Airbnb Analysis

```
#New data set as Airbnb
library(readr)
library(csv)
Airbnb <- read_csv('https://raw.githubusercontent.com/Shubham619/Airbnb_Analysis/master/Airbnb.csv')
```

```
## Parsed with column specification:
## cols(
##   room_id = col_integer(),
##   host_id = col_integer(),
##   room_type = col_character(),
##   borough = col_character(),
##   neighborhood = col_character(),
##   reviews = col_integer(),
##   overall_satisfaction = col_double(),
##   accommodates = col_double(),
##   bedrooms = col_double(),
##   price = col_double(),
##   minstay = col_double(),
##   latitude = col_double(),
##   longitude = col_double(),
##   last_modified = col_datetime(format = "")
## )
```

```
#summary
summary(Airbnb)
```

```
##     room_id           host_id          room_type
##  Min.   :    105   Min.   :     500   Length:30001
##  1st Qu.:1763333   1st Qu.: 2538621   Class :character
##  Median :4204485   Median : 7858210   Mode  :character
##  Mean   :3960630   Mean   :11985356
##  3rd Qu.:6145124   3rd Qu.:19971797
##  Max.   :7823353   Max.   :41112492
##
##    borough          neighborhood          reviews
##  Length:30001       Length:30001       Min.   :  0.00
##  Class :character   Class :character   1st Qu.:  0.00
##  Mode  :character   Mode  :character   Median :  3.00
##                                        Mean   : 11.91
##                                        3rd Qu.: 13.00
##                                        Max.   :255.00
##
##  overall_satisfaction  accommodates       bedrooms          price
##  Min.   :1.000        Min.   : 1.000   Min.   : 0.000   Min.   :    10.0
##  1st Qu.:4.500        1st Qu.: 2.000   1st Qu.: 1.000   1st Qu.:    85.0
##  Median :4.500        Median : 2.000   Median : 1.000   Median :   129.0
##  Mean   :4.623        Mean   : 2.491   Mean   : 1.144   Mean   :   241.6
##  3rd Qu.:5.000        3rd Qu.: 3.000   3rd Qu.: 1.000   3rd Qu.:   199.0
##  Max.   :5.000        Max.   :16.000   Max.   :10.000   Max.   :140000.0
##  NA's   :8374         NA's   :2419     NA's   :133
##     minstay           latitude       longitude
```

```
##  Min.   :  1.000   Min.   :40.51   Min.   :-74.24
##  1st Qu.:  1.000   1st Qu.:40.70   1st Qu.:-73.99
##  Median :  2.000   Median :40.73   Median :-73.96
##  Mean   :  3.395   Mean   :40.73   Mean   :-73.96
##  3rd Qu.:  3.000   3rd Qu.:40.76   3rd Qu.:-73.94
##  Max.   :999.000   Max.   :40.91   Max.   :-73.73
##  NA's   :1144
##  last_modified
##  Min.   :2015-08-10 21:44:27
##  1st Qu.:2015-08-10 21:44:27
##  Median :2015-08-10 21:44:27
##  Mean   :2015-08-11 01:32:57
##  3rd Qu.:2015-08-11 01:16:09
##  Max.   :2015-08-12 01:20:18
##
```

```
library(dplyr)
glimpse(Airbnb)
```

```
## Observations: 30,001
## Variables: 14
## $ room_id              <int> 269404, 2451438, 2905768, 1624665, 354204...
## $ host_id              <int> 1411399, 10193030, 12095101, 8638841, 178...
## $ room_type            <chr> "Private room", "Private room", "Entire h...
## $ borough              <chr> "Queens", "Manhattan", "Manhattan", "Broo...
## $ neighborhood         <chr> "Astoria", "Harlem", "Hell's Kitchen", "W...
## $ reviews              <int> 18, 2, 2, 24, 14, 99, 7, 7, 33, 5, 0, 50,...
## $ overall_satisfaction <dbl> 4.5, 5.0, 4.5, 5.0, 4.5, 5.0, 4.5, 4.5, 4...
## $ accommodates         <dbl> 2, 2, 2, 1, 4, 2, 3, 2, 3, 2, 2, 3, 2, 2,...
## $ bedrooms             <dbl> 1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 1,...
## $ price                <dbl> 2492, 85, 140, 80, 299, 96, 250, 100, 70,...
## $ minstay              <dbl> 30, 3, 20, 3, 3, 1, 4, 2, 3, 2, 1, 3, 1, ...
## $ latitude             <dbl> 40.76684, 40.82349, 40.76363, 40.70553, 4...
## $ longitude            <dbl> -73.91211, -73.94193, -73.99309, -73.9272...
## $ last_modified        <dttm> 2015-08-10 21:44:27, 2015-08-10 21:44:27...
```

```
#first 6 and last 6 observations
head(Airbnb)
```

```
## # A tibble: 6 x 14
##   room_id  host_id       room_type   borough     neighborhood reviews
##     <int>    <int>           <chr>     <chr>            <chr>   <int>
## 1  269404  1411399    Private room    Queens          Astoria      18
## 2 2451438 10193030    Private room Manhattan           Harlem       2
## 3 2905768 12095101 Entire home/apt Manhattan   Hell's Kitchen       2
## 4 1624665  8638841    Private room  Brooklyn     Williamsburg      24
## 5 3542044 17830235 Entire home/apt  Brooklyn     Williamsburg      14
## 6  589089   335272    Private room  Brooklyn Prospect Heights      99
## # ... with 8 more variables: overall_satisfaction <dbl>,
## #   accommodates <dbl>, bedrooms <dbl>, price <dbl>, minstay <dbl>,
## #   latitude <dbl>, longitude <dbl>, last_modified <dttm>
```

```
tail(Airbnb)
```

```
## # A tibble: 6 x 14
##   room_id  host_id       room_type   borough  neighborhood reviews
```

```
##     <int>    <int>           <chr>     <chr>         <chr>   <int>
## 1   54486   255748    Private room  Brooklyn Crown Heights      79
## 2 4645969  1472210 Entire home/apt Manhattan  East Village      19
## 3 7583719 39769353    Private room  Brooklyn Crown Heights       0
## 4 2774189 14187996 Entire home/apt Manhattan      Kips Bay      62
## 5 4662840  4114392    Private room Manhattan        Harlem       1
## 6 4309180  2653079 Entire home/apt  Brooklyn    Park Slope       2
## # ... with 8 more variables: overall_satisfaction <dbl>,
## #   accommodates <dbl>, bedrooms <dbl>, price <dbl>, minstay <dbl>,
## #   latitude <dbl>, longitude <dbl>, last_modified <dttm>
```

```
#columns
names(Airbnb)
```

```
##  [1] "room_id"              "host_id"              "room_type"
##  [4] "borough"              "neighborhood"         "reviews"
##  [7] "overall_satisfaction" "accommodates"         "bedrooms"
## [10] "price"                "minstay"              "latitude"
## [13] "longitude"            "last_modified"
```

```
#checking if any NA values
any(is.na(Airbnb))
```

```
## [1] TRUE
```

```
#Viewing NA values columns
Airbnb[!complete.cases(Airbnb),]
```

```
## # A tibble: 11,163 x 14
##    room_id  host_id       room_type   borough       neighborhood reviews
##      <int>    <int>           <chr>     <chr>              <chr>   <int>
##  1 2268888 11424459 Entire home/apt Manhattan Washington Heights       0
##  2  628100  2162167 Entire home/apt  Brooklyn       Williamsburg       8
##  3 3603287   146825    Private room Manhattan             Nolita       0
##  4 5188840 26377263    Private room    Queens       Far Rockaway       0
##  5 2228296 11371978 Entire home/apt Manhattan            Midtown       0
##  6 5335792  7877557    Private room Manhattan       East Village       0
##  7 2070042  8119315 Entire home/apt    Queens          Ridgewood       0
##  8 4947430 25495114 Entire home/apt Manhattan       East Village       0
##  9 3278516 16572416 Entire home/apt Manhattan               SoHo       0
## 10 3554019 17893521    Private room    Queens            Maspeth       0
## # ... with 11,153 more rows, and 8 more variables:
## #   overall_satisfaction <dbl>, accommodates <dbl>, bedrooms <dbl>,
## #   price <dbl>, minstay <dbl>, latitude <dbl>, longitude <dbl>,
## #   last_modified <dttm>
```

```
#Omiting all NA values and new data set without NA values
Airbnb_cleaned <- na.omit(Airbnb)

#rechecking NA
any(is.na(Airbnb_cleaned))
```

```
## [1] FALSE
```

```
names(Airbnb_cleaned)
```

```
##  [1] "room_id"              "host_id"              "room_type"
##  [4] "borough"              "neighborhood"         "reviews"
```

```
##  [7] "overall_satisfaction" "accommodates"         "bedrooms"
## [10] "price"                "minstay"              "latitude"
## [13] "longitude"            "last_modified"
```

```
#rows
count(Airbnb_cleaned)
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1 18838
```

```
#summary of new dataset
summary(Airbnb_cleaned)
```

```
##     room_id           host_id          room_type
##  Min.   :   2515   Min.   :    1039   Length:18838
##  1st Qu.:1225130   1st Qu.: 2127987   Class :character
##  Median :3461000   Median : 6394483   Mode  :character
##  Mean   :3374035   Mean   :10278135
##  3rd Qu.:5127433   3rd Qu.:16857761
##  Max.   :7808182   Max.   :40739836
##    borough          neighborhood          reviews
##  Length:18838       Length:18838       Min.   :  1.00
##  Class :character   Class :character   1st Qu.:  2.00
##  Mode  :character   Mode  :character   Median :  7.00
##                                        Mean   : 16.15
##                                        3rd Qu.: 19.00
##                                        Max.   :253.00
##  overall_satisfaction  accommodates       bedrooms         price
##  Min.   :1.00         Min.   : 1.000   Min.   :0.000   Min.   :   10.0
##  1st Qu.:4.50         1st Qu.: 2.000   1st Qu.:1.000   1st Qu.:   80.0
##  Median :4.50         Median : 2.000   Median :1.000   Median :  125.0
##  Mean   :4.63         Mean   : 2.538   Mean   :1.021   Mean   :  194.5
##  3rd Qu.:5.00         3rd Qu.: 3.000   3rd Qu.:1.000   3rd Qu.:  180.0
##  Max.   :5.00         Max.   :16.000   Max.   :6.000   Max.   :15000.0
##     minstay          latitude       longitude
##  Min.   :  1.00   Min.   :40.54   Min.   :-74.17
##  1st Qu.:  1.00   1st Qu.:40.70   1st Qu.:-73.99
##  Median :  2.00   Median :40.73   Median :-73.96
##  Mean   :  3.27   Mean   :40.73   Mean   :-73.96
##  3rd Qu.:  3.00   3rd Qu.:40.76   3rd Qu.:-73.95
##  Max.   :999.00   Max.   :40.91   Max.   :-73.73
##  last_modified
##  Min.   :2015-08-10 21:44:27
##  1st Qu.:2015-08-10 21:44:27
##  Median :2015-08-10 21:44:27
##  Mean   :2015-08-11 00:23:20
##  3rd Qu.:2015-08-10 21:44:27
##  Max.   :2015-08-12 01:20:18
```

```
glimpse(Airbnb_cleaned)
```

```
## Observations: 18,838
## Variables: 14
## $ room_id              <int> 269404, 2451438, 2905768, 1624665, 354204...
```

```
## $ host_id              <int> 1411399, 10193030, 12095101, 8638841, 178...
## $ room_type            <chr> "Private room", "Private room", "Entire h...
## $ borough              <chr> "Queens", "Manhattan", "Manhattan", "Broo...
## $ neighborhood         <chr> "Astoria", "Harlem", "Hell's Kitchen", "W...
## $ reviews              <int> 18, 2, 2, 24, 14, 99, 7, 7, 33, 5, 50, 19...
## $ overall_satisfaction <dbl> 4.5, 5.0, 4.5, 5.0, 4.5, 5.0, 4.5, 4.5, 4...
## $ accommodates         <dbl> 2, 2, 2, 1, 4, 2, 3, 2, 3, 2, 3, 2, 2, 2,...
## $ bedrooms             <dbl> 1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 1, 1, 1,...
## $ price                <dbl> 2492, 85, 140, 80, 299, 96, 250, 100, 70,...
## $ minstay              <dbl> 30, 3, 20, 3, 3, 1, 4, 2, 3, 2, 3, 1, 4, ...
## $ latitude             <dbl> 40.76684, 40.82349, 40.76363, 40.70553, 4...
## $ longitude            <dbl> -73.91211, -73.94193, -73.99309, -73.9272...
## $ last_modified        <dttm> 2015-08-10 21:44:27, 2015-08-10 21:44:27...
```

```
#selecting and grouping
Airbnb_cleaned %>%
  group_by(room_type,minstay)
```

```
## # A tibble: 18,838 x 14
## # Groups:   room_type, minstay [94]
##    room_id  host_id       room_type   borough     neighborhood reviews
##      <int>    <int>           <chr>     <chr>            <chr>   <int>
##  1  269404  1411399    Private room    Queens          Astoria      18
##  2 2451438 10193030    Private room Manhattan           Harlem       2
##  3 2905768 12095101 Entire home/apt Manhattan   Hell's Kitchen       2
##  4 1624665  8638841    Private room  Brooklyn     Williamsburg      24
##  5 3542044 17830235 Entire home/apt  Brooklyn     Williamsburg      14
##  6  589089   335272    Private room  Brooklyn Prospect Heights      99
##  7  264017  1385139 Entire home/apt Manhattan           Nolita       7
##  8 3190765  4844197    Private room  Brooklyn     Williamsburg       7
##  9  215172     5926    Private room Manhattan           Harlem      33
## 10 2091991 10362675 Entire home/apt Manhattan     East Village       5
## # ... with 18,828 more rows, and 8 more variables:
## #   overall_satisfaction <dbl>, accommodates <dbl>, bedrooms <dbl>,
## #   price <dbl>, minstay <dbl>, latitude <dbl>, longitude <dbl>,
## #   last_modified <dttm>
```

```
#new column actual_rate = price divide with minstay
actual_rate<-(Airbnb_cleaned$price/Airbnb_cleaned$minstay)
#rounding values
actual_rate <- round(actual_rate)
#column binding to dataset
Airbnb_cleaned <-cbind(actual_rate,Airbnb_cleaned)
```

