

Skunked Beer Tasting Experiment

Ashley Levato, Alex Smith, Andrew Burrage

Friday, April 29, 2016

Contents

Abstract	2
Introduction	2
Experimental Overview	2
Hypotheses	2
Practical Check	3
Experimental Design	3
Randomization	4
Pilot Study	5
Experiment Recruitment	5
Analysis and Results	6
Variable Descriptions and Distribution	6
Main Hypothesis Testing	11
Impressions of Results	14
Additional Analysis	14
Conclusions and Future Work	18
Appendix: Beer Tasting Survey	18

Abstract

The objective of this study was to determine if people can taste the difference between skunked and non-skunked beer. To evaluate this, we experimentally skunked Corona and had participants blindly taste samples of Corona that was exposed to different amounts of sunlight. Participants tasted and scored three Coronas bottled in clear glass and three Coronas in aluminum cans. Glass bottles and aluminum cans were exposed to varying amounts of sunlight: zero minutes, 10 minutes and 30 minutes. The taste tests were conducted in three different locations: California, Georgia and Virginia. Friends and family were recruited to be participants for this study. The results of our study indicate that, on average, participants were not able to distinguish the difference between skunked beer and non-skunked beer. However, on average, participants preferred the Corona in cans over bottles.

Introduction

Skunked beer is documented and common knowledge among beer drinkers; however, can an average beer drinker actually taste the difference? The science of skunked beer is documented. Beers bottled in clear and green bottles are notorious for being skunked, whereas, beers bottled in brown bottles or aluminum cans are known to be resistant to skunking. However, sources vary on the time it takes to skunk a beer.

Skunked beer smells like the name implies, similar to what the furry striped mammal sprays. Not only does skunked beer have a bad aroma, it doesn't taste good either. According to an article that appeared in Popular Science, isohumulone is a bittering agent that ultraviolet light can degrade. With the help of riboflavin, visible light can also degrade isohumulone. When beer is exposed to light, chemical reactions start taking place and isohumulone becomes 3-methylbut-2-ene-1-thiol, 3MBT, which is chemically similar to compounds found in skunk spray. Beer gets a lot of its flavor from hops. When hops are boiled, they release isoalpha acids into the beer. If exposed to sunlight these isoalpha acids start to break down. The resulting compounds bind with proteins that contain sulfur to create a new chemical, 3MBT. When a beer becomes light-struck in a bottle with clear, green, or blue glass, the skunking reaction begins because clear, green, or blue glass does not filter out ultraviolet light. However, brown bottles are effective at keeping out ultraviolet and aluminum cans are even better at blocking ultraviolet light. Due to the problem of skunking, many beer manufacturers choose to sell their beer in brown bottles/cans.

Experimental Overview

Hypotheses

This experiment was designed to test two hypotheses:

- (1) Can individuals taste the difference in skunked beer?
- (2) Can individuals taste the difference between beer from a bottle and from a can?

In this experiment, we are testing for normal or average individuals, not beer sommelier or experts in this field.

The null hypothesis (H_0) for each of these experimental questions are:

- (1) Individuals rate beer that has not been exposed to sun and beer that has been exposed to sun the same
- (2) Individuals rate beer that is stored in a bottle and beer that is stored in a can the same.

Practical Check

Before committing to a full experiment, we need to determine if skunking beer is feasible. In order to test the feasibility, we needed to determine the thresholds to use in the experiment. In the first round of tests, we used Heineken and varied the sun exposure from 2 minutes to 120 minutes. We found that beers with sun exposure greater than 20 minutes tasted skunked. The second round then tested a narrower time window, since some literature suggested only a few minutes of exposure is needed. In addition, we decided to explore if individuals can taste a difference between the normal beer, a slightly skunked beer, and a fully skunked beer. The second round tested both Heineken and Corona at intervals of 0, 10, 20, and 30 minutes of sun exposure. The results are shown in the table below.

Table 1: Beer Skunking Tests

Location	Beer	Time (min)	Skunked
CA	Heineken	30	Yes
CA	Heineken	60	Yes
VA	Heineken	0	No
VA	Heineken	10	Yes
VA	Heineken	20	Yes
VA	Heineken	30	Yes
VA	Corona	0	No
VA	Corona	10	Partial
VA	Corona	20	Yes
VA	Corona	30	Yes
GA	Heineken	0	No
GA	Heineken	2	No
GA	Heineken	20	Yes
GA	Heineken	120	Yes

The Corona beers had more distinguishing characteristics at each time level than the Heineken. Based on these observations the time exposures were set to 0 minutes for the control beers, 10 minutes for the lightly skunked beer, and 30 minutes for the completely skunked beer. Additionally, tests were conducted by skunking in bright sunlight, overcast and rainy weather, and through a window. All methods were successful in skunking the beers as UV exposure is the key component in skunking.

Experimental Design

The overall experimental design is a 3x2. We only use Corona beer since it is available in both clear bottles and cans, available in all locations, and had varying levels of “skunked-ness” based on the feasibility testing. The experimental categories are shown graphically below and are the three stages of beer exposure to ultraviolet light via sunlight at 0 minutes, 10 minutes, and 30 minutes, and the two container styles of clear bottles and aluminum cans.

