HEDONIC REGRESSION ANALYSIS COMPARING NAPA TO PASO ROBLES RED WINE PRICES WITH AN EMPHASIS ON SCREW CAP CLOSURE METHODS

Presented to the

Faculty of the Agribusiness Department

California Polytechnic State University

In Partial Fulfillment

Of the Requirements for the Degree

Bachelor of Science

By

Todd Patrick Griffin

August 2009

Approval Page

TITLE: Hedonic Regression Analysis Comparing Napa to Paso Robles Red Wine Prices with an Emphasis on Screw Cap Closure Methods

AUTHOR: Todd P. Griffin

DATE SUBMITTED: August 2009

Dr. Jenni James
Senior Project Advisor

Abstract

The purpose of the senior project was to test and estimate quantitative and qualitative relationships between wine bottle characteristics in two different wine consuming regions of California. A hedonic regression analysis was performed to evaluate the significance on price fluctuations based on varietals of wine that contain labels of origin from two American Viticultural Areas (AVA's) in California. A hedonic regression analysis divided the price of the product into variables that composed the product. The division of variables allowed the researcher to examine and observe whether they are of high or low significance to the estimation of price. The senior project also performed a complex evaluation of regional, marketing, and bottling parameters according to price. In this study, screw cap bottling parameters were evaluated because of the latest trends in market innovation for the wine industry.

Table of Contents

Chapter **Page**

I. INTRODUCTION……. …. 1

Problem Statement……2

Hypothesis……. 2

Objectives of the Study……..2

Justification……2

II. REVIEW OF THE LITERATURE……. 3

Branding AVA's……3

Napa AVA……. 3

Paso Robles AVA…….. 4

Hedonic Regression Analysis…… 4

Quality Ratings…….. 6

Hedonic Study on Honey……..... 7

Screw Caps……7

Perceptions of Screw Caps…….……8

III. METHODOLOGY……. 10

Procedures for Data Collection……..10

Procedures for Data Analysis……12

Assumptions…….. … 14

Limitations…….… 14

IV. DEVELOPMENT OF THE STUDY…….. … 15

Data Collection Problems…….. 15

Analysis……. 15

V. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS……. 24

Summary…….....24

Conclusion……. 24

Recommendations……..25

References Cited……..27

APPENDIX……. 29

LIST OF TABLES

Table **Page**

1 Data Collection Sheet 1-20…….……. 16

2 First Regression…….…….. 17

3 Second Regression…….…….. 18

4 Third Regression……..... 19

5 Progressive Model of Coefficients…….....….20

LIST OF FIGURES

Figure **Page**

1 Region and Varietal Number of Observations……21

2 Average Price by Varietal and Region USD…….....21

3 Market Location and Varietal Average Price USD……. 22

4 Label Attribute: Number of Observations……..... 23

Chapter 1

INTRODUCTION

Wine is a drink, a life, and a culture made from fermenting grapes that are grown in many different climates of the world. In California, grapes are grown across the entire state. The climates differ from the warm, dry, desert like southern regions, to the Mediterranean climates of the central coast and the drier, humid growing conditions of the northern central valley. In the northern, central, and southern most regions of California, the land provides necessary amounts of water for farmers' using professional harvesting techniques for the growth of high quality grapes that develop into high quality wines.

However, quality wines come at a price. Many prices are observed in the grape growing or production process, others in the distribution and wholesale process, and many prices are observed in the grocery store through the retail process. Consumers can be seen observing prices on the shelves of grocery stores and purchasing those bottles of wine. But what do these prices mean? Surely the production and distribution costs are added into each bottle but how are they calculated with one bottle being much more expensive than another? Does the quality of a professional taster, a region of origin grown, a nice looking label, or special closure method add that much more value to the price on the shelf?

Problem Statement:

Does the labeling of origin "Napa" on a Cabernet Sauvignon, Merlot, Pinot Noir, or Zinfandel wine bottle contain a price premium > $6 when comparing the region to Paso Robles?

Hypothesis:

No, labeling the origin "Napa" has no effect on the price of the Cabernet Sauvignon, Merlot, Pinot Noir, or Zinfandel when comparing to the region of Paso Robles.

Objectives:

1) To evaluate the relationship between prices for Napa and Paso Robles wine regions.
2) To assess qualitative characteristics according to label attributes on wine bottles.
3) To explain the variables affecting the price on the retail shelf.

Justification:

The purpose of this senior project is to examine whether there exists a strong correlation between the prices of wine at the retail level based on characteristics of that wine bottle. Organizations and viticulture related businesses with research and development teams, marketing teams, business owners, and the American Association of Wine Economists (AAWE) would find this information useful as it will be presented to the AAWE in Reims, France. The project can be redone in the future to further provide whether correlations in price and qualitative characteristics exist between regions in California.

Chapter 2

LITERATURE REVIEW

Branding AVA's

American Viticultural Areas (AVA's) are designated wine-grape growing regions within the United States. Region specific wines have gained popularity by claiming one AVA from another. Farquar (1989) suggests that the region of origin can be used as a brand to the product. In this case, "the product is something that offers a functional benefit whereas a brand (name, symbol, design, or mark) enhances the value of the product surpassing the functional value" (Farquar, 1989). This form of branding would allow consumers to try new varietals and labels within a desired region

Orth, Wolf, and Dodd (2005) suggest that wine brands such as Napa Valley are emerging but are fighting to gain a larger portion of the market because of larger regions and stronger producing competitors. The larger regions can be those labeled "West Coast" instead of a more specific region like "Napa Valley." Walker (2003) suggested that wine marketing executives have begun to utilize the region of origin as equity for political battles amongst the rules of labeling wine from a specific AVA.

Napa AVA

One AVA in particular is located just about 50 miles east of San Francisco and has long been dubbed the wine capital of California. The Napa AVA has five million tourists a year and ranks above Yosemite for popular destinations in California (Farnham, 2003). According to the

Wine Institute (2001), Napa contains approximately 225,000 viticulture acres and hundreds of wineries. Napa Valley has been long regarded as one of the best places in the world for producing high quality wines.

Paso Robles AVA

High quality wines have also come from other great wine producing Mediterranean climates in California. The Paso Robles AVA is halfway between San Francisco and Los Angeles and has had an increased popularity for its wines in recent years. 60 vineyards attract a million tourists per year (Farnham 2003). The Wine Institute (2001) shows Paso Robles more than doubles the acreage of Napa with approximately 666,000 acres of wine producing land. Its climate has a long growing season with warm days and cool nights that is ideal for growing red wine varietals.

