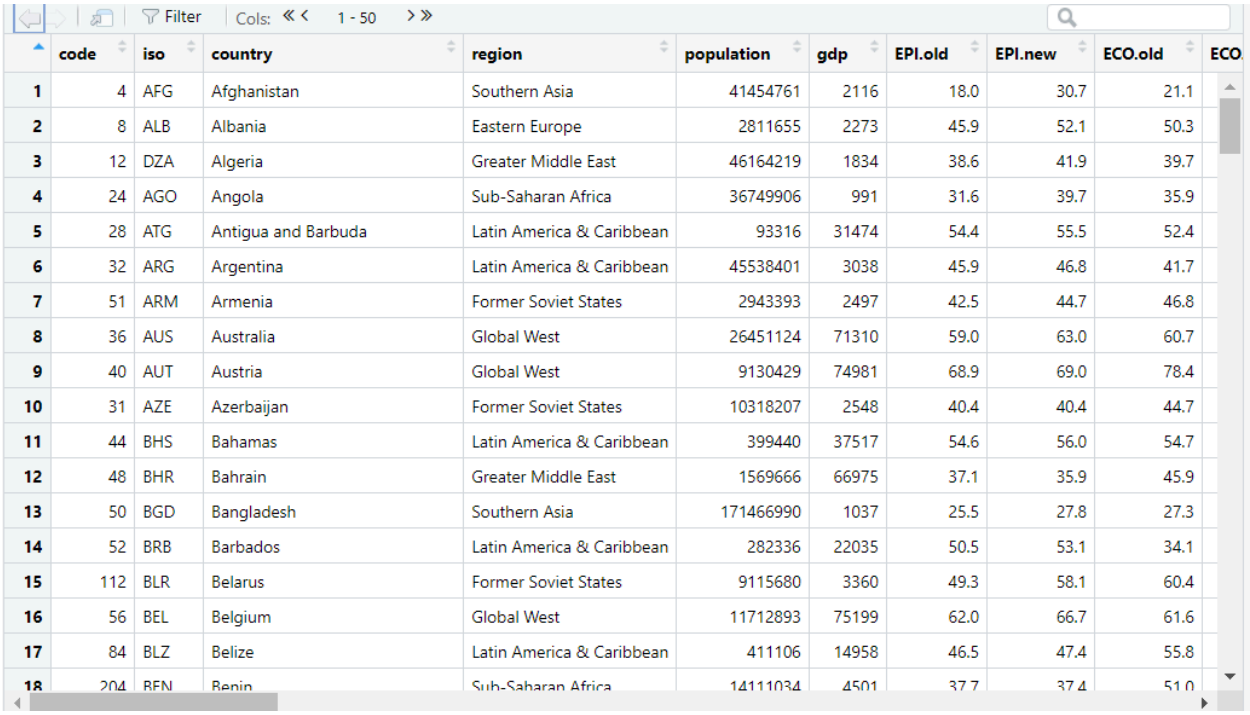
Lab 1 Exercise 1

Line 11 - Displays a view of the csv file



Filter | Cols: « < 1 - 50 > »

	code	iso	country	region	population	gdp	EPI.old	EPI.new	ECO.old	ECO
1	4	AFG	Afghanistan	Southern Asia	41454761	2116	18.0	30.7	21.1	
2	8	ALB	Albania	Eastern Europe	2811655	2273	45.9	52.1	50.3	
3	12	DZA	Algeria	Greater Middle East	46164219	1834	38.6	41.9	39.7	
4	24	AGO	Angola	Sub-Saharan Africa	36749906	991	31.6	39.7	35.9	
5	28	ATG	Antigua and Barbuda	Latin America & Caribbean	93316	31474	54.4	55.5	52.4	
6	32	ARG	Argentina	Latin America & Caribbean	45538401	3038	45.9	46.8	41.7	
7	51	ARM	Armenia	Former Soviet States	2943393	2497	42.5	44.7	46.8	
8	36	AUS	Australia	Global West	26451124	71310	59.0	63.0	60.7	
9	40	AUT	Austria	Global West	9130429	74981	68.9	69.0	78.4	
10	31	AZE	Azerbaijan	Former Soviet States	10318207	2548	40.4	40.4	44.7	
11	44	BHS	Bahamas	Latin America & Caribbean	399440	37517	54.6	56.0	54.7	
12	48	BHR	Bahrain	Greater Middle East	1569666	66975	37.1	35.9	45.9	
13	50	BGD	Bangladesh	Southern Asia	171466990	1037	25.5	27.8	27.3	
14	52	BRB	Barbados	Latin America & Caribbean	282336	22035	50.5	53.1	34.1	
15	112	BLR	Belarus	Former Soviet States	9115680	3360	49.3	58.1	60.4	
16	56	BEL	Belgium	Global West	11712893	75199	62.0	66.7	61.6	
17	84	BLZ	Belize	Latin America & Caribbean	411106	14958	46.5	47.4	55.8	
18	204	BEN	Benin	Sub-Saharan Africa	14111034	4501	37.7	37.4	51.0	

Line 14 - summary of variables in dataframe

```
> summary(epi.data$EPI.new)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  24.50   38.25   45.50   46.84   53.10   75.30
```

Line 17 - print values in variable

```
> # print values in variable
> epi.data$EPI.new
  [1] 30.7 52.1 41.9 39.7 55.5 46.8 44.7 63.0 69.0 40.4 56.0 35.9 27.8 53.1 58.1 66.7 47.4 37.4
 [19] 43.3 44.9 45.6 49.0 53.0 48.5 56.3 41.5 33.0 37.9 31.0 38.1 61.1 38.3 35.2 50.0 35.5 49.4
 [37] 37.9 55.5 42.5 62.6 52.3 54.0 65.6 39.0 67.9 32.2 49.2 47.6 51.2 43.8 41.5 41.6 28.6 75.3
 [55] 38.5 35.8 45.8 73.7 67.1 53.1 37.1 46.9 74.6 36.6 67.4 46.0 32.6 36.2 41.6 48.6 36.2 40.2
 [73] 60.1 64.3 27.6 33.8 41.6 30.4 65.7 48.1 60.5 48.5 61.7 47.5 47.5 36.9 44.1 44.9 42.2 26.1
 [91] 59.9 40.1 36.6 34.1 63.9 75.0 29.9 34.9 41.2 38.1 33.9 66.6 42.6 34.2 47.3 44.7 40.6 45.6
[109] 37.0 47.6 39.7 38.6 26.9 43.8 32.9 67.2 57.7 47.4 39.2 37.5 50.0 70.0 51.9 25.5 52.9 36.5
[127] 39.0 46.6 32.0 64.4 62.2 47.2 41.2 57.2 46.5 33.4 51.0 54.1 46.8 35.9 42.6 43.3 49.3 48.2
[145] 39.7 53.8 65.0 62.5 41.8 42.9 51.0 64.2 38.7 38.6 56.6 70.5 68.0 50.3 31.9 43.1 45.4 49.7
[163] 35.2 40.2 52.1 45.7 37.6 40.7 35.4 54.6 52.0 72.7 57.3 43.9 42.9 44.6 53.1 24.5 46.1 51.7
```

Line 27 - print values in variable

```
> EPI.new
  [1] 30.7 52.1 41.9 39.7 55.5 46.8 44.7 63.0 69.0 40.4 56.0 35.9 27.8 53.1 58.1 66.7 47.4 37.4
 [19] 43.3 44.9 45.6 49.0 53.0 48.5 56.3 41.5 33.0 37.9 31.0 38.1 61.1 38.3 35.2 50.0 35.5 49.4
 [37] 37.9 55.5 42.5 62.6 52.3 54.0 65.6 39.0 67.9 32.2 49.2 47.6 51.2 43.8 41.5 41.6 28.6 75.3
 [55] 38.5 35.8 45.8 73.7 67.1 53.1 37.1 46.9 74.6 36.6 67.4 46.0 32.6 36.2 41.6 48.6 36.2 40.2
 [73] 60.1 64.3 27.6 33.8 41.6 30.4 65.7 48.1 60.5 48.5 61.7 47.5 47.5 36.9 44.1 44.9 42.2 26.1
 [91] 59.9 40.1 36.6 34.1 63.9 75.0 29.9 34.9 41.2 38.1 33.9 66.6 42.6 34.2 47.3 44.7 40.6 45.6
[109] 37.0 47.6 39.7 38.6 26.9 43.8 32.9 67.2 57.7 47.4 39.2 37.5 50.0 70.0 51.9 25.5 52.9 36.5
[127] 39.0 46.6 32.0 64.4 62.2 47.2 41.2 57.2 46.5 33.4 51.0 54.1 46.8 35.9 42.6 43.3 49.3 48.2
[145] 39.7 53.8 65.0 62.5 41.8 42.9 51.0 64.2 38.7 38.6 56.6 70.5 68.0 50.3 31.9 43.1 45.4 49.7
[163] 35.2 40.2 52.1 45.7 37.6 40.7 35.4 54.6 52.0 72.7 57.3 43.9 42.9 44.6 53.1 24.5 46.1 51.7
```

Line 40 - print true if there are NAs in variable, false otherwise (printed 0 which is false)

```
> # find NAs in variable - outputs vector of logical values, true if NA, false otherwise
> NAs <- is.na(EPI)
> EPI[which(NAs)]
numeric(0)
```

Line 45 - print values in variable

```
> # print values in variable
> MHP <- epi.data$MHP.new
> MHP
  [1]    NA  24.4   0.0   0.0  33.1  33.0    NA  42.5    NA  20.0  17.4  59.4  61.1   0.0    NA
 [16]  80.0  36.7   0.0    NA    NA  31.0    NA  39.9   1.1   0.0    NA    NA   0.0  17.6  75.6
 [31]  26.9    NA    NA  37.8   2.5  51.0  23.3  37.5   2.4  33.4  33.9  16.0    NA  59.6  17.3
 [46]   0.0   3.2  43.5  66.3  27.5  38.9  25.3   0.0  82.6    NA    NA  16.9  44.7  58.0  50.0
 [61]  56.4  14.3  95.1   0.1  39.7   3.4  19.2   7.7  21.4   0.0  26.8  25.6    NA  29.2   0.2
 [76]  13.0  22.2   0.4  46.9  50.0  41.8  20.3  58.7   0.0  41.7  20.3   8.0   7.4    NA    NA
 [91]  42.9   0.0    NA   4.5 100.0    NA  14.8    NA  12.4   1.3    NA  57.9   5.1  20.7  17.1
[106]  48.6   0.0    NA    NA   0.0   6.6  16.6   2.1   7.7    NA  53.2  24.5  40.6    NA   0.4
[121]    NA  11.3  65.8   0.7  23.5   3.5    NA  16.6  12.5  79.5  30.5  36.2  22.7  90.0   4.4
[136]    NA  13.2  10.7   5.2   0.0   5.8  18.4    NA  50.2  39.1   0.0    NA  27.8   7.3  54.6
[151]  15.1  47.4   1.5  28.0  43.4  47.5    NA  10.5    NA  37.8  28.6  17.7   0.0   3.7   3.4
[166]  25.7  21.0  12.5    NA  34.2  31.1  54.2  46.1  22.1    NA   1.8  14.9   7.0    NA    NA
```

Line 51 - Prints a list of NAs from variable

```
> # find NAs inv variavle - outputs vector of logical values, true if NA, false otherwise
> NAs <- is.na(MHP)
> # print NAs
> MHP[which(NAs)]
 [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[31] NA NA NA NA NA NA
```

Line 56 - Prints a list of non NA values from variable

```
> # take subset of NOT NAs from variable
> MHP.noNA <- MHP[!NAs]
> MHP.noNA
  [1]  24.4   0.0   0.0  33.1  33.0  42.5  20.0  17.4  59.4  61.1   0.0  80.0  36.7   0.0  31.0
 [16]  39.9   1.1   0.0   0.0  17.6  75.6  26.9  37.8   2.5  51.0  23.3  37.5   2.4  33.4  33.9
 [31]  16.0  59.6  17.3   0.0   3.2  43.5  66.3  27.5  38.9  25.3   0.0  82.6  16.9  44.7  58.0
 [46]  50.0  56.4  14.3  95.1   0.1  39.7   3.4  19.2   7.7  21.4   0.0  26.8  25.6  29.2   0.2
 [61]  13.0  22.2   0.4  46.9  50.0  41.8  20.3  58.7   0.0  41.7  20.3   8.0   7.4  42.9   0.0
 [76]   4.5 100.0  14.8  12.4   1.3  57.9   5.1  20.7  17.1  48.6   0.0   0.0   6.6  16.6   2.1
 [91]   7.7  53.2  24.5  40.6   0.4  11.3  65.8   0.7  23.5   3.5  16.6  12.5  79.5  30.5  36.2
[106]  22.7  90.0   4.4  13.2  10.7   5.2   0.0   5.8  18.4  50.2  39.1   0.0  27.8   7.3  54.6
[121]  15.1  47.4   1.5  28.0  43.4  47.5  10.5  37.8  28.6  17.7   0.0   3.7   3.4  25.7  21.0
[136]  12.5  34.2  31.1  54.2  46.1  22.1   1.8  14.9   7.0
```

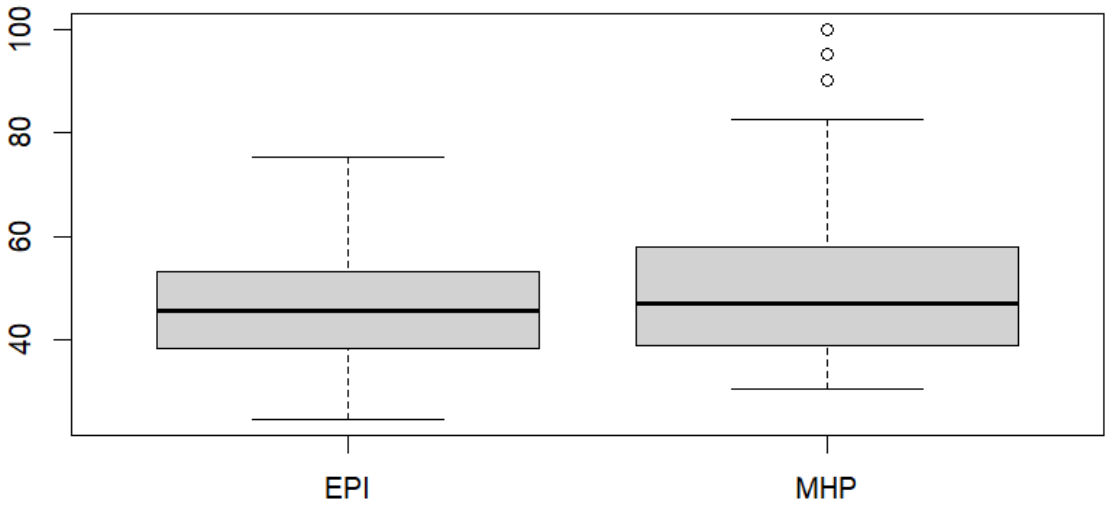
Line 61 - Prints list of values above 30 from variable

```
> # filter for only values above 30
> MHP.above30 <- MHP.noNA[MHP.noNA>30]
> MHP.above30
 [1]  33.1  33.0  42.5  59.4  61.1  80.0  36.7  31.0  39.9  75.6  37.8  51.0  37.5  33.4  33.9
[16]  59.6  43.5  66.3  38.9  82.6  44.7  58.0  50.0  56.4  95.1  39.7  46.9  50.0  41.8  58.7
[31]  41.7  42.9 100.0  57.9  48.6  53.2  40.6  65.8  79.5  30.5  36.2  90.0  50.2  39.1  54.6
[46]  47.4  43.4  47.5  37.8  34.2  31.1  54.2  46.1
```

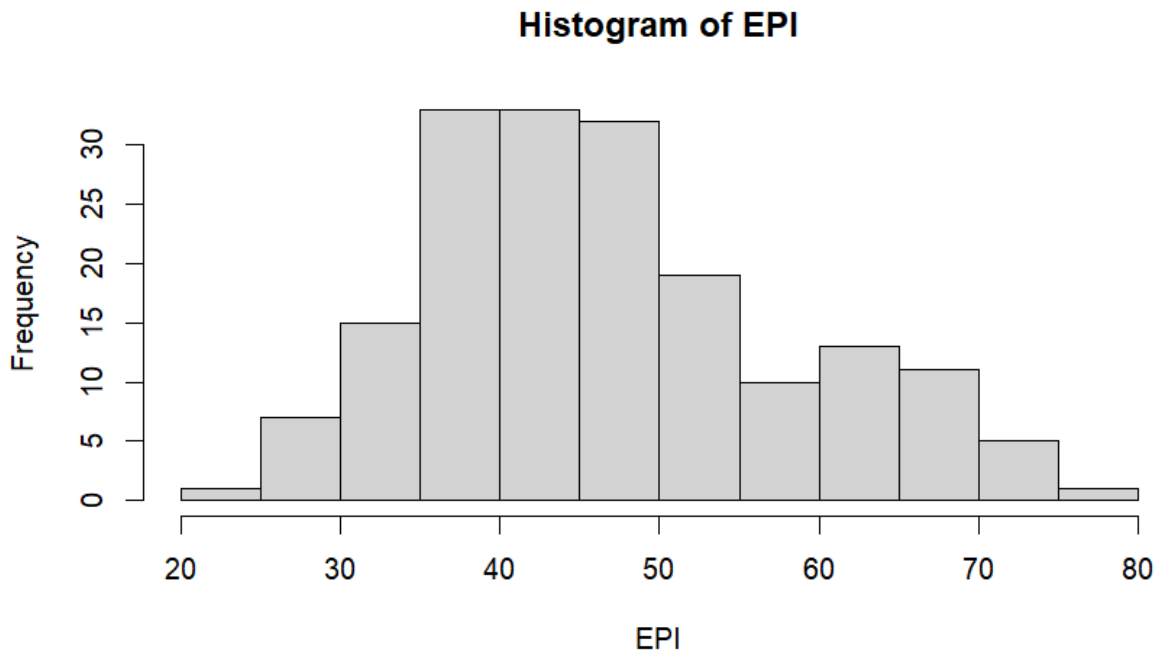
Line 64 - summary of values above 30 in the variable

```
> # stats
> summary(MHP.above30)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  30.50   38.90   46.90   50.77   58.00  100.00
```

