ADVANCED STATISTICAL METHODS.

ANALYZING CHOCOLATE BRAND WITH CUSTOMERY SURVEY.

VIPUL JADHAV.

ANALYZING CHOCOLATE BRAND WITH CUSTOMERY SURVEY.

INTRODUCTION

The stepwise selection process suggests that the best model, based on AIC, is the one with only the frequency variable

The confectionery industry is characterized by fierce competition, with chocolate brands vying for consumer attention and loyalty. Understanding customer preferences is paramount for companies seeking to gain a competitive edge in this lucrative market. Through comprehensive customer surveys, we aim to delve into the nuanced aspects of consumer behavior and discern their perceptions towards various chocolate brands.

This study focuses on analyzing the relationship between chocolate brands and customer survey responses. By examining factors such as taste, texture, packaging preferences, and price considerations, we aim to uncover insights that can inform strategic decision-making for chocolate manufacturers and marketers.

The significance of this research lies in its potential to shed light on the underlying drivers of consumer preferences within the chocolate industry. By identifying which attributes resonate most with consumers and how they vary across different brands, companies can tailor their product offerings and marketing strategies to better align with customer expectations.

Through a combination of descriptive analysis, comparative examination, and effect size calculation, we seek to provide a comprehensive understanding of the factors influencing consumer perceptions of chocolate brands. Ultimately, this research aims to contribute to the broader discourse on consumer behavior and assist industry stakeholders in optimizing their product offerings to meet evolving consumer demands.

#OBJECTIVE

- **DESCRIPTIVE ANALYSIS**
  - **MEAN**
  - **MEDIAN**
  - **STANDARD DEVIATION**
  - **VARIANCE**
  - **SKEWNESS**
  - **KURTOSIS**
- **PREDICTIVE ANALYSIS**
  - **LOGISTIC REGRESSION**
  - **MULTIPLE REGRESSION**
  - **CONTINGENCY TABLE CHI-SQUARE TEST**
  - **MANOVA (HYPOTHESIS)**
  - **ANOVA**
- **VISUALIZATION**
  - **BARPLOT**
  - **HISTOGRAM**
  - **PIE CHART**

➢ METHODOLOGY

WE CONDUCTED A SURVEY USING GOOGLE FORM AND GOT RESPONSES FORM THE PEOPLE AND COVNVERTED THE DATA IN THE .CSV FILE.

GOOGLE FORM LINK:- https://forms.gle/XNo6TcLqfmAncawt5

1	Gender	Age Group	chocolat e brand	frequenc y	products	cadbury taste	nestle taste	cadbury texture	nestle texture	cadbury packagin g	nestle packagin g	price preferen ce	price range	liked feature	recomme ndation
2	Male	18-25	Nestle	Few times	Nestle Kit	3	5	4	5	Yes	Yes	Nestle	20-40	Texture	Nestle
3	Female	18-25	Cadbury	Few times	Cadbury D	5	4	4	4	Yes	Yes	Cadbury	20-40	Mood enh	Cadbury
4	Female	18-25	Cadbury	Few times	Nestle Kit	3	3	4	4	Yes	Yes	Nestle	20-40	Taste	Cadbury
5	Female	18-25	Nestle	Few times	Nestle Kit	4	5	4	5	Yes	Yes	Cadbury	20-40	Taste	Nestle
6	Male	18-25	Cadbury	Few times	Cadbury S	4	4	4	4	Yes	Yes	Nestle	20-40	Taste	Cadbury
7	Female	18-25	Cadbury	Few times	Cadbury D	3	2	2	2	No	No	Cadbury	50-90	Taste	Cadbury
8	Male	18-25	Cadbury	Rarely	Cadbury S	5	4	5	4	Yes	Yes	Cadbury	20-40	Taste	Cadbury
9	Male	18-25	Cadbury	Few times	Cadbury D	4	4	4	4	Yes	No	Cadbury	100 & abo	Taste	Cadbury
10	Male	18-25	Cadbury	Rarely	Cadbury S	3	3	3	2	No	Yes	Nestle	20-40	Texture	Cadbury
11	Female	18-25	Cadbury	Few times	Cadbury D	4	5	4	4	Yes	Yes	Nestle	20-40	Taste	Cadbury
12	Male	18-25	Nestle	Rarely	Nestle Kit	3	3	3	3	No	Yes	Nestle	5-10 rs	Package	Nestle
13	Male	18-25	Cadbury	Few times	Cadbury B	5	4	5	4	Yes	No	Cadbury	20-40	Taste	Cadbury
14	Male	18-25	Cadbury	Few times	Cadbury D	3	2	3	3	No	No	Cadbury	20-40	Texture	Cadbury
15	Male	18-25	Cadbury	Few times	Nestle Kit	4	3	4	3	No	Yes	Cadbury	20-40	Texture	Cadbury
16	Male	18-25	Nestle	Few times	Nestle Kit	4	5	4	5	Yes	Yes	Nestle	20-40	Taste	Nestle
17	Female	18-25	Cadbury	Rarely	Nestle Kit	3	3	2	2	Yes	Yes	Cadbury	20-40	Taste	Nestle
18	Female	18-25	Cadbury	Rarely	Cadbury D	3	3	4	3	Yes	Yes	Cadbury	20-40	Mood enh	Cadbury
19	Male	18-25	Nestle	Daily	Nestle Kit	1	5	1	5	No	Yes	Nestle	100 & abo	Taste	Nestle
20	Female	18-25	Cadbury	Few times	Cadbury B	4	3	4	4	Yes	Yes	Cadbury	20-40	Taste	Cadbury

