Assignment_1

Venkata Naga Siddartha Gutha

2022-09-12

```
#Using rent prediction data set
# Source:https://www.kaggle.com/datasets/iamsouravbanerjee/house-rent-prediction-dataset

#Importing data set into r

rent_dataset<-read.csv("C:/Users/sidda/Desktop/KSU_Fall/FML/House_Rent_Dataset.csv")

# Descriptive statistics
summary(rent_dataset)
```

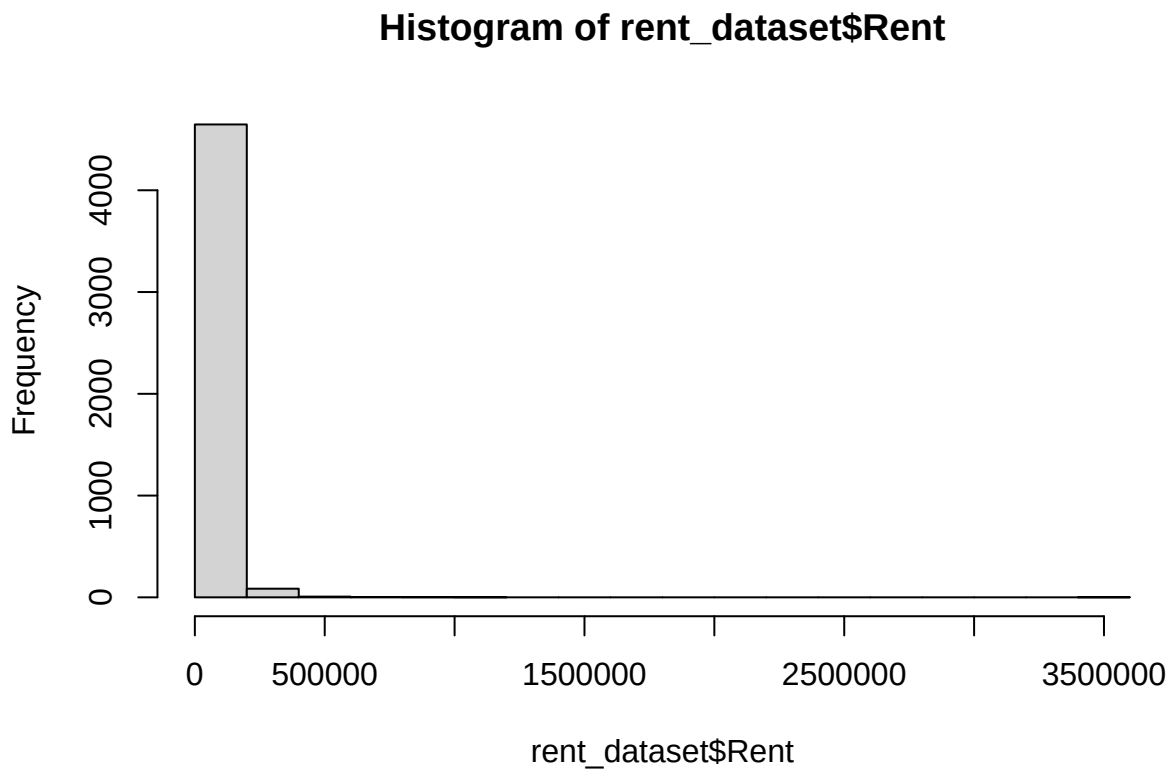
```
##   Posted.On              BHK             Rent              Size
##  Length:4746        Min.   :1.000   Min.   :   1200   Min.   :  10.0
##  Class :character   1st Qu.:2.000   1st Qu.:  10000   1st Qu.: 550.0
##  Mode  :character   Median :2.000   Median :  16000   Median : 850.0
##                     Mean   :2.084   Mean   :  34993   Mean   : 967.5
##                     3rd Qu.:3.000   3rd Qu.:  33000   3rd Qu.:1200.0
##                     Max.   :6.000   Max.   :3500000   Max.   :8000.0
##     Floor            Area.Type         Area.Locality          City
##  Length:4746        Length:4746        Length:4746        Length:4746
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##  Furnishing.Status  Tenant.Preferred      Bathroom      Point.of.Contact
##  Length:4746        Length:4746        Min.   : 1.000   Length:4746
##  Class :character   Class :character   1st Qu.: 1.000   Class :character
##  Mode  :character   Mode  :character   Median : 2.000   Mode  :character
##                                        Mean   : 1.966
##                                        3rd Qu.: 2.000
##                                        Max.   :10.000
```

```
mean_rent=mean(rent_dataset$Rent)
median_rent=median(rent_dataset$Rent)
mode_rent=mode(rent_dataset$Rent)
sd_rent=sd(rent_dataset$Rent)


#Transforming Rent and size variables of the dataset

transformed_rent=(rent_dataset$Rent-median_rent)/sd_rent + mean_rent
```

```
new_size = rent_dataset$Size/2


#plotting rent

library(esquisse)

esquisse::esquisser(rent_dataset)
```