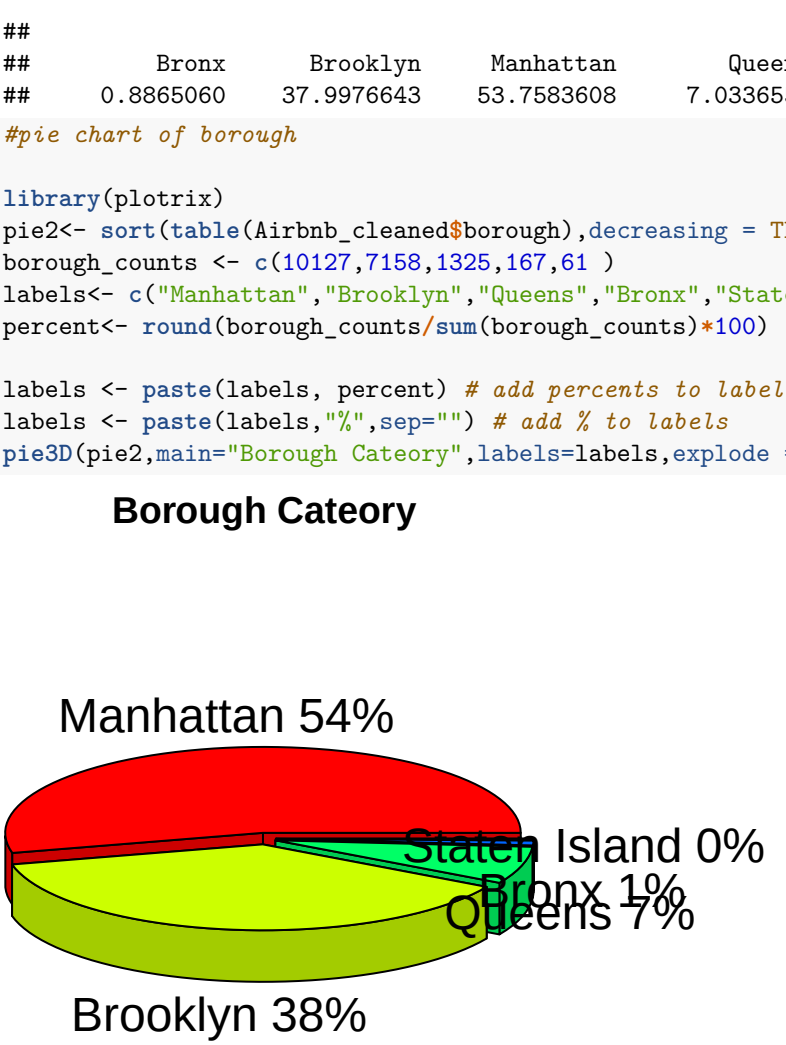
```
#mean of actual rate
actual_rate_mean <-mean(Airbnb_cleaned$actual_rate)

#dividing rooms rates as less than mean and more than mean
Rooms_withless_rates <- Airbnb_cleaned[Airbnb_cleaned$actual_rate <= actual_rate_mean,]
View(Rooms_withless_rates)

Rooms_withmore_rates <- Airbnb_cleaned[Airbnb_cleaned$actual_rate > actual_rate_mean,]
View(Rooms_withmore_rates)


cheap_room_satisfaction <-mean(Rooms_withless_rates$overall_satisfaction)
```

```
luxury_room_satisfaction <- mean(Rooms_withmore_rates$overall_satisfaction)
```

```
Borough<- table(Airbnb_cleaned$borough)
prop.table(Borough)*100
```

```
##
##         Bronx      Brooklyn     Manhattan        Queens Staten Island
##     0.8865060    37.9976643    53.7583608     7.0336554     0.3238136
```

```
#pie chart of borough

library(plotrix)
pie2<- sort(table(Airbnb_cleaned$borough),decreasing = TRUE)
borough_counts <- c(10127,7158,1325,167,61 )
labels<- c("Manhattan","Brooklyn","Queens","Bronx","Staten Island")
percent<- round(borough_counts/sum(borough_counts)*100)

labels <- paste(labels, percent) # add percents to labels
labels <- paste(labels,"%",sep="") # add % to labels
pie3D(pie2,main="Borough Cateory",labels=labels,explode = 0.045)
```