CODE:-

```
data<- read.csv("cadbury.csv")
head(data)
#for frequency
data$frequency=factor(data$frequency,levels = c('Daily','Few times a week',
                              'Few times a month','Rarely','Never'),
             labels = c('1','2','3','4','5'))
data$frequency=as.numeric(data$frequency)
str(data$frequency)
hist(data$frequency,labels = c('Daily','Few times a week',
                   'Few times a month','Rarely','Never'))
library(moments)
kurtosis(data$frequency)
skewness(data$frequency)
summary(data$frequency)
```

```
sd(data$frequency)
var(data$frequency)
#for chocolate brand
barplot(table(data$chocolate.brand))
pie(table(data$chocolate.brand))
#for products
data$products=factor(data$products,levels = c('Cadbury Dairy Milk','Cadbury Silk','Cadbury
Bournville',
                                          'Nestle Kitkat','Nestle Crunch','Nestle Aero'),
                   labels = c('1','2','3','4','5','6'))
data$products=as.numeric(data$products)
str(data$products)
barplot(table(data$products),col=c('black','green','red','yellow','gray','blue'))
labels=c('Cadbury Dairy Milk','Cadbury Silk','Cadbury Bournville',
          'Nestle Kitkat','Nestle Crunch','5 star')
pie(table(data$products),labels = labels)
# Rating for cadbury taste
barplot(table(data$cadbury.taste))
#rating for nestle taste
pie(table(data$nestle.taste),col=c('black','green','red','yellow','blue'))
#rating for cadbury texture
pie(table(data$cadbury.texture),col=c('black','green','red','yellow','blue'))
#rating for nestle texture
barplot(table(data$nestle.texture),col=c('black','green','red','yellow','blue'))
#cadbury packing
pie(table(data$cadbury.packaging),col=c('black','green'))
#nestle packing
pie(table(data$nestle.packaging),col=c('red','blue'))
#price of chocolates range
barplot(table(data$price.range))
#liked feature
barplot(table(data$liked.feature))
#recomendation to friends
pie(table(data$recommendation))

#Logistic Regression
#for frequency
data$frequency=factor(data$frequency,levels = c('Daily','Few times a week',
                                          'Few times a month','Rarely','Never'),
                   labels = c('1','2','3','4','5'))
```

```
data$frequency=as.numeric(data$frequency)
str(data$frequency)
data$Gender=factor(data$Gender,levels = c('Male','Female'),labels = c('1','2'))
data$Gender=as.numeric(data$Gender)
data$liked.feature=factor(data$liked.feature,levels = c('Taste',
            'Texture','Mood enhancement','Package'),labels = c('1','2','3','4'))
data$liked.feature=as.numeric(data$liked.feature)
data$price.range=factor(data$price.range,levels = c('5-10 rs','20-40','50-90','100 & above'),
             labels = c('1','2','3','4'))
data$price.range=as.numeric(data$price.range)
data$chocolate.brand=factor(data$chocolate.brand,levels = c('Cadbury','Nestle'),
               labels = c('1','0'))
data$chocolate.brand=as.numeric(data$chocolate.brand)
str(data$chocolate.brand)
# Check unique values in chocolate.brand
unique(data$chocolate.brand)
# Recode chocolate.brand to binary (0 and 1)
data$chocolate.brand <- ifelse(data$chocolate.brand == 2, 0, 1)
model<- glm(data$chocolate.brand ~ data$frequency +
        data$Gender + data$liked.feature + data$price.range,family = binomial)
summary(model)
#Linear Regression with Selection method
library(MASS)
fullmodel<- lm(formula = data$chocolate.brand ~ data$frequency +
         data$Gender + data$liked.feature + data$price.range)
stepmodel<- stepAIC(fullmodel,direction = 'both')
#for manova test
data$chocolate.brand=factor(data$chocolate.brand,levels = c('Cadbury','Nestle'),
               labels = c('1','0'))
data$chocolate.brand=as.numeric(data$chocolate.brand)
str(data$chocolate.brand)
# Check unique values in chocolate.brand
unique(data$chocolate.brand)
# Recode chocolate.brand to binary (0 and 1)
data$chocolate.brand <- ifelse(data$chocolate.brand == 2, 0, 1)
data$liked.feature=factor(data$liked.feature,levels = c('Taste',
                        'Texture','Mood enhancement','Package'),labels =
c('1','2','3','4'))
H0='chocolate brand means are equal'
H1='chocolate brand means are not equal'
```

```
model<- lm(cbind(data$cadbury.taste,data$nestle.taste,
        data$cadbury.texture,data$nestle.texture,data$liked.feature)
     ~data$chocolate.brand)
model
a=manova(model)
summary(a,test='Pillai')
Fcal<- summary(a,test='Pillai')$stats['data$chocolate.brand','approx F']
Fcal
ftab=qf(0.95,5,128)
ftab
if(Fcal>ftab){
  print('we reject H0')
}else{
  print("Accept H1")
}
#contingency table
H0=' There is no association between Gender and Chocolate Brand preference.'
H1='There is an association between Gender and Chocolate Brand preference.'
# Create contingency table
contingency_table <- table(data$Gender, data$chocolate.brand)
# Display the contingency table
print(contingency_table)
# Perform chi-squared test
chi_squared <- chisq.test(contingency_table)
# Print the results
print(chi_squared)

#anova
# Perform multi-factor ANOVA
anova_result <- aov(data$chocolate.brand ~ data$cadbury.taste +data$nestle.taste+
         data$cadbury.texture+ data$nestle.texture +data$liked.feature)

# Summary of ANOVA
summary(anova_result)
```

