STAT253 - Homework#1

Osman MANTICI - 150117505

29 Mart 2019

1-) Systolic bp, and diastolic bp and age variables are multivariate, discrete and quantitative. The sex variable is qualitative, continuous, and bivariate.

2-) Dotplot, histogram, boxplot, stem-and-leaf plot, can be used when this data set read.

When the questions are want to learn about the range, median and mean, maximum and minimum values, where is the collapses or outliers, answers are found easily. Furthermore, boxplot gives us about the five number summary rather than the others. Stem-and-leaf plot and dotplot also gives more the exact information.

3-)

a. Means of men and women

Results are in that order, men systolic blood pressure mean, women systolic blood pressure mean, men diastolic blood pressure mean, women diastolic blood pressure mean.

```
## [1] 117.72
```

```
## [1] 110.21
```

```
## [1] 75.47
```

```
## [1] 70.32
```

b. Sample variance of men and women

Results are in that order men systolic blood pressure sample variance women systolic blood pressure sample variance men diastolic blood pressure sample variance women diastolic blood pressure sample variance

```
## [1] 175.9208
```

```
## [1] 150.7938
```

```
## [1] 105.9082
```

```
## [1] 75.04808
```

c. Standard deviation of men and women

Results are in that order

men systolic blood pressure standard deviation women systolic blood pressure standard deviation men diastolic blood pressure standard deviation women diastolic blood pressure standard deviation

```
## [1] 13.26351
```

```
## [1] 12.27981
```

```
## [1] 10.29117
```

```
## [1] 8.66303
```

d. 1. Lower quartiles of men and women

Results are in that order

men systolic blood pressure lower quartiles women systolic blood pressure lower quartiles men diastolic blood pressure lower quartiles women diastolic blood pressure lower quartiles

```
## 25%
## 110
```

```
##   25%
## 101.5
```

```
## 25%
##  70
```

```
##  25%
## 65.5
```

d. 2. Upper quartiles of men and women

Results are in that order men systolic blood pressure upper quartiles women systolic blood pressure upper quartiles men diastolic blood pressure upper quartiles women diastolic blood pressure upper quartiles

```
## 75%
## 126
```

```
## 75%
## 118
```

```
## 75%
##  80
```

```
## 75%
##  78
```

e. 1. Maximum values of men and women

Results are in that order men systolic blood pressure maximum women systolic blood pressure maximum men diastolic blood pressure maximum women diastolic blood pressure maximum

```
## [1] 148
```

```
## [1] 158
```

```
## [1] 100
```

```
## [1] 96
```

e. 2. Minimum values of men and women

Results are in that order men systolic blood pressure minimum women systolic blood pressure minimum men diastolic blood pressure minimum women diastolic blood pressure minimum

```
## [1] 148
```

```
## [1] 86
```

```
## [1] 56
```

```
## [1] 40
```

f. Range values of men and women

Results are in that order men systolic blood pressure range women systolic blood pressure range men diastolic blood pressure range women diastolic blood pressure range

```
## [1] 78
```

```
## [1] 72
```

```
## [1] 44
```

```
## [1] 56
```

g. Comparing range and standart deviations, Range values how many standard deviation of men and women

Results are in that order men systolic blood pressure number of standart deviations women systolic blood pressure number of standart deviations men diastolic blood pressure number of standart deviations women diastolic blood pressure number of standart deviations

```
## [1] 5.880794
```

```
## [1] 5.863281
```

```
## [1] 4.27551
```

```
## [1] 6.464251
```

h. Median values of men and women

Results are in that order men systolic blood pressure median women systolic blood pressure median men diastolic blood pressure median women diastolic pressure median

```
## [1] 120
```

```
## [1] 110
```

```
## [1] 76
```

```
## [1] 70
```

i. IQR's of men and women

Results are in that order men systolic blood pressure IQR's women systolic blood pressure IQR's men diastolic blood pressure IQR's women diastolic blood pressure IQR's

```
## [1] 16
```

```
## [1] 16.5
```

```
## [1] 10
```

```
## [1] 12.5
```

j. 5-number-summary of men and women

Results are in that order men systolic blood pressure 5-number-summary women systolic blood pressure 5-number-summary men diastolic blood pressure 5-number-summary women diastolic blood pressure 5-number-summary

