

# Wine Critic Scores and Consumer Behavior in a Major USA Metropolitan Market

Andrew P. Snow

School of Information and Telecommunication Systems  
Ohio University  
Athens, Ohio  
e-mail: asnow@ohio.edu

Gary R. Weckman

Department of Industrial and Systems Engineering  
Ohio University  
Athens, Ohio  
e-mail: weckmang@ohio.edu

**Abstract**— In this paper, we investigated three questions. First, to what extent do wine critic scores and descriptions influence consumer-buying decisions? Second, to what extent this influence varies with price? Third, how do demographics affect consumer decisions? The experimental design consisted of convenience samples from four different stores in a major United States (US) metropolitan market, with random assignment of consumers to different groups, who completed a total of 240 survey questionnaires. The dependent variable was likelihood to buy wine when presented with varying amount of wine critic information (a control group and three experimental groups with different levels of information). Independent variables included wine price, age, gender, wine interest, store type and location. Major findings include some surprises. For a \$20 bottle of wine, the critic information was not a factor on likelihood to purchase, while critic information was a factor for a \$50 bottle. These findings varied somewhat based on demographics such as wine club membership, gender, type store, and store location. The findings have implications for wine producer, distributor and retailer marketing methods.

**Keywords**— wine critic; consumer behavior; wine marketing; tukey-kramer

## I. INTRODUCTION

Total world wine production in 2014 was estimated at 271 million hectoliters [1]. The world is producing more wine than ever and Europe is not the only player; new world countries, such as the United States of America (USA), Argentina, Chile, Australia and South Africa are producing large quantities of wine. With China on the horizon, the world production will continue to rise, and there are now thousands of consumer wine choices. Consumers are exposed to a range of wine advertising including magazine ads, emails, website popups and newsletters. Extensive store displays lend perception that wines are of quality and must be purchased. At the end of the day, price range is certainly an important factor consumer wine purchasing decisions. As people only have a certain budget for wine expenditures, a natural question is, how important are wine critic scores in consumer choices?

This is a very important question because of the proliferation of wine ratings in the marketplace. There is evidence to suggest that consumers might be willing to spend

a bit more because of a wine critic score. Whether it is a 97-point score from The Wine Advocate or landing in the top 10 of The Wine Spectator's top 100, significant positive attention is great for a wine's image and marketability. Retailers and restaurants may seek the wine and distributors could very well change their purchasing strategy to meet market demand. These scores also affect wineries because higher scores usually lead to higher prices for retailers and direct sales. Also, wine critics are known to have certain specific tastes and it can be argued that if a winery can tailor their wine to that taste and get a high score, perhaps they could increase their price. The purpose of this research is to test how critic's wine scores and descriptions affect consumer behavior in a large USA metropolitan market.

Experimental research in a major USA metropolitan marketplace was conducted to investigate three questions regarding the impact of wine critic information in buying decisions:

- R1. How do critic scores and descriptions affect consumer decisions?
- R2. If wine critic scores/descriptions do affect consumer decisions, how sensitive are decisions to price?
- R3. How do consumer demographics like gender, age and wine club membership affect consumer decisions?

In Section 2, a background on wine assessment and critic scoring. In Section 3, the methodology of the testing is proposed and explained in detail. In Section 4, the results are compared against each other. Finally, in Section 5, the major findings and overall conclusion is stated.

## II. LITERATURE REVIEW

Wine is quite arguably the most complex alcoholic beverage in existence. While beer and spirits are produced from grains and vegetables, there are thousands of grape varieties. Beer and spirits can be produced year-round, but for wine production, grapes are only available once a year.

How good a taster is the modern wine critic and what are the necessary qualifications to become a critic? Jancis Robinson is a very influential critic and holds the title of Master of Wine, which is one of the highest accolades attainable in the wine business. Ironically, Robert Parker has no formal wine training or certifications however, for his efforts during his career; he has received the highest honor

possible from Italy and France (Commendatore and Knighthood in the Legion of Honor, respectively). Parker has also authored over a dozen books translated into multiple languages. This suggests that a professional certification in wine may not be necessary and experience can be just as valuable as accreditation. Perhaps people would buy a wine highly recommended by a critic because the perception is that the critic knows more about wine than the average person does and have better tastes.

How do people assess wine? There is little evidence to suggest that everyone will enjoy a very expensive wine. A study of 6,000 blind tastings [2] between inexpensive and expensive wine showed that experienced tasters (minority of the population) enjoyed the more expensive wine; however, the inexperienced wine tasters more often enjoyed the less expensive wine.

A test conducted in the USA on how Generation Y buys European wine [3] claims that Generation Y accounts for 72% of food and beverage sales in the USA, and that European wineries should target this group. Their test concluded that when Generation Y purchases European wines, brand familiarity and experience are key factors in the decision process. Interestingly, researchers did not test wine critic reviews for this peer group, and the questionnaire did not ask the subjects if they subscribed to wine publications. Additionally, the test was limited to European.

The problem is that with many more countries producing wine, there are a myriad of choices for consumers. Also, rules regarding labeling vary from country to country (percentage of grape variety/vintage to be used in labeling a wine) which always leads to the infamous traditionalist versus modern debate. A major criticism of European wine labels is that they are hard to understand. Many labels do not have an indication of grape variety and give little clue on the back label as to what is in the bottle. This is confirmed by a survey published in Wines & Vines [4], an organization that compiles wine metrics and information about the wine business. According to a survey conducted in 2005, wine labels confused 36% of USA wine drinkers, 51% of drinkers found imported wine labels very difficult to read/comprehend, and 81% of wine drinkers want labels that are easy to read and understand. Interestingly, 51% of wine drinkers liked humorous wine labels.