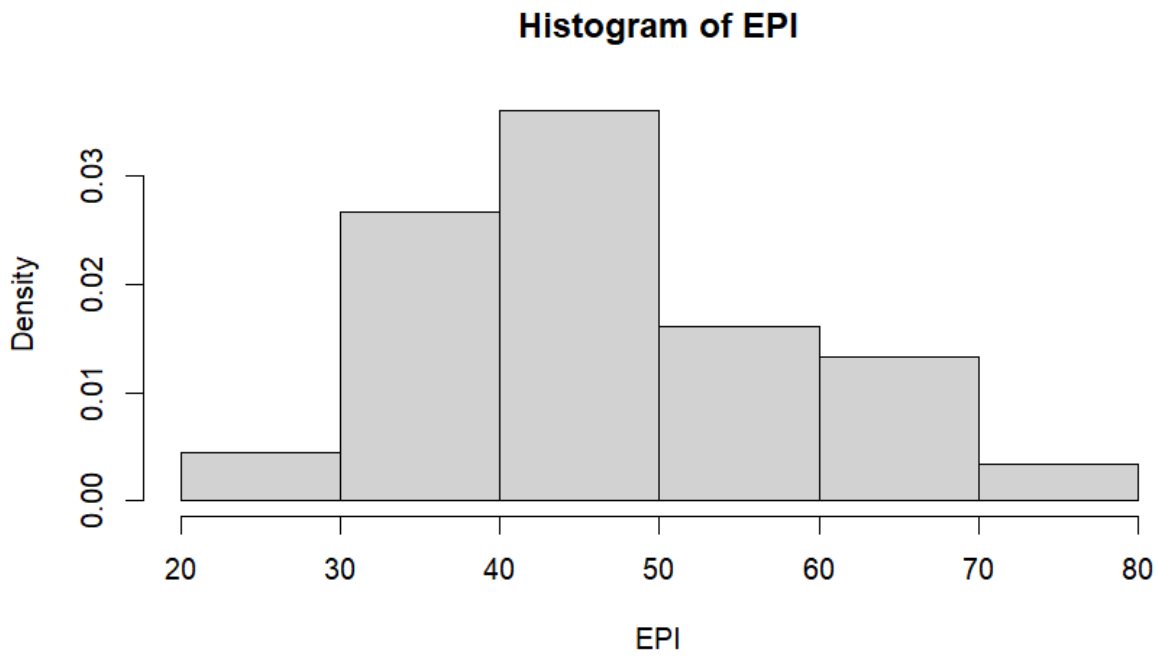
Line 67 - Create a boxplot of the variables



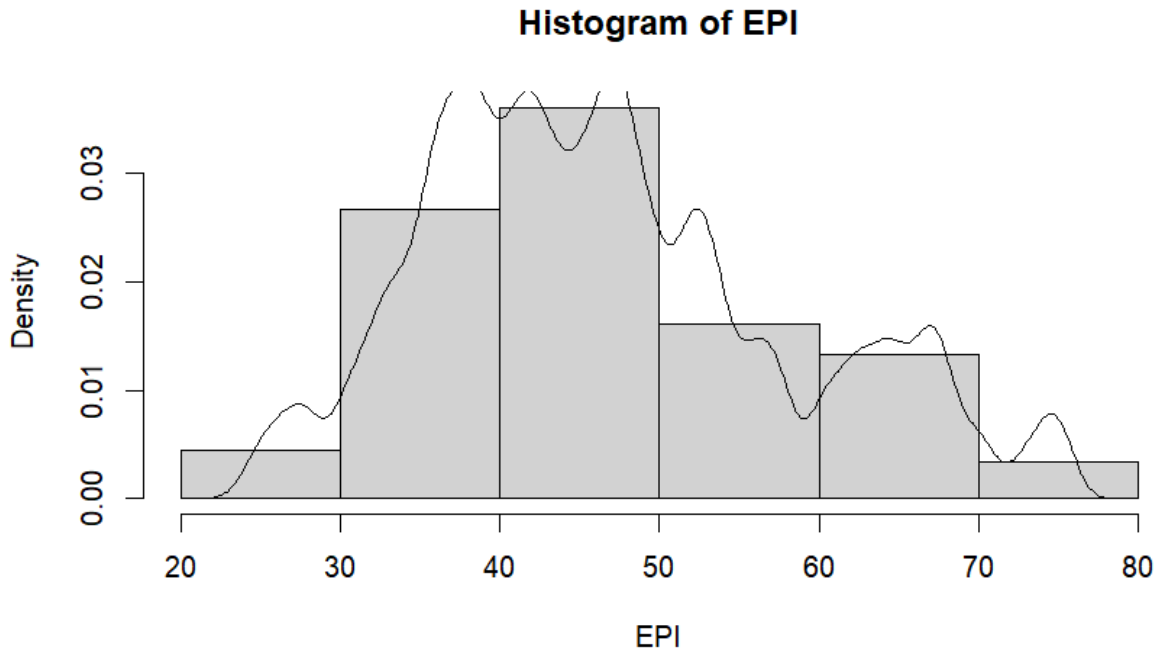
Line 73 - Create a histogram of the EPI variable



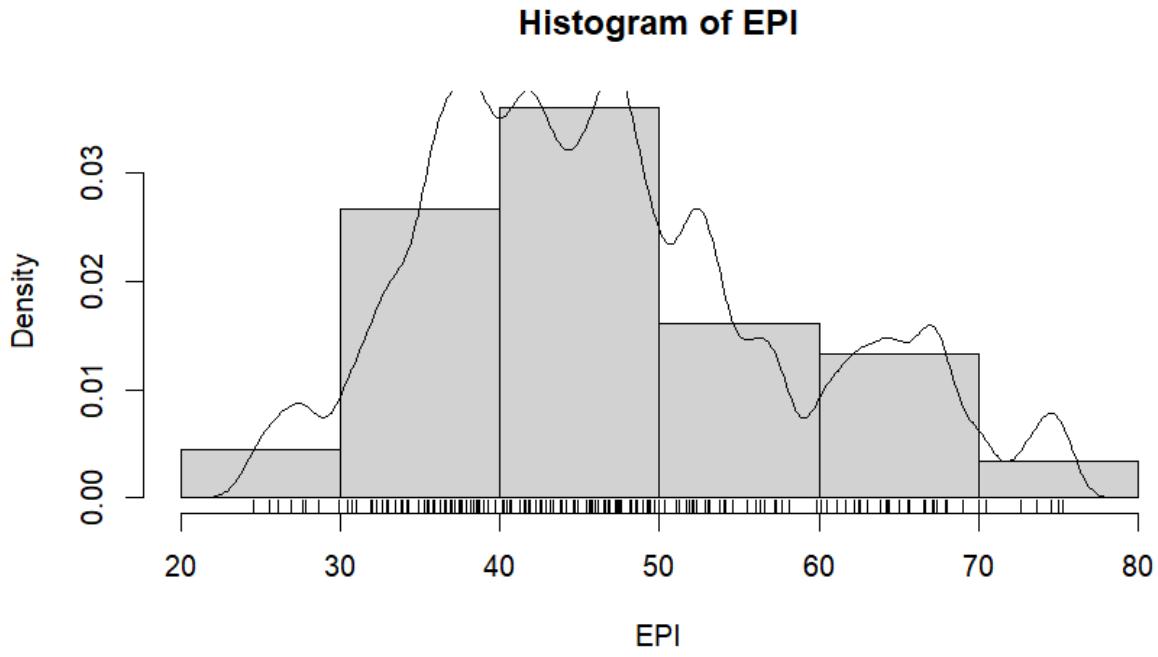
Line 79 - Create a histogram of the EPI variable over the range 20-80 and in segments of 10



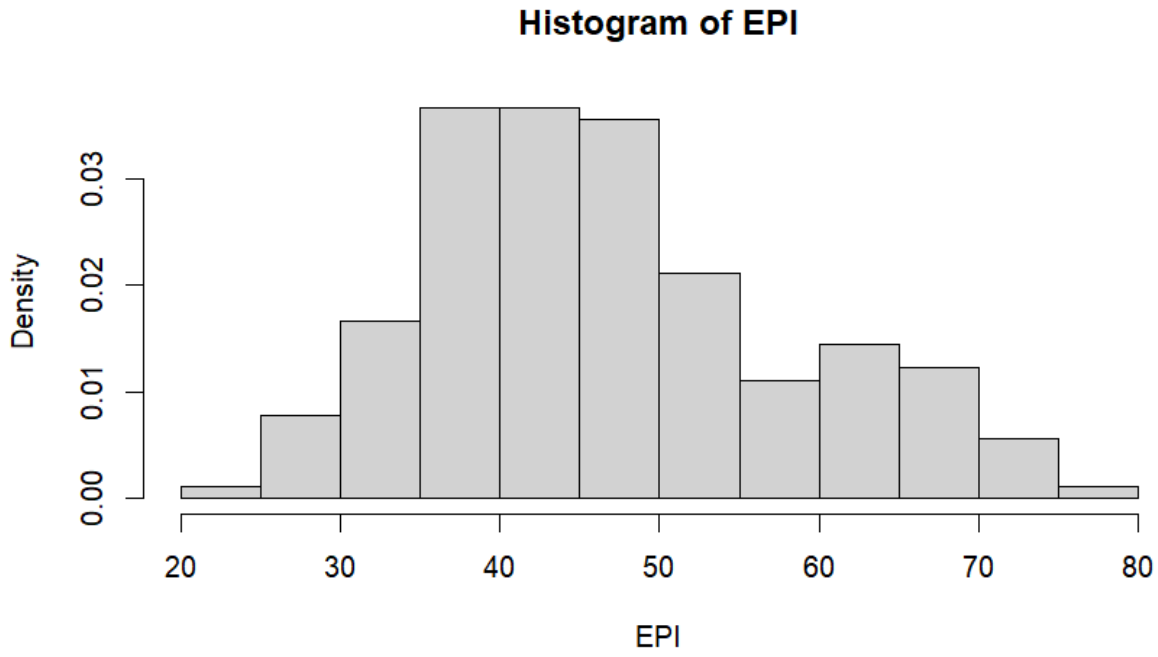
Line 82 - Print estimated density curve for histogram



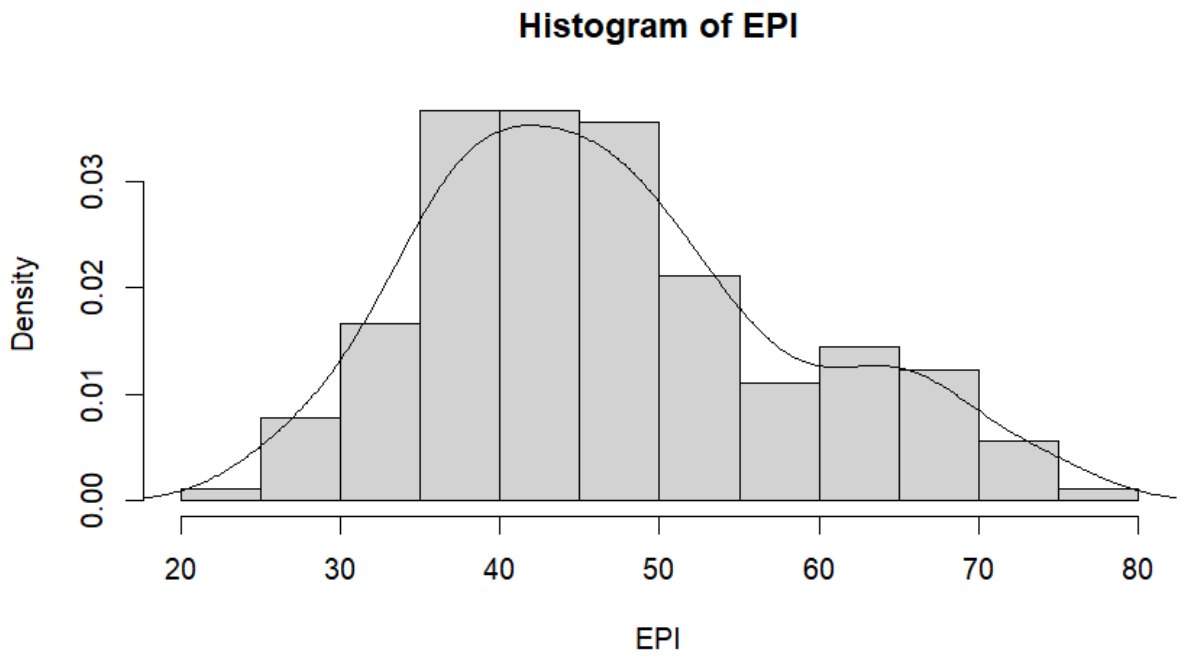
Line 85 - Prints the rug for histogram



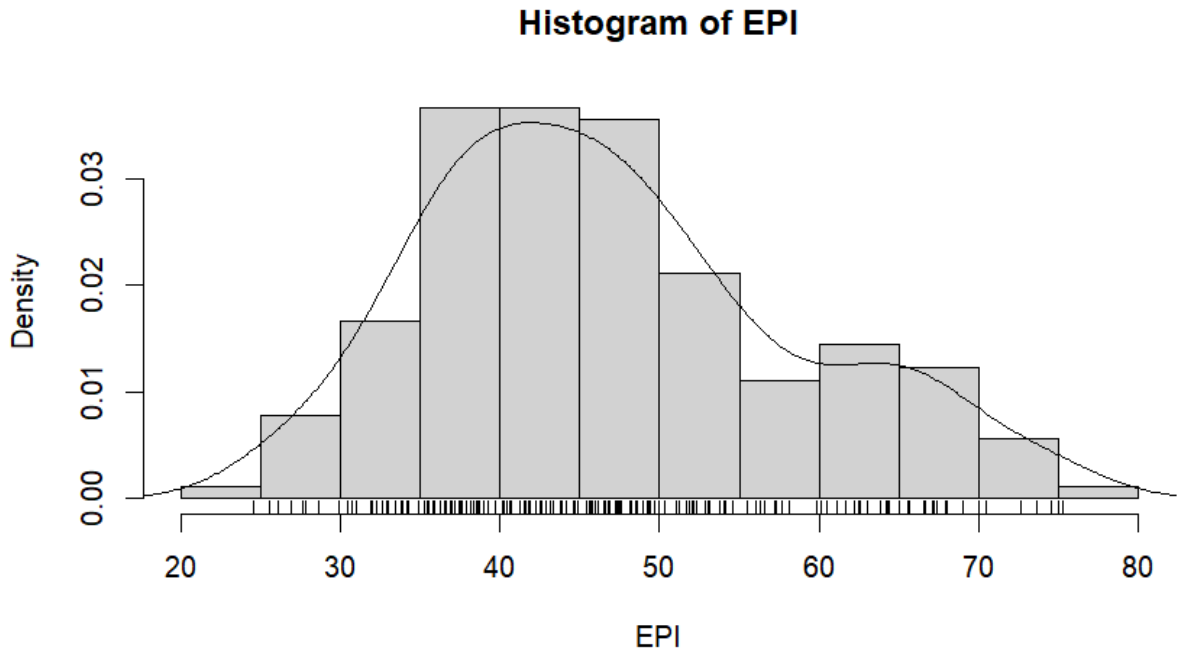
Line 90 - Create a histogram of the EPI variable over the range 20-80 and in segments of 10



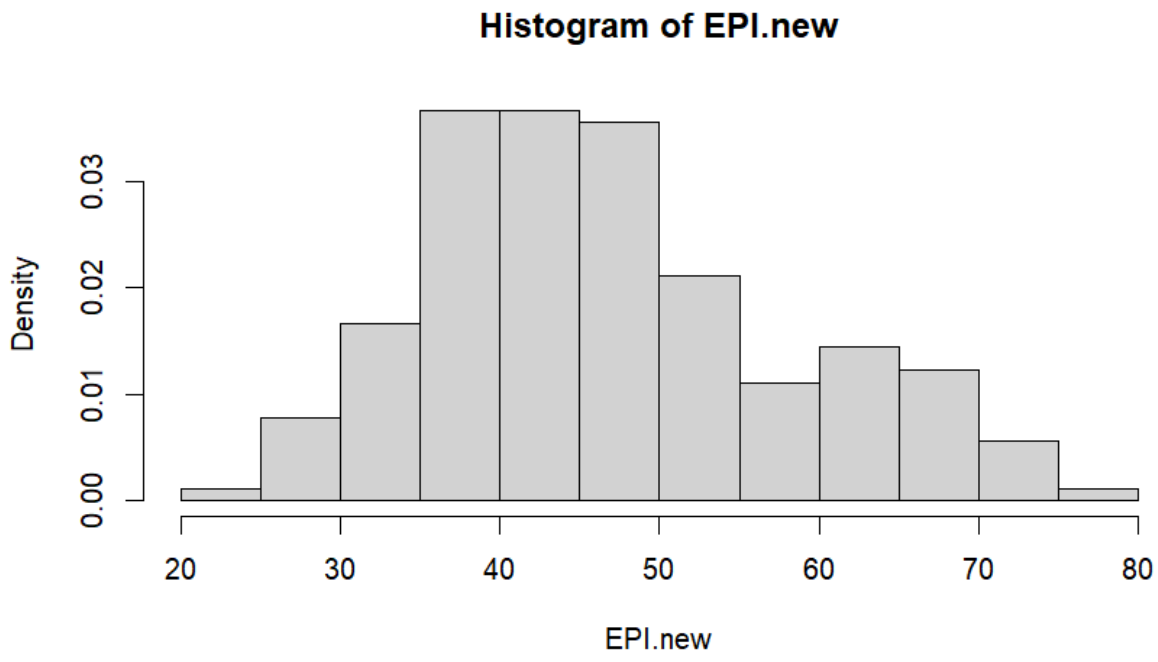
Line 93 - Print estimated density curve for variable



