



PRELIMINARY WINE ANALYSIS

TO: E. and J. Gallo Winery
FROM: Daniel Monheit
SUBJECT: Case 1 Analysis
DATE : October 17, 2025

Introduction

This document was created to analyze wine data in order to identify how wine type and region affect both price and rating. The dataset includes over 12,000 wines, containing information on type, country, price, and rating. The purpose of this report is to determine whether differences in average prices and ratings are statistically significant across categories and regions. Section 1 examines average prices and ratings by wine type, using descriptive statistics and boxplots to show how the data is distributed. Section 2 compares South American wines to other Southern Hemisphere wines by running hypothesis tests on both price and rating to see if there is a significant difference between the two groups. Section 3 uses Tableau visualizations to present the average rating by type of wine and the map of average rating by country to give a clearer picture of global patterns. The findings show that sparkling wines have the highest average ratings, South American wines differ significantly in rating from other Southern Hemisphere wines, and certain countries, such as Moldova, tend to receive higher ratings than others, such as China. The main recommendation is that producers should continue to emphasize quality and regional reputation to strengthen consumer perception and maintain higher ratings across markets.

Data and Analysis

The dataset includes 12,499 wines with information on type, country of origin, price, and user rating on a 0–5 scale. The purpose of this analysis was to identify how price and rating differ across wine types and regions, and whether those differences are statistically significant. Descriptive statistics and boxplots were used to show averages, distribution, and outliers, while two-sample t-tests were conducted to test for significance between groups. Three questions guided the analysis: whether price and rating differ by wine type, whether South American wines differ from other Southern Hemisphere wines, and which countries show the highest and lowest ratings.

The descriptive results show that sparkling wines had the highest average rating (4.05), followed by red, white, and rosé. Prices were right-skewed, with most wines inexpensive and a few extremely costly, while ratings were tightly grouped between 3.5 and 4.5. Hypothesis testing found no significant difference in price ($p = 0.055$) but a significant difference in rating ($p = 0.00015$) between South American and other Southern Hemisphere wines. Tableau visualizations confirmed these findings, showing small gaps among wine types and notable rating contrasts between countries, with Moldova among the highest and China among the lowest (3.20).

Overall, the results indicate that both type and country of origin influence perception, suggesting producers should emphasize consistency and regional reputation to strengthen consumer trust.

Section 1 – Discussion of Price and Rating

Wine Type	Count of Wines	Average Price	Average Rating
Red	8,379	\$39.84	3.90
Rosé	371	\$10.15	3.70
Sparkling	266	\$51.69	4.05
White	3,483	\$21.74	3.84
Overall	12,499	\$34.17	3.88

Table 1: Wine type counts, average prices, and average ratings.

The largest difference in average price is between Sparkling and Rosé wines (\$41.54), followed by Sparkling and White wines (\$29.95). The smallest price difference is between Rosé and White wines (\$11.59), with the next smallest between Red and Sparkling wines (\$11.85). For ratings, the biggest gap is between Sparkling and Rosé (0.35), while the smallest is between Red and White (0.06).

The boxplot for price shows that wine prices are very spread out and right-skewed, meaning most wines are on the cheaper side while a few are extremely expensive. There are several outliers above the main group, with some prices going well over \$3,000, showing that a small number of wines cost much more than average. In contrast, the boxplot for rating looks more balanced and tightly grouped, with most ratings between about 3.5 and 4.5. This suggests that wine ratings don't vary as much as prices do, and there are only a few outliers, mostly on the lower end, showing that most wines tend to get similar ratings overall.

Section 2 – Comparing South America vs Southern Hemisphere Average Wine Rating and Price

Comparing Average Price. NOTE: South America is defined as Group 1.

Ha: There is a significant difference between average price of South American wines compared to the average price of Southern Hemisphere wines.

Ho: There is no significant difference between average price of South American wines compared to average price of Southern Hemisphere wines.

As seen in Figure 8, the p-value of 0.055 is greater than the significance level (5%), which means we cannot reject the Ho, hence we cannot accept the Ha.

Because our allowed error of 5% (significance level) is less than the actual error of 5.5% (p-value) we cannot reject the Ho. Thus, we cannot accept that there is a difference in pricing between South American and Southern Hemisphere wines.

Ha: There is a significant difference between average rating of South American wines compared to the average rating of Southern Hemisphere wines.

Ho: There is no significant difference between average rating of South American wines compared to average rating of Southern Hemisphere wines.

As seen in Figure 9, the p-value of 0.00015 is less than the significance level, which means we can reject the Ho, and accept the Ha.