Hedonic Regression Analysis

In efforts to analyze a regression analysis on red wine varietals, examples are needed to support this study. In recent years, hedonic analysis has been used to examine traditional factors affecting wine prices. Traditional factors that affect price can include the quality of the wine, which is determined by a professional wine taster. One professional wine taster is Robert Parker (2009) who is considered one of the most famous and influential critics of French and other international wines in the world. Parker is famous for awarding scores to wines that fit within many detailed, desirable and preferable parameters to the palate. Since the parameters are very subjective to the individual, parameter variables can be difficult in evaluating the regression analysis for qualitative indicators on price.

According to another study, Brooks (2001) analyzed wine characteristics including: country-of-origin, blind tasted quality, vintage, year of price observation, year that the wine is

best consumed and grape variety. Such variables were used as dummy variables in calculating the significance to the prices of international wines. Each variable was paired with a "1" if the quality characteristic was present and "0" if the quality characteristic was absent from each sample. Dummy variables (2008) are often used in statistical regressions to indicate the absence or presence of some qualitative effect that may be expected to shift the outcome, in this case, price.

Analysis showed possible biases in the data. By controlling the variety, vintage, and quality rating, analysis indicated one variable of high significance to price. Studies showed "country-of-origin" significantly affected prices of samples. Two countries in the analysis showed significant results. By indicating their country-of-origin, "France" and/or "Italy", the French and Italian wines benefited tremendously from the study (Brooks 92).

Perrouty, d'Hauteville and Lockshin also explored the idea that region of origin contained equity on the bottle of wine. According to Perrouty et al (2006), when doing an international survey of approximately 1,162 European wine purchasers, the region of origin was significant to the consumer purchasing decisions. Even more interesting, region of origin was more important to consumers who felt they were "more expert" than the consumers who felt they were novices or "less expert."

Gunter and Anderson also conducted a hedonic study of wine prices from Australia and New Zealand. The variables of the study were region, varietal, quality, vineyard reputation, and wine ratings. Gunter and Anderson (2003) noticed that when examining the dummy variables for region of origin, the region became increasingly significant overtime. In contrast, the premium that consumers were willing to pay for higher-rated wines trended downwards since the 1990's (Gunter, Anderson 2003). Analysis of the study concluded that Gunter and Anderson (2003)

were able to discern that both Australian and New Zealand consumers were constant in becoming more confident in purchasing wines based on their own ability to distinguish quality of wines, therefore less inclined to use expert ratings.

Quality Ratings

Schiefer (2008) discusses variations in ratings from an expert to a non-expert consumer view. One way of rating or "award-labeling" wines has led to controversial indications for wine quality. Schiefer argues that by placing an award-label on a bottle, a low involvement consumer or novice consumer will be more likely to purchase the wine. In this case, the novice wine consumer will assume that they have the same or similar taste buds as the expert consumer.

Schiefer (2008) suggests that the expert and novice consumer tasting behavior and their approach to sensory evaluation differ substantially. The article claims that a wine's typicality for its variety, origin and vintage highly affect its expert ratings. In turn, novices merely rate a particular wine as "like" or "dislike." Because of this, quality evaluation by consumers and experts may lead to substantially different results since their opinions are based upon final outcomes for taste.

Schiefer also did another study comparing novice-to-novice consumer sensory evaluation of wines. All comparisons of novice consumer tastings were highly variable in the rating of wines. The analysis of the study showed a weak correlation between expert rating and novice ratings. Since consumers commonly have differing opinions, one might ask whether award-labeling or wine ratings can be helpful to a majority of consumers looking for wine on the retail shelf especially if a vast amount of wines do not have award labels.

Hedonic Study on Honey

In addition to wine, other retail industries have used hedonic analysis for similar studies. Unnevehr's article "Retail Premiums for Honey" shows how hedonic regression analysis is used for honey. According to Unnevehr (1996-1998) retail demand for commodities like honey is estimated by using retail scanner data. Retail scanner data is electronically calculated at the register by quantifying consumer purchases. This quantifiable data is then used in regression analysis to form equations to better understand why consumers purchase specific commodities.

The problem is that honey prices, much like wine prices, are extremely competitive. In the Retail Premiums article, information gathered from A.C. Nielsen shows that hedonic regression analysis was conducted using similar quality characteristics like the ones used in the Brooks study. These characteristics were based on container type, flavor, and form in efforts of gathering unit sales, dollar sales and average prices. These figures showed that the total retail market for honey was worth $127.9 million in 1994 to 1995. Of that $127.9 million, 28% of total dollar volume came from store labels, 64% consisted of 48 different brands, while 205 brands accounted for 8% of the total dollar volume (Unnevehr).

With information from hedonic analysis, producers are able to capture larger portions of the consumer market through product differentiation. Much like the variable "country-of-origin" that France and Italy benefitted from, prices do vary according to container, product variables, flavor, and form.

Screw Caps

To examine the purpose of hedonic analysis on Cabernet Sauvignon, Merlot, Pinot Noir, or Zinfandel, multiple wine bottle characteristics must be examined. In the case study, "Rodney Strong Winery: The Great Cork Debate," Tom Atkin discusses the winery's dilemma when cork

taint was ruining some of their bottled wine. The winery explored the concept of using screw cap sealing methods rather than the natural cork method in order to eliminate this problem. Atkin compares the negative American consumer perceptions of screw cap wines with the more accepting, positive perceptions of European and Australian consumers. The cost issues of cork versus non-cork closure methods are also discussed in the case study, and it allows the reader to understand the reasoning behind natural corking methods and non-corking methods.

Cork taint is caused by an aromatic chemical compound known as 2, 4, 6-trichloroanisole commonly referred to as TCA (Boyd). High levels of TCA in the cork of a wine bottle have lead to moldy, musty aromas in the wine, and have affected approximately 10% of wine produced worldwide (Atkin). Not only does cork taint destroy quality and pleasant characteristics of the wine, but it can also destroy positive consumer perception, and economically devastate the wine brand. Since the wine industry is consumer driven, producers of the wine have to adapt and evolve so producers won't lose business due to a tainted cork.

Screw cap sealing methods will eliminate cork taint issues, in addition to allowing consumers to reseal the bottles after opening. Cork taint issues can be eliminated with the help of data analysis from Europe and Australia that show that screw caps and synthetic closures sell better and are more cost friendly than natural cork. Each year, costs associated with cork taint exceed 10 billion dollars worldwide (Atkin). Overall, synthetic and screw cap closure provide the best solution to this financial problem in the wine industry.

Perceptions of Screw Caps

Atkin's also argues that consumer responses to screw cap wines are considerably different between the market the United States (US) and the market in Australia and New Zealand. Trends in response to alternate methods of wine closure indicate that the majority of

people are willing to try screw cap wines. However, this majority comes from Australia and New Zealand. American consumers are still unwilling to convert to new methods of innovation in the wine industry.

