Assignment 1 PM566

Jazmin Hernandez

Assignment 1

Question 1

Data 2002

Looking at the 2002 data (data1) and summarizing the results, the data has 22 columns 15,976 observations. Based on the headers and footers, the first and last 6 rows of the data show no deviations from normality. The key variable in question is Daily Mean PM2.5 Concentration and is characterized as a numeric variable. There are no missing values and the min and max values are all within reasonable range.

```
library(data.table)
data1 <- fread ("2002data.csv")
data2 <- fread ("2022data.csv")
```

```
dim(data1)
```

```
[1] 15976    22
```

```
head(data1)
```

```
         Date Source  Site ID   POC Daily Mean PM2.5 Concentration    Units
       <char> <char>    <int> <int>                           <num>   <char>
1: 01/05/2002    AQS 60010007     1                            25.1 ug/m3 LC
2: 01/06/2002    AQS 60010007     1                            31.6 ug/m3 LC
3: 01/08/2002    AQS 60010007     1                            21.4 ug/m3 LC
4: 01/11/2002    AQS 60010007     1                            25.9 ug/m3 LC
5: 01/14/2002    AQS 60010007     1                            34.5 ug/m3 LC
6: 01/17/2002    AQS 60010007     1                            41.0 ug/m3 LC
```

```
   Daily AQI Value Local Site Name Daily Obs Count Percent Complete
           <int>          <char>           <int>            <num>
1:             81       Livermore               1              100
2:             93       Livermore               1              100
3:             74       Livermore               1              100
4:             82       Livermore               1              100
5:             98       Livermore               1              100
6:            115       Livermore               1              100
   AQS Parameter Code AQS Parameter Description Method Code
               <int>                     <char>       <int>
1:              88101  PM2.5 - Local Conditions         120
2:              88101  PM2.5 - Local Conditions         120
3:              88101  PM2.5 - Local Conditions         120
4:              88101  PM2.5 - Local Conditions         120
5:              88101  PM2.5 - Local Conditions         120
6:              88101  PM2.5 - Local Conditions         120
                     Method Description CBSA Code
                                <char>     <int>
1: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
2: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
3: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
4: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
5: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
6: Andersen RAAS2.5-300 PM2.5 SEQ w/WINS     41860
                          CBSA Name State FIPS Code      State
                             <char>           <int>     <char>
1: San Francisco-Oakland-Hayward, CA               6 California
2: San Francisco-Oakland-Hayward, CA               6 California
3: San Francisco-Oakland-Hayward, CA               6 California
4: San Francisco-Oakland-Hayward, CA               6 California
5: San Francisco-Oakland-Hayward, CA               6 California
6: San Francisco-Oakland-Hayward, CA               6 California
   County FIPS Code  County Site Latitude Site Longitude
              <int>  <char>         <num>          <num>
1:                1 Alameda      37.68753      -121.7842
2:                1 Alameda      37.68753      -121.7842
3:                1 Alameda      37.68753      -121.7842
4:                1 Alameda      37.68753      -121.7842
5:                1 Alameda      37.68753      -121.7842
6:                1 Alameda      37.68753      -121.7842
```

```
tail(data1)
```

```
         Date Source  Site ID   POC Daily Mean PM2.5 Concentration    Units
       <char> <char>    <int> <int>                           <num>   <char>
1: 12/10/2002    AQS 61131003     1                              15 ug/m3 LC
2: 12/13/2002    AQS 61131003     1                              15 ug/m3 LC
3: 12/22/2002    AQS 61131003     1                               1 ug/m3 LC
4: 12/25/2002    AQS 61131003     1                              23 ug/m3 LC
5: 12/28/2002    AQS 61131003     1                               5 ug/m3 LC
6: 12/31/2002    AQS 61131003     1                               6 ug/m3 LC
   Daily AQI Value      Local Site Name Daily Obs Count Percent Complete
