FML_Assignment_1

Bhargavi kundrapu

2023-09-09

```
knitr::opts_chunk$set(echo = TRUE)
```

Downloaded the dataset

The data conatins ecommerce transactions

https://www.kaggle.com\datasets\aliessamali\ecommerce

Imported the data set

```
ecommerce = read.csv("C:/Users/user/Downloads/dataset.csv", header = TRUE,
sep = ",")
```

Descriptive Statistics for Quantitative variables

```
head(ecommerce)
```

```
##   InvoiceNo StockCode                          Description Quantity
## 1    536365    85123A  WHITE HANGING HEART T-LIGHT HOLDER        6
## 2    536365     71053                  WHITE METAL LANTERN        6
## 3    536365    84406B       CREAM CUPID HEARTS COAT HANGER        8
## 4    536365    84029G KNITTED UNION FLAG HOT WATER BOTTLE        6
## 5    536365    84029E       RED WOOLLY HOTTIE WHITE HEART.        6
## 6    536365     22752          SET 7 BABUSHKA NESTING BOXES       2
##          InvoiceDate UnitPrice CustomerID        Country
## 1 01-12-2010 08:26      2.55      17850 United Kingdom
## 2 01-12-2010 08:26      3.39      17850 United Kingdom
## 3 01-12-2010 08:26      2.75      17850 United Kingdom
## 4 01-12-2010 08:26      3.39      17850 United Kingdom
## 5 01-12-2010 08:26      3.39      17850 United Kingdom
## 6 01-12-2010 08:26      7.65      17850 United Kingdom
```

```
median(ecommerce$UnitPrice)
```

```
## [1] 2.08
```

```
min(ecommerce$UnitPrice)
```

```
## [1] 0
```

```r
max(ecommerce$UnitPrice)
```

```
## [1] 38970
```

```r
min(ecommerce$Quantity)
```

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

```r
max(ecommerce$Quantity)
```

```
## [1] 80995
```

```r
mean(ecommerce$Quantity)
```

```
## [1] 11.34049
```

```r
summary(ecommerce)
```

```
##    InvoiceNo           StockCode           Description           Quantity
##  Length:541909      Length:541909      Length:541909      Min.   :    1.00
##  Class :character   Class :character   Class :character   1st Qu.:    1.00
##  Mode  :character   Mode  :character   Mode  :character   Median :    3.00
##                                                           Mean   :   11.34
##                                                           3rd Qu.:   10.00
##                                                           Max.   :80995.00
##
##   InvoiceDate          UnitPrice          CustomerID        Country
##  Length:541909      Min.   :    0.00   Min.   :12346      Length:541909
##  Class :character   1st Qu.:    1.25   1st Qu.:13953      Class :character
##  Mode  :character   Median :    2.08   Median :15152      Mode  :character
##                     Mean   :    4.65   Mean   :15288
##                     3rd Qu.:    4.13   3rd Qu.:16791
##                     Max.   :38970.00   Max.   :18287
##                                        NA's   :135080
```

