This report is an excerpt of my part to the overall market intelligence report on the Ice Cool Brand in Singapore. The marketing research problems, survey, and conclusion were conducted by different team members. I did not include their part. I only include the part I fully contributed to.

3.4 Data Analysis

3.4.1 Summary

This section reveals the graphical, numerical, and statistical methods used to analyze the survey data. Please refer to Appendix section 5.1 Acknowledgements and Data Collection for more details related to the background of the survey data and its usage in this analysis. Please refer to section 5.2 Variables for explanations related to each created variable used in the analysis. Please refer to sections 5.3-5.5 for the graphical and statistical representations used in this section.

3.4.2 Correlation Results

The following are moderate to strong correlation results for each relevant variable. Weak correlation results are not presented in this report. Please refer to 5.4.4 Significance Testing Results Analysis, which builds on top of the following results.

Strong correlation between price_importantce and discount_importance is 0.7343963. This result is most notable as it may prove that the importance of price of a beverage of the individual correlates with the discount or promotional sale of a product. Please refer to 3.4.5 for a deep dive into regression analysis between variables.

Moderate correlation between flitered_brandswitch and filtered_unique_values is -0.5784075. This result implies that there is a negative correlation between individuals who switch brands and have unique brands with Ice Cool. This result solidifies a deeper understanding that the respondents with no unique values with Ice Cool are truly willing to switch brands. This is true

because many respondents said they are willing to switch brands if another brand were on sale. Thus, given the opportunity, respondents with no unique value given to Ice Cool would switch to another brand, hence the negative correlation result.

Moderate correlation between filtered_marketing and filtered_unique_values is 0.5270463. This result implies that there is a positive correlation between respondents who do not see Ice Cool marketing advertisements and have unique brands with Ice Cool. As the majority of the sample size for filtered_marketing does not observe any marketing advertisements from Ice Cool, it is possible to say that those with respondents do not have unique values with Ice Cool.

Moderate correlation between filtered_marketing and age is 0.5447048. This result may seem impressive but cannot be thoroughly explained as much of the survey data is given to one age group. Thus, this moderate correlation, while positive, does not fully describe the effects between age and the observation of Ice Cool marketing.

Moderate correlation between filtered_marketing and occupation is 0.6085247. This result may seem impressive but cannot be thoroughly explained as much of the survey data is given to one occupation group. Thus, this moderate correlation, while positive, does not fully describe the effects between age and the observation of Ice Cool marketing.

Moderate correlation between info_packaging and sweetness is 0.5678971. This result does not provide any meaningful explanation as info_packaging measures the value given to informative packaging when purchasing a beverage like coconut water, while sweetness measures the level of sweetness of Ice Cool coconut water. Thus, this correlation result is not relevant.

Moderate correlation between info_packaging and taste is 0.6325991. This result does not provide any meaningful explanation as info_packaging measures the value given to informative packaging when purchasing a beverage like coconut water, while taste is the average score

between the sweetness, natural_ingre, and with_coconut variables. Thus, this correlation result is not relevant.

Moderate correlation between filtered_unique_values and age is 0.5421152. This result is relevant because it describes that age may play a role in the unique values given to Ice Cool. However, as many of the survey respondents are students, it is unclear if this distinction is valid to understanding other age groups and their unique values with Ice Cool. It is valid to say that students may not have strong values with Ice Cool, as the correlation is moderately positive.

3.4.3 Histograms, Pie Charts, and Other Distribution Methods

Histograms are used to illustrate discrete or continuous data that are measured on an interval scale (Statistics Canada). In this analysis, histograms show the distribution of the survey data. Pie Charts are used to organize and show data as a percentage of a whole (Tableau).

It is clear that price is important to respondents when looking at sections 5.3.4, 5.3.6-9, and 5.3.21.

It is also interesting to find that marketing campaigns and advertisements are influential in consumer decisions. It is clear that Ice Cool does not put its budget into marketing campaigns and advertisements as shown in 5.3.11, where 70% of respondents have not seen Ice Cool advertisements.

It is evident that packaging plays a huge role in purchasing decisions. As shown in 5.3.13, the vast majority of respondents find packaging design as important. It is also clear that Ice Cool packaging is not visually appealing, as the vast majority of respondents find the packaging design visually unappealing, in 5.3.14. Lastly, in 5.3.15, many respondents value informative packaging (e.g., health benefits, nutritional info) when purchasing a beverage like coconut water.

It is apparent the majority of respondents value the importance of natural ingredients in coconut water. It is also interesting that 62.5% of respondents say the presence of real coconut pieces in coconut water positively influences consumer decisions as shown in 5.3.17.

3.4.4 Significance Testing

In this section, t-testing is used to compute the statistical significance of each testable variable. The following list is considered variable good for significance testing in section 5.4.2 Testable Variables. The criteria for significance testing are also listed below in section 5.4.3 Significance Testing Criteria. Please refer to 5.2 Variables for the full description and explanation of each variable. Please refer to explanations on each tested variable in 5.4.4 Significance Testing Results Analysis.

3.4.5 Regression Analysis

This section seeks to understand the relationship between variables. Through multiple linear regression, it is evident which variables from correlation effects on price sensitivity, influential marketing, package design importance, among other characteristics. For analysis purposes, “strong significance” implies the variable is statistically significant at $p < 0.01$, “moderate significance” implies the variable is statistically significant at $p < 0.05$, while “weak significance” implies the variable is statistically significant at $p < 0.1$ (also known as marginal significance). Results can be seen in 5.5 Regression Analysis Results.

3.4.6 Other Regression Methods and Algorithms

A continuation to the previous analysis led to conducting a comprehensive analysis utilizing Support Vector Regression (SVR), Random Forest, Logistic Regression, Classification, and Decision Tree techniques. K-means is a notable regression method with outcomes of low silhouette scores depending on the combination of variables. The outcomes revealed a notable

challenge, as negative R-squared values, and comparatively high Mean Squared Error (MSE) results were consistently observed. These findings suggest the presence of suboptimal clustering results and potential issues in data pre-processing procedures. This is likely due to a small sample size, among other data collection practices. Thus, these methods and results will not be present. It is also evident that the data does not follow a normal distribution, rendering the utilization of advanced statistical methods such as supervised machine learning methods impossible for decision making.

5. Appendices

5.1 Acknowledgements and Data Collection

The conducted survey consists of a variety of questions used for data collection to support the hypotheses in 3.3 Survey. Data collection primarily targets students. The sample size is 40. Many variables are not used in the analysis because the sample size is not representative for each distinct group. The following sections illustrate the Variables, Analysis, and Conclusions.

5.2 Variables

By creating a variable for each question, performing data analysis and statistical testing is efficient and effective. The following list of variables correspond to the order of the survey questions:

- consumption (Question 1)
    - This variable denotes the frequency of consumption related to coconut water. The values given each option response are listed below:
        - 0 denotes individuals who selected the “Never” option
        - 1 denotes individuals who selected the “1-3 times a week” option
        - 2 denotes individuals who selected the “4-6 times a week” option

        - 3 denotes individuals who selected the “More than 7 times a week” option
        - 4 denotes individuals who selected the “Once every two weeks” option
        - 5 denotes individuals who selected the “Less than once a month” option
        - 6 denotes individuals who selected the “Once a month” option
- consumption_reason (Question 2)
    - This variable denotes the reason for consumption related to coconut water. The values given each option response are listed below:
        - 0 denotes individuals who selected the “I don't consume coconut water” option
        - 1 denotes individuals who selected the “Dietary preference” option
        - 2 denotes individuals who selected the “Refreshment” option
        - 3 denotes individuals who selected the “Taste” option
        - 4 denotes individuals who selected the “Health benefits” option
- brand_frequency (Question 3)
    - This variable denotes the number of brands the individual sticks to regarding coconut water. The values given each option response are listed below:
        - 0 denotes individuals who selected the “No preference” option
        - 1 denotes individuals who selected the “Stick to one brand” option
        - 2 denotes individuals who selected the “Try different brands” option
- brand_switching (Question 4)
    - This variable denotes the likelihood of the individual switching from their usual brand when another brand is on sale regarding coconut water. The values given each option response are listed below:

        - 0 denotes individuals who selected the “I don’t have a usual brand” option
        - 1 denotes individuals who selected the “Definitely not” option
        - 2 denotes individuals who selected the “Probably not” option
        - 3 denotes individuals who selected the “Might or might not” option
        - 4 denotes individuals who selected the “Probably” option
        - 5 denotes individuals who selected the “Definitely” option
- usual_brand (Question 5)
    - This variable denotes the factor by which the individual would most likely switch from their usual brand regarding coconut water. The values given each option response are listed below:
        - 0 denotes individuals who selected the “Price reduction/sale” option
        - 1 denotes individuals who selected the “Recommended by a friend or family” option
        - 2 denotes individuals who selected the “Better taste” option
        - 3 denotes individuals who selected the “Health benefits” option
        - 4 denotes individuals who selected the “Packaging appeal” option
        - 5 denotes individuals who selected the “Refreshing” option
- price_importantce (Question 6)
    - This variable denotes the importance of price when making a purchasing decision regarding beverages. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Unimportant” option
        - 2 denotes individuals who selected the “Somewhat Unimportant” option
        - 3 denotes individuals who selected the “Neutral” option

        - 4 denotes individuals who selected the “Somewhat Important” option
        - 5 denotes individuals who selected the “Important” option
- discount_importance (Question 7)
    - This variable denotes the likelihood to purchase a new beverage when on discount or promotion. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Unlikely” option
        - 2 denotes individuals who selected the “Somewhat Unlikely” option
        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Likely” option
        - 5 denotes individuals who selected the “Likely” option
- stu_price_sense (Question 8)
    - This variable denotes the price-sensitive behavior when buying beverages if the individual is a student. The values given each option response are listed below:
        - 0 denotes individuals who selected the “No” option
        - 1 denotes individuals who selected the “Yes” option
        - 2 denotes individuals who selected the “Unsure” option
- first_hear (Question 9)
    - This variable denotes the type of advertisement the individual hears when they first encounter the Ice Cool brand. The values given each option response are listed below:
        - 0 denotes individuals who selected the “I haven't heard of the Ice Cool brand” option
        - 1 denotes individuals who selected the “In-store/shelf presence” option

        - 2 denotes individuals who selected the “Word of mouth” option
        - 3 denotes individuals who selected the “Online advertisement” option
        - 4 denotes individuals who selected the “Social media” option
- familiarity (Question 10)
    - This variable denotes familiarity of the individual with the Ice Cool brand when presented with an image of its main product. The values given each option response are listed below:
        - 0 denotes individuals who selected the “No” option
        - 1 denotes individuals who selected the “Yes” option
        - 2 denotes individuals who selected the “Unsure” option
- marketing (Question 11)
    - This variable denotes the attention the individual gives to marketing campaigns or advertisements from Ice Cool. The values given each option response are listed below:
        - 0 denotes individuals who selected the “No” option
        - 1 denotes individuals who selected the “Yes” option
        - 2 denotes individuals who selected the “Unsure” option
- influence_marketing (Question 12)
    - This variable denotes the level of influence marketing campaigns and advertisements have on the beverage purchasing decisions of the individual. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Not Influential” option
        - 2 denotes individuals who selected the “Somewhat Not Influential” option

        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Influential” option
        - 5 denotes individuals who selected the “Influential” option
- package_design (Question 13)
    - This variable denotes the importance of packaging design. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Unimportant” option
        - 2 denotes individuals who selected the “Somewhat Unimportant” option
        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Important” option
        - 5 denotes individuals who selected the “Important” option
- visua_appeal (Question 14)
    - This variable denotes the level of packaging visual appeal of the Ice Cool product shown in the survey. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Unappealing” option
        - 2 denotes individuals who selected the “Somewhat Unappealing” option
        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Appealing” option
        - 5 denotes individuals who selected the “Appealing” option
- info_packaging (Question 15)
    - This variable denotes the value of informative packaging when the individual purchases a beverage like coconut water. The values given each option response are listed below:

        - 1 denotes individuals who selected the “Not Informative” option
        - 2 denotes individuals who selected the “Somewhat Not Informative” option
        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Informative” option
        - 5 denotes individuals who selected the “Informative” option
- sweetness (Question 16)
    - This variable denotes the level of sweetness the individual associates with Ice Cool coconut water. The values given each option response are listed below:
        - 0 denotes individuals who selected the “Haven't tried it” option
        - 1 denotes individuals who selected the “Lacks sweetness” option
        - 2 denotes individuals who selected the “Not sweet enough” option
        - 3 denotes individuals who selected the “Just right” option
        - 4 denotes individuals who selected the “Sweet” option
        - 5 denotes individuals who selected the “Too sweet” option
- natural_ingre (Question 17)
    - This variable denotes the importance that coconut water is made from natural ingredients. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Unimportant” option
        - 2 denotes individuals who selected the “Somewhat Unimportant” option
        - 3 denotes individuals who selected the “Neutral” option
        - 4 denotes individuals who selected the “Somewhat Important” option
        - 5 denotes individuals who selected the “Important” option

- with_coconut (Question 18)
    - This variable denotes real coconut pieces may influence the decision of the individual to purchase coconut water. The values given each option response are listed below:
        - 1 denotes individuals who selected the “Negatively Influence” option
        - 2 denotes individuals who selected the “No Influence” option
        - 3 denotes individuals who selected the “Positively Influence” option
- first_word (Question 19)
    - This variable denotes an open-ended response from the individual of the first word they think of when thinking about Ice Cool. The values are string data.
- unique_values (Question 20)
    - This variable denotes the association of any unique values or benefits the individual may have with the Ice Cool brand. The values given each option response are listed below:
        - 0 denotes individuals who selected the “No” option
        - 1 denotes individuals who selected the “Yes” option
        - 2 denotes individuals who selected the “Unsure” option
- standout (Question 21)
    - This variable denotes what the individual believes makes a coconut water brand stand out from its competitors. The values are string data.
- age (Question 22)
    - This variable denotes the age group the individual belongs to. The values given each option response are listed below:

        - 0 denotes individuals who selected the “17 or Under” option
        - 1 denotes individuals who selected the “18-24” option
        - 2 denotes individuals who selected the “25-34” option
        - 3 denotes individuals who selected the “35-44” option
        - 4 denotes individuals who selected the “45-54” option
        - 5 denotes individuals who selected the “55-64” option
        - 6 denotes individuals who selected the “65 and over” option
- gender (Question 23)
    - This variable denotes the gender group the individual belongs to. The values given each option response are listed below:
        - 0 denotes individuals who selected the “Male” option
        - 1 denotes individuals who selected the “Female” option
- occupation (Question 24)
    - This variable denotes the occupation group the individual belongs to. The values given each option response are listed below:
        - 0 denotes individuals who selected the “Prefer not to say” option
        - 1 denotes individuals who selected the “Student” option
        - 2 denotes individuals who selected the “Working (On-site)” option
        - 3 denotes individuals who selected the “Working (Hybrid)” option
        - 4 denotes individuals who selected the “Working (Remote)” option
        - 5 denotes individuals who selected the “Unemployed” option
        - 6 denotes individuals who selected the “Retired” option

There are also newly created variables which average out the scores for each individual who answered the survey. Using this approach, it is easier to perform data analysis and conduct further research. The following variables are the newly created variables:

- price_sense
    - This variable denotes the average score for each individual between the price_importance and discount_importance variables.
- brand_aware
    - This variable denotes the average score for each individual between the stu_price_sense, familiarity, and marketing, and influence_marketing variables.
- packaging
    - This variable denotes the average score for each individual between the package_design, visua_appeal, and info_packaging variables.
- taste
    - This variable denotes the average score for each individual between the sweetness, natural_ingre, and with_coconut variables.