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    70.0   110.0   120.0   117.7   126.0   148.0
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    86.0   101.5   110.0   110.2   118.0   158.0
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   56.00   70.00   76.00   75.47   80.00  100.00
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   40.00   65.50   70.00   70.32   78.00   96.00
```

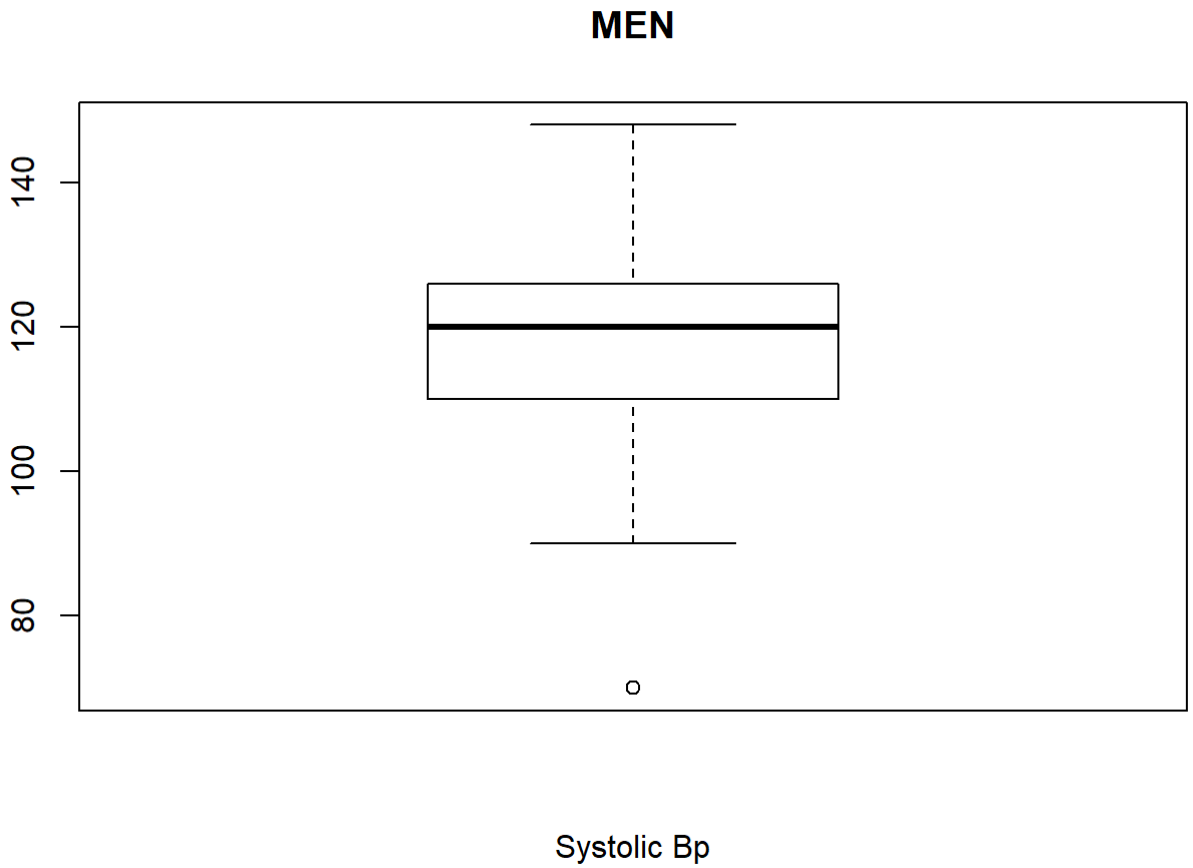
k. Boxplot of men and women

Results are in that order men systolic blood pressure boxplot # No outliers and the median of this data seen as about 120, minimum and maximum value seen about 90 and 148 respectively, and 1. quartile is about 110, 3. quartile is about 127.

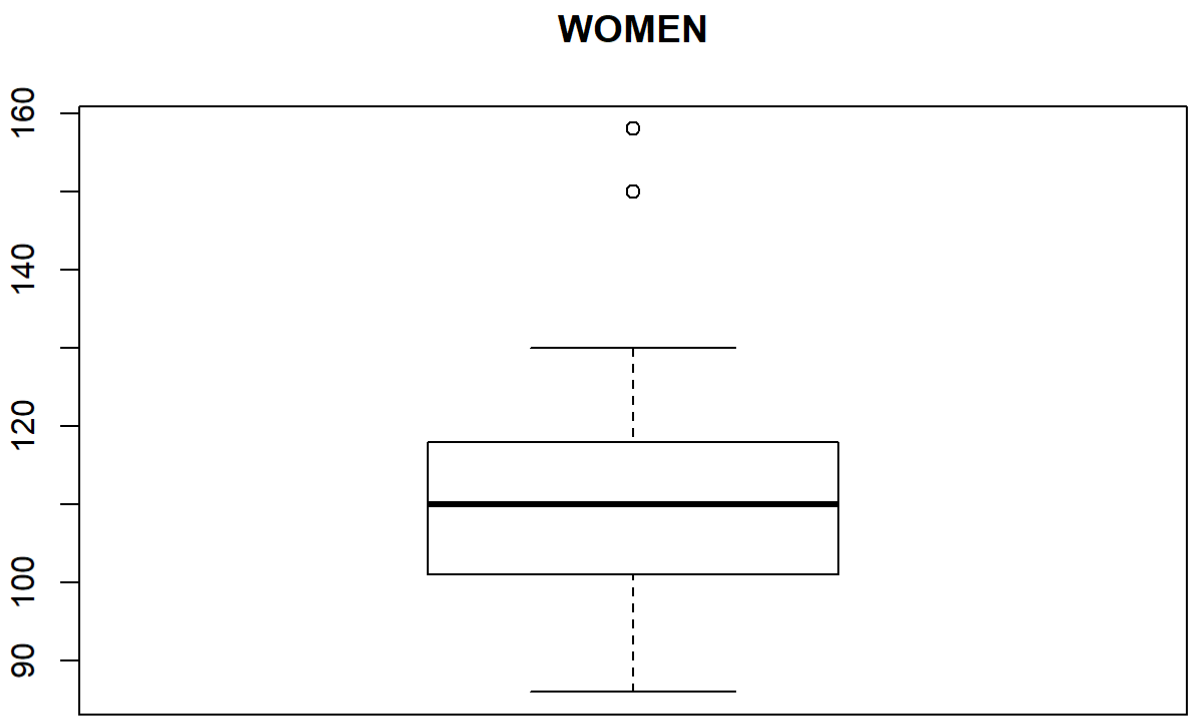
women systolic blood pressure boxplot #There are outliers and median of this data seen as about 110, minimum and maximum value seen about 88 and 130 respectively, and 1. quartile is about 100, 3. quartile is about 120. Outliers are about 150 and 160. men diastolic blood pressure boxplot #There are outliers and median of this data seen as about 77, minimum and maximum value seen about 57 and 95 respectively, and 1.

quartile is about 70, 3. quartile is about 80. Outliers are about 97 and 100. women diastolic blood pressure boxplot #There are outliers and median of this data seen as about 70, minimum and maximum value seen about 55 and 98 respectively, and 1. quartile is about 53, 3. quartile is about 78. Outlier is about 40.

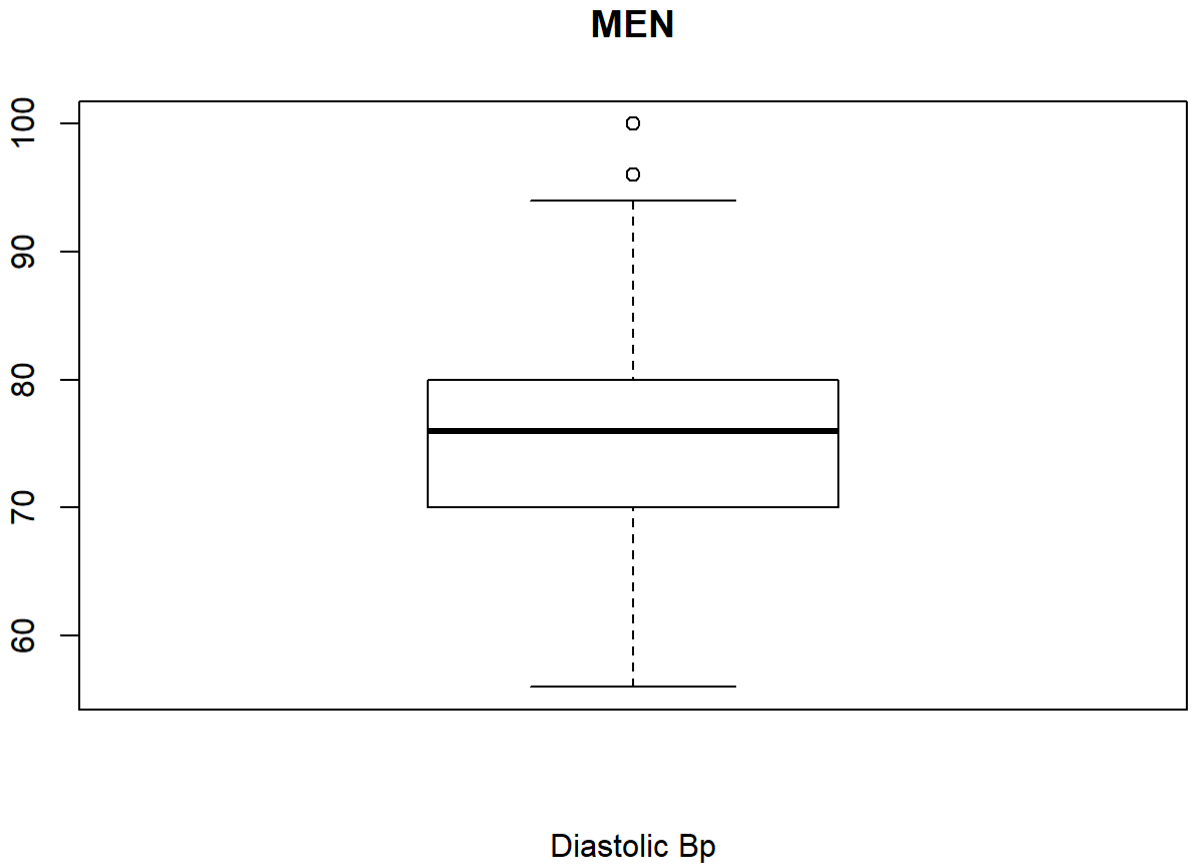
```
boxplot(SYSBP[GENDER==0],main="MEN", xlab="Systolic Bp", horizontal = FALSE)
```



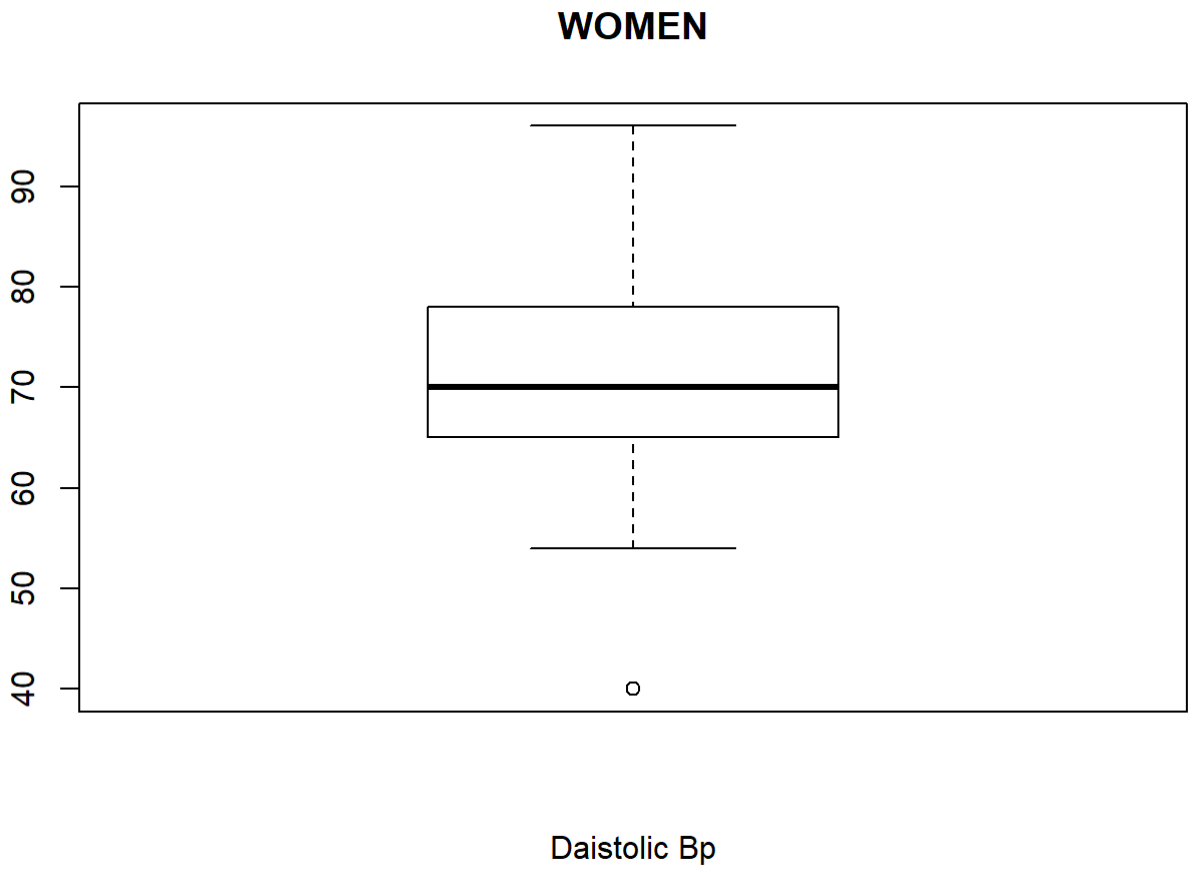