Descriptive Statistics for Categorical variable

```r
table(ecommerce$Country)
```

```
##
##             Australia              Austria              Bahrain
##                  1259                  401                   19
##               Belgium               Brazil               Canada
##                  2069                   32                  151
##       Channel Islands               Cyprus       Czech Republic
##                   758                  622                   30
##               Denmark                 EIRE   European Community
##                   389                 8196                   61
##               Finland               France              Germany
##                   695                 8557                 9495
##                Greece            Hong Kong              Iceland
##                   146                  288                  182
```

```
##                    Israel             Italy             Japan
##                       297               803               358
##                   Lebanon         Lithuania             Malta
##                        45                35               127
##               Netherlands            Norway            Poland
##                      2371              1086               341
##                  Portugal               RSA      Saudi Arabia
##                      1519                58                10
##                 Singapore             Spain            Sweden
##                       229              2533               462
##               Switzerland United Arab Emirates   United Kingdom
##                      2002                68            495478
##               Unspecified               USA
##                       446               291
```
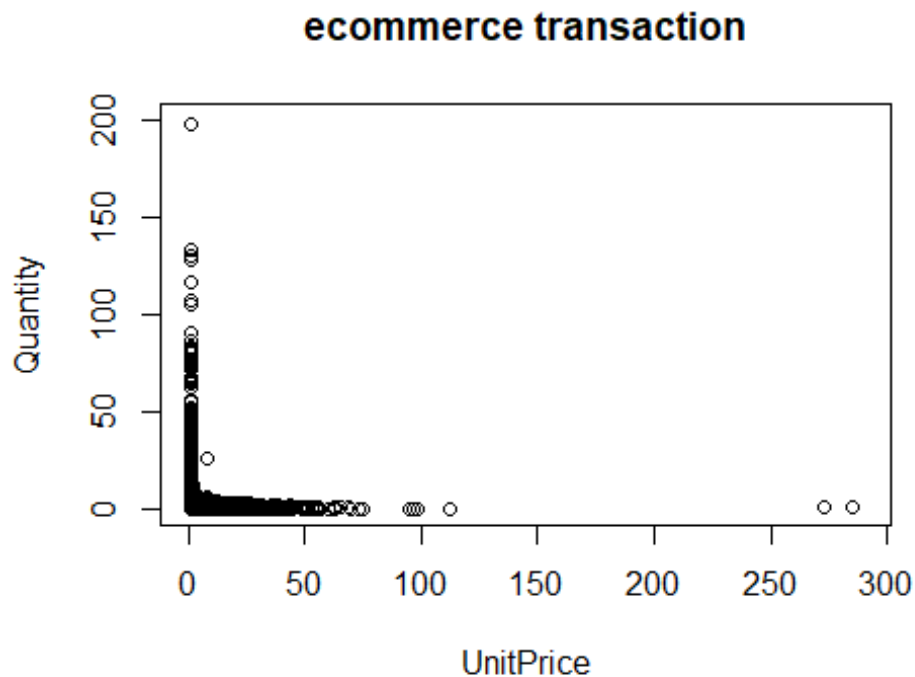
Transformation of Unitprice Variable

```
unitprice_rs = (ecommerce$UnitPrice) * 82
head (unitprice_rs)

## [1] 209.10 277.98 225.50 277.98 277.98 627.30
```

Scatterplot for Qunatity and UnitPrice

```
square_quan = sqrt(ecommerce$Quantity)
square_uni = sqrt(ecommerce$UnitPrice)
plot(x = square_quan, y = square_uni,
     xlab = "UnitPrice",
     ylab = "Quantity",
     xlim = c(0,290),
     ylim = c(0,200),
     main = "ecommerce transaction")
```

ecommerce transaction

Barplot for no of invoices produced in various countries

```
Country_tab = table(ecommerce$Country)
print(Country_tab)
```

```
##
##           Australia              Austria             Bahrain
##                1259                  401                  19
##             Belgium               Brazil              Canada
##                2069                   32                 151
##      Channel Islands              Cyprus      Czech Republic
##                 758                  622                  30
##             Denmark                 EIRE  European Community
##                 389                 8196                  61
##             Finland               France             Germany
##                 695                 8557                9495
##              Greece            Hong Kong             Iceland
##                 146                  288                 182
##              Israel                Italy               Japan
##                 297                  803                 358
##             Lebanon            Lithuania               Malta
##                  45                   35                 127
##         Netherlands               Norway              Poland
##                2371                 1086                 341
##            Portugal                  RSA        Saudi Arabia
##                1519                   58                  10
```

```
##          Singapore                Spain              Sweden
##              229                 2533                462
##        Switzerland United Arab Emirates     United Kingdom
##             2002                   68              495478
##        Unspecified                  USA
##              446                  291

barplot(Country_tab, main = "BARPLOT",
    xlab = "Country",
    ylab = "Frequency", ylim= c(0, 10000))
```