Table 2: Beer Numbering for Data Recording

	Bottle	Can
No Sun Exposure	1	4
10 min Sun Exposure	2	5
30 min Sun Exposure	3	6

Each location purchased all of the beer to run its experiment from the same store, at room temperature, and in one batch. This removes any supplier variation in storage or heat cycling that could be introduced if the beer was purchased cold or from various locations. The beer in each location was skunked the day of the experiment and then put in a cooler or refrigerator to chill before serving.

The study is single blind because the subjects did not know anything about the beers or the experiment. Subjects did not see the beer being poured or know whether the beer comes from cans or bottles. We poured samples of 2-3 ounces for each tasting. Participants were asked to fill out a survey (available in the Appendix) with some demographic information. The researchers told the participants that they were tasting similar beers and to eat a cracker in between each tasting sample. Participants were asked to rate each beer individually on a scale of 1-5, with 1 being extremely dislike and 5 being really enjoy. This scale was chosen to inhibit individuals from ranking the beers in preference order.

Randomization

In order to mitigate effects of drinking order, the order of beers tasted was randomized for each individual. The random orders were creating using the following R code.

```
x<-c('A', 'B', 'C', 'D', 'E', 'F')
replicate (50, sample(x))
```

In case there is any association of particular letters with positive or negative connotations, like A being the best or F being the worst, we randomized the lettering of the beers based on location. The 4 tables below show the labeling for each location.

Table 3: Beer Randomization: CA

	Bottle	Can
No Sun Exposure	A	D
10 min Sun Exposure	B	E
30 min Sun Exposure	C	F

Table 4: Beer Randomization: GA

	Bottle	Can
No Sun Exposure	E	F
10 min Sun Exposure	A	B
30 min Sun Exposure	C	D

Table 5: Beer Randomization: VA

	Bottle	Can
No Sun Exposure	F	C
10 min Sun Exposure	B	D
30 min Sun Exposure	E	A

Table 6: Beer Randomization: UCB

	Bottle	Can
No Sun Exposure	D	E
10 min Sun Exposure	C	A
30 min Sun Exposure	F	B

Pilot Study

To ensure we can experimentally skunk beer, test the mechanics of running the experiment and check if we can detect a treatment effect we conducted a pilot study in California. We solicited the help of 12 participants comprised of family, friends and coworkers. Participants were asked to fill out the survey as described in the experimental design. There was a total of six samples presented to the participants. The order of beer sampling was randomized to eliminate rating bias based on order. We exposed Corona in aluminum cans and glass bottles to three different sun exposure levels: zero, ten minutes and thirty minutes.

The pilot study is intended not only to test the mechanics of running the experiment, but also to see if we can detect anything before proceeding to the full experiment. We will walk through the analysis in the full experimental analysis section, but the main components are an analysis of variance test (ANOVA) test to test against our null hypothesis that there is no difference between the ratings of skunked and non-skunked beers and a regression analysis.

For part of the analysis we used an ANOVA because we want to make multiple comparisons without making a Bonferroni correction. When performing multiple statistical comparisons, one increases the chances of incorrectly rejecting the null hypothesis by chance. For example, when doing a single t-test comparing two states and using a standard of $p=0.05$, we only have a 5% chance of incorrectly rejecting the null hypothesis. However, once we compare another set of states, even if we keep our p-value at 0.05, we now have a 9.75% chance of incorrectly rejecting the null hypothesis. We wanted to compare variables between 4 locations. This would have been 6 different comparisons. Using a p-value of 0.05 would have resulted in a 26.49% chance of incorrectly rejecting the null hypothesis at least once. Instead, we use an ANOVA test to determine if the variance between the variables by states were statistically significant. The results from the ANOVA test were not statistically significant, but this is a relatively small sample size.

We used regression analysis using within subject design to test if there was a causal relationship between the sun exposure or vessel type (bottles vs cans). The regression analysis equation is shown below. We were able to detect a statistically significant result between cans versus bottles and a slight significant result with the fully skunked beer. So we decided to proceed with the full experiment.

$$Rating = \beta_0 + \beta_1 Bottle + \beta_2 Sun10 + \beta_3 Sun30 + \beta_4 Bottle \cdot Sun10 + \beta_5 Bottle \cdot Sun30$$

