Week 4 – Thursday – PREPARATION SOLUTIONS

Please try on your own before looking at this!

Quantitative Reasoning 2020

```
x <- read.csv("India_AirQuality.csv")
```

Challenge 1: Consider Delhi. Calculate the following: the minimum, maximum, mean, median, and standard deviation of the Delhi PM2.5 measurements. This involves using functions: `min`, `max`, `mean`, `median`, and `sd`.

```
min(x$Delhi)
```

```
## [1] 6
```

```
max(x$Delhi)
```

```
## [1] 129
```

```
mean(x$Delhi)
```

```
## [1] 49.06528
```

```
median(x$Delhi)
```

```
## [1] 48
```

```
sd(x$Delhi)
```

```
## [1] 19.08148
```

Challenge 2: Now consider one of the other cities. Are there any missing PM2.5 measurements for this city? If so, how many?

```
# For example, Mumbai:
sum(is.na(x$Mumbai))     # Yes, looks like 70 missing values!
```

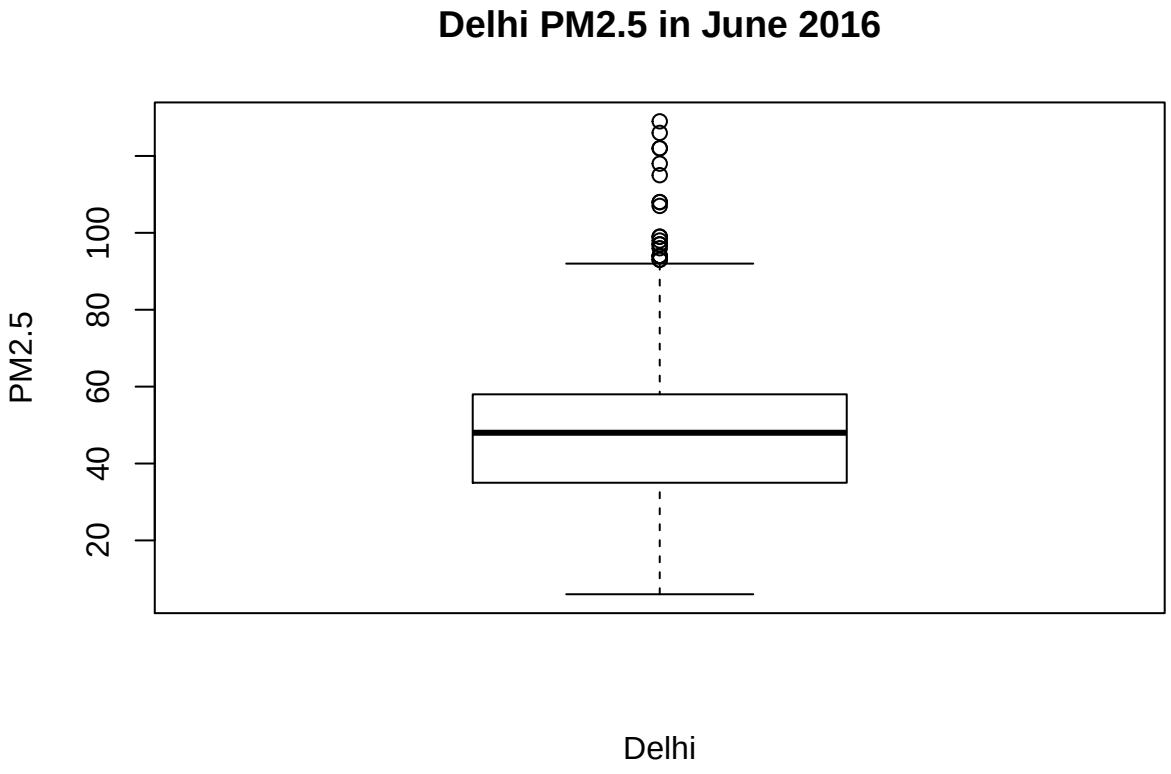
```
## [1] 0
```

```
table(is.na(x$Mumbai))   # And 650 non-missing values.
```

```
## < table of extent 0 >
```

