W241 Final Report

Vous voulez du vin? - Does the use of foreign language advertising in wine increase customer's purchase likelihood?

Rory Liu, Charlotte Swavola, Sharad Varadarajan

Introduction

When walking through the grocery store, one can easily spot products with foreign-language labels - "Vins de France" on a wine bottle, "Wirklich gut!" on a sausage pack, etc. Advertisers often confront consumers with foreign languages, such as German or French [1]. In fact, the use of foreign language in advertising is a well-studied topic. According the study done by France Leclerc, Bernd Schmitt and Laurette Dubé,showing French pronunciation of a brand name affects the perceived hedonism of the products, attitudes toward the brand, and attitudes toward the brand name[2]. Furthermore, Jos Hornikx and Frank van meurs in their study that the use of foreign languages serve as a strong cue for a product's country of origin, and the associations that the foreign language evoke and those that the country-of-origin evoke are similar[3].

While scholars seem to agree that foreign language advertising shapes product perception by implicitly giving consumers cues on product's country of origin, few of the existing studies dived deep into the causal link between foreign language and consumer's purchase likelihood. This is an important causal link to establish. At the end of the day, the success of an advertising campaign is, or should be measured by the additional sales the campaign generates. Without understanding of whether having foreign language in product communications increases sales, advertiser will be ill-guided in their decision to apply such advertising techniques.

In our study, we seek to understand the relationship between foreign language advertising and consumer willingness to pay. To narrow down our scope for a manageable experiement given our time and resource constaints, we decided to focus on one particular product: wine. The high variance of wine prices, and the strong association between place of origin and the signal on quality makes it a great subject for our study.

Experimental Design

To fully understand the relationship between foreign language advertising and will purchases, we drafted a multi-factorial design for our experiement. There are three main factors that we want to investigate: advertising language, wine's country of origin, and length of flavor profile description.

For advertising language, we will test 3 languages: English (as baseline), French and German. We selected French because France is typically perceived as a premium wine production country, and

[1]Jos Hornikx, Frank van Meurs & Robert-Jan Hof (2013) The Effectiveness of Foreign-Language Display in Advertising for Congruent versus Incongruent Products, Journal of International Consumer Marketing, 25:3, 152-165, DOI: 10.1080/08961530.2013.780451

[2]Leclerc, F., B. H. Schmitt, and L. Dubé. 1994. Foreign branding and its effects on product perceptions and attitudes. Journal of Marketing Research 31 (2):263–270

[3]Jos Hornikx & Frank van Meurs (2017) Foreign Languages in Advertising as Implicit Country-of-Origin Cues: Mechanism, Associations, and Effectiveness, Journal of International Consumer Marketing, 29:2, 60-73, DOI: 10.1080/08961530.2016.1243996

		Page Language					
		English		French		German	
		Flavor Profile Description					
		Long			Short		
Country of Origin	US	1	2	3	4	5	6
	France	7	8	9	10	11	12

Figure 1: Groups Design

we hypothesize that advertsizing in French will increase consumer's willingness to pay. We also selected German as a comparison language. Germany is not typically associated with wine, and therefore, by including German, we will be able to see whether the effect of having foreign language is limited to the language of the country associated with a product, or whether it expands to other languages as well.

For wine's country of origin, we included US and France. Here we are particularly interested in the interaction between country of origin and advertsing language. Prior studies have found that foreign-language advertising serve as country of origin cues. By including country of origin as a variable, we are able to see whether foreign language has any additional impact when country of origin is also given.

For flavor profile description, we have two versions - long and short. We believe that this could be a proxy for foreign language "dosage". A long flavor profile in French might be more noticeable than a short French tagline, and may amplify the effect that we find.

To summarize, we have a 3 x 2 x 2 design, which results in 12 groups of participants. (See summary in figure 1).

