***Market Street Wine***

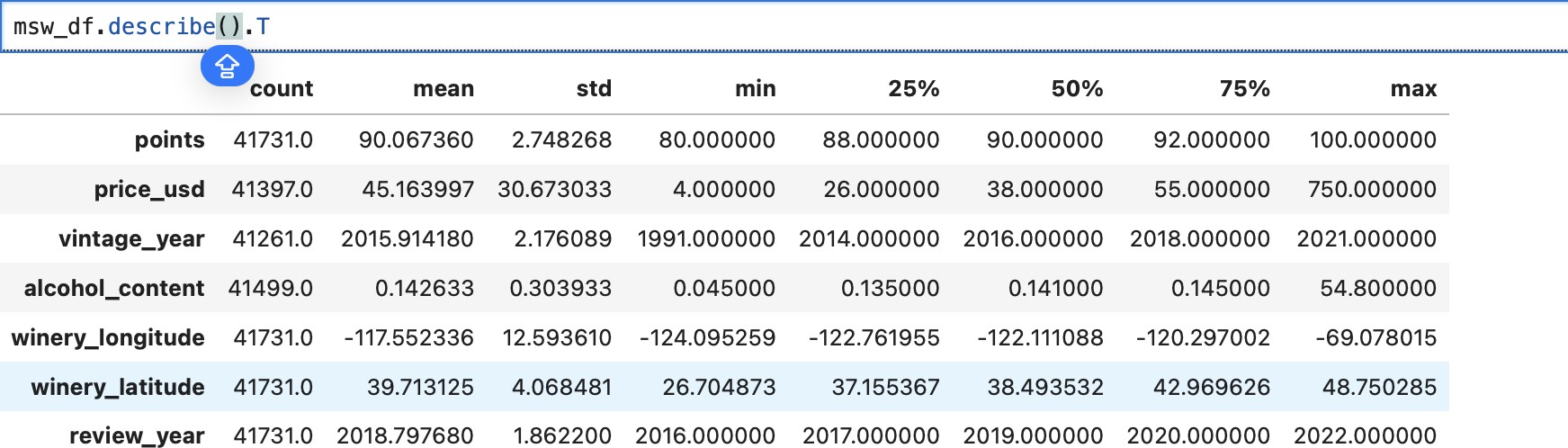
***Data Exploration***

***Observations after reading the case study***

1. Popular wine-producing states in the USA are California, Washington, Oregon, New York, Idaho, Colorado, Ohio, and Virginia.
2. Since opening there were Friday night tastings which had fun and interesting wines, craft beers and a small selection of food.
3. Wine alternatives apart from traditional French and Italian labels ? (Question to answer)
4. Other wine-producing regions that are more easily available? (Question to answer)
5. Any good wines under 20$? (Question to answer)
6. Due to Covid-19 prices were high because of supply chain disruptions and that’s why people preferred cheap wines. Also, the method of purchasing changed from in-person to online.
7. The wine shop had to close during the pandemic and it became more of a warehouse.
8. The customer base evolved when new owners took over. (generally it was mid 40s-50s but shifted to 30s and 40s). There were 3 groups of people according to their wine knowledge.
9. 3 tier structure of distribution. (manufacturers, wholesalers, retailers)

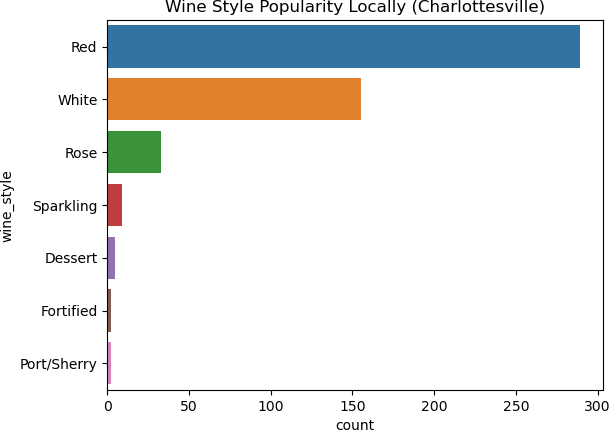
***Observations based on the dataset***

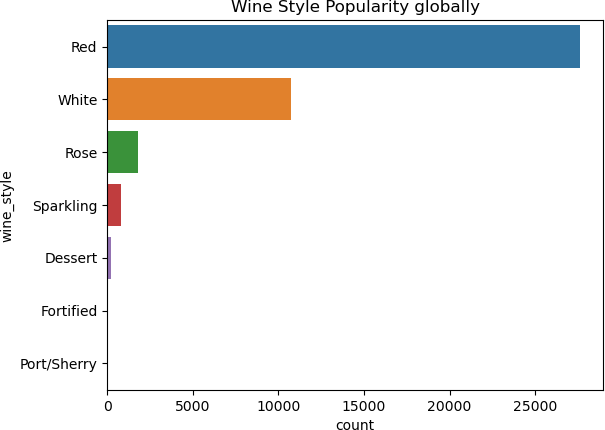
Only continuous variables are considered for the below summary.