             <int>               <char>           <int>            <num>
1:              62 Woodland-Gibson Road               1              100
2:              62 Woodland-Gibson Road               1              100
3:               6 Woodland-Gibson Road               1              100
4:              77 Woodland-Gibson Road               1              100
5:              28 Woodland-Gibson Road               1              100
6:              33 Woodland-Gibson Road               1              100
   AQS Parameter Code AQS Parameter Description Method Code
                <int>                    <char>       <int>
1:              88101  PM2.5 - Local Conditions         117
2:              88101  PM2.5 - Local Conditions         117
3:              88101  PM2.5 - Local Conditions         117
4:              88101  PM2.5 - Local Conditions         117
5:              88101  PM2.5 - Local Conditions         117
6:              88101  PM2.5 - Local Conditions         117
                      Method Description CBSA Code
                                  <char>     <int>
1: R & P Model 2000 PM2.5 Sampler w/WINS     40900
2: R & P Model 2000 PM2.5 Sampler w/WINS     40900
3: R & P Model 2000 PM2.5 Sampler w/WINS     40900
4: R & P Model 2000 PM2.5 Sampler w/WINS     40900
5: R & P Model 2000 PM2.5 Sampler w/WINS     40900
6: R & P Model 2000 PM2.5 Sampler w/WINS     40900
                                CBSA Name State FIPS Code      State
                                   <char>           <int>     <char>
1: Sacramento--Roseville--Arden-Arcade, CA               6 California
2: Sacramento--Roseville--Arden-Arcade, CA               6 California
3: Sacramento--Roseville--Arden-Arcade, CA               6 California
4: Sacramento--Roseville--Arden-Arcade, CA               6 California
5: Sacramento--Roseville--Arden-Arcade, CA               6 California
6: Sacramento--Roseville--Arden-Arcade, CA               6 California
```

```
   County FIPS Code County Site Latitude Site Longitude
              <int> <char>         <num>          <num>
1:              113   Yolo      38.66121      -121.7327
2:              113   Yolo      38.66121      -121.7327
3:              113   Yolo      38.66121      -121.7327
4:              113   Yolo      38.66121      -121.7327
5:              113   Yolo      38.66121      -121.7327
6:              113   Yolo      38.66121      -121.7327
```

```
str(data1)
```

```
Classes 'data.table' and 'data.frame':  15976 obs. of  22 variables:
 $ Date                           : chr  "01/05/2002" "01/06/2002" "01/08/2002" "01/11/2002"
 $ Source                         : chr  "AQS" "AQS" "AQS" "AQS" ...
 $ Site ID                        : int  60010007 60010007 60010007 60010007 60010007 6001000
 $ POC                            : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Daily Mean PM2.5 Concentration: num  25.1 31.6 21.4 25.9 34.5 41 29.3 15 18.8 37.9 ...
 $ Units                          : chr  "ug/m3 LC" "ug/m3 LC" "ug/m3 LC" "ug/m3 LC" ...
 $ Daily AQI Value                : int  81 93 74 82 98 115 89 62 69 107 ...
 $ Local Site Name                : chr  "Livermore" "Livermore" "Livermore" "Livermore" ...
 $ Daily Obs Count                : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Percent Complete               : num  100 100 100 100 100 100 100 100 100 100 ...
 $ AQS Parameter Code             : int  88101 88101 88101 88101 88101 88101 88101 88101 8810
 $ AQS Parameter Description      : chr  "PM2.5 - Local Conditions" "PM2.5 - Local Conditions"
 $ Method Code                    : int  120 120 120 120 120 120 120 120 120 120 ...
 $ Method Description             : chr  "Andersen RAAS2.5-300 PM2.5 SEQ w/WINS" "Andersen RA
 $ CBSA Code                      : int  41860 41860 41860 41860 41860 41860 41860 41860 4186
 $ CBSA Name                      : chr  "San Francisco-Oakland-Hayward, CA" "San Francisco-Oa
 $ State FIPS Code                : int  6 6 6 6 6 6 6 6 6 6 ...
 $ State                          : chr  "California" "California" "California" "California"
 $ County FIPS Code               : int  1 1 1 1 1 1 1 1 1 1 ...
 $ County                         : chr  "Alameda" "Alameda" "Alameda" "Alameda" ...
 $ Site Latitude                  : num  37.7 37.7 37.7 37.7 37.7 ...
 $ Site Longitude                 : num  -122 -122 -122 -122 -122 ...
 - attr(*, ".internal.selfref")=<externalptr>
```

