⊖ OPEN ACCESS

Analysis of Wine Reviews Using Latin Squares

Apurva Kenekar, Debjani Ganguli, Paul D. Berger[*]

Bentley University Waltham, MA 02452 USA

Abstract	Original Research Article

This paper considers the impact of three factors on the score that a wine is given by wine experts. The study is limited to wines produced in the United States. The "points" awarded to a specific wine, in terms of variety and year and other factors, is a subjective evaluation of the quality of the wine. The three factors this paper focuses on are the Province (location) of the vineyard, the Variety of the wine (Cabernet Sauvignon, Merlot, or Pinot Noir), and the approximate retail Price of a bottle of the wine. Year is held constant. Our findings indicate that Price is a dominant factor that affects "points," but the other two factors cannot be said to have an impact on "points."

Keywords: Variety of wine, Province of wine, Price of wine, Analysis of variance (ANOVA), SPSS, Latin square

INTRODUCTION

As per [1], The United States produced over 800 million gallons of wine in 2016, which accounted for nearly 12 percent of the global wine production volume. The country's wine production is mainly concentrated in California (3,674 wineries), which accounted for about 90 percent of entire U.S. wine production in 2017. According to [2], there were about 7,762 wineries in the United States as of March 2018. Over the past few years, wine sales have held a share of about 17 percent of total retail alcohol sales in the United States [3].

Average wine consumption per United States resident was 2.94 gallons in 2016, up from 2.34 gallons in 2005. U.S. wine consumption from 2005 through 2016 is shown in figure 1. Among wine drinkers, Merlot, Chardonnay, and Cabernet were the three most purchased varieties.
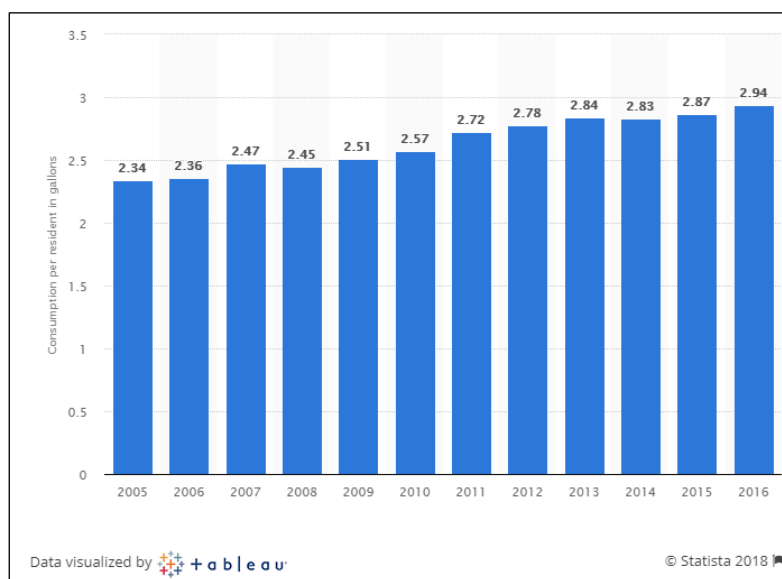

Fig-1: Wine consumption over time in the United States [3]

Our objective is to determine the impact of three factors on the score that a wine is given by wine experts. The study is limited to wines produced in the United States. The "points" awarded to a specific wine, in terms of variety and year and other factors, is a subjective evaluation of the quality of the wine. The three factors this paper focuses on are the Province (location) of the vineyard, the Variety of the wine (Cabernet Sauvignon, Merlot, or Pinot Noir), and the approximate retail Price of a bottle of the wine.

LITERATURE REVIEW

The impact of a wine critics has been tremendous, and has been influencing the way wine-lovers' taste and purchase their wines, for some time now. Several articles (e.g., [4]) and papers, such as [5], have been published which have studied the relationship between consumer wine buying behavior with respect to the medium of wine sales (online, retail, winery), as well as other factors such as lifestyle of the consumers and brand effects of the wine bottle, among others, in recent years.

The topic that we study in this paper is based on *Wine Review*, released by Wine Enthusiast Magazine, and the data are sourced from [6]. A wine rating is a score assigned by wine critics, after tasting a wine, as a summary of that critic's judgement of that wine. A wine rating is, therefore, a subjective numerical quality score given to a specific wine. This is our dependent variable. American wine critic, Robert Parker is credited with popularizing the use of numerical wine ratings.

We first describe the dataset, then our model selection, and next our methodology. We then present our findings and conclusions.