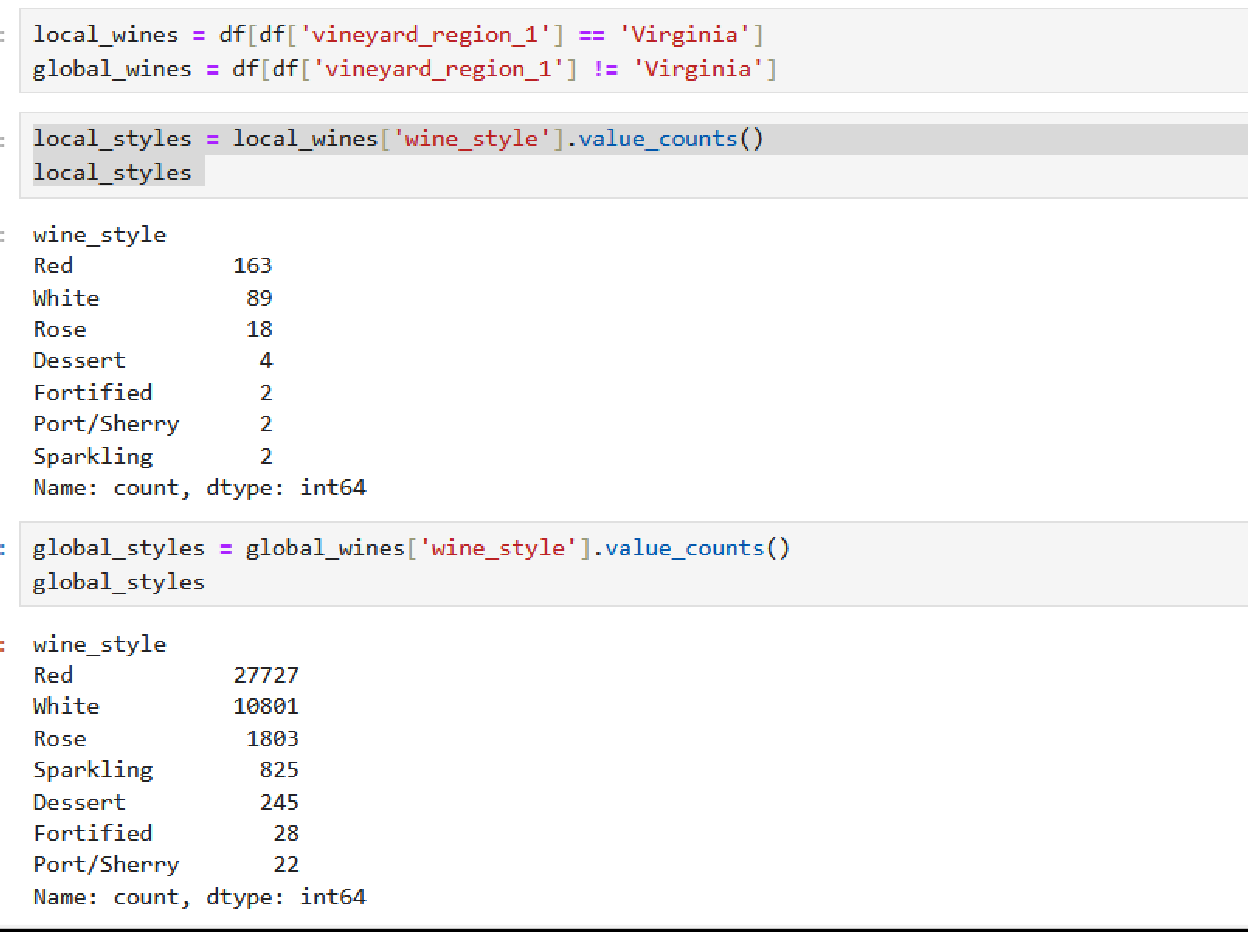
1. The dataset has 41731 rows and 65 columns, the count is not similar in all the columns which indicates that there are missing values.
2. (Assuming the default range for the points variable is 0-100) All the wines in the dataset are of high quality since the points are ranging from 80-100.
3. The range in price suggests that there are very cheap as well as very expensive wines too indicating the variety of wines catering to different customer tastes and price preferences.
4. Vintage year indicates that there are old wines as well as new ones which are from the year 1991 and the new ones are of the year 2021 and older the wine the more exquisite and more expensive.
5. The mean alcohol content of 0.142 tells that most of the wines have a lower alcohol content. Whereas the maximum is 54 which can also be an outlier since the value of Q1, Q2 and Q3 is also similar to the mean.
6. There’s one more column called **bottle\_size** which is a continuous variable but is not included in this summary because it’s datatype is object and not a number (int/float)
7. Badge column has 35138 missing values, price\_usd has 334 missing values, vintage\_year has 470, alcohol\_content has 232, designation has 9931, and vineyard\_region has 1180.
8. **Reviewing the wine trends locally in Charlottesville and globally, what would customers be open to trying?**

Based on Winery\_state





Based on Vineyard region

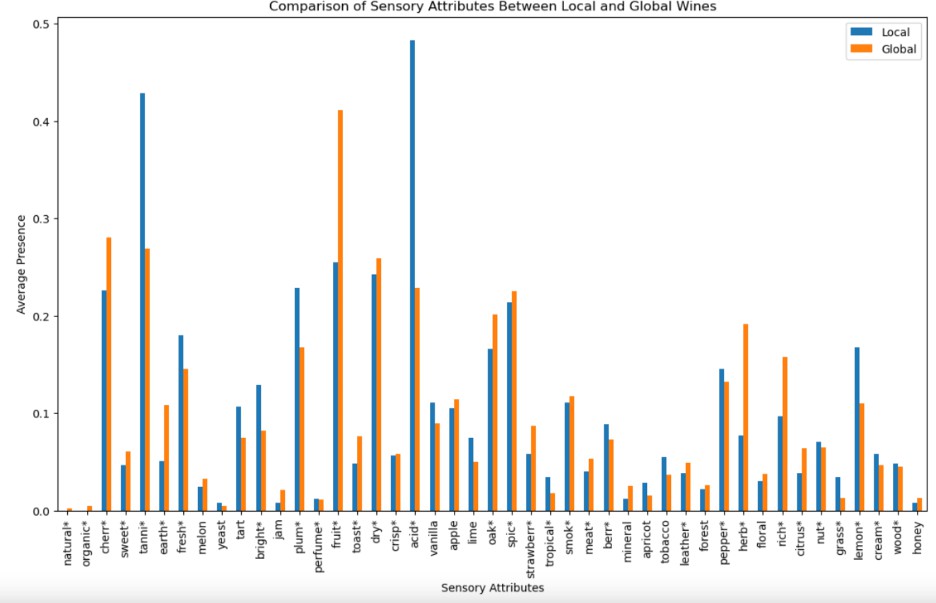


Comparing wine styles based on the winery states and vineyard regions, locally and globally, indicates that red wine is quite popular if compared to any other locally as well as globally across the country..