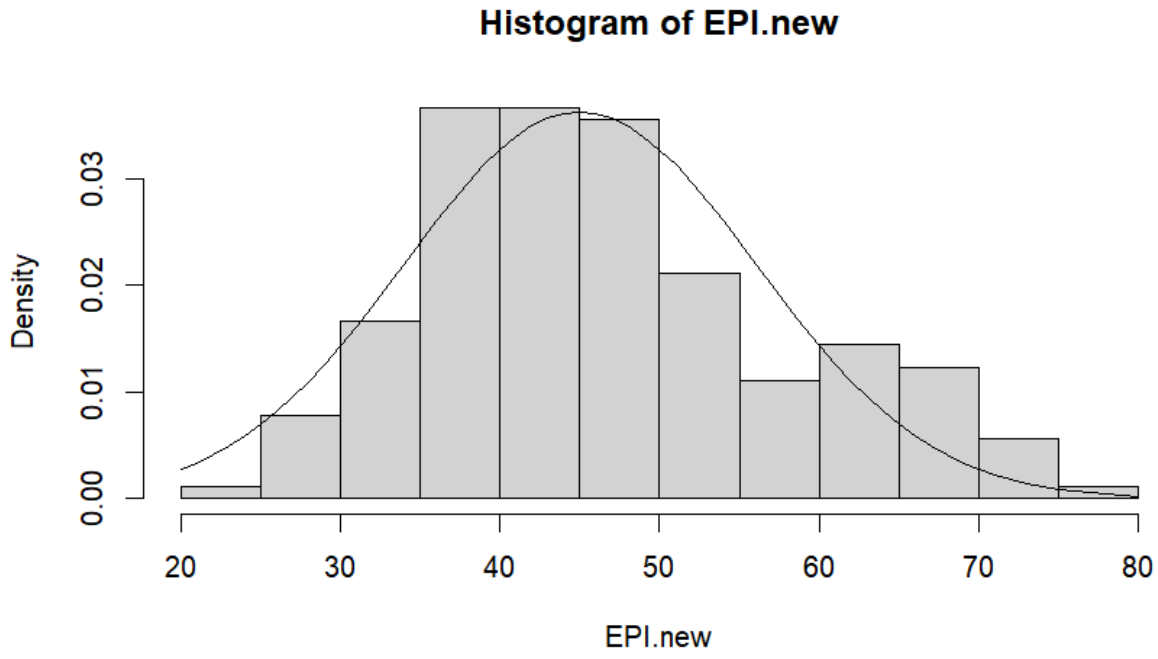
Line 96 - Print the rug for histogram



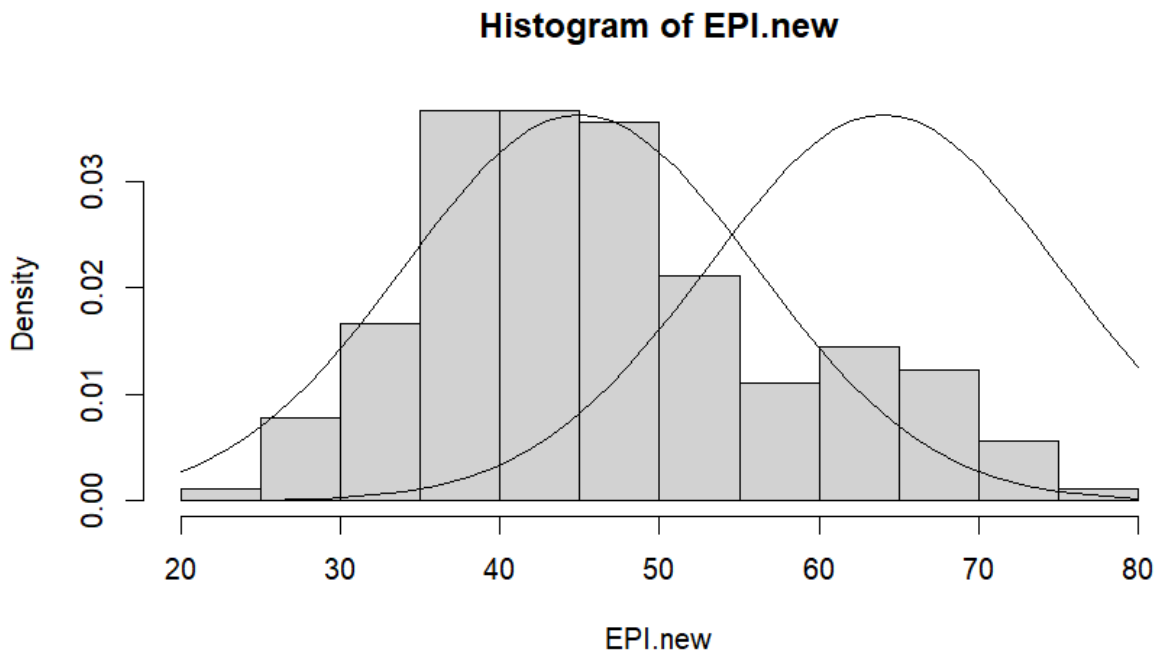
Line 100 - Create a histogram of the EPI.new variable over the range 20-80 and in segments of 10



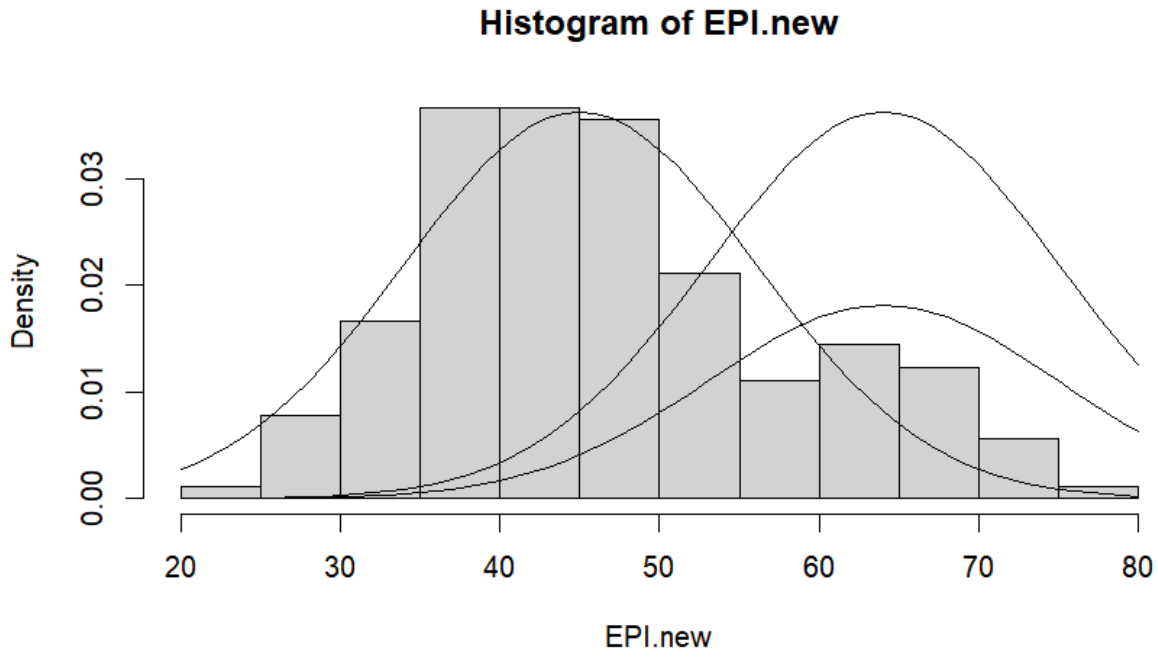
Line 109 - Print density curve for variable



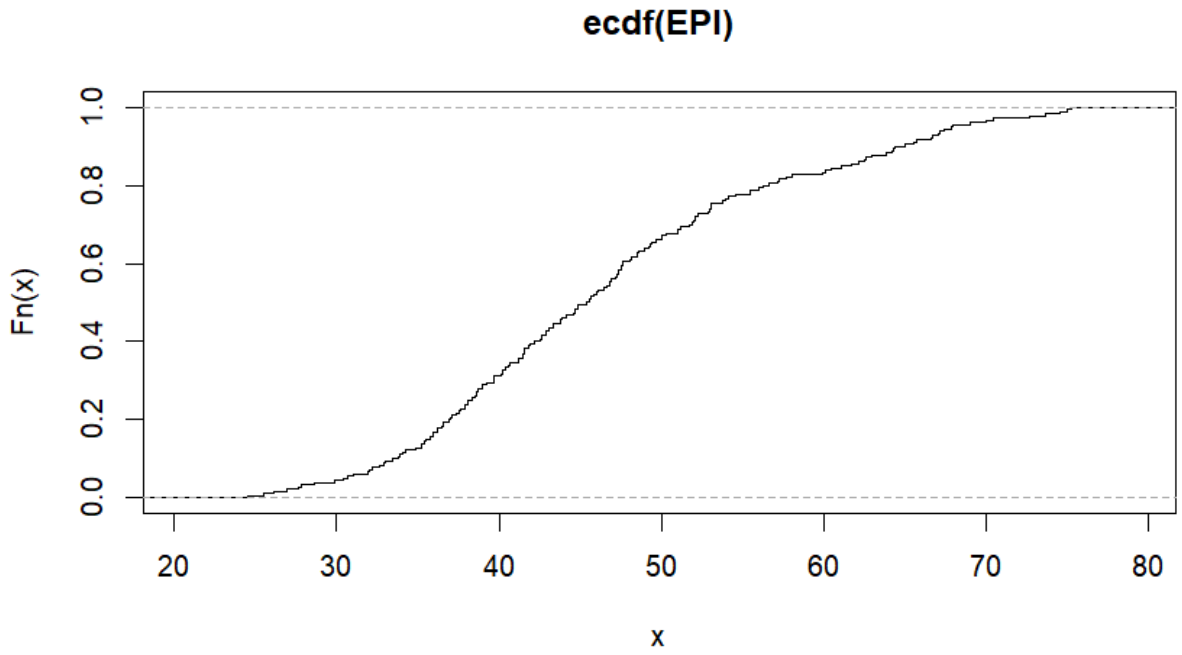
Line 115 - Print 2nd density curve



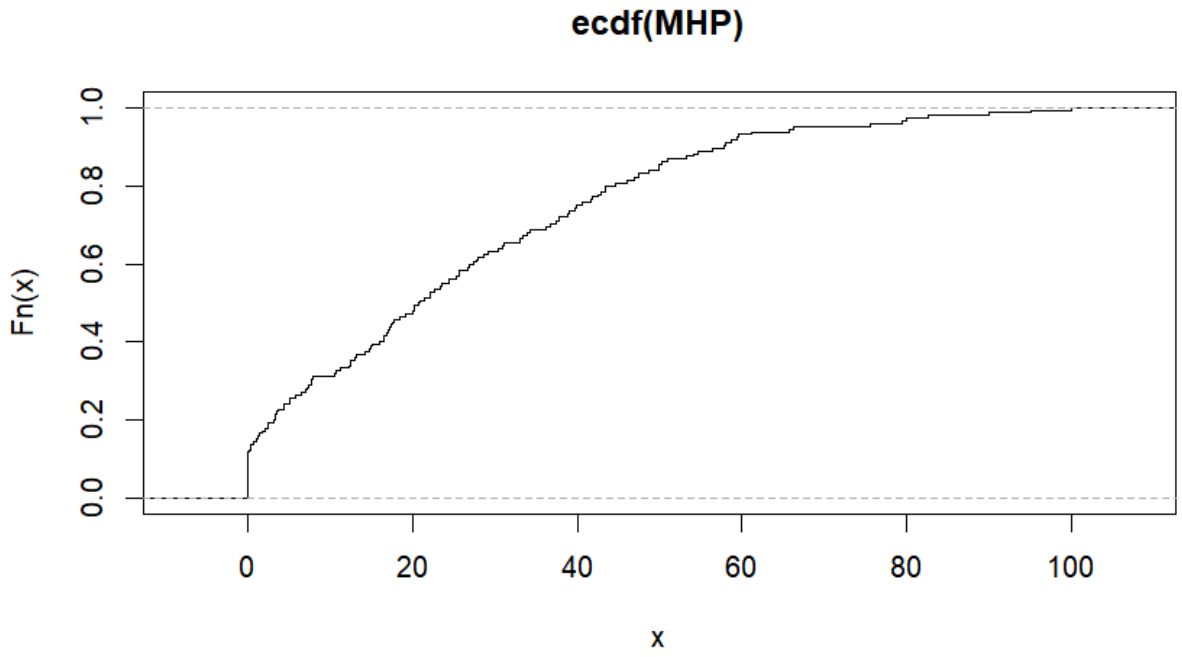
Line 118 - Print 3rd density curve



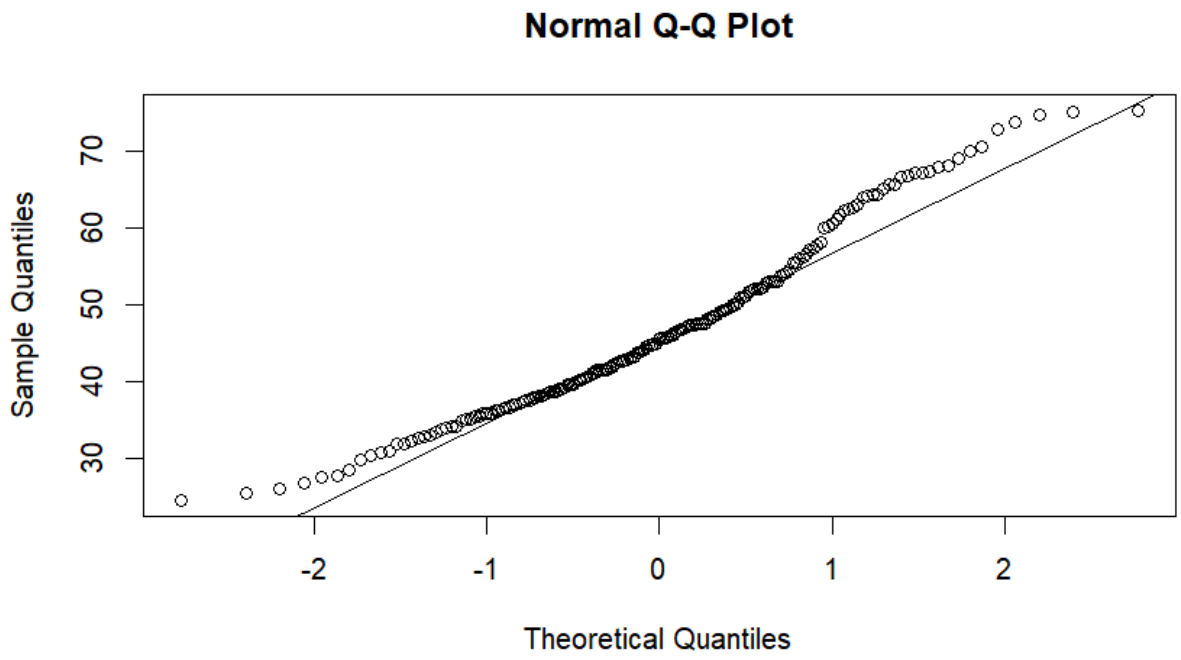
Line 123 - Print empirical cumulative distribution function (ecdf) plot for EPI



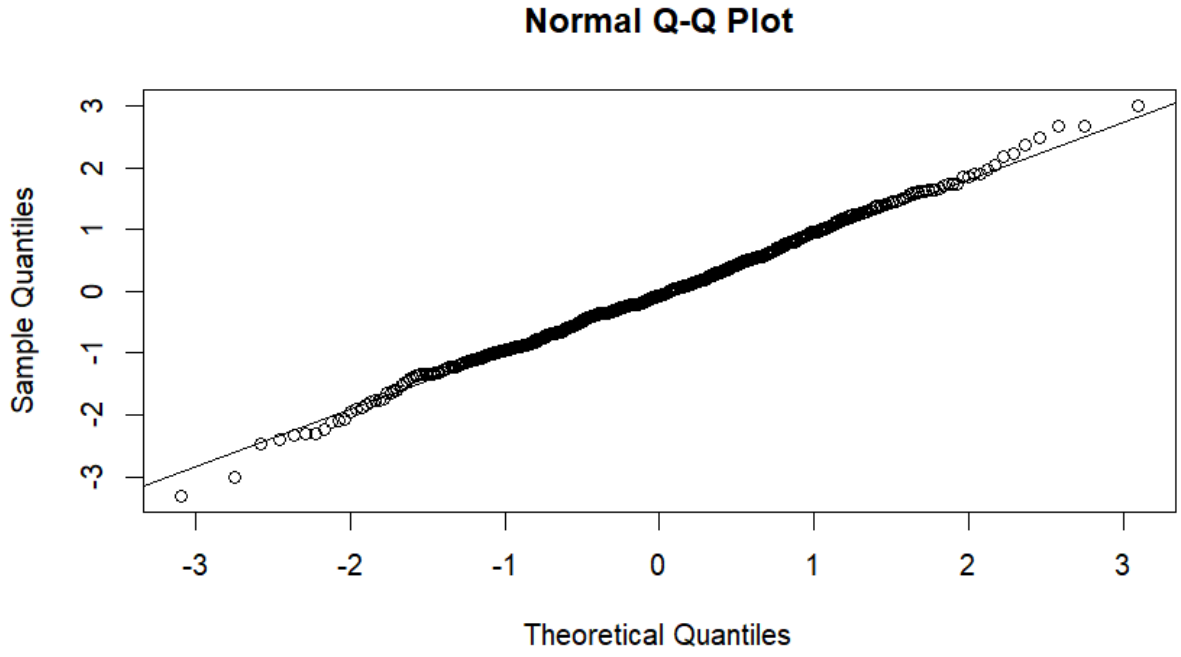
Line 125 - Print empirical cumulative distribution function (ecdf) plot for MHP