Because the p-value of 0.00015 is smaller than the 5% significance level, we reject the null hypothesis. This indicates there is a statistically significant difference in the average rating between South American and Southern Hemisphere wines. In other words, the probability that this difference occurred by random chance is extremely low, confirming that the two groups' ratings differ in a meaningful way.

Section 3 – Tableau Average Rating by Type and Map of Average Rating by Country

[Tableau Public Link](#)

Access the Tableau site using the hyperlink above. As seen in the first worksheet, the average rating by type of wine shows that Sparkling wines received the highest average rating of 4.05 stars, followed by Red wines at 3.90, White wines at 3.84, and Rosé wines at 3.72. While the differences are relatively small, the results suggest that consumers tend to rate sparkling wines more favorably. This could be due to their association with special occasions, consistency in production, or a perceived higher quality. In contrast, rosé wines show slightly lower ratings, which may reflect broader variability in style and consumer preference. Overall, wine type appears to have a minor but noticeable impact on average ratings.

The second worksheet, which focuses on average rating by country, highlights how wine ratings differ by country of origin. The visualization shows that Moldova had one of the highest average ratings, while China had one of the lowest at 3.20 stars. This variation suggests that regional factors such as production methods, reputation, and reviewer expectations may influence how wines are perceived. Established wine-producing nations often receive higher ratings due to familiarity and brand confidence, while emerging markets may face greater variability in quality perception. Together, these findings demonstrate that both wine type and country of origin contribute to differences in consumer ratings, offering insight into global wine preferences and trends.

Conclusion

The analysis found that Sparkling wines received the highest average ratings, while Rosé wines rated the lowest. Price differences among regions were not statistically significant, but ratings between South American and other Southern Hemisphere wines showed a meaningful difference. Additionally, countries such as Moldova tended to have higher average ratings than countries like China, suggesting regional quality and reputation play important roles in consumer perception. Producers are encouraged to focus on maintaining consistent quality and promoting regional strengths to improve consumer confidence and market performance. For any questions or further discussion, please contact Daniel Monheit at demonheit@arizona.edu.

Appendix

Overall Wine Descriptive Statistics

<i>Price</i>		<i>Rating</i>	
Mean	34.16601648	Mean	3.881542523
Standard Error	0.640733466	Standard Error	0.002619202
Median	16.69	Median	3.9
Mode	9.9	Mode	3.8
Standard Deviation	71.63331379	Standard Deviation	0.292823954
Sample Variance	5131.331644	Sample Variance	0.085745868
Kurtosis	473.5221338	Kurtosis	0.119795335
Skewness	14.96480335	Skewness	0.001388222
Range	3407.09	Range	2.4
Minimum	3.7	Minimum	2.5
Maximum	3410.79	Maximum	4.9
Sum	427041.04	Sum	48515.4
Count	12499	Count	12499
Figure 1			
Daniel Monheit			

Red Wine Descriptive Statistics



<i>Price</i>		<i>Rating</i>	
Mean	39.83914906	Mean	3.901205394
Standard Error	0.915053964	Standard Error	0.0033196
Median	18.9	Median	3.9
Mode	9.9	Mode	3.8
Standard Deviation	83.76118292	Standard Deviation	0.303865803
Sample Variance	7015.935764	Sample Variance	0.092334427
Kurtosis	372.5146214	Kurtosis	0.051263595
Skewness	13.55004639	Skewness	-0.059267894
Range	3404.84	Range	2.3
Minimum	5.95	Minimum	2.5
Maximum	3410.79	Maximum	4.8
Sum	333812.23	Sum	32688.2
Count	8379	Count	8379
Figure 2			
Daniel Monheit			

Rosé Wine Descriptive Statistics

<i>Price</i>		<i>Rating</i>	
Mean	10.1489488	Mean	3.71940701
Standard Error	0.23178281	Standard Error	0.01309537
Median	8.84	Median	3.7
Mode	7.9	Mode	3.8
Standard Deviation	4.46445223	Standard Deviation	0.25223471
Sample Variance	19.9313338	Sample Variance	0.06362235
Kurtosis	2.84889239	Kurtosis	0.19720367
Skewness	1.58602871	Skewness	-0.2650422
Range	25.2	Range	1.8
Minimum	3.7	Minimum	2.7
Maximum	28.9	Maximum	4.5
Sum	3765.26	Sum	1379.9
Count	371	Count	371
Figure 3			
Daniel Monheit			

