R Notebook

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```
# Final Project

#Upload the data
#Find the dimensions of this data set.
#Summarize the data and showcase the structure of this dataset.
getwd()
```

```
## [1] "/cloud/project"
```

```
fb.df<-read.csv("facebook.csv")
str(fb.df)
```

```
## 'data.frame':    1143 obs. of  11 variables:
##  $ ad_id               : int  708746 708749 708771 708815 708818 708820 708889 708895 708953 708958 .
##  $ xyz_campaign_id     : int  916 916 916 916 916 916 916 916 916 916 ...
##  $ fb_campaign_id      : int  103916 103917 103920 103928 103928 103929 103940 103941 103951 103952 .
##  $ age                 : chr  "30-34" "30-34" "30-34" "30-34" ...
##  $ gender              : chr  "M" "M" "M" "M" ...
##  $ interest            : int  15 16 20 28 28 29 15 16 27 28 ...
##  $ Impressions         : int  7350 17861 693 4259 4133 1915 15615 10951 2355 9502 ...
##  $ Clicks              : int  1 2 0 1 1 0 3 1 1 3 ...
##  $ Spent               : num  1.43 1.82 0 1.25 1.29 ...
##  $ Total_Conversion    : int  2 2 1 1 1 1 1 1 1 1 ...
##  $ Approved_Conversion: int  1 0 0 0 1 1 0 1 0 0 ...
```

```
summary(fb.df)
```

```
##      ad_id          xyz_campaign_id fb_campaign_id       age
##  Min.   : 708746   Min.   : 916    Min.   :103916   Length:1143
##  1st Qu.: 777632   1st Qu.: 936    1st Qu.:115716   Class :character
##  Median :1121185   Median :1178    Median :144549   Mode  :character
##  Mean   : 987261   Mean   :1067    Mean   :133784
##  3rd Qu.:1121804   3rd Qu.:1178    3rd Qu.:144658
##  Max.   :1314415   Max.   :1178    Max.   :179982
##     gender            interest       Impressions          Clicks
##  Length:1143        Min.   :  2.00   Min.   :     87   Min.   :  0.00
##  Class :character   1st Qu.: 16.00   1st Qu.:   6504   1st Qu.:  1.00
##  Mode  :character   Median : 25.00   Median :  51509   Median :  8.00
##                     Mean   : 32.77   Mean   : 186732   Mean   : 33.39
##                     3rd Qu.: 31.00   3rd Qu.: 221769   3rd Qu.: 37.50
##                     Max.   :114.00   Max.   :3052003   Max.   :421.00
##      Spent        Total_Conversion Approved_Conversion
##  Min.   :  0.00   Min.   : 0.000   Min.   : 0.000
```

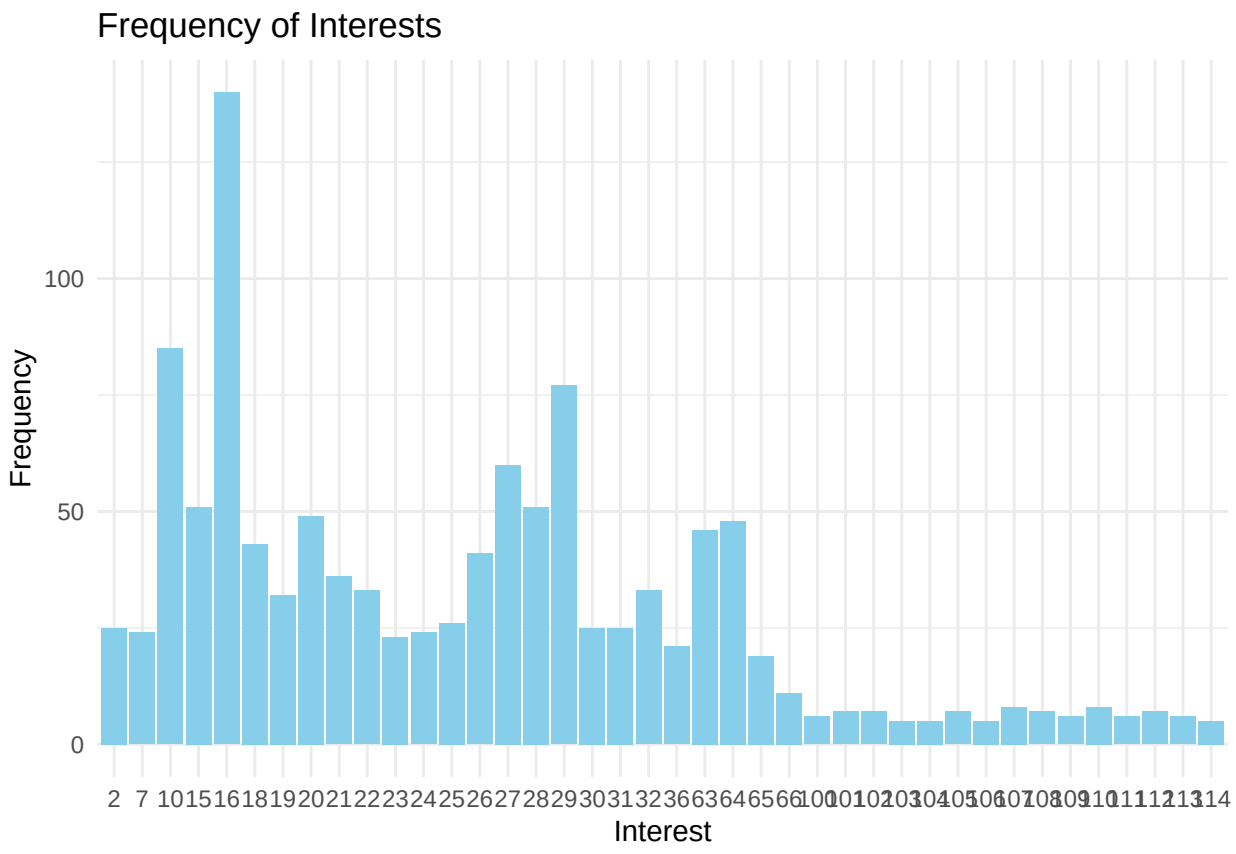
```
##  1st Qu.:  1.48   1st Qu.: 1.000   1st Qu.: 0.000
##  Median : 12.37   Median : 1.000   Median : 1.000
##  Mean   : 51.36   Mean   : 2.856   Mean   : 0.944
##  3rd Qu.: 60.02   3rd Qu.: 3.000   3rd Qu.: 1.000
##  Max.   :639.95   Max.   :60.000   Max.   :21.000
```