Consumer attitudes about wine closures affect which method the winery will use. Wine consumer attitudes can be measured in the five barriers of purchasing wine: (1) usage - for what purpose it will be used, i.e. dinner, events, parties; (2) value - the importance in quality of wine; (3) risk - likelihood wine will be an unfavorable choice; (4) tradition - custom of drinking wine; and (5) image - the way one is perceived drinking the wine. These barriers can affect the chances of US consumers purchasing screw cap wines (Atkin).

If a winery tries to market an innovation product before it has been accepted, then the winery will be increasing its risk of failure (Atkin). Failure is not an option for producers, so the winery must do research on barriers to see if US consumers will purchase their products. History shows that US consumers will not purchase a wine that they are uncomfortable drinking. Therefore, until the barriers (usage, value, risk, tradition, and image) are lifted, consumer perception of screw cap wine in America will remain negative rather than positive.

Chapter 3

METHODOLOGY

Procedures for Data Collection

The project researcher, a Cal Poly student, will gather data with two other Cal Poly researchers to evaluate the impact of region on price. The data will be collected based on the scope of the study. The scope includes 750 ml red wine bottle characteristics at the main retail grocery stores in San Luis Obispo and Redwood City all under $50. In San Luis Obispo, the researcher will collect data from Albertson's, Ralphs, Vons, and BevMo. In Redwood City, the researcher will collect data from Safeway, K & L Wine Merchants, Lucky's, and again BevMo.

Data will come from the stores that have major similarities between the central and northern California geographical distribution and demographics of these retail grocery store areas. The researcher will be able to gather sound data from the third largest wine growing region in California, San Luis Obispo, which provides $113 million in wine related tourism for San Luis Obispo County (Taranto). Redwood City has a population very similar to San Luis Obispo and its propensity to generate wine revenues is very comparable to San Luis Obispo. Both locations for data collection, San Luis Obispo and Redwood City, will be outside of the regions "Napa" and "Paso Robles" so biases will not strongly affect data results of the study. The collection of red wine data will begin on April 24th, 2009 and continue to May 1st, 2009. This will allow enough time for clear and accurate data collection from the retail grocery stores.

Characteristics are chosen based on simple features of the wine bottles. These include the red wine varietals, labeling, screw cap versus natural cork closure methods, and a detailed

evaluation of the picture on bottle. Red wine varietals that will be used are Cabernet Sauvignon, Merlot, Pinot Noir, and Zinfandel. Closure type will be evaluated as screw cap or standard corking (includes synthetic corking). Labeling will be evaluated to see if the label is in calligraphy (cursive) and the researcher will look for words on the bottle that show "estate", "reserve", "old vines", "sustainable" (on front or back), year, location (Napa or Paso Robles), "organic", and position on shelf (ex. top, middle, bottom shelf). A sample of 500 bottles will help validate any possible fluctuations in price for the red table wine labeled with these two regions. The sample size of 500 is an estimate because of possible limiting factors of location and the number of bottles on the shelves representative for Napa and Paso Robles.

The method of price collection will come from the retail price that is on the shelf. The "consumer" is out of the equation, because whoever purchases the bottles is irrelevant to bottle prices based upon characteristics for this regression. In essence, it is geared toward prices that retail grocery stores are charging for their wines. Whether or not the wines sell is not a part of the study. However, more data should be analyzed to see if those bottles sold have the desired variables to strengthen the results.

Secondary data will come from sources such as *A Multinational Study of the Diffusion of a Discontinuous Innovation*, and *The History and revival of Screw Caps* as well as *Retail Premiums for Honey Characteristics*. These sources will allow the researcher to analyze other forms of regression analysis and compare them to the analysis for this senior project. The secondary data will also enable the researcher to analyze other markets and commodities in those markets and compare them to differing regions in California.

Procedures for Data Analysis

After data is collected, it will be entered into a Microsoft Office 2008 Excel spreadsheet and regressed on a Macintosh laptop computer. With this program, a statistical regression analysis will be the tool for the senior project. In the regression selection, a binary numbering system (1, 0's), known as dummy variables will be utilized for a regression analysis. Data Input will be programmed into the regression as shown: type "1" if the bottle does have a certain characteristic or type "0" into the excel function if the bottle does not have that certain characteristic.

The statistical analysis will include the number of observations, Analysis of Variance (ANOVA), t-stat, and p-value for strong or weak correlations in values. Significant variables will include the positive and negative coefficients showing a response to the intercept or base value being regressed. For example, a P-value indicates whether there is or isn't a significant impact on the dependent variable. A P-value close to 0.1 or lower would indicate a greater than 90% Confidence Interval (CI) of explanation between independent and dependent variables. The P-value will also be significant at the 95% and 99% CI's, $P \leq 0.05$ and $P \leq 0.01$, for strong correlations in the price of the qualitative characteristic being analyzed. So, the closer the p-value is to zero, the more significant the p-value will be for the analysis.

In contrast, the R^2 value usually needs to be closer to 1 to show significant results for independent variables affecting the dependent variables. In other words, the R^2 indicates the percentage of variation of the dependent variable that has been explained by the variation in the independent variable. However, this type of data will not rely heavily on the R^2 value because of cross tabulation of data.

T-stats are just as important as the P-value because of the direct relationship with each statistic/value. A statistically significant T-stat would indicate a numerical value usually greater than 1.96 that would indicate a strong relationship between price and the qualitative characteristic being analyzed.

The spreadsheet that the bottle information is gathered on displays a blank box for store prices and qualitative characteristics being analyzed. After data is compiled into a spreadsheet, an equation must be set up to run a regression on pricing. The equation is set up so product prices are based on the sum of values of the attributes within the wine. The use of β's or attributes will explain variations in prices. The estimates β's minimize the SSE (sum of squared differences) between actual prices and estimated prices. Next the researcher can make an educated guess on the value of individual attributes by solving for estimated values where price is a function of those attributes. The equation for the analysis will look like this:

Price = ∑ value of attribute * attribute

Price = $\sum \beta_{attribute}$ * attribute

$P = \sum \beta * A$ **price is a function of attributes*

$e = p - p\hat{}$ **estimated price*

The regression will show whether prices of wine decrease or increase based on the qualitative variables used for the wine bottles. Intuitively, factors affecting the output are the dependent and independent variables. The independent variables are Cabernet Sauvignon, Merlot, Pinot Noir, and Zinfandel, screw cap closure, calligraphy, estate, reserve, old vines,

sustainable (on front or back), year, location, organic, and position on shelf. The dependent variable will be the "price" (P) of each type of wine because P is affected by independent characteristics of the wine. All data will be directly evaluated against the base variables of: Cabernet Sauvignon (varietal), Paso Robles (region), and San Luis Obispo (data collection). The regression is the best method for evaluating this project because the regression will give results on unbiased factual variables from the equation.