Experiment Recruitment

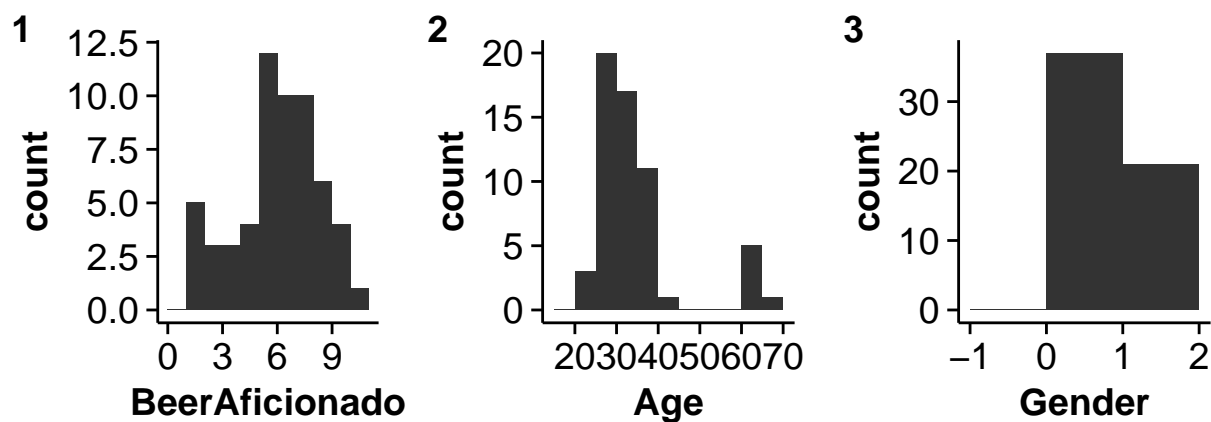
The experimental design was implemented in four locations from subsamples of the population composed of family, friends, and co-workers. Each location aimed to recruit a range of beer aficionado participants. The experiment was run in multiple settings including small groups of individuals as well as large gatherings. We assigned each person to a random drink order and wrote this order on the survey. Since the recruitment population is focused on individuals the researchers know, we expect there may be some covariates that are not representative of the overall population. The exploratory analysis in the following section highlights areas that could improve generalizability of results and provides possible characteristics to look for in recruitment for any follow on up expanded experiments.

Analysis and Results

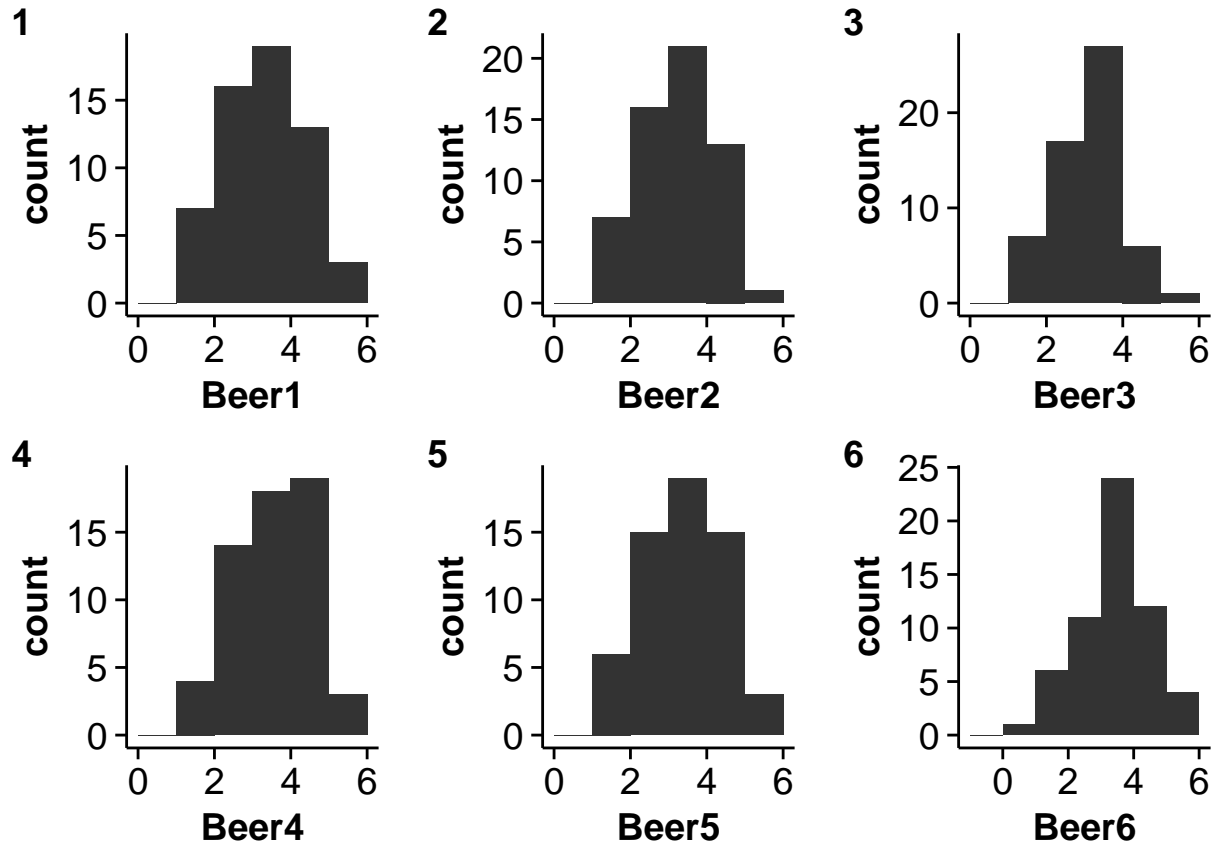
Variable Descriptions and Distribution

After running the full experiment, we start the full analysis by examining some variable distributions and performing some exploratory data analysis. [Note: the full `r` code is available in the R-markdown version of this document.]