```r
#Demographic Information: Age and gender of the target audience.
age.gender.table<- table(fb.df$age, fb.df$gender)
prop.table(age.gender.table)
```

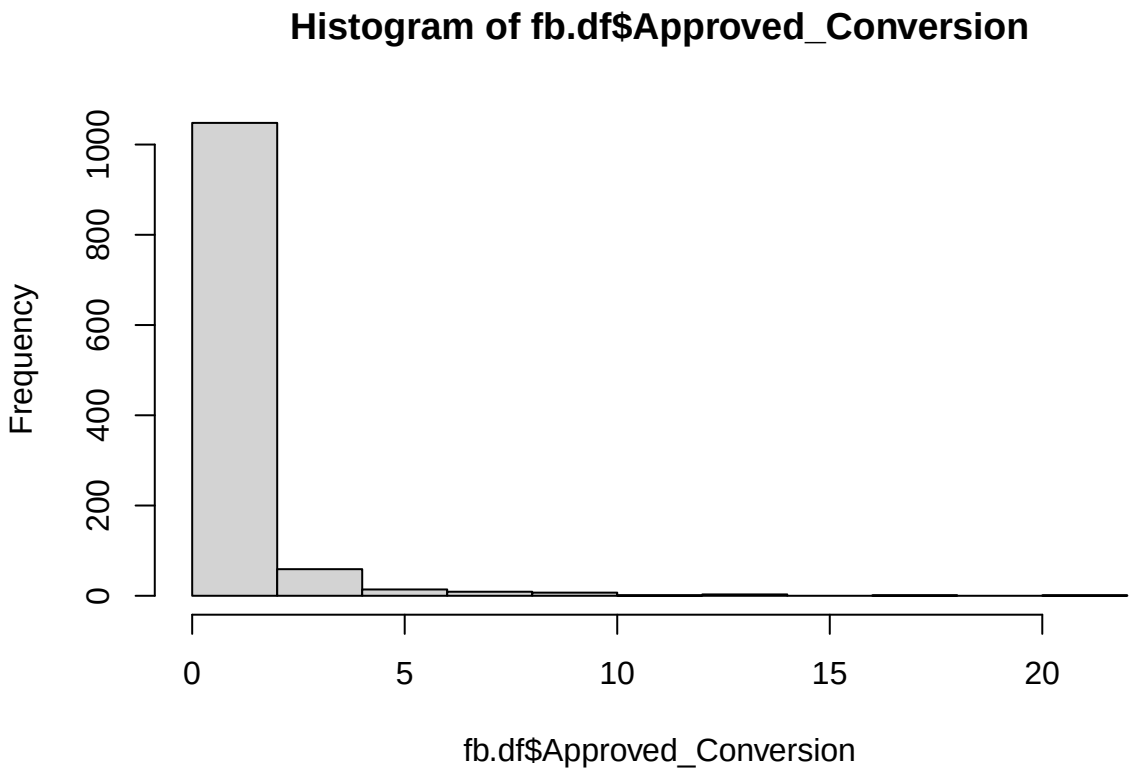
```
##
##                   F          M
##   30-34 0.17235346 0.20034996
##   35-39 0.09536308 0.12160980
##   40-44 0.09361330 0.09011374
##   45-49 0.12073491 0.10586177
```

