

```

1 #####BOSTON DATASET#####
2 install.packages("MASS") # installing the MASS package
3 library(MASS) # load the library MASS
4 attach(Boston) # attaching the dataset
5

```

Installing package into ‘/usr/local/lib/R/site-library’
(as ‘lib’ is unspecified)

```

1 ?Boston # help function with "?"
2 head(Boston) # show the head of the dataset
3

```

A data.frame: 6 × 14

	crim	zn	indus	chas	nox	rm	age	dis	rad	tax	ptratio
	<dbl>	<dbl>	<dbl>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<dbl>	<dbl>
1	0.00632	18	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3
2	0.02731	0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8
3	0.02729	0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8
4	0.03237	0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7
5	0.06905	0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7
6	0.02985	0	2.18	0	0.458	6.430	58.7	6.0622	3	222	18.7

```

1 dim(Boston) # dimensions of the dataset
2

```

506 · 14

```

1 names(Boston) # column names
2

```

'crim' · 'zn' · 'indus' · 'chas' · 'nox' · 'rm' · 'age' · 'dis' · 'rad' · 'tax' · 'ptratio' · 'black' · 'lstat' · 'medv'

```

1 str(Boston) # str function shows the structure of the dataset
2

```

```

1 nrow(Boston) # function shows the number of rows
2

```

506

```
1 ncol(Boston) # function shows the number of columns
2
```

14

```
1 summary(Boston) # summary() function shows the summary
2
3
```

crim	zn	indus	chas
Min. : 0.00632	Min. : 0.00	Min. : 0.46	Min. : 0.00000
1st Qu.: 0.08205	1st Qu.: 0.00	1st Qu.: 5.19	1st Qu.: 0.00000
Median : 0.25651	Median : 0.00	Median : 9.69	Median : 0.00000
Mean : 3.61352	Mean : 11.36	Mean : 11.14	Mean : 0.06917
3rd Qu.: 3.67708	3rd Qu.: 12.50	3rd Qu.: 18.10	3rd Qu.: 0.00000
Max. : 88.97620	Max. : 100.00	Max. : 27.74	Max. : 1.00000

nox	rm	age	dis
Min. : 0.3850	Min. : 3.561	Min. : 2.90	Min. : 1.130
1st Qu.: 0.4490	1st Qu.: 5.886	1st Qu.: 45.02	1st Qu.: 2.100
Median : 0.5380	Median : 6.208	Median : 77.50	Median : 3.207
Mean : 0.5547	Mean : 6.285	Mean : 68.57	Mean : 3.795
3rd Qu.: 0.6240	3rd Qu.: 6.623	3rd Qu.: 94.08	3rd Qu.: 5.188
Max. : 0.8710	Max. : 8.780	Max. : 100.00	Max. : 12.127

rad	tax	ptratio	black
Min. : 1.000	Min. : 187.0	Min. : 12.60	Min. : 0.32
1st Qu.: 4.000	1st Qu.: 279.0	1st Qu.: 17.40	1st Qu.: 375.38
Median : 5.000	Median : 330.0	Median : 19.05	Median : 391.44
Mean : 9.549	Mean : 408.2	Mean : 18.46	Mean : 356.67
3rd Qu.: 24.000	3rd Qu.: 666.0	3rd Qu.: 20.20	3rd Qu.: 396.23
Max. : 24.000	Max. : 711.0	Max. : 22.00	Max. : 396.90

lstat	medv
Min. : 1.73	Min. : 5.00
1st Qu.: 6.95	1st Qu.: 17.02
Median : 11.36	Median : 21.20
Mean : 12.65	Mean : 22.53
3rd Qu.: 16.95	3rd Qu.: 25.00
Max. : 37.97	Max. : 50.00

```
1 summary(Boston$crim) # summary of the "crime" column in the
2
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00632	0.08204	0.25651	3.61352	3.67708	88.97620

```
1 #####MY experiments: Display summary by eliminating the NA values from computation
```

```
2
```

```
3
```

```
1 summary(Boston$rad, na.rm=T) #####for the rad column
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	4.000	5.000	9.549	24.000	24.000

 **Generate**

Using ...

a slider using jupyter widgets



Close

```
1 summary(Boston$rad) #####for the rad column
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.000	4.000	5.000	9.549	24.000	24.000

```
1 #####From here we work on the EPI DATASET#####
```

```
1 library(ggplot2)
```

```
2 #####READ THE DATASET#####
```

```
3 EPI<-read.csv("EPI_data.csv")
```

```
4 #####ACCESS THE EPI VARIABLE FROM THE DATAFRAME
```

```
5
```

```
6 summary(EPI$EPI)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
32.10	48.60	59.20	58.37	67.60	93.50	68

```
1 #####DO not consider the missing values in the summary statistics calculation
```

```
2 summary(EPI$EPI, na.rm = T)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
32.10	48.60	59.20	58.37	67.60	93.50	68

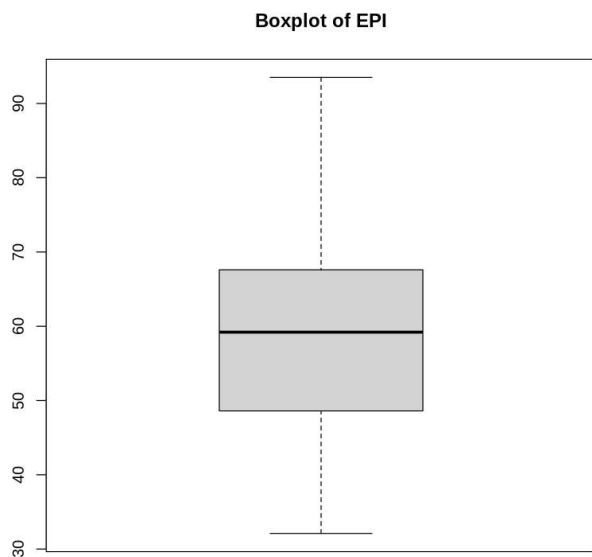
```
1 #####DISPLAY THE MEAN, MEDIAN, Minimum, Lower Hinge (Q1),Median (Q2), Upper Hinge (Q3) and Maximum
```

```
2 fivenum(EPI$EPI, na.rm = T)
```

```
32.1 • 48.6 • 59.2 • 67.6 • 93.5
```

```
1 #####TO PLOT THE BOXPLOT OF THE VARIABLE EPI
```

```
2 boxplot(EPI$EPI, main = "Boxplot of EPI",names = c("EPI"))
```



1 Start coding or [generate](#) with AI.

✓ 0s completed at 8:46 PM

