fml-assignment1

2023-09-10

#Importing dataset by giving the path

```
library("readxl")
retail_sales_dataset<-read.csv("C:/Users/gandu/OneDrive/retail_sales_dataset.csv")
View(retail_sales_dataset)
```

#descriptive statistics for selection of quantitative and categorical variables

```
summary(retail_sales_dataset)
```

```
##  Transaction.ID       Date           Customer.ID          Gender
##  Min.   :   1.0   Length:1000        Length:1000        Length:1000
##  1st Qu.: 250.8   Class :character   Class :character   Class :character
##  Median : 500.5   Mode  :character   Mode  :character   Mode  :character
##  Mean   : 500.5
##  3rd Qu.: 750.2
##  Max.   :1000.0
##       Age        Product.Category      Quantity     Price.per.Unit
##  Min.   :18.00   Length:1000        Min.   :1.000   Min.   : 25.0
##  1st Qu.:29.00   Class :character   1st Qu.:1.000   1st Qu.: 30.0
##  Median :42.00   Mode  :character   Median :3.000   Median : 50.0
##  Mean   :41.39                      Mean   :2.514   Mean   :179.9
##  3rd Qu.:53.00                      3rd Qu.:4.000   3rd Qu.:300.0
##  Max.   :64.00                      Max.   :4.000   Max.   :500.0
##   Total.Amount
##  Min.   :  25
##  1st Qu.:  60
##  Median : 135
##  Mean   : 456
##  3rd Qu.: 900
##  Max.   :2000
```

```
sd(retail_sales_dataset$Transaction.ID)
```

```
## [1] 288.8194
```

```
var(retail_sales_dataset$Age)
```

```
## [1] 187.1815
```

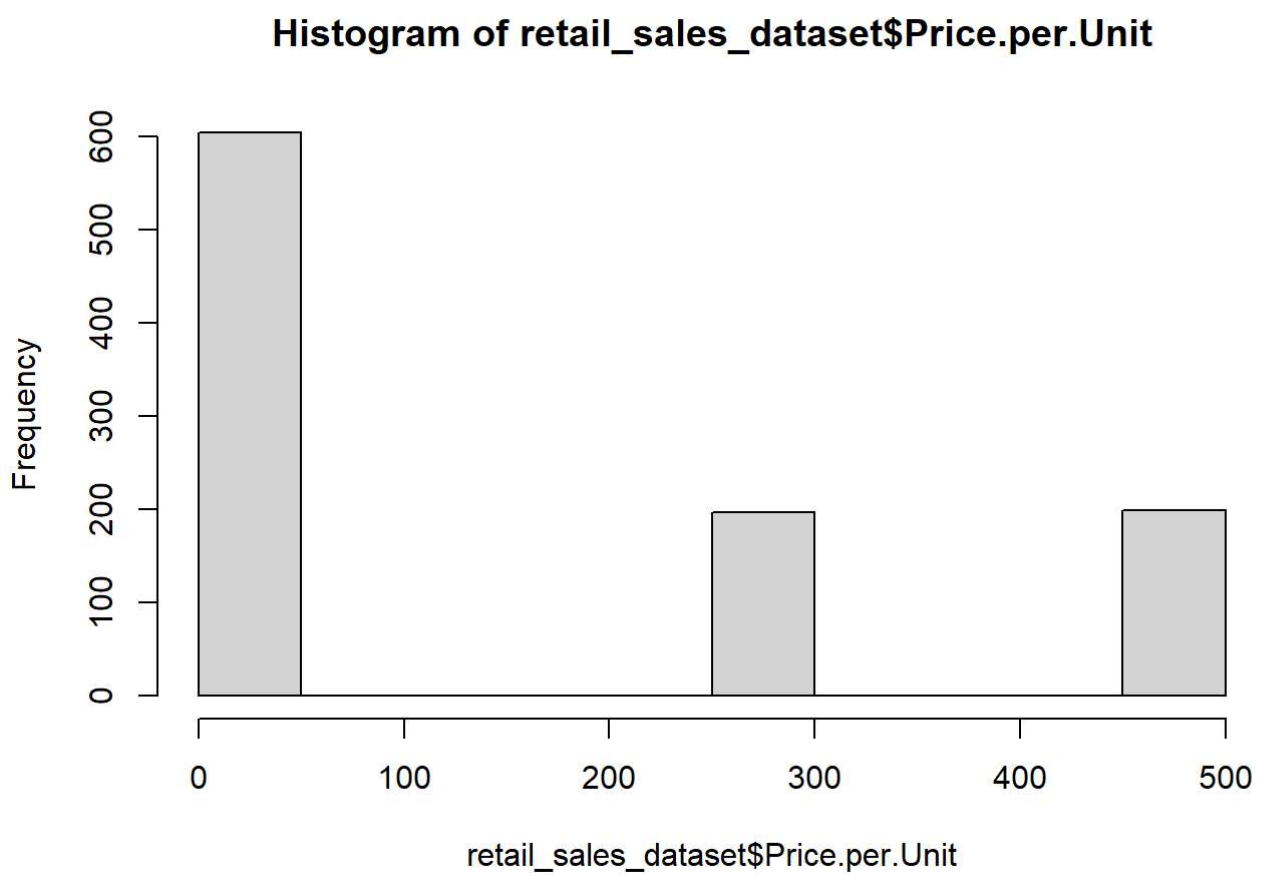
#Transform at least one variable

```
retail_sales_dataset$Transaction.ID<-(retail_sales_dataset$Transaction.ID+1)
summary(retail_sales_dataset)
```

```
##  Transaction.ID       Date           Customer.ID          Gender
##  Min.   :   2.0   Length:1000        Length:1000        Length:1000
##  1st Qu.: 251.8   Class :character   Class :character   Class :character
##  Median : 501.5   Mode  :character   Mode  :character   Mode  :character
##  Mean   : 501.5
##  3rd Qu.: 751.2
##  Max.   :1001.0
##       Age        Product.Category      Quantity     Price.per.Unit
##  Min.   :18.00   Length:1000        Min.   :1.000   Min.   : 25.0
##  1st Qu.:29.00   Class :character   1st Qu.:1.000   1st Qu.: 30.0
##  Median :42.00   Mode  :character   Median :3.000   Median : 50.0
##  Mean   :41.39                      Mean   :2.514   Mean   :179.9
##  3rd Qu.:53.00                      3rd Qu.:4.000   3rd Qu.:300.0
##  Max.   :64.00                      Max.   :4.000   Max.   :500.0
##   Total.Amount
##  Min.   :  25
##  1st Qu.:  60
##  Median : 135
##  Mean   : 456
##  3rd Qu.: 900
##  Max.   :2000
```