But there is a huge difference in the kind of red wines that are marketed globally and locally in Charlottesville. Therefore, people would be more open to trying Red wine because it has more varieties as compared to other styles.

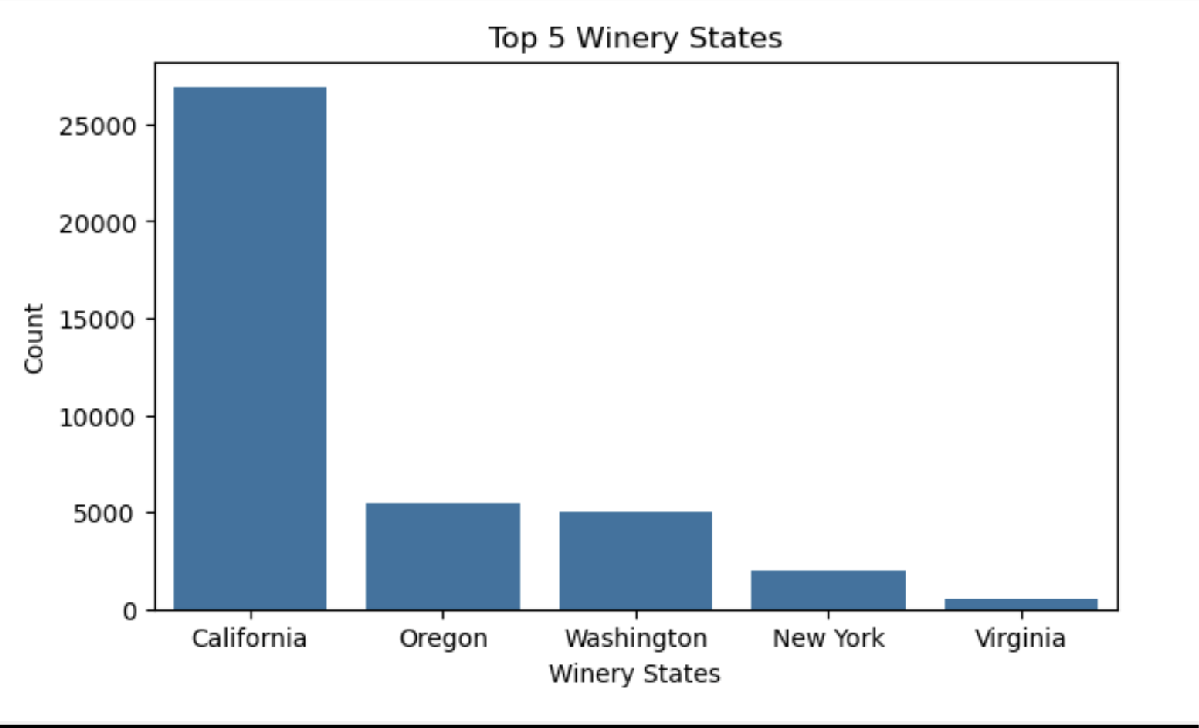
For other wine styles like Sparkling, Dessert, etc. are still consumed by few people, that indicates there is a small population which likes to try different styles of wine both locally and across the country. The least popular one remains the same i.e. Port/Sherry.

Based on the presence of flavors available in the wines, the average presence of acid and tanni is more in local wines, whereas global wines are more herby and high on the fruity note. Overall, in both segments the wines are having a similar amount of crisp level and the natural and organic attributes are close to 0.