Because of our time and resource constraint, we could not conduct a real-life experiement with actual wines in actual stores. Instead, we decided to conduct the study through a Qualtrics survey and gather responses through Mechanical Turk. The survey is structured into the following sections:

We start with some demographics question about participants primary language, income and gender. These wil serve as covariates in our later model. Then, we ask participants about their wine-related behavior, including how often they drink wine, whether they like Cabernet Sauvignon, and how often they purchase bottles of wine. We think that these wine-related behavior will explain some of the variation in willingness-to-pay, and therefore will be good covariates that help us reduce our standard error. The third section is the main part of our experiement where we introduce the treament. Here, we randomly assign the participants into one of twelve groups mentioned above, and show them a simulated purchasing page (example seen in figure 3).

In the simulated wine page, we translate all texts into the groups assigned language, and provide english translation for key elements (e.g. flavor profile, 'top wine of the year' tag) to make sure that

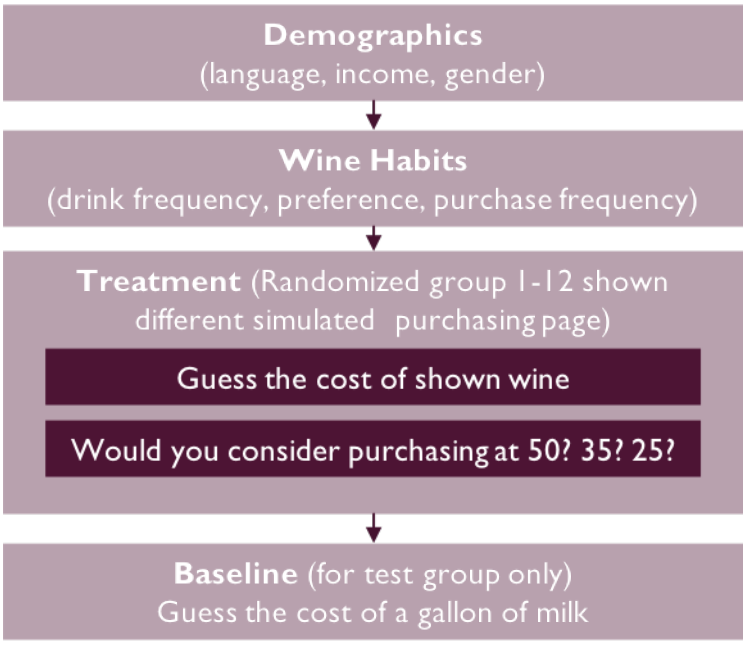


Figure 2: Survey Flow

Figure 3: Page Sample

the respondent understand the information. All groups are shown the exact same wine (which was a hypothetical wine fabricated by us), and the same descriptive information.

We then ask the respondents to guess the cost of the wine. This is an un-anchored direct pricing question, aimed to understand respondents' first perception of the wine shown. Afterwards, we use a simplified version of the Gabor Granger method developed by André Gabor and C. W. J. Granger in the 1960s [4]. In this method, we ask respondents whether they would consider purchase the shown wine at $50. If respondents say now, we ask the same question at a lower price point of $35. If they still answers no, we lower the price further to $20. Through this series of 3 questions, we seek to understand where their true willingness-to-pay lies. We selected the 3 price points to cover a wide range of potential willingness-to-pay, and validated our choice through a pilot study of 100 respondents. From the pilot, we saw that most respondents' willingness to pay fall into the 20-50 price range.

Our survey is concluded after the treatment section for our main sample of 1200 respondents. We added a baseline question asking respondents to estimate the price of a gallon of milk in our second launch. We aim to use the data from our first launch as "training set", where we test different regression models, and methods, and use the second launch data as "testing set" to validate that our findings are real. The baseline milk question helps us establish a reference price points for respondents price guesses, and adjust for any overall inflated / deflated guesses.