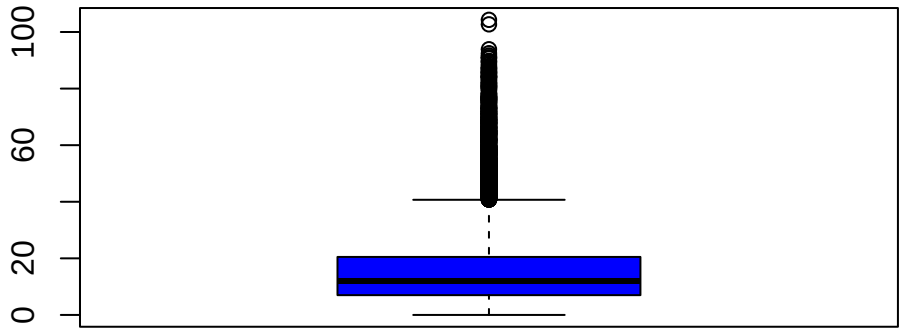
```
summary(data1$`Daily Mean PM2.5 Concentration`)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    7.00   12.00   16.12   20.50  104.30
```

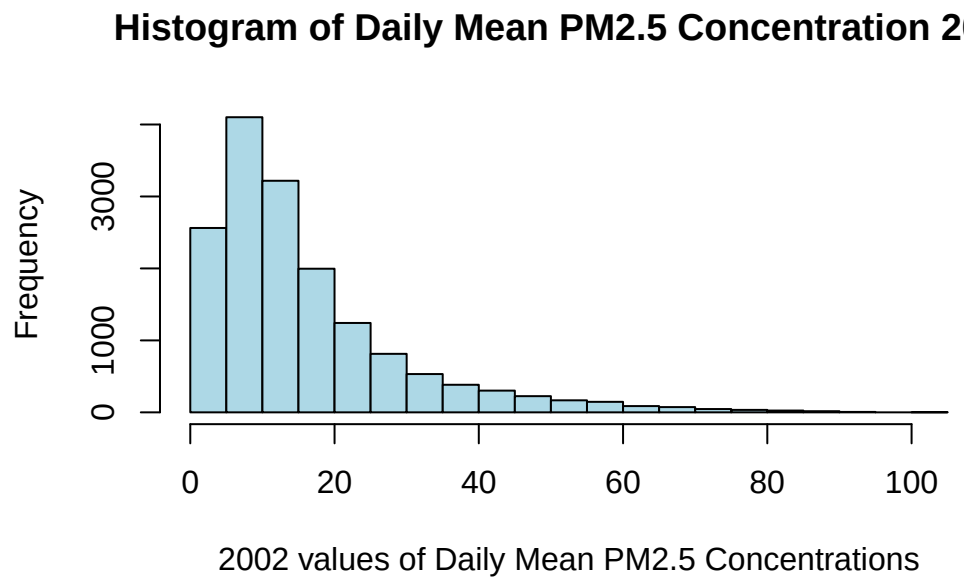
```r
mean(is.na(data1$`Daily Mean PM2.5 Concentration`))
```

```
[1] 0
```

```r
boxplot(data1$`Daily Mean PM2.5 Concentration`, col = "blue")
```




```r
hist(data1$`Daily Mean PM2.5 Concentration`,
      main = "Histogram of Daily Mean PM2.5 Concentration 2002",
     xlab = "2002 values of Daily Mean PM2.5 Concentrations",
     ylab = "Frequency",
     col = "lightblue",
     border = "black")
```