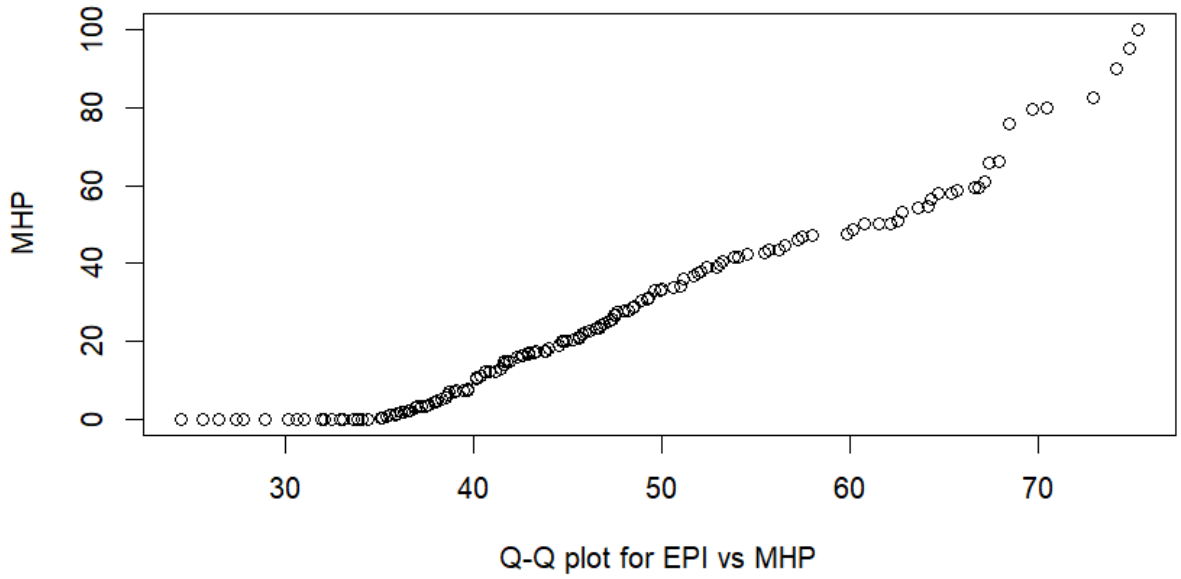
Line 131 - print quantile quantile plot for variable with theoretical normal distribution

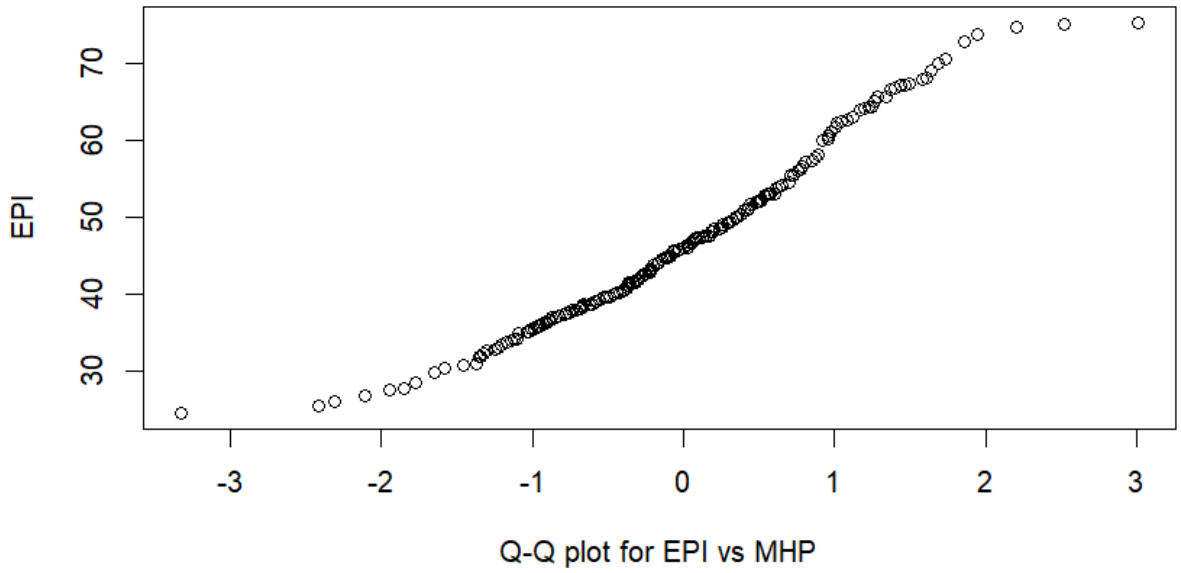


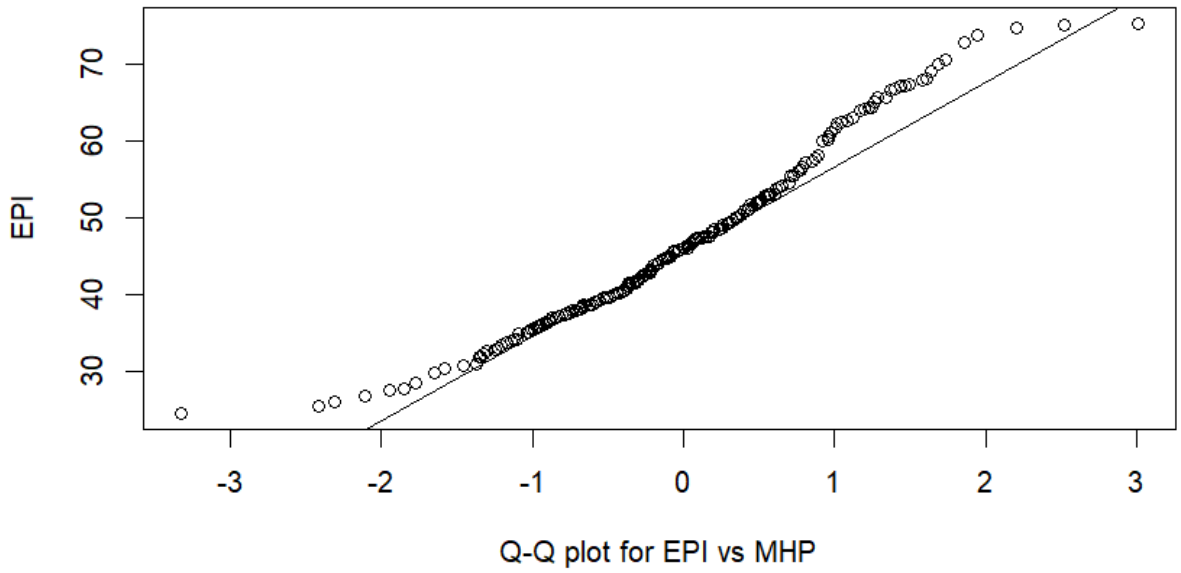
Line 136 - Print quantile quantile plot for random numbers from a normal distribution with theoretical normal distribution

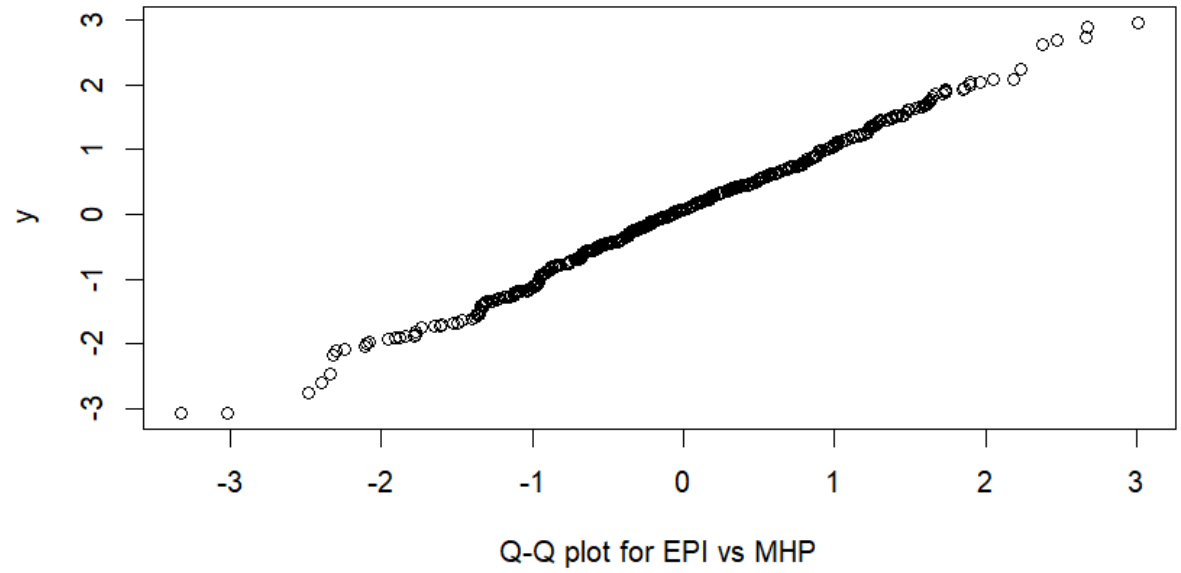


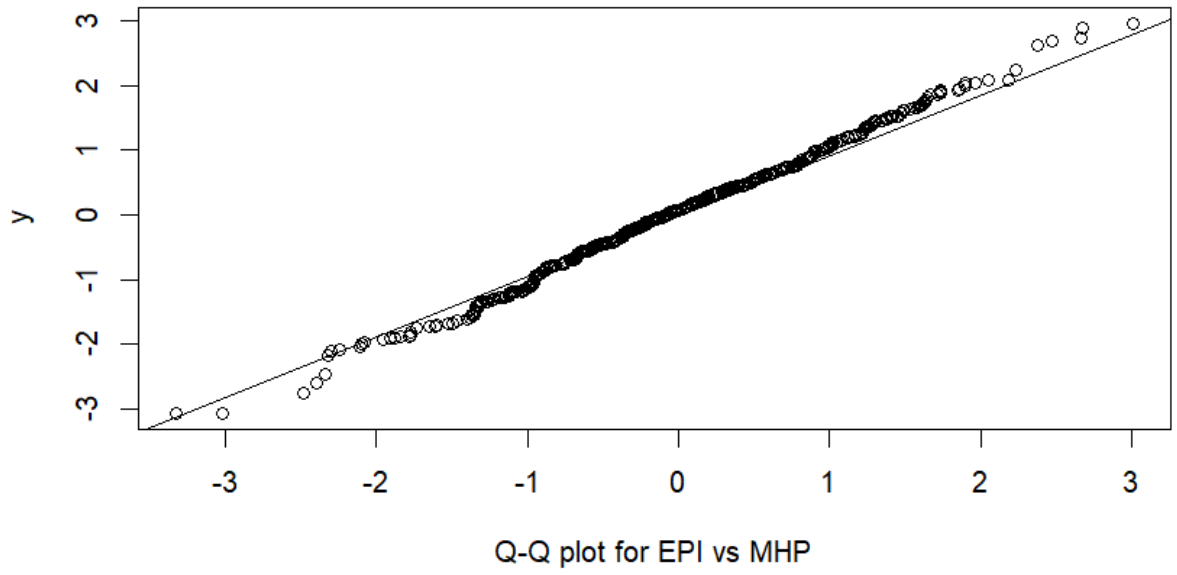
Line 144, 146, 147, 151, 152 - Print quantile quantile value for 2 variables (EPI vs MHP)


MHP
0
20
40
60
80
100
30
40
50
60
70
Q-Q plot for EPI vs MHP

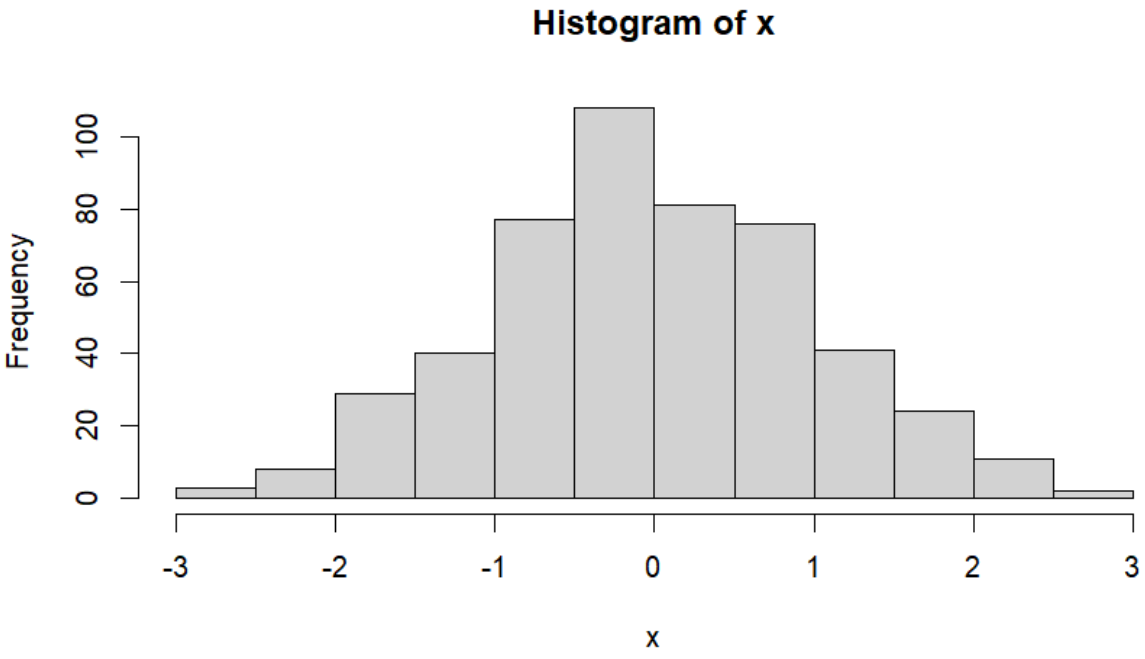

EPI
30
40
50
60
70
-3
-2
-1
0
1
2
3
Q-Q plot for EPI vs MHP


EPI
30
40
50
60
70
-3
-2
-1
0
1
2
3
Q-Q plot for EPI vs MHP

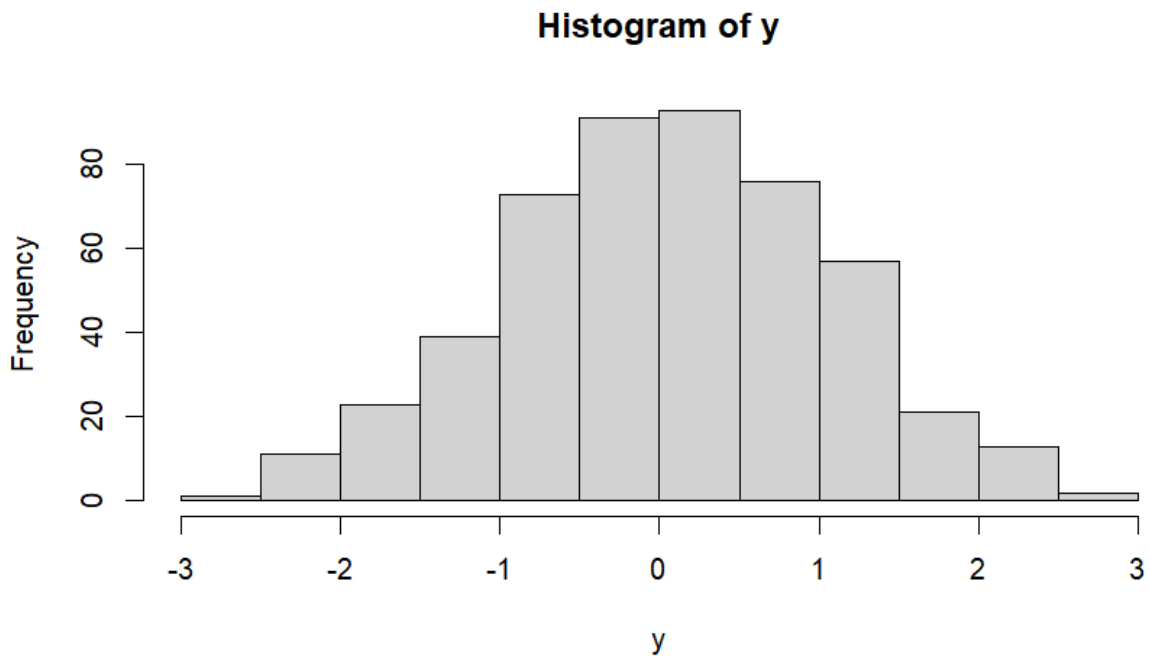

y
-3
-2
-1
0
1
2
3
-3
-2
-1
0
1
2
3
Q-Q plot for EPI vs MHP



Line 160 - Histogram of X variable



Line 161 - Histogram of Y variable



Line 163 - Print Shapiro test for X variable

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.99726, p-value = 0.5789
```

Line 164 - Print Shapiro test for Y variable

```
> shapiro.test(y)

        Shapiro-Wilk normality test

data:  y
W = 0.99594, p-value = 0.2263
```

Line 166 and 167 - Produce errors, could not find function “ad.test()”

Line 169 - Print Kolmogorov-Smirnov test for variables x and y

```
> ks.test(x,y)

        Asymptotic two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.056, p-value = 0.4131
alternative hypothesis: two-sided
```

Line 171 - Print wilcox test for variables x and y

```
> wilcox.test(x,y)

        Wilcoxon rank sum test with continuity correction

data:  x and y
W = 118830, p-value = 0.1767
alternative hypothesis: true location shift is not equal to 0
```

Line 173 - Prints F test to compare variances in x and y variables

```
> var.test(x,y)

