FML assignment 1

shiva gadila

2023-01-29

```
library(tinytex)
```

#1. Downloaded the Dataset from Kaggle.

#2. IMPORTING THE DATASET:

```
Hotel.Reservations <- read.csv("~/Desktop/Hotel Reservations.csv")
head(Hotel.Reservations)
```

```
##   Booking_ID no_of_adults no_of_children no_of_weekend_nights
no_of_week_nights
## 1   INN00001            2              0                    1
2
## 2   INN00002            2              0                    2
3
## 3   INN00003            1              0                    2
1
## 4   INN00004            2              0                    0
2
## 5   INN00005            2              0                    1
1
## 6   INN00006            2              0                    0
2
##   type_of_meal_plan required_car_parking_space room_type_reserved
lead_time
## 1       Meal Plan 1                          0         Room_Type 1
224
## 2      Not Selected                          0         Room_Type 1
5
## 3       Meal Plan 1                          0         Room_Type 1
1
## 4       Meal Plan 1                          0         Room_Type 1
211
## 5      Not Selected                          0         Room_Type 1
48
## 6       Meal Plan 2                          0         Room_Type 1
346
##   arrival_year arrival_month arrival_date market_segment_type
repeated_guest
## 1         2017            10            2             Offline
0
## 2         2018            11            6              Online
```

```
0
## 3            2018              2              28              Online
0
## 4            2018              5              20              Online
0
## 5            2018              4              11              Online
0
## 6            2018              9              13              Online
0
##   no_of_previous_cancellations no_of_previous_bookings_not_canceled
## 1                            0                                    0
## 2                            0                                    0
## 3                            0                                    0
## 4                            0                                    0
## 5                            0                                    0
## 6                            0                                    0
##   avg_price_per_room no_of_special_requests booking_status
## 1              65.00                      0    Not_Canceled
## 2             106.68                      1    Not_Canceled
## 3              60.00                      0        Canceled
## 4             100.00                      0        Canceled
## 5              94.50                      0        Canceled
## 6             115.00                      1        Canceled
```

#3. DESCRIPTIVE STATISTICS:

```
summary(Hotel.Reservations)

##   Booking_ID          no_of_adults    no_of_children    no_of_weekend_nights
## Length:36275        Min.   :0.000   Min.   : 0.0000   Min.   :0.0000
## Class :character    1st Qu.:2.000   1st Qu.: 0.0000   1st Qu.:0.0000
## Mode  :character    Median :2.000   Median : 0.0000   Median :1.0000
##                     Mean   :1.845   Mean   : 0.1053   Mean   :0.8107
##                     3rd Qu.:2.000   3rd Qu.: 0.0000   3rd Qu.:2.0000
##                     Max.   :4.000   Max.   :10.0000   Max.   :7.0000
##  no_of_week_nights type_of_meal_plan  required_car_parking_space
## Min.   : 0.000     Length:36275       Min.   :0.00000
## 1st Qu.: 1.000     Class :character   1st Qu.:0.00000
## Median : 2.000     Mode  :character   Median :0.00000
## Mean   : 2.204                        Mean   :0.03099
## 3rd Qu.: 3.000                        3rd Qu.:0.00000
## Max.   :17.000                        Max.   :1.00000
##  room_type_reserved   lead_time        arrival_year   arrival_month
## Length:36275        Min.   :  0.00   Min.   :2017   Min.   : 1.000
## Class :character    1st Qu.: 17.00   1st Qu.:2018   1st Qu.: 5.000
## Mode  :character    Median : 57.00   Median :2018   Median : 8.000
##                     Mean   : 85.23   Mean   :2018   Mean   : 7.424
##                     3rd Qu.:126.00   3rd Qu.:2018   3rd Qu.:10.000
##                     Max.   :443.00   Max.   :2018   Max.   :12.000
##   arrival_date  market_segment_type repeated_guest
```

```
## Min.   : 1.0   Length:36275        Min.   :0.00000
## 1st Qu.: 8.0   Class :character    1st Qu.:0.00000
## Median :16.0   Mode  :character    Median :0.00000
## Mean   :15.6                       Mean   :0.02564
## 3rd Qu.:23.0                       3rd Qu.:0.00000
## Max.   :31.0                       Max.   :1.00000
## no_of_previous_cancellations no_of_previous_bookings_not_canceled
## Min.   : 0.00000             Min.   : 0.0000
## 1st Qu.: 0.00000             1st Qu.: 0.0000
## Median : 0.00000             Median : 0.0000
## Mean   : 0.02335             Mean   : 0.1534
## 3rd Qu.: 0.00000             3rd Qu.: 0.0000
## Max.   :13.00000             Max.   :58.0000
## avg_price_per_room no_of_special_requests booking_status
## Min.   :  0.00     Min.   :0.0000         Length:36275
## 1st Qu.: 80.30     1st Qu.:0.0000         Class :character
## Median : 99.45     Median :0.0000         Mode  :character
## Mean   :103.42     Mean   :0.6197
## 3rd Qu.:120.00     3rd Qu.:1.0000
## Max.   :540.00     Max.   :5.0000
```