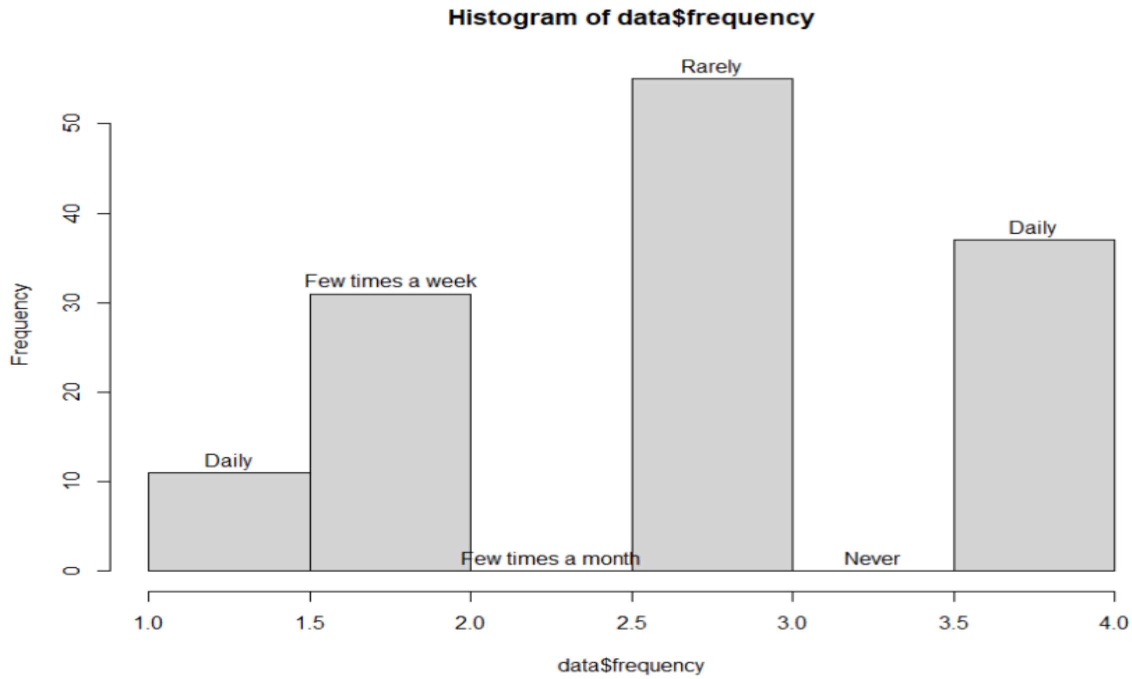
#DATASET.

```
> data<- read.csv("cadbury.csv")
> head(data)
  Gender Age.Group chocolate.brand         frequency          products cadbury.taste
1   Male     18-25          Nestle  Few times a week      Nestle Kitkat            3
2 Female     18-25         Cadbury Few times a month Cadbury Dairy Milk            5
3 Female     18-25         Cadbury  Few times a week      Nestle Kitkat            3
4 Female     18-25          Nestle Few times a month      Nestle Kitkat            4
5   Male     18-25         Cadbury Few times a month       Cadbury Silk            4
6 Female     18-25         Cadbury  Few times a week Cadbury Dairy Milk            3
  nestle.taste cadbury.texture nestle.texture cadbury.packaging nestle.packaging
1            5               4              5               Yes              Yes
2            4               4              4               Yes              Yes
3            3               4              4               Yes              Yes
4            5               4              5               Yes              Yes
5            4               4              4               Yes              Yes
6            2               2              2                No               No
  price.preference price.range    liked.feature recommendation
1           Nestle       20-40          Texture         Nestle
2          Cadbury       20-40 Mood enhancement        Cadbury
3           Nestle       20-40            Taste        Cadbury
4          Cadbury       20-40            Taste         Nestle
5           Nestle       20-40            Taste        Cadbury
6          Cadbury       50-90            Taste        Cadbury
```