        F test to compare two variances

data:  x and y
F = 0.99646, num df = 499, denom df = 499, p-value = 0.9685
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.835919 1.187842
sample estimates:
ratio of variances
         0.9964635
```

Line 174 - Prints Welch 2 sample t test for variables x and y

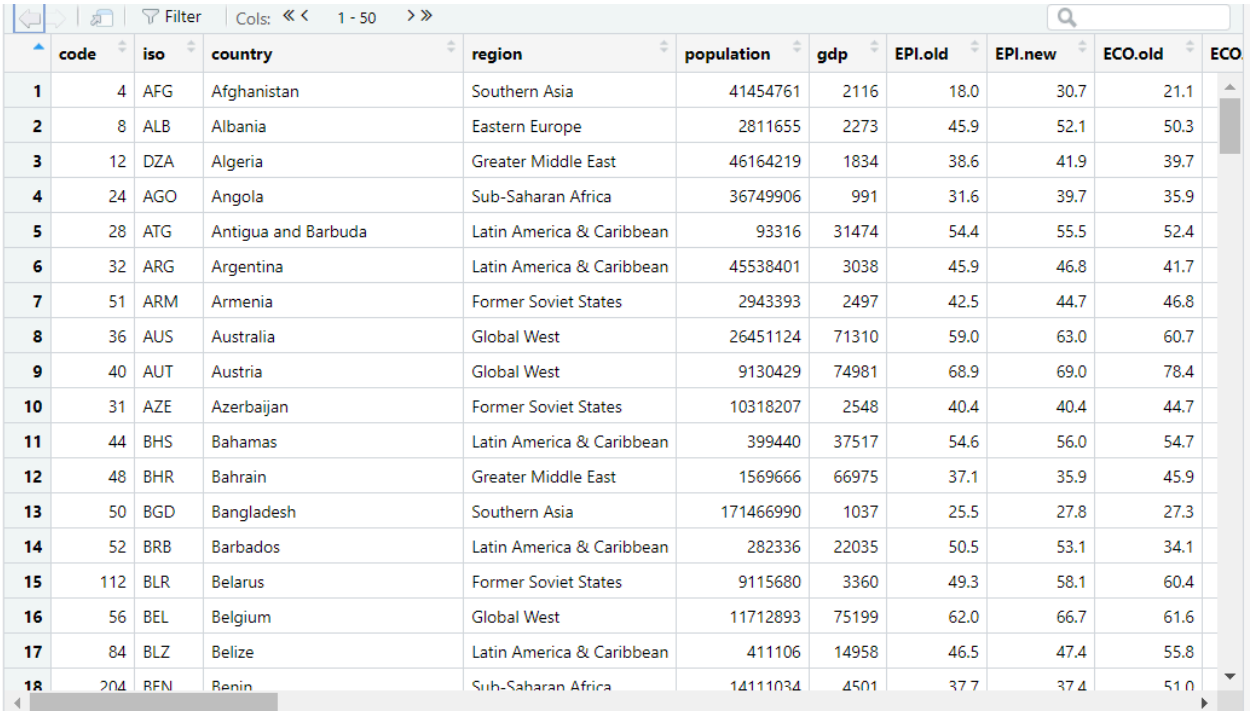
```
> t.test(x,y)

        Welch Two Sample t-test

data:  x and y
t = -1.2321, df = 998, p-value = 0.2182
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.20402060  0.04664221
sample estimates:
  mean of x   mean of y
-0.01772948  0.06095971
```

Lab 1 Exercise 2

Line 11 - Displays a view of the csv file



Filter | Cols: « < 1 - 50 > »

	code	iso	country	region	population	gdp	EPI.old	EPI.new	ECO.old	ECO
1	4	AFG	Afghanistan	Southern Asia	41454761	2116	18.0	30.7	21.1	
2	8	ALB	Albania	Eastern Europe	2811655	2273	45.9	52.1	50.3	
3	12	DZA	Algeria	Greater Middle East	46164219	1834	38.6	41.9	39.7	
4	24	AGO	Angola	Sub-Saharan Africa	36749906	991	31.6	39.7	35.9	
5	28	ATG	Antigua and Barbuda	Latin America & Caribbean	93316	31474	54.4	55.5	52.4	
6	32	ARG	Argentina	Latin America & Caribbean	45538401	3038	45.9	46.8	41.7	
7	51	ARM	Armenia	Former Soviet States	2943393	2497	42.5	44.7	46.8	
8	36	AUS	Australia	Global West	26451124	71310	59.0	63.0	60.7	
9	40	AUT	Austria	Global West	9130429	74981	68.9	69.0	78.4	
10	31	AZE	Azerbaijan	Former Soviet States	10318207	2548	40.4	40.4	44.7	
11	44	BHS	Bahamas	Latin America & Caribbean	399440	37517	54.6	56.0	54.7	
12	48	BHR	Bahrain	Greater Middle East	1569666	66975	37.1	35.9	45.9	
13	50	BGD	Bangladesh	Southern Asia	171466990	1037	25.5	27.8	27.3	
14	52	BRB	Barbados	Latin America & Caribbean	282336	22035	50.5	53.1	34.1	
15	112	BLR	Belarus	Former Soviet States	9115680	3360	49.3	58.1	60.4	
16	56	BEL	Belgium	Global West	11712893	75199	62.0	66.7	61.6	
17	84	BLZ	Belize	Latin America & Caribbean	411106	14958	46.5	47.4	55.8	
18	204	BEN	Benin	Sub-Saharan Africa	14111034	4501	37.7	37.4	51.0	

Line 14 - summary of variables in dataframe

```
> # print summary of variables in dataframe
> summary(epi.data$ECO.new)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  23.10   42.27   49.75   51.10   60.08   83.60
```

Line 15 - Print sd value for ECO.new

```
> sd(epi.data$ECO.new)
[1] 13.16001
```

Line 17 - Print values in variable

```
> epi.data$ECO.new
  [1] 31.2 51.8 42.2 43.2 53.2 47.1 47.8 63.3 78.2 44.4 53.9 38.6 31.4 35.0 68.1 69.1 57.3 54.6
 [19] 59.5 55.7 51.3 74.0 63.8 48.9 70.8 56.9 48.1 23.1 44.0 48.1 60.6 56.1 49.4 58.4 35.9 56.4
 [37] 44.4 62.5 48.9 72.8 49.8 57.2 78.0 53.1 63.5 24.4 40.6 56.8 56.7 51.7 44.7 43.7 25.1 76.6
 [55] 38.7 45.9 36.9 68.4 68.4 64.8 35.5 50.2 80.5 46.9 67.9 40.4 38.6 47.4 44.8 56.2 36.8 49.0
 [73] 73.8 60.9 30.5 39.3 45.9 33.1 67.5 43.6 63.4 49.7 59.9 50.1 49.2 49.1 45.0 54.8 43.3 40.0
 [91] 68.6 38.1 46.3 30.0 74.3 83.6 27.7 53.8 39.7 39.4 43.2 65.5 34.9 33.7 34.3 47.7 27.6 48.4
[109] 54.4 50.2 40.7 47.8 26.6 62.0 47.4 67.8 51.3 58.5 57.0 43.3 55.1 72.6 65.6 29.4 59.1 35.5
[127] 43.6 56.4 33.7 79.3 63.4 57.4 63.8 68.4 48.2 44.5 45.1 50.6 37.2 31.6 50.3 50.9 56.1 48.3
[145] 38.9 55.9 77.8 67.7 30.2 49.8 49.9 68.5 39.3 42.3 63.9 67.3 69.4 51.8 46.2 59.6 50.8 47.6
[163] 45.9 34.0 48.9 48.1 35.6 40.8 44.5 58.5 63.5 73.3 54.1 39.1 49.3 36.5 61.0 27.7 65.3 66.3
```

Line 27 - Print values in variables

```
> ECO.new
  [1] 31.2 51.8 42.2 43.2 53.2 47.1 47.8 63.3 78.2 44.4 53.9 38.6 31.4 35.0 68.1 69.1 57.3 54.6
 [19] 59.5 55.7 51.3 74.0 63.8 48.9 70.8 56.9 48.1 23.1 44.0 48.1 60.6 56.1 49.4 58.4 35.9 56.4
 [37] 44.4 62.5 48.9 72.8 49.8 57.2 78.0 53.1 63.5 24.4 40.6 56.8 56.7 51.7 44.7 43.7 25.1 76.6
 [55] 38.7 45.9 36.9 68.4 68.4 64.8 35.5 50.2 80.5 46.9 67.9 40.4 38.6 47.4 44.8 56.2 36.8 49.0
 [73] 73.8 60.9 30.5 39.3 45.9 33.1 67.5 43.6 63.4 49.7 59.9 50.1 49.2 49.1 45.0 54.8 43.3 40.0
 [91] 68.6 38.1 46.3 30.0 74.3 83.6 27.7 53.8 39.7 39.4 43.2 65.5 34.9 33.7 34.3 47.7 27.6 48.4
[109] 54.4 50.2 40.7 47.8 26.6 62.0 47.4 67.8 51.3 58.5 57.0 43.3 55.1 72.6 65.6 29.4 59.1 35.5
[127] 43.6 56.4 33.7 79.3 63.4 57.4 63.8 68.4 48.2 44.5 45.1 50.6 37.2 31.6 50.3 50.9 56.1 48.3
[145] 38.9 55.9 77.8 67.7 30.2 49.8 49.9 68.5 39.3 42.3 63.9 67.3 69.4 51.8 46.2 59.6 50.8 47.6
[163] 45.9 34.0 48.9 48.1 35.6 40.8 44.5 58.5 63.5 73.3 54.1 39.1 49.3 36.5 61.0 27.7 65.3 66.3
```

Line 40 - print true if there are NAs in variable, false otherwise (printed 0 which is false)

```
> # find NAs in variable - outputs vector of logical values, true if NA, false otherwise
> NAs <- is.na(ECO)
> ECO[which(NAs)]
numeric(0)
```

Line 45 - print values in variable

```
> BDH
  [1] 32.1 50.6 33.0 41.5 52.8 35.0 47.4 55.4 74.4 36.9 46.8 26.0 29.1 12.5 70.3 66.4 58.4 63.7
 [19] 67.2 63.6 45.7 85.8 62.2 47.4 69.1 73.3 51.7 19.8 57.3 45.0 52.1 71.0 60.1 42.5  9.5 55.5
 [37] 49.9 63.9 56.3 69.8 43.4 51.6 78.7 51.8 53.3 18.1 27.8 57.6 50.3 47.4 36.3 45.6 16.8 78.8
 [55] 30.7 46.0 16.2 58.7 61.6 63.9 38.3 44.3 82.5 45.0 62.7 21.4 36.8 61.4 54.2 56.1 30.6 51.0
 [73] 67.0 54.8 11.4 31.5 42.9 20.2 62.9 30.0 58.9 38.9 47.5 32.9 50.0 43.9 31.4 50.6 36.5 50.8
 [91] 68.3 24.1 60.4 26.5 74.8 84.9 27.0 68.3 28.8 12.0 50.8 67.7 13.4 36.2 14.6 32.5  5.0 53.3
[109] 63.4 42.1 30.4 57.3 23.4 69.8 55.6 61.0 39.6 66.1 69.7 47.1 52.8 71.6 56.7 25.7 57.0 19.8
[127] 47.8 48.9 25.6 81.3 60.4 50.2 71.4 71.9 41.0 49.9 30.3 35.6 25.9 33.6 45.4 56.5 53.5 51.8
[145] 46.0 36.6 81.8 64.8 13.2 40.1 32.8 67.3 33.7 39.1 64.3 59.9 60.0 39.4 53.6 65.3 46.2 45.3
[163] 54.7 15.3 49.5 38.3 20.1 39.5 51.9 53.8 59.3 71.9 41.0 29.4 44.4 18.1 61.3 25.4 83.7 70.5
```

Line 51 - Prints a list of NAs from variable (there were none)

```
> # find NAs inv variavle - outputs vector of logical values, true if NA, false otherwise
> NAs <- is.na(BDH)
> # print NAs
> BDH[which(NAs)]
numeric(0)
```

Line 56 - Prints a list of non NA values from variable

```
> BDH.noNA
  [1] 32.1 50.6 33.0 41.5 52.8 35.0 47.4 55.4 74.4 36.9 46.8 26.0 29.1 12.5 70.3 66.4 58.4 63.7
 [19] 67.2 63.6 45.7 85.8 62.2 47.4 69.1 73.3 51.7 19.8 57.3 45.0 52.1 71.0 60.1 42.5  9.5 55.5
 [37] 49.9 63.9 56.3 69.8 43.4 51.6 78.7 51.8 53.3 18.1 27.8 57.6 50.3 47.4 36.3 45.6 16.8 78.8
 [55] 30.7 46.0 16.2 58.7 61.6 63.9 38.3 44.3 82.5 45.0 62.7 21.4 36.8 61.4 54.2 56.1 30.6 51.0
 [73] 67.0 54.8 11.4 31.5 42.9 20.2 62.9 30.0 58.9 38.9 47.5 32.9 50.0 43.9 31.4 50.6 36.5 50.8
 [91] 68.3 24.1 60.4 26.5 74.8 84.9 27.0 68.3 28.8 12.0 50.8 67.7 13.4 36.2 14.6 32.5  5.0 53.3
[109] 63.4 42.1 30.4 57.3 23.4 69.8 55.6 61.0 39.6 66.1 69.7 47.1 52.8 71.6 56.7 25.7 57.0 19.8
[127] 47.8 48.9 25.6 81.3 60.4 50.2 71.4 71.9 41.0 49.9 30.3 35.6 25.9 33.6 45.4 56.5 53.5 51.8
[145] 46.0 36.6 81.8 64.8 13.2 40.1 32.8 67.3 33.7 39.1 64.3 59.9 60.0 39.4 53.6 65.3 46.2 45.3
[163] 54.7 15.3 49.5 38.3 20.1 39.5 51.9 53.8 59.3 71.9 41.0 29.4 44.4 18.1 61.3 25.4 83.7 70.5
```

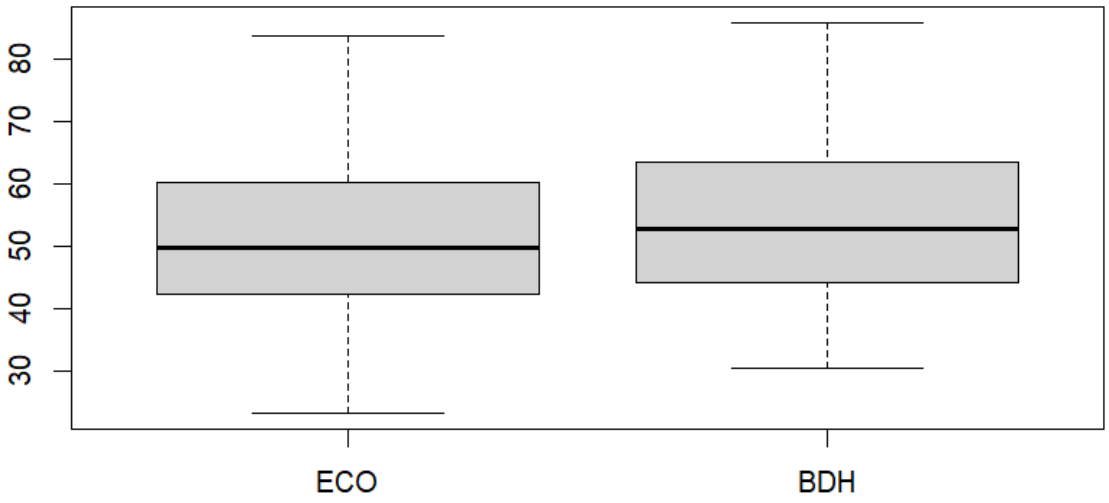
Line 61 - Prints list of values above 30 from variable

```
> BDH.above30
  [1] 32.1 50.6 33.0 41.5 52.8 35.0 47.4 55.4 74.4 36.9 46.8 70.3 66.4 58.4 63.7 67.2 63.6 45.7
 [19] 85.8 62.2 47.4 69.1 73.3 51.7 57.3 45.0 52.1 71.0 60.1 42.5 55.5 49.9 63.9 56.3 69.8 43.4
 [37] 51.6 78.7 51.8 53.3 57.6 50.3 47.4 36.3 45.6 78.8 30.7 46.0 58.7 61.6 63.9 38.3 44.3 82.5
 [55] 45.0 62.7 36.8 61.4 54.2 56.1 30.6 51.0 67.0 54.8 31.5 42.9 62.9 58.9 38.9 47.5 32.9 50.0
 [73] 43.9 31.4 50.6 36.5 50.8 68.3 60.4 74.8 84.9 68.3 50.8 67.7 36.2 32.5 53.3 63.4 42.1 30.4
 [91] 57.3 69.8 55.6 61.0 39.6 66.1 69.7 47.1 52.8 71.6 56.7 57.0 47.8 48.9 81.3 60.4 50.2 71.4
[109] 71.9 41.0 49.9 30.3 35.6 33.6 45.4 56.5 53.5 51.8 46.0 36.6 81.8 64.8 40.1 32.8 67.3 33.7
[127] 39.1 64.3 59.9 60.0 39.4 53.6 65.3 46.2 45.3 54.7 49.5 38.3 39.5 51.9 53.8 59.3 71.9 41.0
[145] 44.4 61.3 83.7 70.5
```

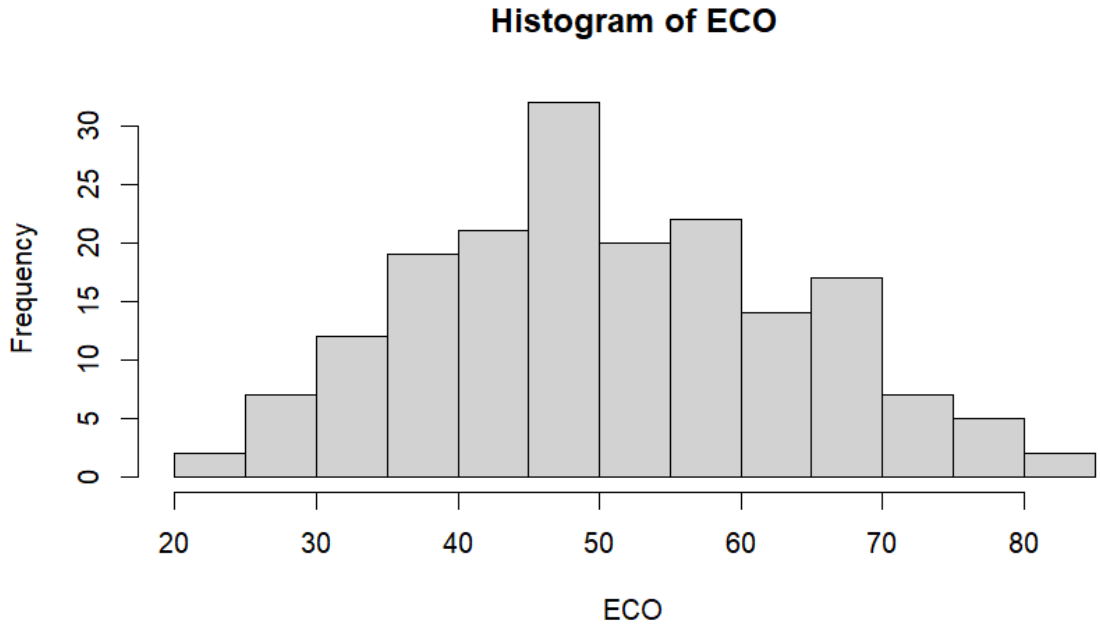
Line 64 - summary of values above 30 in the variable