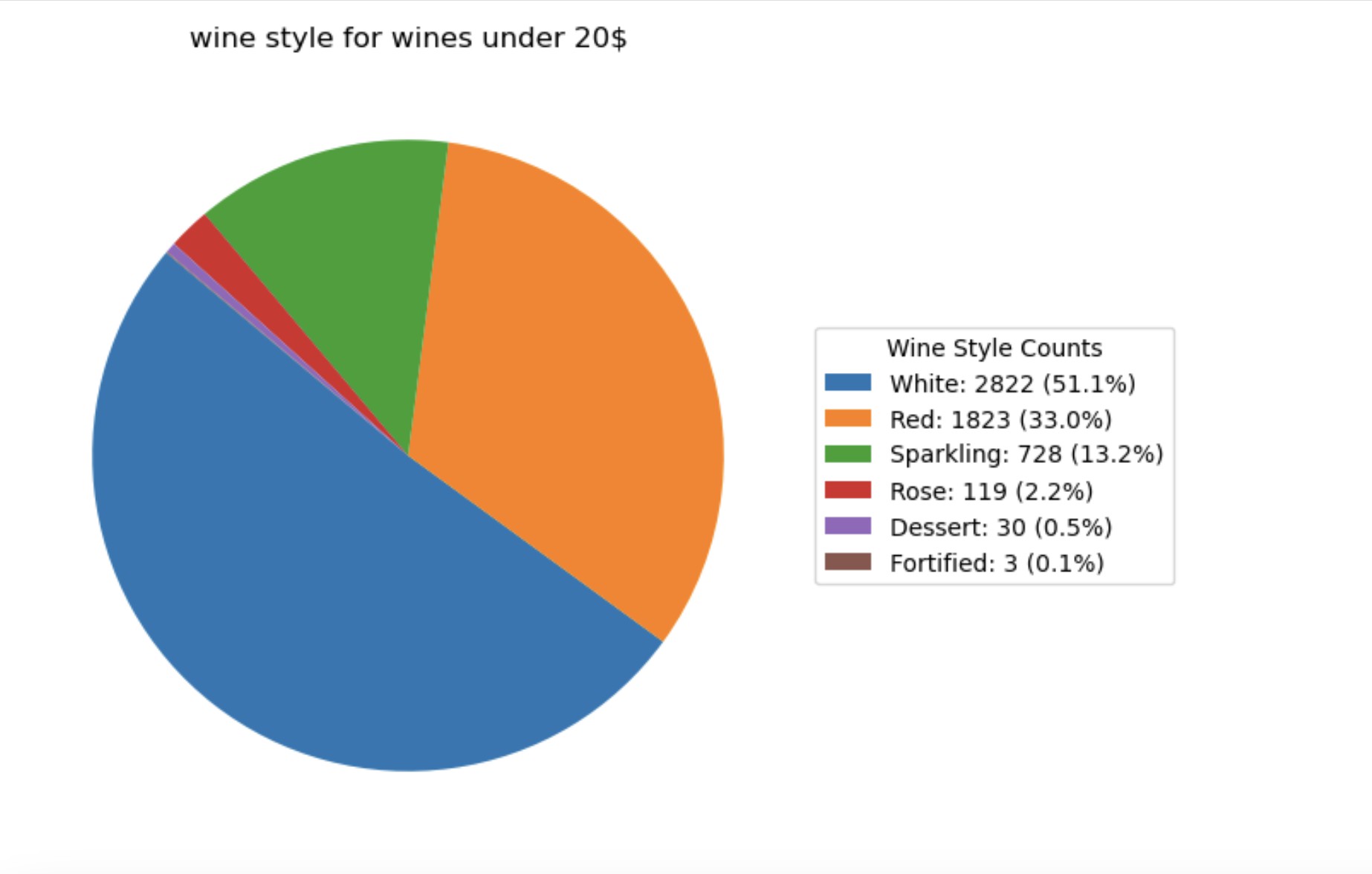


# Are there wine-producing regions more easily available in the US that wine drinkers might be interested in?

Majorly the wine is produced in California, compared to all other regions, This can be the main source to get the stock for the varieties based on the wine style as it is more accessible due to its high distribution, popular demand, or presence in mainstream markets.