Assumptions

It is assumed that the retail grocery stores in San Luis Obispo and Redwood City will be willing to assist with the senior project. This data is assumed to be factual for the study to work properly. It is assumed that variables entered into Excel will be chosen according to bottle characteristics for the regression to work properly. Variables such as type of closure method, varietal, name brand, and labeling characteristics are assumed to be of major significance for price fluctuations in the wines chosen for the senior project.

Limitations

While it is assumed that the stores will allow for data collection, systematic data is usually hard to attain from retail operations unless prior consent is authorized. If the stores are not willing to allow for data collection, then the internet will be used for price collection. Web sites will hold the quality and quantity dynamics for each bottle. However, Quality can be a matter of subjectivity which has to be left out of the regression because it will throw off the data. Subjective attributes include labeling which has many other variables like colors (hues, contrasts) and features (animal picture) that can be difficult to use in an unbiased regression.

Chapter IV

DEVELOPMENT OF THE STUDY

Data Collection Problems

There was one major difficulty with the data collection. The senior project put an emphasis on price fluctuations, particularly on screw caps from Napa Valley and Paso Robles. However, after collecting 569 bottles of wine, the amount of bottles from Napa and Paso Robles that contained screw caps was minimal. In fact, when analyzing the data, only 10 bottles of red wine that contained screw cap closure methods were observed. The low sample size of screw caps left one to assume that screw caps are not a very popular method for wine closures in California. Therefore, with a very low number of screw cap bottles, sample data was not significant for the study on price.

Also, gaining statistical data from grocery was very difficult. Some stores would allow for data collection while other stores would not allow for data collection. Even though the store was told that the project was for educational purposes, some store managers did not feel comfortable with the primary data collection process.

Analysis

This is an example of a chart used to collect data from both San Luis Obispo and Redwood City. The Brand is the Label on the 175 mL bottle of red wine. The dependent variables on the chart are: Price-Regular and Price-Discount. The independent variables are: Location, Varietal, Year, Screw Cap, Animal, Calligraphy, Old Vines, Reserve, Estate, Sustainability (front and back), Organic, and Shelf Location.

Table 4-1: Data Collection Sheet 1-20

#	Brand	Price - Regular	Price - Discounted	Location	Varietal	Year	Screw Cap	Animal	Caligraphy	"Old Vines"	"Reserve"	"Estate"	Sustainability on Front	Sustainability on Back	Organic Seal	# of shelves from bottom
1	Peachy Canyon	11.99	8.98	P	Z	06	0	0	0	0	0	0	0	0	0	Top 4/5
2	Robert Hall	16.99	14.48	P	M	06	0	0	1	0	0	0	0	0	0	4/5
3	Wild Horse	21.99	15.98	P	M	05	0	1	0	0	0	0	0	0	0	4/5
4	Turn Bull	37.99	26.59	N	M	01	0	1	0	0	0	1	0	0	0	4/5
5	Cameron Hughes	16.99	13.98	N	C	05	0	0	0	0	0	0	0	0	0	4/5
6	Folie a Deux	27.99	19.59	N	C	06	0	0	1	0	0	0	0	0	0	3/5
7	Merryvale Starmont	27.99	19.59	N	C	05	0	0	0	0	0	0	0	0	0	3/5
8	Nine Points	28.54	19.98	N	C	06	0	1	0	0	0	0	0	0	0	3/5
9	J Lohr Carols Vineyard	29.99	20.99	N	C	05	0	0	1	0	0	0	0	0	0	3/5
10	Napa Cellars	24.99	17.49	N	C	05	0	0	0	0	0	0	0	0	0	3/5
11	Louis M Martini	23.99	16.79	N	C	02	0	1	1	0	0	0	0	0	0	3/5
12	Raymond R Collection	18.99	15.98	N	C	05	0	0	1	0	0	0	0	0	0	3/5
13	J Lohr Seven Oaks	17.99	12.98	P	C	07	0	0	0	0	0	0	0	0	0	3/5
14	Rutherford Ranch	17.99	14.98	P	C	06	0	0	1	0	0	0	0	0	0	3/5
15	BV	19.99	12.98	N	C	05	0	0	1	0	0	0	0	0	0	3/5
16	Rabbit Ridge	17.99	14.98	P	C	05	0	1	0	0	0	0	0	0	0	3/5
17	Screw Kappa Napa	12.99	12.98	N	C	05	1	0	0	0	0	0	0	0	0	3/5
18	Screw Kappa Napa	12.99	12.98	N	M	05	1	0	0	0	0	0	0	0	0	3/5
19	Castle Rock	13.99	9.98	N	C	06	0	0	0	0	0	0	0	0	0	3/5
20	Liberty School	13.99	10.98	P	C	06	0	0	0	0	0	0	0	0	0	3/5

From Table 4-1, one can see that ones and zeros were used for certain variables while other variables had multiple choices; therefore letters were used to identify each characteristic.

Table 4-2: First Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.517648
R Square	0.267959
Adjusted R Square	0.261458
Standard Error	8.959359
Observations	569

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	16542.3	3308.46	41.21659	3.7E-36
Residual	563	45192.07	80.27011		
Total	568	61734.37			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	20.62	0.81	25.44	0.00	19.03	22.21	19.03	22.21
Dummy for Napa	9.60	0.85	11.32	0.00	7.94	11.27	7.94	11.27
Dummy Redwood City	2.58	0.79	3.27	0.00	1.03	4.13	1.03	4.13
Dummy for Merlot	-5.26	0.88	-5.99	0.00	-6.99	-3.54	-6.99	-3.54
Dummy for Zin	-3.25	1.09	-2.97	0.00	-5.40	-1.10	-5.40	-1.10
Dummy for Pinot	-0.88	1.55	-0.57	0.57	-3.92	2.16	-3.92	2.16

This was the first regression of many. Most of the variables appear to be of high significance. With a base intercept of Cabernet Sauvignon from Paso Robles, some characteristics appear to increase with price. Napa on the label increases price by $9.60, Redwood City has a premium of $2.58 over San Luis Obispo. Merlot, Zinfandel, and Pinot Noir appear to have lower prices compared to the base price. However, the Dummy for Pinot does not appear to have a strong

influence on price as its p-value is not significant. Possible reasons include the sample size of Pinot Noir collected that was very low, making it not a great indicator of price.