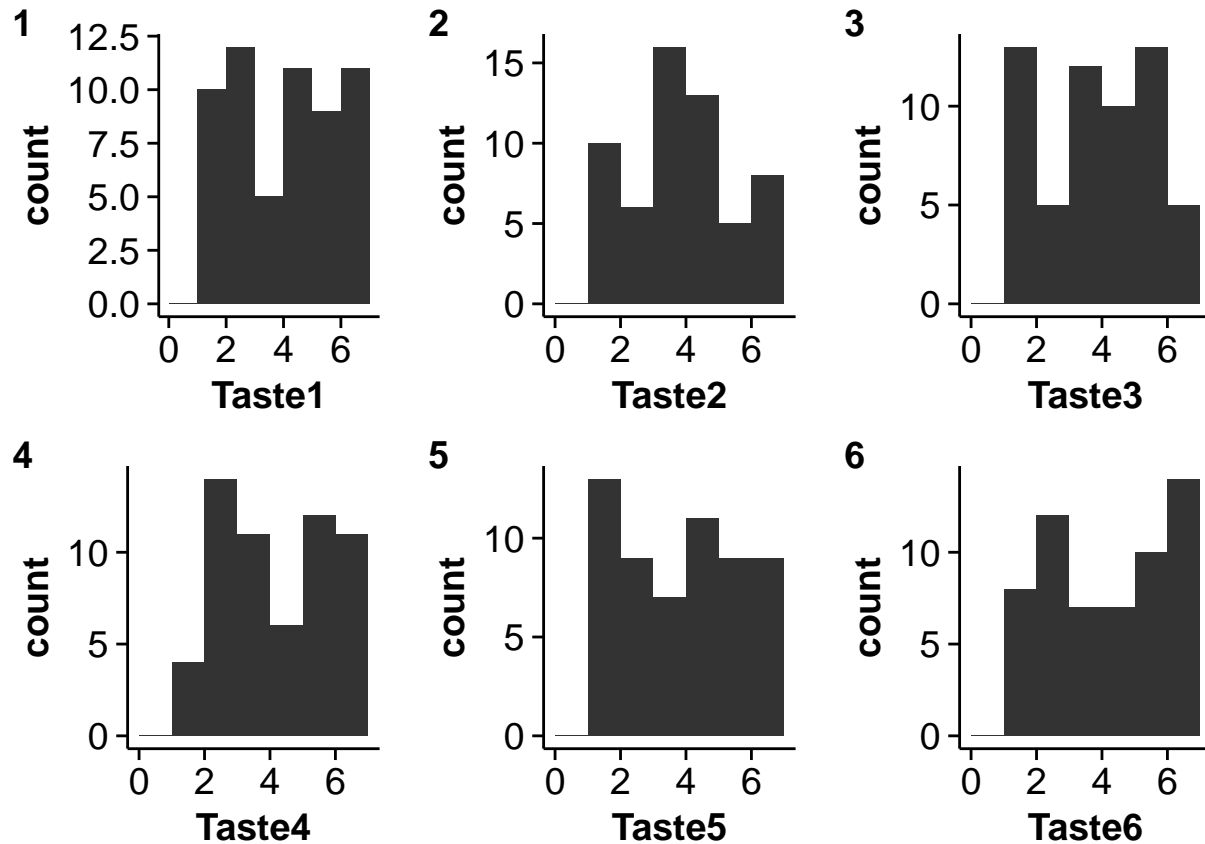
In looking at the histograms of *BeerAficionado*, *Age*, and *Gender*, we see some possible areas of improvement for recruitment if we were to expand this study. The *BeerAficionado* has a mean of 5.52 which is pretty close to the middle of the scale. This distribution is a little heavy on the low end of the scale but overall appears reasonably normal. On the other hand, the *Age* variable is positively skewed, so in any follow on or additional studies, a wider range of age's should be recruited. We also notice that we have more males in the study than females, so additional work to recruit females should be included in any follow on work.



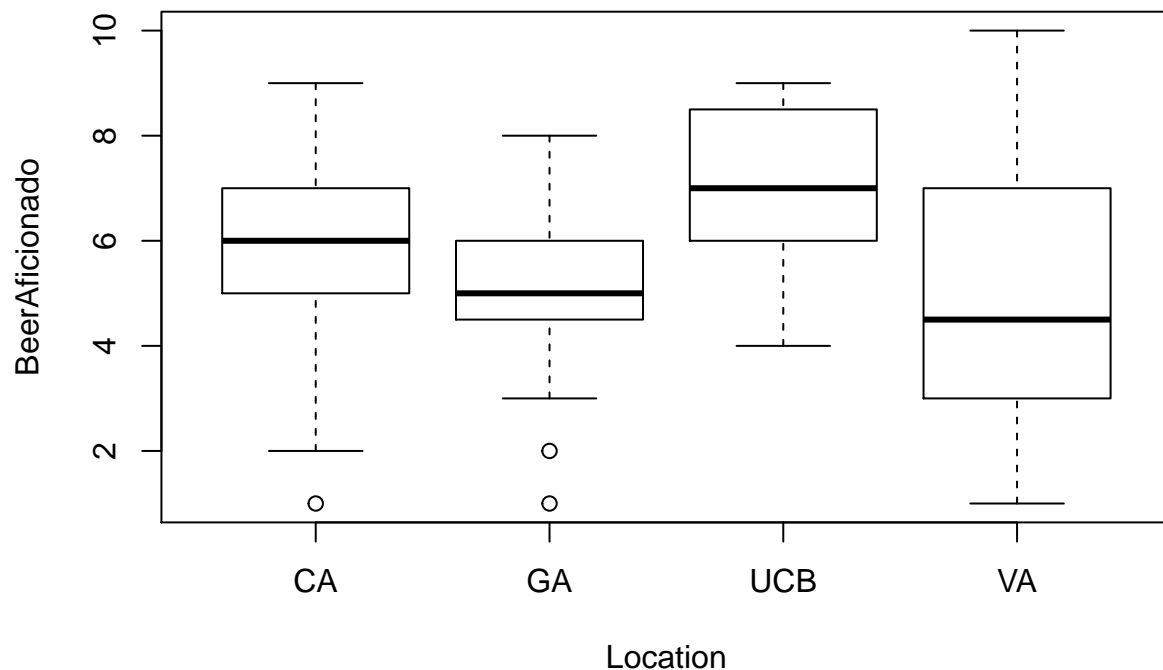
Next, exploring the histograms of the beer ratings overall can provide some initial impressions of what results we might expect. Most of the histograms have the highest number of ratings at 3 except for *Beer4* (the No Sun/Can beer), which is rated higher. Additionally, beers 1-3 have more observations in the lower ratings. These observations may indicate there is a preference for cans over bottles. We will explore this idea in the regression analysis later on.