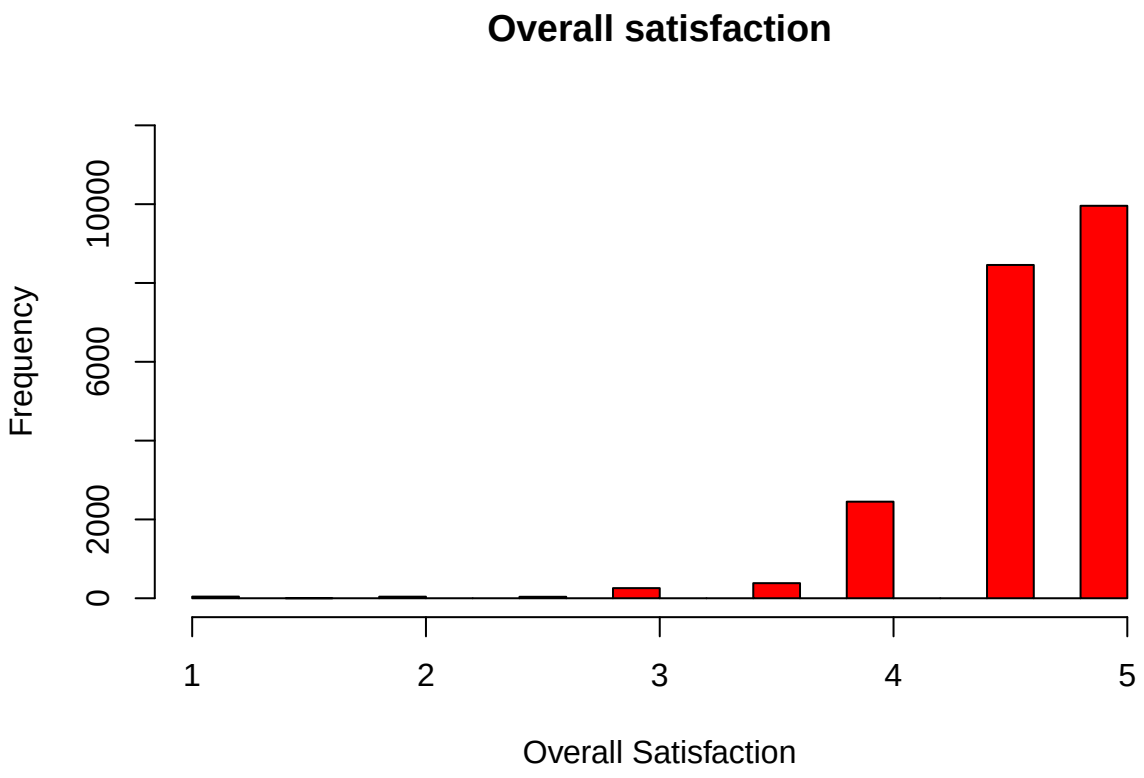



```
Neighborhood <- table(Airbnb_cleaned$neighborhood)
```

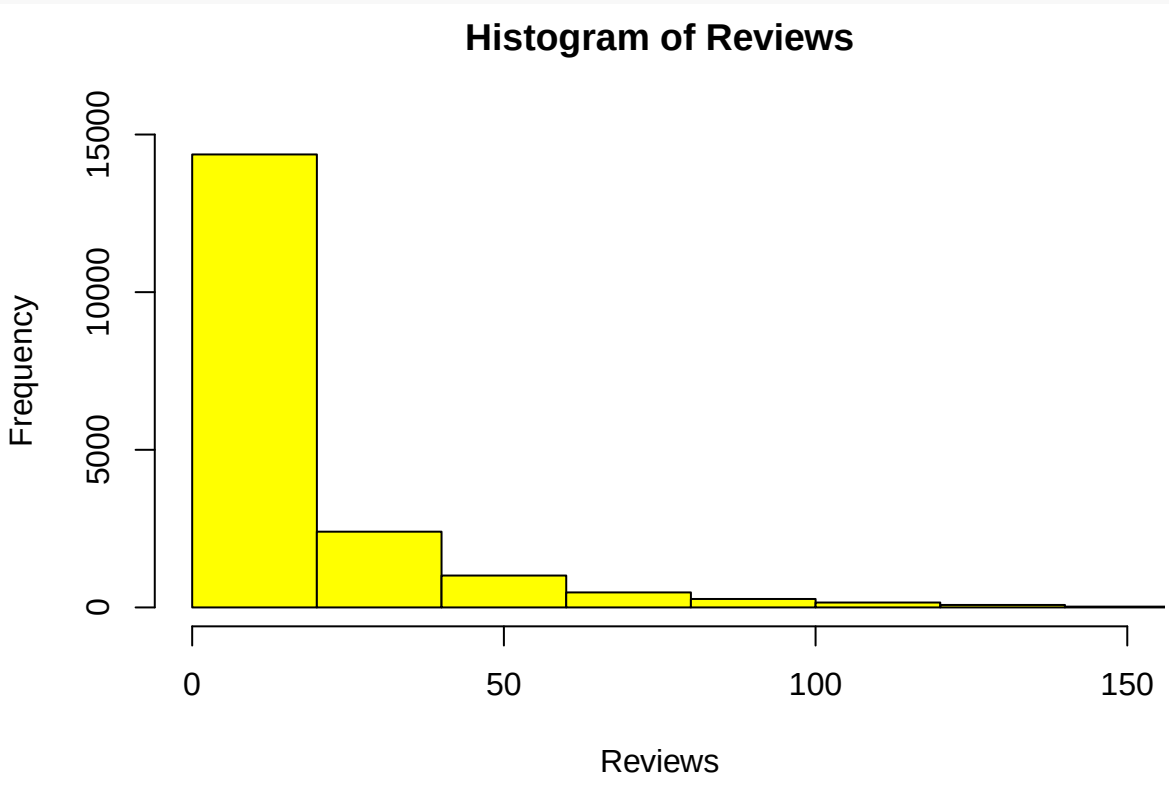
```
#table for room_type
table(Airbnb_cleaned$room_type)
```

```
##
## Entire home/apt    Private room     Shared room
##           10416            7933             489
```

