script.R

frank

Sat Feb 09 21:42:35 2019

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------------------------- tidyverse 1.2
```

```
## v ggplot2 3.1.0     v purrr   0.3.0
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
```

```
## -- Conflicts -------------------------------------------------------------------- tidyverse_conflicts()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
```

```
## The following object is masked from 'package:tidyr':
##
##     smiths
```

```
library(expss)
```

```
##
## Attaching package: 'expss'
```

```
## The following objects are masked from 'package:stringr':
##
##     fixed, regex
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, compute, first, last, na_if, recode, vars
```

```
## The following objects are masked from 'package:purrr':
##
##     keep, modify, modify_if, transpose
```

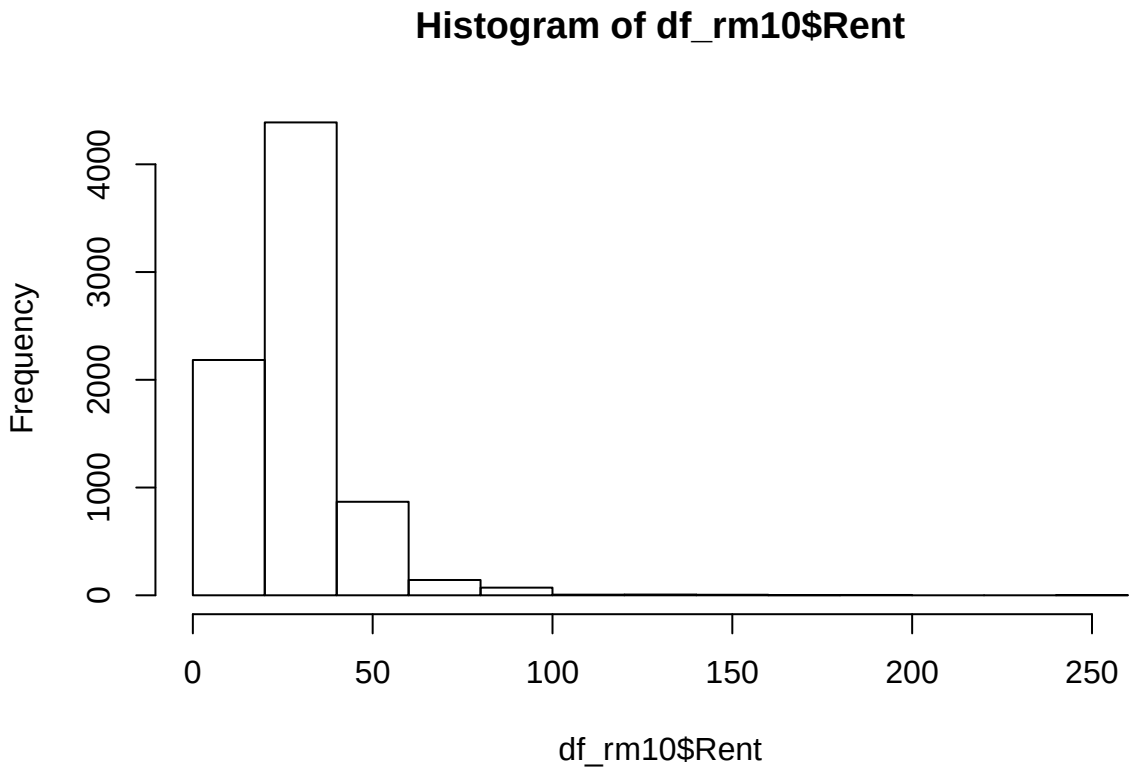
```
## The following object is masked from 'package:tidyr':
##
##     nest
```

```
## The following object is masked from 'package:ggplot2':
##
##     vars
```

```
df <- read.csv('greenbuildings.csv')

# Replicating the Analyst's results
df_rm10 <- df[df$leasing_rate > 10,] #this removes rows where leasing_rate is < 10
df_rm10_green <- df[df$green_rating == 1, ]
df_rm10_ngreen <- df[df$green_rating != 1, ]
```

```r
rent <- median (df_rm10$Rent) #rent of all buildings
g_rent <- median (df_rm10_green$Rent) #rent of >10 occupancy green buildings
ng_rent <- median (df_rm10_ngreen$Rent) #rent of >10 occupancy non-green buildings
d_rent <- g_rent-ng_rent #difference in rent from green and non-green

n <- 30 #years of operation
o_rate <- 0.90 #occupancy rate
sqft <- 250000 #square footage of building
c_cost <- 100000000 #cost of building
g_cert <- 0.05 #percentage to certify green

#payoff analysis
p1 <- (c_cost*g_cert)/(d_rent*sqft) #at 100% occupancy
p2 <- (c_cost*g_cert)/(d_rent*sqft*o_rate) #at 90% occupancy

t1 <- (n-p1)*d_rent*sqft #total payoff under senario 1
t2 <- (n-p2)*d_rent*sqft*o_rate #total payoff under senario 2

full_rate <- c(100:1)/100 #a range of values from 0-100 for occupancy rates
full_payoff <- (c_cost*g_cert)/(d_rent*sqft*full_rate) #pay off times based on occupancy rates

df_plot <- data.frame(full_payoff,full_rate)

analyst_plot1 <- ggplot(df_plot)+
  geom_point(mapping = aes(x = full_payoff, y = full_rate), color='black')+
  labs(title="Payoff Analysis",x="Years to Payoff",y="Occupancy Rate")