- **RESULT & ANALYSIS**

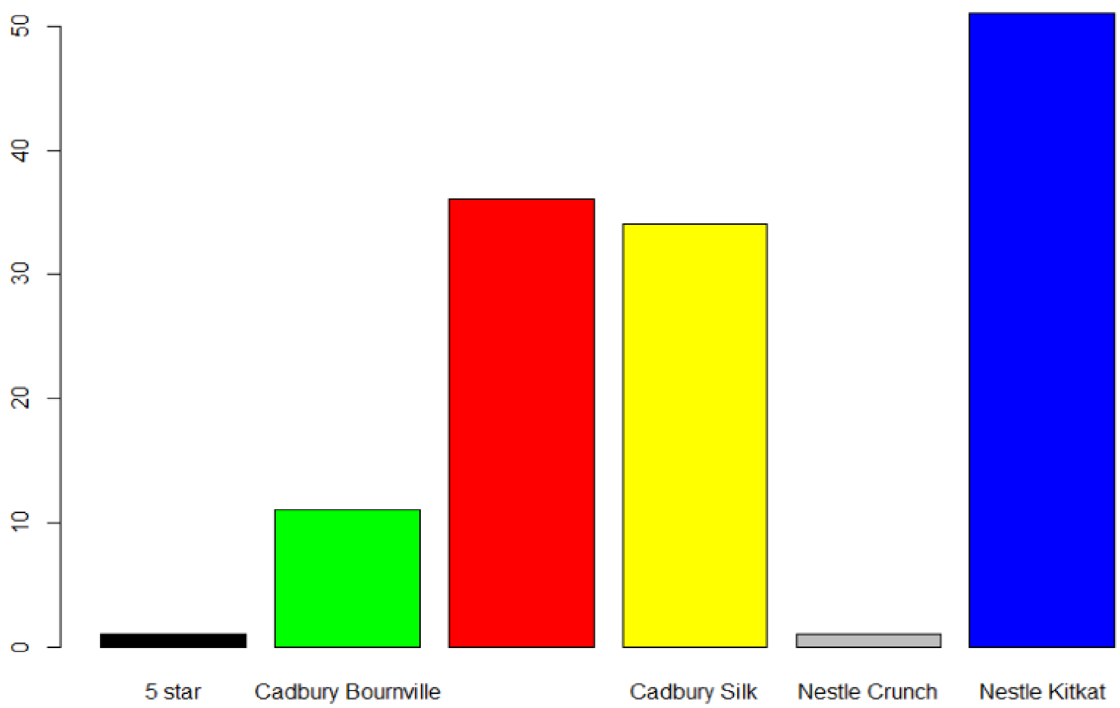
#DESCRIPTIVE ANALYSIS.

```
> #for frequency
> data$frequency=factor(data$frequency,levels = c('Daily','Few times a week',
+                                                  'Few times a month','Rarely','Never'),
+                       labels = c('1','2','3','4','5'))
> data$frequency=as.numeric(data$frequency)
> str(data$frequency)
 num [1:134] 2 3 2 3 3 2 4 2 4 3 ...
> hist(data$frequency,labels = c('Daily','Few times a week',
+                                'Few times a month','Rarely','Never'))
```

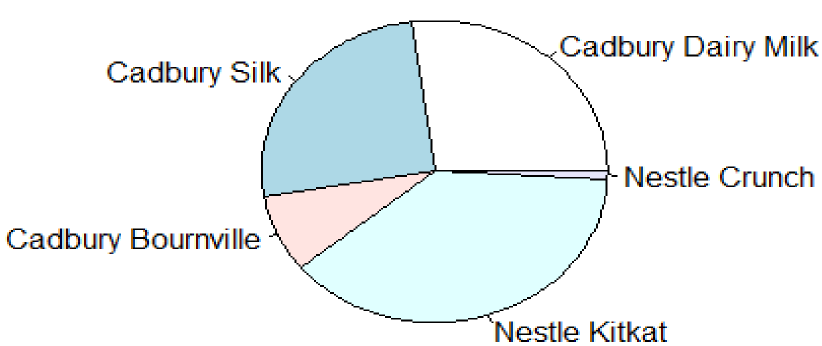



```
> library(moments)
> kurtosis(data$frequency)
[1] 2.369766
> skewness(data$frequency)
[1] -0.424279
> summary(data$frequency)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  1.000   2.000   3.000   2.881   4.000   4.000 
> sd(data$frequency)
[1] 0.9098027
> var(data$frequency)
[1] 0.827741
```

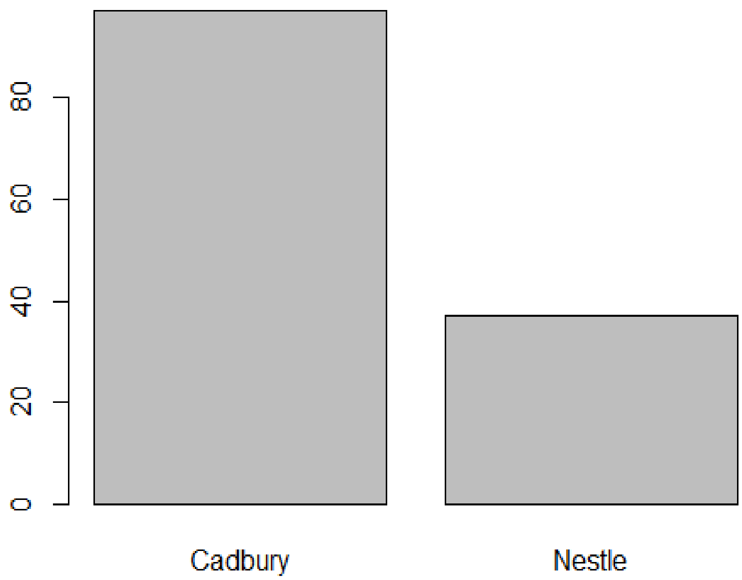
```
> barplot(table(data$products),col=c('black','green','red','yellow','gray','blue'))
```

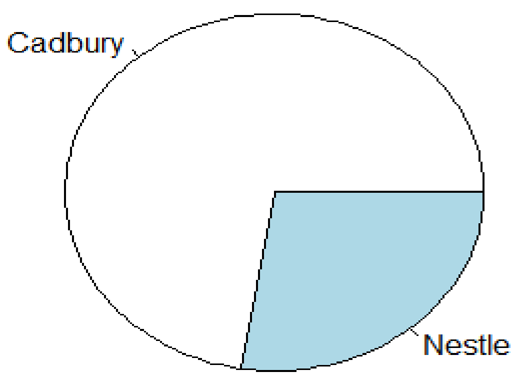



```
> #for products
> data$products=factor(data$products,levels = c('Cadbury Dairy Milk','Cadbury Silk','Cadbury Bournville',
+                                               'Nestle Kitkat','Nestle Crunch','Nestle Aero'),
+                      labels = c('1','2','3','4','5','6'))
> data$products=as.numeric(data$products)
> str(data$products)
 num [1:134] 4 1 4 4 2 1 2 1 2 1 ...
> labels=c('Cadbury Dairy Milk','Cadbury Silk','Cadbury Bournville',
+          'Nestle Kitkat','Nestle Crunch','5 star')
> pie(table(data$products),labels = labels)
```

