Title: Analyzing Burger King Survey & Customer Preferences

Mithil Sanjay kadam
23MBD012

Mithil Sanjay Kadam
23MBD012

I. Introduction

Burger King is best known for its flame-grilled burgers, offering a diverse menu that caters to various tastes and preferences. From the classic Whopper, featuring flame-grilled beef topped with fresh lettuce, tomatoes, onions, pickles, and creamy mayonnaise, to innovative creations like the Impossible Whopper, made with a plant-based patty, Burger King continues to innovate while staying true to its roots.

Main Objective:

The purpose of this survey is to gather valuable feedback from our loyal customers like you. Your insights will help us understand your preferences, expectations, and areas where we can improve our products and services. Whether you visit us for our iconic Whopper, our savory chicken sandwiches, or our delightful desserts, your feedback will guide us in providing you with the best possible Burger King experience.

Here are some customer preferences for burger king.

Burger Preferences:

Favorite burger options. Preferences for burger toppings . Interest in plant-based burger options (e.g., Impossible Whopper).

Menu Variety and Innovation:

Interest in limited-time menu offerings or promotions. Preferences for new menu items or innovations.

Value and Pricing: Perceived value for money spent. Interest in value meal deals or combo options.

Customization Options: Interest in customizable menu options (e.g., "Build Your Own Burger"). Preferences for sauce options and condiments.

II. Objective:

The objective of this study is to analyze a burger king survey to gain people's preference

Descriptive Analysis:

- Identify a Mean, Median, Mode, Variance, Standard Deviation Ratings of different variable (i.e is fav items and satisfaction).
- Identify Kurtosis and Skewness Ratings of different variable (i.e is rating & fav-items)

Graphical Representation:

- How does a bar plot illustrate the frequency of menu list and the frequency of Order Methods
- How does a pie chart represent the distribution of pizza consumption and restaurants among survey
- How does a boxplot show the frequency of Order method ad Menu list among survey
- How does a histogram display the frequency of menu list and restaurants among survey

Regression Analysis:

- Can we predict the probability of brand based on ingredients and rating (Logistic Regression)
- How does Rating, Quality ,satisfaction and ingredients based on brand using(MLR)
- How closely does the linear model fit the relationship between decision-to-visit and brand using (Least square method)

Hypotheses testing:

- Is there a significant difference in the mean frequency of Quality rating and brand test the hypothesis using (ANOVA)
- Are there Mean differences in the frequency of Restaurants based on satisfaction and Quality test with different types of tests of hypothesis (MANOVA)

Contingency Table:

- Contingency table with respect to gender and ingredients
- Contingency table with respect to brand and satisfaction
- Contingency table with respect to quality and most appealing & etc

III. Methodology:

a) Data Collection:

The data for dataset was collected by circulating google forms

Dataset of collected data in excel sheet:

B	C	D	E	F	G	H	I	J	K	L	M	N
Day	decision-t	favourite	satisfactio	most-appe	fav-item	rating	brand	quality	recipe	whopper	ingredient	gender
Once in a r	Menu opti	Whopper	Food quali	Combo Me	French frie	4	Burger Kin	Burger Kin	standardiz	4	Toppings	Male
Never	Proximity	Whopper	Food quali	Discounts	French frie	3	McDonald	McDonald	customize	4	Bun	Female
Once in a r	Menu opti	Hershey's	Food quali	Combo Me	Milkshakes	3	McDonald	Burger Kin	customize	4	Toppings	Male
Once in a r	Pricing	Whopper	Food quali	Combo Me	Softdrinks	4	Burger Kin	Burger Kin	standardiz	4	Meat	Female
Twice in a	Menu opti	Whopper	Food quali	Combo Me	French frie	4	Burger Kin	Burger Kin	standardiz	4	Toppings	Male
Once in a r	Proximity	Whopper	Food quali	Discounts	French frie	4	Burger Kin	Burger Kin	customize	3	Toppings	Female
Once in a r	Pricing	Whopper	Service spe	Combo Me	Softdrinks	4	McDonald	McDonald	standardiz	5	Meat	Male
Never	Menu opti	Whopper	Food quali	Combo Me	French frie	3	Burger Kin	Burger Kin	customize	4	Meat	Female
Once in a r	Menu opti	Crispy Chic	Food quali	Combo Me	Softdrinks	4	McDonald	Burger Kin	standardiz	3	Toppings	Female
Twice in a	Pricing	Crispy Chic	Food quali	Combo Me	Softdrinks	3	Burger Kin	Burger Kin	customize	3	Toppings	Female
Never	Pricing	Whopper	Cleanlines	Limited tin	Softdrinks	2	McDonald	McDonald	standardiz	3	Meat	Male
Never	Pricing	Crispy Chic	Food quali	Combo Me	French frie	3	McDonald	McDonald	standardiz	1	Toppings	Male
Once in a r	Proximity	Chicken Nu	Food quali	Combo Me	French frie	3	Burger Kin	Burger Kin	customize	3	Toppings	Female
Once in a r	Pricing	Chicken Nu	Food quali	Combo Me	Milkshakes	1	Burger Kin	Burger Kin	standardiz	1	Meat	Female
Once in a r	Menu opti	Chicken Nu	Food quali	Combo Me	Milkshakes	3	McDonald	McDonald	standardiz	4	Toppings	Male
Never	Menu opti	Whopper	Service spe	Combo Me	Milkshakes	1	McDonald	McDonald	standardiz	1	Bun	Male
Twice in a	Pricing	Crispy Chic	Food quali	Combo Me	French frie	5	Burger Kin	Burger Kin	customize	4	Meat	Female
Twice in a	Menu opti	Bacon King	Food quali	Combo Me	French frie	4	McDonald	McDonald	customize	2	Toppings	Male
Twice in a	Proximity	Bacon King	Cleanlines	Limited tin	Milkshakes	5	McDonald	McDonald	customize	3	Bun	Male
Never	Menu opti	Hershey's	Food quali	Combo Me	French frie	4	McDonald	Burger Kin	standardiz	3	Toppings	Female
Twice in a	Proximity	Whopper	Service spe	Discounts	Softdrinks	4	Burger Kin	Burger Kin	standardiz	4	Bun	Male
Once in a r	Menu opti	Whopper	Food quali	Combo Me	Milkshakes	1	McDonald	McDonald	customize	2	Bun	Female
Never	Menu opti	Chicken Nu	Cleanlines	Discounts	Salad	3	Burger Kin	Burger Kin	customize	3	Meat	Female
Once in a r	Pricing	Chicken Nu	Service spe	Discounts	Milkshakes	2	McDonald	McDonald	customize	2	Meat	Female
Once in a r	Menu opti	Whopper	Food quali	Combo Me	French frie	4	McDonald	McDonald	customize	4	Toppings	Male
Twice in a	Pricing	Whopper	Food quali	Combo Me	French frie	4	Burger Kin	Burger Kin	standardiz	5	Toppings	Female
Once in a r	Proximity	Whopper	Food quali	Combo Me	French frie	4	Burger Kin	Burger Kin	customize	4	Meat	Male
Never	Menu opti	Hershey's	Cleanlines	Combo Me	Softdrinks	4	Burger Kin	Burger Kin	standardiz	5	Meat	Male

