Assignment 2 - DY (submission)

Daniel Yim

2023-03-03

#1. Install the airquality dataset from the library datasets. (20 Points)

#(a) Display the first 6 rows of the airquality dataset.

```
library(datasets)
data(airquality)
head(airquality,6)
```

```
##   Ozone Solar.R Wind Temp Month Day
## 1    41     190  7.4   67     5   1
## 2    36     118  8.0   72     5   2
## 3    12     149 12.6   74     5   3
## 4    18     313 11.5   62     5   4
## 5    NA      NA 14.3   56     5   5
## 6    28      NA 14.9   66     5   6
```

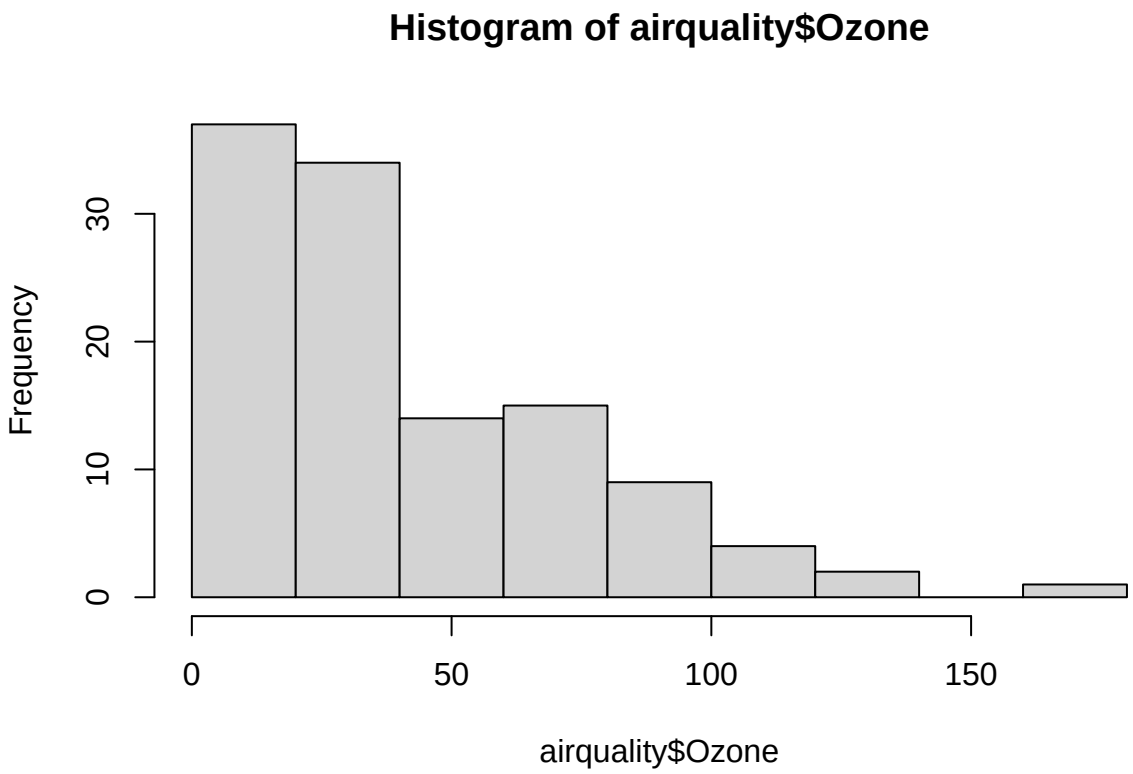
#(b) Display the class of each column of the airquality dataset.

```
sapply(airquality,class)
```

```
##     Ozone   Solar.R      Wind      Temp     Month       Day
## "integer" "integer" "numeric" "integer" "integer" "integer"
```

#(c) Use a histogram to assess normality of the Ozone variable. Does it #appear normally distributed?

```
library(ggplot2)
hist(airquality$Ozone)
```