## Could they find good ones at the popular price point of under $20?

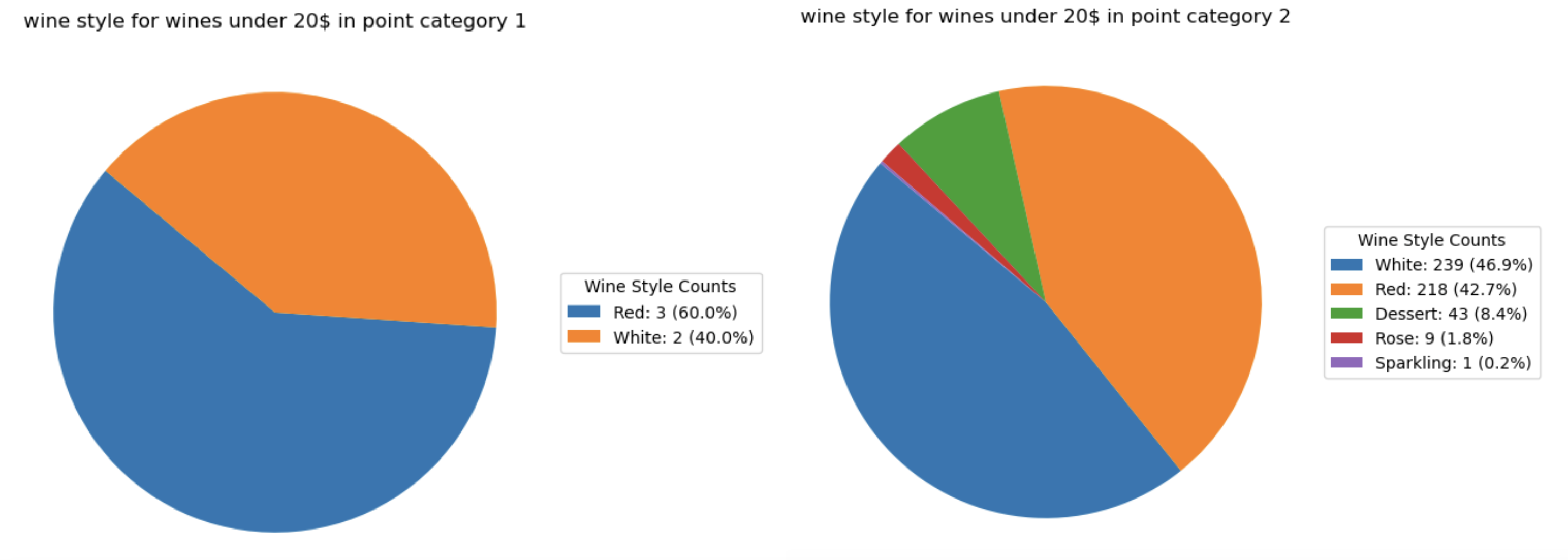
****

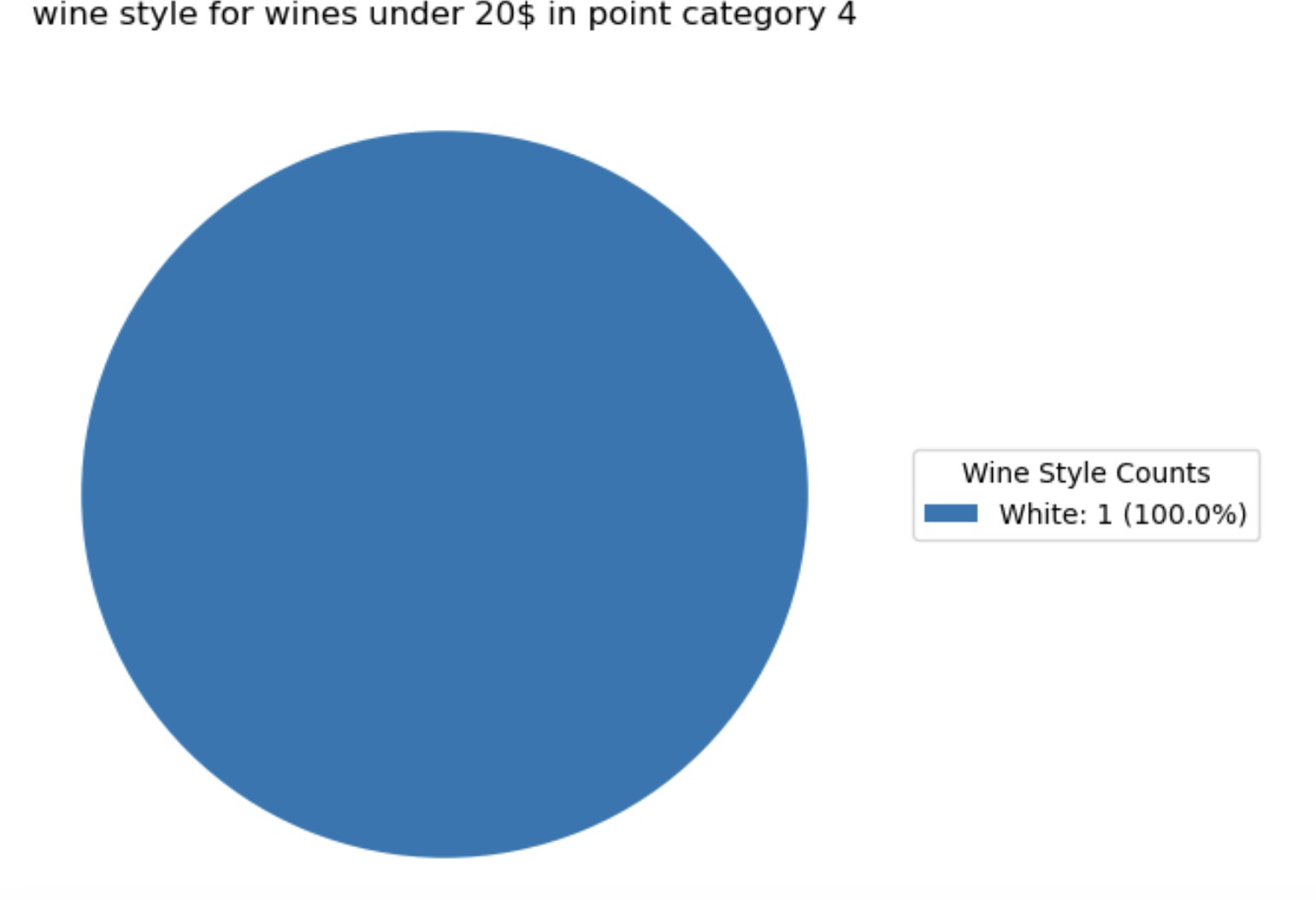
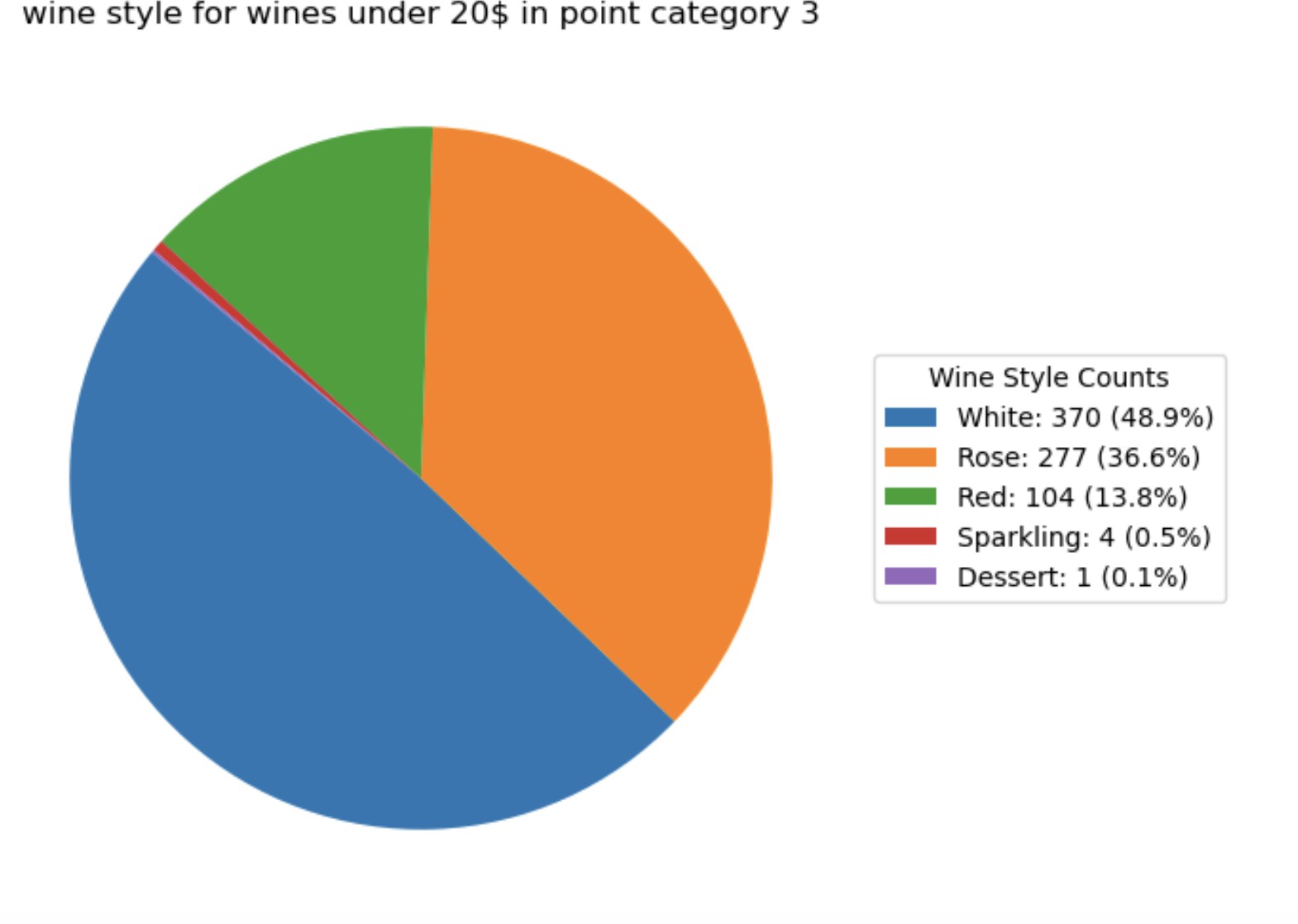
Distribution of wine styles which are below $20 price point indicates more than 50% of the market is taken by White wine and 33% of red wines come under the $20 price and only 2.5% is taken collectively by Rose, Dessert and Fortified (being the least contribution).

