# In Class Assignment 09.02.2025

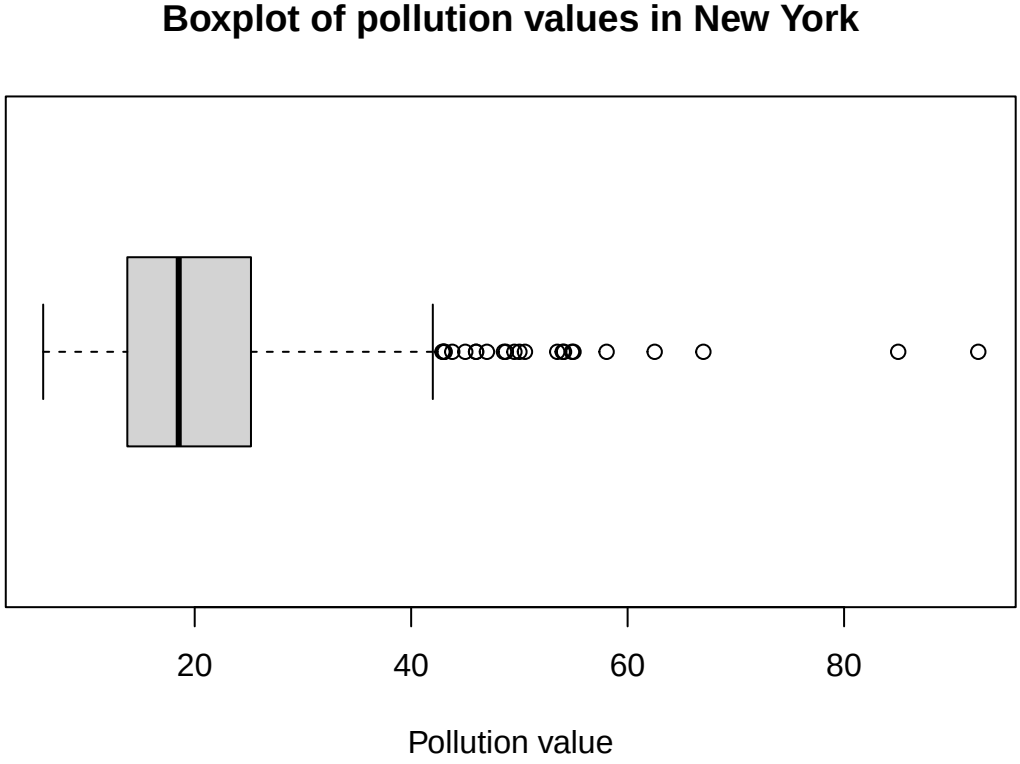
Akash Rajeev

2025-09-02

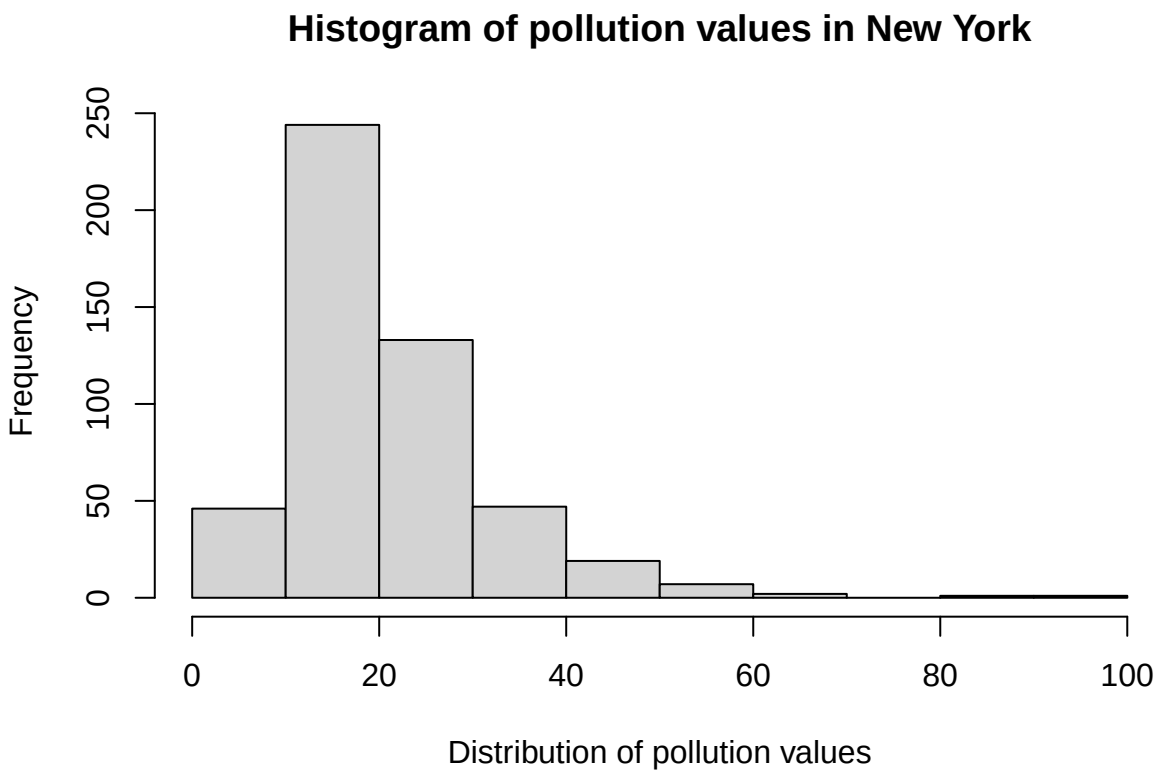
```
nypol = fread("ny_pollution.csv.gz")
nypol %>%
  head(10)
```

```
##           date death pollution
##         <IDat> <int>     <num>
##  1: 1997-01-01   211        NA
##  2: 1997-01-02   229        NA
##  3: 1997-01-03   231        NA
##  4: 1997-01-04   204  25.28571
##  5: 1997-01-05   208        NA
##  6: 1997-01-06   227        NA
##  7: 1997-01-07   189        NA
##  8: 1997-01-08   213        NA
##  9: 1997-01-09   237        NA
## 10: 1997-01-10   219  25.25000
```

```
nypolbox = boxplot(nypol$pollution, horizontal = TRUE, xlab = "Pollution value", main = "Boxplot of poll
```



```
nypolhist = hist(nypol$pollution, xlab = "Distribution of pollution values", main = "Histogram of pollu
```




```
mean(nypol$pollution, na.rm = T)
```

```
## [1] 21.0248
```

```
median(nypol$pollution, na.rm = T)
```

```
## [1] 18.525
```

## Statement and Supporting Premise

Statement: The Pollution data in this dataset is right skewed.

Supporting Premise: The histogram and boxplot demonstrate that the data may exhibit a skew to the right. Additionally, the mean is greater than the median, which may also indicate a slight right skew.