```
boxplot(SYSBP[GENDER==1],main="WOMEN", xlab="Systolic Bp", horizontal = FALSE)
```




```
boxplot(DIASBP[GENDER==0],main="MEN", xlab="Diastolic Bp", horizontal = FALSE)
```



```
boxplot(DIASBP[GENDER==1],main="WOMEN",  xlab="Daistolic Bp", horizontal = FALSE)
```



I. Stem-and-leaf plot of men and women

Results are in that order men systolic blood pressure Stem-and-leaf plot women systolic blood pressure Stem-and-leaf plot men diastolic blood pressure Stem-and-leaf plot women diastolic blood pressure Stem-and-leaf plot

```
stem(SYSBP[GENDER==0])
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##    7 | 0
##    8 |
##    9 | 000488
##   10 | 00000235556888
##   11 | 000000000224444555667888
##   12 | 00000000000000000222224456666
##   13 | 000000000000002257
##   14 | 022458
```

```
stem(SYSBP[GENDER==1])
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##    8 | 668
##    9 | 00002446888
##   10 | 000000000002222444466688889
##   11 | 00000000000000000222444444444468888
##   12 | 00000000002446688
##   13 | 0000
##   14 |
##   15 | 08
```

```
stem(DIASBP[GENDER==0])
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##    5 | 6688
##    6 | 00000000234
##    6 | 55588
##    7 | 00000000000000000022222224444
##    7 | 566666666688888888
##    8 | 000000000000022
##    8 | 6668
##    9 | 0000000444
##    9 | 6666
##   10 | 0
```

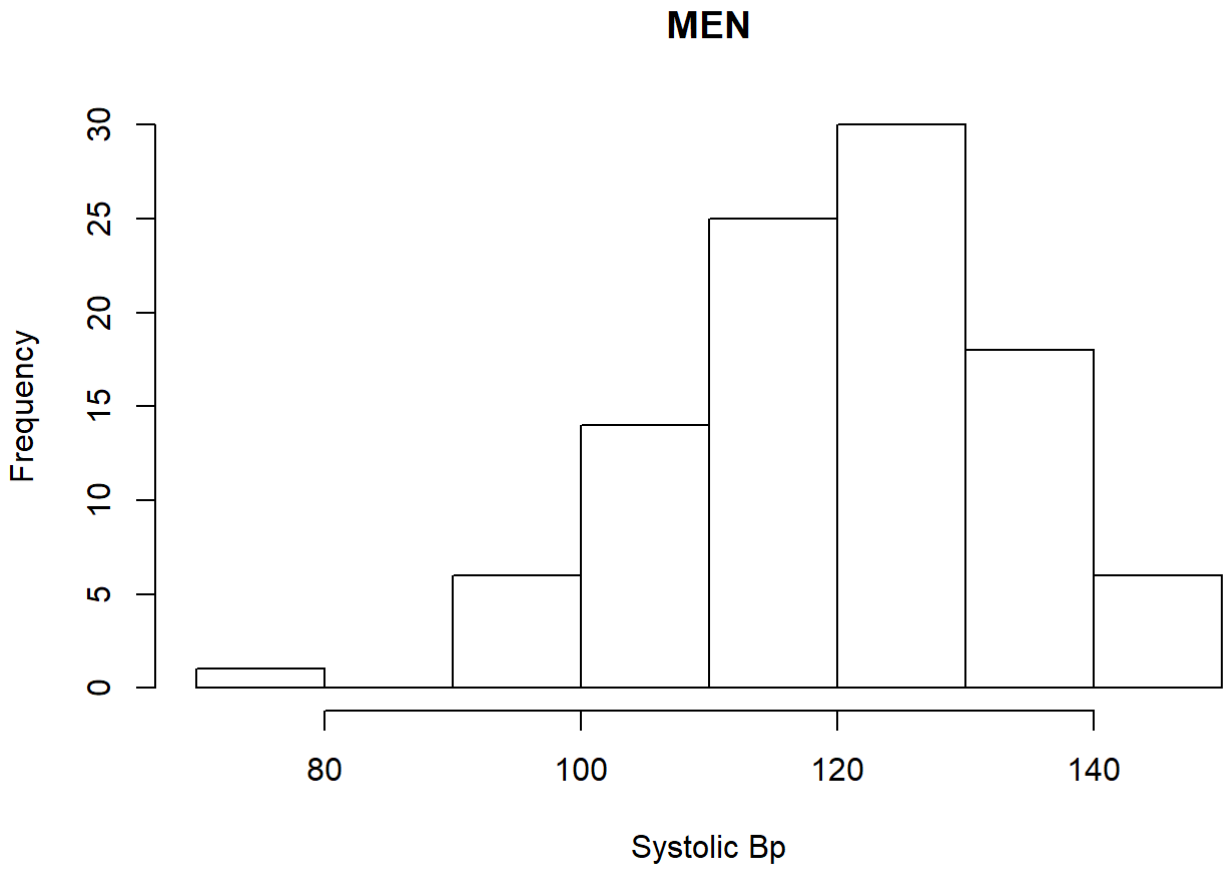
```
stem(DIASBP[GENDER==1])
```