The last set of histograms depict the tasting order of the beers. Here each bin is a different beer following Table 2 from the experimental design. Since the order was randomly assigned we do not expect the distributions to be normal, but instead would expect more even or constant frequencies. The plots below indicate that overall the random assignment performed ok, but there are some that are less evenly distributed. For instance, *Taste2* was either *Beer3* (30 min Sun/Bottle) or *Beer4* (No Sun/Can) for about half of the participants. In future experiments, a larger sample size should help even out these distributions and decrease some of these spikes we notice in the plots below.

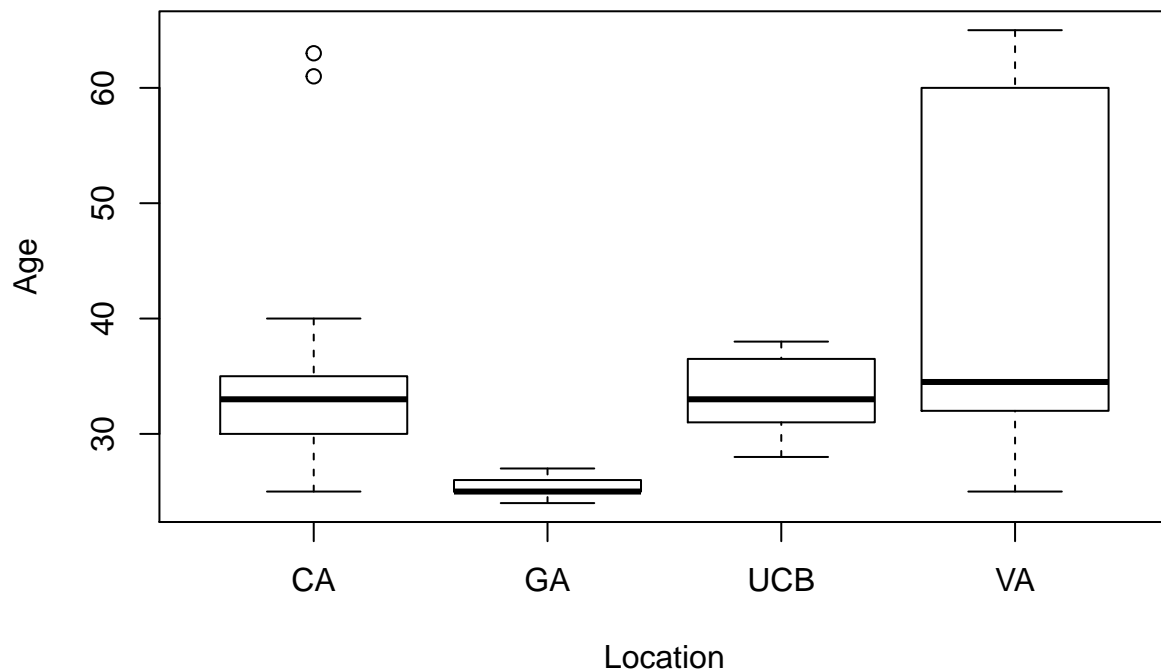


Additional exploratory analysis looked at state or location level differences. By looking at how the study populations broke down across states, we hoped to identify interesting differences that would help us understand our results. We first looked at beer aficionado score because it is intuitive that this variable will correlate with how well individuals can determine the difference between non-skunked and skunked beer. People who self-describe as knowing more about beer are more likely to discern between different beer tastes. We notice that the UCB MIDS faculty had the highest mean beer aficionado score, the Georgia group had the lowest mean and smallest variance, and the Virginia group had the widest range. An ANOVA test is run on this variable to determine if there is statistically significant differences in the variance. The ANOVA results indicate no statistical significance. In follow on experiments, recruitment of a wider spectrum of beer aficionados may decrease the between group variation while increasing the within group variation.



```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Location    3  28.96   9.653    1.92  0.137
## Residuals  54 271.52   5.028
```