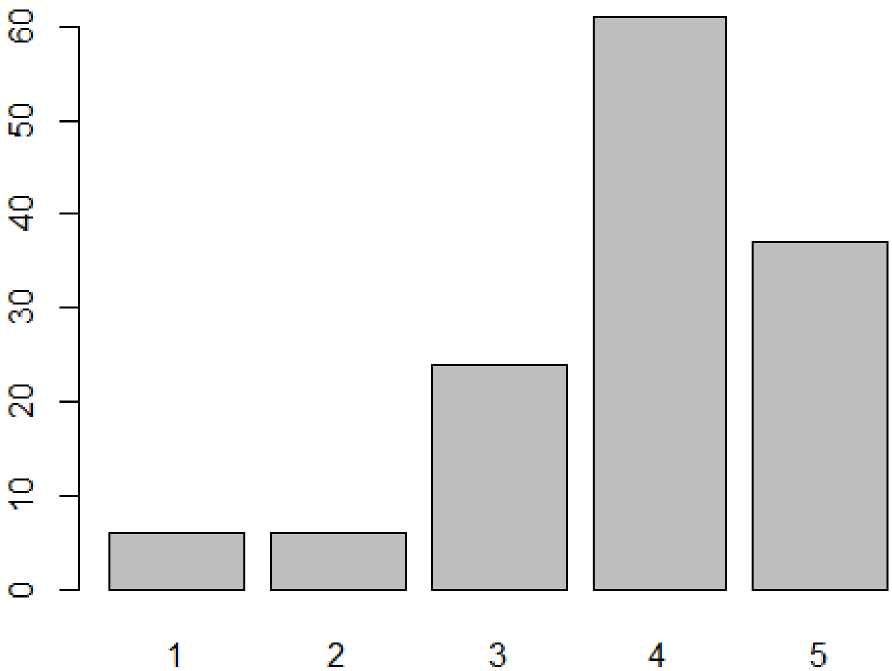


```
> #for chocolate brand
> barplot(table(data$chocolate.brand))
> pie(table(data$chocolate.brand))
```

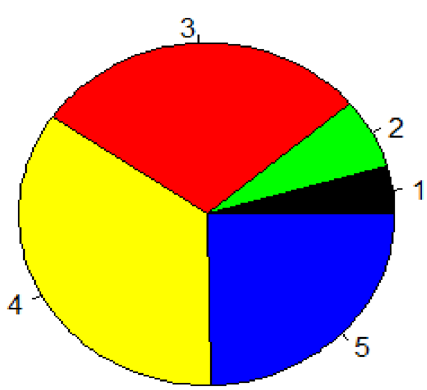




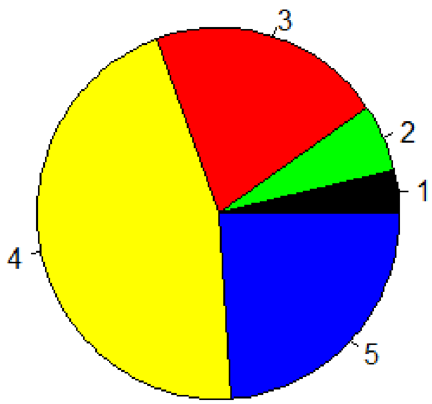

```
> # Rating for cadbury taste
> barplot(table(data$cadbury.taste))
```



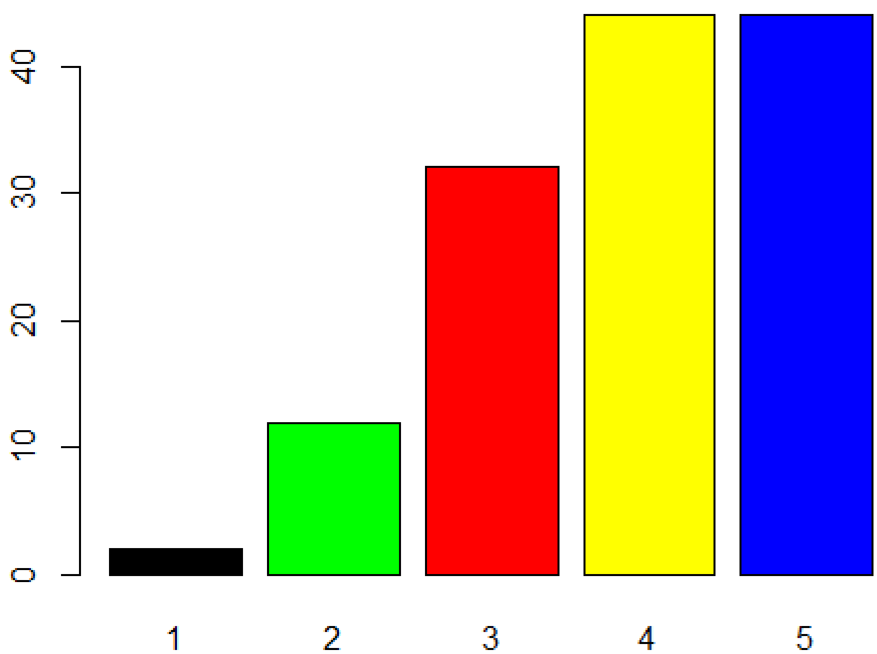

```
> #rating for nestle taste
> pie(table(data$nestle.taste),col=c('black','green','red','yellow','blue'))
```