#one quantitative variable

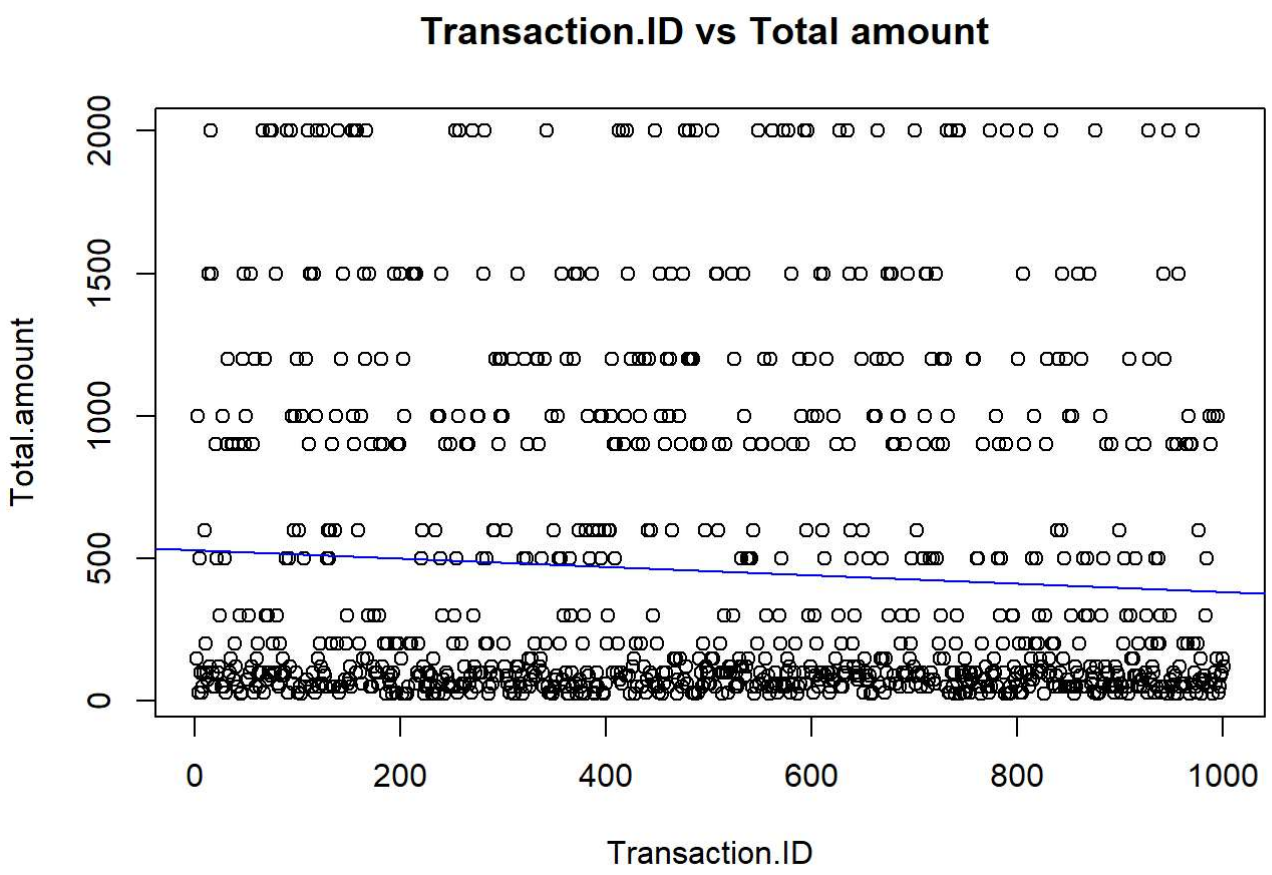
```
hist(retail_sales_dataset$Price.per.Unit)
```



#one

scatterplot

```
x<-(retail_sales_dataset$Transaction.ID)
y<-(retail_sales_dataset$Total.Amount)
# corrected code
plot(x,y,main ="Transaction.ID vs Total amount ",xlab ="Transaction.ID",ylab="Total.amount")
abline(lm(y~x),col="blue")
```


Transaction.ID vs Total amount
Total.amount
0
500
1000
1500
2000
0
200
400
600
800
1000
Transaction.ID