sboyapa1

Sai Prasad

2024-01-26

Here uploding my dataset House_rent_Dataset

```
rental_data <- read.csv("C:\\Users\\desineni\\Downloads\\House_Rent_Dataset (1).csv")
```

```
summary(rental_data)
```

```
##   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
```

Here printing descriptive statistics for a selection of quantitative from dataset which was “Size”

```
summary(rental_data$Size)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    10.0   550.0   850.0   967.5  1200.0  8000.0
```

Here printing descriptive statistics for a selection of categorical variable is Area.Locatlity

```
summary(rental_data$Area.Locality)
```

```
##    Length     Class      Mode
##      4746 character character
```

Here I am creating variable “coupon” And assigning it to square root of a rent. Finally applying to new varaible “new_rent” where it means rent after applying the coupon

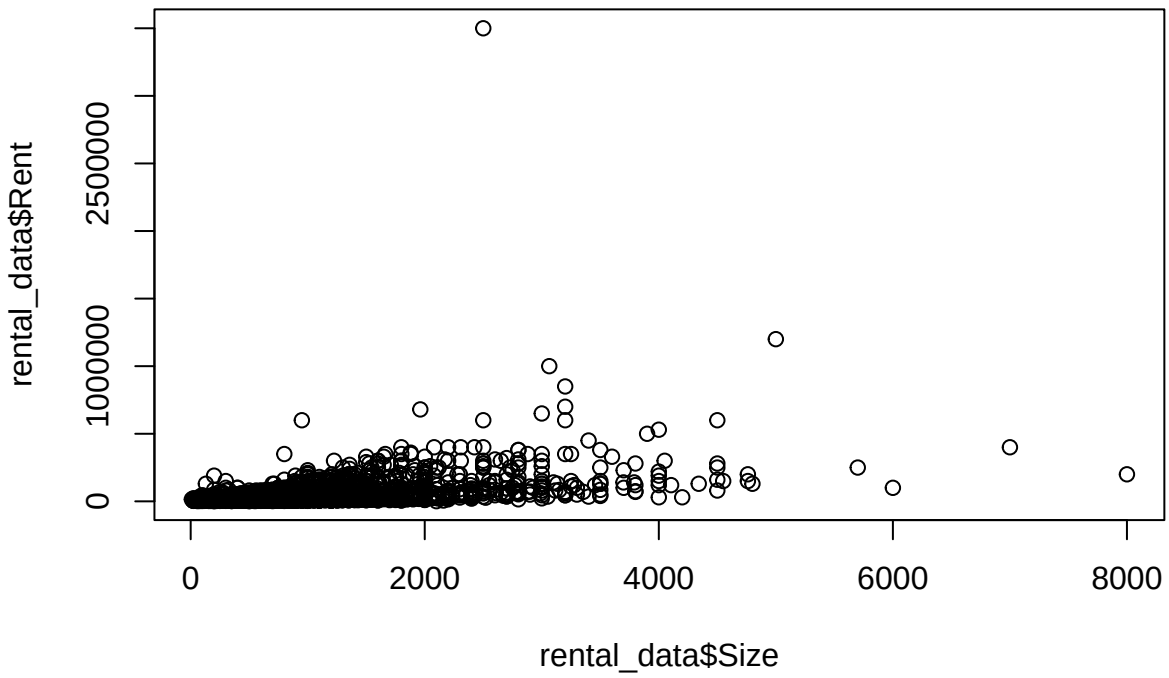
```
coupon <- sqrt(rental_data$Rent)
View(rental_data)
View(coupon)
```

```
# Transforming Rent variables of the dataset
new_rent <- (rental_data$Rent-coupon)
head(new_rent)
```

```
## [1]  9900.000 19858.579 16869.616  9900.000  7413.397  6916.334
```

Here I am plotting a scatter plot Size vesus Rent. Because Rent depends on the Size. Scale x-axis is size of the House y-axis is rent of the House

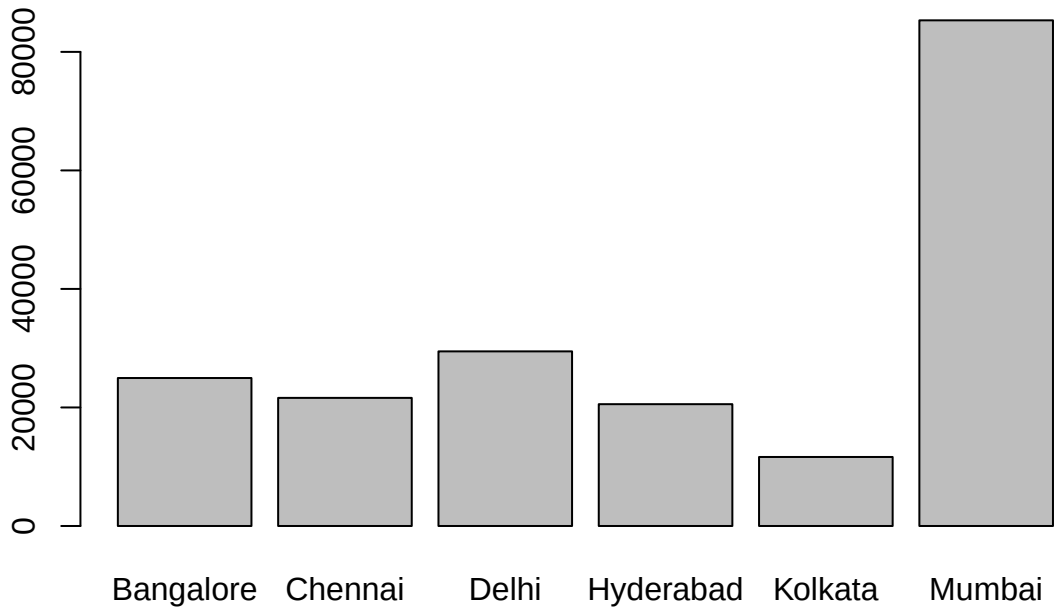
```
plot(rental_data$Size,rental_data$Rent)
```



```r
library(ggplot2)
```

```r
rental_mean <- aggregate(Rent ~ City, data = rental_data, FUN = mean)

# Creating Barplot Rent Versus City.
barplot(rental_mean$Rent, names.arg = rental_mean$City)
```



the people living Mumbai have highest Rent whereas the people living Kolkata have least Rent.

Source of the data used for this task: https://www.kaggle.com/code/prashantverma13/house-rent-prediction-in-depth-analysis-models/input?select=House_Rent_Dataset.csv