It can be argued that the modern labelling simplifies the wine process for the consumer because the labels are easy to understand and indicate grape variety. However, new world labels can be misleading because in the USA, only 75% of a grape variety is required to label that wine as a single variety. Thus, a wine with 76% zinfandel and seven other grapes can be labeled solely as a zinfandel. It could be argued that such labeling is misleading. Either way (traditional or modern), there are many choices, which lead one to question whether consumers are confused. Another study [5] suggests that too many choices can be positive, increasing the likelihood of satisfying customers.

At the top end, there are educated and wealthy connoisseurs who are not afraid to spend for what they know or want. However, sales of Romanée Conti and first growth Bordeaux represent a miniscule amount of wine purchases.

At the other end, \$10 for a bottle may be too expensive for some people – price is their selection criterion. This leaves a gaping hole of consumers in the middle that are potentially very confused. Consumers may choose to buy a wine for a multitude of reasons (price, prestige, style, country of origin, etc.) and a critic review is something to consider.

A study done in Australia [6] focused on how consumers assess wine. Their report concluded Australian consumers peruse a wine display shelf for one minute. Their research focused on shelf information in front of the bottle (shelf-talker) because although there are many sources of opinion in Australia, there is no major dominating critic influence. Their study concluded that wines with shelf information increased the choice of a particular wine by 7.4%. It is common to find shelf information in the USA; however, the USA is home to some of the world's most influential critics (The Wine Advocate, The Wine Spectator) and their information may be used as well. It is important to note that while the influential critics in the USA taste and rate some of the worlds' most sought-after wines; they also do the same for inexpensive wines. In fact, The Wine Advocate and The Wine Spectator have value reports in their publications that focus on inexpensive/good value wines with good ratings. These publications review many kinds/prices of wine and target professional, collector, intermediate and novice consumers.

Are wine critic reviews always effective? Someone went so far as to say the wine critics are "BS artists" [7]. Quandt bashes critics because scores vary considerably for the same wine, sometimes critics contradict themselves, but most importantly, he attacks the critics for their vocabulary, which no one can understand. What does scorched earth taste like? Or how do zesty minerals differ from minerals – are crushed rocks really delicious? Quant argues that reviews do not tell you much about a wine in terms that an ordinary person can comprehend.

Adding to the argument, tests were conducted by Weil. In one test [8], Weil provided subjects with three wines, of which two were the same. For those who can correctly identify which one is different, they are then presented with the critic review to see if they can match the wine to the verbiage. Only 51% of the subjects could correctly identify the correct review. Weil conducted another test [9] to see if people agree with the wine critic scores. The test was the same, except for those who correctly identify the different one, they were then asked to select which one they prefer. The wines used for the test were from the same producer, one was the entry level, and the other was the reserve. His test concluded that for those who could identify the difference, 52% liked the reserve wine better. Consequently, Weil claims that the average person does not benefit from a critic review because they may not like the same style wine as a critic.

While the previous tests may suggest that wine critic reviews are not effective, there is evidence and reason to suggest that they are perceived to be valuable. The Wine Advocate has at least 40,000 subscribers in every state in the USA and have subscribers in 37 foreign countries [10]. A paper on the impact of wine critics [11] suggest that these

reviews can be very valuable for Bordeaux lovers, as wine critics taste almost all Bordeaux wines (at the en priemeur tasting held every April). Most wine consumers cannot attend these events; thus, they may rely on the insight of wine critics. A test was conducted on Bordeaux wines rated by Robert Parker [11] and it was concluded that a “Parker effect” existed and affected pricing for Bordeaux wine by as much as 3 Euro per bottle (highly graded wines were affected most, and the effect diminished for lowly rated wines).

The effectiveness of wine critic scores was tested in the USA in a national grocery chain, targeted in northern California [12]. Over a 2-month period, 32% of all the wines that had ratings were selected for tasting. These wines had critic expert opinions and scores in front of them on the shelves. Using scores from 78 to 89, the test concluded that wines with expert opinion and scores increased demand by 25%, while lower scoring decreased demand. This test may help explain the aforementioned study about consumer confusion because in a grocery store chain, the selections are quite large and within a category, there are dozens of choices within a similar price range. Interestingly, wines that were not selected for testing did not change significantly in sales, illustrating that not everyone may pay attention to reviews or might have already developed brand loyalty before the test was conducted. Although this test was only conducted for two months and in a very small area, it does shed light on the possible effectiveness of wine critic scores.

### III. METHODOLOGY

In the research presented here, we focused testing in a setting where the sole purpose of the store visit is to buy wine from a wine outlet. We chose to test the effectiveness of wine critic reviews in the Baltimore, Maryland (MD) to Washington District of Columbia (DC) corridor. The stores ranged from small boutique selections to larger stores that cater to everyone in terms of selection and price.

To address the aforementioned research questions, we generated three sets of major null and alternative hypotheses:

#### Wine Critic Score Influence

- H1<sub>0</sub>: The WA score has no effect on consumer decisions on selecting wine to buy.
- H1<sub>a</sub>: The WA score influences consumer decisions on selecting wine to buy.