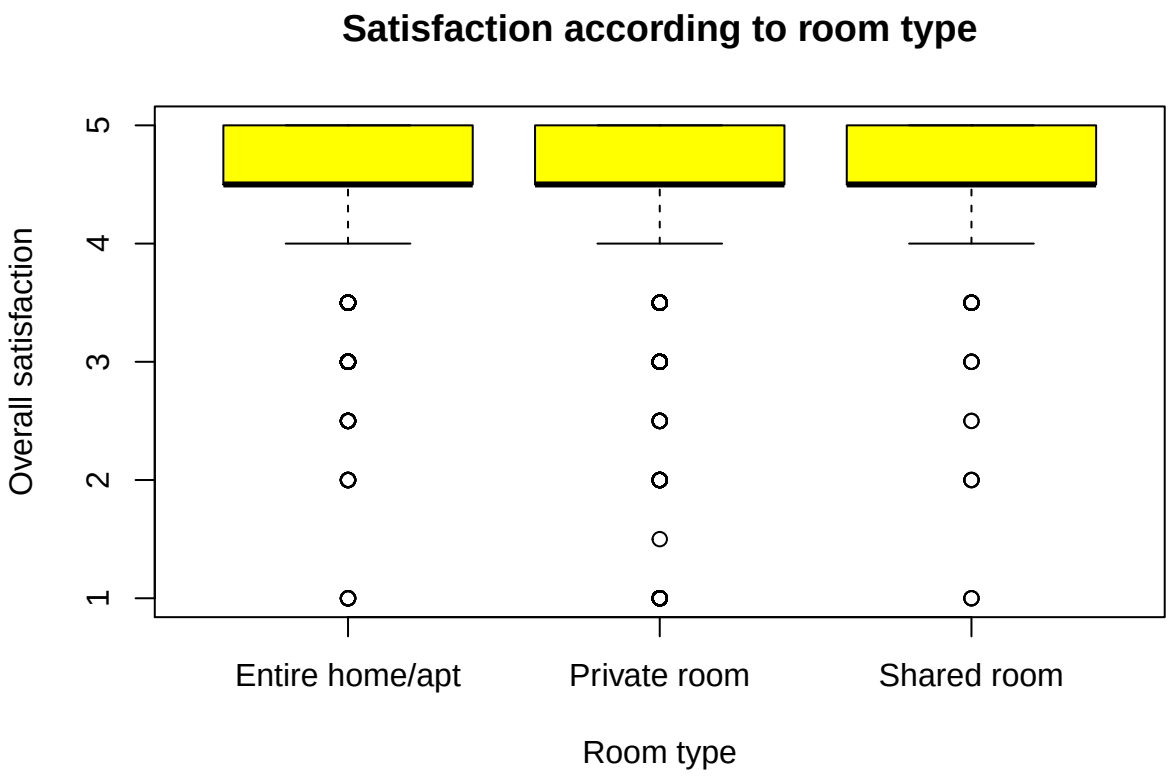
```
#room satisfaction histogram
hist(Airbnb$overall_satisfaction,col='red',ylim=c(0,12000),xlab="Overall Satisfaction",ylab="Frequency"
```



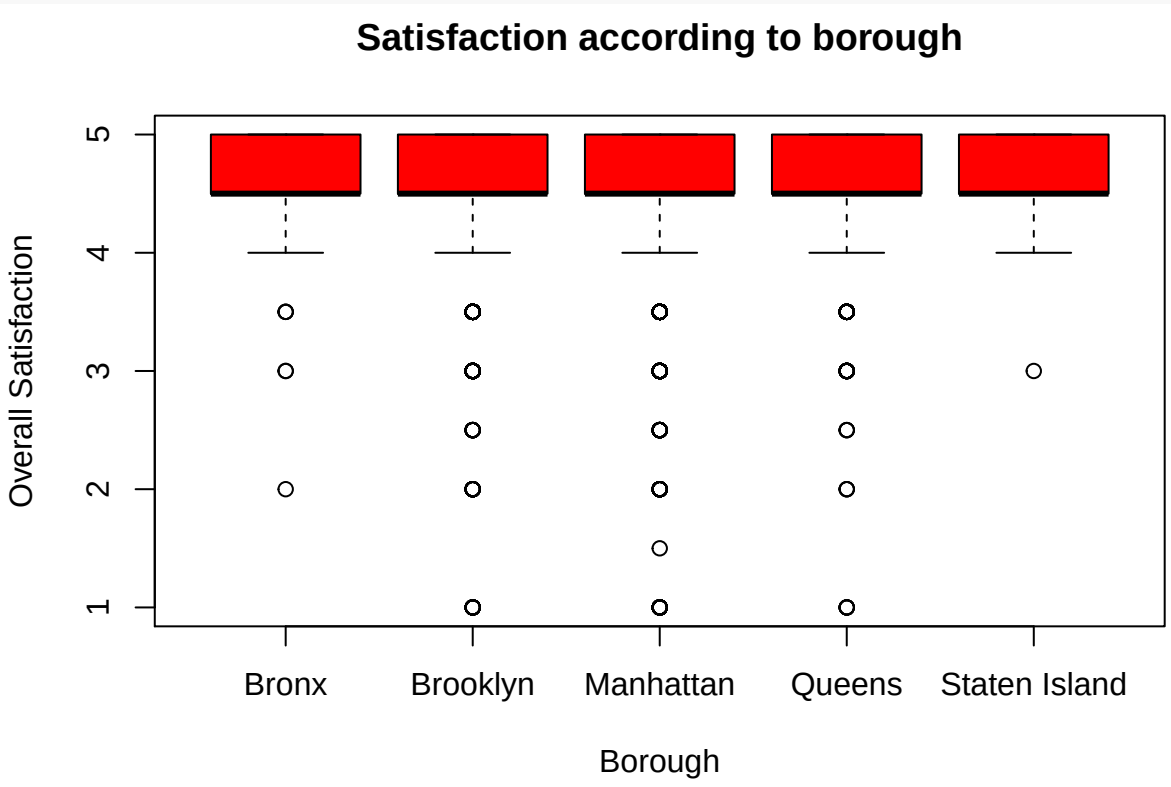

```
#room reviews histogram
hist(Airbnb_cleaned$reviews,col="yellow",xlim = c(0,150),ylim=c(0,15000),breaks = 15,ylab="Frequency",x
```




```
#room satisfaction by room_type
boxplot(overall_satisfaction ~ room_type,data= Airbnb_cleaned,col="yellow",xlab="Room type",ylab="Overal
```



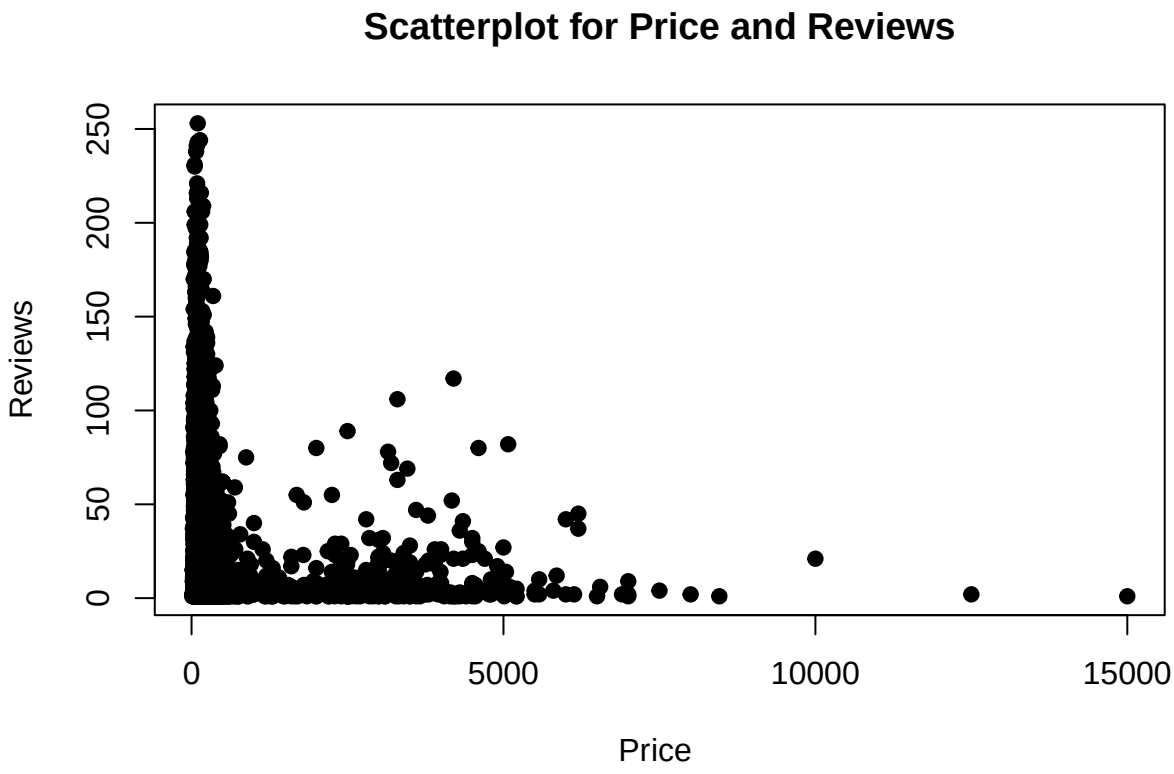

```
#room satisfaction by borough
boxplot(overall_satisfaction ~ borough,data=Airbnb_cleaned,col="red",xlab="Borough",ylab="Overall Satisf
```