Reading data from excel sheet

R code:

```
> data=read.csv("survey.csv")
> data
    X              Day decision.to.visit               favourite  satisfaction          most.appealing     fav.item rating       brand     quality
1  NA  Once in a month      Menu options                 Whopper  Food quality              Combo Meal French fries      4 Burger King Burger King
2  NA            Never         Proximity                 Whopper  Food quality               Discounts French fries      3  McDonald's  McDonald's
3  NA  Once in a month      Menu options    Hershey's Sundae Pie  Food quality              Combo Meal   Milkshakes      3  McDonald's Burger King
4  NA  Once in a month           Pricing                 Whopper  Food quality              Combo Meal   Softdrinks      4 Burger King Burger King
5  NA Twice in a month      Menu options                 Whopper  Food quality              Combo Meal French fries      4 Burger King Burger King
6  NA  Once in a month         Proximity                 Whopper  Food quality               Discounts French fries      4 Burger King Burger King
7  NA  Once in a month           Pricing                 Whopper Service speed              Combo Meal   Softdrinks      4  McDonald's  McDonald's
8  NA            Never      Menu options                 Whopper  Food quality              Combo Meal French fries      3 Burger King Burger King
9  NA  Once in a month      Menu options Crispy Chicken Sandwich  Food quality              Combo Meal   Softdrinks      4  McDonald's Burger King
10 NA Twice in a month           Pricing Crispy Chicken Sandwich  Food quality              Combo Meal   Softdrinks      3 Burger King Burger King
11 NA            Never           Pricing                 Whopper   Cleanliness Limited time menu items   Softdrinks      2  McDonald's  McDonald's
12 NA            Never           Pricing Crispy Chicken Sandwich  Food quality              Combo Meal French fries      3  McDonald's  McDonald's
13 NA  Once in a month         Proximity         Chicken Nuggets  Food quality              Combo Meal French fries      3 Burger King Burger King
14 NA  Once in a month           Pricing         Chicken Nuggets  Food quality              Combo Meal   Milkshakes      1 Burger King Burger King
15 NA  Once in a month      Menu options         Chicken Nuggets  Food quality              Combo Meal   Milkshakes      3  McDonald's  McDonald's
16 NA            Never      Menu options                 Whopper Service speed              Combo Meal   Milkshakes      1  McDonald's  McDonald's
17 NA  Twice in a week           Pricing Crispy Chicken Sandwich  Food quality              Combo Meal French fries      5 Burger King Burger King
18 NA  Twice in a week      Menu options              Bacon King  Food quality              Combo Meal French fries      4  McDonald's  McDonald's
19 NA Twice in a month         Proximity              Bacon King   Cleanliness Limited time menu items   Milkshakes      5  McDonald's  McDonald's
20 NA            Never      Menu options    Hershey's Sundae Pie  Food quality              Combo Meal French fries      4  McDonald's Burger King
21 NA Twice in a month         Proximity                 Whopper Service speed               Discounts   Softdrinks      4 Burger King Burger King
```

```
> head(data)
   X             Day decision.to.visit           favourite satisfaction most.appealing     fav.item rating       brand     quality             recipe whopper
1 NA  Once in a month      Menu options             Whopper Food quality     Combo Meal French fries      4 Burger King Burger King standardized recipe       4
2 NA           Never         Proximity             Whopper Food quality      Discounts French fries      3  McDonald's  McDonald's    customize recipe       4
3 NA  Once in a month      Menu options Hershey's Sundae Pie Food quality     Combo Meal   Milkshakes      3  McDonald's Burger King    customize recipe       4
4 NA  Once in a month           Pricing             Whopper Food quality     Combo Meal   Softdrinks      4 Burger King Burger King standardized recipe       4
5 NA Twice in a month      Menu options             Whopper Food quality     Combo Meal French fries      4 Burger King Burger King standardized recipe       4
6 NA  Once in a month         Proximity             Whopper Food quality      Discounts French fries      4 Burger King Burger King    customize recipe       3
  ingredients gender
1    Toppings   Male
2         Bun Female
3    Toppings   Male
4        Meat Female
5    Toppings   Male
6    Toppings Female
> |
```