#### Wine Critic Descriptions Influence

- H2<sub>0</sub>: Wine descriptions have no effect on consumer decisions on selecting wine to buy.
- H2<sub>a</sub>: Wine descriptions influence consumer decisions on selecting wine to buy.

#### Wine Critic Score versus Description

- H3<sub>0</sub>: Critic wine scores are given no more weight than wine descriptions by consumers in wine purchasing decisions.
- H3<sub>a</sub>: Critic wine scores are given more weight than wine descriptions by consumers in wine purchasing decisions.

For the test, we developed a one-page questionnaire for consumers buying wine from wine outlets. Each subject was presented with certain information about two wines. Each subject was asked how likely they would be to buy each wine, by rating on an anchored scale of 1 (lowest) to 7 (highest). Four questionnaires were developed with different information about each wine. All subjects were asked the same questions for each wine, regardless of the experimental group.

Three of these groups were experimental groups and one was a control group. Sampling (store selection) was based on convenience and there was random assignment of subjects to the four different groups.

- Experimental Group A - Subjects were presented with Châteaux, price, critic score and a critic description of the wine.
- Experimental Group B - Subjects were presented with Châteaux, price and critic score, but no critic description of the wine.
- Experimental Group C - Subjects were presented with Châteaux, price and critic description of the wine, but not the critic score.
- Control Group D - Subjects were only presented with Châteaux and the price. This is the control group.

In order to make statistical comparisons, questionnaires were distributed to four retail stores in the greater Washington D.C. and Baltimore area. A total of 240 responses were administered, resulting in 60 responses per experimental group. The stack of questionnaires provided to each store were ordered by group and inconspicuously arranged and labelled (A, B, C, D, A, B, C, D, .. , A, B, C, D) to ensure random subject assignment. Researchers also knew which store questionnaires come from, allowing for additional comparisons as to store type.

Besides demographic information, each questionnaire asked the likelihood of buying two differently priced bottles of wine (no tasting was involved). This allowed testing of whether likelihood to buy was sensitive to price. Subjects were presented with the questionnaire at the cashier during checkout. It was expected that some would choose not to do the survey, but everyone was asked until 60 were completed at each store. Certain demographics were not targeted, e.g., gender, age, as assignment was random and sequential.

As each questionnaire contained different amounts of data, the following legend was used to facilitate an analysis and comparison of the data for groups:

- P – Price
- S – Score
- D – Description

For each questionnaire, châteaux names were provided for each of the two wines. However, for questionnaire A, the châteaux, price, score and description were provided. For questionnaire D, only the châteaux and price were provided. The groups are now referred to as:

- PSD – price/châteaux, score and description (Group A)
- PS – price/châteaux and score (Group B)
- PD – price/châteaux and description (Group C)
- P – price/châteaux (Group D)

To control for possible confounding variables, the following techniques were used:

- Budget - Certain subjects will have more discretionary income than others. Telling everyone that they have a wine budget of \$70 for this purchase helped control for this variable.
- Appellations – This was controlled by only using one appellation (Bordeaux).
- Critic – All scores came from one critic source (The Wine Advocate).
- Châteaux name – The name of the two Chateaus were fictional. Real Châteaux names could bias answers if a subject knew of or had tasted these wines.

An example questionnaire is shown in Figure 1 for Group PS (Group B). This questionnaire contains information about several Bordeaux wines. First, assume that you like Bordeaux. Second, assume you have a budget for this purchase of \$70. Please indicate how likely you are to purchase a 750ml bottle, based on the information provided. You are not obliged to purchase anything by participating in this survey.

Wine 1: Chateau Segay 2012  
 Price (\$ per 750ml): \$20  
 Wine Advocate (Robert Parker) Score: 90 points

UNLIKELY TO PURCHASE                      NEUTRAL                      VERY-LIKELY TO PURCHASE

1-----2-----3-----4-----5-----6-----7

Wine 2: Chateau Chelios 2012  
 Price (\$ per 750ml): \$50  
 Wine Advocate (Robert Parker) Score: 95 points

UNLIKELY TO PURCHASE                      NEUTRAL                      VERY-LIKELY TO PURCHASE

1-----2-----3-----4-----5-----6-----7

	Strongly Agree	Neutral					Strongly Disagree
	1	2	3	4	5	6	7
Wine is an important part of my life	1	2	3	4	5	6	7
I have a strong interest in wine	1	2	3	4	5	6	7
I am a member of a wine club or wine-related organisation	YES	NO					

Please indicate your **gender**  Male  Female

Please indicate your **age group**  21 – 30  31 – 40  41 – 50  51 – 60  60+

Figure 1. Example of questionnaire for Group PS.

The wines selected for the questionnaire were two Bordeaux wines from the 2012 vintage. To eliminate bias during the test, the châteaux names were changed to fictional ones.

- Wine #1 – 2012 Chateau Joanin Becot, Cotes de Castillon (\$20)
- Wine #1 pseudonym in the questionnaire – 2012 Chateau Segay
- Wine #2 – 2012 Chateau D’Isaan, Margaux (\$50)
- Wine #2 pseudonym in the questionnaire – 2012 Chateau Chelios

Other factors involved were:

- Wine Score – rated by Robert Parker for the 2012 vintage; sourced from his website.
- Wine description – exact wine review from Robert Parker sourced from his website. The reviews were only modified to remove the property name, winemakers/consultants and components of the final assemblage to remove any potential bias.
- Price – the price is the average bottle price in the USA that was sourced from www.wine-searcher.com. This site collects data from wine retailers all over the world.