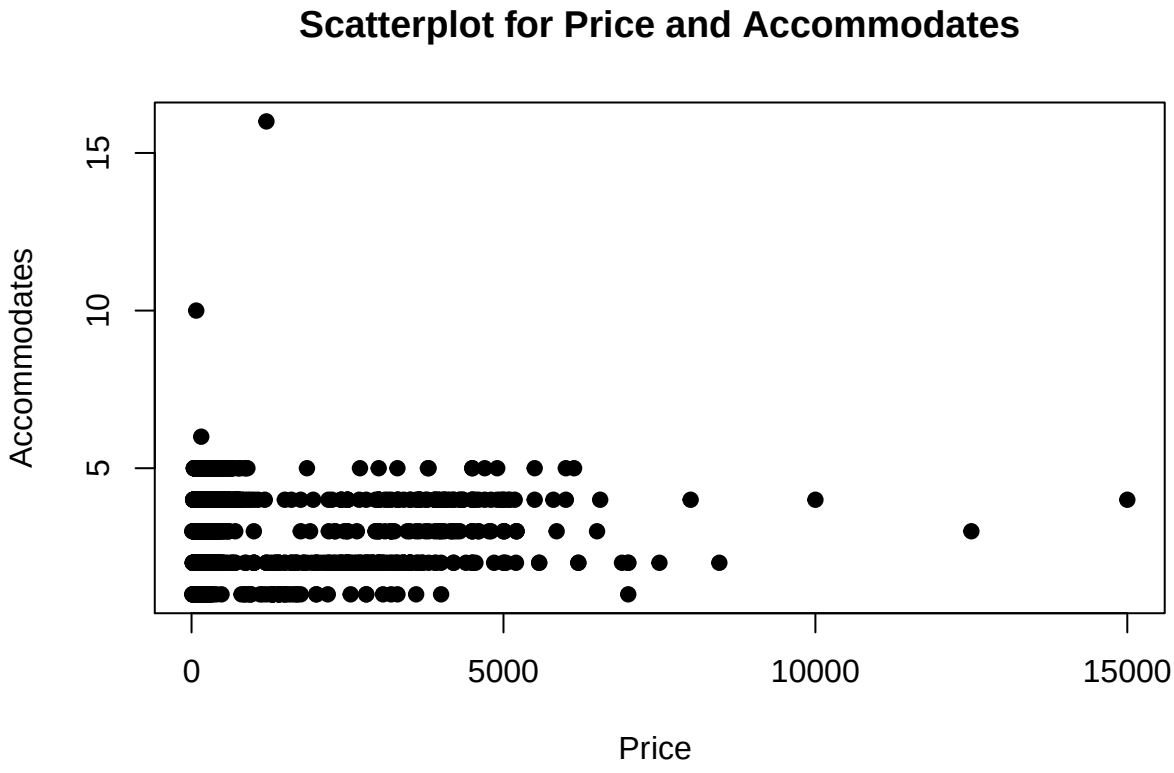

```
#plot price and bedrooms scatterplot
plot(Airbnb_cleaned$price, Airbnb_cleaned$bedrooms, main="Scatterplot for Price and Bedrooms",
```

```
     xlab="Price ", ylab="Bedrooms", pch=19)
```

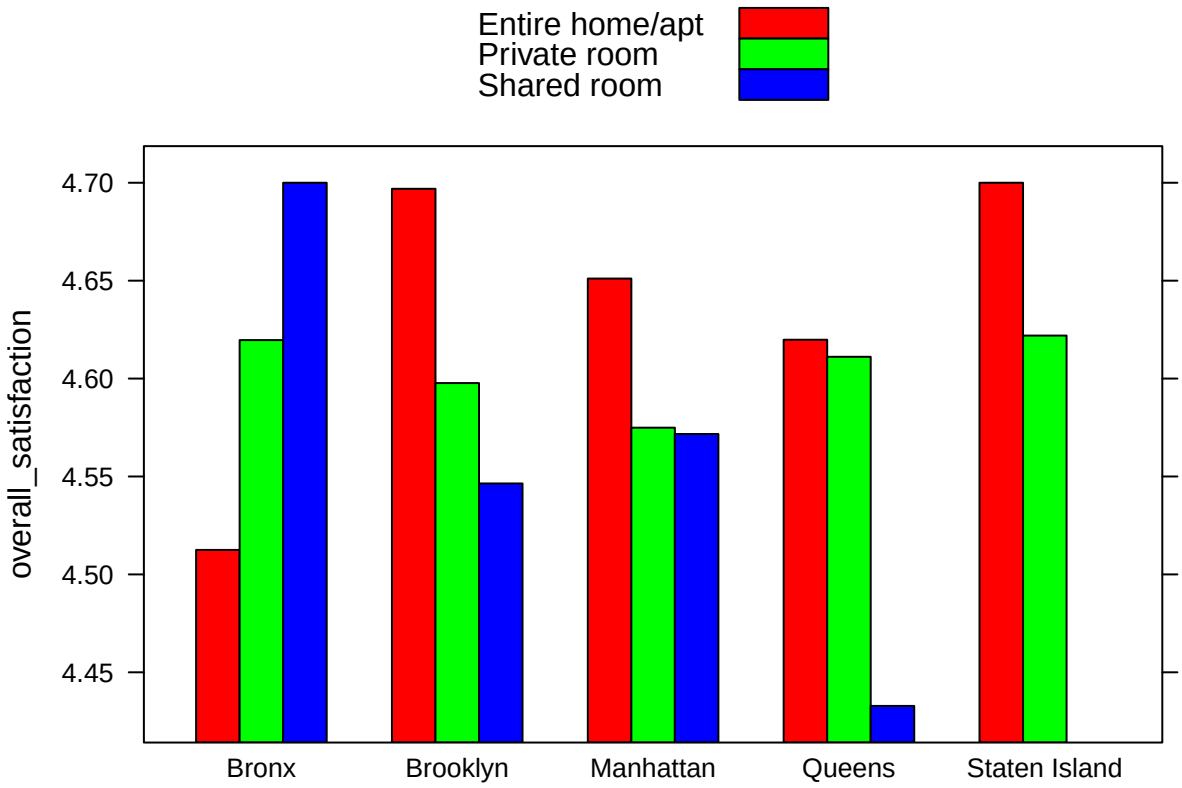

```
#plot price and reviews scatterplot

plot(Airbnb_cleaned$price, Airbnb_cleaned$reviews, main="Scatterplot for Price and Reviews",
     xlab="Price ", ylab="Reviews", pch=19)
```



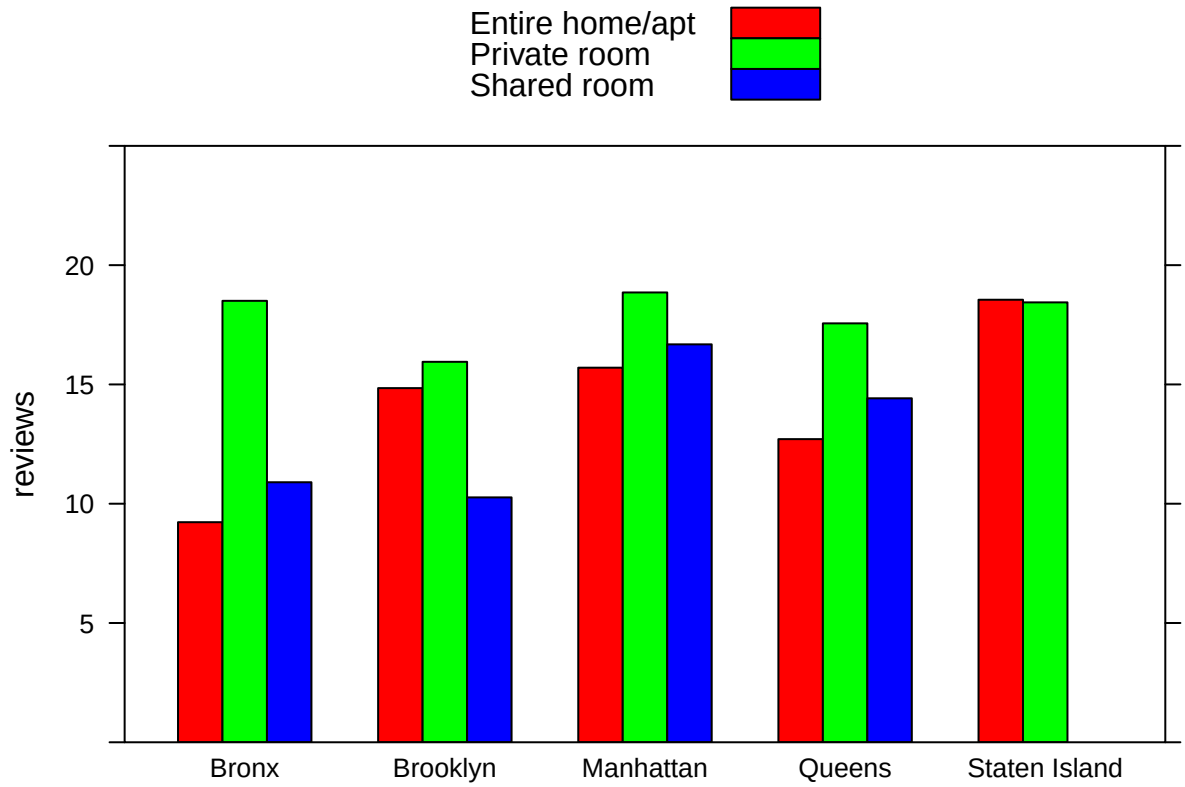

```
#plot price and accommodates scatterplot
plot(Airbnb_cleaned$price, Airbnb_cleaned$accommodates, main="Scatterplot for Price and Accommodates",
     xlab="Price ", ylab="Accommodates", pch=19)
```