```
> # stats
> summary(BDH.above30)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  30.30   44.20   52.80   53.62   63.45   85.80
```

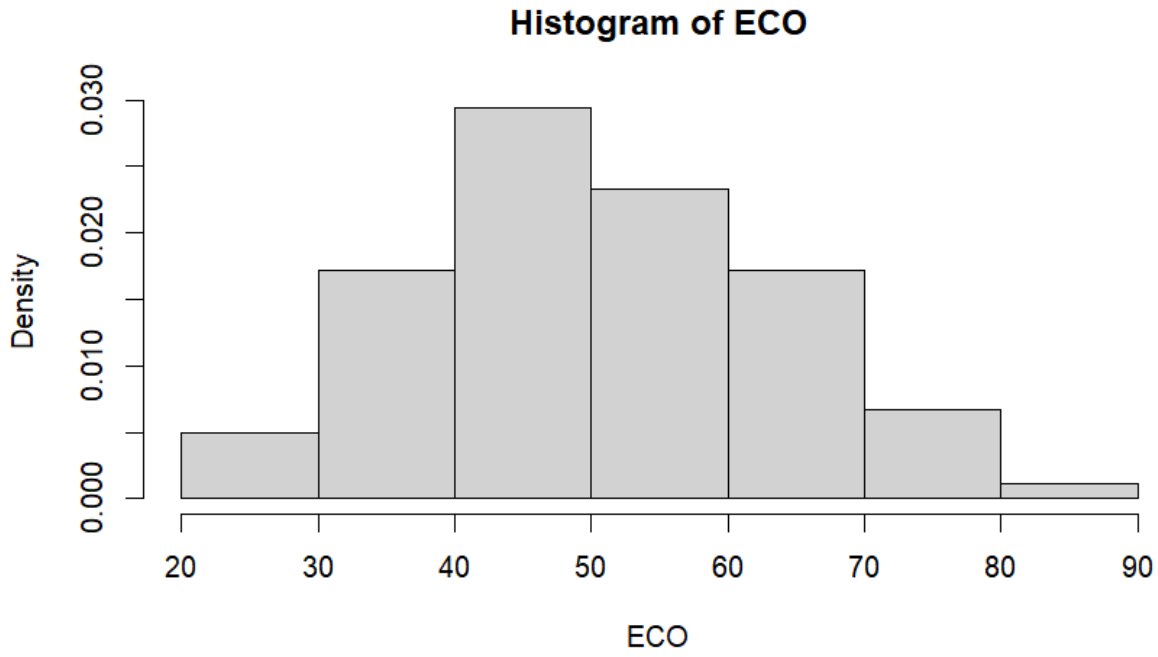
Line 67 - Create a boxplot of the variables



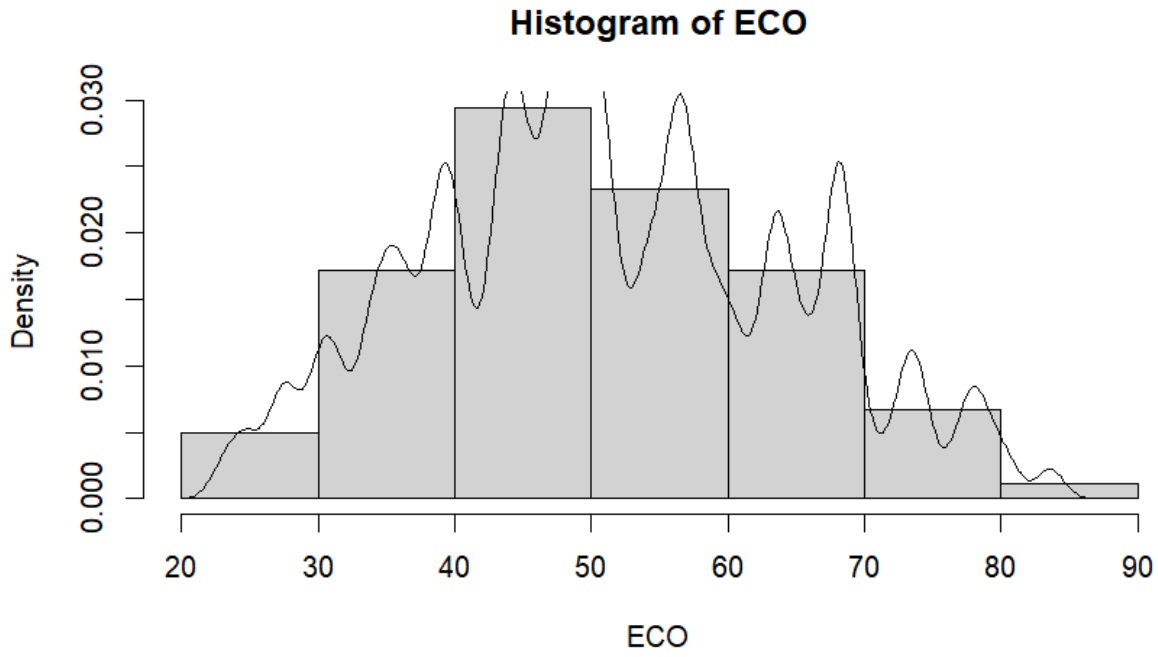
Line 73 - Create a histogram of the ECO variable



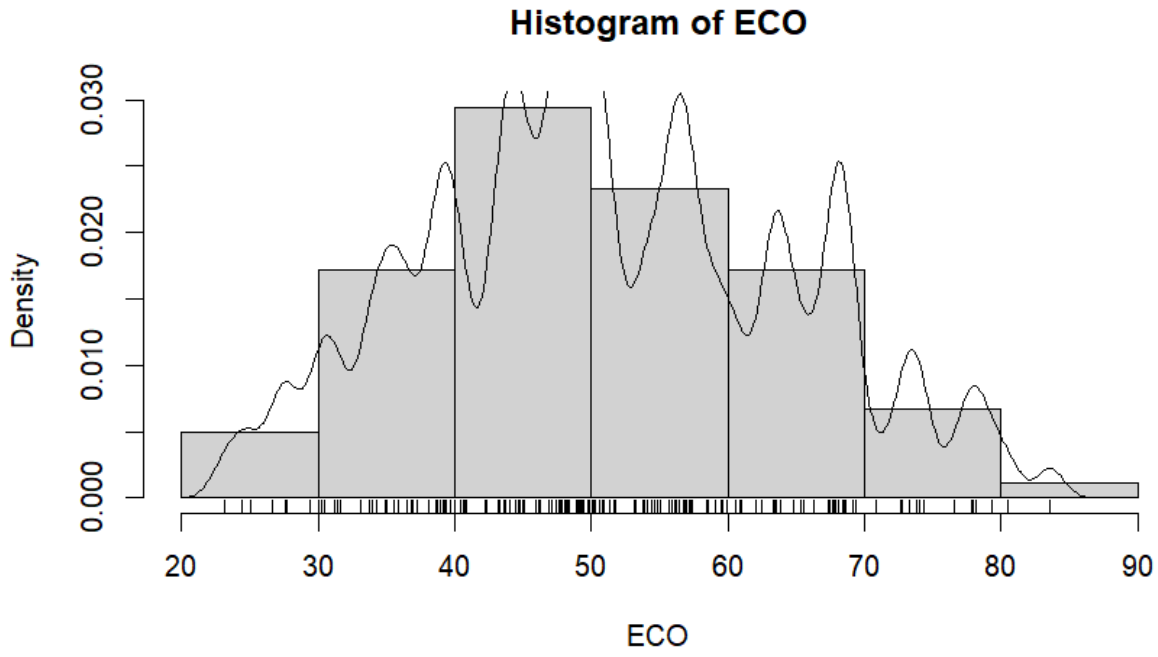
Line 79 - Create a histogram of the ECO variable over the range 20-90 and in segments of 10



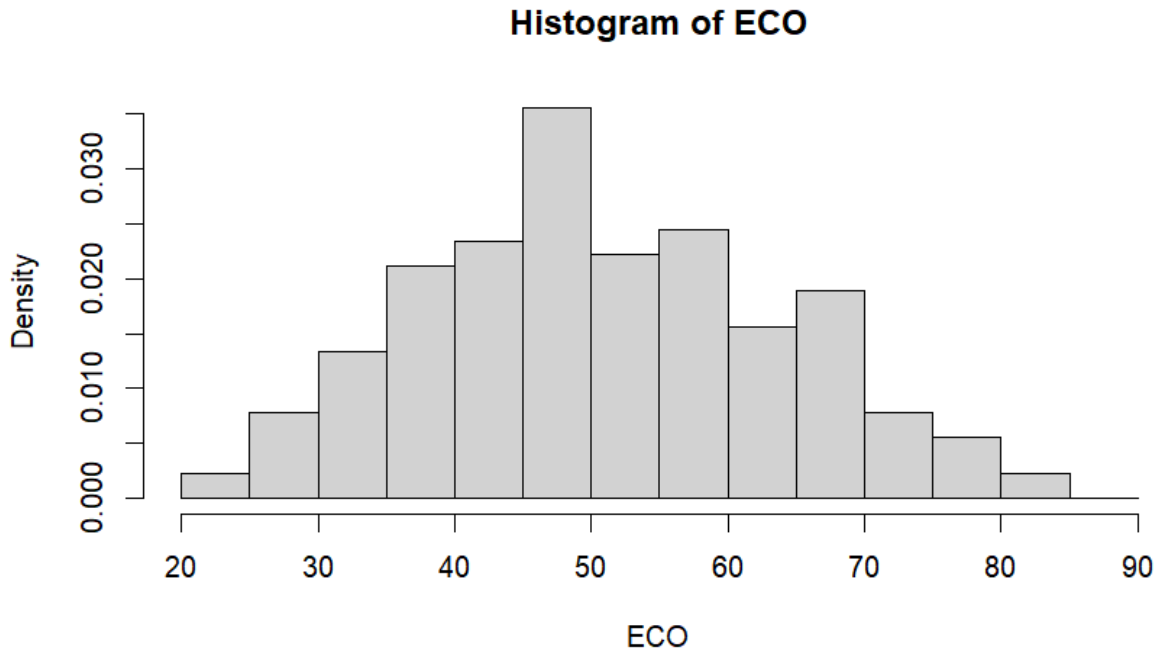
Line 82 - Print estimated density curve for histogram



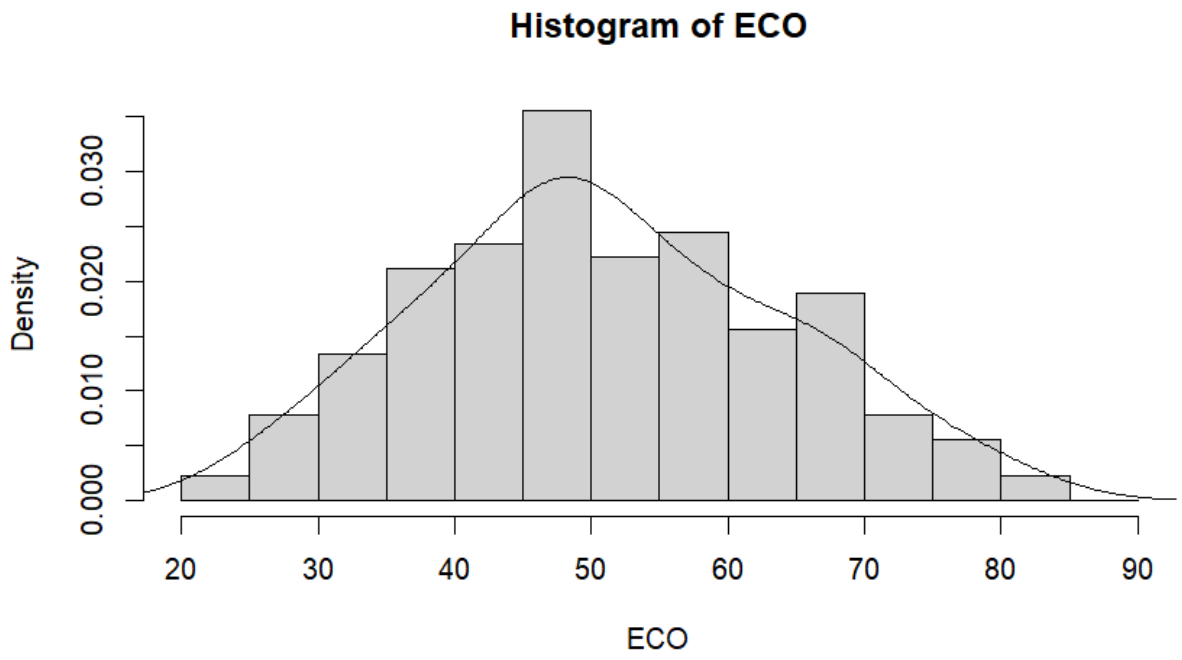
Line 85 - Prints the rug for histogram



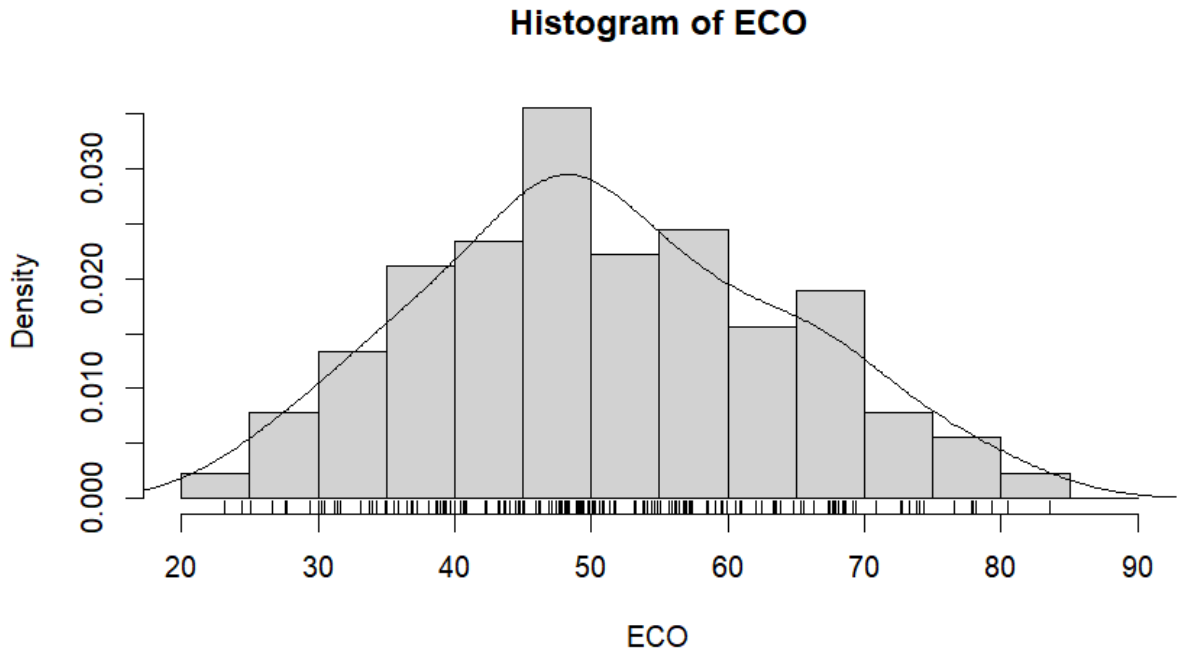
Line 90 - Create a histogram of the ECO variable over the range 20-90 and in segments of 10



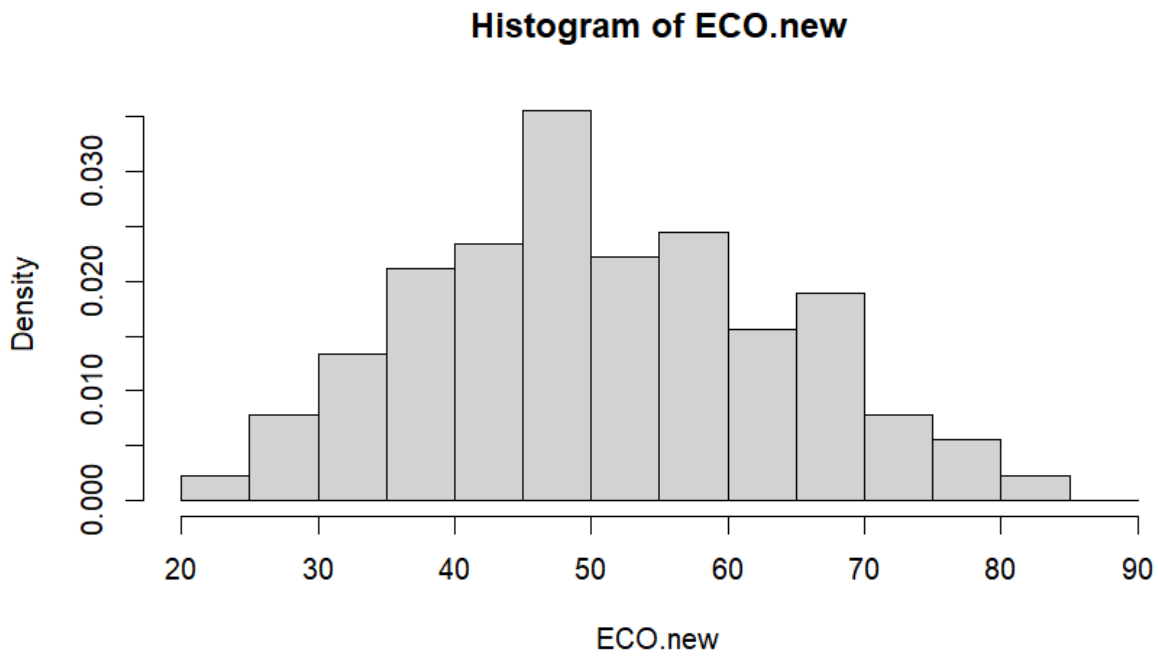
Line 93 - Print estimated density curve for variable



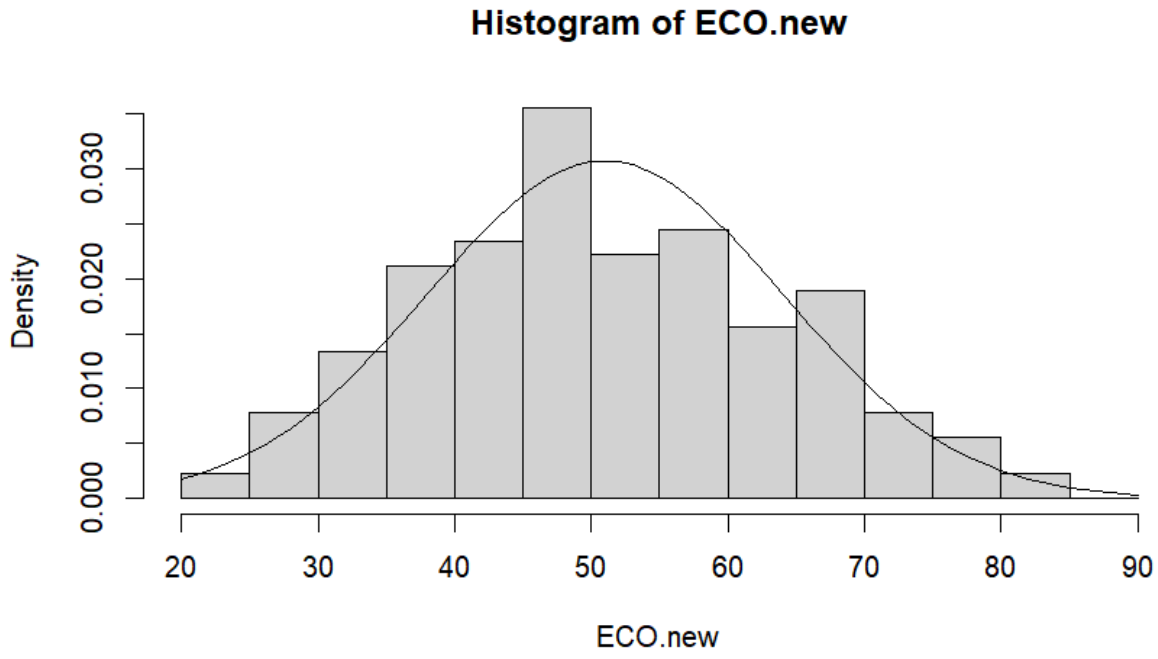
Line 96 - Print the rug for histogram



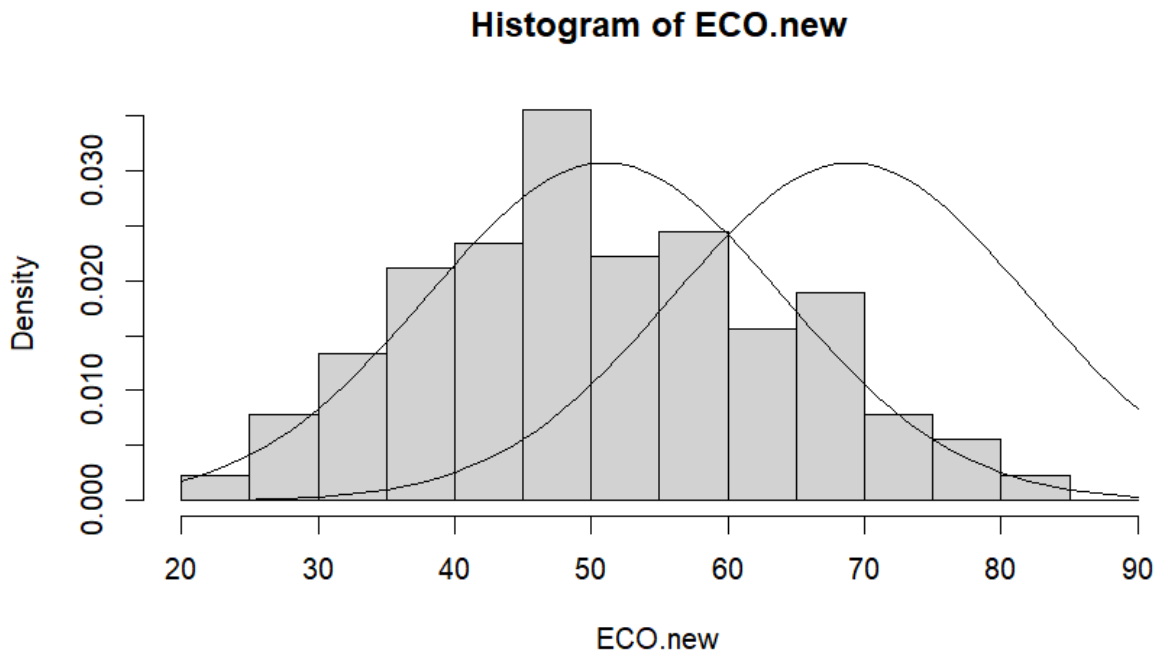
Line 100 - Create a histogram of the ECO.new variable over the range 20-90 and in segments of 10



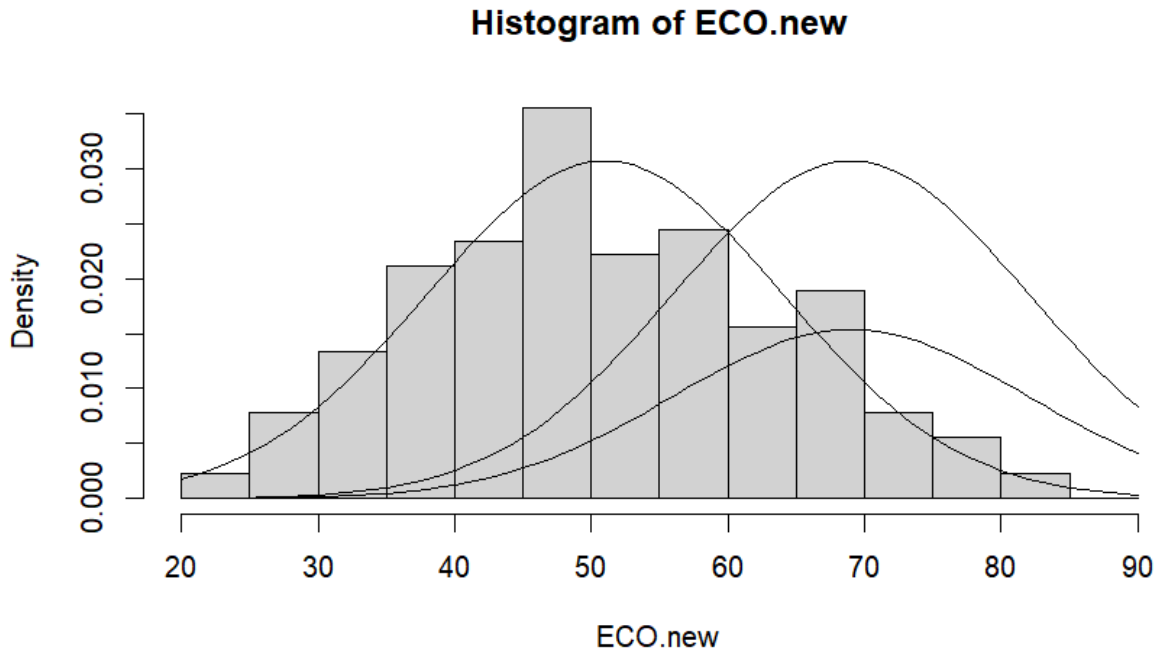
Line 109 - Print density curve for variable



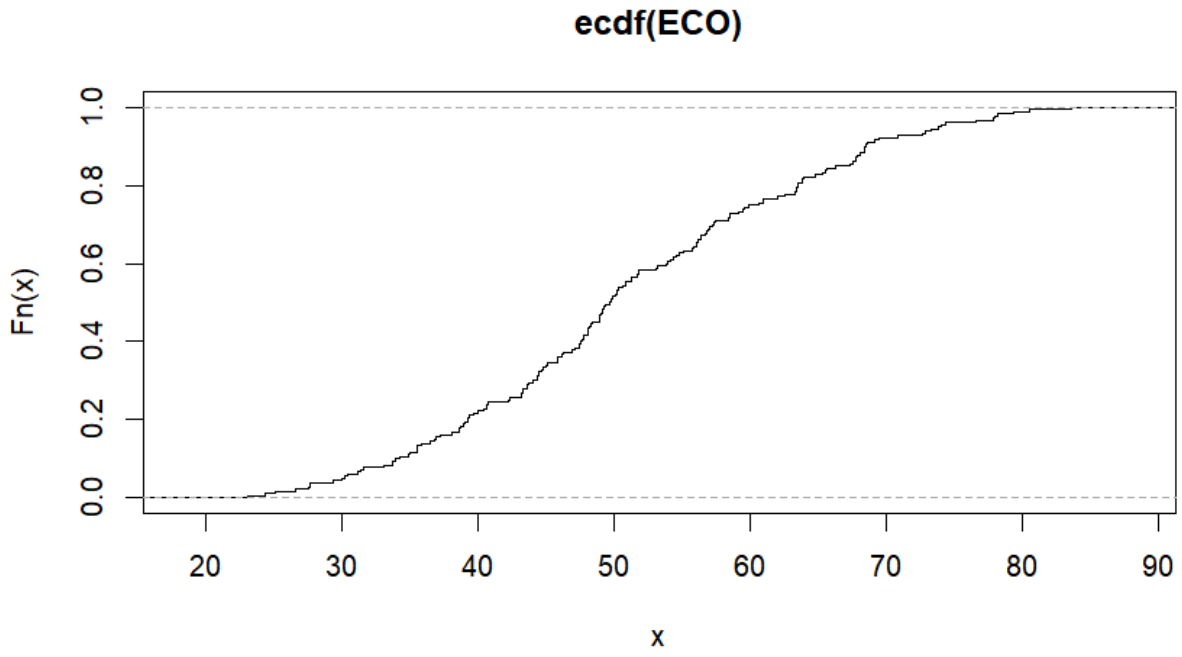
Line 115 - Print 2nd density curve



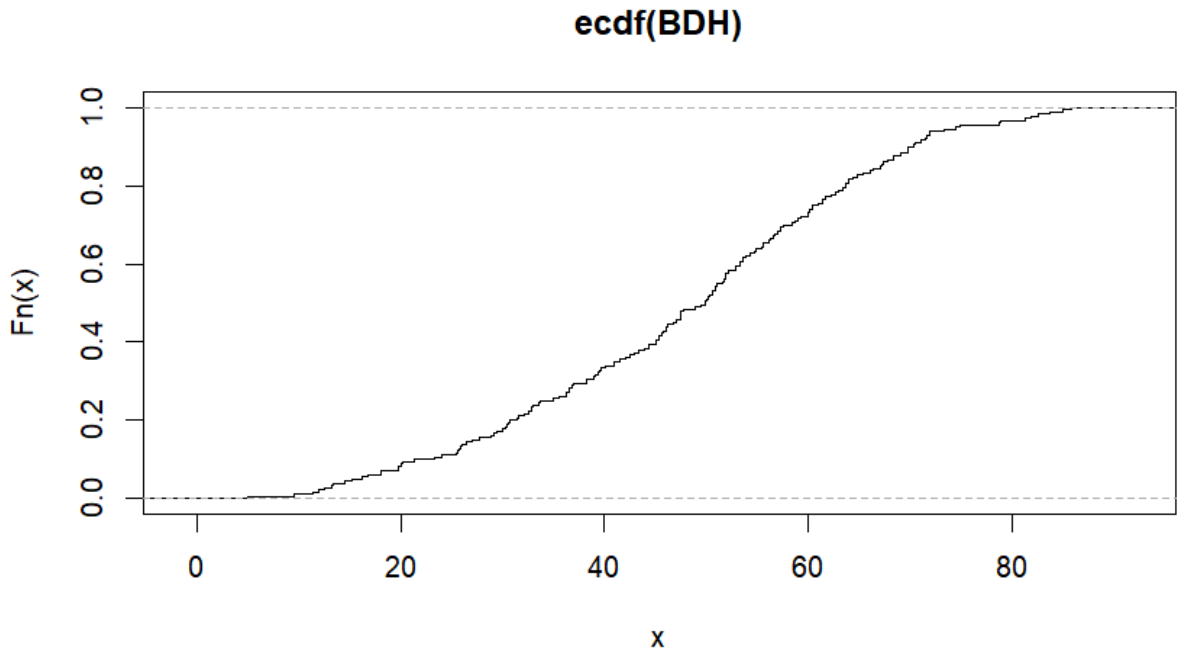