Similarly, *Age* differences by location were examined since we noticed a positive skew in the age histogram from above and intuitively, we may expect differences in age to cause differences in the ability to discern beer skunkedness. We can imagine a number of relationships between age and beer tasting ability. Perhaps older people have consumed a greater quantity of beer in their lives and are therefore better at discerning the difference. Or, perhaps, taste buds degrade with age and so older people are less likely to accurately identify skunked beer. While we were not sure of the exact relationship between age and tasting ability, we knew this was a variable important to explore. The box and whisker plot below shows Georgia with the lowest mean Age and the smallest variance and Virginia with the highest mean and largest variance. However, the means from California, MIDS Faculty, and Virginia are close. An ANOVA test resulted in a statistically significant p-value indicating we can reject the null hypothesis that the mean *Age* is consistent across locations. As mentioned above in regards to the *Age* histogram, a wider variety of ages in each location should be recruited for future efforts.



```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Location    3   1704    568.0   6.152 0.00112 **
## Residuals  54   4985     92.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Main Hypothesis Testing

We have a within subject design experiment. This means we have to fulfill two additional assumptions, no-anticipation and no-persistence. These were taken into account in the experimental design, but we verified them using the study data. The no anticipation assumption is fulfilled by the experimental design by randomizing the order of tastings for each individual. The no-persistence assumption is addressed by requiring participants to eat a cracker between each tasting. In order to verify these with the experiment data, we conducted both individual analysis and group level analysis to see if the skunked beer influenced the ratings of the beers after it was tasted, or if there was any spillover of the bad taste into subsequent samples. We did not find any statistically significant differences on the individual or group levels and thereby fulfill the no-persistence assumption.

In order to illustrate the no-persistence assumption, we examine both within subject ratings and the overall average ratings before and after the skunked beer. For this analysis, we start by identifying the location of the fully skunked beer (*Beer3*). For each individual we compare the ratings of the beers they tasted prior to this beer and those they tasted after this beer. Note: those participants who tasted *Beer3* first or last will have NA values for this analysis. An additional dummy variable *Lower* is created with a result of 1 if the ratings are lower after the skunked beer and 0 otherwise. Without any skunked beer persistence, we would expect 50% of the individuals to rate the beers lower, and we can see in the analysis below that about 43% ranked them lower, not a significant deviation from what we expected. Additionally, a t-test was conducted on the mean of the ratings prior to skunked beer and the mean after the skunked beer. The results indicate no significant difference with a p-value of 0.3. These two tests demonstrate there was no persistence of the skunked beer on future tastings with one cracker eaten between tastes.

```
#Now to see the results
```

```
#The mean will give us the average of those who rated the beer after the skunked beer lower  
mean(d$Lower, na.rm=TRUE)
```

```
## [1] 0.4347826
```

```
#the t-test will compare the muPrior and muPost to test if the means are statistically different  
t.test(d$muPrior, d$muPost)
```

```
##  
## Welch Two Sample t-test  
##  
## data: d$muPrior and d$muPost  
## t = -1.034, df = 101.088, p-value = 0.3036  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.4961092 0.1561290  
## sample estimates:  
## mean of x mean of y  
## 2.693082 2.863072
```

In addition to the no-persistence assumption and tests run above, we also checked for first-taste bias to see if the first beer that was tasted for each individual was rated the highest. This is also to show there was no anticipation and no persistence. After breaking out the data into separate dataframe slices for each beer where it was tasted first and then not tasted first, we ran a t-test for each beer to see if there was a statistically significant change in the average rating if it was tasted first. The t-test results show no significant p-values, indicating there was no first-taste bias in our data.

With the above analysis indicating no-persistence and no-anticipation we then remove the tasting order variables and proceed with the regression analysis. The first thing we have to do is reorganize the data

by breaking out the beer types into the binary variables *bottle*, *Sun10*, and *Sun30*. We also added in the continuous variable *Expose* with values of 0, 10, 30 to indicate the sun exposure. However, we do not use this continuous variable in subsequent analysis because we can be more explicit with the binary variables.

We can start by running an ANOVA on the beer ratings by beer to compare the means, which indicates statistically significant results. However, the ANOVA analysis does not identify which beer ratings differ or any possible correlated covariates.

```
# let's do an anova test to compare the Beer tests
beers <- aov(rating ~ beer, data=d2)
summary(beers)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## beer           9  274.2   30.471    28.61 <2e-16 ***
## Residuals     546   581.4     1.065
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 24 observations deleted due to missingness
```

In order to explore the possible causal relationships, we proceed with regression analysis. We conducted a linear model with results shown in the table below. The regression is focused on the sun exposure and vessel type (bottle versus can) with their interaction terms, and described by the following equation. The first model we run without leveraging the within subject design. The second model takes advantage of the within subject design by using fixed effects for each individual.

$$Rating = \beta_0 + \beta_1 Bottle + \beta_2 Sun10 + \beta_3 Sun30 + \beta_4 Bottle \cdot Sun10 + \beta_5 Bottle \cdot Sun30$$

```
#run the baseline model
model0<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30, data=d2)
#Now we cluster by individual to leverage the within subject design
model1<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30 + factor(id), data=d2)
```