Survey data cleaning and analysis

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':
##
##     between, first, last

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Please cite as:
```

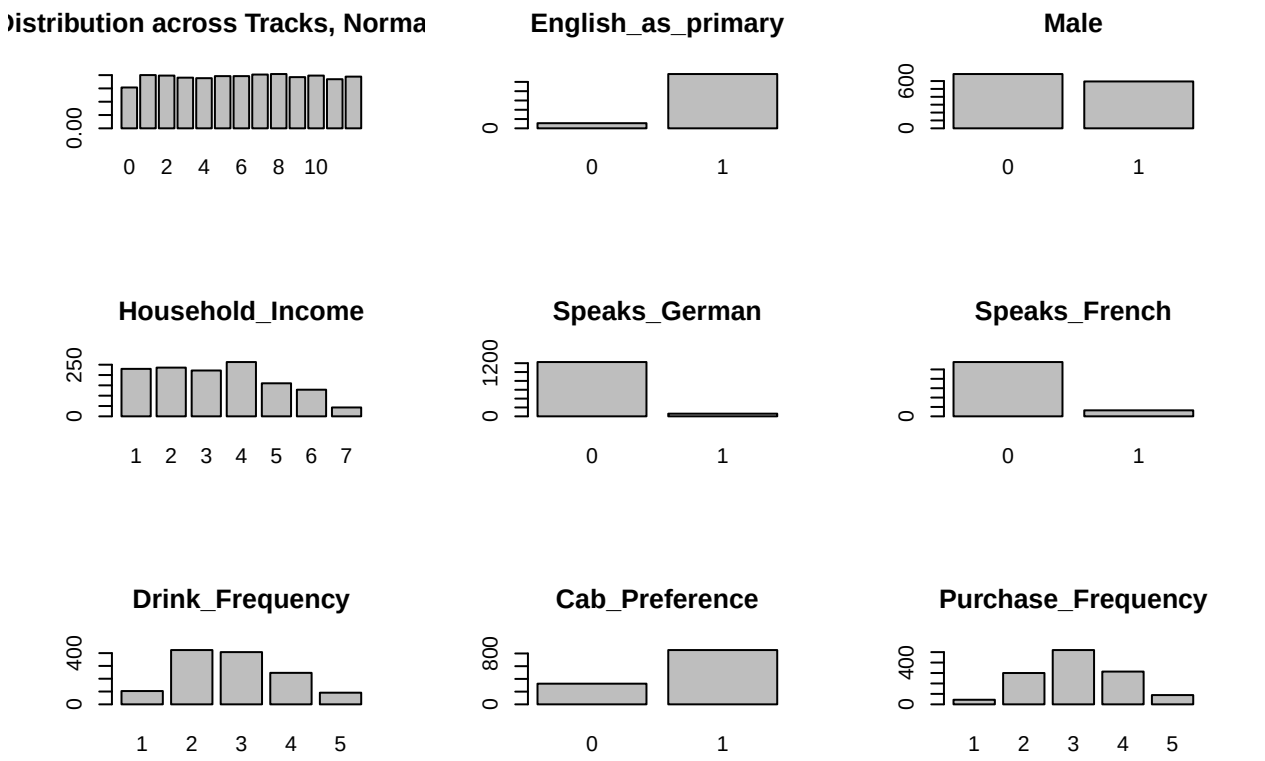
[4] Gabor, A. and Granger, C. (1966). Price as an Indicator of Quality: Report on an Enquiry. Economica, 33(129), p.43.

```
##  Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2.2. https://CRAN.R-project.org/package=stargazer
```

1. EDA and Cleaning

Pre-cleaning

```
demographics=data.table(d[,.(English_as_primary,Male,Household_Income,Speaks_German,Speaks_Fre
par(mfrow=c(3,3))
#look at randomization/tracks
barplot(prop.table(table(d$Track)), main="Distribution across Tracks, Normalized", names.arg =
#Important to note- 0 on barchart is N/A or attrition.
#demographic summary

for (i in (1:length(names(demographics)))){
  #print(summary(demographics[,..i]))
  barplot(table(demographics[,..i]),
          main=names(demographics)[i]
         #,names.arg = demographics_labels[i]
         )}
```




```
#colnames(d)
summary(d$Guess_the_Cost)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
```