2022 Data Set

For the 2022 data (data2) and summarizing the results, the data has 22 columns 59,756 observations. Based on the headers and footers, the first and last 6 rows of the data show no deviations from normality. There are no missing values in this data set but looking at the min and max values, we can see that the min PM2.5 concentration is -6.7 which is highly unlikely.

```
dim(data2)
```

```
[1] 59756    22
```

```
head(data2)
```

```
         Date Source  Site ID   POC Daily Mean PM2.5 Concentration    Units
       <char> <char>    <int> <int>                           <num>   <char>
1: 01/01/2022    AQS 60010007     3                            12.7 ug/m3 LC
2: 01/02/2022    AQS 60010007     3                            13.9 ug/m3 LC
3: 01/03/2022    AQS 60010007     3                             7.1 ug/m3 LC
4: 01/04/2022    AQS 60010007     3                             3.7 ug/m3 LC
5: 01/05/2022    AQS 60010007     3                             4.2 ug/m3 LC
```

```
6: 01/06/2022    AQS 60010007     3                          3.8 ug/m3 LC
   Daily AQI Value Local Site Name Daily Obs Count Percent Complete
           <int>          <char>           <int>            <num>
1:              58       Livermore               1              100
2:              60       Livermore               1              100
3:              39       Livermore               1              100
4:              21       Livermore               1              100
5:              23       Livermore               1              100
6:              21       Livermore               1              100
   AQS Parameter Code AQS Parameter Description Method Code
               <int>                    <char>       <int>
1:              88101  PM2.5 - Local Conditions         170
2:              88101  PM2.5 - Local Conditions         170
3:              88101  PM2.5 - Local Conditions         170
4:              88101  PM2.5 - Local Conditions         170
5:              88101  PM2.5 - Local Conditions         170
6:              88101  PM2.5 - Local Conditions         170
                    Method Description CBSA Code
                                <char>     <int>
1: Met One BAM-1020 Mass Monitor w/VSCC     41860
2: Met One BAM-1020 Mass Monitor w/VSCC     41860
3: Met One BAM-1020 Mass Monitor w/VSCC     41860
4: Met One BAM-1020 Mass Monitor w/VSCC     41860
5: Met One BAM-1020 Mass Monitor w/VSCC     41860
6: Met One BAM-1020 Mass Monitor w/VSCC     41860
                         CBSA Name State FIPS Code      State
                            <char>           <int>     <char>
1: San Francisco-Oakland-Hayward, CA               6 California
2: San Francisco-Oakland-Hayward, CA               6 California
3: San Francisco-Oakland-Hayward, CA               6 California
4: San Francisco-Oakland-Hayward, CA               6 California
5: San Francisco-Oakland-Hayward, CA               6 California
6: San Francisco-Oakland-Hayward, CA               6 California
   County FIPS Code  County Site Latitude Site Longitude
              <int>  <char>         <num>          <num>
1:                1 Alameda      37.68753      -121.7842
2:                1 Alameda      37.68753      -121.7842
3:                1 Alameda      37.68753      -121.7842
4:                1 Alameda      37.68753      -121.7842
5:                1 Alameda      37.68753      -121.7842
6:                1 Alameda      37.68753      -121.7842
```

```
tail(data2)
```

```
         Date Source  Site ID   POC Daily Mean PM2.5 Concentration    Units
       <char> <char>    <int> <int>                           <num>   <char>
1: 12/01/2022    AQS 61131003     1                             3.4 ug/m3 LC
2: 12/07/2022    AQS 61131003     1                             3.8 ug/m3 LC
3: 12/13/2022    AQS 61131003     1                             6.0 ug/m3 LC
4: 12/19/2022    AQS 61131003     1                            34.8 ug/m3 LC
5: 12/25/2022    AQS 61131003     1                            23.2 ug/m3 LC
6: 12/31/2022    AQS 61131003     1                             1.0 ug/m3 LC
   Daily AQI Value      Local Site Name Daily Obs Count Percent Complete
             <int>               <char>           <int>            <num>
1:              19 Woodland-Gibson Road               1              100