Line 118 - Plot 3rd density curve



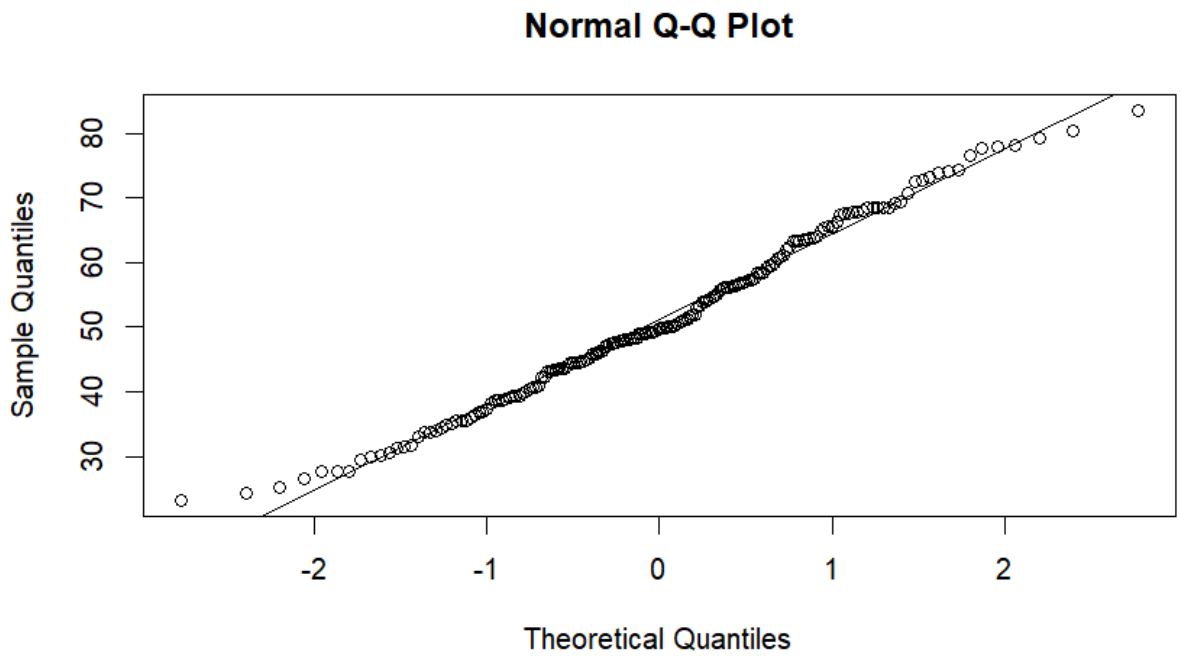
Line 123 - Print empirical cumulative distribution function (ecdf) plot for ECO



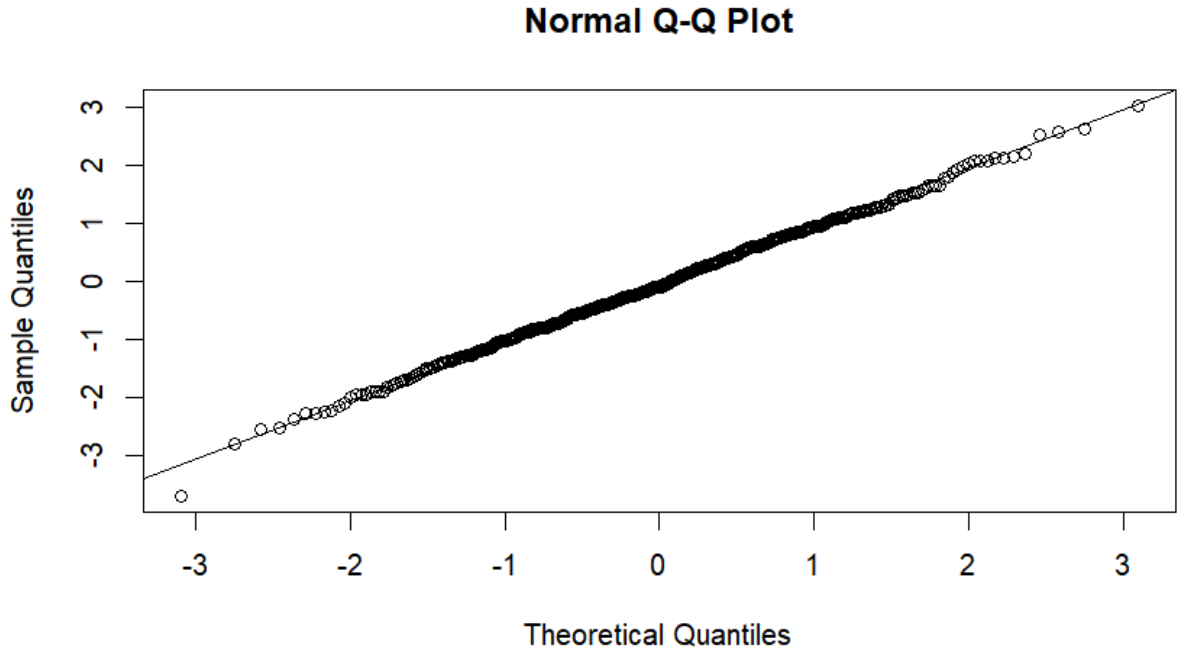
Line 125 - Print empirical cumulative distribution function (ecdf) plot for BDH



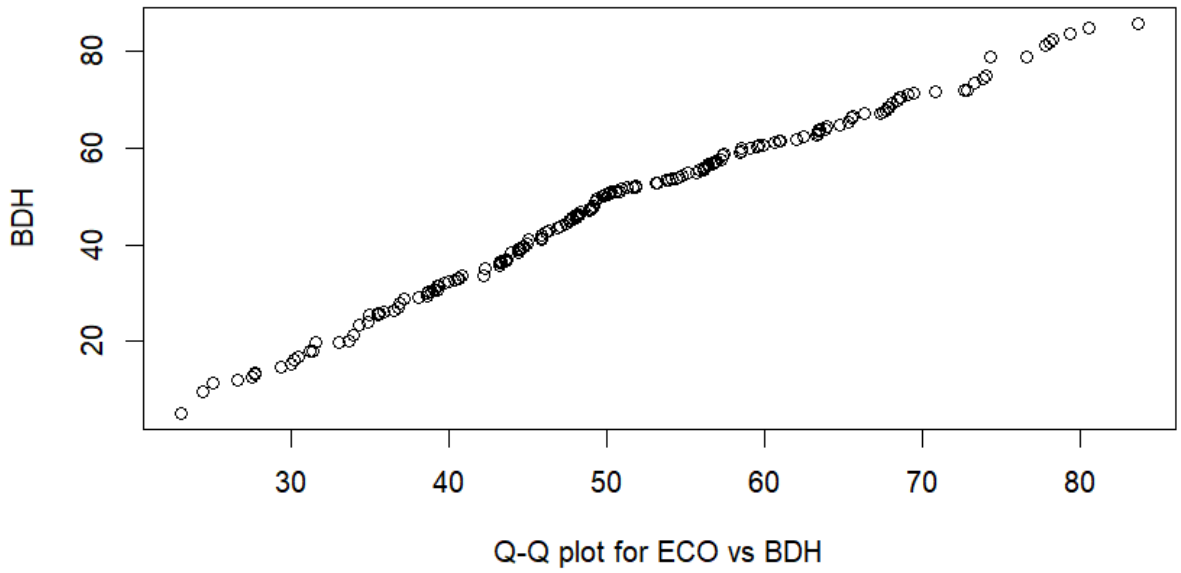
Line 131 - print quantile quantile plot for variable with theoretical normal distribution

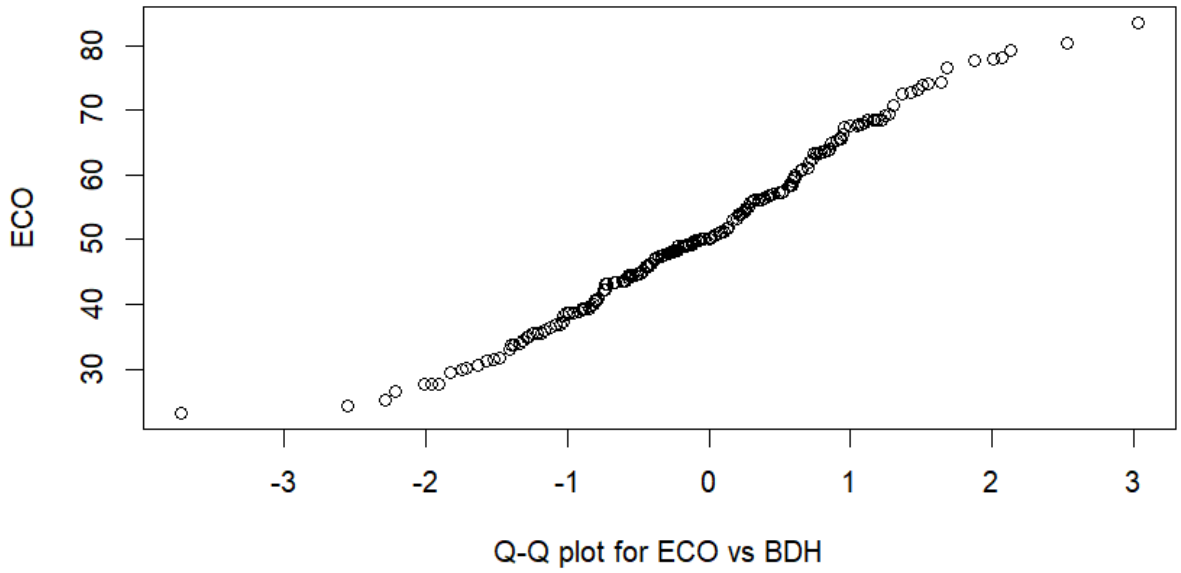


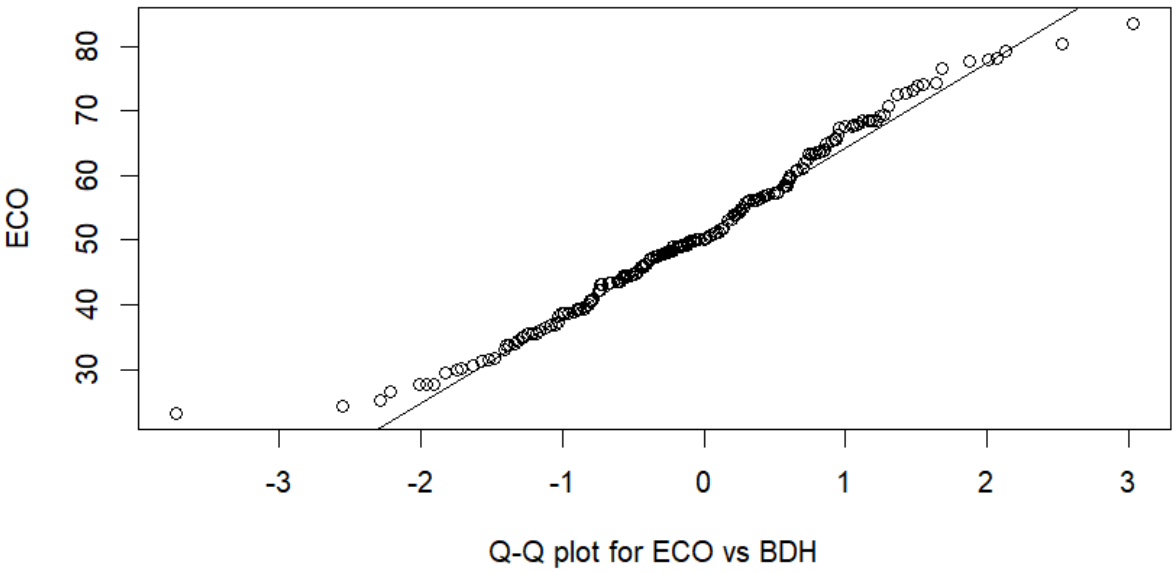
Line 136 - Print quantile quantile plot for random numbers from a normal distribution with theoretical normal distribution

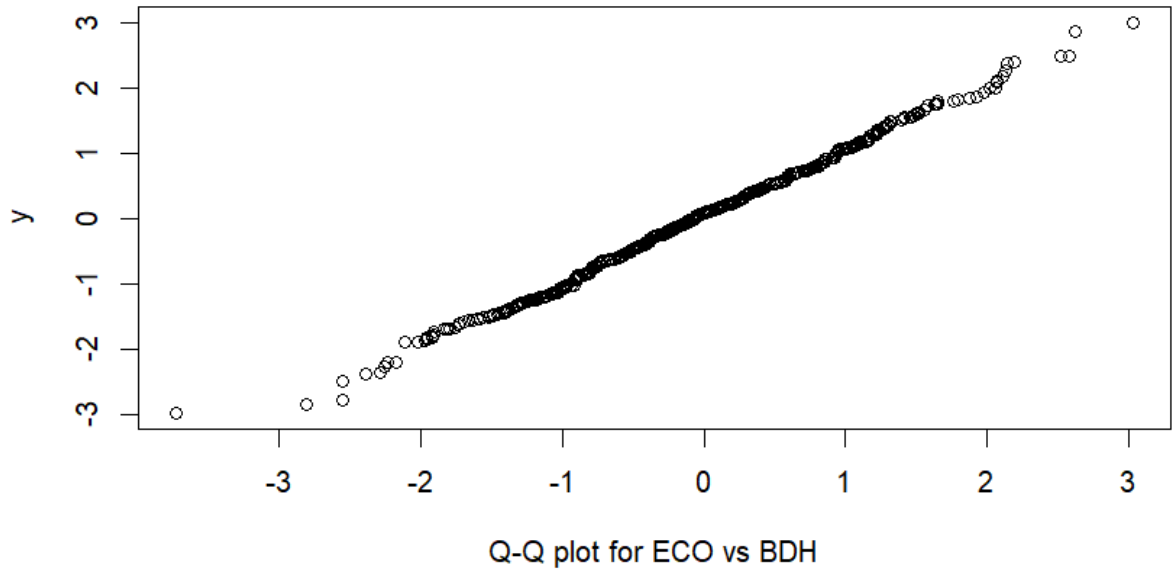


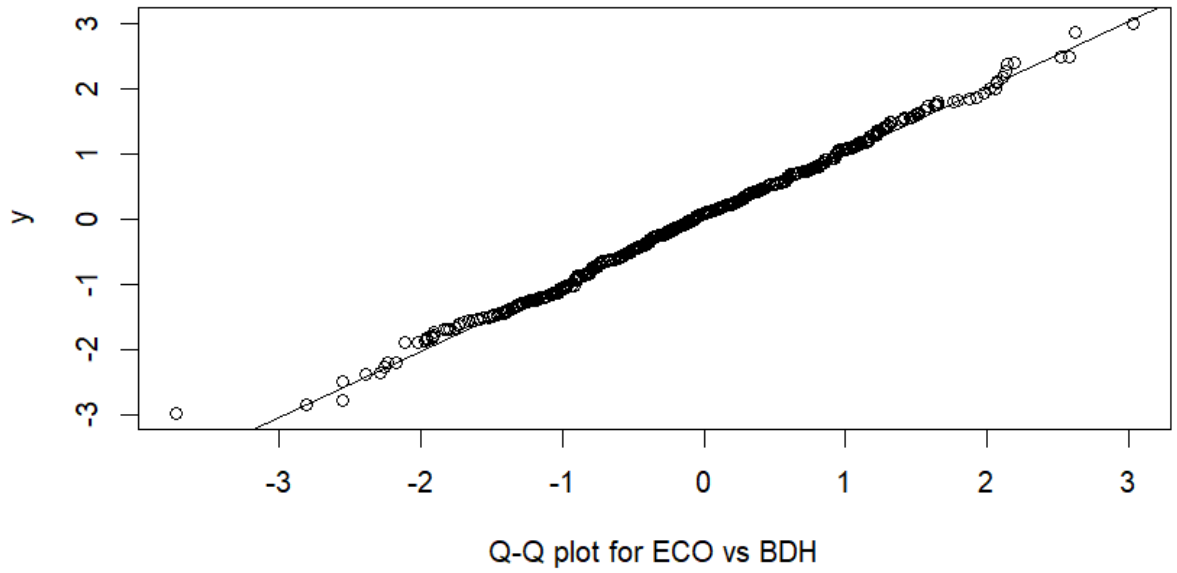
Line 144, 146, 147, 151, 152 - Print quantile quantile value for 2 variables (EPI vs MHP)


BDH
20
40
60
80
30
40
50
60
70
80
Q-Q plot for ECO vs BDH

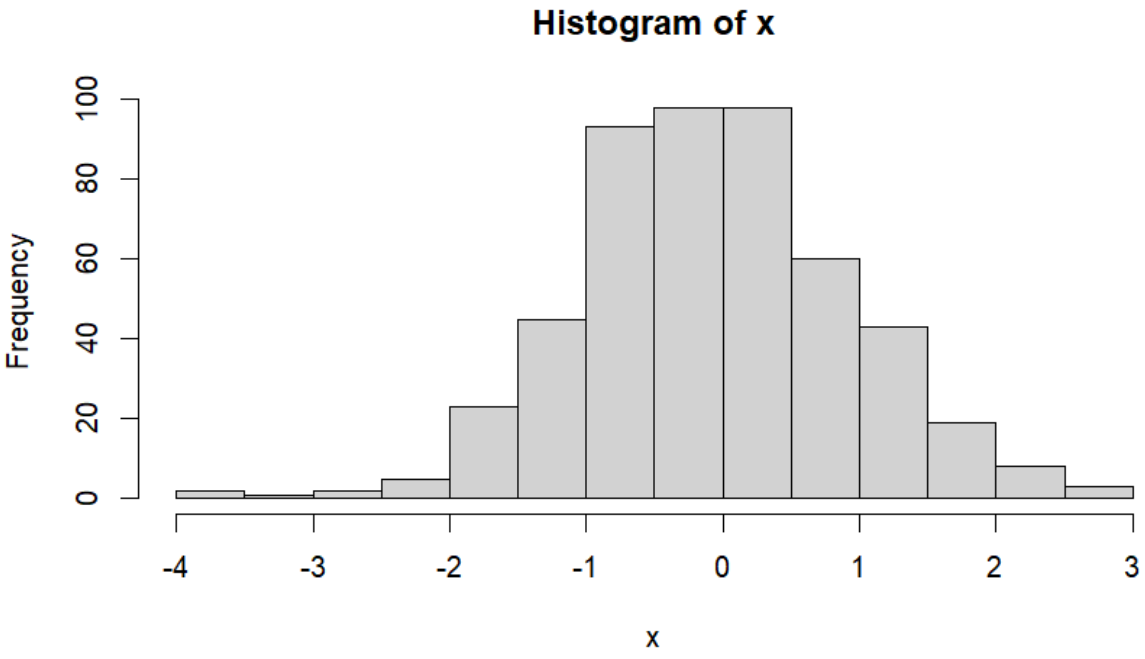

ECO
30
40
50
60
70
80
-3
-2
-1
0
1
2
3
Q-Q plot for ECO vs BDH


ECO
30
40
50
60
70
80
-3
-2
-1
0
1
2
3
Q-Q plot for ECO vs BDH

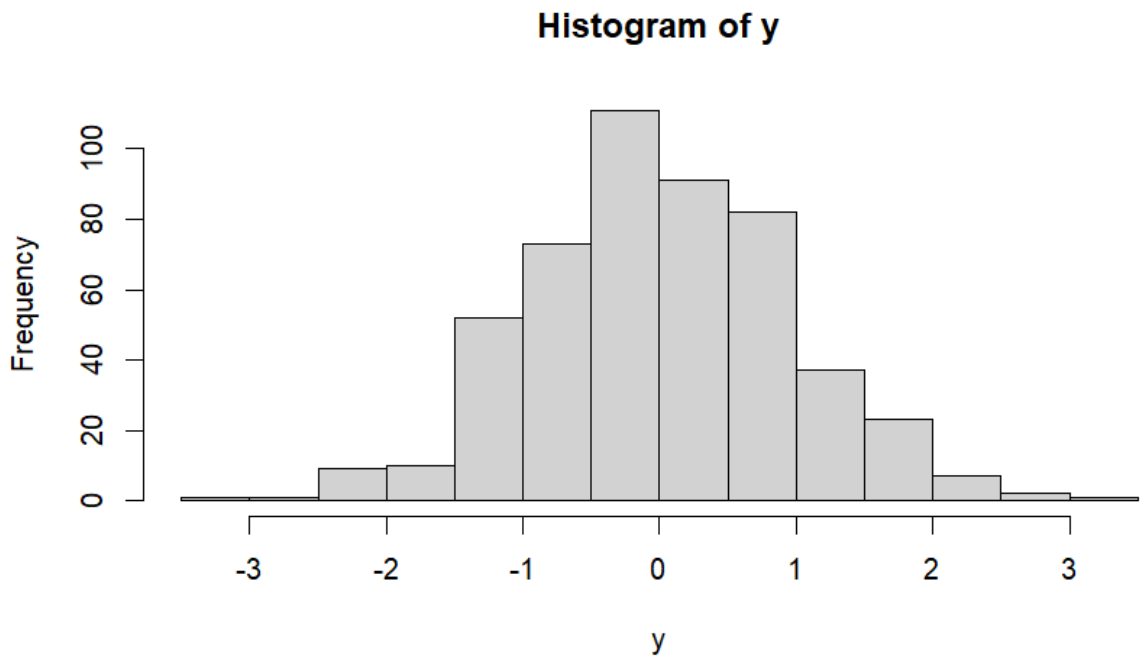

y
-3
-2
-1
0
1
2
3
-3
-2
-1
0
1
2
3
Q-Q plot for ECO vs BDH



Line 160 - Histogram of X variable



Line 161 - Histogram of Y variable



Line 163 - Print Shapiro test for X variable

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.99522, p-value = 0.1266
```