Lastly, for accurate representations of variables which provided “Unsure,” “I don’t know,” or similar response options, new filtered variables were created. These variables remove the third option responses from the binary question (“Yes or No” questions). The following are filtered variables:

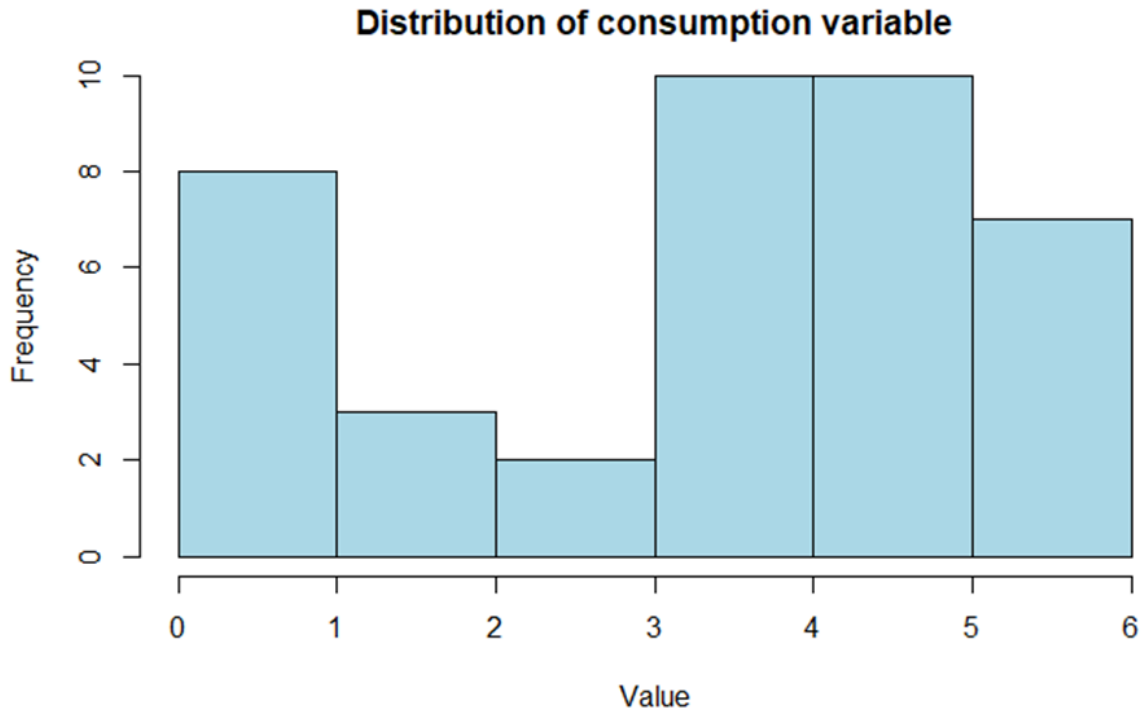
- flitered_brandswitch
    - This variable denotes a filtered version of the brand_switching variable.
- filtered_stu_price_sense
    - This variable denotes a filtered version of the stu_price_sense variable.

- flitered_familiarity
    - This variable denotes a filtered version of the familiarity variable.
- filtered_marketing
    - This variable denotes a filtered version of the marketing variable.
- filtered_sweetness
    - This variable denotes a filtered version of the sweetness variable.
- filtered_unique_values
    - This variable denotes a filtered version of the unique_values variable.
- filtered_brand_aware
    - This variable denotes a filtered version of the brand_aware variable.

5.3 Histograms, Pie Charts, and Other Distribution Methods for Each Variable

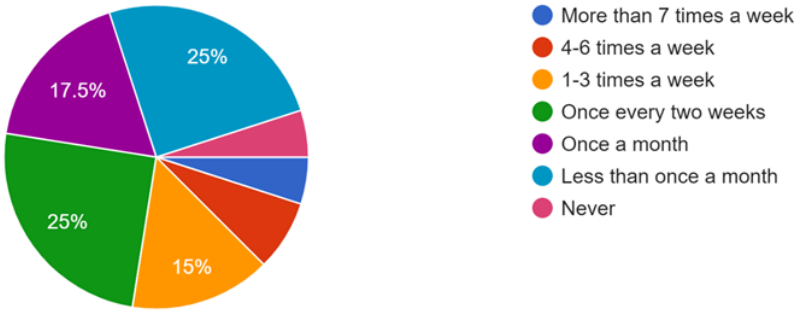
The following section reveals each detailed histogram, pie chart, or graphical method for each variable. Please refer to the 3.4.2 Variables section for variable explanations and documentation.

5.3.1 consumption

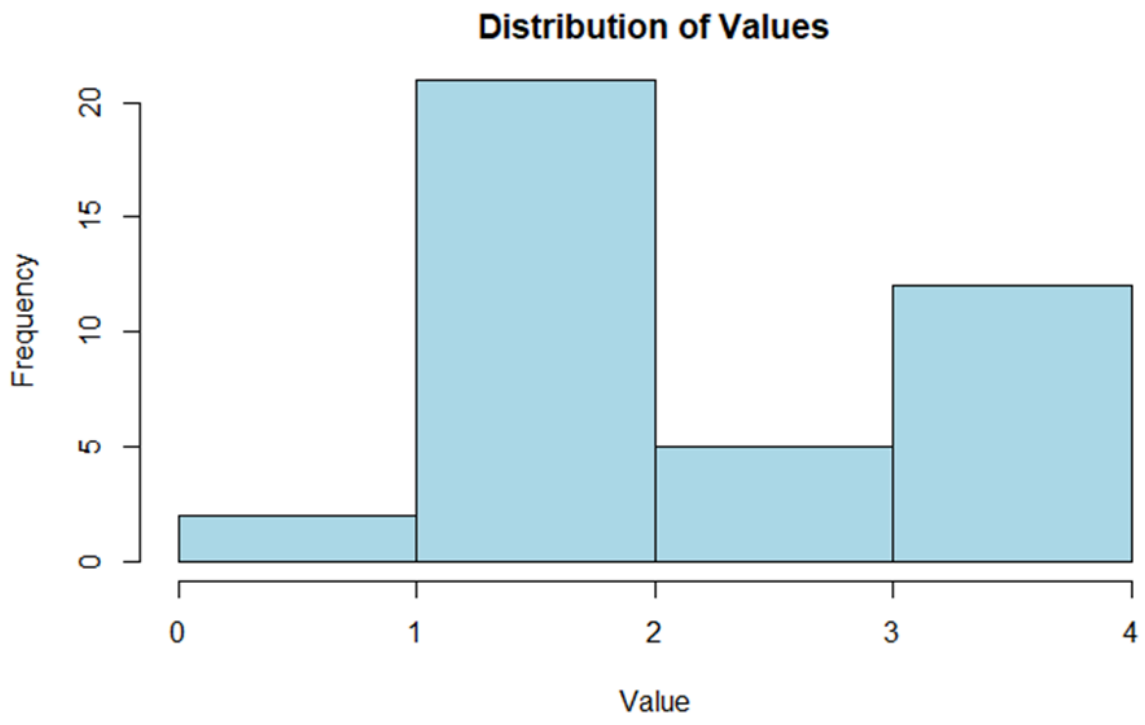


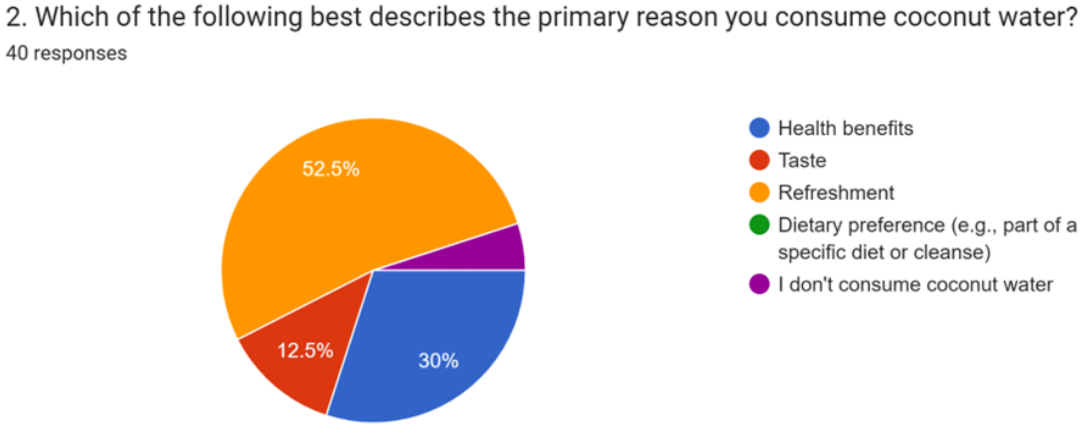
1. How often do you consume coconut water in a typical month?