Based on the factors above, the wines for the questionnaire were chosen:

- Wine Score – a difference of 5 points between the wines was assumed sufficient to clearly distinguish the ratings. The Wine Advocate’s rating system says that scores from 90 – 95, “are an outstanding wine of exceptional complexity and character. In short, these are terrific wines” [10]. While both of these wines are in the outstanding category, they fall at opposite ends of the exceptional spectrum. Parker rated the \$20 bottle as 90 while the \$50 bottle was rated 95.
- Wine Description – Both of the wine descriptions have the classic Parker vocabulary. The wines are described as having plump fruit, intense spices, inky/opaque color and being flashy/opulent. Also, there is mention of the wines being “over achievers” or being underrated.
- Price – Bordeaux can be painfully expensive, so price was a major factor in determining which wines to select. Once above 95 points, the price was too high to use for the test. In fact, we were surprised to find a wine-rated 95 points for \$50 because other wines with the same rating score were much more expensive. By having a price difference of \$30 between the two wines, the wines were clearly separated from each other.

### A. Hypotheses

The dependent variable investigated for each experimental group was likelihood to buy the stipulated wine (on a scale of 1 to 7). The expectation regarding the mean dependent variable responses for the groups is given by:

$$\mu_{PSD} > \mu_{PS} > \mu_{PD} > \mu_P$$

We expected that consumers, when given all relevant information, will depend heavily on that information, especially the critic score.

After the questionnaires were completed, the likelihoods to buy for each group was compiled and the mean responses for each group compared by the following tests:

- Single Analysis of Variance (ANOVA) test (F-Test) – performed to determine if there was a statistical equivalence of experimental group means, using a critical value of 0.05. If the means were equivalent, no further testing was necessary. If the means were different, further testing was required.
- Tukey-Kramer test – The Single ANOVA test above indicates if at least one on the means is different, but it does not provide the answer as to which means are different. Additionally, some of the means could be

statistically equal and others could be statistically greater/less than each other. The test for this situation is the Tukey-Kramer multiple comparison procedure, also at the 0.05 level of significance. With four control and experimental groups, there are six combinations tested for equivalence.

$$\begin{aligned} \mu_{PSD} &= \mu_{PS} & \mu_{PSD} &= \mu_{PD} \\ \mu_{PSD} &= \mu_P & \mu_{PS} &= \mu_{PD} \\ \mu_{PS} &= \mu_P & \mu_{PD} &= \mu_P \end{aligned}$$

For the Tukey-Kramer test, the absolute differences in means for each pair were computed and then compared to a critical range, computed by:

$$w_{ij} = Q_{0.05,(k,N-k)} \sqrt{\left(\frac{1}{n_i} + \frac{1}{n_j}\right)MSW/2} \quad (1)$$

where  $Q$  is the 0.05 critical value of the studentized range distribution,  $k$  is the number of groups (in our case 4),  $n_i, n_j$  the number of observations in the subpopulations  $i$  and  $j$  associated with each mean,  $N$  is the number of total observations ( $n_i + n_j$ ) and mean-square-within (MSW) is the value from the single factor ANOVA [13].

**B. Statistical Test Hypotheses**

For the tests conducted, all hypotheses were accepted or rejected based on critical value of 0.05. For all tests performed for the \$20 and \$50 bottle, the hypotheses are listed below:

Single ANOVA Test (Means across groups)  
 $H_0: \mu_{PSD} = \mu_{PS} = \mu_{PD} = \mu_P$   
 $H_a: \text{At least one of the means is unequal.}$

Tukey-Kramer Test (Pairwise means across groups)  
 $H_0: \text{two means are equal.}$   
 $H_a: \text{two means are unequal.}$

T – Test (Means between two sub-demographics)  
 $H_0 - \text{two demographic pairs are equal.}$   
 $H_a - \text{two 2 demographic pairs are unequal.}$

**IV. RESULTS**

**A. Single ANOVA Test**

All the raw data was compiled, and a Single ANOVA Test was performed to determine the mean responses for each of the two bottles of wine. We also conducted ANOVA on different demographic groups. The summary table is listed in Table 1 along with the number of subjects in each demographic.

All colors highlighted orange indicate that there is no statistical difference in the means across the four experimental groups. For the \$20 bottle, all but one category had no statistical differences in the means, thus, we accepted  $H_0$ . Although some of the means appear to be different, the critical value of 0.05 was used to determine if statistically significant differences existed. Figure 2 is an example of the

Single ANOVA Test performed on the category “Wine Club” for the \$20 bottle.

TABLE 1. SINGLE ANOVA TABLE

CATEGORY (n)	\$20				\$50			
	PSD	PS	PD	P	PSD	PS	PD	P
ALL (240)	4.3	4.8	4.1	4.3	3.5	3.4	2.0	1.8
AGE > 40 (79)	4.4	5.0	3.6	3.7	4.4	3.6	2.0	1.4
AGE ≤ 40 (161)	4.3	4.6	4.4	4.6	3.1	3.3	2.1	2.0
WINE CLUB MEMBER (64)	4.8	5.7	5.0	4.6	4.8	4.9	2.6	2.3
NOT CLUB MEMBER (176)	4.0	4.5	3.8	4.2	2.7	3.0	1.8	1.7
BOTIQUE STORE (120)	4.3	4.8	4.4	4.6	3.3	3.0	2.0	1.9
LARGE STORE (120)	4.4	4.7	4.0	4.0	3.8	3.8	2.0	1.7
DC (120)	3.9	4.3	4.4	4.63	3.4	3.0	2.3	2.3
MD (120)	4.8	5.1	4.0	4.0	3.7	3.8	1.7	1.2
MALE (140)	4.4	5.2	4.0	4.4	4.0	3.9	1.7	1.9
FEMALE (100)	4.2	4.1	4.3	4.1	2.9	2.8	2.4	1.5
Legend for ANOVA Results								
No Statistical Difference in Means					At Least One Mean Statistically Different			