```
#run the clustered se fucntion on the model
model1c<-cl(model1, d2$id)
```

% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
 % Date and time: Fri, Apr 29, 2016 - 10:00:02 AM

From the table we can see that the only coefficient that is significant is Bottle. This indicates that on average participants were not able to distinguish the difference between skunked beer and non-skunked beer. However, on average participants preferred the beer in cans over beer in bottles. The rating difference between canned and bottled beer was about half a point on the scale.

Table 7:

	<i>Dependent variable:</i>	
	rating	
	(1)	(2)
Bottle	−0.549*** (0.178)	−0.505*** (0.169)
Sun10	−0.155 (0.229)	−0.155 (0.216)
Sun30	−0.155 (0.229)	−0.155 (0.216)
Bottle:Sun10	0.394 (0.290)	0.350 (0.274)
Bottle:Sun30	0.256 (0.290)	0.212 (0.274)
Constant	3.052*** (0.162)	3.034*** (0.396)
Individual Fixed effects	No	Yes
Robust SE	1.037	0.884

Impressions of Results

Based on our regression results we are slightly skeptical of the reported outcomes and would suggest repeating the experiment with a wider recruitment base and larger sample size. We are not as surprised by the container preference because the three tastings of Corona from cans should all have tasted un-skunked, so we would expect that overall the cans would be rated higher. Along the same lines, we are not surprised to see no statistically significant effect on just the *Sun10* or *Sun30* coefficients because those dummy variables include both the bottles and cans, of which half should have tasted un-skunked. We are skeptical of the interaction term results. We would expect to see a negative coefficient on the interaction terms to indicate that the bottle * Sun effects would produce lower results. We are surprised this effect is not present because while running the experiment several participants remarked at least one beer had a very unpleasant taste immediately when tasting. It was almost a gut reaction of participants to just put it down and eat a cracker. Overall, the results are interesting but could be more generalizable with some of the recruitment and experimental modifications mentioned throughout the rest of the paper.

Additional Analysis

We already calculated if people as a whole could tell the difference between skunked and non-skunked beer. We failed to find that people could tell such a difference. This additional analysis is not to identify causal relationships but to explore possible avenues for future work and tease out possible modifications to this experiment. We also wanted to explore the possibility that in some location people actually could tell the difference between the beers. We worried that these people might be obscured in noisy data from all the locations. To combat this worry, we created a number of variables: (1) semi-sensitive, (2) sensitive, (3) samsies, and (4) semi-samsies. Semi-sensitive identified people could tell the difference between the non-skunked beer and the heavily skunked beer. Sensitive people could identify the difference between non-skunked beer and lightly skunked beer. Samsies people rated all 3 canned beers as the same. Semi-samsies people rated 2 of the 3 canned beers as the same. People who fell into these categories could identify the differences between beers or could accurately tell that all canned beer tasted the same. We tested using an ANOVA, as a simple measure, to see if one location had more of these types of people. We did not find a difference for these variables between any location.

```
# let's see what portion of people are semi-sensitive in each location
```

```
mean(d$semi_sensitive[d$Location=='VA'])
```

```
## [1] 0.5
```

```
mean(d$semi_sensitive[d$Location=='CA'])
```

```
## [1] 0.3333333
```

```
mean(d$semi_sensitive[d$Location=='GA'])
```

```
## [1] 0.5333333
```

```
mean(d$semi_sensitive[d$Location=='UCB'])
```

```
## [1] 0.25
```

We conduct an ANOVA test to test for differences between means by location and we see....

```
# let's wrap it up with an anova test for the difference between means
semi_sensitive_diff <- aov(semi_sensitive ~ Location, data = d)
summary(semi_sensitive_diff)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Location      3  0.669  0.2230   0.899  0.448
## Residuals    54 13.400  0.2482
```

Now we repeat this for sensitive people (people who can tell the difference between non-skunked and lightly skunked beer.

```
# let's see what portion of people are sensitive to skunk in each location
mean(d$sensitive[d$Location=='VA'])
```

```
## [1] 0.5
```

```
mean(d$sensitive[d$Location=='CA'])
```

```
## [1] 0.1428571
```

```
mean(d$sensitive[d$Location=='GA'])
```

```
## [1] 0.4
```

```
mean(d$sensitive[d$Location=='UCB'])
```

```
## [1] 0.375
```

```
# let's wrap it up with an anova test for the difference between means by state
sensitive_diff <- aov(sensitive ~ Location, data = d)
summary(sensitive_diff)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Location      3  1.229  0.4098   1.917  0.138
## Residuals    54 11.546  0.2138
```

Additionally we conducted some testing to see if individuals identify that all three cans were the same, or even if they could identify that two of the cans were the same.

```
# let's compare the means
samsies_diff <- aov(samsies ~ Location, data = d)
summary(samsies_diff)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Location      3  0.599  0.1996   1.538  0.215
## Residuals    54  7.005  0.1297
```

```
# let's compare the means
semi_samsies_diff <- aov(semi_samsies ~ Location, data = d)
summary(semi_samsies_diff)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Location      3  0.186  0.06207    0.294   0.83
## Residuals    54 11.400  0.21111
```

The results are not statistically significant which is not surprising because previous experiments have been performed showing that participants have a difficult time determining the samples are the same.

To increase statistical power in detecting differences by location, we can conduct some additional exploration using regression analysis. We used the overall casual regression equation and added in state fixed effects instead of individual fixed effects. In order to tease out all of the different state comparisons, we conducted the model with each state as the base case. As we can see in the table below, Georgia is statistically significant when compared to California and weakly significant when compared to Virginia. This indicates that there may be something different in Georgia than the other locations. Further analysis would be needed to tease out exactly what is leading to this state difference, but based on our initial histograms and data exploration it is may be due to the slightly skewed population from the Georgia cluster.