To remove one unwanted column from data set

```
> data<-subset(data,select= -X)
> data
```

```
> head(data)
             Day decision.to.visit           favourite satisfaction most.appealing     fav.item rating       brand     quality             recipe whopper
1  Once in a month      Menu options             Whopper Food quality     Combo Meal French fries      4 Burger King Burger King standardized recipe       4
2           Never         Proximity             Whopper Food quality      Discounts French fries      3  McDonald's  McDonald's    customize recipe       4
3  Once in a month      Menu options Hershey's Sundae Pie Food quality     Combo Meal   Milkshakes      3  McDonald's Burger King    customize recipe       4
4  Once in a month           Pricing             Whopper Food quality     Combo Meal   Softdrinks      4 Burger King Burger King standardized recipe       4
5 Twice in a month      Menu options             Whopper Food quality     Combo Meal French fries      4 Burger King Burger King standardized recipe       4
6  Once in a month         Proximity             Whopper Food quality      Discounts French fries      4 Burger King Burger King    customize recipe       3
  ingredients gender
1    Toppings   Male
2         Bun Female
3    Toppings   Male
4        Meat Female
5    Toppings   Male
6    Toppings Female
```

To convert some categorical values into numeric values

```
data$favourite=factor(data$favourite,
                      levels = c('Whopper','Bacon King','Impossible Whopper','Crispy Chicken Sandwich'
                                 ,'Chicken Nuggets','Hersheys Sundae Pie'),
                      labels = c(4,5,6,7,8,9))
data$favourite=as.numeric(data$favourite)
data$decision.to.visit=factor(data$decision.to.visit,
                                  levels = c('Proximity','Menu options','Pricing'),
                                  labels = c(1,2,3))
data$decision.to.visit=as.numeric(data$decision.to.visit)

data$favourite=factor(data$favourite,
                      levels = c('Whopper','Bacon King','Impossible Whopper','Crispy Chicken Sandwich'
                                 ,'Chicken Nuggets','Hersheys Sundae Pie'),
                      labels = c(4,5,6,7,8,9))
data$favourite=as.numeric(data$favourite)
```

b) Data Analysis:

Descriptive Analysis: Summarize the data using descriptive statistics such as mean, median, mode, range, and standard deviation.

Hypothesis testing: Formulate hypotheses based on the relationships observed in the EDA. Use statistical tests such as chi-square test, ANOVA and MANOVA to test the hypotheses and determine if the relationship are statistically significant.

Regression Analysis: Perform regression analysis to understand the impact of independent variables on dependent variable. Use multiple linear regression for analyzing relationships involving multiple independent variable Use linear regression for analyzing relationships involving one independent variable.

Graphical Representation: Create visualization (e.g, charts, graphs, etc.) to present key findings and insight based on the survey data.

C) Visualization:

This contains all the visualization tools which I have used under this analysis.

1- Bar graph

A bar graph also known as a bar chart. It is a graphical representation of data using rectangular bars or columns. Each bar represents a category or group, and the height or length of the bar corresponds to the value or frequency of the data it represents. Bar graphs are commonly used to compare and display information across different categories or to show changes in data over time.

2- Pie chart

A pie chart is a circular statistical graphic that is divided into slices to illustrate numerical proportions. Each slice represents a proportionate part of the whole, and the size of each slice is proportional to the quantity it represents. Pie charts are commonly used to visualize the distribution of a single categorical variable or to show the relative proportions of different categories within a dataset.

3- Box Plot

A box plot, also known as a box-and-whisker plot, is a graphical representation of the distribution of a dataset along a single numerical axis. It provides a concise summary of the central tendency, dispersion, and skewness of the data, as well as the presence of outliers. Box plots are particularly useful for comparing the distributions of multiple datasets or for identifying patterns.

4- Histogram plot

A histogram plot is a graphical representation of the distribution of numerical data. It displays the frequencies or counts of observations within different intervals, or "bins," along a single numerical axis. Histograms are particularly useful for understanding the shape, central tendency, and spread of a dataset, as well as identifying patterns and outliers.

5- Line Graph

A line graph is a type of chart that displays data points connected by straight lines. It is particularly useful for showing trends and changes in data over time or across different categories. Line graphs consist of two axes – a horizontal axis (x-axis) and a vertical axis (y-axis) – with data points plotted at specific coordinates corresponding to the values ovariables being measured.

IV. Results and Analysis

1- Descriptive Analysis

- Analysis based on rating column

#Calculate Mean

```
> mean(data$rating)
[1] 3.435644
```

#Calculate Median

```
> median(data$rating)
[1] 4
```

#Calculate Mode

```
> mode<-function(v){
+   uniqv<-unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> result<-mode(data$rating)
> print(result)
[1] 4
```

#Calculate Interquartile Range

```
> IQR(data$rating)
[1] 1
>
```

#Calculate Quantile

```
> quantile(data$rating)
  0%  25%  50%  75% 100%
   1    3    4    4    5
>
```

#Calculate difference

```
> diff(range(data$rating))
[1] 4
```

#Calculate Sandard Deviation

```
> sd(data$rating)
[1] 1.00415
>
```

#Calculate Variance

```
> var(data$rating)
[1] 1.008317
```

#Calculate Skewness and Kurtorsis

```
> # Load the moments package
> library(moments)
> a<- data$rating
> # Calculate skewness and kurtosis for the rating feature
> skew <- skewness(a)
> kurt <- kurtosis(a)
> print(paste("Skewness of rating:", skew))
[1] "Skewness of rating: -0.477734071247052"
> print(paste("Kurtosis of rating:", kurt))
[1] "Kurtosis of rating: 2.99427856582634"
```

Conclusion:

The mean rating is of 3 and the median is 4. The data is negatively skewneed and kurtosis is equal to 3 which means that the data is mesokurtic which indicates that the rating received by the burger king is overall good. With the help of chi-square test we come to know that actual rating and expected rating where almost same.