```r
#User Interests: Categorized interests that align with the target audience.
interests_table <- table(fb.df$interest)
interests.df <- as.data.frame(interests_table)
names(interests.df) <- c("Interest", "Frequency")

library(ggplot2)
ggplot(interests.df, aes(x = Interest, y = Frequency)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "Frequency of Interests", x = "Interest", y = "Frequency") +
  theme_minimal()
```



```r
#Visualize the Approved conversion using histogram
hist(fb.df$Approved_Conversion)
```



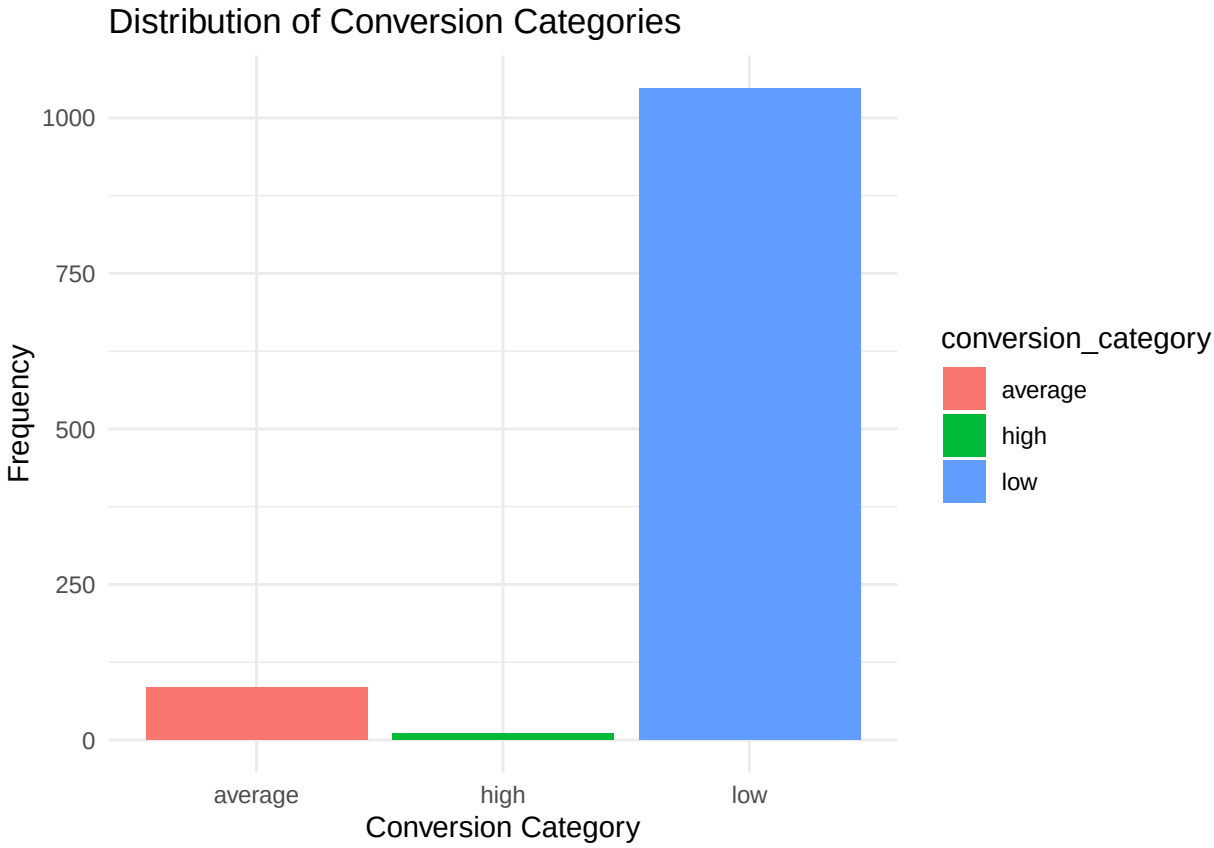