Table 4-3: Second Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.523758494
R Square	0.274322961
Adjusted R Square	0.26263943
Standard Error	8.952190223
Observations	569

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	9	16935.16	1881.6839	23.47946	4.315E-34
Residual	559	44799.22	80.14171		
Total	568	61734.37			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	17.90	0.98	18.31	0.00	15.98	19.83
Dummy for Napa	10.29	0.81	12.69	0.00	8.70	11.88
Dummy Redwood City	0.43	0.80	0.53	0.59	-1.15	2.00
Dummy for Merlot	-0.11	0.89	-0.12	0.90	-1.87	1.64
Dummy for Zin	-0.20	1.09	-0.19	0.85	-2.34	1.94
Dummy for Pinot	2.92	1.60	1.83	0.07	-0.22	6.06
Animal	0.12	1.03	0.12	0.91	-1.91	2.15
"Estate"	2.26	1.23	1.84	0.07	-0.15	4.67
Dummy Safeway or Vons	3.57	0.98	3.64	0.00	1.64	5.51
Dummy BevMo	-2.34	0.93	-2.51	0.01	-4.16	-0.51

This regression also showed significant and insignificant results to the study. Variables left the reader to have a generalized assumption that the price may be affected by certain variables however, a high P-value greater than 0.10 leaves one to conclude there would not be a significant correlation between price and label for animal, Zin, Merlot and Redwood City. However, Napa, Pinot Noir, “estate”, Safeway, and BevMo were significant at the 90% Confidence Interval.

Table 4-4: Third Regression
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.584057
R Square	0.341122
Adjusted R Square	0.32811
Standard Error	8.545517
Observations	569

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	11	21058.97	1914.45	26.22	0.00
Residual	557	40675.40	73.03		
Total	568	61734.37			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	19.75	0.87	22.73	0.00	18.04	21.45
Dummy for Napa	10.13	0.83	12.13	0.00	8.49	11.77
Dummy Redwood City	4.14	1.05	3.96	0.00	2.08	6.19
Dummy for Merlot	-5.23	0.85	-6.17	0.00	-6.90	-3.57
Dummy for Zin	-3.13	1.07	-2.92	0.00	-5.23	-1.03
Dummy for Pinot	-1.58	1.49	-1.06	0.29	-4.51	1.36
Animal	2.18	1.00	2.19	0.03	0.23	4.14
"Reserve"	-0.39	1.93	-0.20	0.84	-4.17	3.39
"Estate"	6.29	1.17	5.37	0.00	3.99	8.59
Dummy BevMo	-1.46	0.97	-1.51	0.13	-3.36	0.44
Dummy K&L	-1.68	1.33	-1.26	0.21	-4.28	0.93
Dummy BevMo*Redwood City	-5.29	1.81	-2.93	0.00	-8.83	-1.74

In the model above, the independent variables that showed high significance and insignificance to price. The dummy variables that had a p-value of 0.00 or less, were Napa, Redwood City, Merlot, Zin, Animal, Estate, BevMo*Redwood City. The p-value shows that these variables are good indicators of explaining the variation in price between both regions in this regression. In contrast, the p-value of 0.84 for “reserve” and 0.29 for Pinot Noir indicates that by including “reserve” and/or Pinot Noir in the regression, the variables were not strong indicators for price. However, a very strong indicator of price was the label “Napa” on bottle. The dummy variable for “Napa” was $10.13 higher than the retail shelf prices labeled “Paso Robles”.

Table 4-5: Progressive Model of Coefficients

	Progressive 1st Model		Progressive 2nd Model		Progressive 3rd Model	
Intercept	20.62	***	17.90	***	19.75	***
Varietal						
Merlot			-0.11		-5.23	***
Zinfandel			-0.20		-3.13	***
Pinot Noir			2.92	*	-1.58	
Region						
Napa	9.60	***	10.29	***	10.13	***
Location						
Redwood City			0.43		4.14	***
Label						
Animal					2.18	**
"Reserve"					-0.39	
"Estate"					6.29	***

*** Significant at the 1% level

Table 4-5 is a progressive model from three different regression trials. From the results, one can infer that the majority of values were significant at the 99% CI. Of the three trials shown in Table 4-4, labeling of Napa was significant at the 99% CI in every regression.

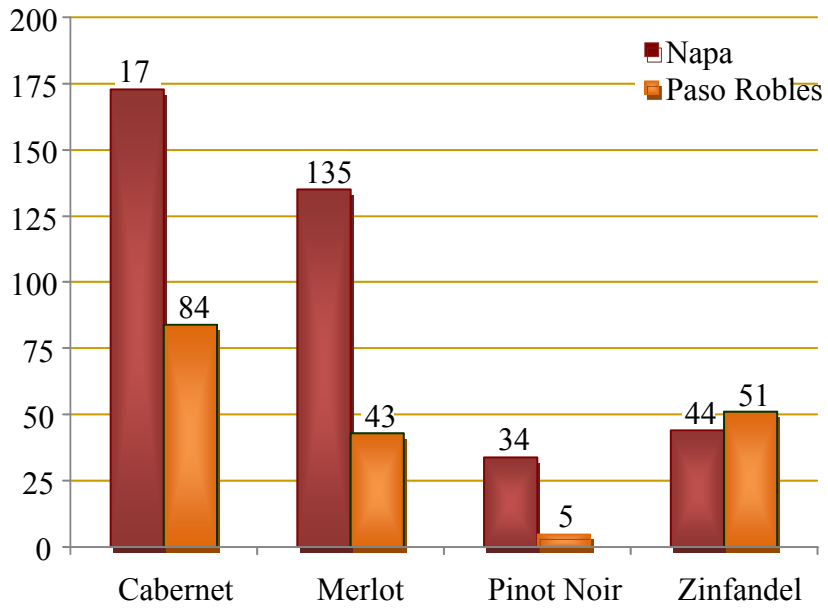

Figure 4-1: Region and Varietal Number of Observations

Figure 4-1 shows most of the samples were collected from Napa and there appears to be 173 Cabernet Sauvignon varietals. The lowest amount was the Paso Robles Pinot Noir with 5 samples. The low number of samples for Pinot Noir came from the low in-store availability when data collection took place. Therefore, the small sample size would not allow for accuracy in the results of the regressions.