```
#In order to cycle through the states as the base case we set Location
#as a factor and then relevel between each model
class(d2$Location)
```

```
## [1] "factor"
```

```
d2$Location<-relevel(d2$Location, 1)
modelCA<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30 +
            factor(Location), data=d2)

d2$Location<-relevel(d2$Location, 2)
modelGA<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30 +
            factor(Location), data=d2)

d2$Location<-relevel(d2$Location, 3)
modelUCB<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30 +
            factor(Location), data=d2)

d2$Location<-relevel(d2$Location, 4)
modelVA<-lm(rating~Bottle + Sun10 + Sun30 + Bottle*Sun10 + Bottle*Sun30 +
            factor(Location), data=d2)
```

```
% Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Fri, Apr 29, 2016 - 10:00:03 AM
```


Table 8:

	<i>Dependent variable:</i>				
	Base Location				
		CA	GA	VA	UCB
	(1)	(2)	(3)	(4)	(5)
Bottle	−0.505*** (0.169)	−0.540*** (0.177)	−0.540*** (0.177)	−0.540*** (0.177)	−0.540*** (0.177)
Sun10	−0.155 (0.216)	−0.155 (0.227)	−0.155 (0.227)	−0.155 (0.227)	−0.155 (0.227)
Sun30	−0.155 (0.216)	−0.155 (0.227)	−0.155 (0.227)	−0.155 (0.227)	−0.155 (0.227)
factor(Location)GA		−0.403*** (0.133)		−0.247* (0.145)	−0.228 (0.174)
factor(Location)CA			0.403*** (0.133)	0.156 (0.137)	0.176 (0.167)
factor(Location)UCB		−0.176 (0.167)	0.228 (0.174)	−0.019 (0.177)	
factor(Location)VA		−0.156 (0.137)	0.247* (0.145)		0.019 (0.177)
Bottle:Sun10	0.350 (0.274)	0.385 (0.288)	0.385 (0.288)	0.385 (0.288)	0.385 (0.288)
Bottle:Sun30	0.212 (0.274)	0.247 (0.288)	0.247 (0.288)	0.247 (0.288)	0.247 (0.288)
Constant	3.034*** (0.396)	3.218*** (0.175)	2.815*** (0.183)	3.062*** (0.185)	3.042*** (0.208)
Individual Fixed effects	Yes	No	No	No	No
State Fixed effects	No	Yes	Yes	Yes	Yes
Robust SE	0.884	1.022	1.022	1.022	1.022

Conclusions and Future Work

Overall, our results are mixed, but we identified several areas where improvements could be made for follow on experiments.

We fail to reject our first hypothesis, that individuals rate beer that has not been exposed to sun and beer that has been exposed to sun the same. On average participants were not able to distinguish the difference between skunked and non-skunked Corona beer. We consider that this might be due to a general dislike for Corona. The science of skunked beer clearly shows it should taste worse than non-skunked beer. One would think this “skunky” taste would lead people to rate it lower compared to non-skunked beer. At this time, our results do not support the science. There are a few issues related to our study that could affect the results: small sample size, not enough locations, too narrow age range and only a single brand of beer. In the future we would like to expand the size of the study to a more representative sample of geographic locations and ages. Besides some recruitment changes for further study, alternative beer brands could be tested because our results could indicate that people cannot tell the difference between skunked and non-skunked Corona and results may differ for other beers. Additionally, the experimental design could be simplified to just include the no sun exposure and the fully skunked beer in case there is tasting fatigue—where all the beers start to taste the same.

The second hypothesis yielded more interesting results and a possible causal relationship. The within subject designed linear regression yielded a statistically significant coefficient for the vessel type; therefore, we can reject our second hypothesis that individuals rate beer that is stored in a bottle and beer that is stored in a can the same. The experimental design intervention and randomization accounts for other possible covariates thereby isolating the beer skunking and beer vessel as the potential causal effects for the beer ratings. The significance of the vessel coefficient leads us to the preliminary conclusion that individuals can taste the difference between cans and bottle and on average prefer Corona from cans. The experimental recruitment improvements mentioned previously are also applicable to this experiment for further study. Additionally, it would be interesting to run a follow on experiment just focusing on can versus bottle taste distinction where participants have two samples and are asked to pick their preference.

Appendix: Beer Tasting Survey

Survey

Beer Tasting

Please answer all of the following questions

1. Age: _____
2. Gender: _____
3. Marital Status: Single / Married / Divorced / Widowed
4. Do you consider yourself a beer drinker or beer aficionado on a scale of 1-10?

Not a Beer Drinker									Beer aficionado
1	2	3	4	5	6	7	8	9	10

Rank the taste of each beer sample

	Extremely Dislike	Dislike	Its Ok	Like	Really Enjoy	Drinking order
Beer A	1	2	3	4	5	
Beer B	1	2	3	4	5	
Beer C	1	2	3	4	5	
Beer D	1	2	3	4	5	
Beer E	1	2	3	4	5	
Beer F	1	2	3	4	5	

Figure 1: Beer Tasting Survey