- Analysis based on whopper column

#Calculate Mean

```
> mean(data$whopper)
[1] 3.306931
>
```

#Calculate Median

```
> median(data$whopper)
[1] 3
```

#Calculate Mode

```
> mode<-function(v){
+   uniqv<-unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> result<-mode(data$whopper)
> print(result)
[1] 4
```

#Calculate Interquartile Range

```
> IQR(data$whopper)
[1] 1
>
```

#Calculate Quantile

```
> quantile(data$whopper)
  0%  25%  50%  75% 100%
   1    3    3    4    5
>
```

#Calculate Difference

```
> diff(range(data$whopper))
[1] 4
```

#Calculate Sandard Deviation

```
> sd(data$whopper)
[1] 1.129093
>
```

#Calculate Variance

```
> var(data$whopper)
[1] 1.274851
```

#Calculate Skewness and Kurtorsis

```
> library(moments)
> a<- data$whopper
> # Calculate skewness and kurtosis for the whopper feature
> skew <- skewness(a)
> kurt <- kurtosis(a)
> print(paste("Skewness of whopper:", skew))
[1] "Skewness of whopper: -0.581643016090925"
> print(paste("Kurtosis of whopper:", kurt))
[1] "Kurtosis of whopper: 2.74879016626293"
```

- Analysis based on satisfaction column

#Calculate Mean

```
> mean(data$satisfaction)
[1] 1.465347
```

#Calculate Median

```
> median(data$satisfaction)
[1] 1
```

#Calculate Mode

```
> mode<-function(v){
+   uniqv<-unique(v)
+   uniqv[which.max(tabulate(match(v,uniqv)))]
+ }
> result<-mode(data$satisfaction)
> print(result)
[1] 1
```

#Calculate Interquartile Range

```
> IQR(data$satisfaction)
[1] 1
```

#Calculate Quantile

```
> quantile(data$satisfaction)
  0%  25%  50%  75% 100%
   1    1    1    2    3
```

#Calculate Difference

```
> diff(range(data$satisfaction))
[1] 2
```

#Calculate Sandard Deviation

```
> sd(data$satisfaction)
[1] 0.7288945
```

#Calculate Variance

```
> var(data$satisfaction)
[1] 0.5312871
```

#Calculate Skewness and Kurtorsis

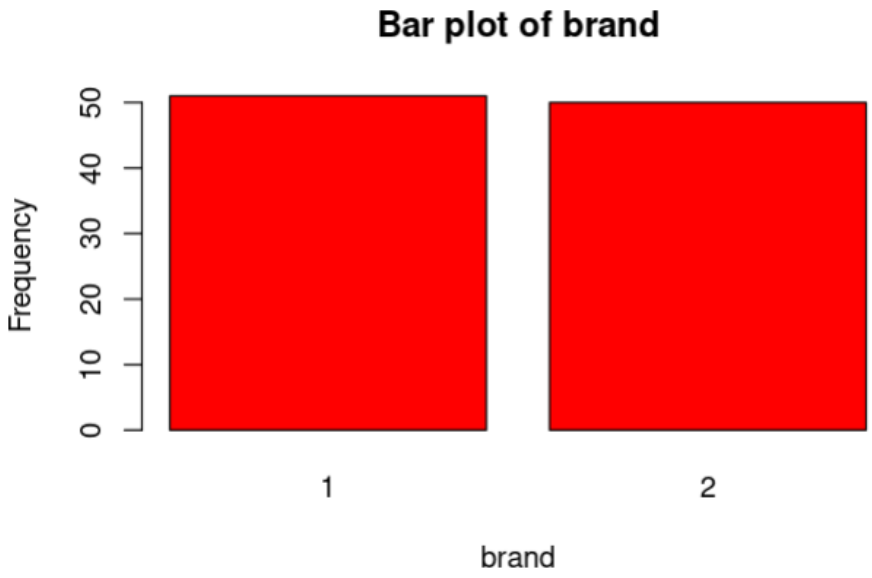
```
> skew <- skewness(a)
> kurt <- kurtosis(a)
> print(paste("Skewness of satisfaction:", skew))
[1] "Skewness of satisfaction: 1.21071389704404"
> print(paste("Kurtosis of satisfaction:", kurt))
[1] "Kurtosis of satisfaction: 2.94828963411627"
>
```

2- Graphical Representation

- Bar Plot

Bar plot based on which brand has more rating (i.e MaDonald's & Burger King

```
barplot(table(data$brand),col = "red",xlab="brand", ylab="Frequency",
        main = 'Bar plot of brand')
```



Conclusion: Above bar graph shows that Burger king has high frequency then Mcdonald's.

#bar plot based on which ingredients has better quality.

```
barplot(table(data$ingredients),col = "green",xlab="ingredients", ylab="Frequency"
        ,main = 'bar plot of ingredients')
```

Conclusion: Above bar graph shows that toppings has high frequency in ingredients**.**

#bar graph on what bases did burger king based on satisfaction

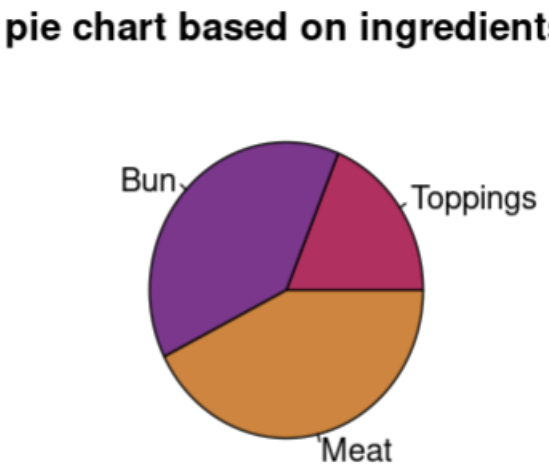
```
barplot(table(data$satisfaction),col = "pink",xlab="satisfaction", ylab="Frequency"
        ,main = 'bar plot of satisfaction')
```

- Pie Plot

```
pie(table(data$Day),labels=c('Once in a month','Twice in a week','Twice in a month','Never',
                             'Once in a week'),col = c('maroon','mediumorchid4',
                            'chocolate1','peru','blue'),main='pie chart based on day')
```