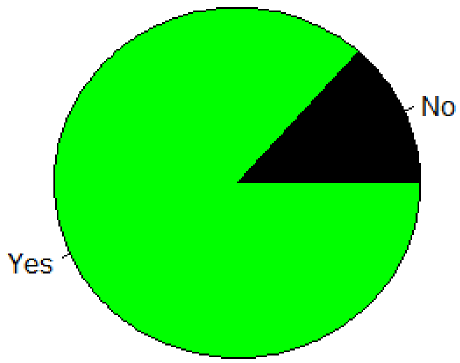

```
> #rating for cadbury texture
> pie(table(data$cadbury.texture),col=c('black','green','red','yellow','blue'))
```



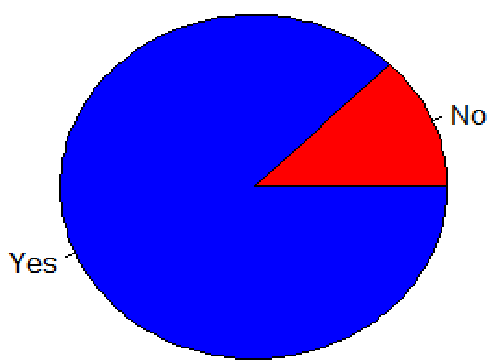

```
> #rating for nestle texture
> barplot(table(data$nestle.texture),col=c('black','green','red','yellow','blue'))
```



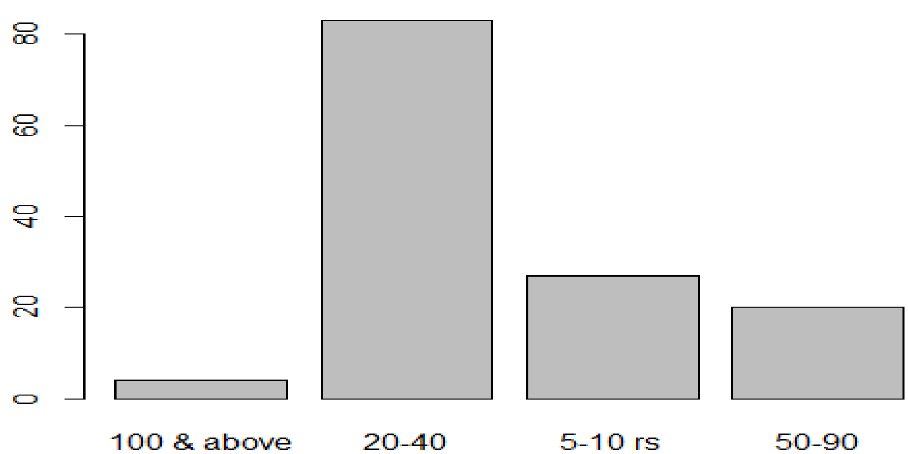
```
> #cadbury packing
> pie(table(data$cadbury.packaging),col=c('black','green'))
```



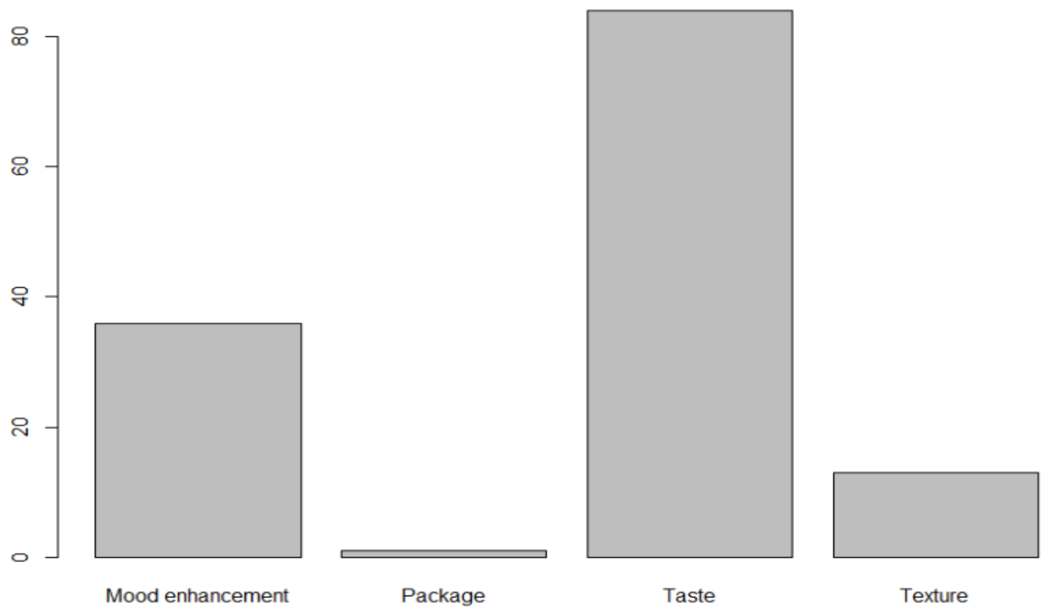

```
> #nestle packing
> pie(table(data$nestle.packaging),col=c('red','blue'))
```



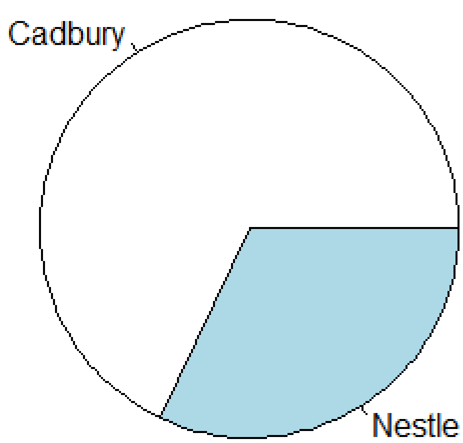

```
> barplot(table(data$price.range))
```



```
> #liked feature
> barplot(table(data$liked.feature))
```




```
> #recomendation to friends
> pie(table(data$recommendation))
```



#PREDICTIVE ANALYSIS.