```
##    1.00   15.00   25.00   75.24   45.00 8600.00      79
```

```
summary(d$Guess_the_Cost[d$French_Purchasing_Page == 1]) #French language only
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00   15.00   25.00   68.04   49.99 5000.00
```

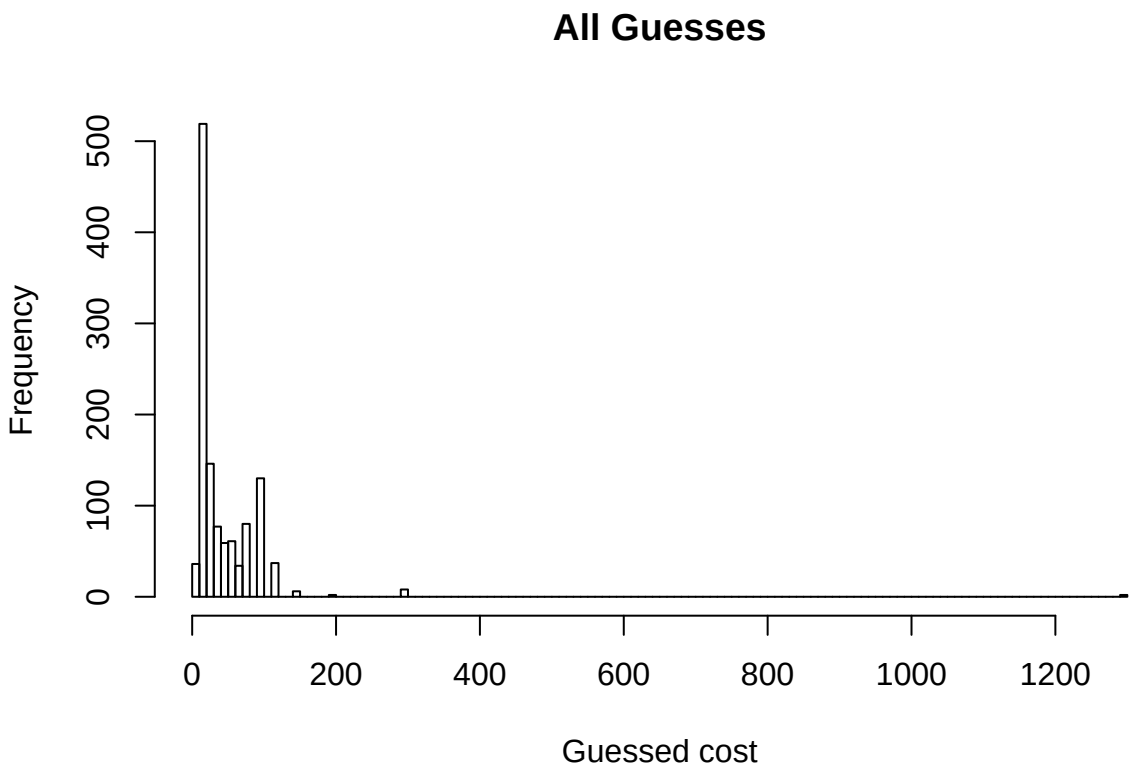
```
summary(d$Guess_the_Cost[d$German_Purchasing_Page == 1]) #German language only
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00   16.00   25.00   71.86   45.00 2050.00
```

```
summary(d$Guess_the_Cost[d$German_Purchasing_Page == 0 & d$French_Purchasing_Page == 0]) #Engl
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   15.00   24.00   85.38   45.00 8600.00
```