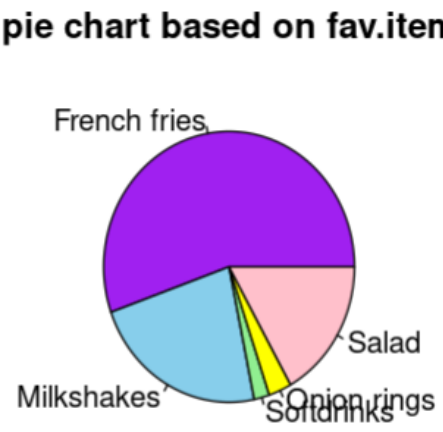
Conclusion: As per above visualization more number of customer goes to burger king Twice in a week.

```
pie(table(data$ingredients),labels=c('Toppings','Bun','Meat')
    ,col = c('maroon','mediumorchid4','peru'),main='pie chart based on ingredients')
```



Conclusion: As per above visualization , from all ingredients meat is good in quality

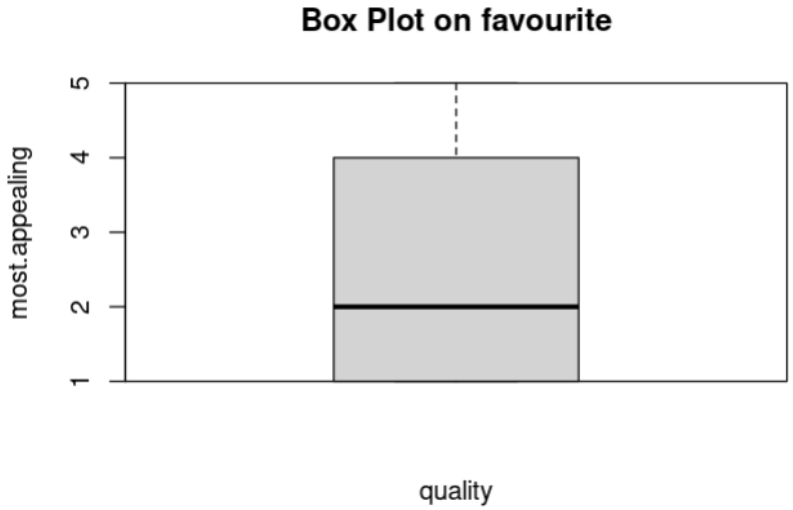
```
pie(table(data$fav.item),labels=c('French fries','Milkshakes','Softdrinks','Onion rings',
                                  'Salad'),col = c('purple','skyblue',
                                 'lightgreen','yellow','pink'),main='pie chart based on fav.item')
```



Conclusion: As per above visualization ,French Fries is more selling with burgers whereas least number of customers preffered Soft drinks.

- Box Plot

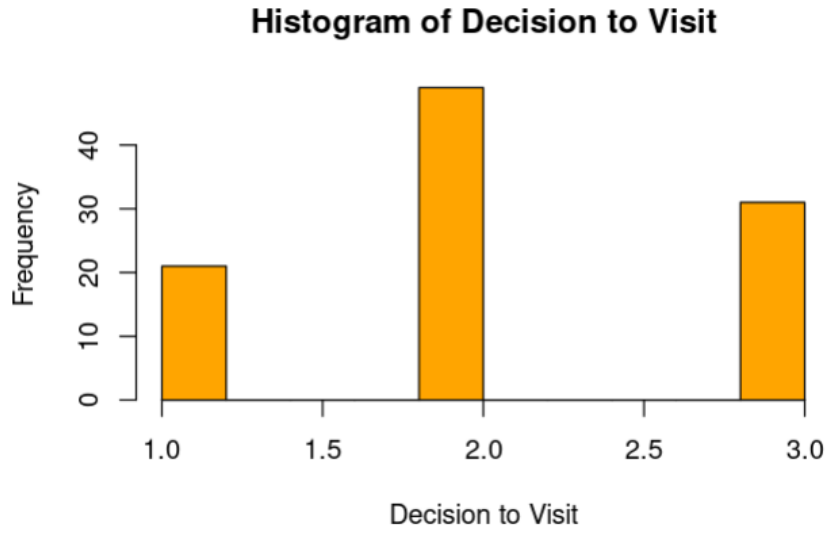
```
> boxplot(data$favourite, main = "Box Plot on favourite", xlab="quality",ylab = "most.appealing")
```



Conclusion: As per the above fig. shown the line lie on 2 which shows the average value for the skewness the curve can be form is Positively Skewed

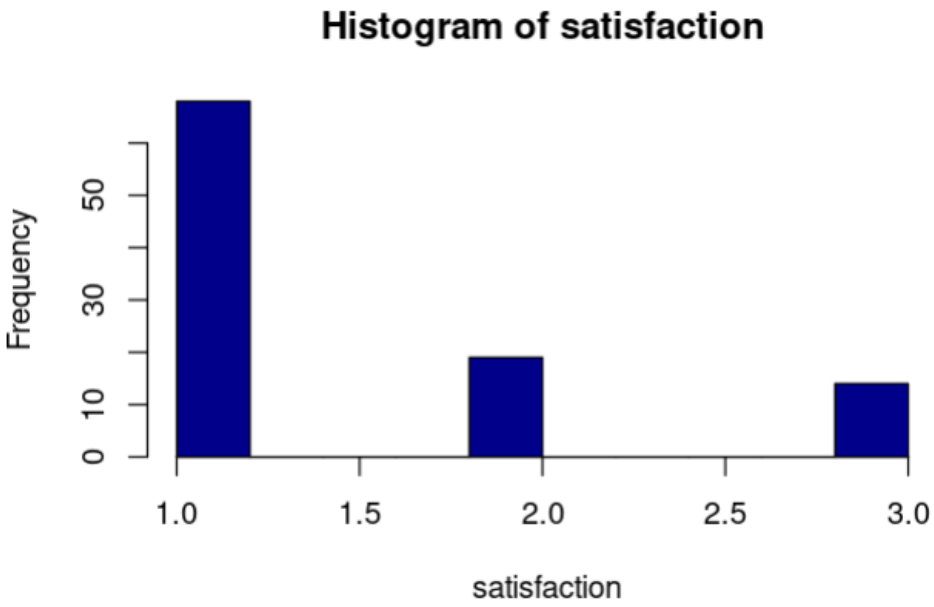
- Histogram plot

```
hist(data$decision.to.visit,xlab = "Decision to Visit",col="orange",
     main = "Histogram of Decision to Visit")
```



Conclusion: So here decision to visit in based on proximity, Menu options & pricing. So the above hist. graph shows that menu options has high frequency for decision to visit burger king.

```
hist(data$satisfaction,xlab = " satisfaction",col="red",
main = "Histogram of satisfaction")
```



Conclusion: Here satisfaction is based on Food quality , Service speed , & Cleanliness. So the above graph shows that Food quality has high frequency.