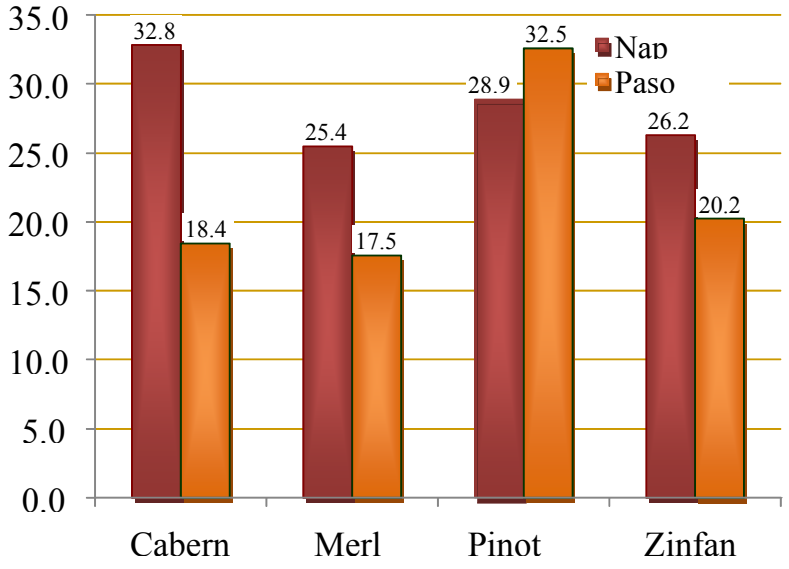

Figure 4-2: Average Price USD by Varietal and Region

Figure 4-2 shows Paso Robles' Pinot Noir nearly $4 more than Napa Pinot Noir. However, the graph also indicates that the majority of bottles observed had higher prices with the label "Napa" than labeling "Paso Robles." In conclusion, based on observation of the graph, most of the bottles had close prices that were within $4-$14 of each other.

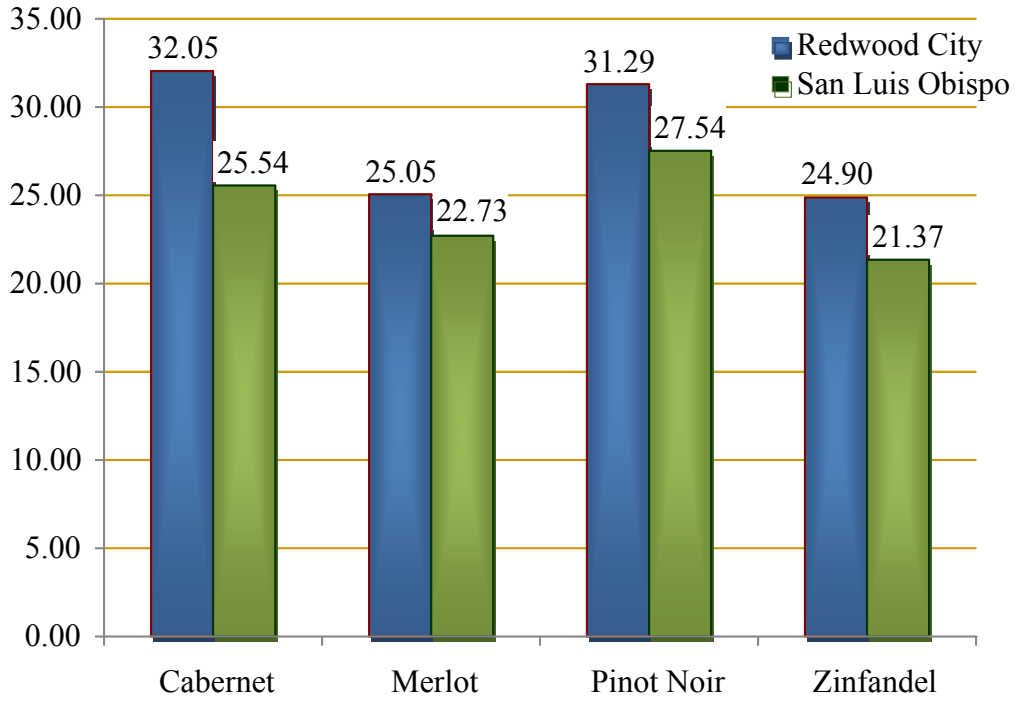

Figure 4-3: Market Location and Varietal with Average Price USD

Figure 4-3 shows that bottle prices in Redwood City tend to be higher than the price on the shelf in San Luis Obispo regardless of region on bottle. This can allow one to assume that the economic standards in Redwood City may be higher for the region and the stores in that region may charge higher prices than San Luis Obispo.

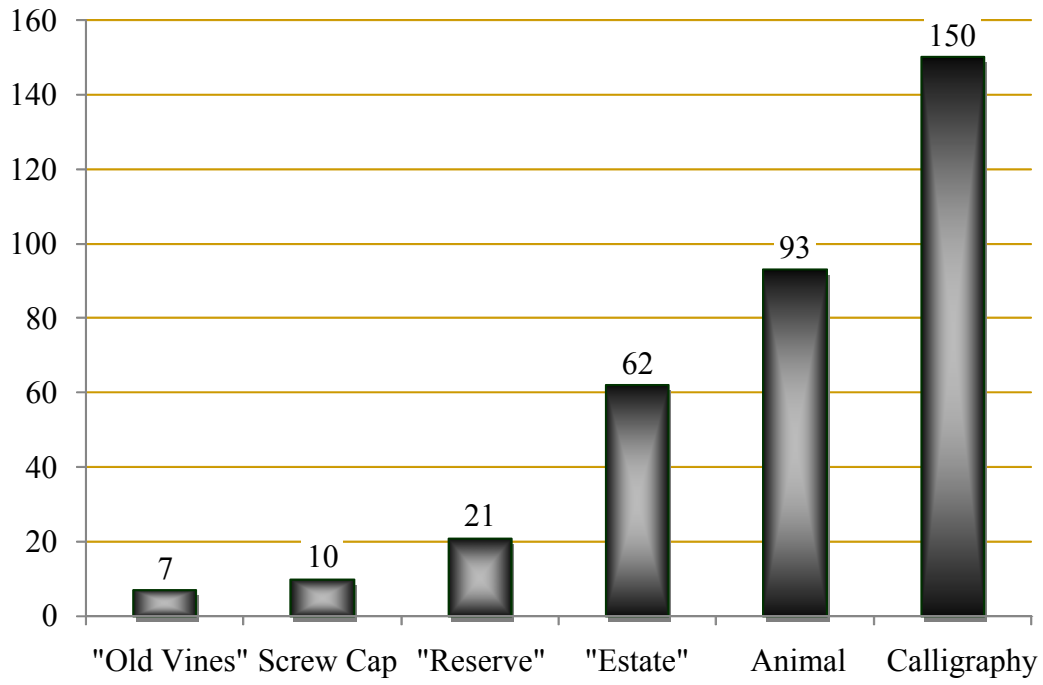

Figure 4-4: Label Attribute: Number of observations

Figure 4-4 shows a quantitative value for labeling characteristics. Calligraphy ranked highest amongst samples collected in San Luis Obispo and Redwood City. In contrast, “Old Vines” and “Screw Caps” were the lowest two sample collected. From the graph, one can see that most bottles contained “Calligraphy” and possibly an animal on the label for Napa and Paso Robles samples taken.