No difference in the means concludes that a buying decision was indifferent between those presented with all information (Group PSD) and those presented with the least information (Group P). This indicates that perhaps wine critic scores and descriptions are not that relevant for inexpensive Bordeaux. However, a difference was identified for the “Maryland” category. The differences in this category were further tested with the Tukey-Kramer test.

Single ANOVA “Wine Club” (\$20 bottle)

ANOVA: Single Factor

Summary				
Groups	Sample size	Sum	Mean	Variance
PSD	24	116.	4.83333	4.23188
PS	13	74.	5.69231	1.73077
PD	16	81.	5.0625	4.4625
P	11	51.	4.63636	4.25455

ANOVA						
Source of Variation	SS	df	MS	F	p-level	F crit
Between Groups	8.35198	3	2.78399	0.73396	0.535840232	2.75808
Within Groups	227.58552	60	3.79309			
Total	235.9375	63				

Figure 2. Single ANOVA “Wine Club” (\$20 bottle)

In contrast, for the \$50 bottle, there was a difference in means for all categories except for “Washington DC”. To summarize, the Single ANOVA Test only indicated that there is a difference in the means. It does not indicate which of the means are equal, less than or greater than each other. The “Maryland” category for the \$20 bottle, and all

categories (except for “Washington D.C.”) were further investigated by way of the Tukey-Kramer Test.

**B. Tukey-Kramer Test Results**

For all differences in means indicated in the Single ANOVA Test, the Tukey-Kramer test was used to investigate the differences. The Single ANOVA only mentions if a difference exists, and the Tukey-Kramer will statistically show which ones are different. For this test, groups are compared to each other (in pairs) to see if the differences were statistically different. A total of 4 groups results in 6 different comparisons. The test was performed for all data for the \$50 bottle (only the \$20 bottle for Maryland), and was also performed categorically. Figure 3 is the result for the \$50 bottle for the category “ALL”:

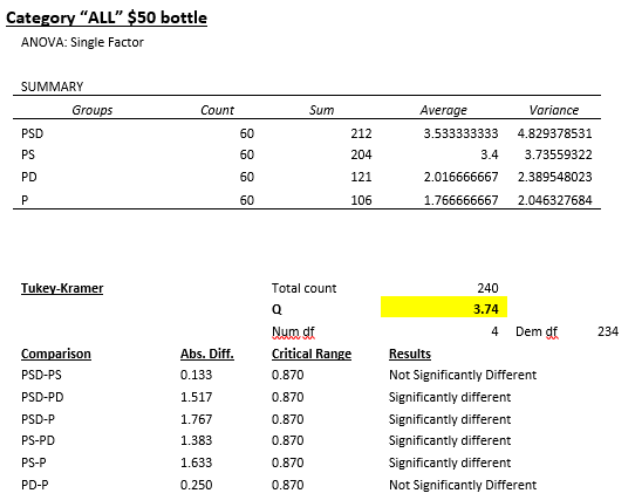


Figure 3. Tukey-Kramer Test “ALL” \$50 bottle

For the Tukey-Kramer test, if the absolute difference of the means was greater than the critical value,  $H_0$  is rejected, noting that the two means are different.

The results are interpreted as follows. For the comparison of “Group PSD” and “Group PS”, the absolute value of the difference of means are not higher than the critical range, thus they are not significantly different. That means that if one person was provided with the price, score and description; and the other person was only provided with the price and score, there was no difference in likelihood as to whether they would purchase the wine. This indicates that the description of the wine was less relevant because there was no difference as to the likelihood of making purchase when one group (Group PSD) had more information than another group (Group PS).

The comparison between “Group PD” and “Group P” were also not significantly different. What this means is that the likelihood of purchasing the wine was not influenced by the fact that one group had the price and the description, while the other group was only provided the price. For this comparison, having an additional piece of information (the description) would not influence the purchase decision.

The significant differences lie in the middle (the other 4 comparisons). Comparing “Group PSD” to “Group PD” shows a significant difference. This indicates that the score is a relevant factor when determining the likelihood of a purchase (ANOVA means for PSD is 3.5 and means for PD are 2.0). When looking at the means from the Single ANOVA Test for the means of “Group PSD” are 3.5 and the means for “Group PS” are 3.4, which are statistically equal. This comparison shows that having more information is a relevant factor in determining the likelihood of a purchase.

The absolute difference of the means is the highest for the comparison of “Group PSD” to “Group P” because “Group P” was only provided with the price and Chateau. The significant difference concludes that for this comparison, having all the information versus the least information influences the decision as to the likelihood of making a purchase. The wine critic score appears to be very important as well in the comparison of “Group PS” to “Group PD”. This comparison shows a significant difference; both groups are provided the price, but the likelihood in determining a purchase is statistically different (the ANOVA means for PS 3.4 and the means for PD are 2.0).