METHODOLOGY

The data used in this paper come from the reviews from the magazine, *Wine Enthusiast*; the data were found online. The dataset is a collection of wine reviews, both qualitative and quantitative. This paper focuses on the quantitative reviews, that is, the scores given to the different types of wines. The original dataset contains the following variables:

Country- The country is which the wine was made
Description – Description of the wine
Designation - The vineyard within the winery where the grapes that made the wine are from
Points - The number of points Wine Enthusiast rated the wine on a scale of 1-100 (though they say they only post reviews for wines that score >=80)
Price - The cost for a bottle of the wine
Province – The state or province of the country the wine is from
Region 1 – Specific region in the province the wine is from
Region 2 – A more specific region in the province in which the wine is from (this field might also be left blank if not applicable)
Tastername – Name of the wine taster
Taster_Twitter Handle – Twitter handle of the wine taster
Title – Title of the wine review in case further details of the review are to be sought
Variety – The type of grapes used to make the wine
Winery – The winery in which the wine was made

The scope of this paper is limited to areas in the U.S., so as to lessen the effects of difference in geography on the quality of wine and to remove any consequent bias. The paper focuses on the three independent variables/factors: Price, Province and Variety. Price is a quantitative variable, and the other two factors are qualitative.

The objective of this study is to understand the impact of the factors: Province, Variety and Price of the wine, on the average number of final points given to the wine by the reviewers. Each factor had 3 possibilities ("levels".) Thus, there were a total of 27 "treatment combinations" (i.e., combinations of different levels of the factors) available (i.e., 3x3x3.)

Table-1: levels of each factor that we chose, traditionally labeled Low, Medium, High, are

	Low	Medium	High
Price (US$)/bottle	20	30	40
Province	California	New York	Washington
Variety	Cabernet Sauvignon	Merlot	Pinot Noir

As we noted earlier, the dependent variable in this study is the points assigned to the different wines by reviewers. The objective of this study is to understand the effect of Price, Province and Variety of the wine on the final score assigned.

We used a Latin-Square experimental design and, hence, are making the assumption that there are no interaction effects among the factors. Since we are using the Latin-Square design, only 9 of the 27 treatment combinations are chosen for study – of course, the 9 chosen must carefully selected conform to what constitutes a Latin-Square design – including the fact that there are three factors (i.e., by definition, a Latin-Square design has three factors) and all factors at the same number of levels, although there can be any number of levels.

Unreplicated Latin Square Analysis

Table-2: Using the following design as a framework, the experiment was conducted

	B_1	B_2	B_3
A_1	C_1	C_2	C_3
A_2	C_2	C_3	C_1
A_3	C_3	C_1	C_2

What the above diagram represents is a set of 9 treatment combinations. We refer to each shaded-in rectangle as a "cell." That is, the upper left cell means that factor A is at level 1, factor B is at level 1 and factor C is at level 1. The (3,3) cell – the lowest on the right, indicates the treatment combination of factor A at level 3, factor B at level 3, and factor C at level 2. In other words, there are no *numerical values* (yet) in the cells. The diagram represents the levels of three factors on a two-dimensional display. Data will, of course, be forthcoming for each cell.

For our study, Province is factor A with California labeled level 1 (A_1), New York level 2 (A_2) and Washington level 3 (A_3). Variety of wine is factor B, similarly coded as Cabernet Sauvignon B_1, Merlot B_2 and Pinot Noir B_3. Finally, the Price is the third factor, C, which has 20 as C_1, 30 as C_2 and 40 as C_3.

Table-3

	Cabernet Sauvignon	Merlot	Pinot Noir
California	$20	$30	$40
New York	$30	$40	$20
Washington	$40	$20	$30

Hence, these 9 treatment combinations form the basis of the study. If the reader looks carefully, he/she will note that among the 9 treatment combinations, each level of each factor is once and only once paired with each level of each other factor. For example, there is only one cell that has the combination (B3, C1) – or (Pinot Noir, $20) - (middle cell of right-most column.) This is one core property of a Latin-Square design, and guarantees unbiased main-effects, under the assumption that there are no interaction effects. Naturally, the core benefit of using a Latin-Square design is that we require only 9 treatment combinations, instead of having to pay for all 27 treatment combinations. The above 3-level Latin-Square design is not unique, but since there are only twelve 3x3 Latin-Square arrangements, the probability of picking a random 9 treatment combinations out of 27 possibilities, given 3 factors at 3 levels each, and having them constitute a Latin Square is very small (close to 1 in 5 million.) The probability would be even smaller if each of the three factors had more than 3 levels.

Analysis and discussion of results

After the data values were collected for the aforementioned 9 treatment combinations (a random wine was chosen which conformed to each of the 9 treatment combinations), the ANOVA was performed using SPSS and the results found are in Table 5. Note that VAR00002 = A, Province, VAR00003 = B, Variety, VAR00004 = C, Price.