40 responses

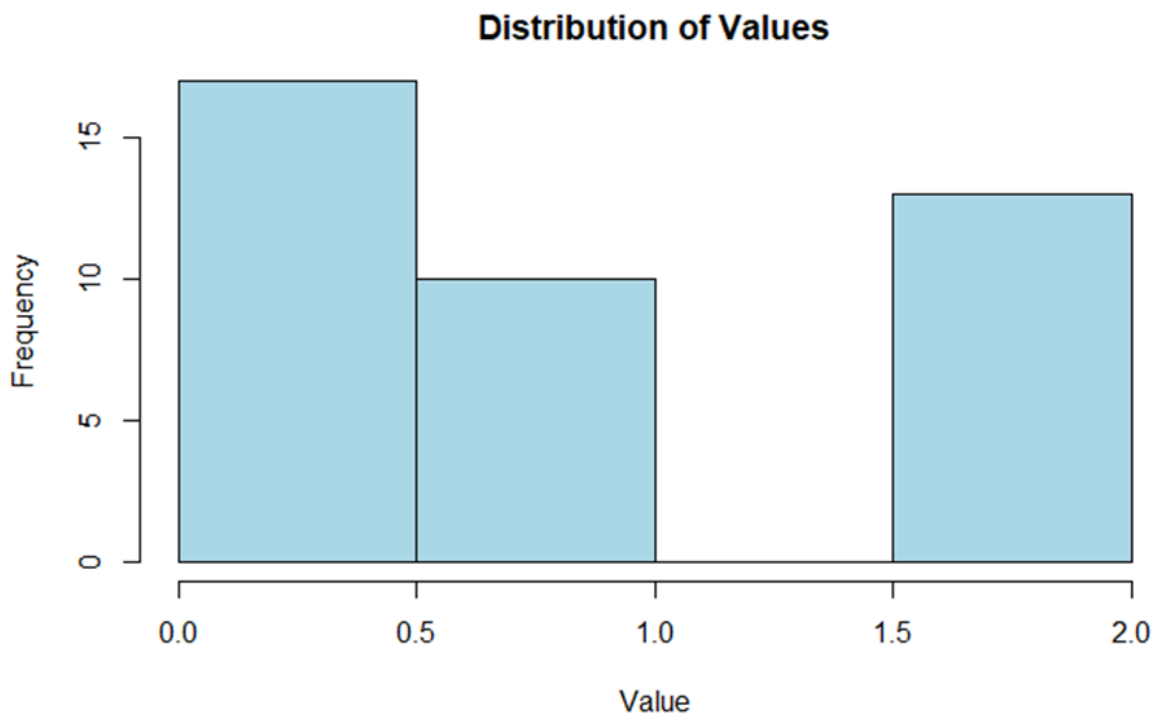


5.3.2 consumption_reason



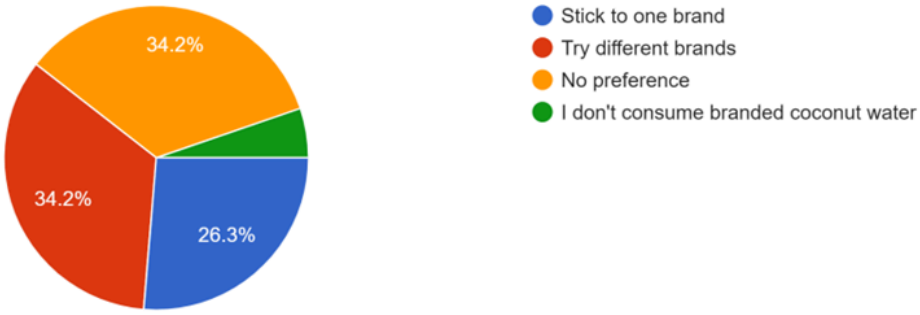


5.3.3 brand_frequency

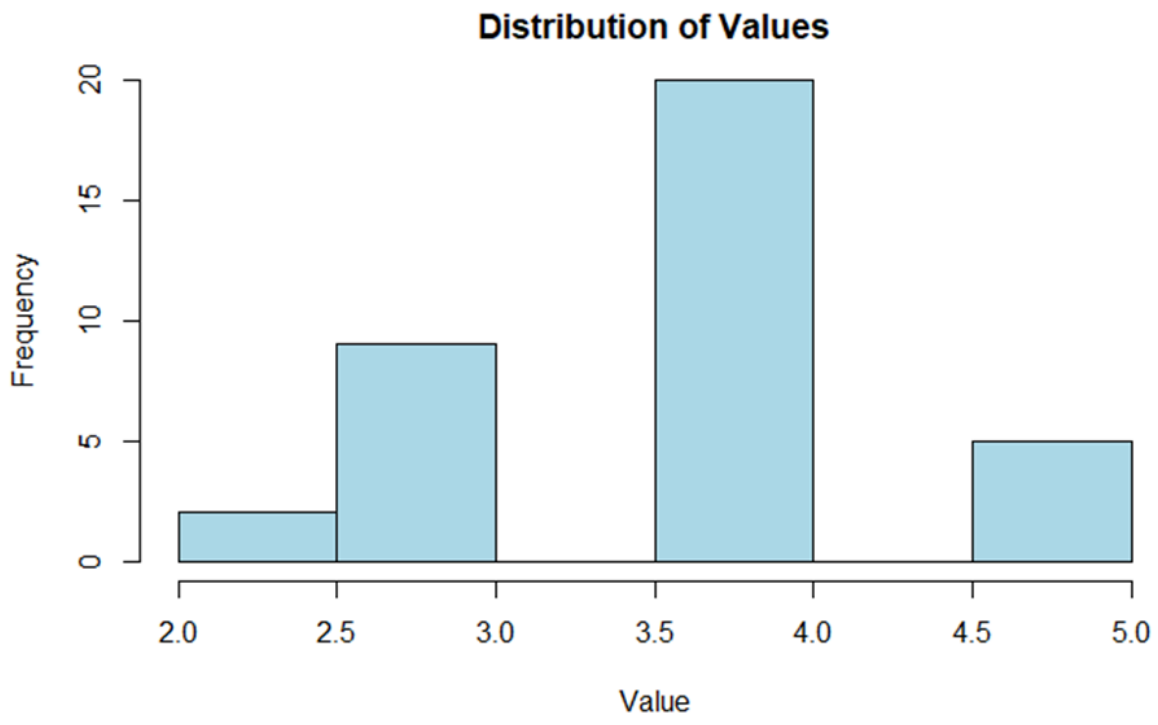


3. Do you usually stick to one brand of coconut water or try different brands?

38 responses

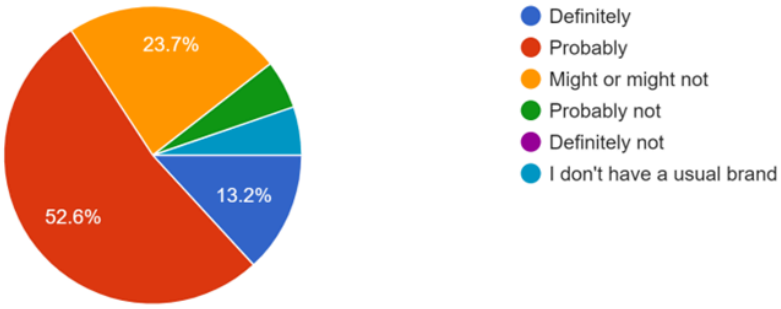


5.3.4 filtered_brand_switching

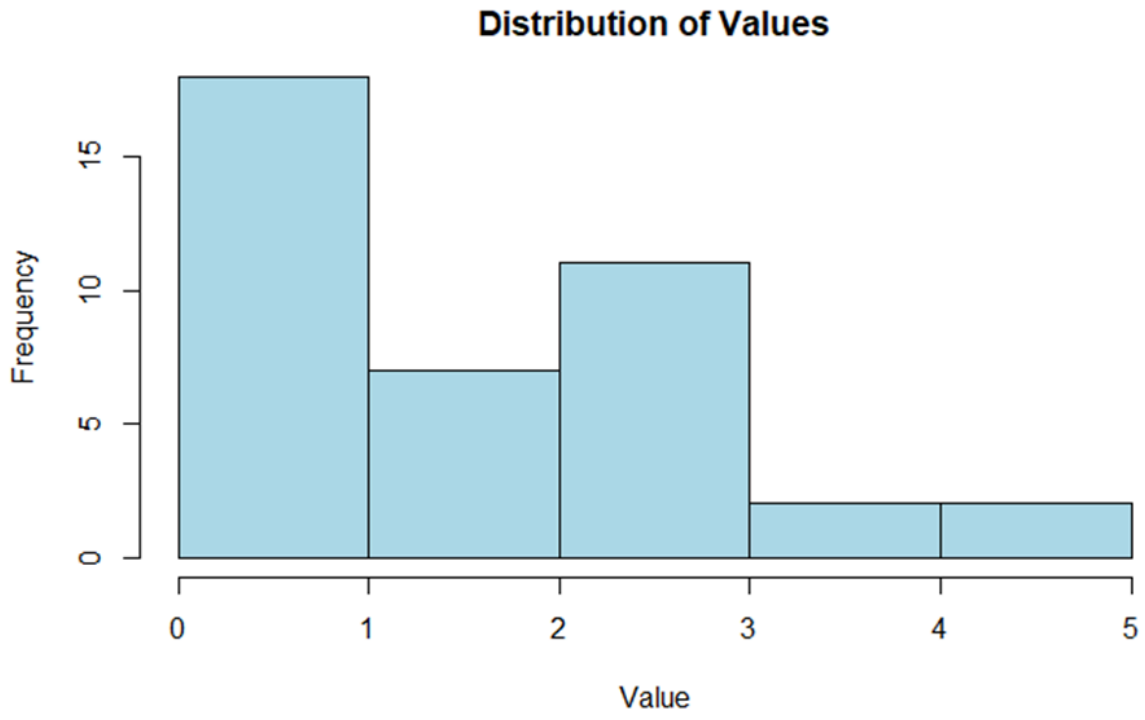


4. If another brand were on sale, would you consider switching from your usual brand?

38 responses

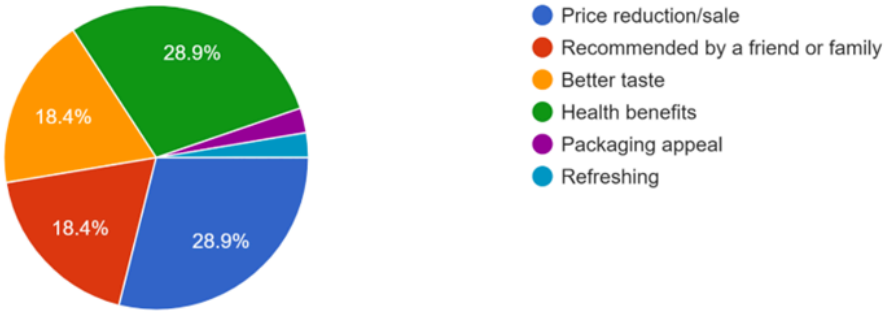


5.3.5 usual_brand

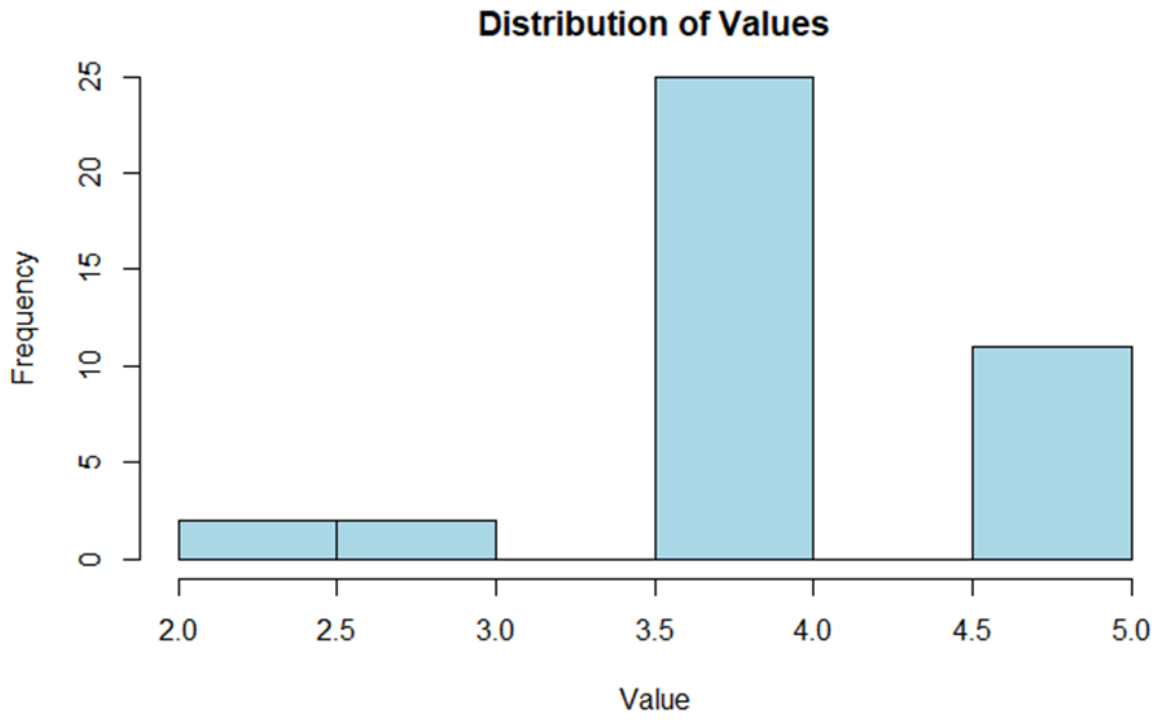


5. Which factor would most likely cause you to switch from your usual coconut water brand to another brand?

38 responses

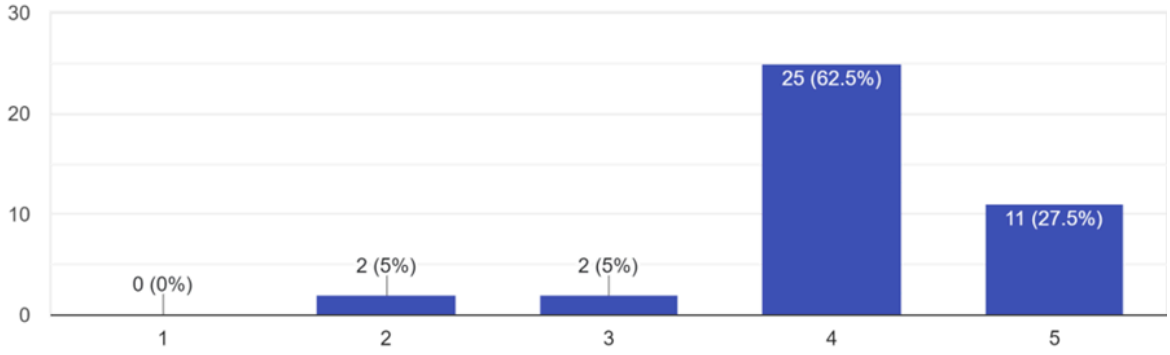


5.3.6 price_importantce

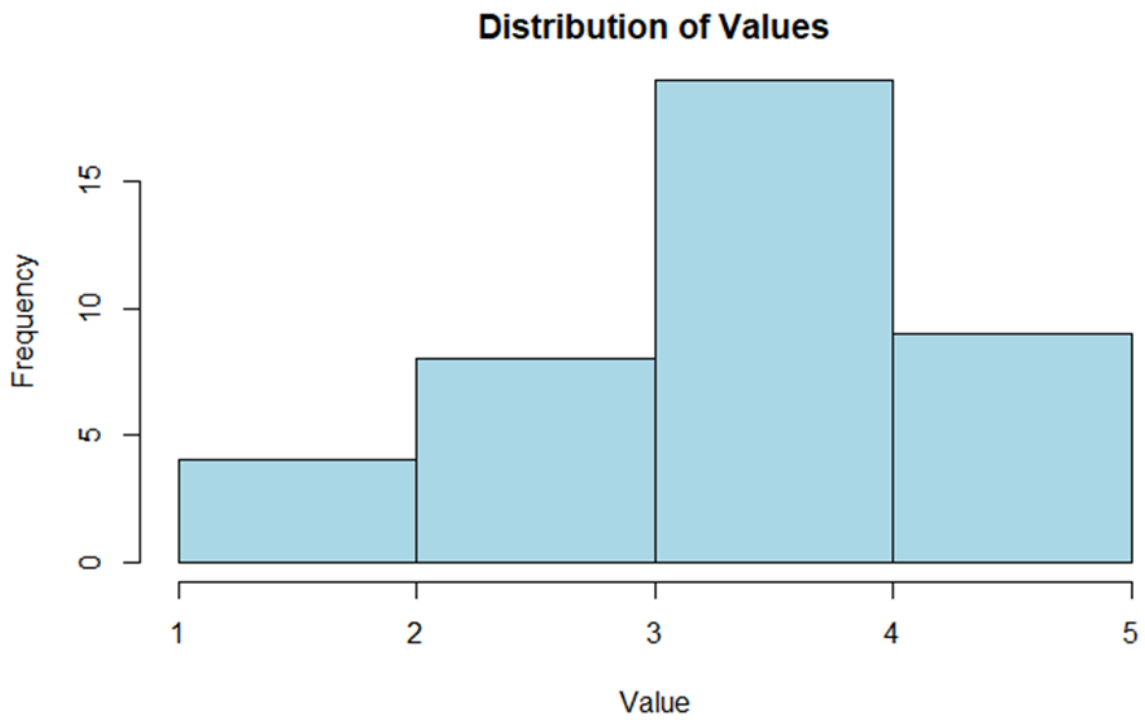


6. On a scale of 1 to 5, how important is the price of a beverage when making a purchasing decision?

40 responses

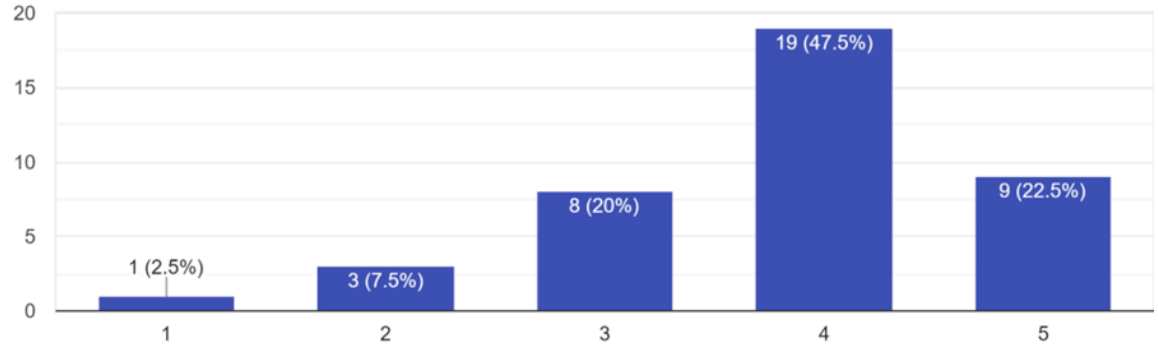


5.3.7 discount_importance

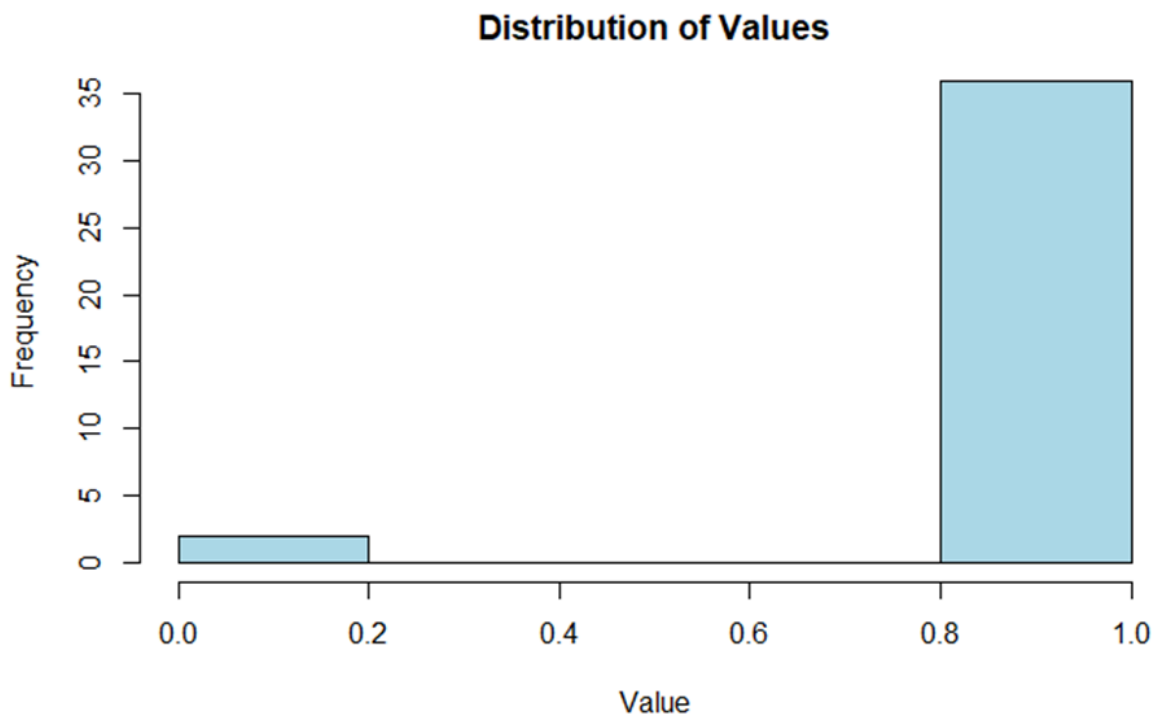


7. On a scale of 1 to 5, how likely are you to purchase a beverage you haven't tried before if there's a discount or promotion?

