HireArt_Exercise

- Reading xlsx file by installing and using library readxl.
- Set working directory.
- Use function read_excel to extract data fromxlsx file.
- Showed top 5 rows of data using function head()

```
#install.packages("readxl")
library(readxl)

## Warning: package 'readxl' was built under R version 3.5.3

setwd('D:/wallpaper')
my_data = read_excel('HireArt - Data Analyst Exercise 10.12.17.xlsx')
head(my_data, n= 5)

## # A tibble: 5 x 3
##   `Account manager` `Client Name`              `Date of Contact`
##   <chr>             <chr>                      <dttm>
## 1 Milan Crona       Wyman, Farrell and Haag    2015-10-31 00:00:00
## 2 Aidan Pouros      Veum, McClure and Schuster 2013-12-15 00:00:00
## 3 Aidan Pouros      Veum, McClure and Schuster 2013-12-07 00:00:00
## 4 Milan Crona       Armstrong Group            2013-11-02 00:00:00
## 5 Aidan Pouros      Lueilwitz, Moore and Hahn  2013-12-25 00:00:00
```

- Checked count of missing values in a data, using sum() and is.na().
- There are 0 missing values in a data.

```
sum(is.na(my_data))

## [1] 0
```

- Extracted month from Date of Contact using months().
- Create dataframe(df_1) and included all the values of my_data and month.
- Displayed top 5 rows using function head().

```
Month_of_Contact = months(my_data$`Date of Contact`, abbreviate = TRUE)

df_1 = data.frame(my_data$`Account manager`, my_data$`Client
Name`,my_data$`Date of Contact`, Month_of_Contact)

head(df_1, n=5)

##   my_data..Account.manager.      my_data..Client.Name.
## 1               Milan Crona    Wyman, Farrell and Haag
## 2              Aidan Pouros Veum, McClure and Schuster
## 3              Aidan Pouros Veum, McClure and Schuster
## 4               Milan Crona            Armstrong Group
## 5              Aidan Pouros  Lueilwitz, Moore and Hahn
##   my_data..Date.of.Contact. Month_of_Contact
```

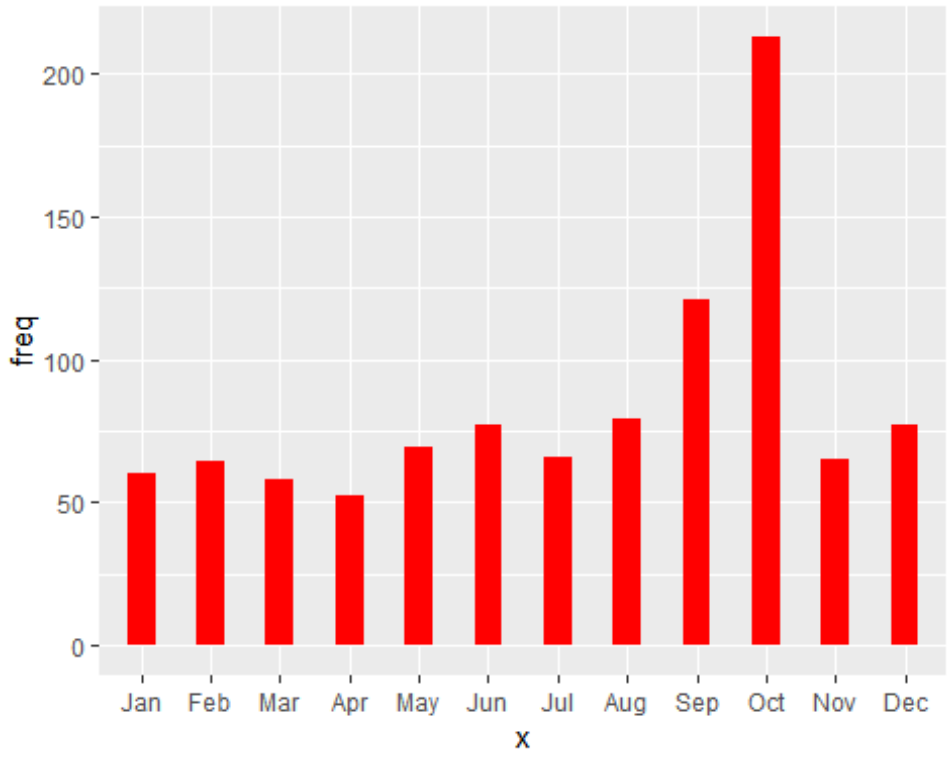
```
## 1                2015-10-31              Oct
## 2                2013-12-15              Dec
## 3                2013-12-07              Dec
## 4                2013-11-02              Nov
## 5                2013-12-25              Dec
```

-created frequency table for months.

```
library(plyr)

## Warning: package 'plyr' was built under R version 3.5.3

Count_month = count(df_1$Month_of_Contact)
Count_month

##      x freq
## 1  Apr   52
## 2  Aug   79
## 3  Dec   77
## 4  Feb   64
## 5  Jan   60
## 6  Jul   66
## 7  Jun   77
## 8  Mar   58
## 9  May   69
## 10 Nov   65
## 11 Oct  213
## 12 Sep  121
```

- Used ggplot2 libraries.
- Utilized ggplot() visualization function to display bar plot of variable : Month and frequency of Month.
- Here we can conclude that October is a month when clients are contacted in a greatest percentage by team

```
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

ggplot(Count_month, aes(x = x, y = freq), size = 1)+ geom_bar(stat =
"Identity", width = .4, fill = "RED") + scale_x_discrete(limits = month.abb)
```


200
150
100
50
0
freq
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
x