```
hist(d$Guess_the_Cost[d$Guess_the_Cost],main="All Guesses",breaks=100,xlab= "Guessed cost")
```



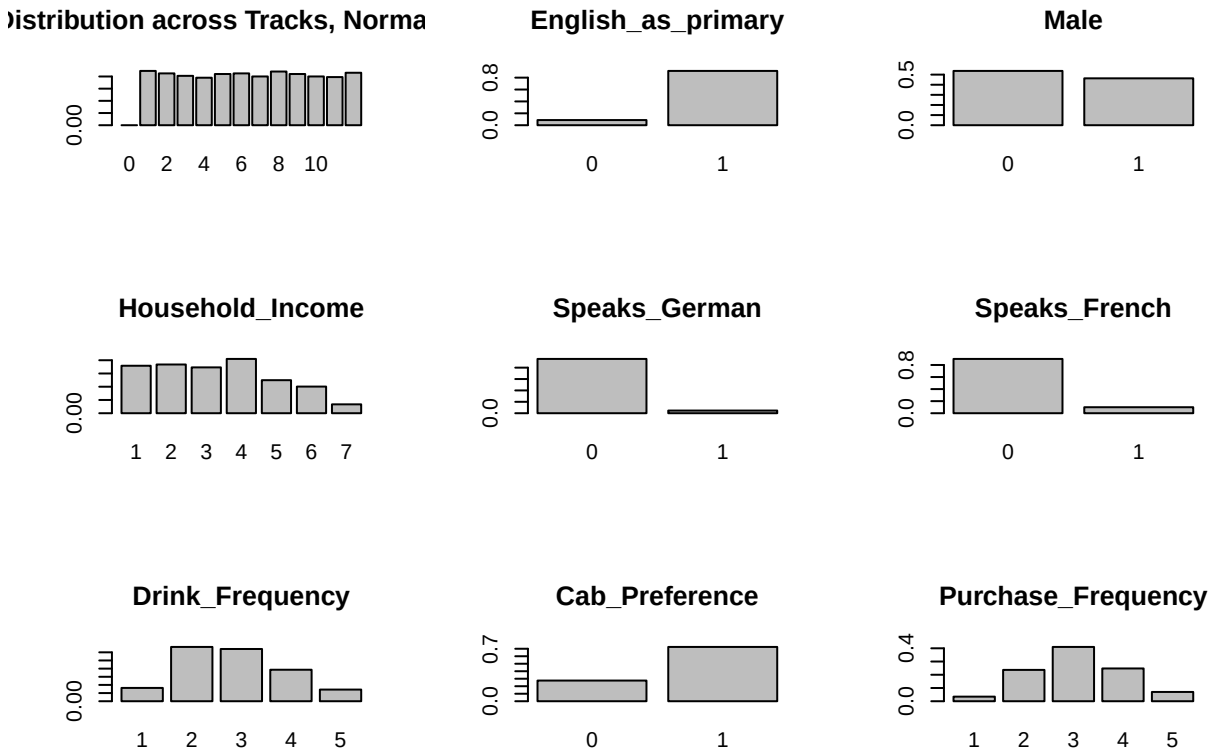
deduping/cleaning

```
library(data.table)
# Cleaning out incomplete responses before treament exposures
d <- d[!is.na(d$Track),]
#summary(d)
# Counting occurances of IP address and Geo-codes
d$geo_code <- paste(d$LocationLatitude, d$LocationLongitude)
```

```
dt = data.table(d)
dt[, `freq_geo` := .N, by = geo_code]
dt[, `freq_ip` := .N, by = IPAddress]
#class(as.data.frame(dt))
#dt
## Creatig 2 dedupped dataset, one with complete dedup where we do not allow any duplicates in
complete_dedup<- dt[dt$freq_ip==1 & dt$freq_geo==1]
#nrow(complete_dedup)
five_or_less_dedup <- dt[dt$freq_ip==1 & dt$freq_geo<=5]
#nrow(five_or_less_dedup)
d<-five_or_less_dedup
```

Post cleaning

```
d<-data.table(d)
par(mfrow=c(3,3))
#look at randomization/tracks
barplot(prop.table(table(d$Track)),main="Distribution across Tracks, Normalized",names.arg = c
#Important to note- 0 on barchart is N/A or attrition.