```
##
##   The decimal point is 1 digit(s) to the right of the |
##
##   4 | 0
##   5 | 48888
##   6 | 00000000000000222446666888888
##   7 | 00000000000000000000000000000002446668888888
##   8 | 0000000000000046668
##   9 | 6
```

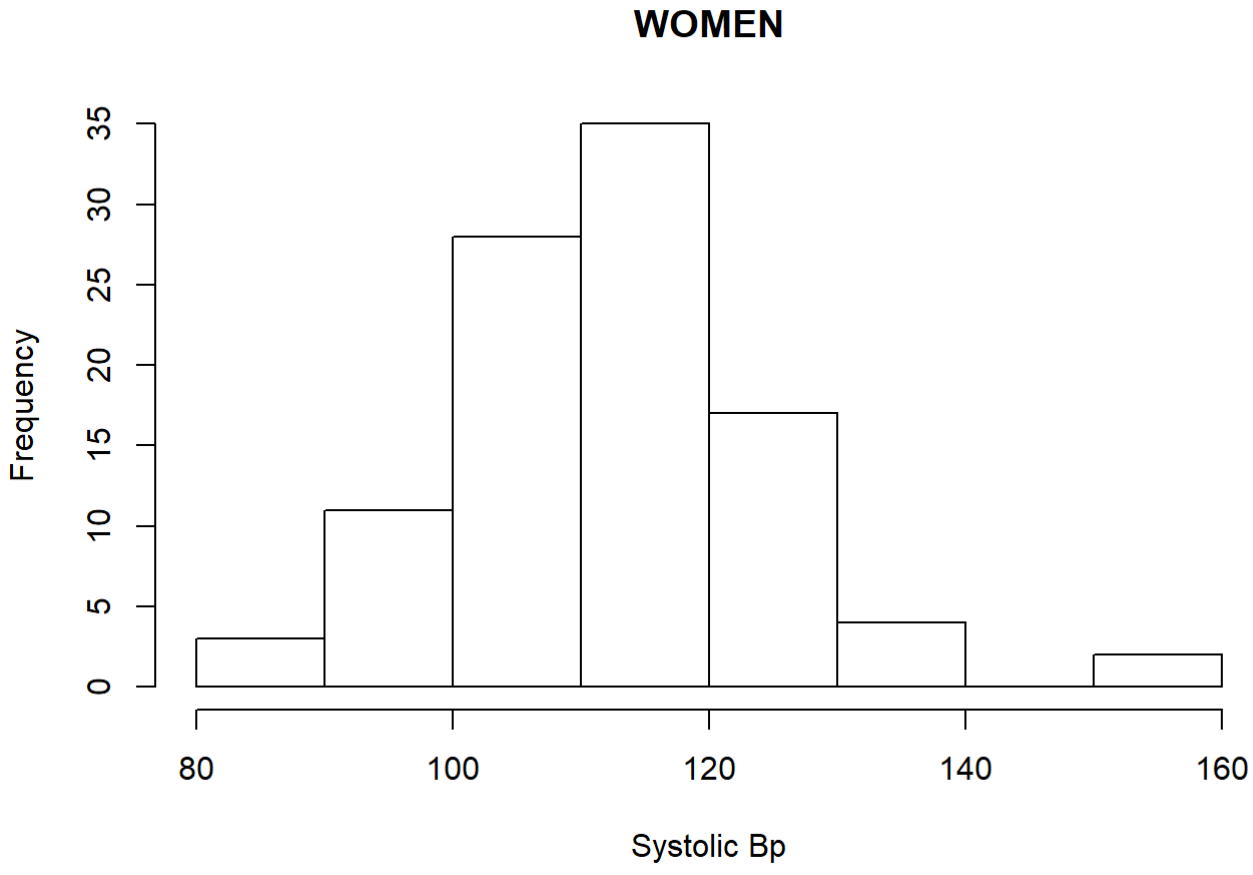
m. Histograms of men and women

Results are in that order men systolic blood pressure histogram women systolic blood pressure histogram men diastolic blood pressure histogram women diastolic blood pressure histogram

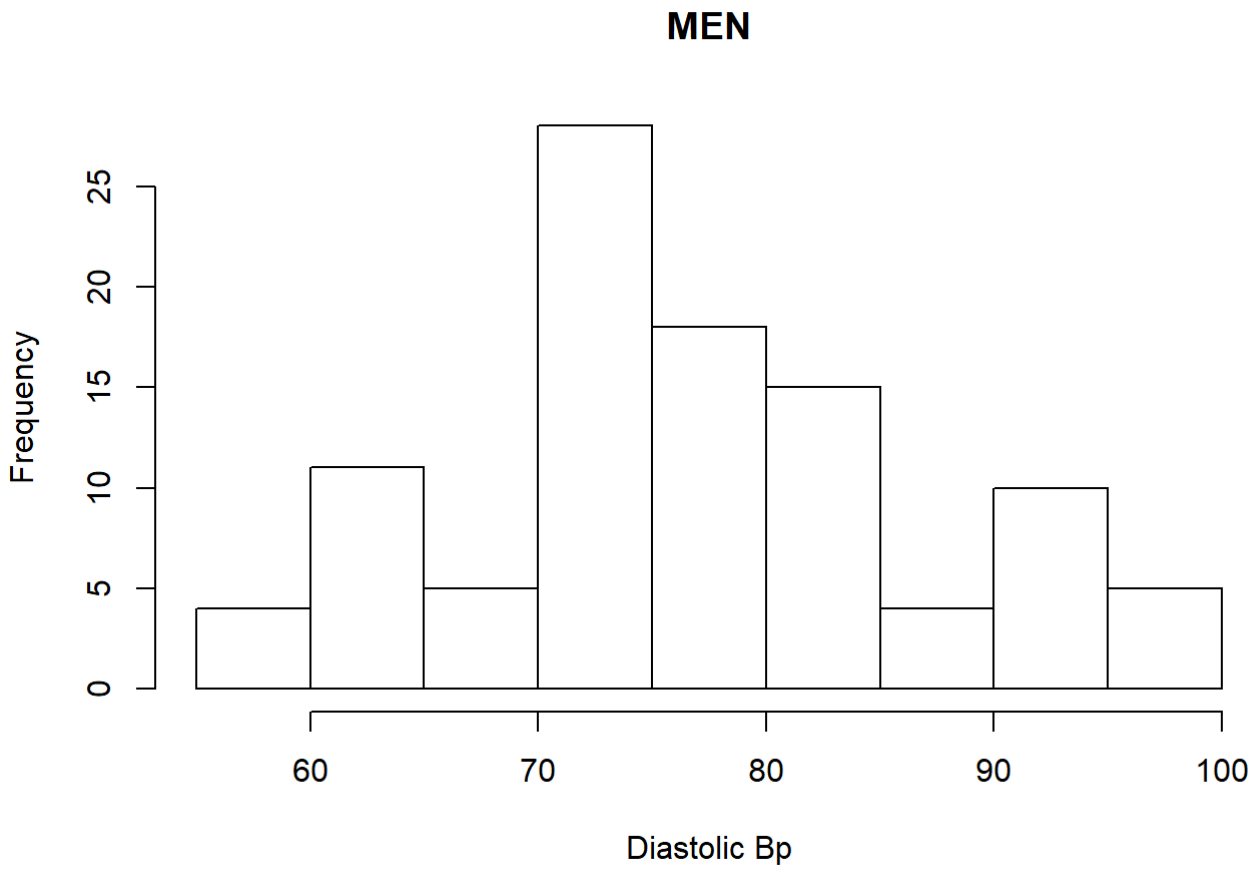
```
hist(SYSBP[GENDER==0], right = FALSE, main = "MEN", xlab = "Systolic Bp")
```



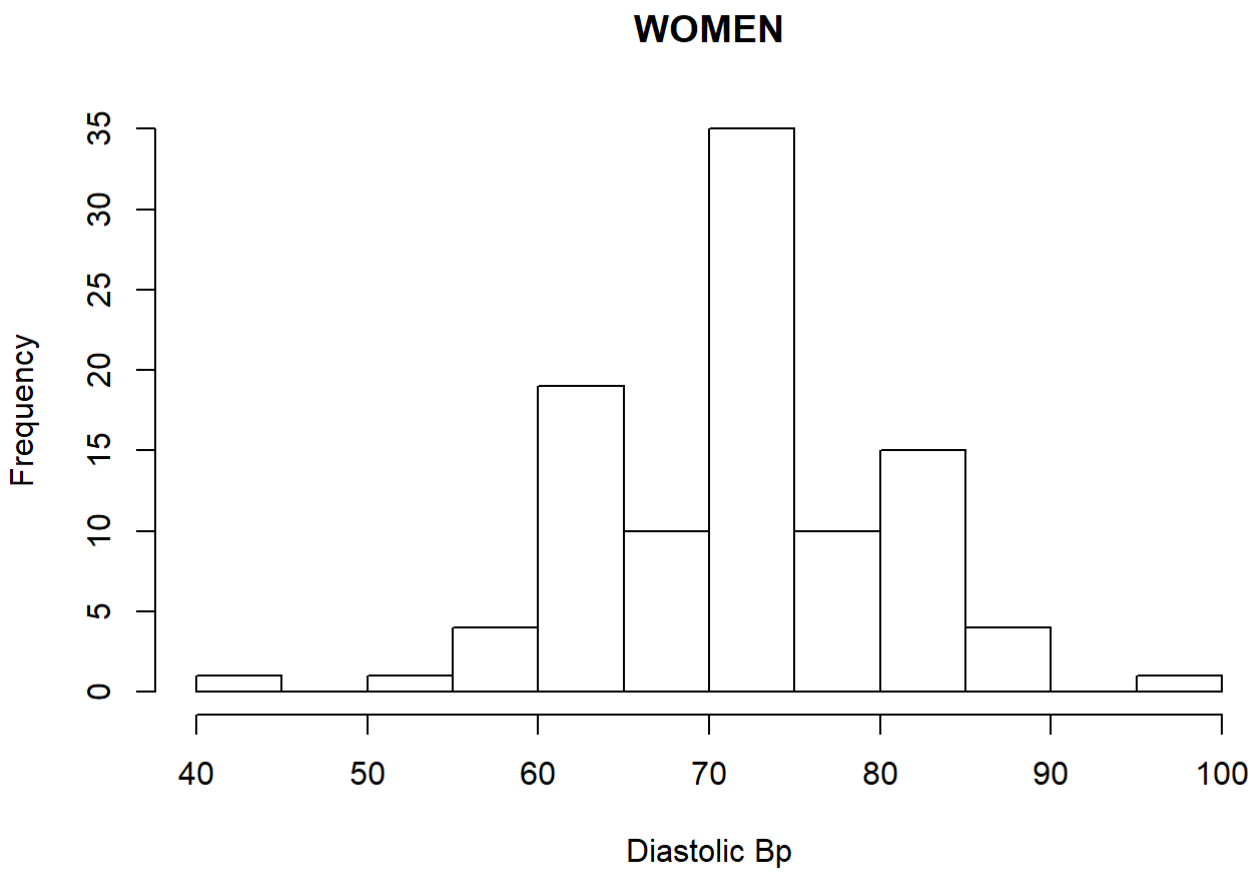

```
hist(SYSBP[GENDER==1], right = FALSE, main = "WOMEN", xlab = "Systolic Bp")
```