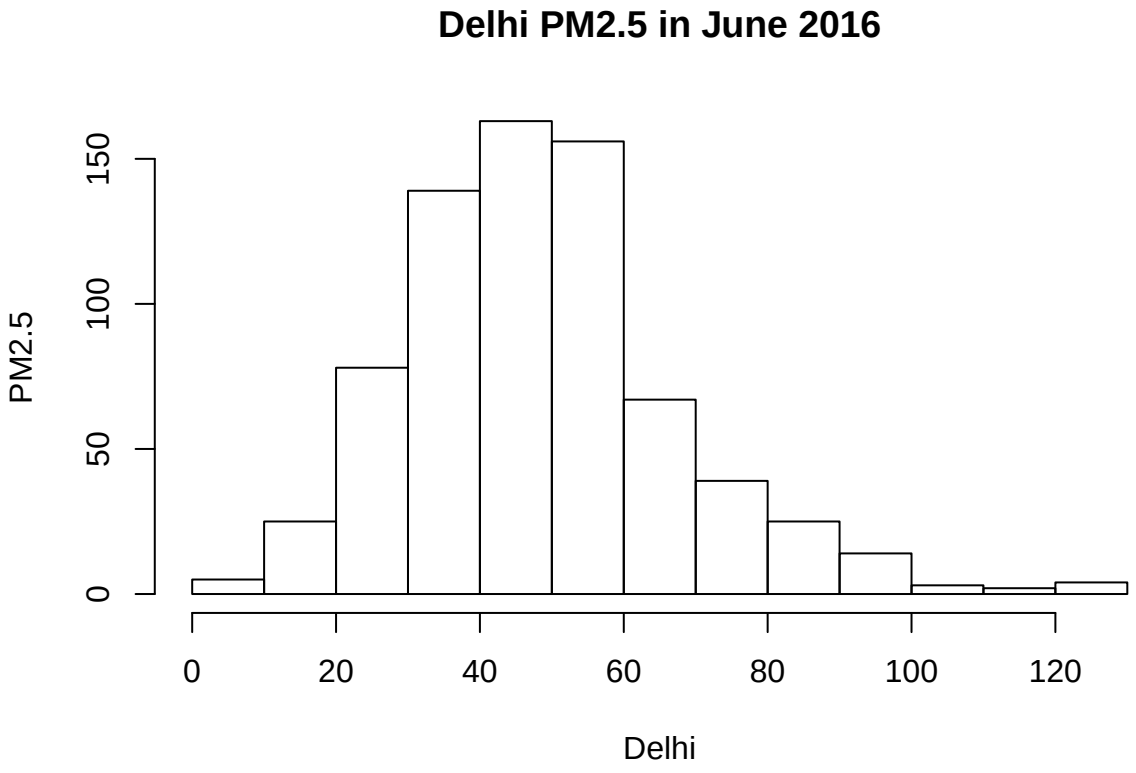
Challenge 3: Use `boxplot()` to explore with a graphic the PM2.5 measurements of Delhi. You might have to refer to the help page for boxplot by typing `?boxplot` in the console.

```
boxplot(x$Delhi, main="Delhi PM2.5 in June 2016",
        xlab="Delhi", ylab="PM2.5")
```



Challenge 4: Use `hist()` to explore with a graphic the PM2.5 measurements of Delhi. You might have to refer to the help page for boxplot by typing `?hist` in the console.

```
# Easy!  I just changed "boxplot" to "hist" here:
hist(x$Delhi, main="Delhi PM2.5 in June 2016",
        xlab="Delhi", ylab="PM2.5")
```



Challenge 5: Are any of Delhi's PM2.5 measurements below 0? In theory, the measurement should not contain any negative value.

```
sum(x$Delhi < 0)              # Phew!  Good to know.
```

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

```
sum(x$Delhi < 0, na.rm=TRUE) # The option is unnecessary because no NAs
```

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