for (i in (1:length(names(demographics)))){
  #print(summary(demographics[,..i]))
  barplot(prop.table(table(demographics[,..i])),
          main=names(demographics)[i])}
```



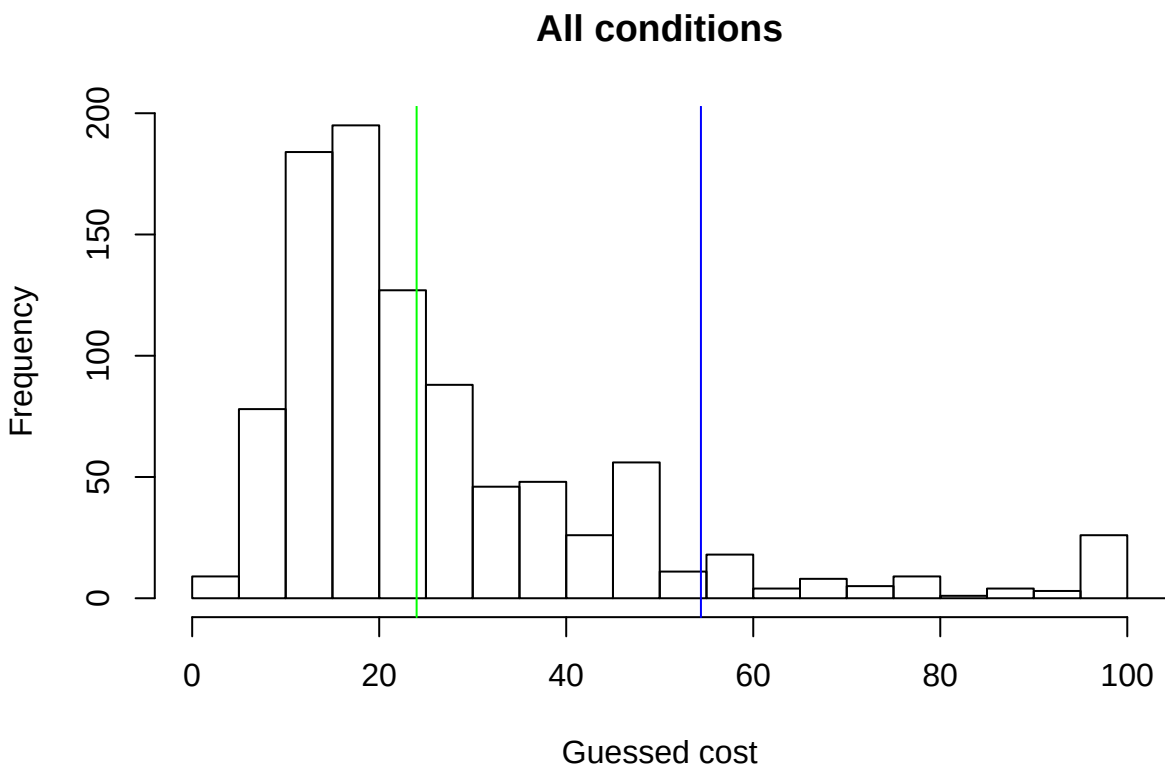
#Demographic/Covariate Balance Check

Now looking at Outcome Variables

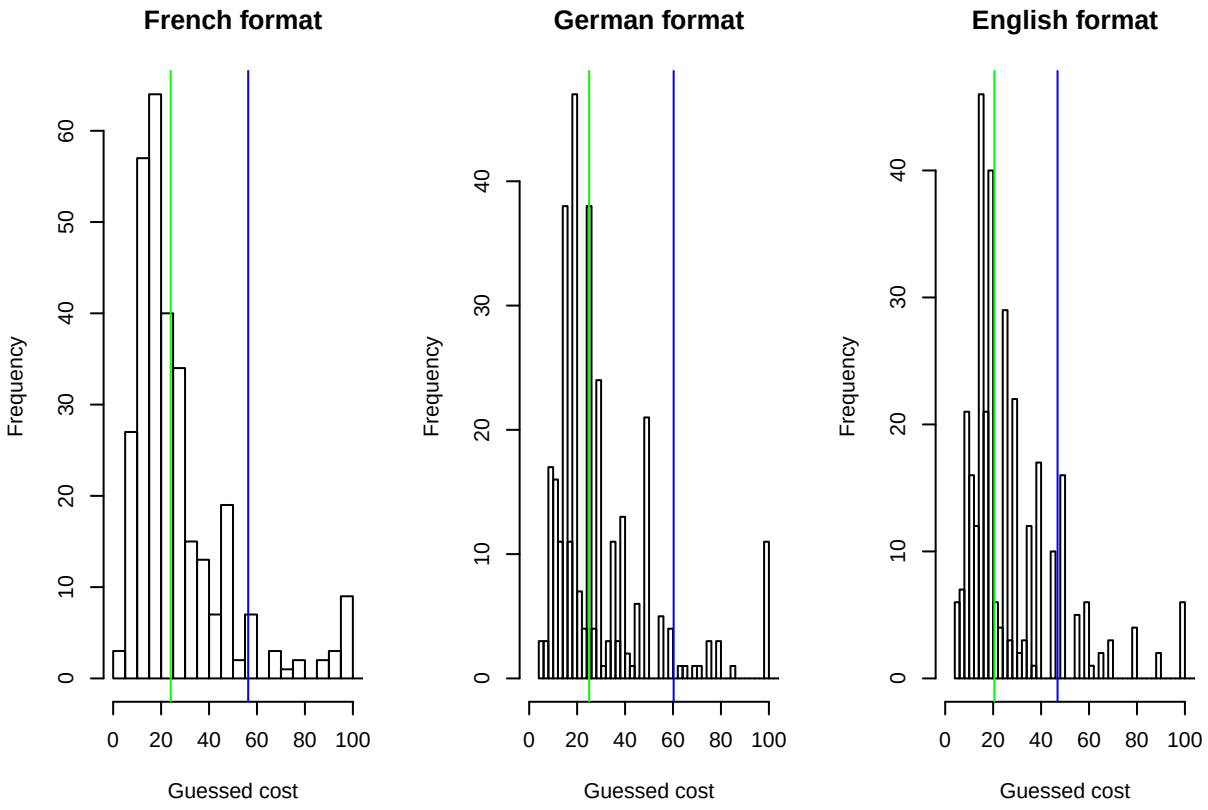
```
summary(d$Guess_the_Cost)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00   15.00   24.00   54.41   40.00 5000.00