2:              21 Woodland-Gibson Road               1              100
3:              33 Woodland-Gibson Road               1              100
4:              99 Woodland-Gibson Road               1              100
5:              77 Woodland-Gibson Road               1              100
6:               6 Woodland-Gibson Road               1              100
   AQS Parameter Code AQS Parameter Description Method Code
                <int>                    <char>       <int>
1:              88101  PM2.5 - Local Conditions         145
2:              88101  PM2.5 - Local Conditions         145
3:              88101  PM2.5 - Local Conditions         145
4:              88101  PM2.5 - Local Conditions         145
5:              88101  PM2.5 - Local Conditions         145
6:              88101  PM2.5 - Local Conditions         145
                                    Method Description CBSA Code
                                                <char>     <int>
1: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
2: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
3: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
4: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
5: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
6: R & P Model 2025 PM-2.5 Sequential Air Sampler w/VSCC     40900
                               CBSA Name State FIPS Code      State
                                  <char>           <int>     <char>
1: Sacramento--Roseville--Arden-Arcade, CA               6 California
2: Sacramento--Roseville--Arden-Arcade, CA               6 California
3: Sacramento--Roseville--Arden-Arcade, CA               6 California
4: Sacramento--Roseville--Arden-Arcade, CA               6 California
5: Sacramento--Roseville--Arden-Arcade, CA               6 California
6: Sacramento--Roseville--Arden-Arcade, CA               6 California
```

```
   County FIPS Code County Site Latitude Site Longitude
              <int> <char>         <num>          <num>
1:              113   Yolo      38.66121      -121.7327
2:              113   Yolo      38.66121      -121.7327
3:              113   Yolo      38.66121      -121.7327
4:              113   Yolo      38.66121      -121.7327
5:              113   Yolo      38.66121      -121.7327
6:              113   Yolo      38.66121      -121.7327
```

```
str(data2)
```

```
Classes 'data.table' and 'data.frame':  59756 obs. of  22 variables:
 $ Date                           : chr  "01/01/2022" "01/02/2022" "01/03/2022" "01/04/2022"
 $ Source                         : chr  "AQS" "AQS" "AQS" "AQS" ...
 $ Site ID                        : int  60010007 60010007 60010007 60010007 60010007 6001000
 $ POC                            : int  3 3 3 3 3 3 3 3 3 3 ...
 $ Daily Mean PM2.5 Concentration: num  12.7 13.9 7.1 3.7 4.2 3.8 2.3 6.9 13.6 11.2 ...
 $ Units                          : chr  "ug/m3 LC" "ug/m3 LC" "ug/m3 LC" "ug/m3 LC" ...
 $ Daily AQI Value                : int  58 60 39 21 23 21 13 38 59 55 ...
 $ Local Site Name                : chr  "Livermore" "Livermore" "Livermore" "Livermore" ...
 $ Daily Obs Count                : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Percent Complete               : num  100 100 100 100 100 100 100 100 100 100 ...
 $ AQS Parameter Code             : int  88101 88101 88101 88101 88101 88101 88101 88101 8810
 $ AQS Parameter Description      : chr  "PM2.5 - Local Conditions" "PM2.5 - Local Conditions"
 $ Method Code                    : int  170 170 170 170 170 170 170 170 170 170 ...
 $ Method Description             : chr  "Met One BAM-1020 Mass Monitor w/VSCC" "Met One BAM-
 $ CBSA Code                      : int  41860 41860 41860 41860 41860 41860 41860 41860 4186
 $ CBSA Name                      : chr  "San Francisco-Oakland-Hayward, CA" "San Francisco-Oa
 $ State FIPS Code                : int  6 6 6 6 6 6 6 6 6 6 ...
 $ State                          : chr  "California" "California" "California" "California"
 $ County FIPS Code               : int  1 1 1 1 1 1 1 1 1 1 ...
 $ County                         : chr  "Alameda" "Alameda" "Alameda" "Alameda" ...
 $ Site Latitude                  : num  37.7 37.7 37.7 37.7 37.7 ...
 $ Site Longitude                 : num  -122 -122 -122 -122 -122 ...
 - attr(*, ".internal.selfref")=<externalptr>
```