40 responses

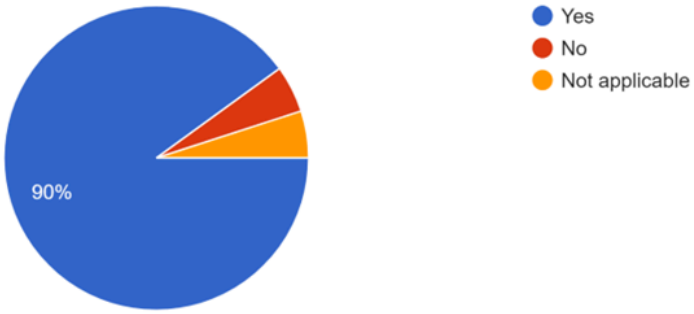


5.3.8 filtered_stu_price_sense

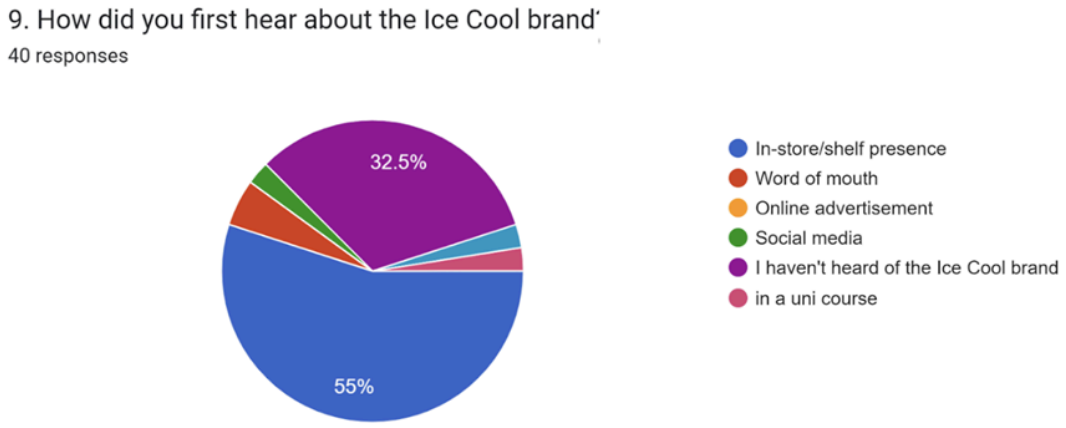


8. If you are a student or have been one recently, do you find yourself being more price-sensitive when buying beverages?

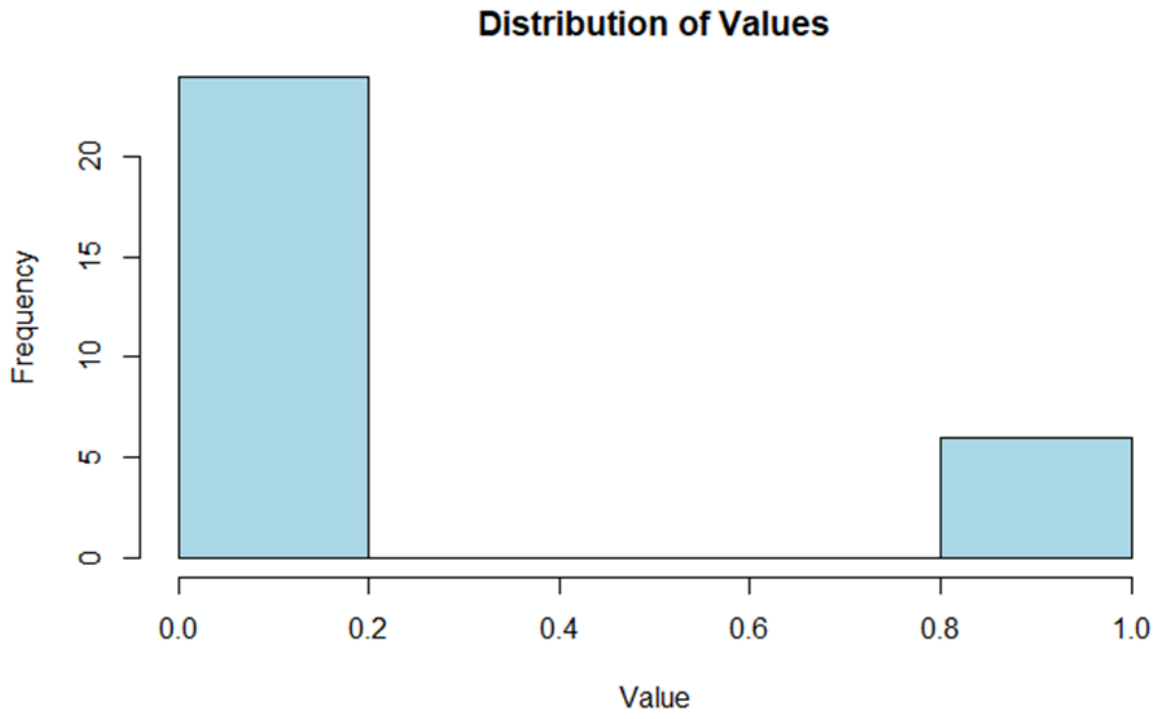
40 responses



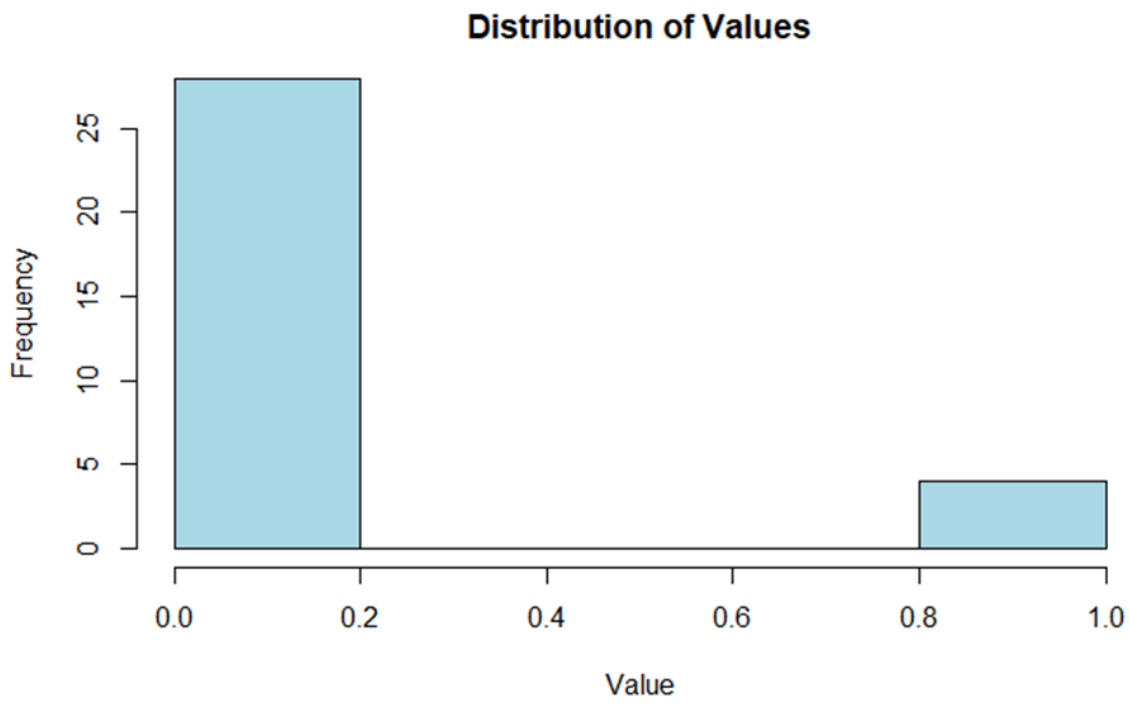
5.3.9 first_hear



5.3.10 filtered_familiarity

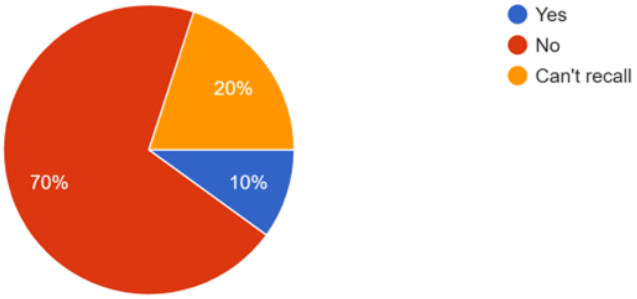


5.3.11 filtered_marketing

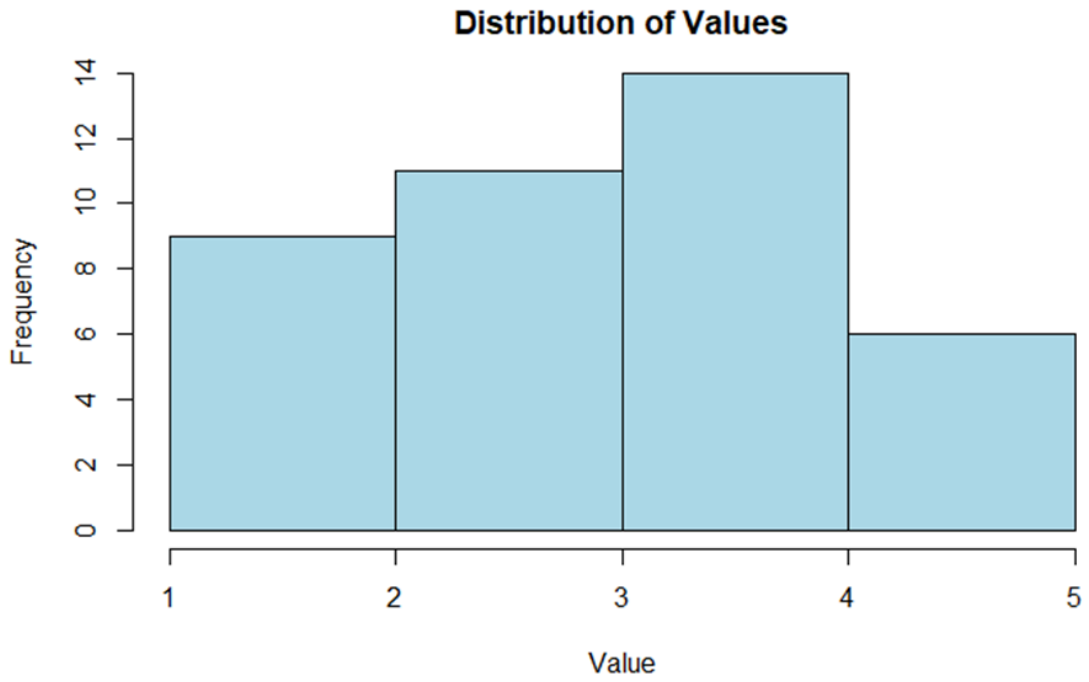


10. Have you ever noticed any marketing campaigns or advertisements from Ice Cool?

40 responses

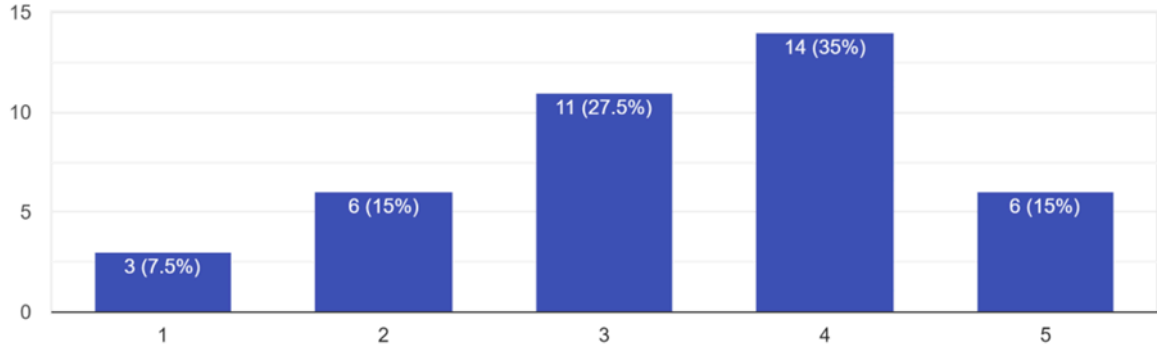


5.3.12 influence_marketing

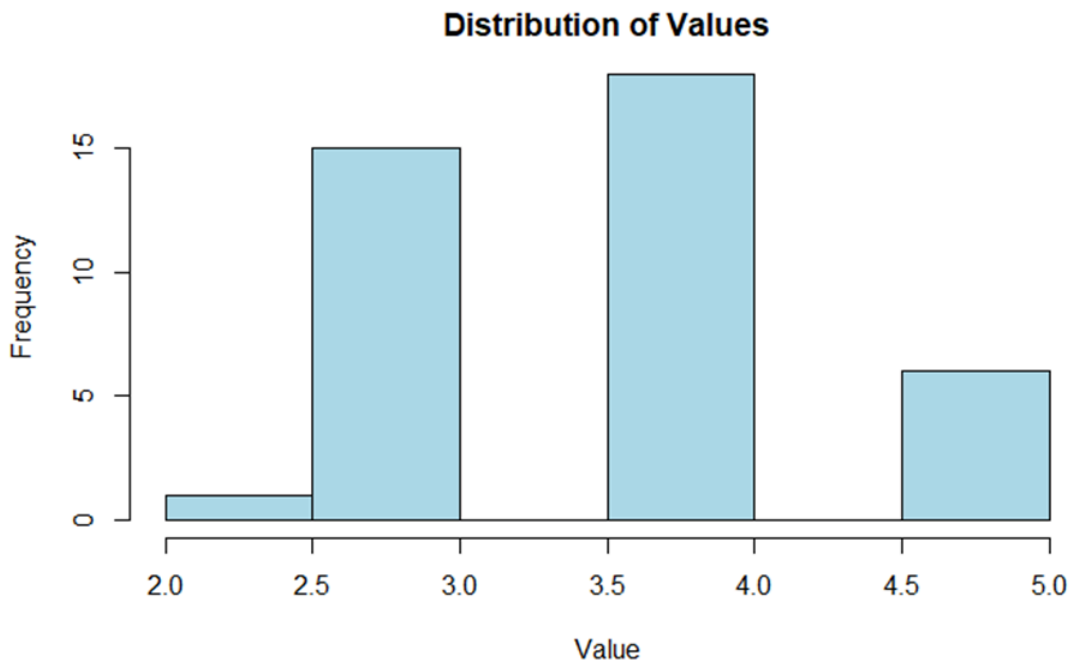


11. On a scale of 1 to 5, how influential are marketing campaigns and advertisements in your beverage purchasing decisions?