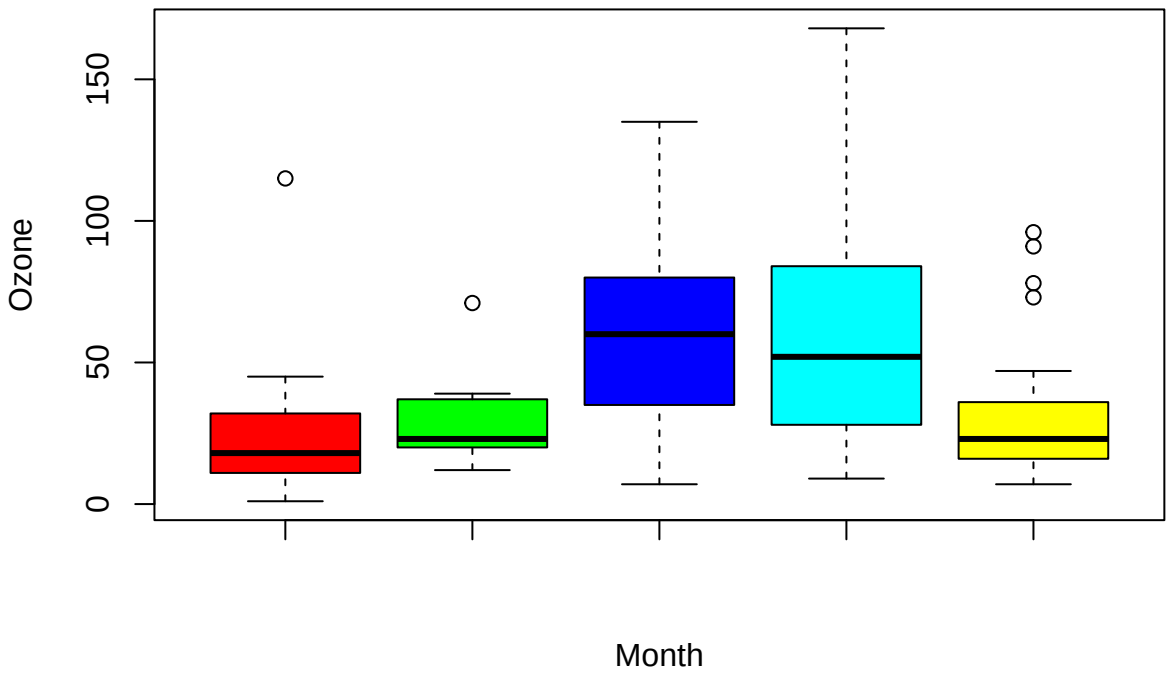


#Comment: it does not appear to be normally distributed. It looks right skewed/positively skewed.

(d) Create a boxplot which shows the distribution of Ozone in each month.

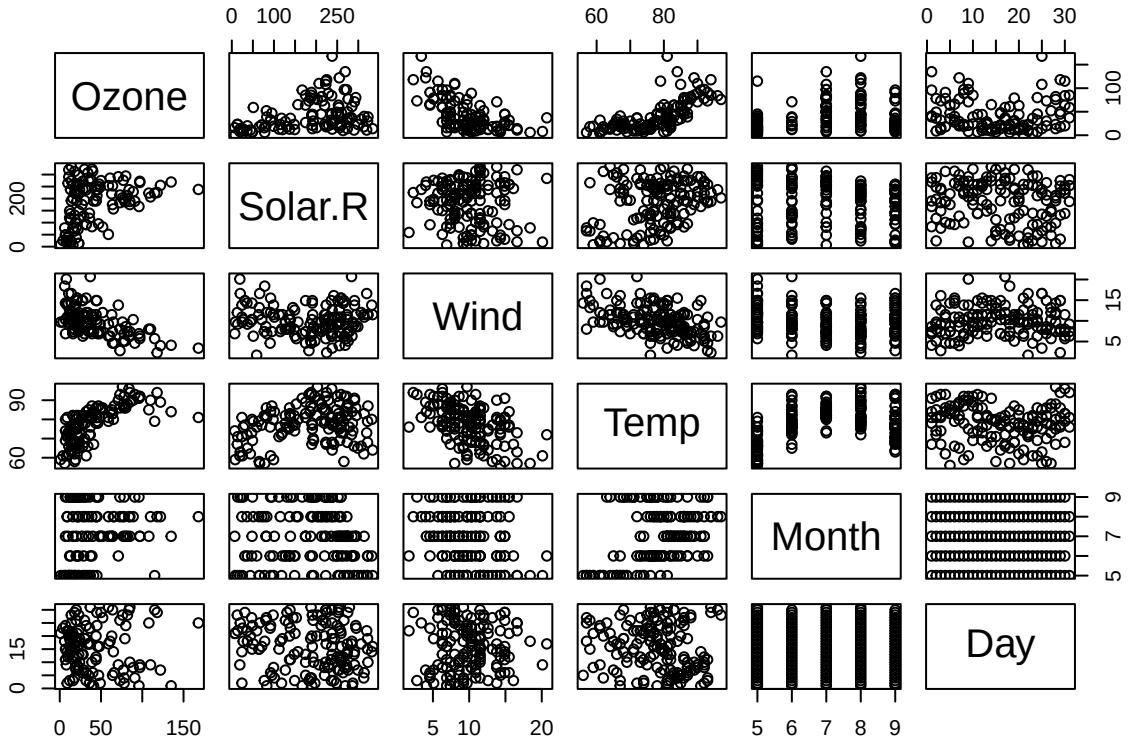
#Use different colors for each month.

```
boxplot(airquality[airquality$Month==5,]$Ozone, airquality[airquality$Month==6,]$Ozone,airquality[airqua
```



#(e) A scatter plot matrix of the numeric variables (Ozone, Solar.R, Wind, #Temp) within the airquality data set. (Hint: pairs())

```
pairs(airquality)
```



#2. Use simulation to estimate the mean and variance of a binomial random variable with n = 18 and p = 0.76. Compare with the theoretical values. (20 Points)

```
simu <-rbinom(n=18,size = 30, prob = 0.76)
mean(simu)
```

```
## [1] 23
```

```
var(simu)
```

```
## [1] 4.470588
```

#3. Estimate the mean and variance of a Poisson random variable whose mean is 7.2 by simulating 10,000 Poisson random numbers. (20 Points)

```
poisten <-rpois(n=10000,lambda=7.2)
mean(poisten)
```

```
## [1] 7.2095
```

```
var(poisten)
```

```
## [1] 7.253535
```

#4. Simulate 100 realizations of a normal random variable having a mean of 51 #and a standard deviation of 5.2. Estimate the mean and standard deviation of your simulated sample. (20 Points)

```
ranorten <- rnorm(10000,mean = 51,sd = 1)
mean(ranorten)
```

```
## [1] 50.99334
```

```
var(ranorten)
```

```
## [1] 1.007302
```