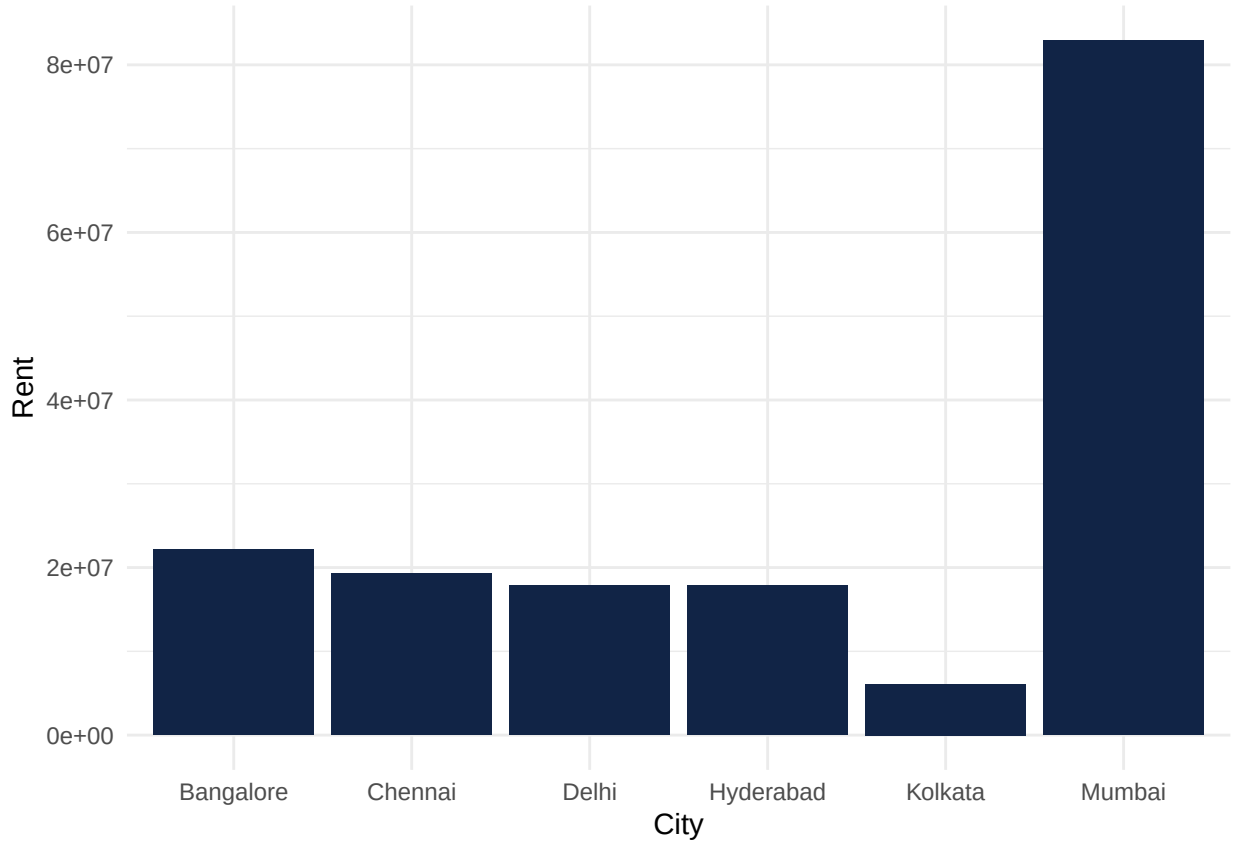
```
## Warning: Esquisse may not work properly, try updating RStudio.

## Loading required package: shiny

##
## Listening on http://127.0.0.1:6791
```

```
#Histogram of rent
hist(rent_dataset$Rent)
```