40 responses

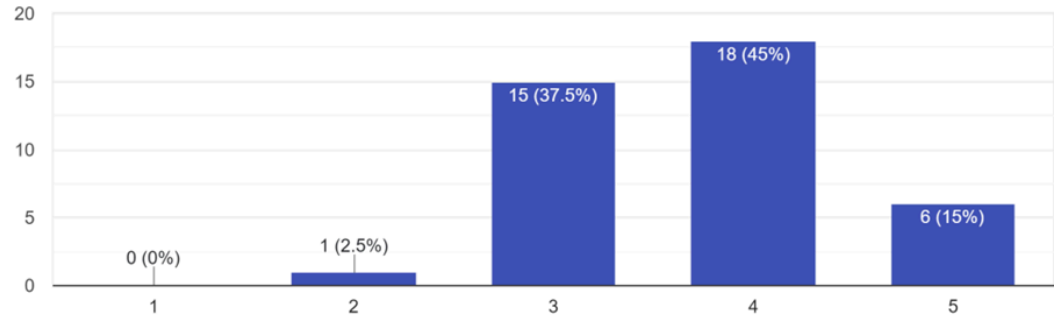


5.3.13 package_design

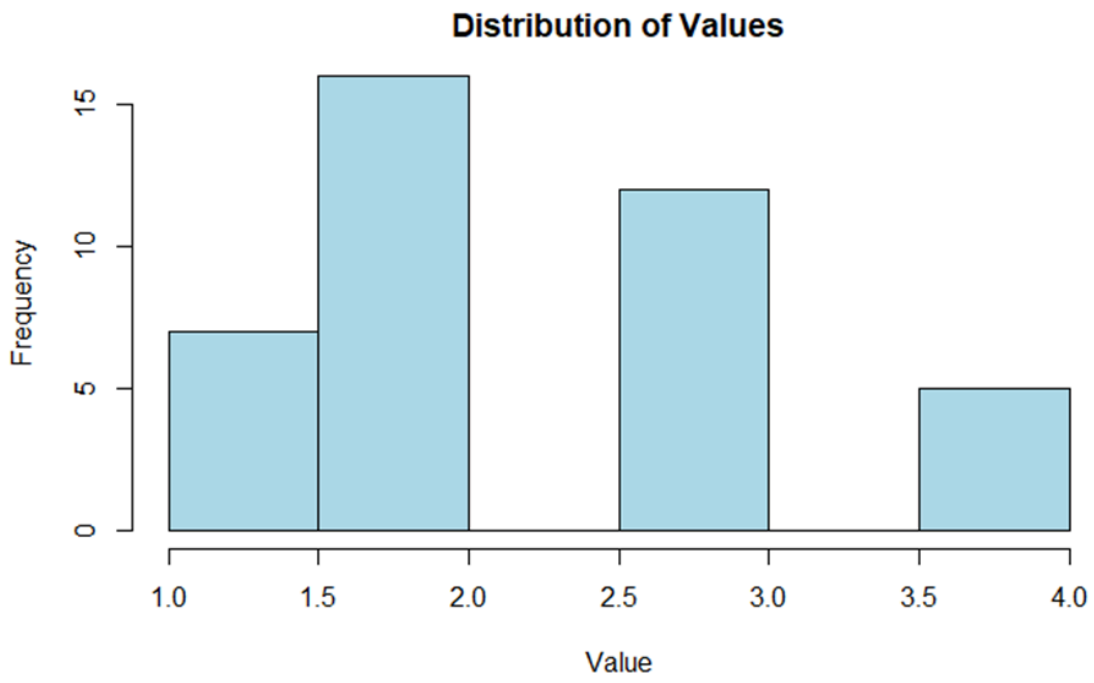


12. On a scale of 1 to 5, how important is the packaging design of a beverage to you?

40 responses

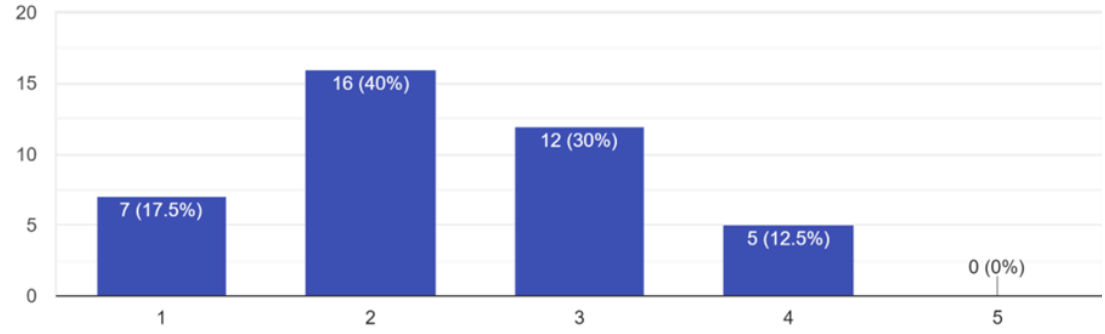


5.3.14 visua_appeal

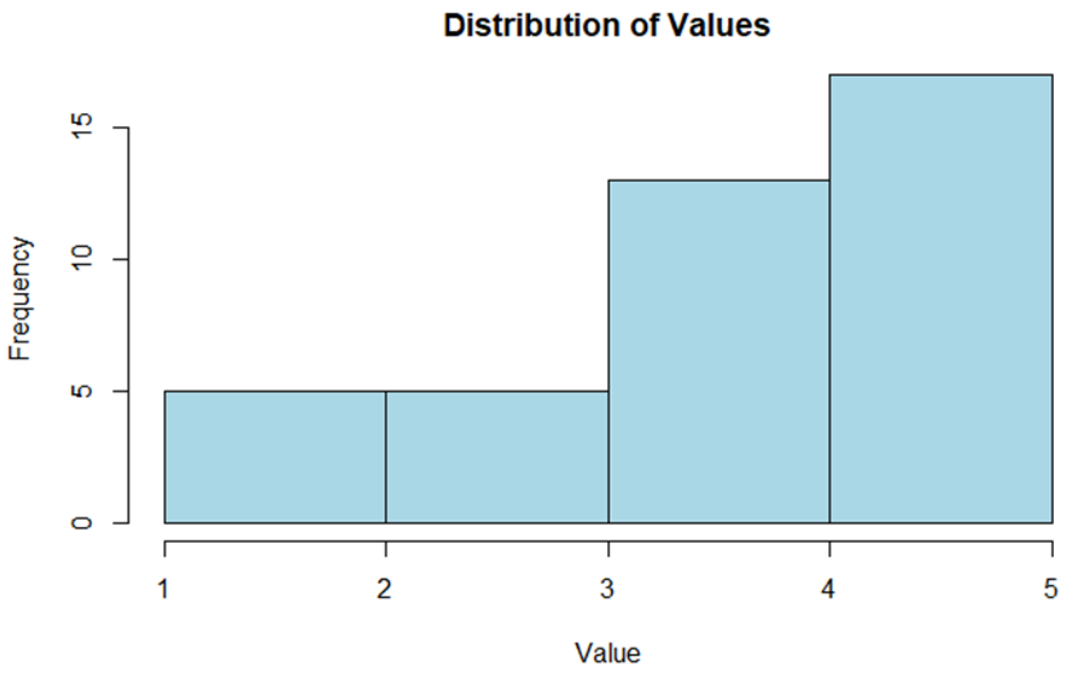


13. On a scale of 1 to 5, how visually appealing do you find the packaging of Ice Cool product below?

40 responses

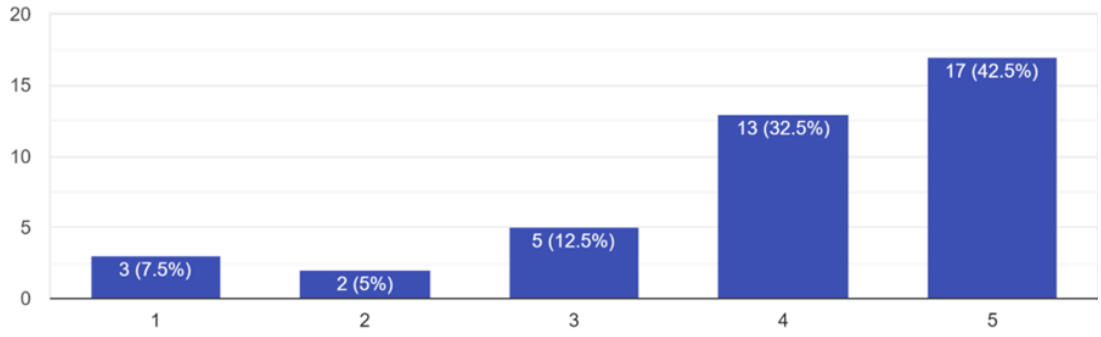


5.3.15 info_packaging

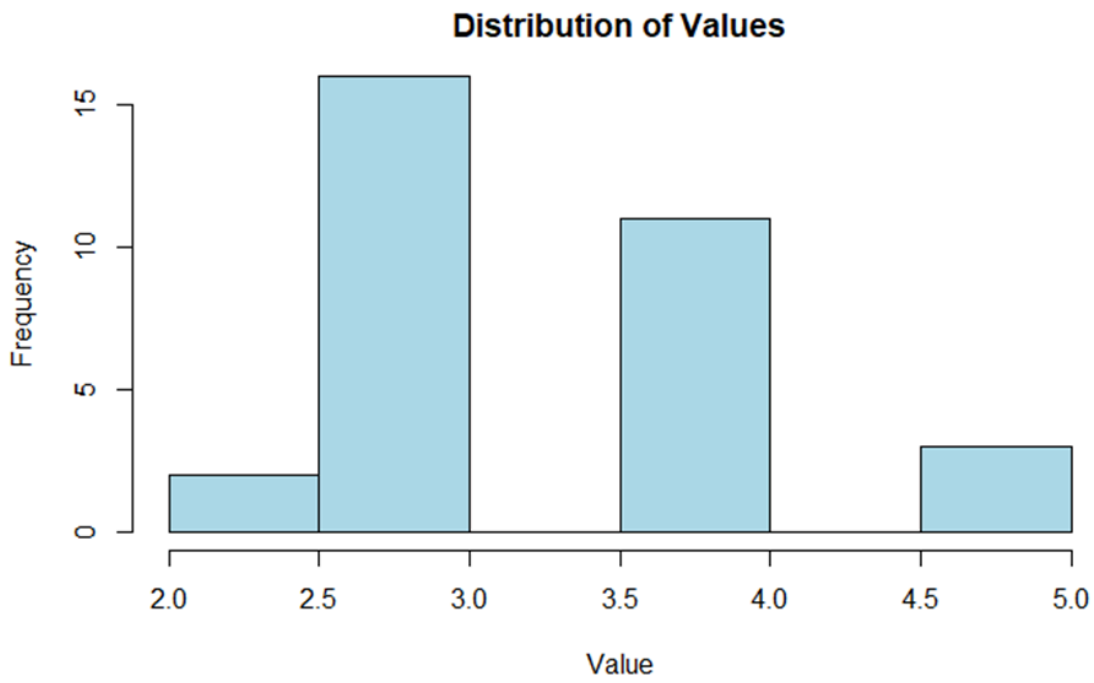


14. On a scale of 1 to 5, how much do you value informative packaging (e.g., health benefits, nutritional info) when purchasing a beverage like coconut water?