#4.TRANSFORMATION OF NO_OF_ADULTS VARIABLES :

```
Hotel.Reservations$no_of_adults<-
mean(Hotel.Reservations$no_of_adults)/sd(Hotel.Reservations$no_of_adults)
head(Hotel.Reservations$no_of_adults)

## [1] 3.556795 3.556795 3.556795 3.556795 3.556795 3.556795
```
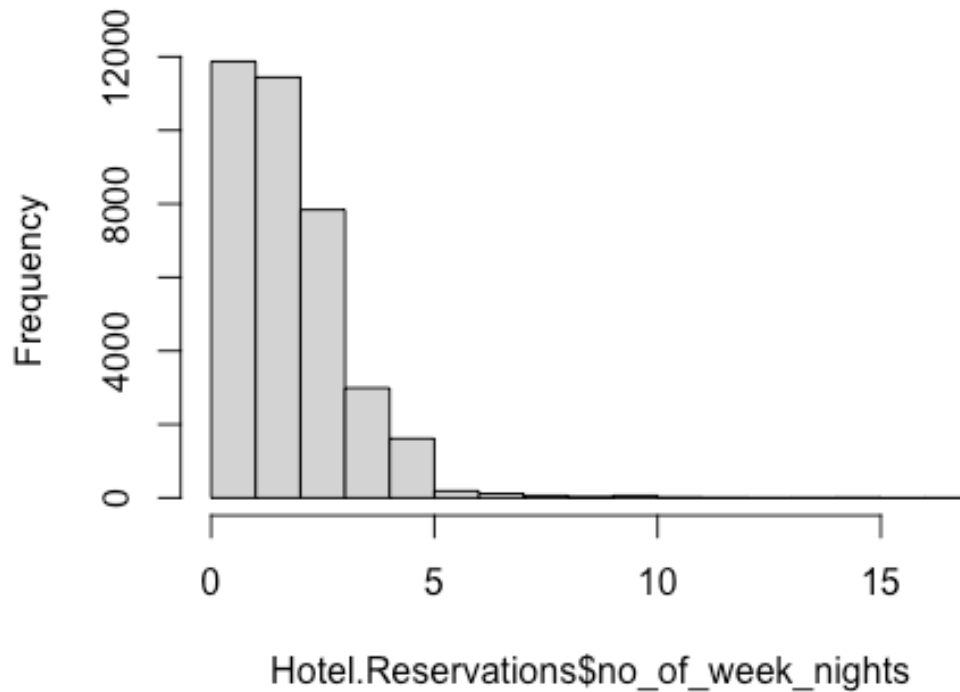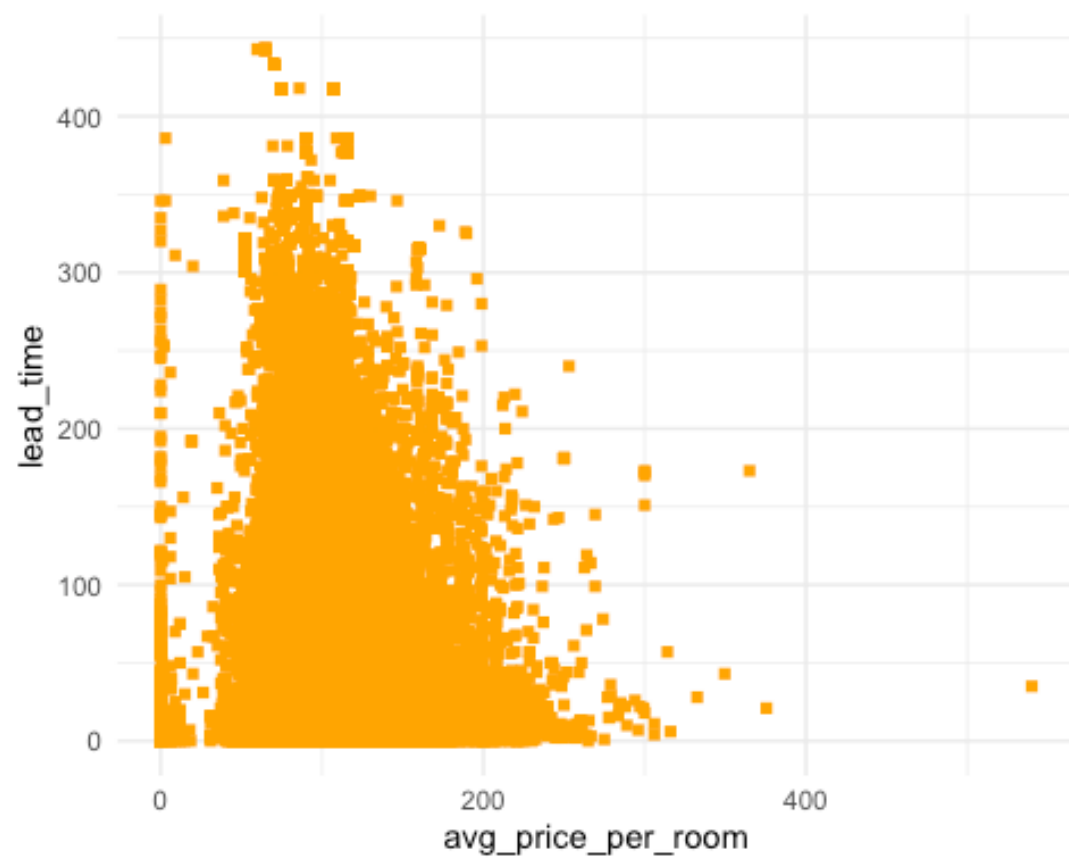
#5.HISTOGRAPH AND SCATTERPLOT:

```
hist(Hotel.Reservations$no_of_week_nights )


library(ggplot2)
```

Histogram of Hotel.Reservations$no_of_week_nigh


```
ggplot(Hotel.Reservations) +
  aes(x =avg_price_per_room , y =lead_time) +
  geom_point(shape = "square", size = 1.5, colour = "orange") +
  theme_minimal()
```

```