```
hist(DIASBP[GENDER==0], right = FALSE, main = "MEN", xlab = "Diastolic Bp")
```



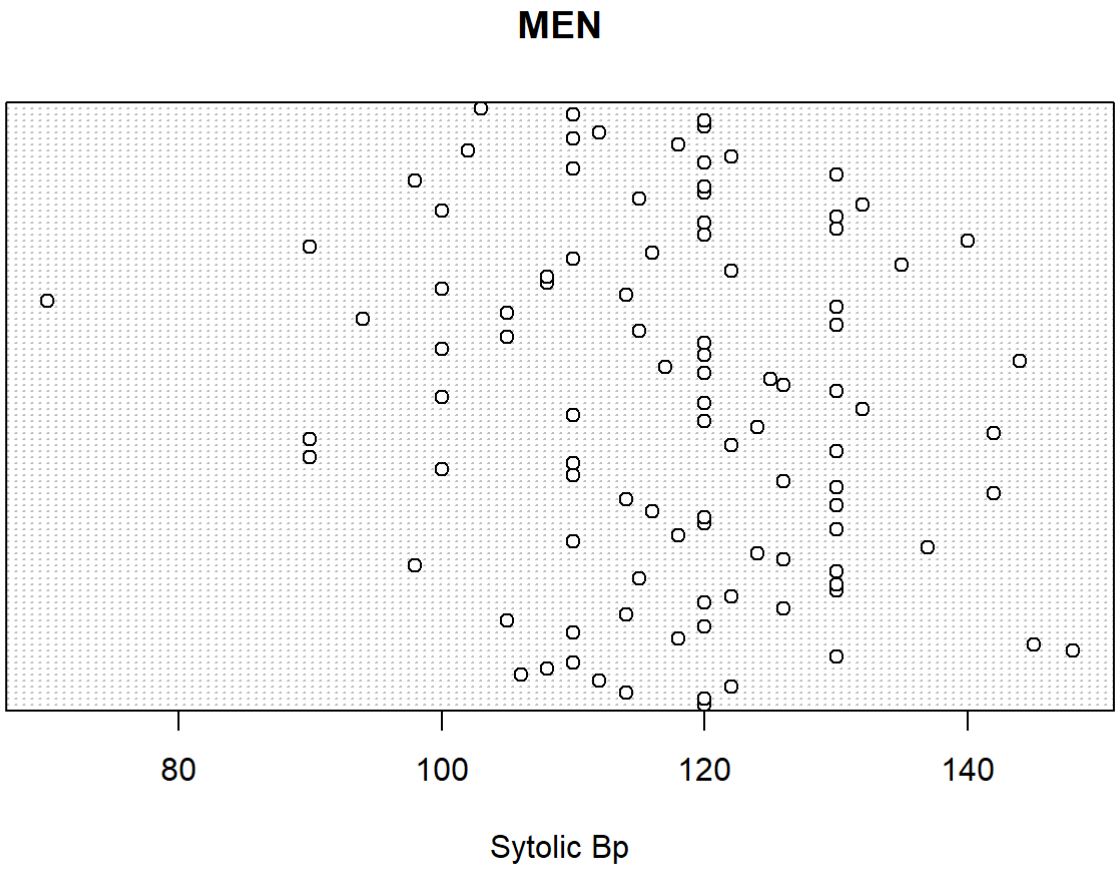

```
hist(DIASBP[GENDER==1], right = FALSE, main = "WOMEN", xlab = "Diastolic Bp")
```



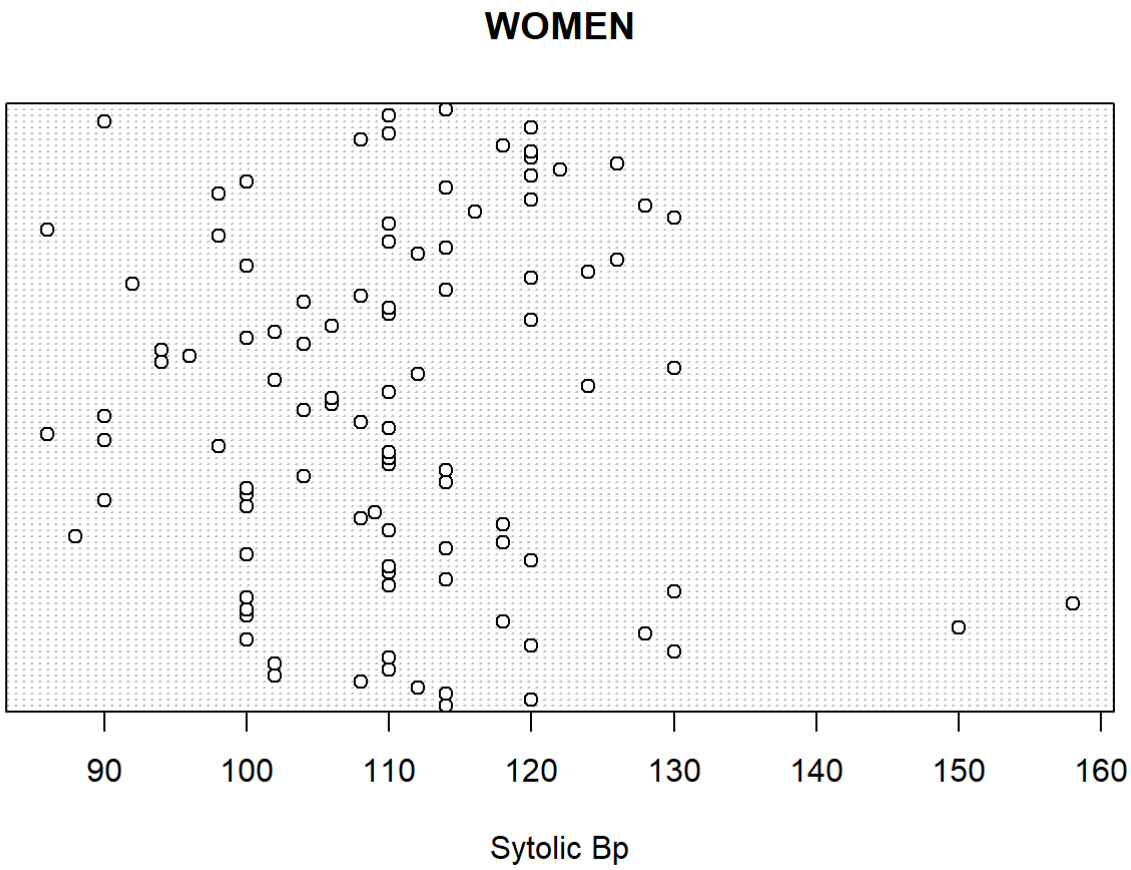
m. Dotplots of men and women

Results are in that order men systolic blood pressure dotplot women systolic blood pressure dotplot men diastolic blood pressure dotplot women diastolic blood pressure dotplot

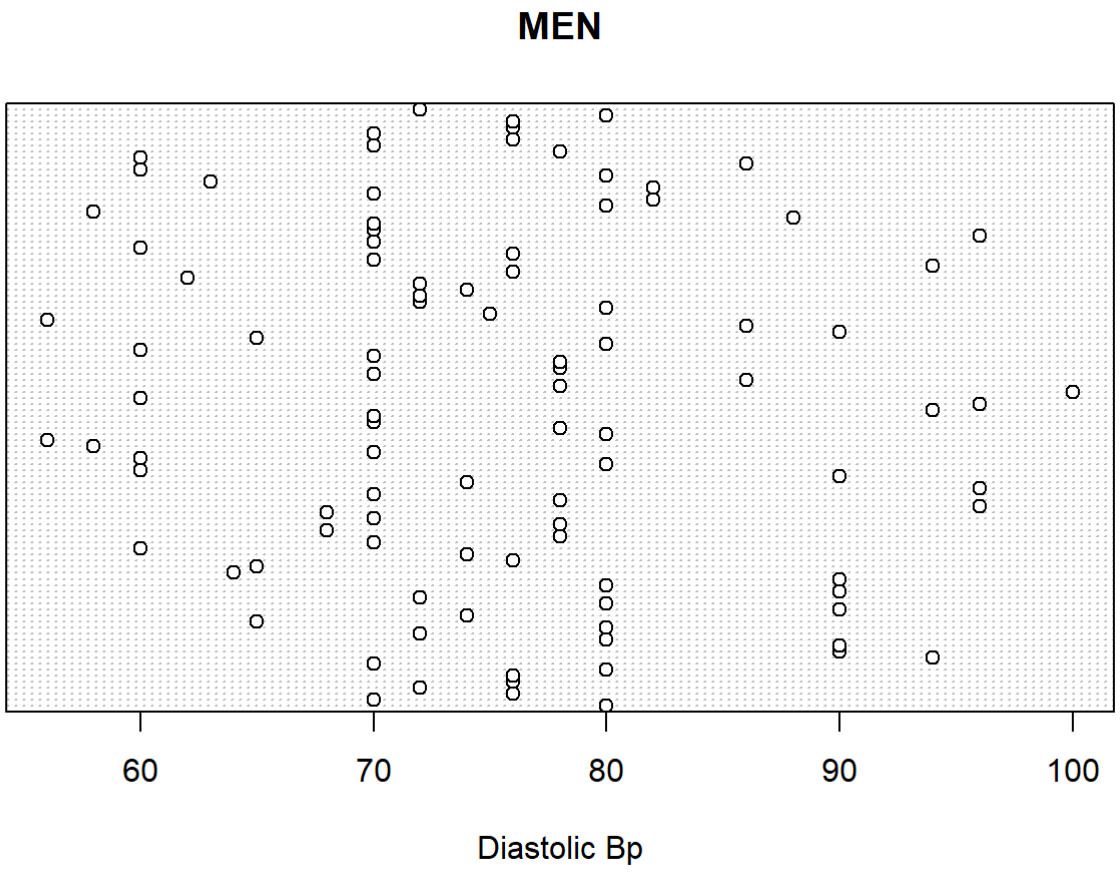
```
dotchart(SYSBP[GENDER==0], main = "MEN", xlab = "Sytolic Bp")
```



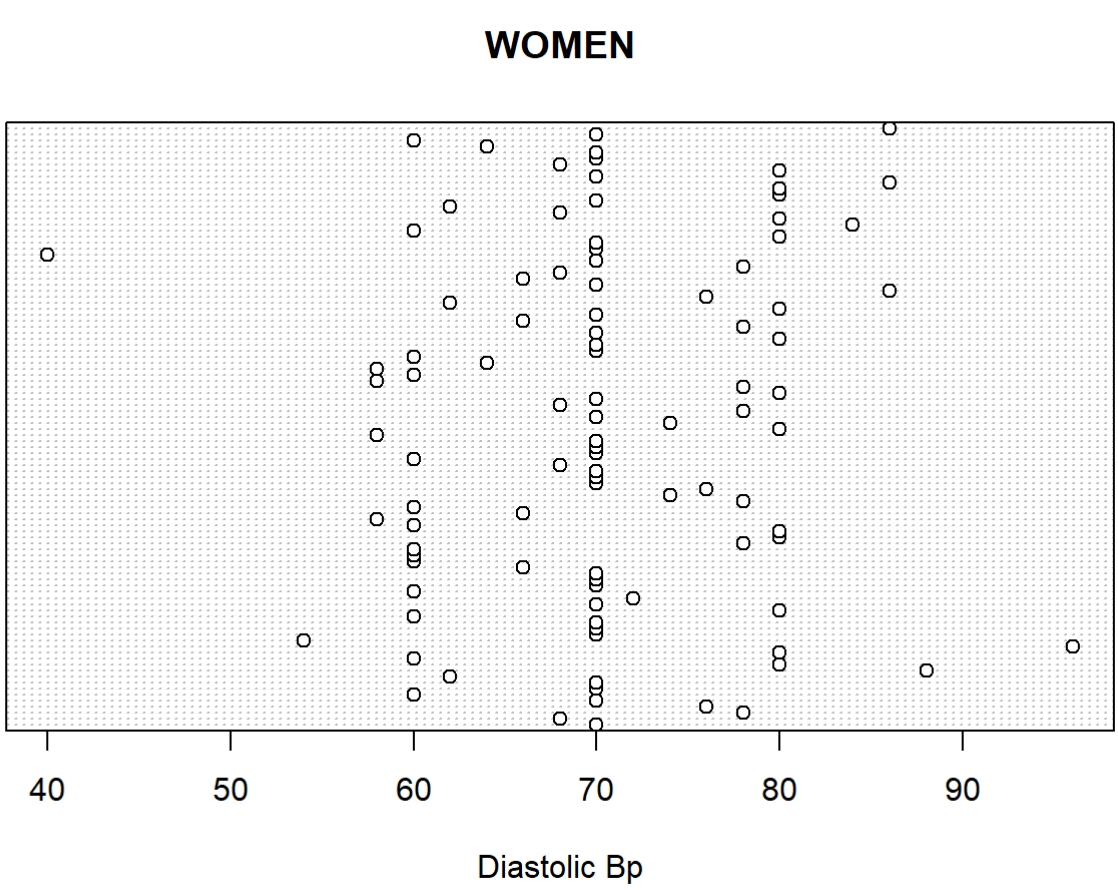

