FML-Assignment1

Sandhya K

2024-02-04

*1.Dataset import:*

# Downloading the Interstellar Travel Customer Satisfaction Analysis dataset from Kaggle.  
url <- "https://www.kaggle.com/datasets/anthonytherrien/interstellar-travel-customer-satisfaction-analysis"  
ITCS\_data <- read.csv("C://Users//sandh//Downloads//ITCS\_data.csv")

# Dataset first few rows  
head(ITCS\_data)

## Age Gender Occupation Travel.Class Destination  
## 1 14 Female Colonist Business Gliese 581  
## 2 22 Male Tourist Economy Alpha Centauri  
## 3 62 Female Businessperson Luxury Alpha Centauri  
## 4 21 Female Colonist Economy Lalande 21185  
## 5 42 Male Explorer Luxury Exotic Destination 10  
## 6 30 Male Other Economy Tau Ceti  
## Star.System Distance.to.Destination..Light.Years.  
## 1 Cunningham Mountains 1.09  
## 2 Hayes Trace 5.70  
## 3 Anna Port 0.37  
## 4 Henry Ville 0.32  
## 5 Graves Mall 6.17  
## 6 Vazquez Tunnel 10.51  
## Duration.of.Stay..Earth.Days. NumberofCompanions Purpose.of.Travel  
## 1 11 5 Tourism  
## 2 23 0 Research  
## 3 4 1 Tourism  
## 4 23 1 Tourism  
## 5 42 1 Colonization  
## 6 60 1 Colonization  
## Transportation.Type Price..Galactic.Credits. Booking.Date DepartureDate  
## 1 Warp Drive 828.9493 17-09-2023 07-01-2025  
## 2 Solar Sailing 488.4691 31-03-2023 26-12-2025  
## 3 Ion Thruster 183.7459 19-05-2022 04-01-2025  
## 4 Warp Drive 358.7540 13-04-2023 09-02-2024  
## 5 Ion Thruster 3073.7599 12-06-2023 15-03-2024  
## 6 Warp Drive 1136.1619 13-04-2023 16-02-2025  
## SpecialRequests LoyaltyProgramMember Month CustomerSatisfactionScore  
## 1 Other No 9 105  
## 2 Other No 3 102  
## 3 None Yes 5 100  
## 4 None No 4 108  
## 5 Special Meal No 6 97  
## 6 Other No 4 105

*2.Printing-Descriptive Statistics:*

# Quantitative variables-Summary  
summary(ITCS\_data$Age)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0 16.0 27.0 31.1 43.0 99.0

# Categorical variable-Frequency table  
table(ITCS\_data$Gender)

##   
## Female Male   
## 228563 319005

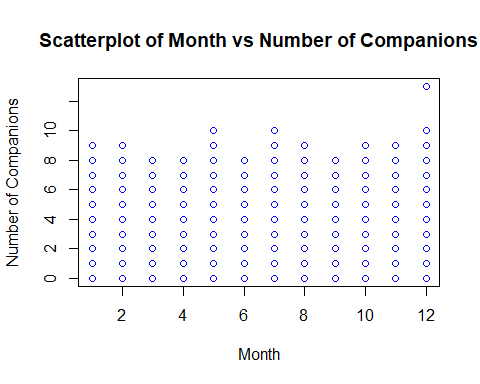
*3. Transforming the Variable:*

#Square root transformation for one quantitative variable  
ITCS\_data$Month\_log <- log(ITCS\_data$Month)  
# Print or visualize the transformed data  
head(ITCS\_data)

## Age Gender Occupation Travel.Class Destination  
## 1 14 Female Colonist Business Gliese 581  
## 2 22 Male Tourist Economy Alpha Centauri  
## 3 62 Female Businessperson Luxury Alpha Centauri  
## 4 21 Female Colonist Economy Lalande 21185  
## 5 42 Male Explorer Luxury Exotic Destination 10  
## 6 30 Male Other Economy Tau Ceti  
## Star.System Distance.to.Destination..Light.Years.  
## 1 Cunningham Mountains 1.09  
## 2 Hayes Trace 5.70  
## 3 Anna Port 0.37  
## 4 Henry Ville 0.32  
## 5 Graves Mall 6.17  
## 6 Vazquez Tunnel 10.51  
## Duration.of.Stay..Earth.Days. NumberofCompanions Purpose.of.Travel  
## 1 11 5 Tourism  
## 2 23 0 Research  
## 3 4 1 Tourism  
## 4 23 1 Tourism  
## 5 42 1 Colonization  
## 6 60 1 Colonization  
## Transportation.Type Price..Galactic.Credits. Booking.Date DepartureDate  
## 1 Warp Drive 828.9493 17-09-2023 07-01-2025  
## 2 Solar Sailing 488.4691 31-03-2023 26-12-2025  
## 3 Ion Thruster 183.7459 19-05-2022 04-01-2025  
## 4 Warp Drive 358.7540 13-04-2023 09-02-2024  
## 5 Ion Thruster 3073.7599 12-06-2023 15-03-2024  
## 6 Warp Drive 1136.1619 13-04-2023 16-02-2025  
## SpecialRequests LoyaltyProgramMember Month CustomerSatisfactionScore  
## 1 Other No 9 105  
## 2 Other No 3 102  
## 3 None Yes 5 100  
## 4 None No 4 108  
## 5 Special Meal No 6 97  
## 6 Other No 4 105  
## Month\_log  
## 1 2.197225  
## 2 1.098612  
## 3 1.609438  
## 4 1.386294  
## 5 1.791759  
## 6 1.386294

*4. Scatterplot and Histogram:*

# Scatterplot for two quantitative variables  
plot(ITCS\_data$Month, ITCS\_data$NumberofCompanions,   
 main = "Scatterplot of Month vs Number of Companions",   
 xlab = "Month", ylab = "Number of Companions", col = "blue",  
 xlim = range(na.omit(ITCS\_data$Month)))

 #Creating a Histogram

hist(ITCS\_data$CustomerSatisfactionScore,   
 main = "Histogram of Customer Satisfaction Score",   
 col = "blue",   
 xlab = "Customer Satisfaction Score",   
 ylab = "Frequency")