40 responses

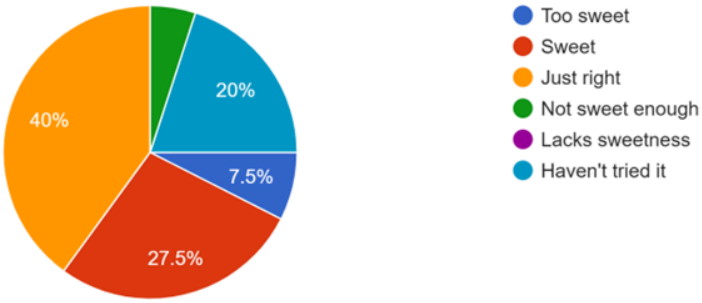


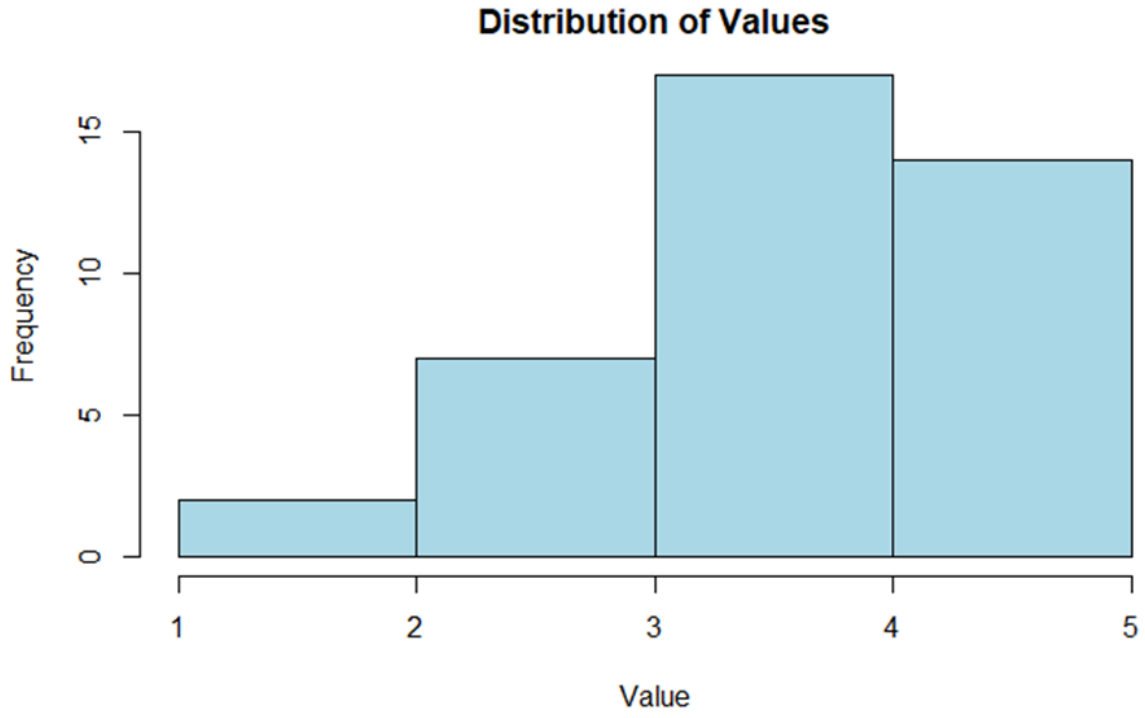
5.3.16 filtered_sweetness



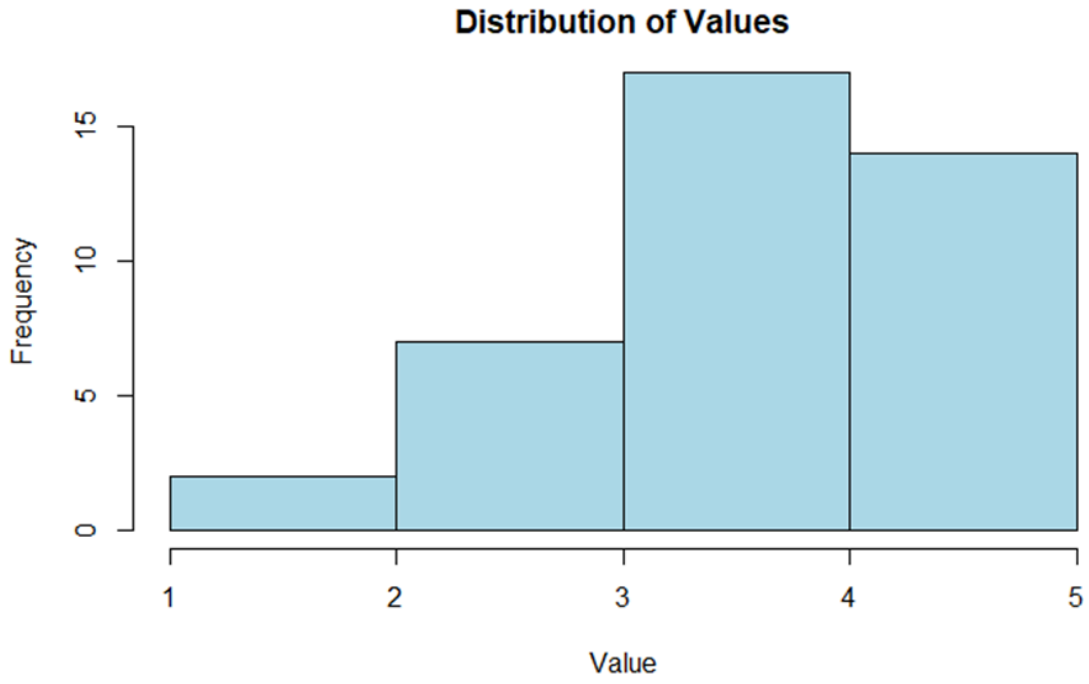
15. How would you rate the sweetness of Ice Cool coconut water?

