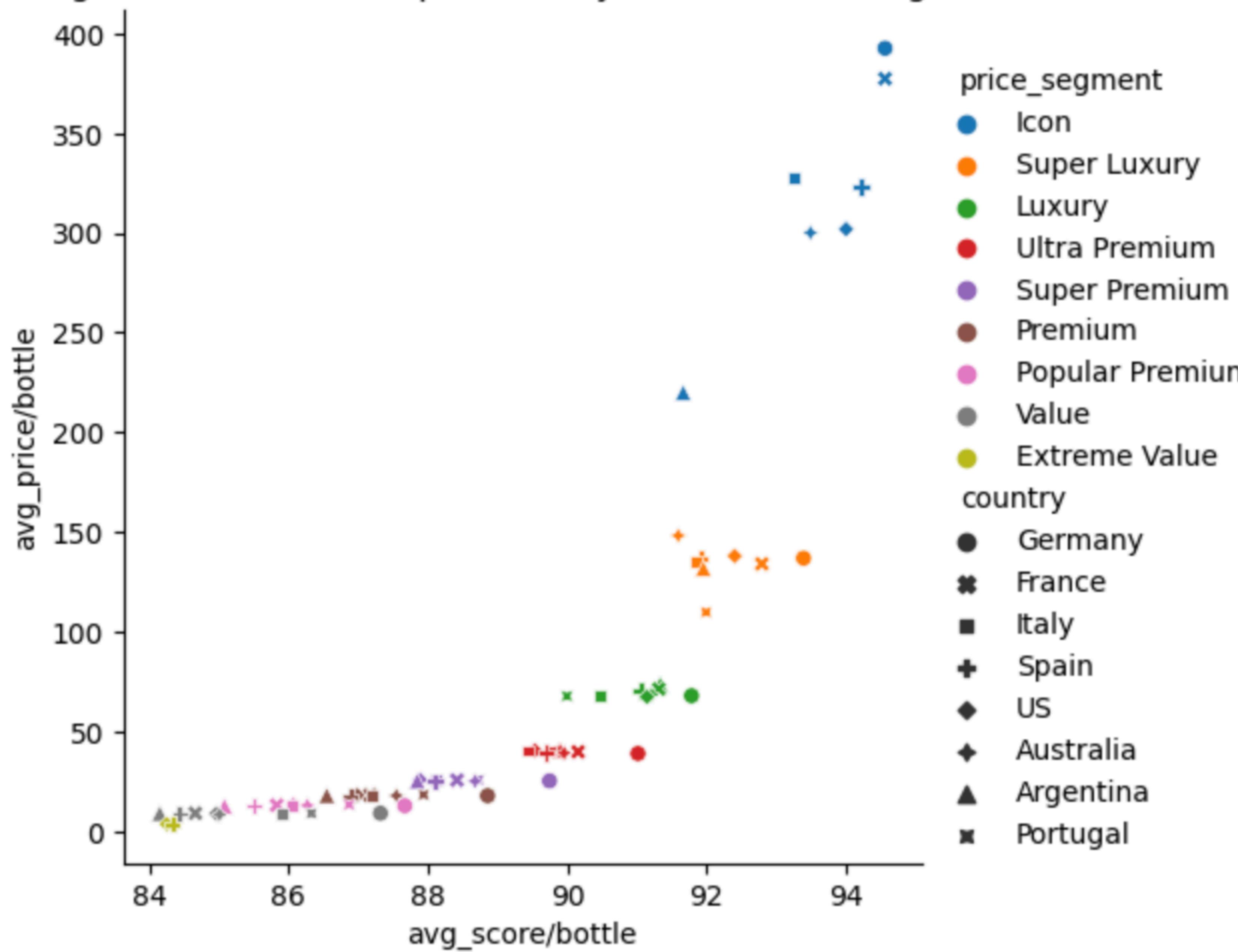


Insights

- The descriptors used for the cheapest price range segments (Extreme Value and Value) tended to emphasize berry and fruit flavors (i.e. 'plum' , 'berry', 'cherry', 'plum', 'apple', etc.). In the higher price range segments, especially Luxury to Icon, descriptor words de-emphasize berry flavors and instead emphasize mouth-feel descriptors such as 'acidity' , 'tannin', 'structure' , 'rich' , 'ripe'.
- The descriptor 'tannin' became more prevalent in judge descriptions after the Popular Premium wine segment. Interestingly, 'tannin' appears to become more important as the wine price segment increases from Extreme Value to Ultra Premium-- then appears to increase more gradually from the Ultra Premium to Icon price segments.
- The descriptors 'ripe' and 'rich' appear significantly more often starting in the Premium price segment, suggesting these terms correspond more closely with higher-priced wines.