Sparkling Wine Descriptive Statistics



<i>Price</i>		<i>Rating</i>	
Mean	51.6940977	Mean	4.04924812
Standard Error	3.2617003	Standard Error	0.01676185
Median	31.84	Median	4
Mode	14.9	Mode	3.9
Standard Dev	53.196722	Standard Dev	0.27337754
Sample Variance	2829.89123	Sample Variance	0.07473528
Kurtosis	3.52938716	Kurtosis	-0.5271124
Skewness	1.94502507	Skewness	0.08135727
Range	252.5	Range	1.3
Minimum	6.5	Minimum	3.3
Maximum	259	Maximum	4.6
Sum	13750.63	Sum	1077.1
Count	266	Count	266
Figure 4			
Daniel Monheit			

White Wine Descriptive Statistics

<i>Price</i>		<i>Rating</i>	
Mean	21.7378467	Mean	3.83870227
Standard Error	0.53653258	Standard Error	0.00433457
Median	13.9	Median	3.8
Mode	9.9	Mode	3.8
Standard Dev	31.6645145	Standard Dev	0.25581291
Sample Variance	1002.64148	Sample Variance	0.06544025
Kurtosis	138.179089	Kurtosis	0.25645168
Skewness	9.61410328	Skewness	0.02004316
Range	674.77	Range	2.1
Minimum	6.6	Minimum	2.8
Maximum	681.37	Maximum	4.9
Sum	75712.92	Sum	13370.2
Count	3483	Count	3483
Figure 5			
Daniel Monheit			

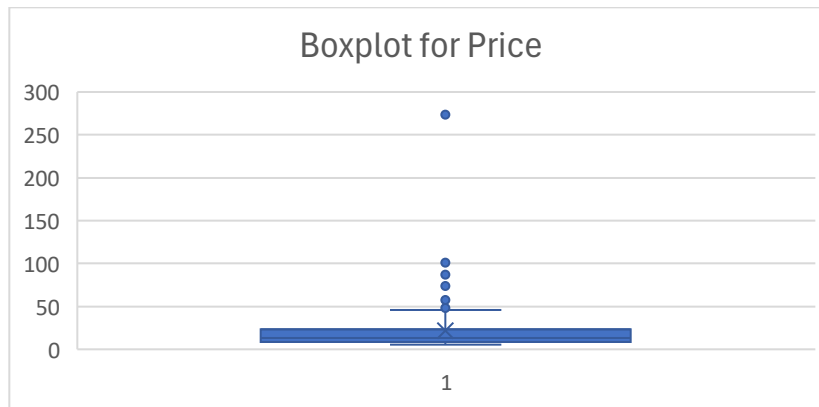


Figure 6

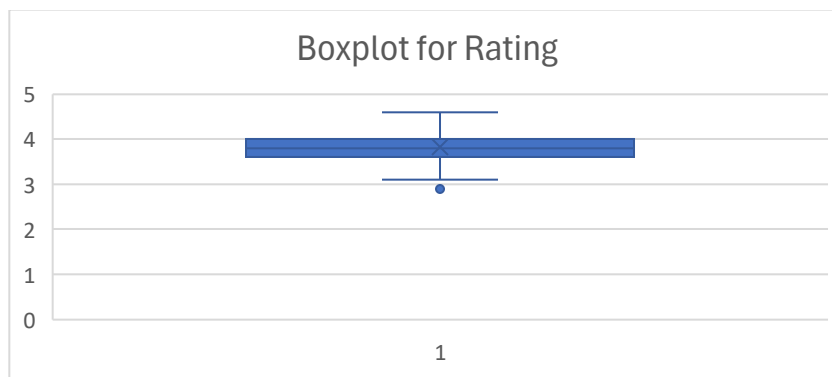


Figure 7

t-Test: Two-Sample Assuming Equal Variances		
	<i>S. America Price</i>	<i>S. Hemisphere Price</i>
Mean	20.43513333	23.77077047
Variance	618.77764	1889.908563
Observations	750	1246
Pooled Variance	1412.437619	
Hypothesized Mean Difference	0	
df	1994	
t Stat	-1.920452394	
P(T<=t) one-tail	0.027471631	
t Critical one-tail	1.645618161	
P(T<=t) two-tail	0.054943261	
t Critical two-tail	1.961154399	
Figure 8		
Daniel Monheit		



t-Test: Two-Sample Assuming Equal Variances		
	<i>S. America Rating</i>	<i>S. Hemisphere Rating</i>
Mean	3.7636	3.818378812
Variance	0.110649239	0.089147892
Observations	750	1246
Pooled Variance	0.097224376	
Hypothesized Mean Difference	0	
df	1994	
t Stat	-3.801319852	
P(T<=t) one-tail	7.41245E-05	
t Critical one-tail	1.645618161	
P(T<=t) two-tail	0.000148249	
t Critical two-tail	1.961154399	
Figure 9		
Daniel Monheit		