Fundamentals of Machine Learning Assignment

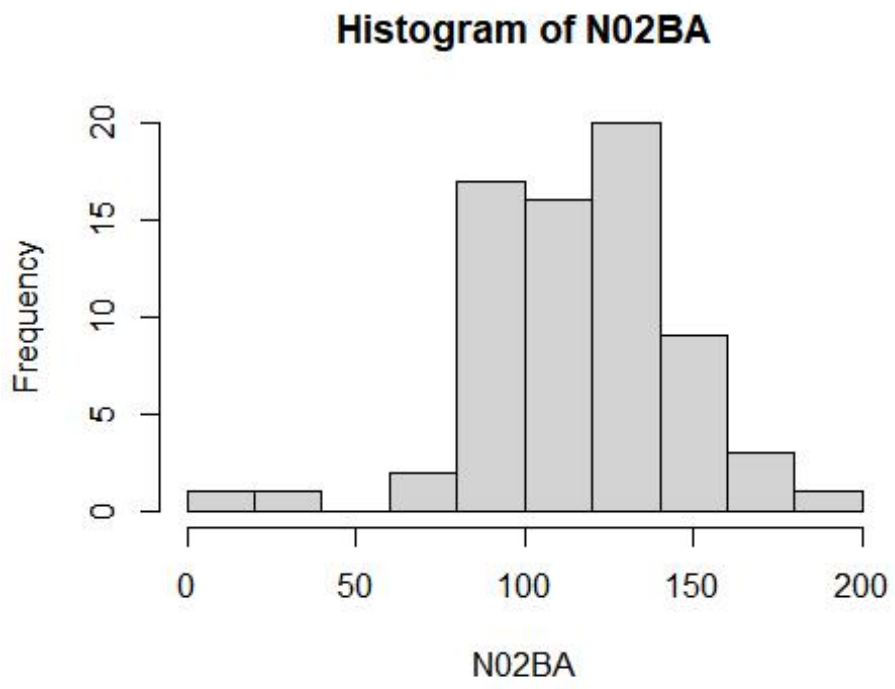
Supriya Mattapelly

2023-09-10

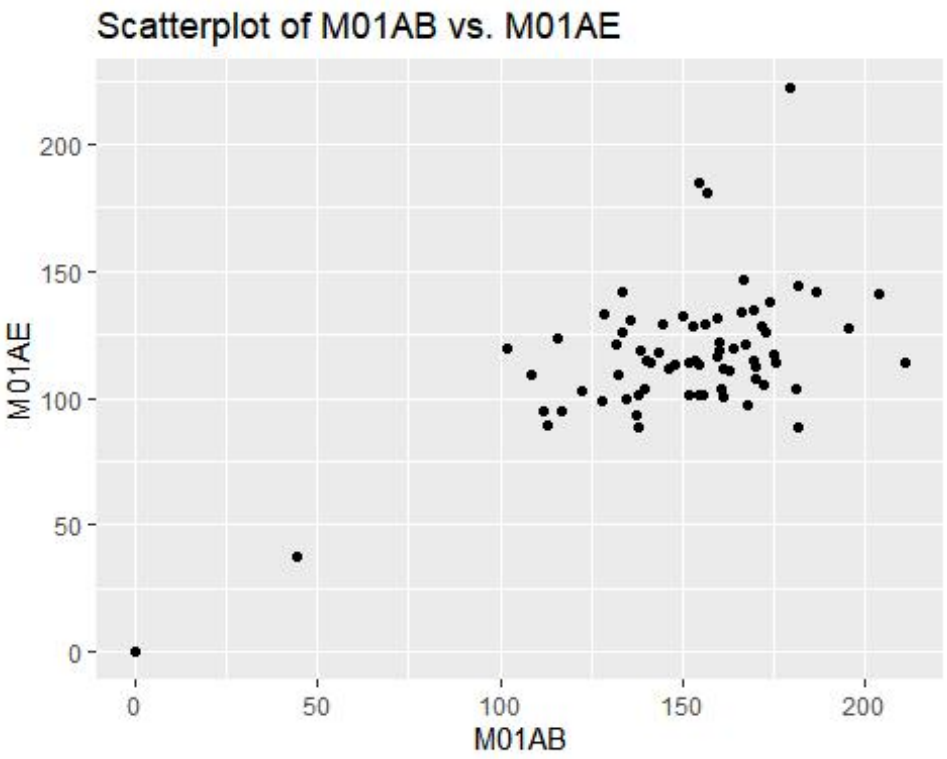
```
knitr::opts_chunk$set(echo = TRUE)

library(readr)
salesmonthly <- read_csv("C:/Users/Supriya mattapelly.000/Downloads/sal
esmonthly.csv")

## Rows: 70 Columns: 9
## ── Column specification ────────────────────────────────────────────
────────────
## Delimiter: ","
## dbl  (8): M01AB, M01AE, N02BA, N02BE, N05B, N05C, R03, R06
## date (1): datum
##
## ℹ Use `spec()` to retrieve the full column specification for this d
ata.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet
 this message.

summary(salesmonthly)

##      datum                M01AB           M01AE           N02BA
##  Min.   :2014-01-31   Min.   :  0.0   Min.   :  0.0   Min.   :  0.00
##  1st Qu.:2015-07-07   1st Qu.:137.5   1st Qu.:103.5   1st Qu.: 94.38
##  Median :2016-12-15   Median :154.6   Median :114.8   Median :117.22
##  Mean   :2016-12-14   Mean   :150.0   Mean   :116.5   Mean   :115.02
##  3rd Qu.:2018-05-23   3rd Qu.:169.0   3rd Qu.:128.4   3rd Qu.:133.84
##  Max.   :2019-10-31   Max.   :211.1   Max.   :222.4   Max.   :191.60
##      N02BE             N05B            N05C             R03
##  Min.   :   0.0   Min.   :  1.0   Min.   : 0.00   Min.   :  0.0
##  1st Qu.: 648.2   1st Qu.:223.8   1st Qu.:12.00   1st Qu.:112.0
##  Median : 865.8   Median :250.3   Median :18.00   Median :160.0
##  Mean   : 892.5   Mean   :262.1   Mean   :17.84   Mean   :167.7
##  3rd Qu.:1061.6   3rd Qu.:293.6   3rd Qu.:23.00   3rd Qu.:218.2
```

```
##  Max.   :1856.8   Max.   :492.0   Max.   :50.00   Max.   :386.0
##       R06
##  Min.   :  0.00
##  1st Qu.: 49.88
##  Median : 74.10
##  Mean   : 86.66
##  3rd Qu.:119.81
##  Max.   :213.04
```

```
# Transforming a variable

salesmonthly$M01AB_squared <- salesmonthly$M01AB^2

#histogram plot of N02BA variable

hist(salesmonthly$N02BA,
     main = "Histogram of N02BA",
     xlab = "N02BA",
     ylab = "Frequency")
```




```
#scatterplot

library(ggplot2)


ggplot(salesmonthly, aes(x = M01AB, y = M01AE)) + geom_point() + labs(t
itle = "Scatterplot of M01AB vs. M01AE", x = "M01AB", y = "M01AE")
```




```
mean_N02BA <- mean(salesmonthly$N02BA)

median_N02BA <- median(salesmonthly$N02BA)

mode_N02BA <- as.numeric(names(sort(table(salesmonthly$N02BA), decreasi
ng = TRUE)[1]))


mean_N02BA

## [1] 115.0208

median_N02BA

## [1] 117.225

mode_N02BA

## [1] 122.1
```