Viz 3: The average price/bottle for each price segment was plotted against the average score/bottle. Another aggregation was made to determine the price:score relationship for each country

Avg Price, Score/Bottle per Country for Each Price Segment

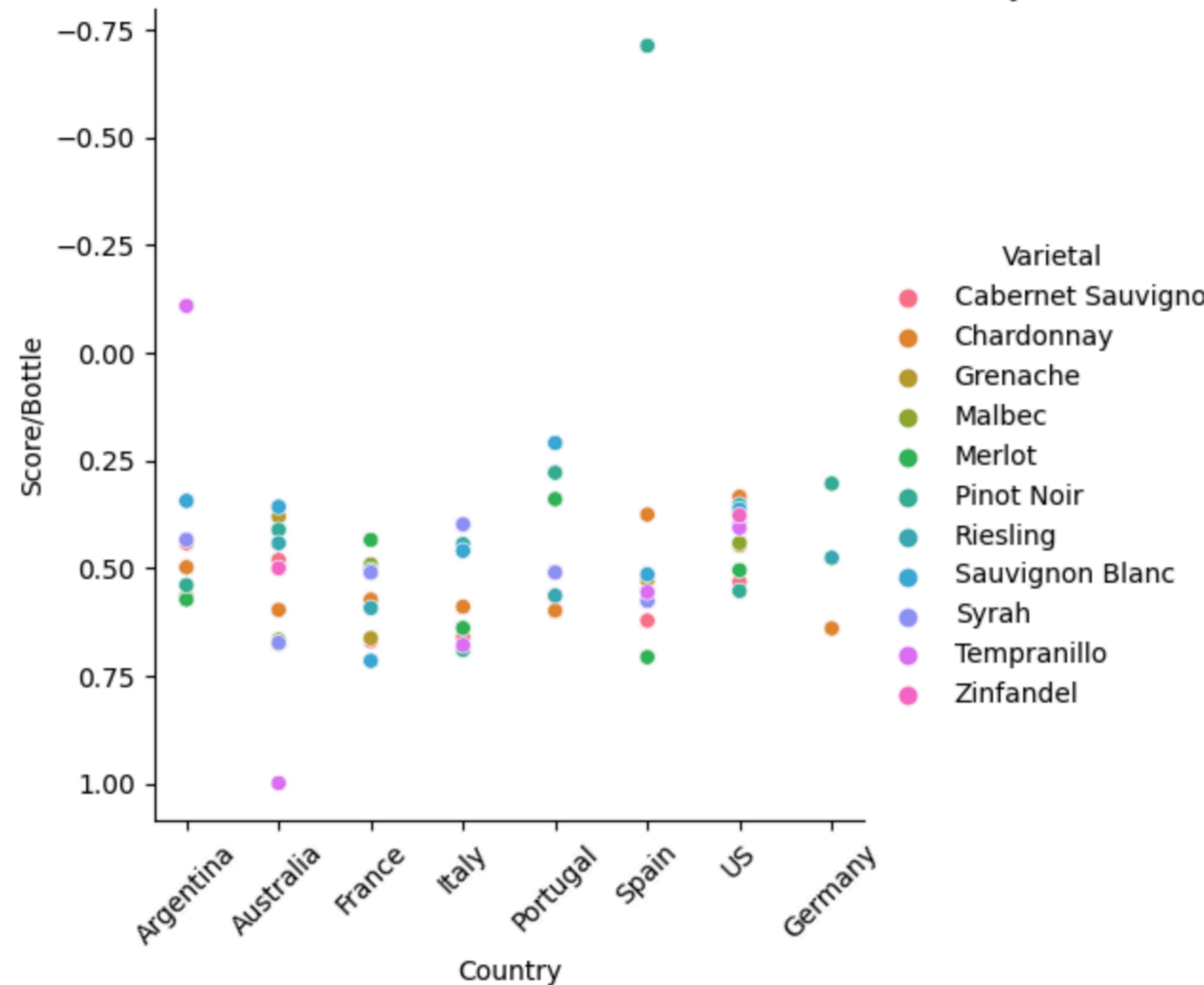


Insights:

- Germany consistently had higher scores on avg in each price segment compared to its peers, except for the Icon price segment.
- Portugal was in the top 2 in avg score/bottle in the Popular Premium and Value price segments. However, Portugal tended to have lower avg score/bottle in both Super Luxury and Luxury segments.
- France, Germany and Spain (all Old World wines) had the highest scores in the Icon segment.
- Old World Wines (France, Italy, Germany, Spain, Portugal) tended to occupy the top 3 spots with respect to avg score/bottle in the Ultra Premium, Luxury, Super Luxury, and Icon Price Segments
- New World Wines (Australia, Argentina, the US) tended to do better in Extreme Value, Value, Popular Premium and Super Premium segments

Viz 4: The correlation between price and score was calculated for each varietal wine, in each country. This was used to determine whether any varietals had high correlation, and whether that correlation was consistent among countries.

Correlation between Price and Score for Varietals in Each Country



Insights:

- r-value of Score to Price (of all wines in the dataframe)= +0.43. This indicates a moderate positive relationship. As price goes up, score is likely to go up as well and vice versa.
- Sauvignon Blanc interestingly had the weakest correlation between score and price for Argentina, Australia, and Portugal