#Logistic Regression

```
> data<- read.csv("cadbury.csv")
> #Logistic Regression
> #for frequency
> data$frequency=factor(data$frequency,levels = c('Daily','Few times a week',
+                                                 'Few times a month','Rarely','Never'),
+                       labels = c('1','2','3','4','5'))
> data$frequency=as.numeric(data$frequency)
> str(data$frequency)
 num [1:134] 2 3 2 3 3 2 4 2 4 3 ...
> data$Gender=factor(data$Gender,levels = c('Male','Female'),labels = c('1','2'))
> data$Gender=as.numeric(data$Gender)
> data$liked.feature=factor(data$liked.feature,levels = c('Taste',
+                       'Texture','Mood enhancement','Package'),labels = c('1','2','3','4'))
> data$liked.feature=as.numeric(data$liked.feature)
> data$price.range=factor(data$price.range,levels = c('5-10 rs','20-40','50-90','100 & above'),
+                         labels = c('1','2','3','4'))
> data$price.range=as.numeric(data$price.range)
> data$chocolate.brand=factor(data$chocolate.brand,levels = c('Cadbury','Nestle'),
+                             labels = c('1','0'))
> data$chocolate.brand=as.numeric(data$chocolate.brand)

> str(data$chocolate.brand)
 num [1:134] 2 1 1 2 1 1 1 1 1 1 ...
> # Check unique values in chocolate.brand
> unique(data$chocolate.brand)
[1] 2 1
> # Recode chocolate.brand to binary (0 and 1)
> data$chocolate.brand <- ifelse(data$chocolate.brand == 2, 0, 1)
> model<- glm(data$chocolate.brand ~ data$frequency +
+               data$Gender + data$liked.feature + data$price.range,family = binomial)
> summary(model)


Call:
glm(formula = data$chocolate.brand ~ data$frequency + data$Gender +
    data$liked.feature + data$price.range, family = binomial)

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)         0.88712    1.31172   0.676    0.499
data$frequency     -0.13409    0.22534  -0.595    0.552
data$Gender         0.12578    0.40152   0.313    0.754
data$liked.feature  0.06735    0.21890   0.308    0.758
data$price.range    0.08860    0.29650   0.299    0.765

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 157.92  on 133  degrees of freedom
Residual deviance: 157.14  on 129  degrees of freedom
AIC: 167.14

Number of Fisher Scoring iterations: 4
```

None of the predictor variables (frequency, Gender, liked.feature, price.range) have p-values less than 0.05, indicating that none of them are statistically significant predictors of chocolate brand.
The intercept term is also not statistically significant.
It appears that the logistic regression model you fitted may not be effectively capturing the relationship between the predictor variables and the response variable.

#Linear Regression with Selection method

```
> #Linear Regression with Selection method
> library(MASS)
> fullmodel<- lm(formula = data$chocolate.brand ~ data$frequency +
+                  data$Gender + data$liked.feature + data$price.range)
> stepmodel<- stepAIC(fullmodel,direction = 'both')

Start:  AIC=-206.52
data$chocolate.brand ~ data$frequency + data$Gender + data$liked.feature +
    data$price.range

                     Df Sum of Sq    RSS     AIC
- data$price.range    1  0.017847 26.647 -208.43
- data$liked.feature  1  0.018659 26.648 -208.43
- data$Gender         1  0.019330 26.649 -208.42
- data$frequency      1  0.070148 26.700 -208.17
<none>                            26.629 -206.52

Step:  AIC=-208.43
data$chocolate.brand ~ data$frequency + data$Gender + data$liked.feature

                     Df Sum of Sq    RSS     AIC
- data$liked.feature  1  0.014773 26.662 -210.36
- data$Gender         1  0.014975 26.662 -210.35
- data$frequency      1  0.090075 26.737 -209.98
<none>                            26.647 -208.43
+ data$price.range    1  0.017847 26.629 -206.52
```

```
Step:  AIC=-210.36
data$chocolate.brand ~ data$frequency + data$Gender

                     Df Sum of Sq    RSS     AIC
- data$Gender         1  0.015419 26.677 -212.28
- data$frequency      1  0.090220 26.752 -211.90
<none>                            26.662 -210.36
+ data$liked.feature  1  0.014773 26.647 -208.43
+ data$price.range    1  0.013962 26.648 -208.43

Step:  AIC=-212.28
data$chocolate.brand ~ data$frequency

                     Df Sum of Sq    RSS     AIC
- data$frequency      1  0.106115 26.784 -213.75
<none>                            26.677 -212.28
+ data$Gender         1  0.015419 26.662 -210.36
+ data$liked.feature  1  0.015218 26.662 -210.35
+ data$price.range    1  0.010102 26.667 -210.33

Step:  AIC=-213.75
data$chocolate.brand ~ 1

                     Df Sum of Sq    RSS     AIC
<none>                            26.784 -213.75
+ data$frequency      1  0.106115 26.677 -212.28
+ data$Gender         1  0.031313 26.752 -211.90
+ data$price.range    1  0.025852 26.758 -211.88
+ data$liked.feature  1  0.015582 26.768 -211.82
```

Start: The initial model includes all predictor variables (frequency, Gender, liked.feature, price.range). The AIC of this model is -206.52.