# A rudimentary analysis of Analyst's parameters gives us a basic understanding of how long it would ta

analyst_plot2 <-ggplot(df_plot)+
  geom_point(mapping = aes(x = full_payoff, y = full_rate), color='black')+
  labs(title="Payoff Analysis",subtitle="30 Year Operational Life",x="Years to Payoff",y="Occupancy Rat
  coord_cartesian(xlim=c(0, 30))+
  geom_hline(yintercept=1, linetype="dashed", color = "red")+
  geom_vline(xintercept=7.6, linetype="dashed", color = "red")+
  geom_hline(yintercept=0.25, linetype="dashed", color = "red")+
  geom_vline(xintercept=30, linetype="dashed", color = "red")


# Here we see a more accurate estimate of the payoff timetable for the building if we built it green. I

# Regardless of the analysis there are a number of points of fault within the Analyst's analysis.

# One: by eliminating buildings where its occupancy rates are below 10%, we exclude important data that

# Two: by utilizing the median rental value of the subsets: green and non-green, we skew the data becau

point2 <- table(df$green_rating)

# Where "0" represent the number of non-green buildings and "1" represent the number of green buildings

# Three: As for the decision to utilize the median instead of the mean value of rent, a cursory assessm
```

```r
hist(df_rm10$Rent)
```




```r
# A histogram of the Analyst's data set provides a view of how the rental prices are positively skewed.

table <- matrix(c(
  median(df_rm10$Rent),
  mean(df_rm10$Rent)
    ),ncol=2,byrow=TRUE)

colnames(table) <- c("Median","Mean")
rownames(table) <- c("Whole Dataset")

table <- as.table(table)
table
```

```
##                 Median     Mean
## Whole Dataset 25.29000 28.58585
```

```r
# However, if we were to compare the price differences within the green versus non-green building subse

table <- matrix(c(
  median(df_rm10$Rent),
  mean(df_rm10$Rent),
  median(df_rm10_green$Rent),
  mean(df_rm10_green$Rent),
  median(df_rm10_ngreen$Rent),
  mean(df_rm10_ngreen$Rent)
```

```
),ncol=2,byrow=TRUE)

colnames(table) <- c("Median","Mean")
rownames(table) <- c("Whole Dataset","Green","Non-Green")

table <- as.table(table)
table
```

```
##                  Median     Mean
## Whole Dataset 25.29000 28.58585
## Green         27.60000 30.01603
## Non-Green     25.00000 28.26678
```

```
## If we were to conduct the analysis with the full set of data and the mean-value, we see a different p

df <- read.csv('greenbuildings.csv')
df_green <- df[df$green_rating == 1, ]
df_ngreen <- df[df$green_rating != 1, ]

rent <- mean (df$Rent) #rent of all buildings
g_rent <- mean (df_green$Rent) #rent of occupancy green buildings
ng_rent <- mean (df_ngreen$Rent) #rent of occupancy non-green buildings
d_rent <- g_rent-ng_rent #difference in rent from green and non-green

n <- 30 #years of operation
o_rate <- 0.90 #occupancy rate
sqft <- 250000 #square footage of building
c_cost <- 100000000 #cost of building
g_cert <- 0.05 #percentage to certify green

#payoff analysis
p1 <- (c_cost*g_cert)/(d_rent*sqft) #at 100% occupancy
p2 <- (c_cost*g_cert)/(d_rent*sqft*o_rate) #at 90% occupancy

t1 <- (n-p1)*d_rent*sqft #total payoff under senario 1
t2 <- (n-p2)*d_rent*sqft*o_rate #total payoff under senario 2

full_rate <- c(100:1)/100 #a range of values from 0-100 for occupancy rates
full_payoff <- (c_cost*g_cert)/(d_rent*sqft*full_rate) #pay off times based on occupancy rates

df_plot <- data.frame(full_payoff,full_rate)

# Begin Analysis of the better data

analyst_plot2 <-ggplot(df_plot)+
  geom_point(mapping = aes(x = full_payoff, y = full_rate), color='black')+
  labs(title="Payoff Analysis",subtitle="30 Year Operational Life",x="Years to Payoff",y="Occupancy Rate
  coord_cartesian(xlim=c(0, 30))+
  geom_hline(yintercept=1, linetype="dashed", color = "red")+
  geom_vline(xintercept=p1, linetype="dashed", color = "red")+
  geom_hline(yintercept=0.37, linetype="dashed", color = "red")+
  geom_vline(xintercept=30, linetype="dashed", color = "red")
```

```
# Here we see a more accurate estimate of the payoff timetable for the building if we built it green us

# In conclusion, we see that a more accurate assessment of the viability of building green is represent
```