```
dotchart(SYSBP[GENDER==1], main = "WOMEN", xlab = "Sytolic Bp")
```



```
dotchart(DIASBP[GENDER==0], main = "MEN", xlab = "Diastolic Bp")
```




```
dotchart(DIASBP[GENDER==1], main = "WOMEN", xlab = "Diastolic Bp")
```



o. Tchebysheff's Theorem

Yes, we can use the Tchebysheff's Theorem because of the usability any data set without a necessity like mound shaped distibution or any kind. So the Tchebysheff's Theorem also employable on this data set.

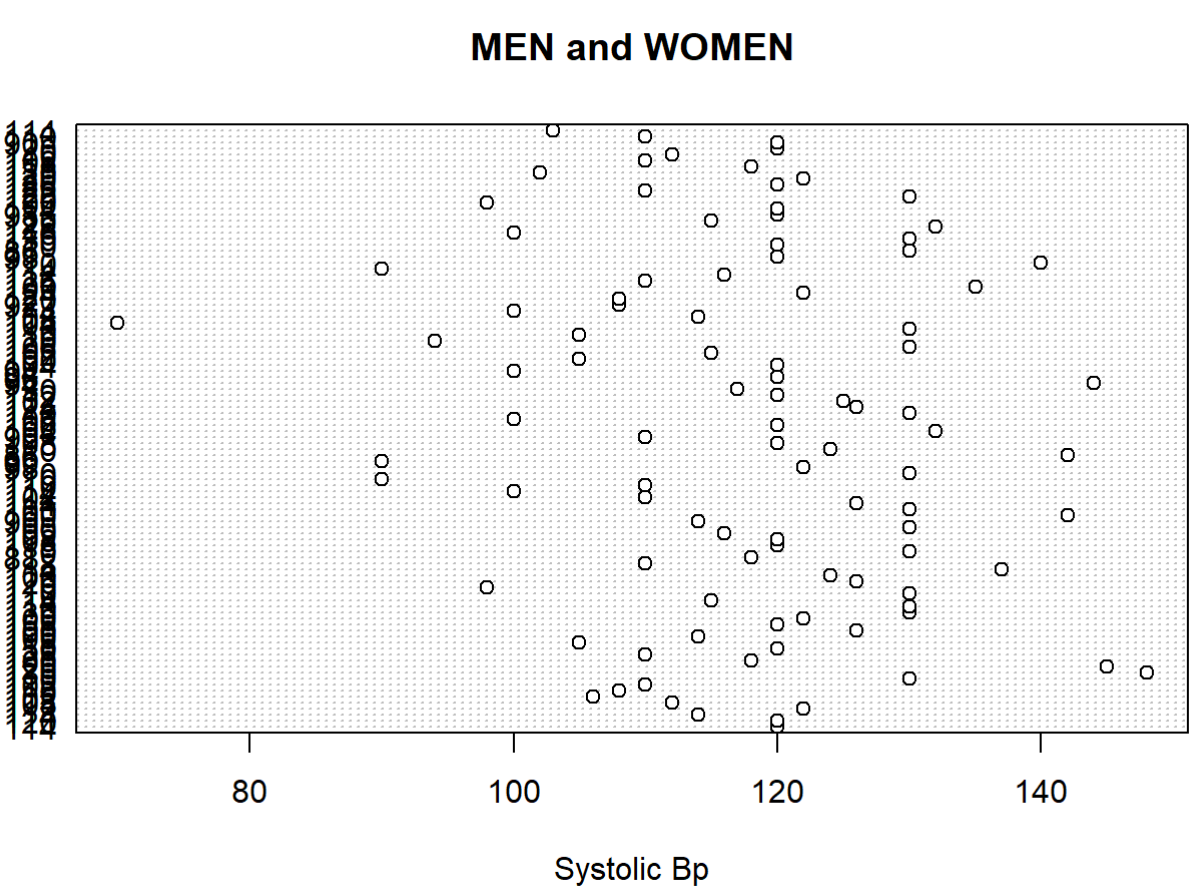
p. Empirical Rule

The Empirical Rule just applicable on mound shaped data set. If we look this dataset, we realize that diastolic blood pressure of men and women are mound shaped, so emprical rule can be apply our data set.

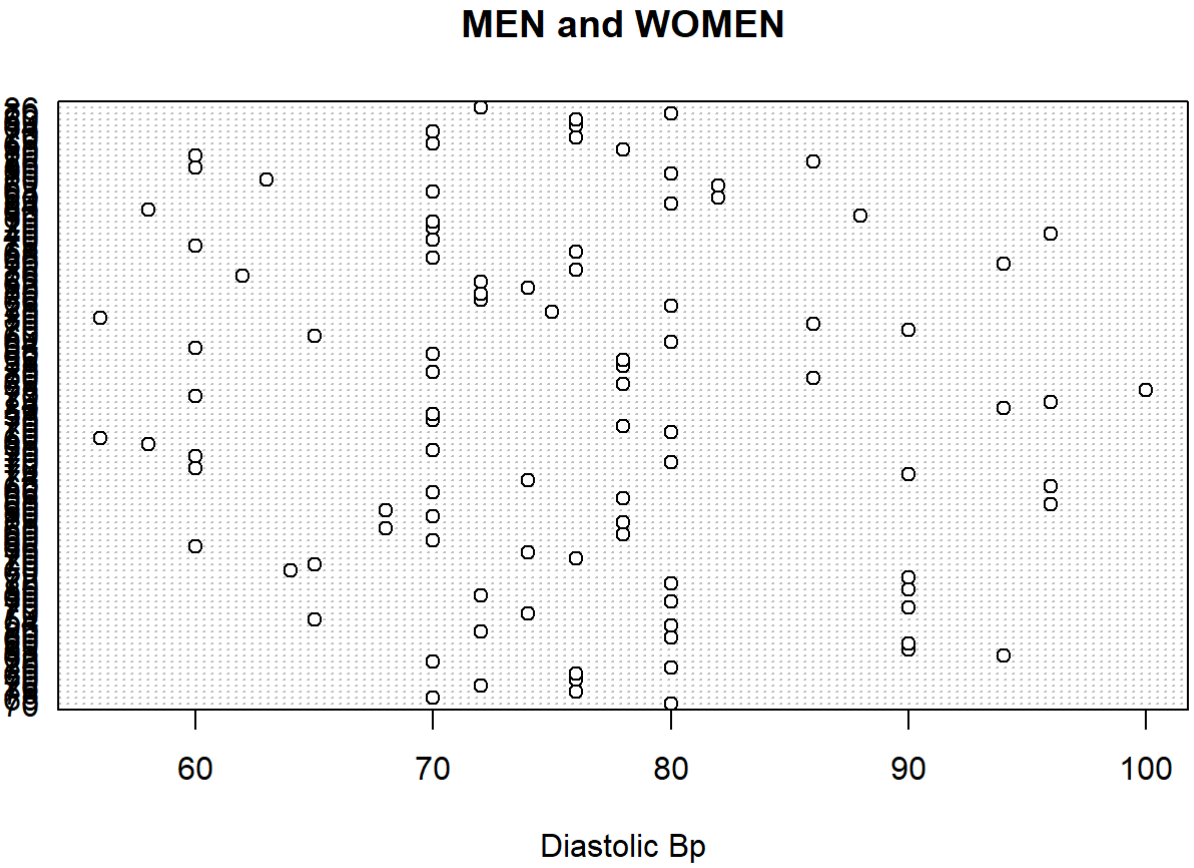
q. Comperative Dotplots of men and women

Results are in that order men and women systolic blood pressure men and women diastolic blood pressure