40 responses



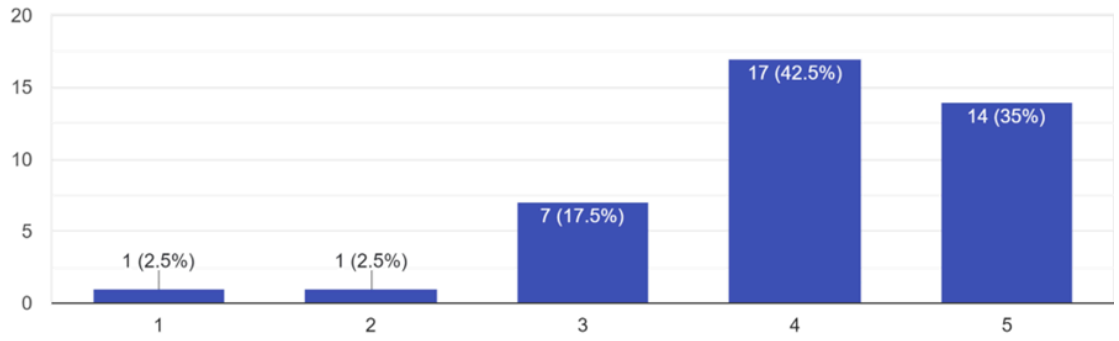


5.3.17 natural_ingre

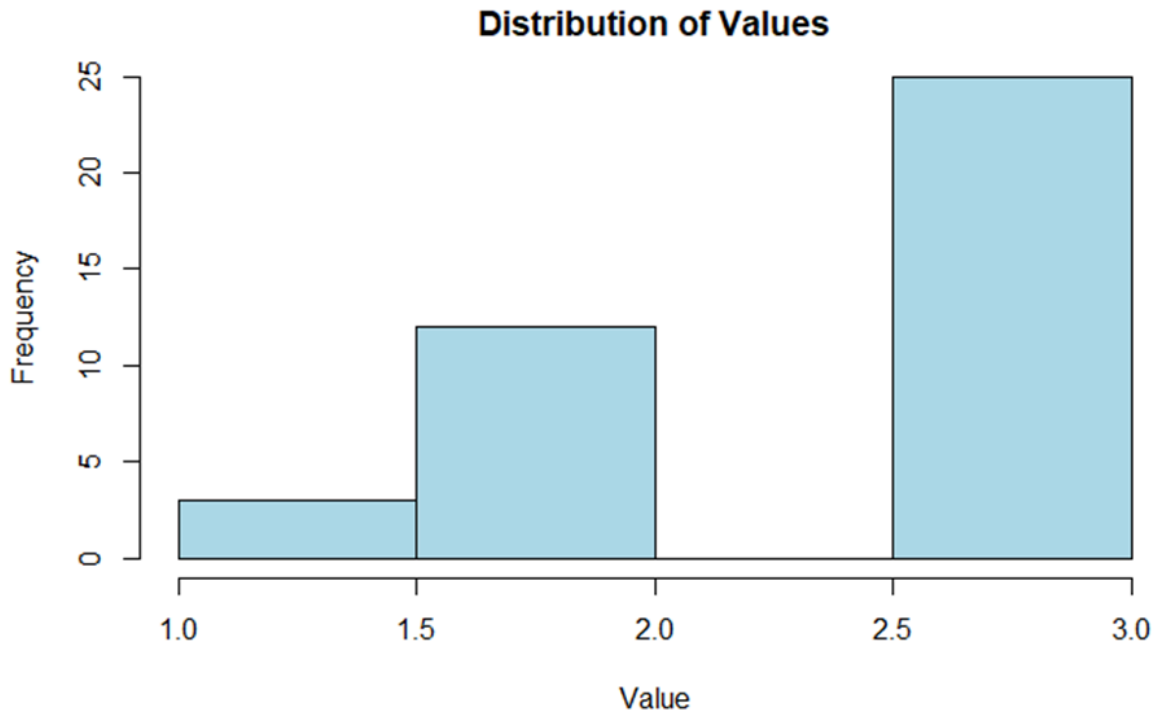


16. On a scale of 1 to 5, how important is it to you that coconut water is made from natural ingredients?

40 responses

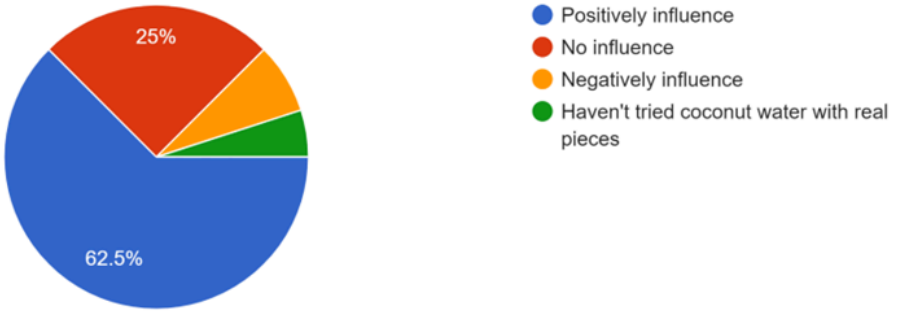


5.3.18 with_coconut



17. How would the presence of real coconut pieces in coconut water influence your decision to purchase it?

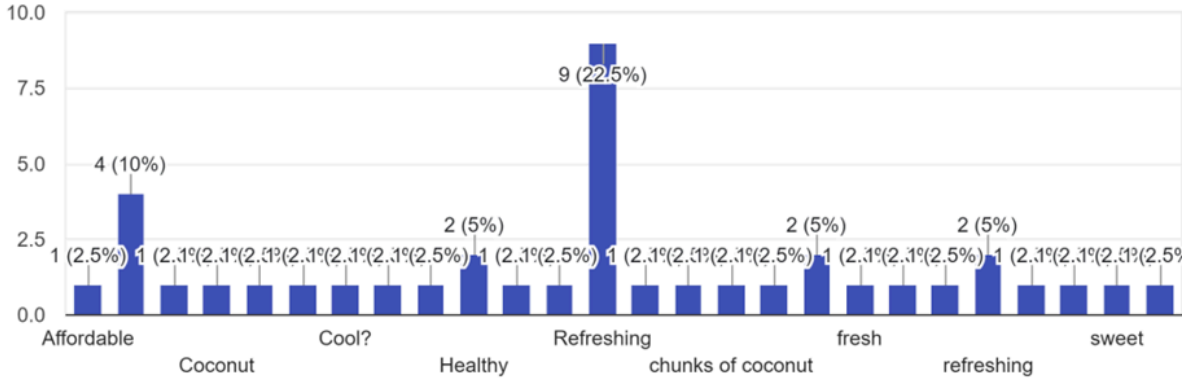
40 responses



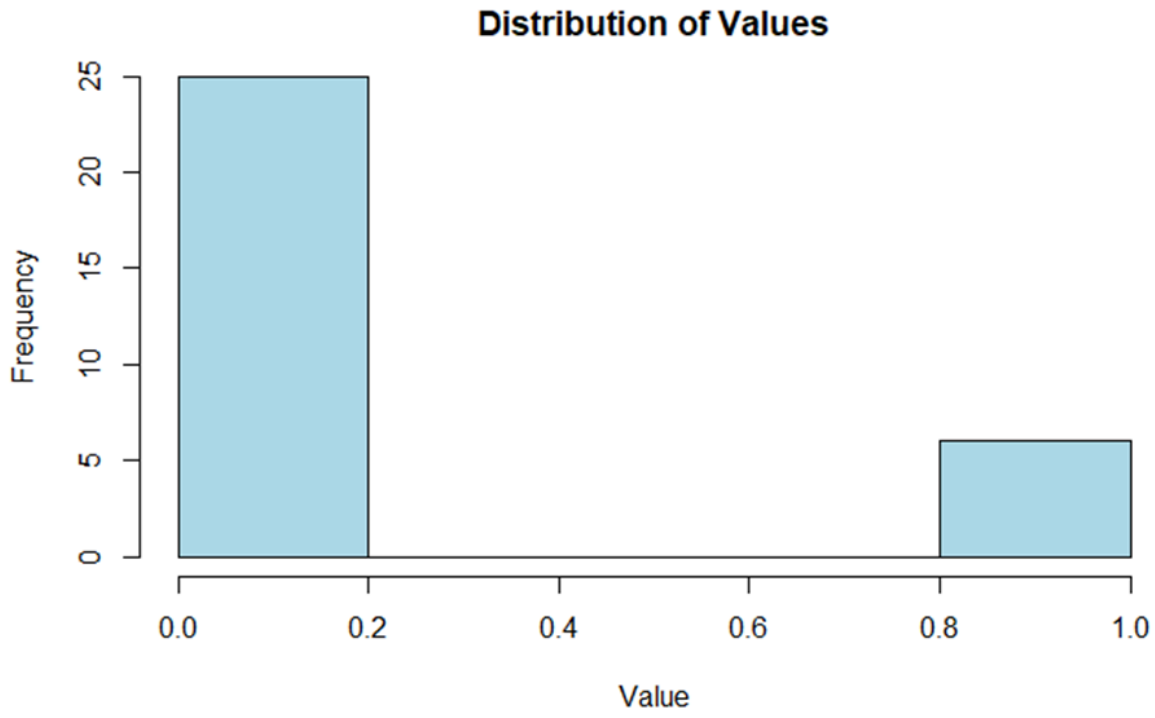
5.3.19 first_word

18. What is the first word that comes to mind when you think about Ice Cool?

40 responses

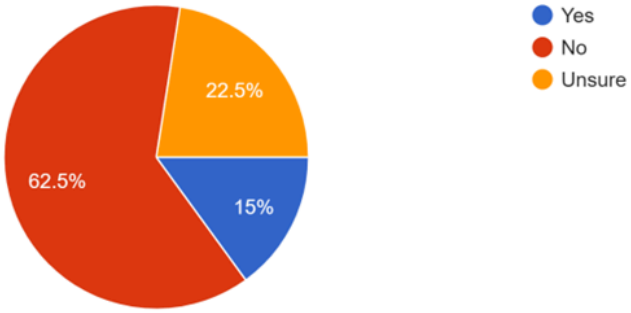


5.3.20 filtered_unique_values



19. Do you associate any unique values or benefits with the Ice Cool brand?

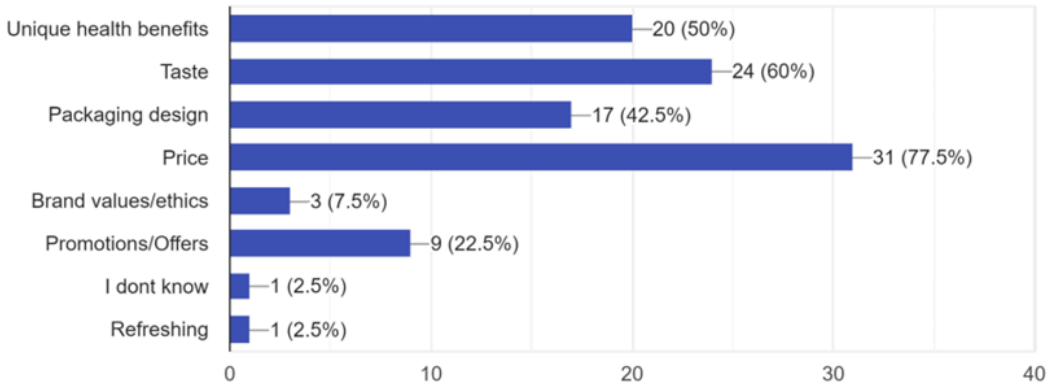
40 responses



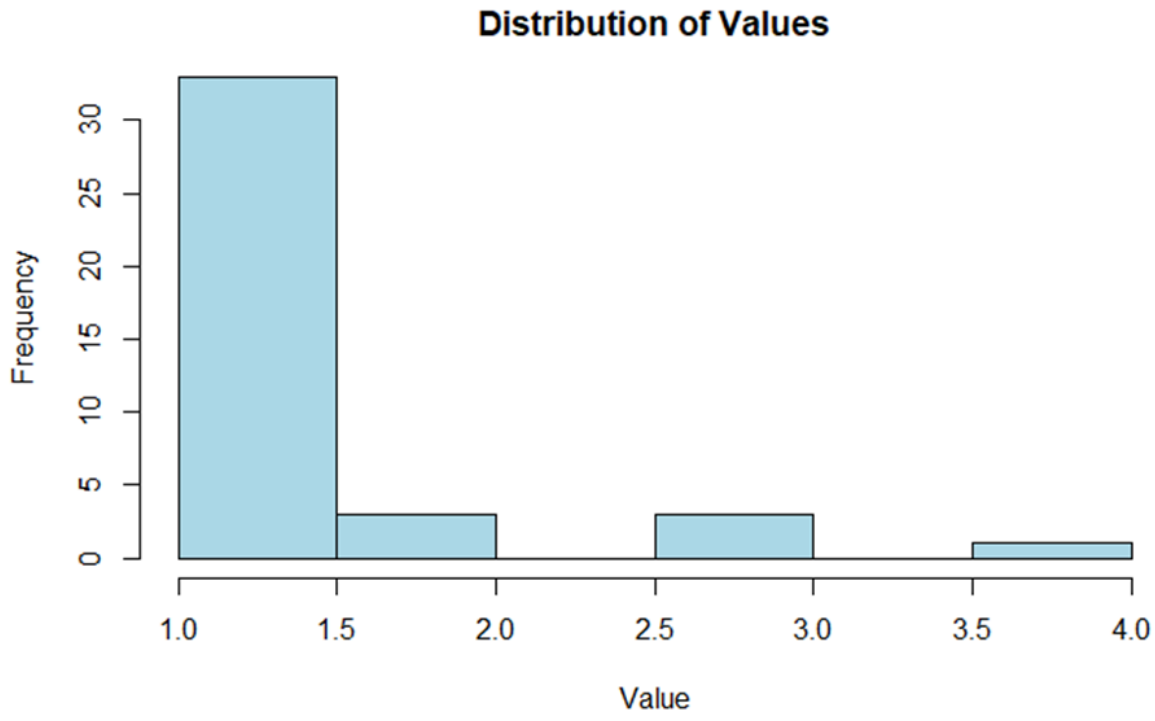
5.3.21 standout

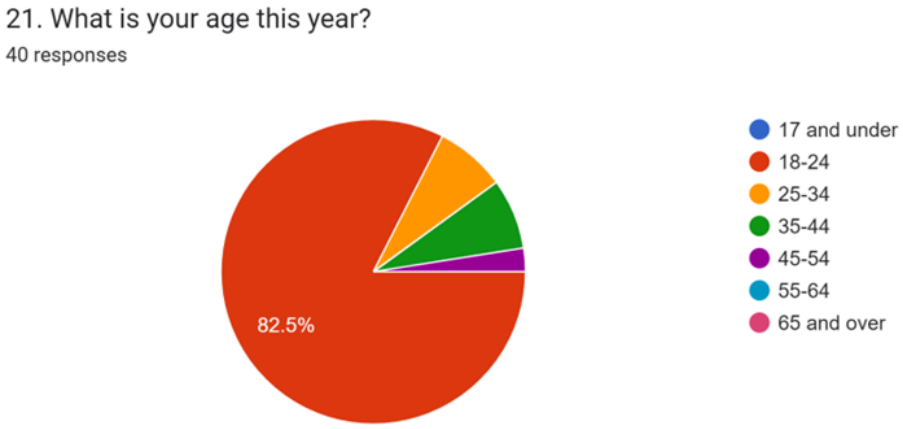
20. In your opinion, what makes a coconut water brand stand out from its competitors? (Select up to three)

40 responses

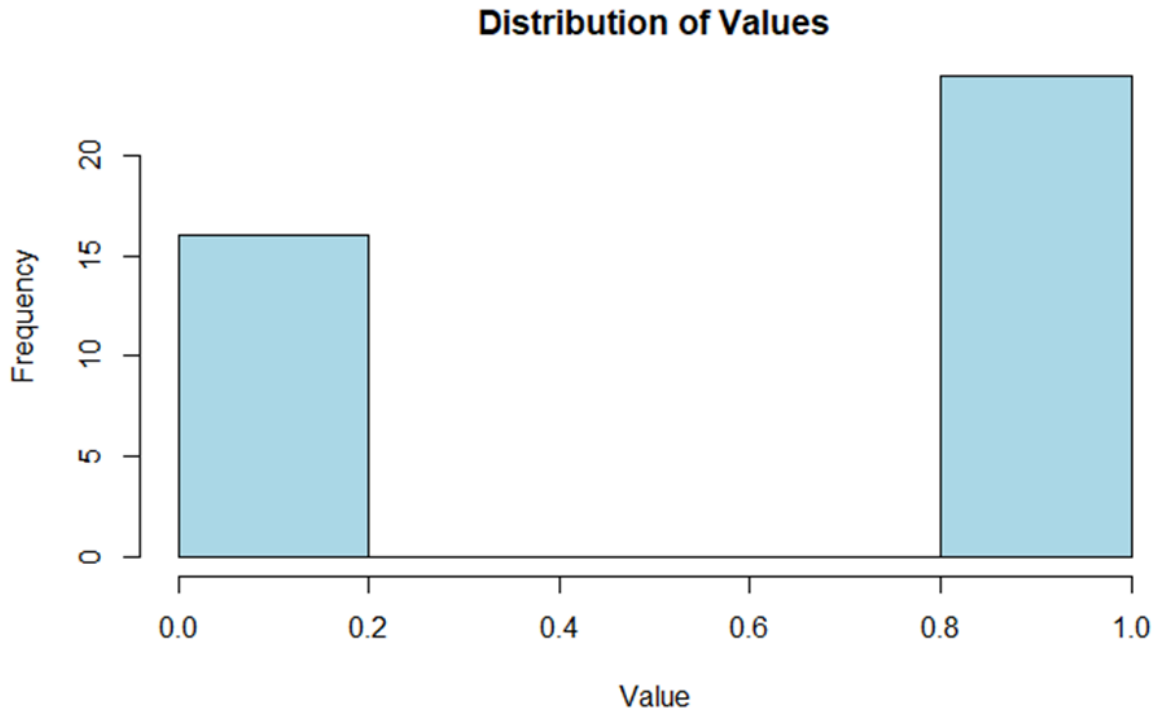


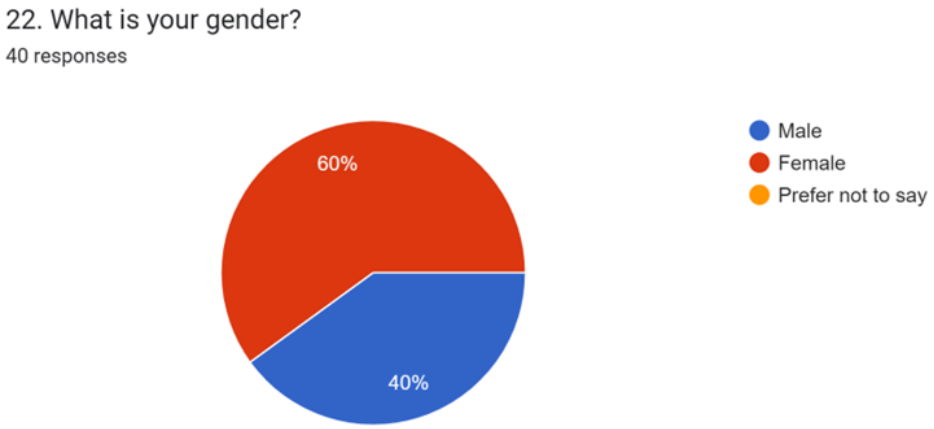
5.3.22 age



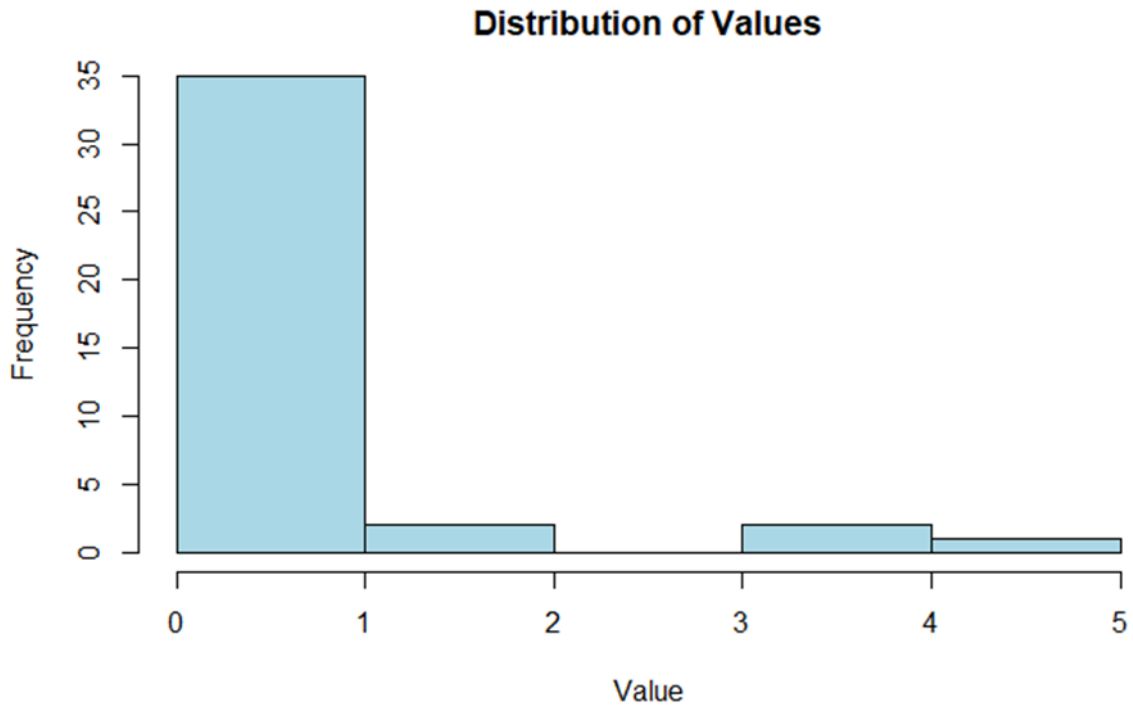

21. What is your age this year?
40 responses
82.5%
17 and under
18-24
25-34
35-44
45-54
55-64
65 and over

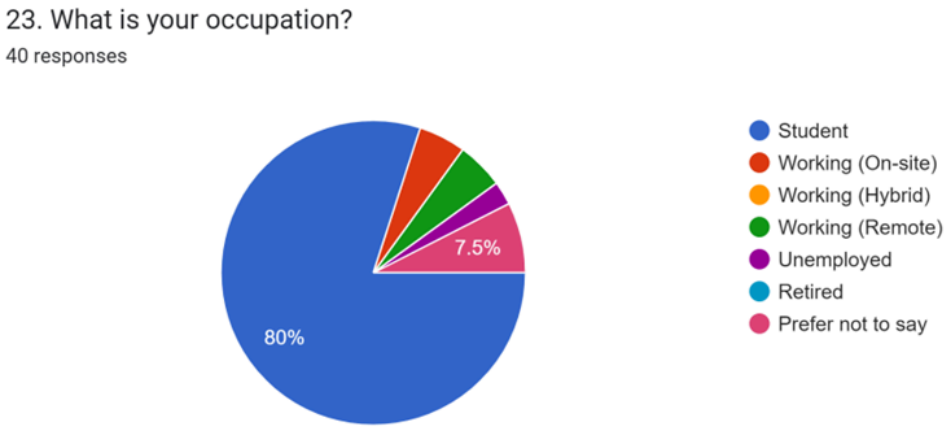
5.3.23 gender


Distribution of Values
Frequency
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0
Value



5.3.24 occupation





5.4 Significance Testing Results

5.4.1.1 filtered_brand_switching

- Testing the alternative that this variable is statistically significant when the mean is 3.

  Accept alternative hypothesis: true mean is greater than 3.

- t = 6.1394, df = 35, p-value = 2.531e-07, mean of x: 3.777778, rejecting the null hypothesis.

5.4.1.2 price_importantce

- Testing the alternative that this variable is statistically significant when the mean is 3.

  Accept alternative hypothesis: true mean is greater than 3.

- t = 9.8439, df = 39, p-value = 1.994e-12, mean of x: 4.125, rejecting the null hypothesis.

5.4.1.3 discount_importance

- Testing the alternative that this variable is statistically significant when the mean is 3.

  Accept alternative hypothesis: true mean is greater than 3.

- t = 5.2372, df = 39, p-value = 2.955e-06, mean of x: 3.8, rejecting the null hypothesis.

5.4.1.4 filtered_stu_price_sense

- Testing the alternative that this variable is statistically significant when the mean is not 0. Accept alternative hypothesis: true mean is not equal to 0.
- t = 25.807, df = 37, p-value < 2.2e-16, mean of x: 0.9473684, rejecting the null hypothesis.

5.4.1.5 filtered_familiarity

- Testing the alternative that this variable is statistically significant when the mean is not equal to 1. Accept alternative hypothesis: true mean is not equal to 1.
- t = -10.77, df = 29, p-value = 1.194e-11, mean of x: 0.2, rejecting the null hypothesis.