```r
#Remove the ID columns
fb.clean <- fb.df[, -c(1:3)]
summary(fb.clean)
```

```
##      age                gender             interest       Impressions
##  Length:1143        Length:1143        Min.   :  2.00   Min.   :     87
##  Class :character   Class :character   1st Qu.: 16.00   1st Qu.:   6504
##  Mode  :character   Mode  :character   Median : 25.00   Median :  51509
##                                        Mean   : 32.77   Mean   : 186732
##                                        3rd Qu.: 31.00   3rd Qu.: 221769
##                                        Max.   :114.00   Max.   :3052003
##      Clicks            Spent         Total_Conversion Approved_Conversion
##  Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   : 0.000
##  1st Qu.:  1.00   1st Qu.:  1.48   1st Qu.: 1.000   1st Qu.: 0.000
##  Median :  8.00   Median : 12.37   Median : 1.000   Median : 1.000
##  Mean   : 33.39   Mean   : 51.36   Mean   : 2.856   Mean   : 0.944
##  3rd Qu.: 37.50   3rd Qu.: 60.02   3rd Qu.: 3.000   3rd Qu.: 1.000
##  Max.   :421.00   Max.   :639.95   Max.   :60.000   Max.   :21.000
```

```r
#Create a new column to indicate if the approved conversion is low, average, or high
#if approved conversion<= 2 consider it as "low"
#if approved conversion<10 consider it as average,
#otherwise consider it as high)
fb.clean$conversion_category <- ifelse(fb.clean$Approved_Conversion<= 2, "low",
                                       ifelse(fb.clean$Approved_Conversion< 10,
                                              "average", "high"))
summary(fb.clean)
```

```
##      age                gender             interest       Impressions
##  Length:1143        Length:1143        Min.   :  2.00   Min.   :     87
##  Class :character   Class :character   1st Qu.: 16.00   1st Qu.:   6504
##  Mode  :character   Mode  :character   Median : 25.00   Median :  51509
##                                        Mean   : 32.77   Mean   : 186732
##                                        3rd Qu.: 31.00   3rd Qu.: 221769
##                                        Max.   :114.00   Max.   :3052003
##      Clicks            Spent         Total_Conversion Approved_Conversion
##  Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   : 0.000
##  1st Qu.:  1.00   1st Qu.:  1.48   1st Qu.: 1.000   1st Qu.: 0.000
##  Median :  8.00   Median : 12.37   Median : 1.000   Median : 1.000
##  Mean   : 33.39   Mean   : 51.36   Mean   : 2.856   Mean   : 0.944
##  3rd Qu.: 37.50   3rd Qu.: 60.02   3rd Qu.: 3.000   3rd Qu.: 1.000
##  Max.   :421.00   Max.   :639.95   Max.   :60.000   Max.   :21.000
##  conversion_category
##  Length:1143
##  Class :character
##  Mode  :character
##
##
##
```

```r
library(ggplot2)
ggplot(fb.clean, aes(x = conversion_category, fill = conversion_category)) +
  geom_bar() +
  labs(title = "Distribution of Conversion Categories",
       x = "Conversion Category",
       y = "Frequency") +
  theme_minimal()
```



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.