```
dotchart(SYSBP[GENDER==0],SYSBP[GENDER==1], main = "MEN and WOMEN", xlab = "Systolic Bp")
```



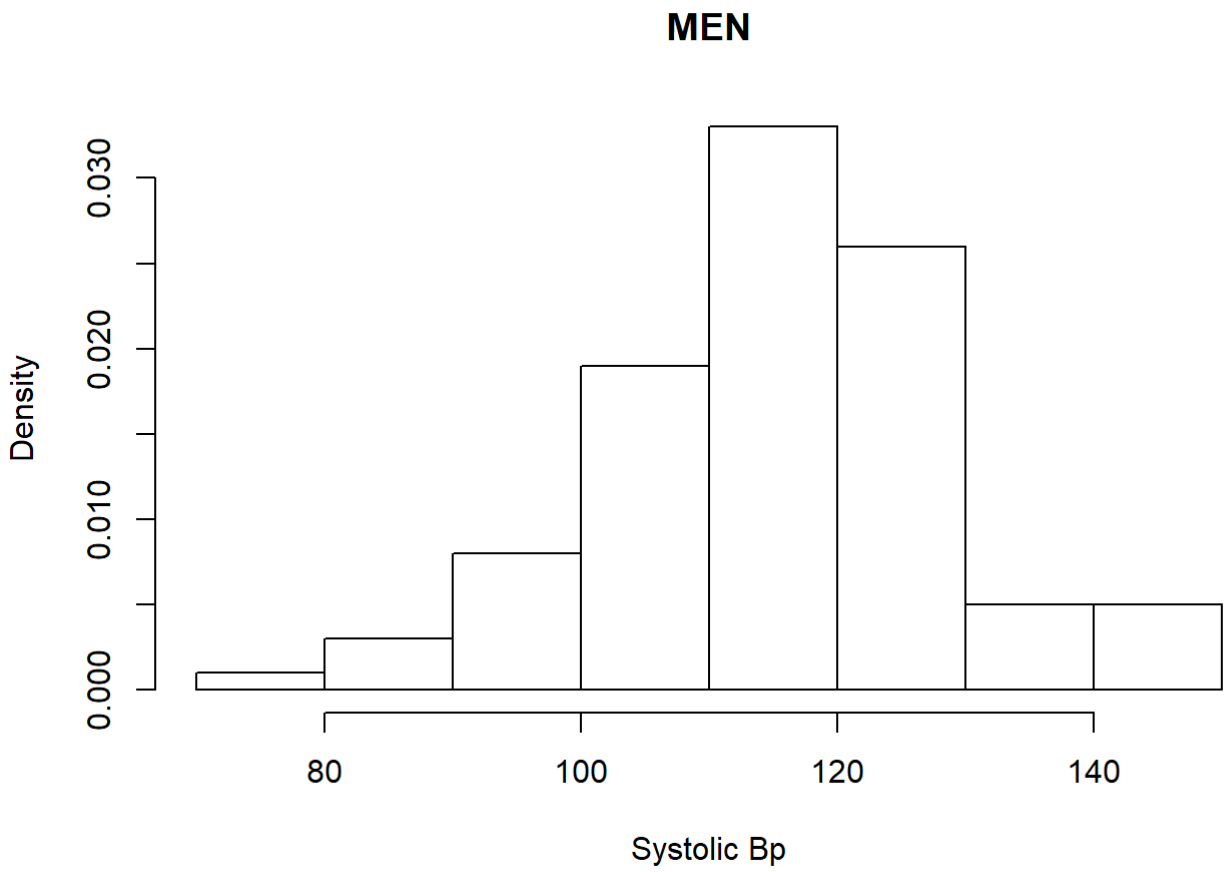