- Line Graph

```
plot(data$rating, type = "l", col = "maroon",
     xlab = "Index", ylab = "rating",
     main = "Line Plot of rating burger king")
```

Conclusion: Here rating is been shown through line graph .It shows the frequency ratings.

3) Performing predictions

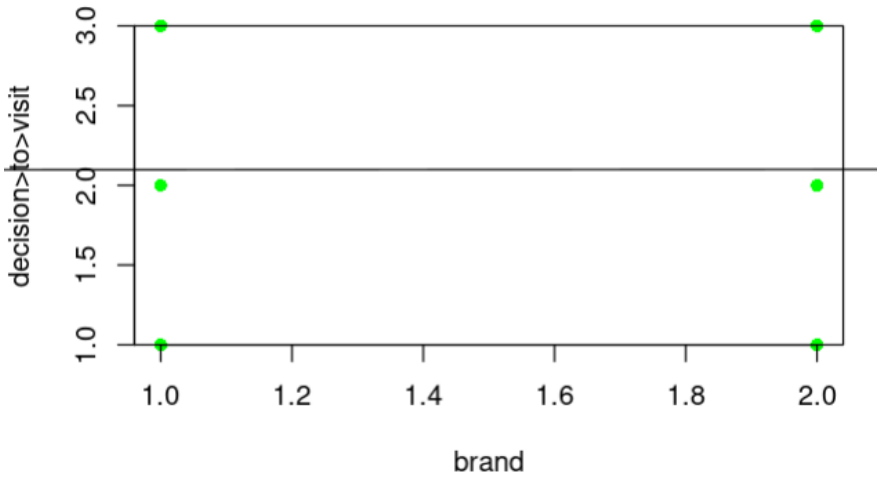
- Least Square Method

```
model=lm(decision.to.visit~brand,data = data)
summary(model)
```

```
Call:
lm(formula = decision.to.visit ~ brand, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-1.10000 -0.10000 -0.09804  0.90000  0.90196

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.096078   0.225204   9.307 3.55e-15 ***
brand       0.001961   0.142856   0.014    0.989
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7178 on 99 degrees of freedom
Multiple R-squared:  1.903e-06, Adjusted R-squared:  -0.0101
F-statistic: 0.0001884 on 1 and 99 DF,  p-value: 0.9891
```

```
plot(data$brand,data$decision.to.visit,pch=16,xlab="brand", ylab="decision>to>visit",
     col='green')
```

```
abline(model)
```



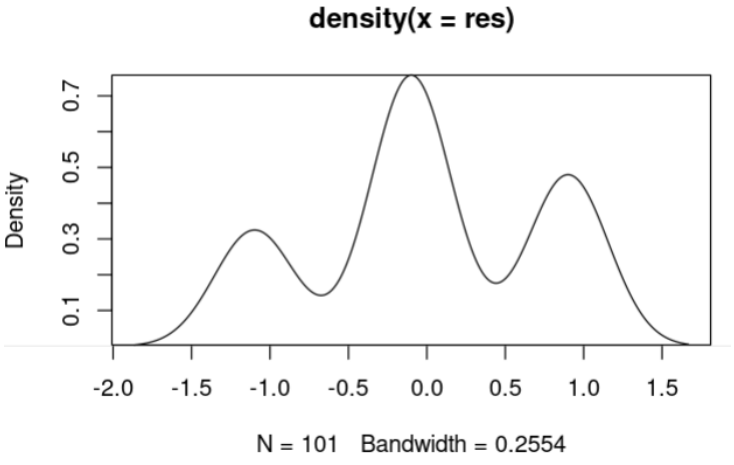

```
> res=resid(model)
> res
          1           2           3           4           5           6           7
-0.09803922 -1.10000000 -0.10000000  0.90196078 -0.09803922 -1.09803922  0.90000000
          8           9          10          11          12          13          14
-0.09803922 -0.10000000  0.90196078  0.90000000  0.90000000 -1.09803922  0.90196078
         15          16          17          18          19          20          21
-0.10000000 -0.10000000  0.90196078 -0.10000000 -1.10000000 -0.10000000 -1.09803922
         22          23          24          25          26          27          28
```

```
> summary(model)

Call:
lm(formula = decision.to.visit ~ brand, data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-1.10000 -0.10000 -0.09804  0.90000  0.90196

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.096078   0.225204   9.307 3.55e-15 ***
brand       0.001961   0.142856   0.014    0.989
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7178 on 99 degrees of freedom
Multiple R-squared:  1.903e-06, Adjusted R-squared:  -0.0101
F-statistic: 0.0001884 on 1 and 99 DF,  p-value: 0.9891
```

Conclusion:

- As for the least squared method variable required to be in numeric type so the above code is to convert categorical value into numeric.
- Here, we find the linear regression between Gender and Menu list on line y=b+ax Equation will be: Y=1.74075-0.01947x
- As per the density plot show the residuals present into data with Number of elements with errors and their Bandwidth.
- These are the residuals which are in the data as per least squared method.

```
plot(density(res))
```



- Multiple Linear Regression(MLR)