Chapter V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

Qualitative characteristics were analyzed by collecting data for 569 bottles of red wine in 8 major retail stores of Redwood City and San Luis Obispo, over a week-long period of time via hedonic regression analysis. After analyzing data, it has been found the qualitative characteristics, such as labeling, do in fact show a strong correlation between increased prices of the red wine. Based upon data gathered, one can understand that by comparing price premiums in two California AVA's, the hedonic results shows that adding Napa Valley on the label adds a price premium of $10.13 for red wines in the Cabernet Sauvignon, Merlot, Pinot Noir and Zinfandel varietal categories. Therefore, the researcher's hypothesis was wrong and the resulting price was >$6. Of the independent variables that were assumed to have significance to price, one major variable such as the screw cap closure method was found not to be significant to price at this time. Possible reasons include those of the sample size being too small to have an effect that can be measured accurately for the data results.

Conclusion

Based upon analysis, by understanding some intuitive statistical results, the individual can see if each variable chosen is significant or insignificant to the study. Many of the original hypothesized important variables ended up not being significant at all. The variables were not significant because the numbers of observations were minimal so a larger sample was needed, or there wasn't a clear relationship between a chosen variable and price.

If a viticulture type company wanted to find out where they could make more money in the market, they could look at this data. Whether the company specializes in growing, harvesting, fermenting, bottling, marketing, branding, buying, selling grapes, or even all those steps combined, they would find this information useful. They could purchase some land in the Napa region or buy a majority of grapes grown there and bottle their wines. Using calligraphy and or an animal on the label showed a correlation to price so that would be a strong part of the branding and labeling strategy.

Recommendations

The project researcher recommends the use of this data because it is representative of both regions sell tons of wine throughout California and other parts of the world. The researcher wanted to see if price was a function of the quality characteristics that were used on a product. After much examination and viewing other studies, the researcher found that price is always a function of the attributes that make up the finished product.

The project did create difficulties for the researchers with time conflicts during the school year. The project researcher had to work with two other individuals and travel to another city to gain data. This made it difficult because it took time to travel, time to gather data, and time to analyze the significance of the data. Now, if this project were redone, the individual could spend more time understanding that other areas or regions can be tested to evaluate bottle characteristics in more detail.

Also, more analysis should be conducted on wine ratings, screw-caps, and vintage variables to strengthen a future regression. The researcher opted out of using a rating system on the regression because multiple companies and/or people rate wine subjectively to their palate. If

the project was redone in the future, an individual might want to simplify data collection by listing “1” if a score was available or “0” if a score was not available. Also, more analysis should be conducted on regions that carry more screw caps and vintage of bottles listed at the retail level to evaluate significance of year and price.

References Cited:

Atkin, T. Dove, Duane. 2006. “Rodney Strong Winery: The Great Cork Debate”. *Journal of the International Academy for Case Studies*. Pg 77-89.

Atkin T., Garcia R., Lockshin L. 2006.“A Multinational Study of the Diffusion of a Discontinuous Innovation.” *Australasian Marketing Journal*; 2006. Pg. 17, 10pgs.

Boyd, Gerald D. "Put a cork in it." Restaurant Hospitality 1 Oct. 1994: 92.

Brooks, Eileen Lizabeth. 2001“Why don't firms export more?” Harvard University, 2001, 97 pages.

Courtney, S., 2001. “The history and revival of screw caps.” Wine of the Week (August).

"Dummy Variable." 2008. Wikipedia, The Free Encyclopedia. 27 Jul 2008, 16:11 UTC. 6 Mar 2009.

Farnham, Alan. 2003. "The Next Napa." Forbes 20 Jan. 2003

Farquar, P.H. (1989), "Managing brand equity", Marketing Research, Vol. 1, September, pp. 24

Gunter Schamel, and Kym Anderson. "Wine quality and varietal, regional and winery reputations: Hedonic prices for Australia and New Zealand." *Economic Record* 79.246 (2003): 357.

Parker, Robert. 2009 “About Robert M. Parker, Jr.” *The Wine Advocate.*

Perrouty, Jean, François d'Hauteville, and Larry Lockshin.2006. "The influence of wine attributes on region of origin equity: An analysis of the moderating effect of consumer's perceived expertise." *Agribusiness* 22.3 (2006): 323.

Schiefer, Jan and Christian Fischer. 2008. "The Gap Between Wine Expert Ratings and Consumer Preferences :Measures, Determinants and Marketing Implications." *International Journal of Wine Business Research* 1 Oct. 2008

Taranto, Christopher. 2007. “2007 San Luis Obispo County Wine Industry Awards Announced.” *Wine Country Press*. July 24.

Walker, L. 2001. "The surge from brand Australia", Wines & Vines, Vol. 84 No. 7, pp. 28-30.

Wine Institute 2001. “American Viticultural Areas” *The Wine Institute*

Ulrich R Orth, Marianne McGarry Wolf, and Tim H Dodd. "Dimensions of wine region equity and their impact on consumer preferences. " *The Journal of Product and Brand Management* 14.2/3 (2005): 88-97.

Unnevehr J Laurian, Fatoumata C Gouzaou. Retail Premiums for Honey Characteristics. Agribusiness (1996-1998); Jan/Feb 1998; 14, 1.