```
dotchart(DIASBP[GENDER==0],DIASBP[GENDER==1], main = "MEN and WOMEN", xlab = "Diastolic Bp")
```



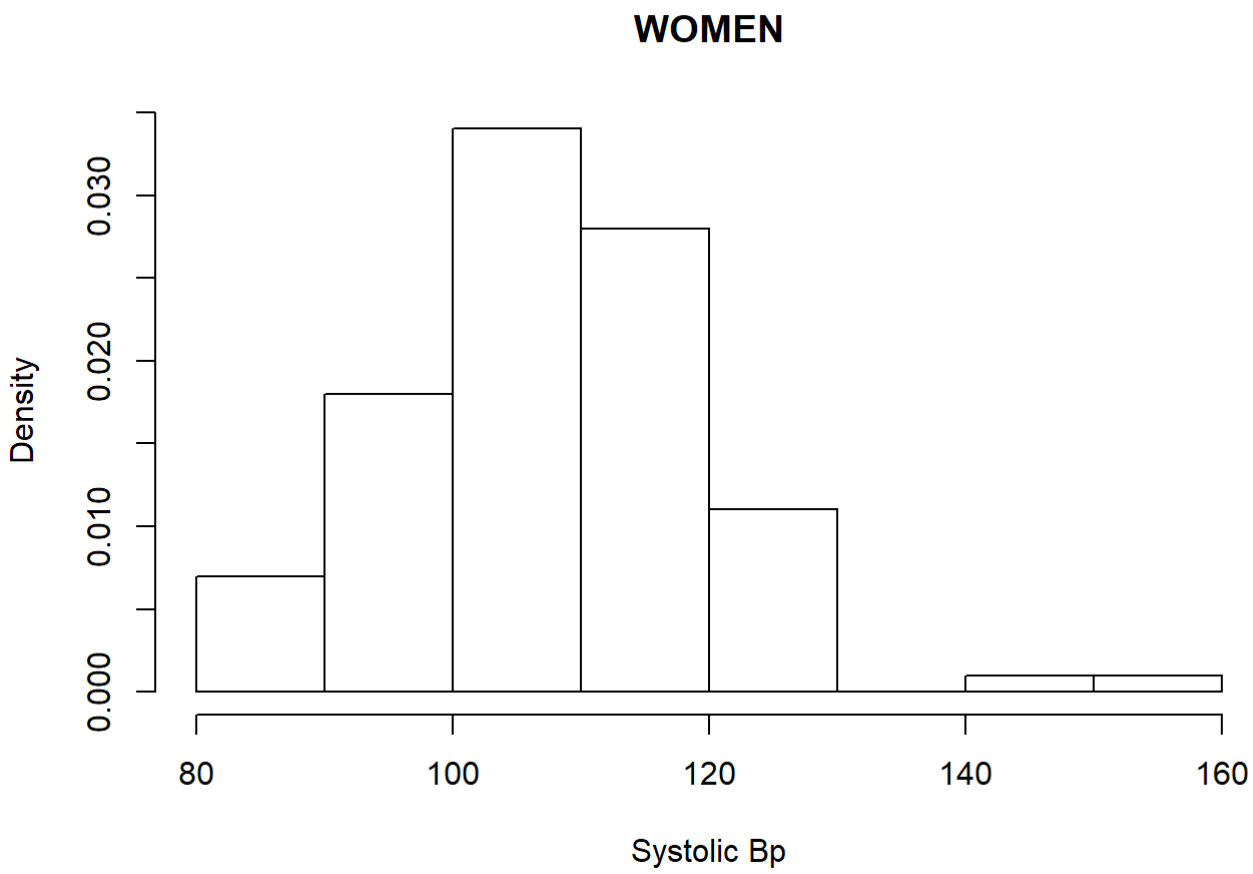
r. Relative frequency histograms of men and women

Results are in that order men systolic blood pressure relative freq. histogram women systolic blood pressure relative freq. histogram men diastolic blood pressure relative freq. histogram women diastolic blood pressure relative freq. histogram

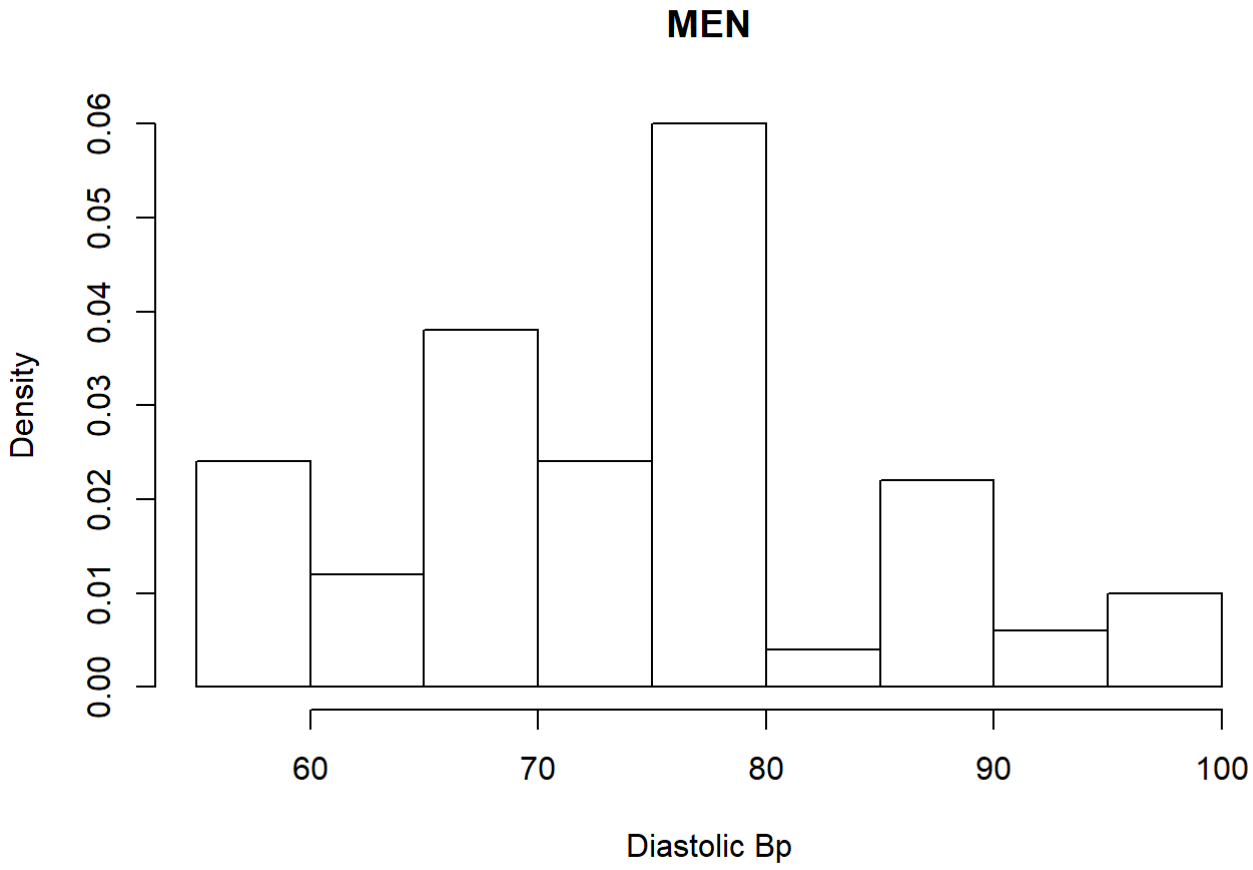
```
# Skewed left
hist(SYSBP[GENDER==0], prob=TRUE, main = "MEN", xlab = "Systolic Bp")
```



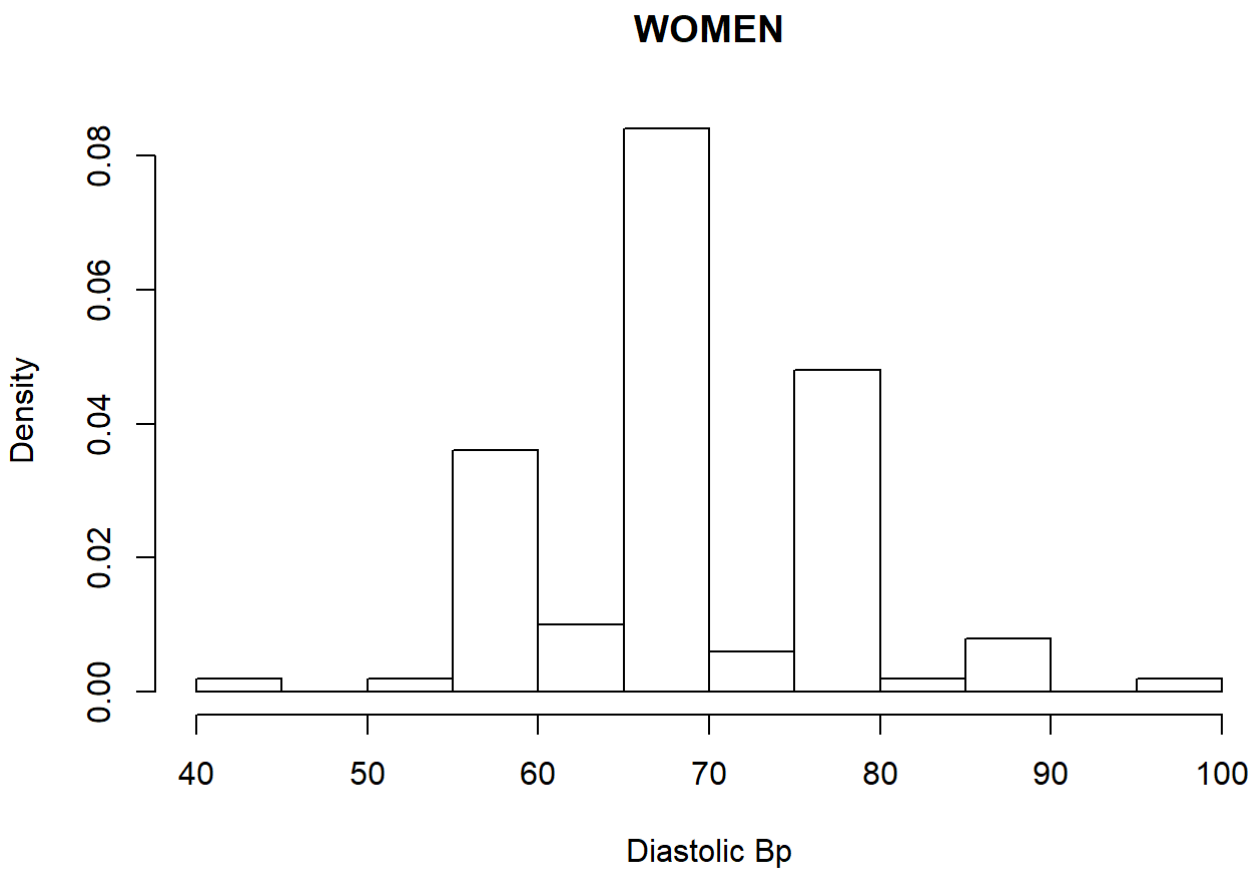

```
# Skewed right
hist(SYSBP[GENDER==1], prob=TRUE, main = "WOMEN", xlab = "Systolic Bp")
```




```
# Mound shaped
hist(DIASBP[GENDER==0], prob=TRUE, main = "MEN", xlab = "Diastolic Bp")
```



```
# Mound shaped
hist(DIASBP[GENDER==1], prob=TRUE, main = "WOMEN", xlab = "Diastolic Bp")
```



s. Z-scores for largest and smallest observation of men and women

Results are in that order
men systolic blood pressure maximum value z-score # It may not be outlier because it is close to 2

men systolic blood pressure minimum value z-score # Outlier because it is smaller than -3

women systolic blood pressure maximum value z-score # Outlier because it is greater than -3

women systolic blood pressure minimum value z-score # Not an outlier because it is between -2 and 2

men diastolic blood pressure maximum value z-score # It may not be outlier because it is close to 2

men diastolic blood pressure minimum value z-score # Not an outlier because it is between -2 and 2

women diastolic blood pressure maximum value z-score # It can be say outlieri it is too close to -3

women diastolic blood pressure minimum value z-score # Outlier because it is smaller than -3

```
## [1] 2.282955
```

```
## [1] -3.59784
```

```
## [1] 3.891753
```

```
## [1] -1.971528
```

```
## [1] 2.383597
```

```
## [1] -1.891913
```

```
## [1] 2.964321
```

```
## [1] -3.49993
```