Step 1: At this step, the variable price.range is removed from the model as it results in the lowest AIC value (-208.43). The model now includes frequency, Gender, and liked.feature.

Step 2: The variable liked.feature is removed from the model as it results in the lowest AIC value (-210.36). The model now includes frequency and Gender.

Step 3: The variable Gender is removed from the model as it results in the lowest AIC value (-212.28). The final model includes only the frequency variable.

Step 4: The final model is a constant model (intercept-only), with the lowest AIC value (-213.75).

The stepwise selection process suggests that the best model, based on AIC, is the one with only the frequency variable.

```
> #for manova test
> data$chocolate.brand=factor(data$chocolate.brand,levels = c('Cadbury','Nestle'),
+                             labels = c('1','0'))
> data$chocolate.brand=as.numeric(data$chocolate.brand)
> str(data$chocolate.brand)
 num [1:134] 2 1 1 2 1 1 1 1 1 1 ...
> # Check unique values in chocolate.brand
> unique(data$chocolate.brand)
[1] 2 1
> # Recode chocolate.brand to binary (0 and 1)
> data$chocolate.brand <- ifelse(data$chocolate.brand == 2, 0, 1)
> data$liked.feature=factor(data$liked.feature,levels = c('Taste',
+                                                          'Texture','Mood enhancement','Packag
e'),labels = c('1','2','3','4'))
> H0='chocolate brand means are equal'
> H1='chocolate brand means are not equal'
> model<- lm(cbind(data$cadbury.taste,data$nestle.taste,
+                  data$cadbury.texture,data$nestle.texture,data$liked.feature)
+            ~data$chocolate.brand)
> model


Call:
lm(formula = cbind(data$cadbury.taste, data$nestle.taste, data$cadbury.texture,
    data$nestle.texture, data$liked.feature) ~ data$chocolate.brand)

Coefficients:
                      [,1]      [,2]      [,3]      [,4]      [,5]
(Intercept)            3.37838   4.24324   3.48649   4.18919   1.62162
data$chocolate.brand   0.68348  -0.76902   0.43104  -0.44692   0.04848

> a=manova(model)
> summary(a,test='Pillai')
                      Df  Pillai approx F num Df den Df    Pr(>F)
data$chocolate.brand   1 0.37468   15.339      5    128 8.434e-12 ***
Residuals            132
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> Fcal<- summary(a,test='Pillai')$stats['data$chocolate.brand','approx F']
> Fcal
[1] 15.33883
> ftab=qf(0.95,5,128)
> ftab
[1] 2.28504
> if(Fcal>ftab){
+   print('we reject H0')
+ }else{
+   print("Accept H1")
+ }
[1] "we reject H0"
```

#CONTINGENCY TABLE WITH CHI-SQUARE TEST

```
> #contingency table
> H0=' There is no association between Gender and Chocolate Brand preference.'
> H1='There is an association between Gender and Chocolate Brand preference.'
> # Create contingency table
> contingency_table <- table(data$Gender, data$chocolate.brand)
> # Display the contingency table
> print(contingency_table)

         Cadbury Nestle
  Female      43     15
  Male        54     22
> # Perform chi-squared test
> chi_squared <- chisq.test(contingency_table)
> # Print the results
> print(chi_squared)

        Pearson's Chi-squared test with Yates' continuity correction

data:  contingency_table
X-squared = 0.040326, df = 1, p-value = 0.8408
```

As, the p-value is greater than 0.05.

Therefore we accept H0 and conclude that, there is no association between the gender and chocolate brand preference.

#ANOVA TESTING

H0:- there is no significant difference between the Chocolate brand preference

H1:- There is significant difference between the chocolate brand preference.

```
> #anova
> # Perform multi-factor ANOVA
> anova_result <- aov(data$chocolate.brand ~ data$cadbury.taste +data$nestle.taste+
+                     data$cadbury.texture+ data$nestle.texture +data$liked.feature)
> # Summary of ANOVA
> summary(anova_result)
                      Df Sum Sq Mean Sq F value   Pr(>F)    
data$cadbury.taste     1  2.449   2.449  19.011 2.67e-05 ***
data$nestle.taste      1  6.875   6.875  53.373 2.77e-11 ***
data$cadbury.texture   1  0.279   0.279   2.166   0.1436    
data$nestle.texture    1  0.407   0.407   3.161   0.0778 .  
data$liked.feature     3  0.544   0.181   1.407   0.2438    
Residuals            126 16.230   0.129                     
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value of Cadbury and nestle taste is less than the LOS(0.05)

Therefore they don’t have much impact on the model and as the texture and liked feature having p-value greater than the LOS we keep them in the model and they have significant effect on the model.

<u>And “hence there is significant difference in the chocolate brand preference according to the taste of the chocolate.”</u>

- ## CONCLUSION:-

In conclusion, our analysis of chocolate brands through customer surveys has provided valuable insights into consumer preferences and perceptions within the confectionery industry. Through descriptive analysis, we observed varying patterns in taste, texture, packaging preferences, and price considerations across different brands. Comparative analysis, including ANOVA testing, revealed significant differences between chocolate brands in terms of consumer perceptions The findings underscore the importance of understanding consumer preferences to remain competitive in the chocolate market. Companies can leverage this knowledge to refine product formulations, enhance packaging designs, and tailor marketing strategies to better resonate with target audiences