Wines are differentiated between good and bad based on the points they have been awarded. The scale of points ranges from 80 to 100 with 100 indicating a really good wine. We can also see badges that indicate awards or recognition for that specific wine as editor’s choice, best buy or cellar selection. These are the wines that have more than or equal to 85 points.

So we can create 4 categories of points and filter them out based on the badges:

* 1. Category 1:- 80-85
  2. Category 2:- 85-90
  3. Category 3:- 90-95
  4. Category 4:- 95-100





When looking at wines under $20, In category 1 (80-85) and category 4(95-100), the selection of wines is limited, with only 5 wines (3 red and 2 white) in category 1 and just

1. white wine in category 4, which are considered as good wines. In contrast, in category
2. and 3 (85 - 95), feature a wider variety of wine styles, though sparkling wines are the least represented.The limited availability in categories 1 and 4 could be because high- quality badges are often awarded to more expensive wines.

Category 1 represents the lowest scoring wines, which suggests they are generally lower in quality. Consequently, it’s understandable that only a limited number of wines in this category receive recognition or awards.

On the other hand, the 95-100 point category typically comprises high-quality, expensive wines that are more likely to receive accolades.

As expected, wines under $20 are sparse in this top category.

The range (**95-100 Category)** includes only a single wine priced below $20

* + *Styring 2014 Whimsy Estate Riesling (Ribbon Ridge)* - White

In contrast, categories 2 and 3 (85-90 and 90-95 points) show a wider selection, with a total of 1,266 wines across various styles. These mid-range categories encompass a greater diversity, although sparkling wines remain the least represented.

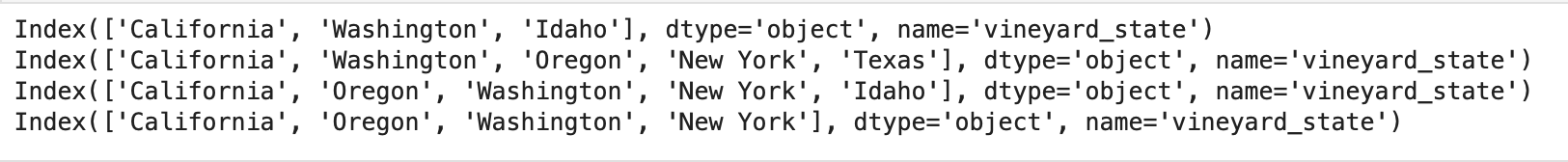
## Wines in the 80-85 Category

* + *Twisted 2014 Cabernet Sauvignon (California)* - Red
  + *Twisted 2013 Moscato (California)*
  + *Hogue 2014 Cabernet Sauvignon (Columbia Valley, WA)* - Red
  + *Colter's Creek 2015 Riesling (Lewis-Clark Valley, ID)*
  + *Happy Camper 2017 Chardonnay (California)*

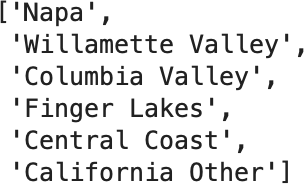
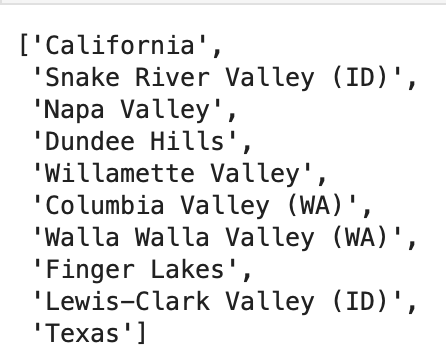
This breakdown highlights the limited recognition for more affordable wines in lower categories and the dominance of high-quality wines in the highest scoring category.

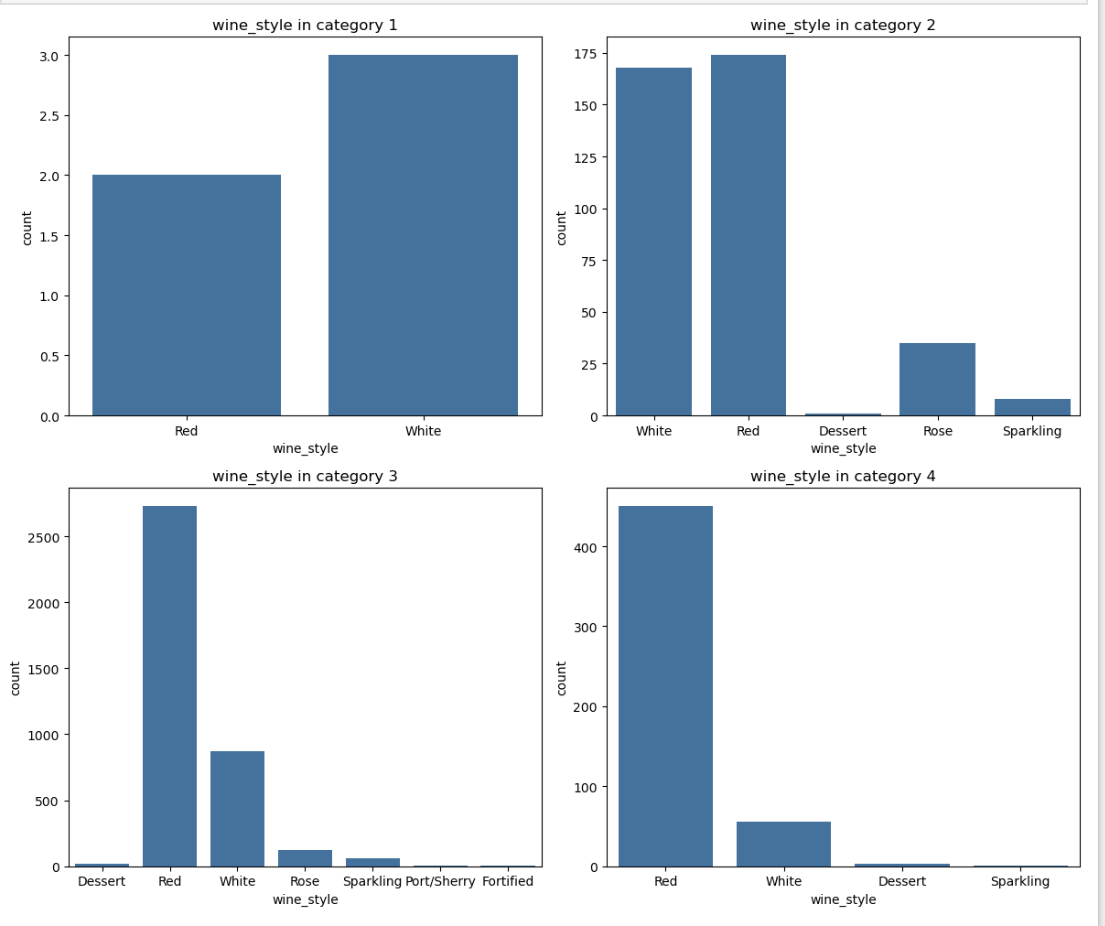
## Are there any specific wines, wine styles, or vineyard regions they should consider?