Overall, 10 categories were selected for the Tukey-Kramer test for the \$50 bottle. Maryland was also tested for the \$20 bottle because of differences noted in the Single ANOVA Test. Of the 10 categories, the results of the comparisons were the same for 7 categories, as those noted in the “All” category above, in which “Group PSD” and “Group PS” were not significantly different. Also, “Group PD” and “Group P” were not significantly different. All other comparisons were significantly different. The categories with the same results (which means were equal) are as follows:

- All (\$50 bottle)
- > 40 (\$50 bottle)
- <= 40 (\$50 bottle)
- Maryland (\$20 bottle)
- Wine Club (\$50 bottle)
- Boutique (\$50 bottle)
- Large (\$50 bottle)
- Male (\$50 bottle)

For these cases, the following is concluded:

$$\mu_{PSD} = \mu_{PS} > \mu_{PD} = \mu_P$$

Other differences were noted than those above in the categories; “No Wine Club”, “Maryland” (\$50 bottle) and “Female”. Each one will be investigated below.

1) Category “No Wine Club” (\$50 bottle)

The differences are almost exactly the same as found in the “All” category above, but there was a significant difference between “Group PSD” and “Group PS. For this comparison, to what extent is the importance of the description? By looking at the means for the two groups from the Single ANOVA, the means for “Group PS” are higher than the means for “Group PSD”. This shows that for the \$50 bottle, which is a high price point; those not belonging to a wine club are going right to the score. For all other significant differences, price is a major factor. For

comparison of “Group PD” and “Group P”, there is no significant difference. Here, we conclude:

$$\mu_{PS} > \mu_{PSD} > \mu_{PD} = \mu_P$$

2) Category “Maryland” (\$50 bottle)

The findings here are mostly consistent with other tables. For comparison of “Group PSD” and Group “PS”, there are no differences, highlighting the fact that the description may not be that important. However, there is a significant difference between the comparison of “Group PD” and “Group P”. The Single ANOVA means are higher for PD than P, but this shows that for a \$50 bottle, description could be a factor in determining the likelihood of a purchase. It is important to note that for the category “Maryland” for the \$20, the same comparison of “Group PD” and “Group P” showed no significant difference. There is a price difference of \$30 and this finding shows that as the bottle price increases, subjects in Maryland were curious about the description. Here, we conclude:

$$\mu_{PSD} = \mu_{PS} > \mu_{PD} > \mu_P$$

3) Category “Female” (\$50 bottle)

The results from this category were very interesting because this is the only category where there was a significant difference between “Group PSD” and “Group PD”. Although the Single ANOVA means are higher for PSD (2.9) than PD (2.4), those means are far closer to each other than any other category. This shows as the price of the bottle becomes more expensive, women pay attention to the wine description and this can be a factor determining the likelihood of a purchase. This is also true when comparing “Group PS” to “Group PD”, as this was the only comparison to be not significantly different amongst the categories. This shows the importance of the wine description to women. This is also true for the “Group PD” and “Group P” comparison. Here, we conclude:

$$\begin{aligned} \mu_{PSD} &= \mu_{PS} & \mu_{PS} &= \mu_{PD} \\ \mu_{PSD} &> \mu_{PD} & \mu_{PD} &> \mu_P. \end{aligned}$$

C. Equal Variance Pairwise Two-tailed T-Test

The results for the pairwise T-Test are presented in Table 2. Presented with the same information, was the likelihood of purchase different for the \$20 and \$50 bottle? The statistical difference is based on a 0.05 significance.

In most cases, when presented the same information, price was a factor in determining the likelihood of a purchase. However, there were some instances in which price was not a factor. Most of the instances occurred when the subjects were presented with the most information:

- Age > 40 - to determine the likelihood of a purchase, age was important.
- Wine Club Member – this was the only category that also had a difference when only the score was presented. Maybe these subjects willing to make a high dollar purchase with only the score.
- Large Store – in a large setting, the consumer has many more choices. Perhaps they depend more on critics when there are so many choices.

- Male - little speculation can be offered other than perhaps status of a more expensive highly rated wine?

TABLE 2. UNEQUAL VARIANCE PAIRWISE T-TEST

CATEGORY	Same Subject difference for \$20 and \$50 Bottle?			
	PSD	PS	PD	P
ALL	0.0113	0.0000	0.0000	0.0000
AGE > 40	0.5000	0.0006	0.0000	0.0001
AGE <= 40	0.0015	0.0000	0.0000	0.0000
WINE CLUB MEMBER	0.5000	0.1050	0.0002	0.0062
NO CLUB MEMBERSHIP	0.0008	0.0000	0.0000	0.0000
BOTIQUE STORE	0.0078	0.0000	0.0000	0.0000
LARGE STORE	0.1457	0.0096	0.0000	0.0000
DC	0.1583	0.0000	0.0000	0.0000
MD	0.0134	0.0012	0.0000	0.0000
MALE	0.2207	0.0004	0.0000	0.0000
FEMALE	0.0022	0.0001	0.0000	0.0000
Statistical Difference Based on T-Test p-Value	No Difference		Difference	

- DC – a small but very affluent city. Also, an international city. Perhaps people in DC are willing to spend more money on wine. City-Data [14] shows that the median household income as of 2013 was about \$25,000 higher in Washington DC versus Baltimore.