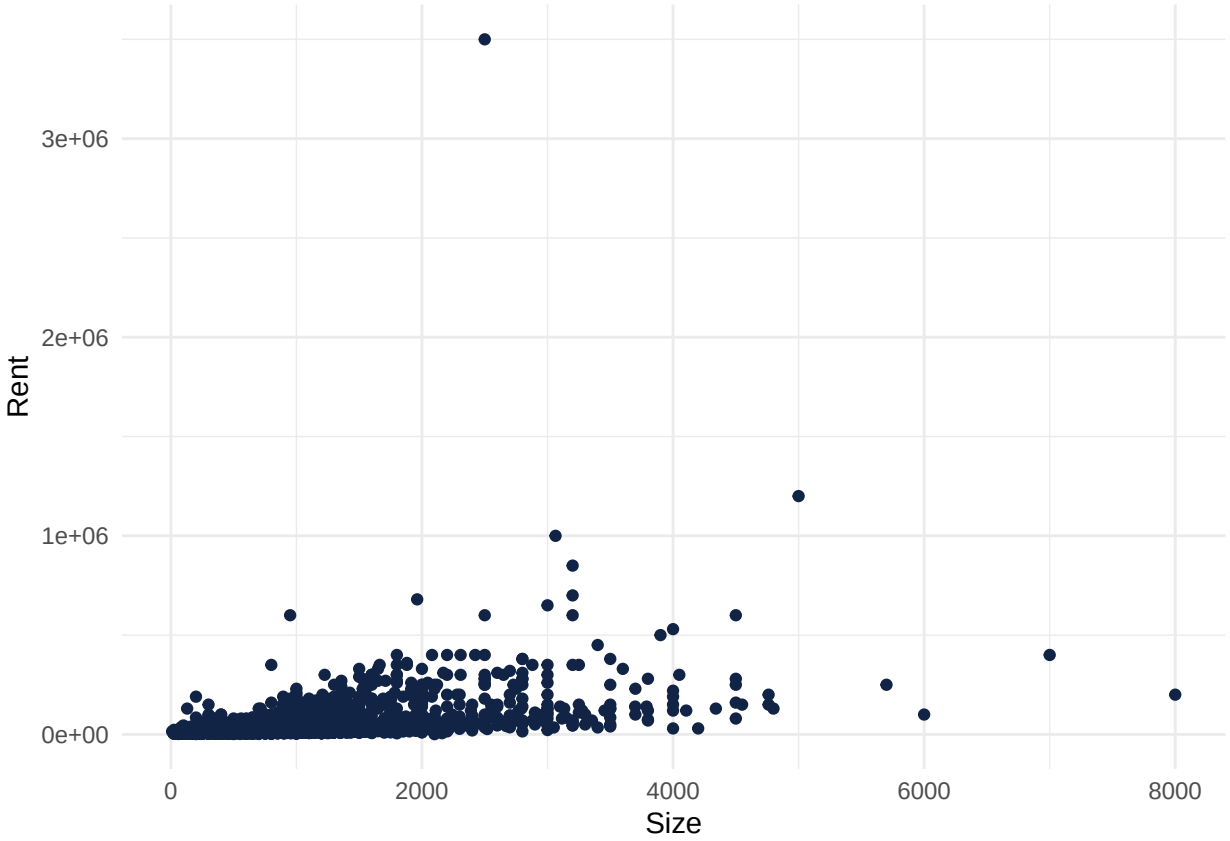



```
# bar graph of Rent and City
library(ggplot2)

ggplot(rent_dataset) +
```

```
aes(x = City, weight = Rent) +
geom_bar(fill = "#112446") +
labs(x = "City", y = "Rent") +
theme_minimal()
```




```
#scatter plot of rent and size of the house

library(ggplot2)

ggplot(rent_dataset) +
  aes(x = Size, y = Rent) +
  geom_point(shape = "circle", size = 1.5, colour = "#112446") +
  labs(x = "Size", y = "Rent") +
  theme_minimal()
```


Rent
3e+06
2e+06
1e+06
0e+00
0
2000
4000
6000
8000
Size