For this we can also consider the same 4 categories that we created above and take the top 5 vineyard states from them. Eliminating the constraint on the price and considering cheap as well as expensive wines.



Extracting the unique vineyard regions from these states results in a total of **367** vineyard regions. And below are the top vineyard regions from each state.





Across all categories, the majority of wines are either Red or White, indicating a strong preference for these styles among customers.

Overall, there are **3,959** unique wines spanning various styles, with Red and White wines making up the largest share.

## Which other current wine trends and strategies should they explore?

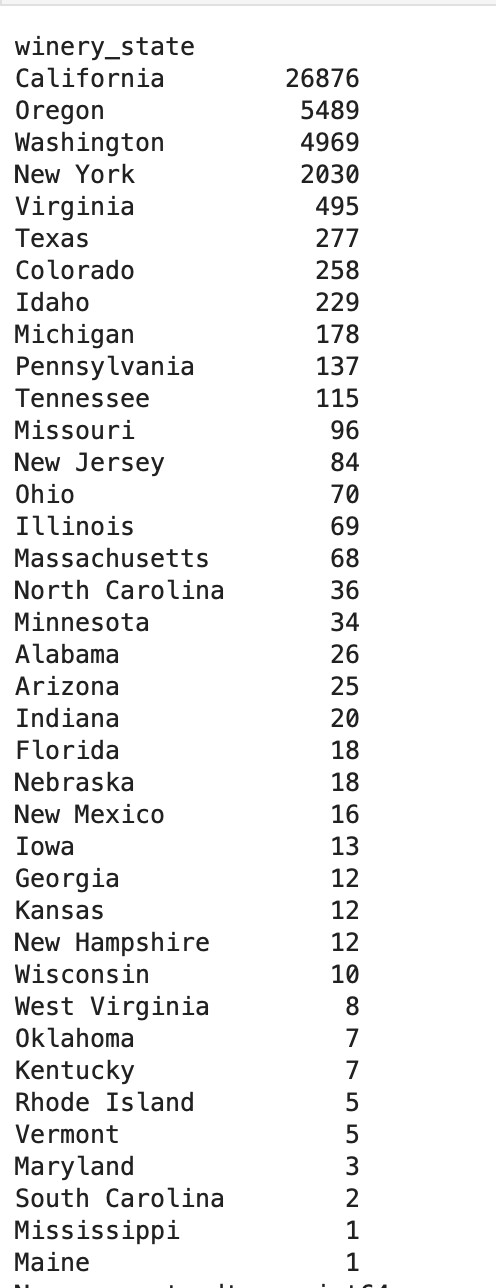
In the past, many people preferred wines with higher alcohol content, but recently there’s been a shift toward wines with lower alcohol levels. Rosé wines have seen a sharp rise in popularity as well. Fans of Bordeaux are now exploring other regions like Burgundy, Barolo, Champagne, Brunello, Rhône, and California. There’s also a growing interest in sustainably produced wines, which are made with environmentally friendly practices. Additionally, supply chain disruptions have led more consumers to try wines from lesser- known regions. This reflects a broader trend of exploration and a desire for quality and sustainability in wine choices.

Based on the information received,people used to prefer more alcohol in wines but due to a movement called “In Pursuit of Balance” in California, they shifted to wines with less alcohol content.

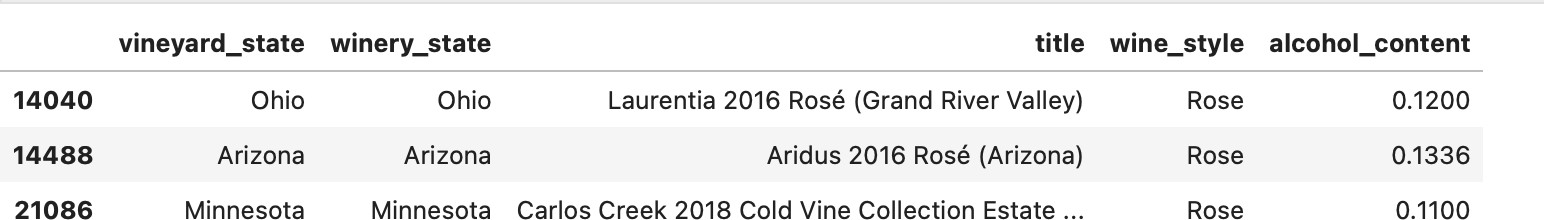
The alcohol content ranges from a minimum of 4.5% to a maximum of 54.8%. To identify low-alcohol wines, consider the 25th percentile as a benchmark, focusing on wines below this threshold. This results in approximately 8,929 unique low-alcohol wines.