5.4.1.6 marketing

- Testing the alternative that this variable is statistically significant at 5% when the mean is 0. Accept alternative hypothesis: true mean is not equal to 0.
- t = 2.1044, df = 31, p-value = 0.04355, mean of x: 0.125, rejecting the null hypothesis.

5.4.1.7 influence_marketing

- Testing the alternative that this variable is statistically significant at 5% when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.
- t = 1.9338, df = 39, p-value = 0.03021, mean of x: 3.35, rejecting the null hypothesis.

5.4.1.8 package_design

- Testing the alternative that this variable is statistically significant when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.
- t = 6.1085, df = 39, p-value = 1.83e-07, mean of x: 3.725, rejecting the null hypothesis.

5.4.1.9 visua_appeal

- Testing the alternative that this variable is statistically significant when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.

- t = -4.273, df = 39, p-value = 6.003e-05, mean of x: 2.375, rejecting the null hypothesis.

5.4.1.10 info_packaging

- Testing the alternative that this variable is statistically significant when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.
- t = 6.1394, df = 35, p-value = 2.531e-07, mean of x: 3.975, rejecting the null hypothesis.

5.4.1.11 sweetness

- Testing the alternative that this variable is statistically significant when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.
- t = 3.4829, df = 31, p-value = 0.0007501, mean of x: 3.46875, rejecting the null hypothesis.

5.4.1.12 natural_ingre

- Testing the alternative that this variable is statistically significant when the mean is 3. Accept alternative hypothesis: true mean is greater than 3.
- t = 7.1228, df = 39, p-value = 7.272e-09, mean of x: 4.05, rejecting the null hypothesis.

5.4.1.13 with_coconut

- Testing the alternative that this variable is statistically significant when the mean is 2. Accept alternative hypothesis: true mean is greater than 2.
- t = 5.4479, df = 39, p-value = 1.512e-06, mean of x: 2.55, rejecting the null hypothesis.

5.4.1.14 filtered_unique_values

- Testing the alternative that this variable is statistically significant when the mean is not equal 1. Accept alternative hypothesis: true mean is not equal to 1.
- t = -11.18, df = 30, p-value = 3.2e-12, mean of x: 0.1935484, rejecting the null hypothesis.

5.4.2 Testable Variables

List of testable variables:

- filtered_brand_switching
- price_importantce
- discount_importance
- filtered_stu_price_sense
- filtered_familiarity
- marketing
- influence_marketing
- package_design
- visua_appeal
- info_packaging
- sweetness
- natural_ingre
- with_coconut
- filtered_unique_values

5.4.3 Significance Testing Criteria

The criteria for significance testing are as followed:

- Can be a binary variable (e.g Yes or No questions)
- Negative and Positive scaling (e.g., Questions scaling 1-5)
- Must not include incomparable options (e.g., The consumption and consumption_reason are categorical variables, while price_importance is a comparable scaling variable)

5.4.4 Significance Testing Results Analysis

The following are the results and interpretations using one-sample t tests with the listed variables in 5.4.2 Testable Variables.

When testing the filtered_brand_switching, it is clear, as shown in 5.4.1, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents are willing to switch brands when another is on sale.

When testing the price_importantce it is clear, as shown in 5.4.2, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents value the price of a beverage when making a purchasing decision.

When testing the discount_importance it is clear, as shown in 5.4.3, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents are likely to purchase a beverage they have not tried before when presented with a discount or promotional sale.

When testing the filtered_stu_price_sense it is clear, as shown in 5.4.4, that the true mean is not 0. This means accepting the alternative hypothesis. This result is significant because it illustrates that students are more price-sensitive when buying beverages.

When testing the filtered_familiarity it is clear, as shown in 5.4.5, that the true mean is not equal to 1. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents are not familiar with the Ice Cool brand before this survey.

When testing the marketing it is clear, as shown in 5.4.6, that the true mean is not equal to 0. This means rejecting the alternative hypothesis at 5% significance level. This result is significant because it illustrates that the respondents have not seen Ice Cool advertisements.

When testing the influence_marketing it is clear, as shown in 5.4.7, that the true mean is greater than 3. This means accepting the alternative hypothesis at 5% significance level. This result is

significant because it illustrates that the respondents are influenced by marketing campaigns and advertisements when faced with beverage purchasing decisions.

When testing the package_design it is clear, as shown in 5.4.8, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents value importance in the packaging design of a beverage.

When testing the visua_appeal it is clear, as shown in 5.4.9, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents value that the packaging of the Ice Cool product is visually appealing when shown an image.

When testing the info_packaging it is clear, as shown in 5.4.10, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents value informative packaging (e.g., health benefits, nutritional info) when purchasing a beverage like coconut water.

When testing the sweetness, it is clear, as shown in 5.4.11, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents believe Ice Cool coconut water is sweet or too sweet.

When testing the natural_ingre it is clear, as shown in 5.4.12, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents value the importance of coconut water made with natural ingredients.

When testing the with_coconut it is clear, as shown in 5.4.13, that the true mean is greater than 3. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents show a positive influence on their decision to purchase coconut water when it contains coconut pieces.

When testing the filtered_unique_values it is clear, as shown in 5.4.14, that the true mean is not equal to 1. This means accepting the alternative hypothesis. This result is significant because it illustrates that the respondents do not associate any unique values or benefits with the Ice Cool brand.

5.5 Regression Analysis Results

5.5.1.1 Price Sensitivity Regression Analysis

Firstly, as shown below, price_sense is the dependent variable, while brand_aware, packaging, and taste are the independent variables. This first model is important because when running the regression model to understand price sensitivity against brand awareness, packaging, and taste variables, it is clear that only brand awareness has a weak statistically significant effect. It is also apparent that brand awareness has a positive effect on price sensitivity. However, when checking the individual brand_frequency and flitered_brandswitch variables, neither have an effect on price sensitivity. It is also apparent that for predictive purposes, the adjusted R-squared is only 9%. This means there are likely better variables to use.

In the second model, as shown below, price_sense is the dependent variable, while brand_aware, packaging, gender are the independent variables. The brand awareness variable is kept in the model to see if, with different variables in the model, brand awareness becomes more important. This second model is important because when removing taste for the gender variable, it is evident that not only brand awareness is now strongly statistically significant, but packaging is also weakly statistically significant, and gender is moderately statistically significant. It is also apparent that for predictive purposes, the adjusted R-squared is now 21.84%. This is a much better result than the previous model, but it is clear that more testing is needed to fully understand the relationships price sensitivity has against important variables.

In this last model as shown below, price_sense is the dependent variable, while brand_aware, packaging, gender, and age are the independent variables. In this model, age is considered a proxy for student status, as the majority of the survey population are students. In this model, it is evident that brand awareness keeps its strong statistical significance, packaging is now moderately statistically significant, gender remains moderately statistically significant, and age is weakly statistically significant. It is clear that the adjusted R-squared increased to 36.31%, another good sign for predictive purposes.

To truly understand the impacts of the models, it is necessary to examine each effect of each variable against price_sense.

The first model shows that the coefficient for brand_aware is 0.3008, and it is statistically significant at the 10% level ($p<0.1$). This means that for every one-unit increase in brand awareness, the price sensitivity (price_sense) increases by 0.3008 units, holding other variables constant. The coefficient for packaging is 0.2725, but it is not statistically significant at the 10% level. This suggests that there is no statistically significant relationship between packaging and price sensitivity in this model. The coefficient for packaging is 0.0073, but it is not statistically significant at the 10% level. This suggests that there is no statistically significant relationship between taste and price sensitivity in this model.

The second model shows that the coefficient for brand_aware is 0.4243, and it is statistically significant at the 1% level ($p<0.01$). This means that for every one-unit increase in brand awareness, the price sensitivity (price_sense) increases by 0.4243 units, holding other variables constant. The coefficient for packaging is 0.3279, and it is statistically significant at the 10% level ($p<0.1$). This suggests a positive relationship between packaging and price sensitivity. The coefficient for gender is -0.4159, and it is statistically significant at the 5% level ($p<0.05$). This

suggests that being female (coded as 1) is associated with lower price sensitivity compared to being male (coded as 0).

The last model shows that the coefficient for brand_aware is 0.4705, and it is statistically significant at the 1% level ($p<0.01$). The interpretation is similar to the previous models. For every one-unit increase in brand awareness, price sensitivity (price_sense) increases by 0.4705 units, holding other variables constant. The coefficient for packaging is 0.4297, and it is statistically significant at the 5% level ($p<0.05$). This suggests a positive relationship between packaging and price sensitivity. The coefficient for gender is -0.3919, and it is statistically significant at the 5% level ($p<0.05$). This suggests that being female (coded as 1) is associated with lower price sensitivity compared to being male (coded as 0). The coefficient for age is -0.3201, and it is statistically significant at the 10% level ($p<0.1$). This suggests that, among the predominantly student population, as age increases, price sensitivity decreases.

Model selection is difficult and tedious, as well as power consumption heavy. Thus, the last model, with a higher adjusted R-squared and more relevant explanations to the survey data is the best model for understanding the price sensitivity characteristics within the survey data. It should be known that many other variables such as consumption or package_design were tested but did not show significant results.

5.5.1.2 Price Sensitivity Regression Analysis Results

	Dependent variable:		
	price_sense		
brand_aware	0.3008*	0.4243**	0.4705**
	(0.1686)	(0.1180)	(0.1494)
packaging	0.2725	0.3279*	0.4297**
	(0.1687)	(0.1725)	(0.1834)
taste	0.0073		
	(0.1176)		
gender		-0.4159**	-0.3919**
		(0.1694)	(0.1600)
age			-0.3201*
			(0.1890)
Constant	1.6439**	1.5819**	1.5864**
	(0.7151)	(0.5679)	(0.5115)
Observations	40	40	40
R2	0.1602	0.2786	0.4284
Adjusted R2	0.0902	0.2184	0.3631
Residual Std. Error	0.5432 (df = 36)	0.5035 (df = 36)	0.4545 (df = 35)
F Statistic	2.2888* (df = 3; 36)	4.6333** (df = 3; 36)	6.5577** (df = 4; 35)
Note:			*p<0.1; **p<0.05; ***p<0.01

5.5.2 Marketing Influence Regression Analysis

Firstly, as shown below, influence_marketing is the dependent variable, while filtered_marketing is the independent variable. This first model is not that important but illustrates that the filtered marketing variable does not have a statistically significant effect on how marketing campaigns and ads influences purchase decision making. This is likely because the majority of respondents indicate that Ice Cool is not effective. It is also apparent that for predictive purposes, the adjusted R-squared is negative. This means there are better variables to use.

In the second model, as shown below, influence_marketing is the dependent variable, while price_importantce and gender are the independent variables. This second model is important because when replacing the filtered marketing variable for the gender and price_importance, it is evident that not only price_importance is strongly statistically significant, but gender is

moderately statistically significant. It is also apparent that for predictive purposes, the adjusted R-squared is now 26%. This is a much better result than the previous model, but it is clear that more testing is needed to fully understand the relationships marketing influence has against important variables.

In the third model, as shown below, influence_marketing is the dependent variable, while price_importantce, gender, filtered_unique_values are the independent variables. In this model, filtered_unique_values can be understood as the unique where respondents do not value Ice Cool, as the majority of the survey population say “No.” In this model, it is evident that price_importantce keeps its strong statistical significance, gender is now strongly statistically significant, and filtered_unique_values is moderately statistically significant. It is clear that the adjusted R-squared increased to 32.92%, another good sign for predictive purposes.

In this last model, as shown below, influence_marketing is the dependent variable, while discount_importance, gender, filtered_unique_values are the independent variables. In this model, discount_importance replaces price_importantce. In this model, it is evident that discount_importance is of strong statistical significance, gender keeps its strong statistical significance, and filtered_unique_values is now strongly statistically significant. It is clear that the adjusted R-squared increased to 47.17%, the best sign for predictive purposes.

To truly understand the impacts of the models, it is necessary to examine each effect of each variable against influence_marketing.

The first model shows that the coefficient for filtered_marketing is 0.0714, but it is not statistically significant ($p>0.1$) at any common significance level. This suggests that filtered_marketing does not have a statistically significant influence on influence_marketing.

The second model shows that the coefficient for price_importantce is 0.7328, and it is statistically significant at the 1% level. This implies that a one-unit increase in price_importantce is associated with an increase of 0.7328 units in influence_marketing. The coefficient for gender is 0.8402, and it is statistically significant at the 5% level. This suggests that being female (coded as 1) is associated with higher levels of influence_marketing compared to being male (coded as 0).

The third model shows that the coefficient for price_importantce is 0.9058, and it is statistically significant at the 1% level. This indicates that an increase in price_importantce is associated with a higher level of influence_marketing. The coefficient for gender is 1.0027, and it is statistically significant at the 1% level. This implies that being female (coded as 1) is associated with significantly higher levels of influence_marketing. The coefficient for filtered_unique_values is 0.5763, and it is statistically significant at the 5% level. This suggests that an increase in filtered_unique_values is associated with higher levels of influence_marketing.

The last model shows that the coefficient for discount_importance is 0.7995, and it is statistically significant at the 1% level. This indicates that an increase in discount_importance is associated with higher levels of influence_marketing. The coefficient for gender is 1.2325, and it is statistically significant at the 1% level. This suggests that being female (coded as 1) is associated with significantly higher levels of influence_marketing. The coefficient for filtered_unique_values is 1.0618, and it is statistically significant at the 1% level. This suggests that an increase in filtered_unique_values is associated with significantly higher levels of influence_marketing.

Model selection is difficult and tedious, as well as power consumption heavy. Thus, the last model, with a higher adjusted R-squared and more relevant explanations to the survey data is the

best model for understanding the marketing influence characteristics within the survey data. It should be known that many other variables such as consumption or package_design were tested but did not show significant results.

5.5.2.1 Marketing Influence Regression Analysis Results

	Dependent variable:			
	influence_marketing			
filtered_marketing	0.0714			
	(0.7785)			
price_importantce		0.7328***	0.9058***	
		(0.2061)	(0.2931)	
discount_importance				0.7995***
				(0.2261)
gender		0.8402**	1.0027***	1.2325***
		(0.3281)	(0.3257)	(0.3015)
filtered_unique_values			0.5763**	1.0618***
			(0.2707)	(0.3053)
Constant	3.4286***	−0.1768	−1.0023	−0.5707
	(0.2249)	(0.9243)	(1.2438)	(0.9993)
Observations	32	40	31	31
R2	0.0004	0.2985	0.3963	0.5245
Adjusted R2	−0.0329	0.2606	0.3292	0.4717
Residual Std. Error	1.1812 (df = 30)	0.9843 (df = 37)	0.9178 (df = 27)	0.8146 (df = 27)
F Statistic	0.0128 (df = 1; 30)	7.8728*** (df = 2; 37)	5.9085*** (df = 3; 27)	9.9283*** (df = 3; 27)
Note:				*p<0.1; **p<0.05; ***p<0.01