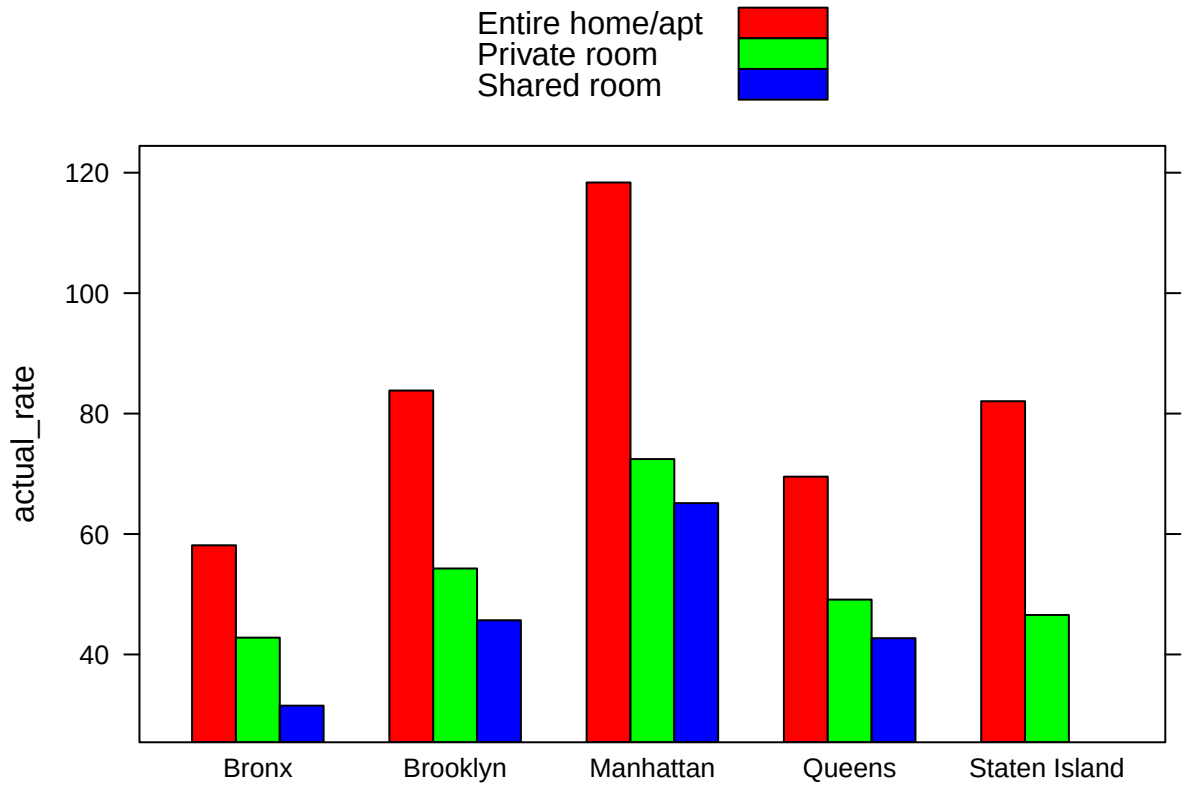
```r
#aggregate and preparing barchart of overall_satisfaction ,borough,room_type by grouping room_type

library(lattice)
seg.agg <- aggregate(overall_satisfaction ~ borough +room_type,data=Airbnb_cleaned,mean)
barchart(overall_satisfaction ~ borough, data=seg.agg,
         groups=room_type, auto.key=TRUE,
         par.settings = simpleTheme(col=c("red", "green","blue")))
```



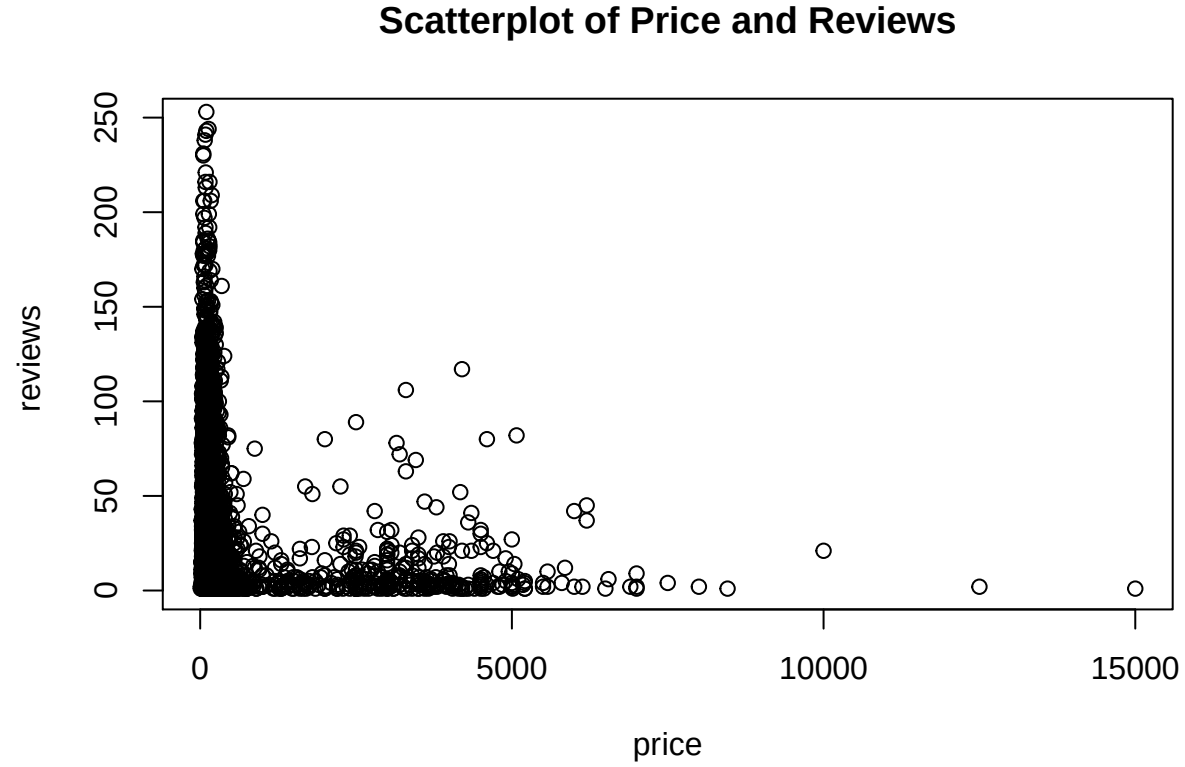

```r
#aggregate and preparing barchart of reviews ,borough,room_type by grouping room_type
seg.agg2 <- aggregate(reviews ~ borough +room_type,data=Airbnb_cleaned,mean)
barchart(reviews ~ borough, data=seg.agg2,ylim=c(0,25),
         groups=room_type, auto.key=TRUE,
         par.settings = simpleTheme(col=c("red", "green","blue")))
```