Appendix

1. Attached is a Microsoft Excel sheet with partial Data collection.

#	Store	City	Brand	Price	Price Disc.	Location	Varietal	Year/Vintage	ScrewCap	Animal	Caligraphy	Old Vines	Reserve	Estate	Sustainability(Frnt)	Sustainability(Back)	Organic Seal	# of Shelves
1	Chain A	San Luis Obispo	Beaulieu Vineyard (Beaulieu Vineyard (BV))	32.99	27.99	N	C	05	0	0	1	0	0	0	0	0	0	4
2	Chain A	San Luis Obispo	Five Rivers	12.99		P	C	06	0	0	0	0	0	0	0	0	0	1
3	Chain A	San Luis Obispo	Heitz Cellar	40.99		N	C	03	0	0	1	0	0	0	0	0	0	4
4	Chain A	San Luis Obispo	Eberle	18.99		P	C	06	0	1	0	0	0	0	0	0	0	4
5	Chain A	San Luis Obispo	Tobin James "Notorious"	15.99		P	C	06	0	0	0	0	0	0	0	0	0	4
6	Chain B	Redwood City	Tobin James "Ballistic"	17.99		P	Z	06	0	0	0	0	0	0	0	0	0	3
7	Chain B	Redwood City	Peachy Canyon Winery "West Side"	17.99		P	Z	06	0	0	0	0	0	0	0	0	0	2
8	Chain B	Redwood City	Kuleto Estate	39.99		N	Z	06	0	0	0	0	0	1	0	0	0	4
9	Chain B	Redwood City	Green & Red Vineyard "Tip Top Vineyard"	26.99		N	Z	04	0	1	0	0	0	0	0	0	0	4
10	Chain B	Redwood City	Frank Family Vineyards	34.99		N	Z	06	0	0	0	0	0	0	0	0	0	1
11	Chain B	Redwood City	Hullabaloo	17.99	15.99	N	Z	06	0	0	0	0	0	0	0	0	0	1
12	Chain C	Redwood City	Buehler	17.99		N	Z	07	0	1	0	0	0	0	0	0	0	
13	Chain C	Redwood City	Frogs Leap	23.99		N	Z	06	0	1	0	0	0	0	0	0	0	2
14	Chain C	Redwood City	Frogs Leap	29.99		N	M	06	0	1	0	0	0	0	0	0	0	
15	Chain C	Redwood City	Heitz Cellars	19.99		N	Z	05	0	0	1	0	0	0	0	0	0	2
16	Chain C	Redwood City	Rosenblum Cellars "Richard Sauret Vineyards"	19.99		P	Z	06	0	0	0	0	0	0	0	0	0	4
17	Chain C	Redwood City	Buehler	17.99		N	Z	07	0	1	0	0	0	0	0	0	0	
18	Chain D	Redwood City	Beaulieu Vineyard (Beaulieu Vineyard (BV))	24.99		N	P	06	0	0	1	0	0	0	0	0	0	4
19	Chain D	Redwood City	EDNA VALLEY VINEYARDS	15.99		P	C	06	0	0	0	0	0	0	0	0	0	1
20	Chain D	Redwood City	ESTANCIA	14.99	9.99	P	C	06	0	0	1	0	0	0	0	0	0	3
21	Chain D	Redwood City	PASO CREEK	18.99		P	C	05	0	1	0	0	0	0	0	0	0	2
22	Chain D	Redwood City	ACACIA	26.99		N	P	06	0	0	0	0	0	0	0	0	0	4
23	Chain D	Redwood City	EDNA VALLEY VINEYARDS	15.99		P	C	06	0	0	0	0	0	0	0	0	0	1
24	Chain E	San Luis Obispo	Beaulieu Vineyard (Beaulieu Vineyard (BV))	22.99	16.99	N	M	05	0	0	0	0	0	0	0	0	0	1
25	Chain E	San Luis Obispo	Edna Valley Vineyard	17.99	14.99	P	C	05	0	0	0	0	0	0	0	0	0	4
26	Chain E	San Luis Obispo	Beringer	22.99	19.99	N	P	06	0	0	0	0	0	0	0	0	0	4
27	Chain E	San Luis Obispo	Acacia	27.99	24.99	N	P	06	0	0	0	0	0	0	0	0	0	4
28	Chain E	San Luis Obispo	Trinchero	13.99	12.99	N	P	06	0	0	0	0	0	0	0	0	0	1
29	Chain E	San Luis Obispo	Aquinas	16.99	13.99	N	P	07	0	0	0	0	0	0	0	0	0	0
30	Chain F	Redwood City	PINE RIDGE CRIMSON CREEK	39.99	32.98	N	M	05	0	0	1	0	0	0	0	0	0	5
31	Chain F	Redwood City	TREFETHEN ESTATES	32.99	24.98	N	M	04	0	0	1	0	0	0	1	0	0	5
32	Chain F	Redwood City	Stags Leap Wine Cellars	36.99	27.98	N	M	05	0	1	1	0	0	0	0	0	0	5
33	Chain F	Redwood City	KEENAN	39.99	34.98	N	M	04	0	0	0	0	0	0	0	0	0	5
34	Chain F	Redwood City	FROG'S LEAP RUTHERFORD	37.99	32.98	N	M	06	0	1	0	0	0	0	0	0	0	5
35	Chain F	Redwood City	SWANSON OAKVILLE	32.99	23.98	N	M	04	0	0	0	0	0	0	0	0	0	5
36	Chain F	Redwood City	CHARLES KRUG PETER MONDAVI	22.99	18.98	N	M	06	0	0	1	0	0	0	0	0	0	4
37	Chain F	Redwood City	Stags Leap Wine Cellars	36.99	27.98	N	M	05	0	1	1	0	0	0	0	0	0	5
38	Chain G	San Luis Obispo	Screw Kappa Napa	12.99	12.98	N	C	05	1	0	0	0	0	0	0	0	0	3/5
39	Chain G	San Luis Obispo	Screw Kappa Napa	12.99	12.98	N	M	05	1	0	0	0	0	0	0	0	0	3/5
40	Chain G	San Luis Obispo	Sequoia Grove	43.99	30.79	N	C	99	0	0	0	0	0	0	0	0	0	top 1/5
41	Chain G	San Luis Obispo	Silverado Vineyards	49.99	34.99	N	C	04	0	0	1	0	0	0	0	0	0	top 1/5
42	Chain G	San Luis Obispo	ST CLEMENT VINEYARDS	44.69	31.28	N	C	04	0	0	1	0	0	0	0	0	0	top 1/5
43	Chain G	San Luis Obispo	St Supery	37.49	26.24	N	C	01	0	1	0	0	0	0	0	0	0	top 1/5
44	Chain G	San Luis Obispo	St Supery	27.99	19.59	N	M	01	0	1	0	0	0	0	0	0	0	top 1/5
45	Chain G	San Luis Obispo	Stags Leap Wine Cellars	48.49	33.94	N	M	04	0	1	0	0	0	0	0	0	0	top 1/5
46	Chain H	San Luis Obispo	Atalon	29.99		N	M	02	0	0	0	0	0	0	0	0	0	4
47	Chain H	San Luis Obispo	Burgess	24.99		N	M	04	0	0	0	0	0	0	0	0	0	4
48	Chain H	San Luis Obispo	Charles Krug	22.99		N	M	04	0	0	1	0	0	0	0	0	0	4
49	Chain H	San Luis Obispo	Frogs Leap	32.99		N	M	05	0	1	0	0	0	0	0	0	0	4
50	Chain H	San Luis Obispo	Ballentine Vineyards	16.99		N	M	03	0	0	0	0	0	1	0	0	0	3