As for wine trends, Rosé wine has seen a surge in popularity, so we filtered out all Rosé wines in the dataset, yielding about 1,821 unique options.

Additionally, recent supply chain disruptions have led consumers to explore wines from lesser-known regions. By examining wine production by vineyard\_state and winery\_state, we can identify lesser-known areas, defined here as regions that produce fewer than 100 wines. This approach allows us to focus on unique and potentially underappreciated wine regions.

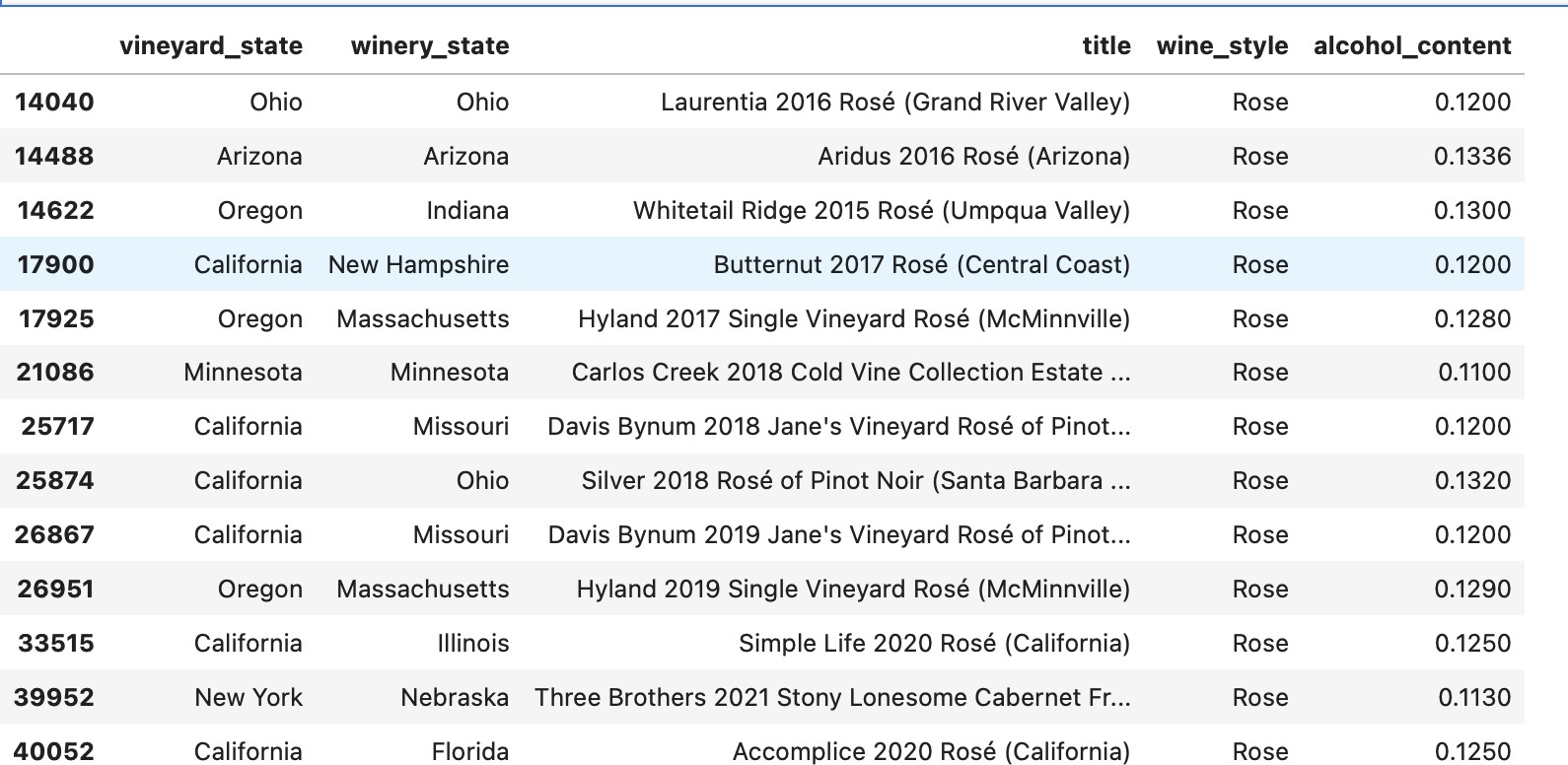
 

Vineyard states Winery states



The selected wines are those with lower alcohol content, fall under the Rosé category, and originate from lesser-known vineyard regions.

Similarly we can also see the same for winery state.



We get the above wines and if observed carefully we can also see that it has the same 3 wines from vineyard state too. All of these wines follow the 3 trends.

**Conclusion and Recommendations -**

In conclusion, comparing wine styles across winery states and vineyard regions reveals that red wine remains the most popular choice both locally and globally. California leads as the primary wine-producing region, providing easy access to a variety of wine styles due to its widespread distribution and high demand in mainstream markets.

This analysis also highlights a trend where high-quality wines dominate the top-scoring category, whereas more affordable wines tend to receive limited recognition. Within the lower price range (below $20), there is a noticeable shift, with white wines making up the majority over reds in categories 2,3 and 4. This suggests that customers looking for good affordable options are more likely to find white wines.

With a total of 367 unique vineyard regions, the market includes 3,959 distinct wines, predominantly red and white, reflecting the broader preferences of wine consumers. There is also a niche for unique options such as Rosé wines with lower alcohol content from lesser- known vineyard regions, offering an intriguing choice for customers seeking something different in their wine experience.

Therefore, the owners can expand their red wine category as it is popular. Along with emphasizing, the point range of 85-95 as those have quality options considering most of the wine styles. They can source from California to remove the constraints of the supply chain as California offers high distribution and production capacity.