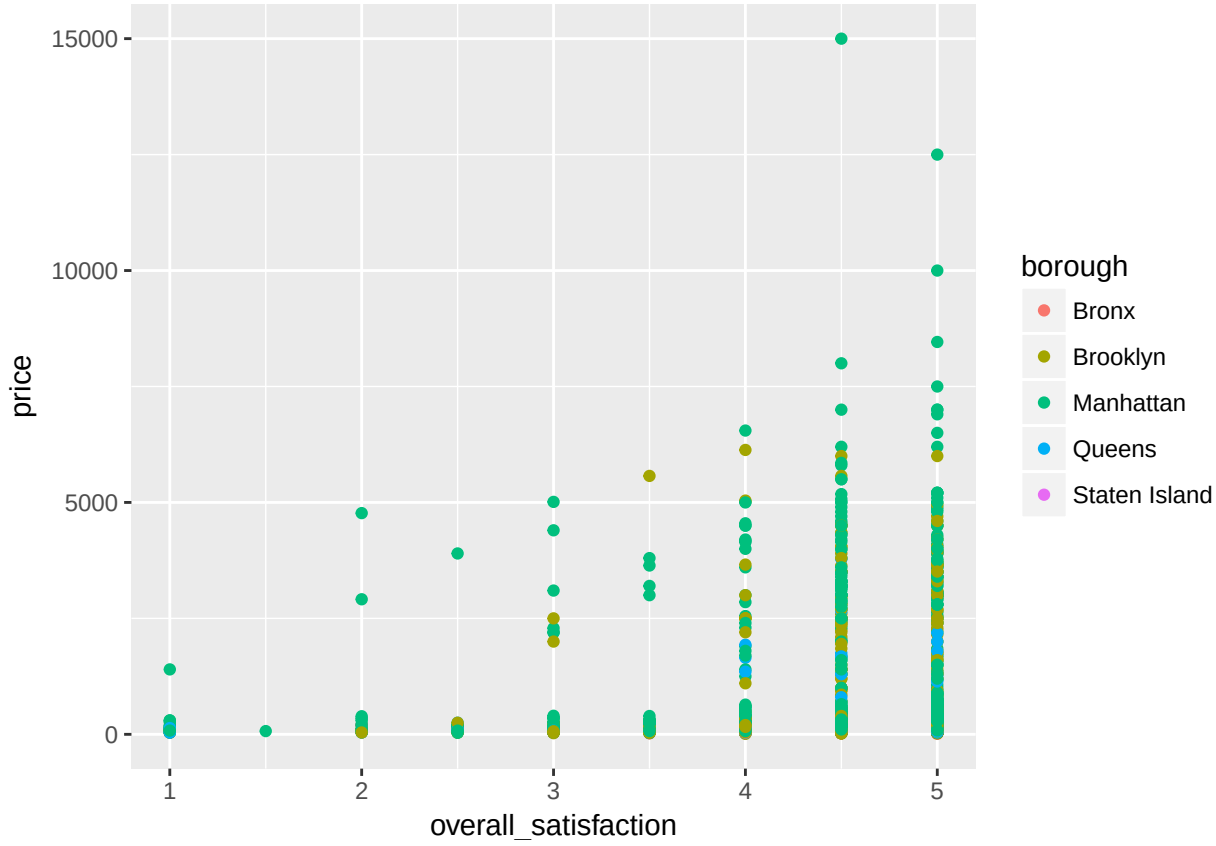

```
#aggregate and preparing barchart of actual_rate,borough,room_type by grouping actual_rate
se.agg3 <- aggregate(actual_rate ~ borough + room_type,data=Airbnb_cleaned,mean)
barchart(actual_rate ~borough,data=se.agg3,
         group=room_type,auto.key=TRUE,par.settings=simpleTheme(col=c("red","green","blue")))
```




```
#scatterplot of price and reviews

plot(Airbnb_cleaned$price,Airbnb_cleaned$reviews, main="Scatterplot of Price and Reviews",
     xlab="price", ylab="reviews ", pch=1,xlim=c(0,15000),ylim=c(0,250))
```

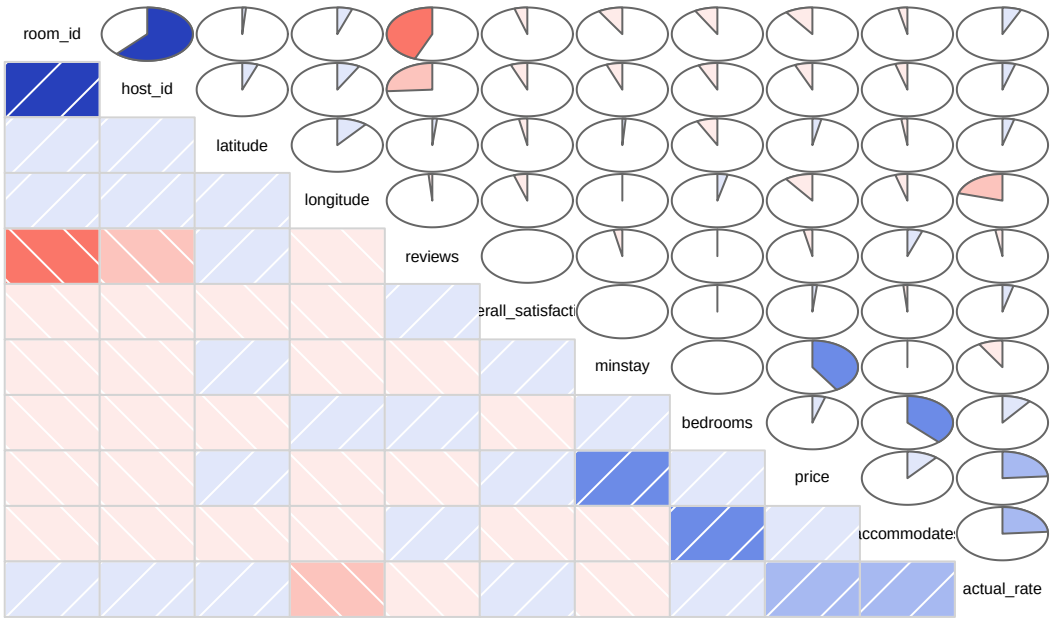


```
#qplot of overall_satisfaction and price by borough
library(ggplot2)

qplot( overall_satisfaction,price, data = Airbnb_cleaned, colour = borough)
```

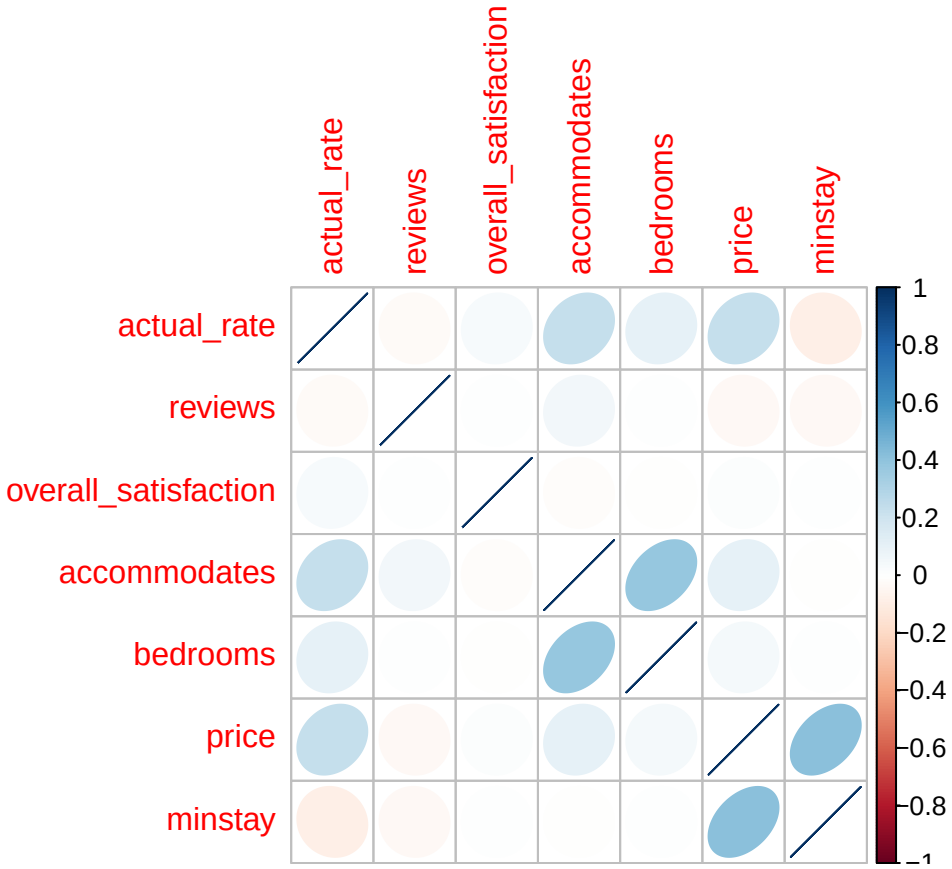



```
#Corrgram of all obs
library(corrgram)
corrgram(Airbnb_cleaned, order=TRUE, lower.panel=panel.shade,
         upper.panel=panel.pie, text.panel=panel.txt,
         main="Corrogram of Airbnb")
```