```

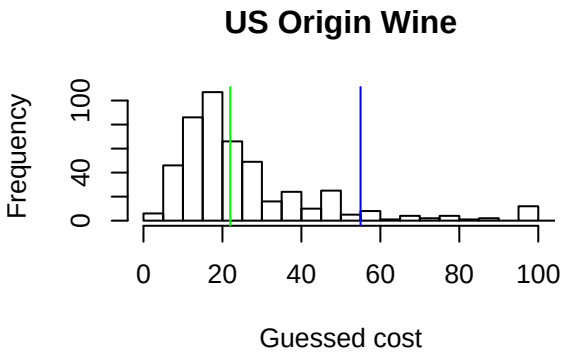
```
#hist, includes all cleaned
hist(d$Guess_the_Cost,main="All conditions",xlab= "Guessed cost", xlim = c(0,100),breaks=1000)
abline(v = mean(d$Guess_the_Cost), col = "blue")
abline(v = median(d$Guess_the_Cost), col = "green")
```

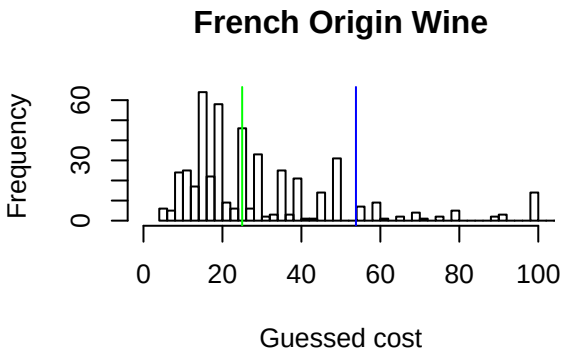


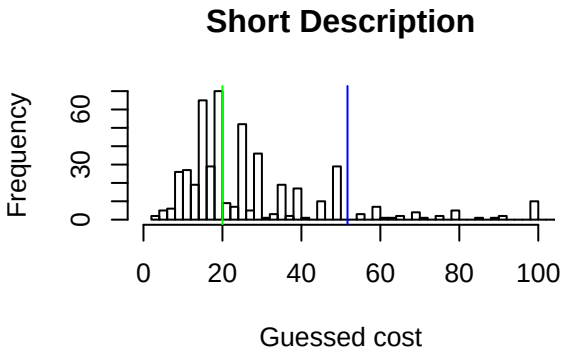