Table-4: Latin-Square designs and the data

	1	2	3	1	2	3
1	A	B	C	82	89	91
2	B	C	A	84	91	81
3	C	A	B	94	80	88
1	C	B	A	95	90	81
2	B	A	C	88	83	87
3	A	C	B	83	94	88
1	B	A	C	91	82	95
2	A	C	B	84	91	92
3	C	B	A	93	91	82

Table-5: ANOVA results for Latin-Square design

Univariate Analysis of Variance

[DataSet0]

Between-Subjects Factors

		N
VAR00002	1.00	3
	2.00	3
	3.00	3
VAR00003	1.00	3
	2.00	3
	3.00	3
VAR00004	1.00	3
	2.00	3
	3.00	3

Tests of Between-Subjects Effects

Dependent Variable: VAR00001

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	190.000ª	6	31.667	4.524	.192
Intercept	67600.000	1	67600.000	9657.143	.000
VAR00002	8.000	2	4.000	.571	.636
VAR00003	.000	2	.000	.000	1.000
VAR00004	182.000	2	91.000	13.000	.071
Error	14.000	2	7.000		
Total	67804.000	9			
Corrected Total	204.000	8			

a. R Squared = .931 (Adjusted R Squared = .725)

From the results, it can be seen that at $\alpha = 0.1$, only Price seems to have an effect on the Points given to the wine by the reviewers, and its p-value is .071 (i.e., not significant at $\alpha = .05$.) The other two factors, Province and Variety, cannot be said to have an effect on the Points given. This was a surprising result, and we decided to replicate the experiment – i.e., add more data. Specifically, we chose two different, additional, Latin-Square designs and ran them also. This gave us 27 data points in total. Table 2 shows the three Latin-Square designs, and the actual data results.

The ANOVA results are shown in Table 6.

Table-6: ANOVA results for design with three Latin Squares

	SS	df	MS	F	p-value	sig
Price	500.667	2	250.333	65.621	0.000	yes
Rows	14	2	7	1.835	0.188	no
Columns	4.667	2	2.333	0.612	0.553	no
Rep	24.667	2	12.333			
Error	68.667	18	3.815			
Total	612.667	26	23.564			

From the results of the ANOVA in Table 6, it can be seen that there is a very significant difference among prices, although, overall, the results otherwise have not changed substantially. Thus, from these experiments it can be said that there is sufficient evidence to conclude that Price has a significant effect on the points given to the wines, whereas there is not sufficient evidence to conclude that Province of the wine, or Variety of the wine, has an influence on the points given to the wine.

Limitations and directions for future research

One major drawback of using Latin Squares is the assumption of no (or negligible) interaction. There could be a possible interaction between the Province in which the vineyard resides and the Price of the wine. Hence, this is a limitation to the research results. In retrospect, the second phase of the study may have been improved by considering all 27 treatment combinations (same number of data points) without any replication, instead of three different and slightly overlapping Latin Squares (in terms of differing treatment combinations). This would have allowed selected interaction effects to be examined (with the assumption of other, higher-order, and interactions being assumed to be zero.) Yet, since the results were very similar in the expanded study – the only material difference being the much, much higher degree of significance of the effect of

Price – it is less likely that any interaction effects would have been significant.

Due to the paucity of data on wine options in the higher price ranges produced in all three provinces and in all three varieties, the maximum price had to be limited to $40, which is not considered a high price in the world of wine consumption. Hence, further research could also be conducted on the higher-priced wines in comparison to the lower-priced wines.

REFERENCES
1. A*verage wine consumption per United States resident from 2005 to 2016 (in gallons)*. Retrieved 06:10, December 5, 2018, from: https://www.statista.com/statistics/233734/average-wine-consuimption-of-the-us-per-resident/
2. Zackthout. *Wine Reviews*. 2017. Retrieved December 5, 2018, from: https://www.kaggle.com/zynicide/wine-reviews
3. *US. Wine Market- Statistics & Facts. Retrieved* 05 December 2018, from: https://www.statista.com/topics/1541/wine-market/
4. Anderson P. *How the impact of critics in the wine industry is changing*. 2014. Retrieved from: https://www.linkedin.com/pulse/20140619231015-65848538-how-the-impact-of-critics-in-the-wine-industry-is-changing/
5. Lockshin L, Corsi AM. Consumer behaviour for wine 2.0: A review since 2003 and future directions. Wine Economics and Policy. 2012 Dec 1;1(1):2-3.
6. Wikipedia contributors. Wine rating. In *Wikipedia, the Free Encyclopedia*. 2018. Retrieved 06:10, December 5, 2018, from: https://en.wikipedia.org/w/index.php?title=Wine_rating&oldid=844999313