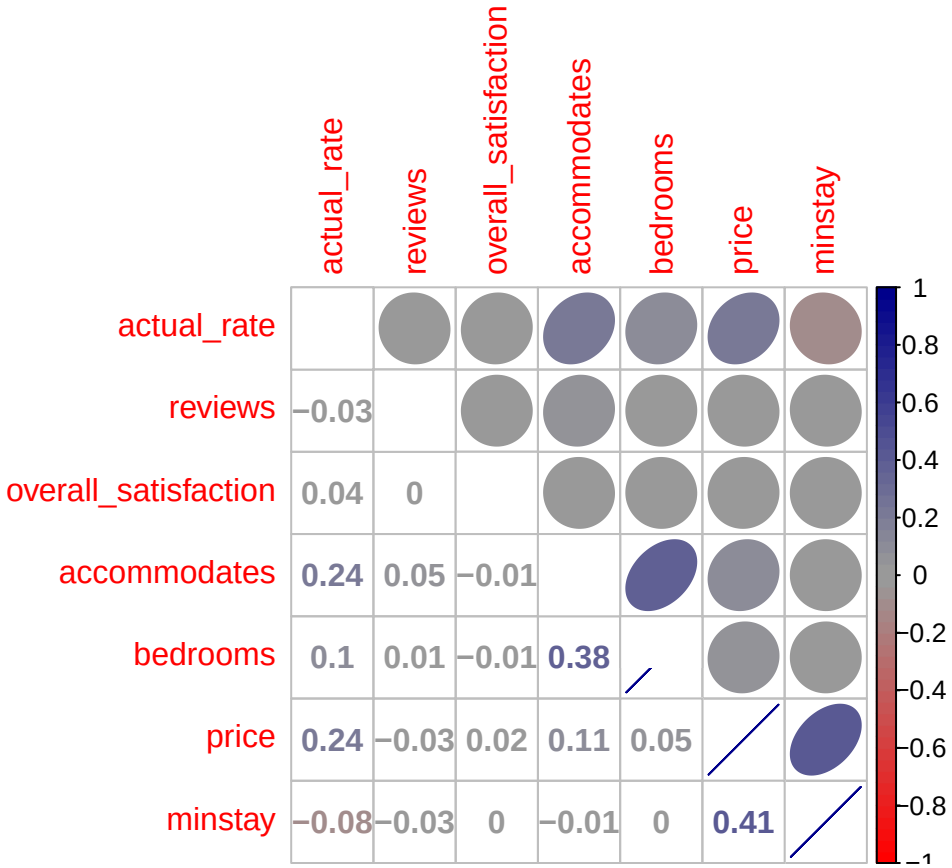
Corrogram of Airbnb




```
#corrplot of actual_rate,overall_satisfaction,reviews,bedrooms,minstay,accommodates,price

library(corrplot)
corrplot(corr=cor(Airbnb_cleaned[ , c(1,7:12)], use="complete.obs"),
         method ="ellipse")
```




```
#corrplot mixed of  actual_rate,overall_satisfaction,reviews,bedrooms,minstay,accommodates,price

library(gplots)
par(mfrow=c(1, 1))
corrplot.mixed(corr=cor(Airbnb_cleaned[ , c(1,7:12)], use="complete.obs"),
               upper="ellipse", tl.pos="lt",
               col = colorpanel(50, "red", "gray60", "blue4"))
```




```
#variance  of overall_satisfaction and price
var(Airbnb_cleaned$overall_satisfaction,Airbnb_cleaned$price, na.rm = TRUE)
```

```
## [1] 3.581694
```

```
#covariance of all_satisfaction,price,accommodates,reviews,bedrooms,minstay
cov(Airbnb_cleaned$overall_satisfaction,Airbnb_cleaned$price+Airbnb_cleaned$accommodates+Airbnb_cleaned
```

```
## [1] 3.602341
```

```
#Multiple linear regression

fit <- lm(Airbnb_cleaned$overall_satisfaction ~ Airbnb_cleaned$borough )


fit2 <- lm(Airbnb_cleaned$overall_satisfaction ~ Airbnb_cleaned$room_type)



fit3 <- lm(overall_satisfaction ~ reviews+bedrooms+accommodates+actual_rate,data=Airbnb_cleaned)


fit4 <-lm(formula = price ~ room_type + neighborhood + accommodates + bedrooms
          + minstay, data = Airbnb_cleaned)
```

```
#fucntions for fit
coefficients(fit)
```

```
##                          (Intercept)      Airbnb_cleaned$boroughBrooklyn
```

```
##                          4.598802395                          0.045860919
##     Airbnb_cleaned$boroughManhattan        Airbnb_cleaned$boroughQueens
##                          0.024235029                          0.006480624
## Airbnb_cleaned$boroughStaten Island
##                          0.048738588
```

```
confint(fit, level=0.95)
```

```
##                                         2.5 %     97.5 %
## (Intercept)                          4.52827866 4.66932613
## Airbnb_cleaned$boroughBrooklyn      -0.02548075 0.11720259
## Airbnb_cleaned$boroughManhattan     -0.04686782 0.09533788
## Airbnb_cleaned$boroughQueens        -0.06835559 0.08131683
## Airbnb_cleaned$boroughStaten Island -0.08760597 0.18508315
```

```
#fucntions for fit2
coefficients(fit2)
```

```
##                          (Intercept) Airbnb_cleaned$room_typePrivate room
##                           4.66441052                          -0.07484794
##  Airbnb_cleaned$room_typeShared room
##                          -0.11635326
```

```
confint(fit2, level=0.95)
```

```
##                                        2.5 %      97.5 %
## (Intercept)                         4.65550925  4.67331179
## Airbnb_cleaned$room_typePrivate room -0.08838547 -0.06131040
## Airbnb_cleaned$room_typeShared room  -0.15838821 -0.07431832
```

```
#t-test
t.test(Airbnb_cleaned$overall_satisfaction,Airbnb_cleaned$actual_rate)
```

```
##
##  Welch Two Sample t-test
##
## data:  Airbnb_cleaned$overall_satisfaction and Airbnb_cleaned$actual_rate
## t = -124.41, df = 18838, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -81.64727 -79.11443
## sample estimates:
## mean of x mean of y
##   4.62987  85.01072
```

```
# paired t-test
t.test(Airbnb_cleaned$overall_satisfaction,Airbnb_cleaned$bedrooms,paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  Airbnb_cleaned$overall_satisfaction and Airbnb_cleaned$bedrooms
## t = 738.99, df = 18837, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.599436 3.618581
## sample estimates:
## mean of the differences
```

```
##                3.609008
```

```
t.test(Airbnb_cleaned$overall_satisfaction,Airbnb_cleaned$actual_rate,paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  Airbnb_cleaned$overall_satisfaction and Airbnb_cleaned$actual_rate
## t = -124.43, df = 18837, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -81.64701 -79.11469
## sample estimates:
## mean of the differences
##               -80.38085
```

```
res <- t.test(Airbnb_cleaned$overall_satisfaction, mu = 25)




# printing the mean
res$estimate
```

```
## mean of x
##   4.62987
```

```
# printing the confidence interval
res$conf.int
```

```
## [1] 4.623229 4.636512
## attr(,"conf.level")
## [1] 0.95
```

```
########################################Insights#################################
```

```
#The table above presents the effects for the predictors.
# 1 ) The second column shows the parameter estimates: a negative estimate means that there is
# a negative impact to the price. The opposite can be said for a positive estimate.
# 2 )The final column indicates the significance of the pattern: the more stars, the more significant
#the effect ( for p < 0.05).
# 3) When testing for significance we use a simple two-sided t-test.
# When testing we test for the variables parameter estimate to be  Since we have a large test size,
# our distribution will be approximately normal distributed.
# 4 ) The critical value for a normal distribution on a 95% confidence interval is 1.96.
# 5) Surprisingly the occupancy of apartments has a positive significant effect on the price.
#The more people who can live in the apartment the more expensive it should be, as the visitors will
# probably split the costs. Also, the accommodation variable could be positive correlated with the size
#of the apartment.which would also lead to higher prices.

#6 )The final two variables does not show any significant influence on the price
```