#Backward Direction

```
> data$brand <- ifelse(data$brand == "McDonald,s", 1, 2)
> m=lm(data$brand~data$ingredients+data$satisfaction+data$rating)
> summary(m)

Call:
lm(formula = data$brand ~ data$ingredients + data$satisfaction +
    data$rating)

Residuals:
       Min         1Q     Median         3Q        Max
-7.402e-16 -5.363e-16 -1.473e-16  3.850e-17  2.168e-14

Coefficients:
                          Estimate Std. Error    t value Pr(>|t|)
(Intercept)              2.000e+00  1.120e-15  1.786e+15   <2e-16 ***
data$ingredientsMeat     3.110e-17  6.304e-16  4.900e-02    0.961
data$ingredientsToppings 5.220e-16  6.193e-16  8.430e-01    0.401
data$satisfaction       -1.548e-16  3.122e-16 -4.960e-01    0.621
data$rating              1.019e-16  2.272e-16  4.490e-01    0.655
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.245e-15 on 96 degrees of freedom
Multiple R-squared:  0.5007,    Adjusted R-squared:  0.4799
F-statistic: 24.06 on 4 and 96 DF,  p-value: 8.348e-14


> #ingredients removed
> m=lm(data$brand~data$satisfaction+data$rating)
> summary(m)

Call:
lm(formula = data$brand ~ data$satisfaction + data$rating)

Residuals:
       Min         1Q     Median         3Q        Max
-4.660e-16 -3.607e-16 -2.555e-16 -8.270e-17  2.195e-14

Coefficients:
                   Estimate Std. Error    t value Pr(>|t|)
(Intercept)       2.000e+00  9.834e-16  2.034e+15   <2e-16 ***
data$satisfaction -1.727e-16  3.106e-16 -5.560e-01    0.579
data$rating        1.052e-16  2.255e-16  4.670e-01    0.642
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.236e-15 on 98 degrees of freedom
Multiple R-squared:  0.5028,    Adjusted R-squared:  0.4927
F-statistic: 49.56 on 2 and 98 DF,  p-value: 1.345e-15
```

```
> m=lm(data$brand~data$rating)
> summary(m)

Call:
lm(formula = data$brand ~ data$rating)

Residuals:
       Min         1Q     Median         3Q        Max
-4.163e-16 -2.914e-16 -1.665e-16 -1.665e-16  2.202e-14

Coefficients:
             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 2.000e+00  7.939e-16 2.519e+15   <2e-16 ***
data$rating 1.249e-16  2.219e-16 5.630e-01    0.575
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.228e-15 on 99 degrees of freedom
Multiple R-squared:  0.5008,    Adjusted R-squared:  0.4958
F-statistic: 99.33 on 1 and 99 DF,  p-value: < 2.2e-16
```

#Using Forward Direction

```
> #forward selection
> m=lm(data$brand~NULL)
> summary(m)

Call:
lm(formula = data$brand ~ NULL)

Residuals:
       Min         1Q     Median         3Q        Max
-2.209e-16 -2.209e-16 -2.209e-16 -2.209e-16  2.209e-14

Coefficients:
             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 2.000e+00  2.209e-16 9.052e+15   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.22e-15 on 100 degrees of freedom
```

```
> m=lm(data$brand~data$ingredients)
> summary(m)

Call:
lm(formula = data$brand ~ data$ingredients)

Residuals:
      Min        1Q    Median        3Q       Max
-5.19e-16 -5.19e-16  0.00e+00  0.00e+00  2.18e-14

Coefficients:
                         Estimate Std. Error   t value Pr(>|t|)
(Intercept)             2.000e+00  5.111e-16 3.913e+15   <2e-16 ***
data$ingredientsMeat    4.514e-30  6.233e-16 0.000e+00      1.0
data$ingredientsToppings 5.190e-16  6.137e-16 8.460e-01      0.4
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.228e-15 on 98 degrees of freedom
Multiple R-squared:  0.4976,    Adjusted R-squared:  0.4874
F-statistic: 48.54 on 2 and 98 DF,  p-value: 2.24e-15
```

```
> m=lm(data$brand~data$satisfaction)
> summary(m)

Call:
lm(formula = data$brand ~ data$satisfaction)

Residuals:
       Min         1Q     Median         3Q        Max
-3.119e-16 -3.119e-16 -3.119e-16 -1.164e-16  2.200e-14

Coefficients:
                   Estimate Std. Error    t value Pr(>|t|)
(Intercept)        2.000e+00  4.996e-16  4.004e+15   <2e-16 ***
data$satisfaction -1.955e-16  3.055e-16 -6.400e-01    0.524
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.227e-15 on 99 degrees of freedom
Multiple R-squared:  0.5028,    Adjusted R-squared:  0.4977
F-statistic: 100.1 on 1 and 99 DF,  p-value: < 2.2e-16
```

```
> m=lm(data$brand~data$rating)
> summary(m)

Call:
lm(formula = data$brand ~ data$rating)

Residuals:
       Min         1Q     Median         3Q        Max
-4.163e-16 -2.914e-16 -1.665e-16 -1.665e-16  2.202e-14

Coefficients:
             Estimate Std. Error   t value Pr(>|t|)
(Intercept) 2.000e+00  7.939e-16 2.519e+15   <2e-16 ***
data$rating 1.249e-16  2.219e-16 5.630e-01    0.575
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.228e-15 on 99 degrees of freedom
Multiple R-squared:  0.5008,    Adjusted R-squared:  0.4958
F-statistic: 99.33 on 1 and 99 DF,  p-value: < 2.2e-16
```

Conclusion:

- As for Multiple linear regression the variable required to be in numeric type so the above code is to convert categorical value into numeric.
- As per the summary provided above in Multiple Linear regression, we remove the variable whose probability value is higher, as per the above we will remove p value which is maximum.
- As per the above we will remove Rating because it has p value which is maximum
- So, as per the above output rating is more suitable with the brand this is called Backward Selection Method.