D. Unequal Variance, Two-Tailed T-Test

Demographic category differences were investigated within the same experimental group. Demographic groups were compared for both \$20 and \$50 bottles. Statistical differences in Table 3 are based on 0.05 significance. When presented with all relevant information for the \$20 bottle, there were only a few instances of demographic differences within the same groups. Most of the differences were between “wine-club versus no-wine-club members” as the information provided decreased (all the way down to price). For each difference (PS, PD and P), the means from the ANOVA table were higher for wine club members, indicating they were more likely to make a purchase. Perhaps the wine club members subscribe to wine publications with ratings? It is also noted that a difference existed between “male versus female” for “Group PS”. The likelihood of a purchase, the means were higher for males.

For the \$50 bottle, there was a difference as those over 40 were more likely to make a purchase. Just like the \$20 bottle, differences were also noted for “wine club versus no wine club” for the \$50 bottle. Differences were only noted in “Group PSD and PS” because as price is much higher for a \$50 bottle, description and price alone were not enough to

make a difference. To note the previous T-Test again, a difference in “Group P” was noted for “DC versus MD”, in which the means were higher for DC. This sheds light that perhaps DC is more affluent than MD and those subjects are willing to make a purchase when only provided the price.

TABLE 3. UNEQUAL VARIANCE, TWO-TAILED T-TEST

CATEGORY	Category Difference in Responses?							
	\$20				\$50			
	PSD	PS	PD	P	PSD	PS	PD	P
AGE > 40	0.826	0.421	0.132	0.156	0.037	0.658	0.8161	0.0831
AGE ≤ 40								
WINE CLUB MEMBER								
NO CLUB MEMBERSHIP	0.124	0.009	0.042	0.002	0.000	0.009	0.1531	0.3920
BOTIQUE STORE								
LARGE STORE	0.857	0.824	0.350	0.244	0.415	0.082	0.8046	0.5925
DC								
MD	0.103	0.085	0.350	0.244	0.561	0.082	0.1581	0.0038
MALE								
FEMALE	0.791	0.015	0.510	0.607	0.056	0.028	0.0951	0.1946
Statistical Difference Based on T-Test p-Value	No Difference				Difference			

For the \$50 bottle, there was a difference as those over 40 were more likely to make a purchase. Just like the \$20 bottle, differences were also noted for “wine club versus no wine club” for the \$50 bottle. Differences were only noted in “Group PSD and PS” because as price is much higher for a \$50 bottle, description and price alone were not enough to make a difference. To note the previous T-Test again, a difference in “Group P” was noted for “DC versus MD”, in which the means were higher for DC. This sheds light that perhaps DC is more affluent than MD and those subjects are willing to make a purchase when only provided the price.

A very interesting finding is the difference in “male versus female” for “Group PS”. This is consistent with the \$20 bottle in which the means are higher for the male to make a purchase. The Tukey-Kramer test indicates women regarded the critic description as very important. Only differences existed where the description was not provided, indicating that perhaps males only care about the score.

It is important to also note the comparison of “male versus female” for Group PSD for the \$50 bottle. Although the T-Test concluded there was no difference, the results were very close (0.056 with a critical difference value of 0.05). The means for males to purchase for Group PSD were much higher than the means for the female. Although no difference was noted, an opposite conclusion is rationale.

## V. MAJOR FINDINGS AND CONCLUSIONS

The likelihood to purchase based on wine critic information was very dependent on price. For a \$20 bottle, the critic information was not a factor in consumer decisions. Conversely, for the \$50 bottle, the scores and description, or just the score are important.

Demographics definitely play a role. When presented with score and description, the same individual is as likely to buy the \$20 bottle as the \$50 bottle if they are male, over the age of 40, belong to a wine membership club, shopping in DC, or in a large store. So, gender, age, membership, store location, and type store matter.

When comparing different individuals demographically, wine critic score is more influential for males versus females, and for club membership/no membership, for both the \$20 and \$50 bottle. In contrast, the wine description matters most to females. The demographic that was most likely to buy wine is club membership, followed by gender, followed by age and location. The type store demographic showed no differences in purchase likelihood.

### A. Research Limitations

Although random assignment of subjects to groups was used, the four stores were selected using convenience sampling. Existing relationships in the region were exploited to recruit stores for this research.

The questionnaire only focused on one wine region (Bordeaux) in the wine-producing world. Even though the questionnaire asked subjects to assume they like Bordeaux, there could be some bias amongst some subjects if they do not happen to like Bordeaux. Additionally, subjects were asked to assume they had \$70 to spend. Some may have had an aversion to buying a \$50 bottle of wine.

Another limitation of the questionnaire is that it only had two different bottle prices. Perhaps, more questions could have been asked with different price points.

Lastly, the test results only pertain to a certain geographical area in the U.S. and may not be representative of other countries/regions.

### B. Implications for Producers, Distributors and Retailers

#### Wine Producers

According to this research, consumers pay little attention to wine critic scores/information for inexpensive bottles. There are of course other factors -- some producers may be great at making wine, but not so great at marketing. Every wine needs a label; in fact, some anecdotal evidence suggests that for inexpensive imported wines, the label is more important than the wine quality. The literature review concluded customers prefer easy to understand labels.

For those producers in the business of selling expensive wines, this research concludes that wine information is important to consumers likely to buy. As the price rises, those with money and the more educated wine-consumer are likely to seek out additional information about the wine. Strategic positioning is very important as well because this research indicates that those in a very affluent area might make a purchase with little or no information at all. Either way, it would make sense for wine producers to get their wines in front of critics because positive attention never hurts. Of course, the risk is a low rating.