Line 164 - Print Shapiro test for Y variable

```
> shapiro.test(y)

        Shapiro-Wilk normality test

data:  y
W = 0.99847, p-value = 0.945
```

Line 166 and 167 - Produce errors, could not find function “ad.test()”

Line 169 - Print Kolmogorov-Smirnov test for variables x and y

```
> ks.test(x,y)

        Asymptotic two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.056, p-value = 0.4131
alternative hypothesis: two-sided
```

Line 171 - Print wilcox test for variables x and y

```
> wilcox.test(x,y)

        Wilcoxon rank sum test with continuity correction

data:  x and y
W = 118921, p-value = 0.1832
alternative hypothesis: true location shift is not equal to 0
```

Line 173 - Prints F test to compare variances in x and y variables

```
> var.test(x,y)

        F test to compare two variances

data:  x and y
F = 1.0716, num df = 499, denom df = 499, p-value = 0.4399
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.8989892 1.2774644
sample estimates:
ratio of variances
          1.071647
```

Line 174 - Prints Welch 2 sample t test for variables x and y

```
> t.test(x,y)

        Welch Two Sample t-test

data:  x and y
t = -1.3476, df = 996.81, p-value = 0.1781
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.20380369  0.03785535
sample estimates:
   mean of x    mean of y
-0.076193691  0.006780477
```