```
#formats
par(mfrow=c(1,3))
#pin(c(2,2))
hist(d$Guess_the_Cost[d$French_Purchasing_Page == 1],main = "French format",xlab= "Guessed cost
abline(v = mean(d$Guess_the_Cost[d$French_Purchasing_Page == 1]), col = "blue")
abline(v = median(d$Guess_the_Cost[d$French_Purchasing_Page == 1]), col = "green")

hist(d$Guess_the_Cost[d$German_Purchasing_Page == 1],main = "German format",xlab= "Guessed cost
abline(v = mean(d$Guess_the_Cost[d$German_Purchasing_Page == 1]), col = "blue")
abline(v = median(d$Guess_the_Cost[d$German_Purchasing_Page == 1]), col = "green")

hist(d$Guess_the_Cost[d$German_Purchasing_Page == 0 & d$French_Purchasing_Page == 0 ],main = "
abline(v = median(d$Guess_the_Cost[d$German_Purchasing_Page == 0 & d$French_Purchasing_Page ==
abline(v = mean(d$Guess_the_Cost[d$German_Purchasing_Page == 0 & d$French_Purchasing_Page == 0
```

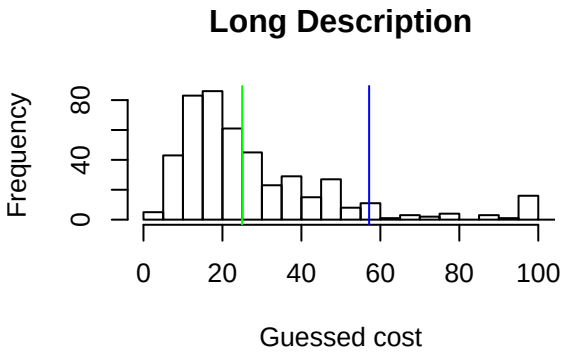



```r
#hist figures for origin, description, cleaned
par(mfrow=c(2,2))

#Origins
d_chart=d$Guess_the_Cost[d$US_Origin == 1]
hist(d_chart,main = "US Origin Wine",xlab= "Guessed cost", xlim = c(0,100),breaks=1000)
abline(v = mean(d_chart), col = "blue")
abline(v = median(d_chart), col = "green")

d_chart=d$Guess_the_Cost[d$US_Origin == 0]
hist(d_chart,main = "French Origin Wine",xlab= "Guessed cost", xlim = c(0,100),breaks=1000)
abline(v = mean(d_chart), col = "blue")
abline(v = median(d_chart), col = "green")

#descriptions
d_chart=d$Guess_the_Cost[d$Long_Description == 0]
hist(d_chart,main = "Short Description",xlab= "Guessed cost", xlim = c(0,100),breaks=1000)
abline(v = mean(d_chart), col = "blue")
abline(v = median(d_chart), col = "green")

d_chart=d$Guess_the_Cost[d$Long_Description == 1]
hist(d_chart,main = "Long Description",xlab= "Guessed cost", xlim = c(0,100),breaks=1000)
abline(v = mean(d_chart), col = "blue")
abline(v = median(d_chart), col = "green")
```


US Origin Wine
Frequency
0
40
100
0
20
40
60
80
100
Guessed cost


French Origin Wine
Frequency
0
30
60
0
20
40
60
80
100
Guessed cost


Short Description
Frequency
0
30
60
0
20
40
60
80
100
Guessed cost


Long Description
Frequency
0
40
80
0
20
40
60
80
100
Guessed cost