#### Wine Retailers

For inexpensive wines, retailers should include shelf information to assist consumers. A wine of the month section



could be created for those wines a retailer would like to highlight. This research also indicates that in large retail stores, drawing attention to particular wine is good.

It is common for retailers to have a few customers that do not mind spending a lot and if they prefer high scoring wines, retailers should be looking at all wine critic publications and identifying these wines for their customers.

This research indicates that wine club members are more involved customers and are maybe willing to spend a bit more for bottles of wine. Retailers could take advantage of this and create their own wine club for their customers. They could do this with wines of the month and can focus a small part of their business on highly involved wine customers. Certain wines could be selected for this wine group and the purpose is to draw more people to the group. As some consumers are less involved but have interest in wine, belonging to club could potentially make them more involved, and perhaps could lead them into possibly raising their budget for wine over time.

#### Wine Distributors

Just like wine retailers, distributors need to look at wine critic scores and publications to be aware of what is happening. They also need to be in constant contact with their customers, so they can obtain wines that certain people may be looking for. For expensive wines, it is important for wine sales representatives to transmit wine critic scores to their customers because a retailer might not always have time to be constantly searching for this information.

For inexpensive wines, distributors need to pay attention to market trends and know what their customers are looking for. Distributors need to buy what they know they can sell, but they have to know the trends in the market place as well. A distributor might determine to bring in a new product because the wine label is great, or the wine has a good story. It is common for distributors to have a marketing department, which can prepare shelf information, display signs and anything a retailer might need to help sell the wine. If a distributor does not have this, they should investigate the possibility of doing so.

#### C. Future Research

Analyzing this dataset using more powerful statistical methods could be useful, such as Artificial Neural Network (ANN) approaches or Principal Component Analysis (PCA), to mine more insights.

It would also make sense to replicate this research using more granular pricing to obtain more information on how price influences purchase decisions. For instance -- to fully investigate this question: At what price point would scores and descriptions not be as relevant? Additionally, there are many major metropolitan markets/cities in the USA and if a similar experiment included more cities, more insights could be gained. Likewise, the research could be extended to study international locations and customers.

## VI. ACKNOWLEDGEMENTS

The authors would like to thank Jeffrey Snow, who independently recruited the four stores, conducted the field research, and provided the data. He also provided us his extensive knowledge and expertise regarding wine markets and consumer behavior.

## REFERENCES

- [1] OIV. State of the vitiviniculture world market. [http://www.oiv.int/oiv/info/en\\_vins\\_effervescents\\_OIV\\_2014](http://www.oiv.int/oiv/info/en_vins_effervescents_OIV_2014). Web.
- [2] J. Almenberg, et al., "Do More Expensive Wines Taste Better? Evidence from a Large Sample of Blind Tastings", American Association of Wine Economists, Working Paper #16, Print. 2008.
- [3] S. Geringer, D. Patterson, and L. Forsythe, "When Generation Y Buys European Wine: A Consumer Decision-Making Model", Academy Of Marketing Studies Journal, vol. 18, no. 1, pp. 167-186, 2014, Business Source Complete, EBSCOhost, viewed 20 January 2015.
- [4] Wines & Vines., 'Designers Take On Label Trends'. N.p., 2015. Web. 9 February. 2015
- [5] B. Scheibehenne, R. Greifeneder, and P. Todd, "Can there ever be Too Many Options? A Meta-Analytic Review of Choice" 2010, Business Source Complete, EBSCOhost, viewed 08 February 2015.
- [6] L. Francis, L. L. Lockshin, S. Mueller, and P. Osidacz, "How does shelf information influence consumers' wine choices?" Australian and New Zealand Wine Industry Journal Aust. N.Z. Wine Ind. J. vol. 24 no. 3, pp. 50-56, 2009. Print.
- [7] R. E. Quandt, "On Wine Bullshit: Some New Software." Journal of Wine Economics vol. 2, no. 2, pp. 129-135, 2007. Print.
- [8] R. L. Weil, "Analysis of Reserve and Regular Bottlings: Why Pay for a Difference Only the Critics Claim to Notice?" Chance vol. 18, no. 3, pp. 9-15, 2005. Print.
- [9] R. L. Weil, "Debunking Critics' Wine Words: Can Economists Distinguish the Smell of Asphalt from the Taste of Cherries?" Journal of Wine Economics vol. 2, no. 2, pp. 136-144, 2007. Print.
- [10] R. Parker, "The Wine Advocate Rating System." Erobotparker.com. <https://www.erobotparker.com/info/wstandards.asp>.
- [11] H. Ali, S. Lecocq, and M. Visser, "The Impact of Gurus: Parker Grades and En Primeur Wine Prices", Economic Journal, vol. 118, no. 529, pp. F158-F173, 2008, Business Source Complete, EBSCOhost, viewed 18 January 2015.
- [12] J. Hilger, G. Rafert, and S. Villas-Boas, "Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market", Review Of Economics & Statistics, vol. 93, no. 4, pp. 1289-1296, 2011, Business Source Complete, EBSCOhost, viewed 1 February 2015.
- [13] J. Hsu, Multiple Comparisons: Theory And Methods. Chapman & Hall/CRC, 1996. Print.
- [14] City-data.com., 'City-Data.Com - Stats About All US Cities - Real Estate, Relocation Info, Crime, House Prices, Cost Of Living, Races, Home Value Estimator, Recent Sales, Income, Photos, Schools, Maps, Weather, Neighborhoods, And More'. N.p., 2015. Web. 8 September. 2015.