- Logistic Regression

```
> #making logistic regession
> data$brand <- ifelse(data$brand == "McDonald,s", 1, 0)
> model=glm(data$brand~data$ingredients+data$rating,data=data,family="binomial")
Warning message:
glm.fit: algorithm did not converge
> model

Call:  glm(formula = data$brand ~ data$ingredients + data$rating, family = "binomial",
    data = data)

Coefficients:
             (Intercept)      data$ingredientsMeat  data$ingredientsToppings
              -2.657e+01                -3.423e-16                -2.231e-15
             data$rating
              -1.394e-15

Degrees of Freedom: 100 Total (i.e. Null);  97 Residual
Null Deviance:      0
Residual Deviance: 5.86e-10     AIC: 8
>
```

```
> prediction=predict(model,type="response")
> head(prediction)
           1            2            3            4            5            6
2.900701e-12 2.900701e-12 2.900701e-12 2.900701e-12 2.900701e-12 2.900701e-12
> accuracy=mean((prediction>0.5)==data$brand)
> accuracy
[1] 1
```

Conclusion:

- Here we have made predictions that how are model is made and is it a good fit or not . we make predictions in model .
- So the output shows that the accuracy value is 1 that means our model is best fit because its value is 1 .
- It it was 0 or close to zero than it will not be a good model.

4) Hypothesis Testing

- Using Manova Test

Importing libraries which are needed for manova testing

```
library(tidyverse)
library(ggpubr)
library(rstatix)
library(car)
library(broom)
library(datarium)
```

```
H0="the three populations have same mean vector"
H1="the three populations do not have same mean vector"
```

Using Wilk's Largest Root

```
> res.man=manova(cbind(data$rating,data$satisfaction)~ingredients,data=df)
> a=summary(res.man,test="Wilks")
> a
            Df   Wilks approx F num Df den Df Pr(>F)
ingredients  2 0.98898  0.26947      4    194 0.8974
Residuals   98
> pvalue=a$stats['ingredients','Pr(>F)']#take P value from a
> pvalue
[1] 0.8973691
> los=0.05
> if(pvalue>los){
+   print('accept H0')
+   print(H0)
+ }else{
+   print('cannot accept H0')
+   print(H1)
+ }
[1] "accept H0"
[1] "the three populations have same mean vector"
> fcal=a$stats['ingredients','approx F']#take approx x value
> fcal
[1] 0.2694653
> ftab=qf(0.95,4, 194)#take los value,Df and den Df
> ftab
[1] 2.418202
> if(fcal<ftab){
+   print('accept H0')
+   print(H0)
+ }else{
+   print('cannot accept H0')
+   print(H1)
+ }
[1] "accept H0"
[1] "the three populations have same mean vector"
>
```

Using Roy Largest Root

```
> b=summary(res.man,test="Roy")
> b
            Df       Roy approx F num Df den Df Pr(>F)
ingredients  2 0.0089065  0.43642      2     98 0.6476
Residuals   98
> #for roy
> pvalue=b$stats['ingredients','Pr(>F)']#take P value from b
> pvalue
[1] 0.6475977
> los=0.05
> if(pvalue>los){
+   print('accept H0')
+   print(H0)
+ }else{
+   print('cannot accept H0')
+   print(H1)
+ }
[1] "accept H0"
[1] "the three populations have same mean vector"
> fcal=b$stats['ingredients','approx F']#take approx x value
> fcal
[1] 0.4364177
> ftab=qf(0.95,2,98)#take los value,Df and den Df
> ftab
[1] 3.089203
> if(fcal<ftab){
+   print('accept H0')
+   print(H0)
+ }else{
+   print('cannot accept H0')
+   print(H1)
+ }
[1] "accept H0"
[1] "the three populations have same mean vector"
```

Conclusion:

- MANOVA testing is used test the impact of Two independent variable on one depending variable.
- This can be done by using various testing type but here I have used only two types of testing that is Wilks and Roy and compare their probability value with level of significance.
- As per to perform test we denoted H0 i.e Null hypothesis and H1 i.e Alternative hypothesis.
- If the pvalue is greater than level of significance than, we accept H0 otherwise reject the H0 and accept H1.

5) Contingency Table

```
> x=table(data$gender,data$ingredients)
> x

         Bun Meat Toppings
  Female  10   18       20
  Male     9   21       23

> y=table(data$brand,data$satisfaction)
> y

    1  2  3
  0 68 19 14
> z=table(data$rating,data$brand)
> z

    0
  1  5
  2 10
  3 35
  4 38
  5 13
> d=table(data$decision.to.visit,data$fav.item)
> d

    French fries Milkshakes Onion rings Salad Softdrinks
  1           13          4           1     0          3
  2           25         14           1     3          6
  3           18          5           0     0          8
> f=table(data$quality,data$most.appealing)
> f

             Combo Meal Discounts Limited time menu items
  Burger King         39        13                     6
  McDonald's          23        14                     6
>
```

Conclusion: The Contingency tables shows the relationship between two two columns .

V. Final Conclusion in summary:

The analysis reveals several key insights about customer behaviour and preferences at burger king:-

Overall Satisfaction: It shows that a significant portion of customers are satisfied with their experience at Burger King. This satisfaction can be attributed to factors such as food quality, menu variety, and customer service.

Influence Factors: The analysis suggests that the primary influencing factor for people choosing to visit Burger King is its menu options. This aligns with Burger King's commitment to innovation and continually evolving its menu to cater to changing consumer tastes**.**

Areas for Improvement: While overall satisfaction is high, there are areas where Burger King can improve. These may include wait times, cleanliness of the restaurants, and the accuracy of orders.

Regression Analysis: The regression model shows a good assoiciation all factor.

Customer Preferences: This reveals insights into customer preferences regarding menu items, promotional offers, and restaurant ambiance. Understanding these preferences can help Burger King tailor its offerings to better meet customer needs.

Recommendations: Recommendations has been made from enhancing the Burger King experience. This may include initiatives to improve service speed, introduce new menu items, or revamp restaurant layouts.

Gender Preferences: The contingency table analysis reveals some differences in preferences between genders. For example, a slightly higher proportion of males visit Burger King once a month compared to females. Additionally, there are variations in favourite menu items and ratings between genders, although the differences are not significant.

Overall, the Burger King survey serves as a valuable tool for gauging customer satisfaction, identifying areas for improvement, and informing strategic decision-making to enhance the overall dining experience.