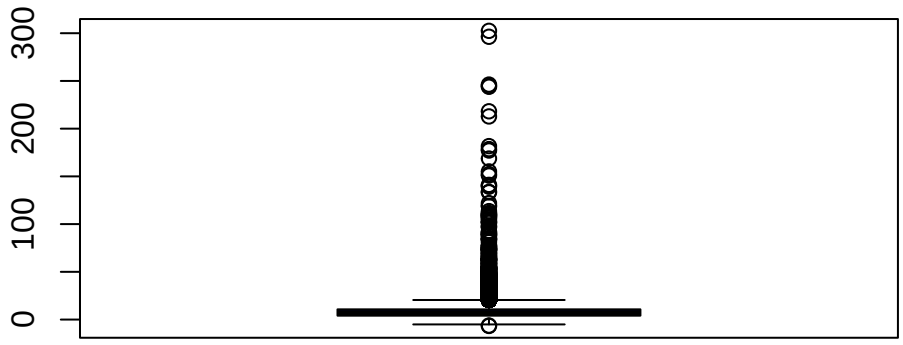
```
summary(data2$`Daily Mean PM2.5 Concentration`)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -6.700   4.100   6.800   8.429  10.700 302.500
```

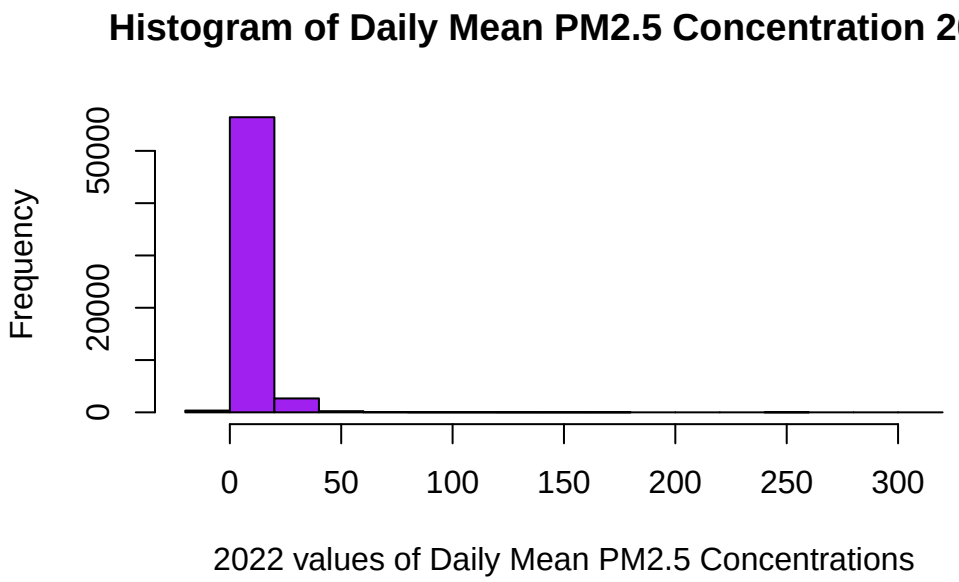
```r
mean(is.na(data2$`Daily Mean PM2.5 Concentration`))
```

```
[1] 0
```

```r
boxplot(data2$`Daily Mean PM2.5 Concentration`, col = "blue")
```




```r
hist(data2$`Daily Mean PM2.5 Concentration`,
      main = "Histogram of Daily Mean PM2.5 Concentration 2022",
     xlab = "2022 values of Daily Mean PM2.5 Concentrations",
     ylab = "Frequency",
     col = "purple",
     border = "black")
```



Question 2

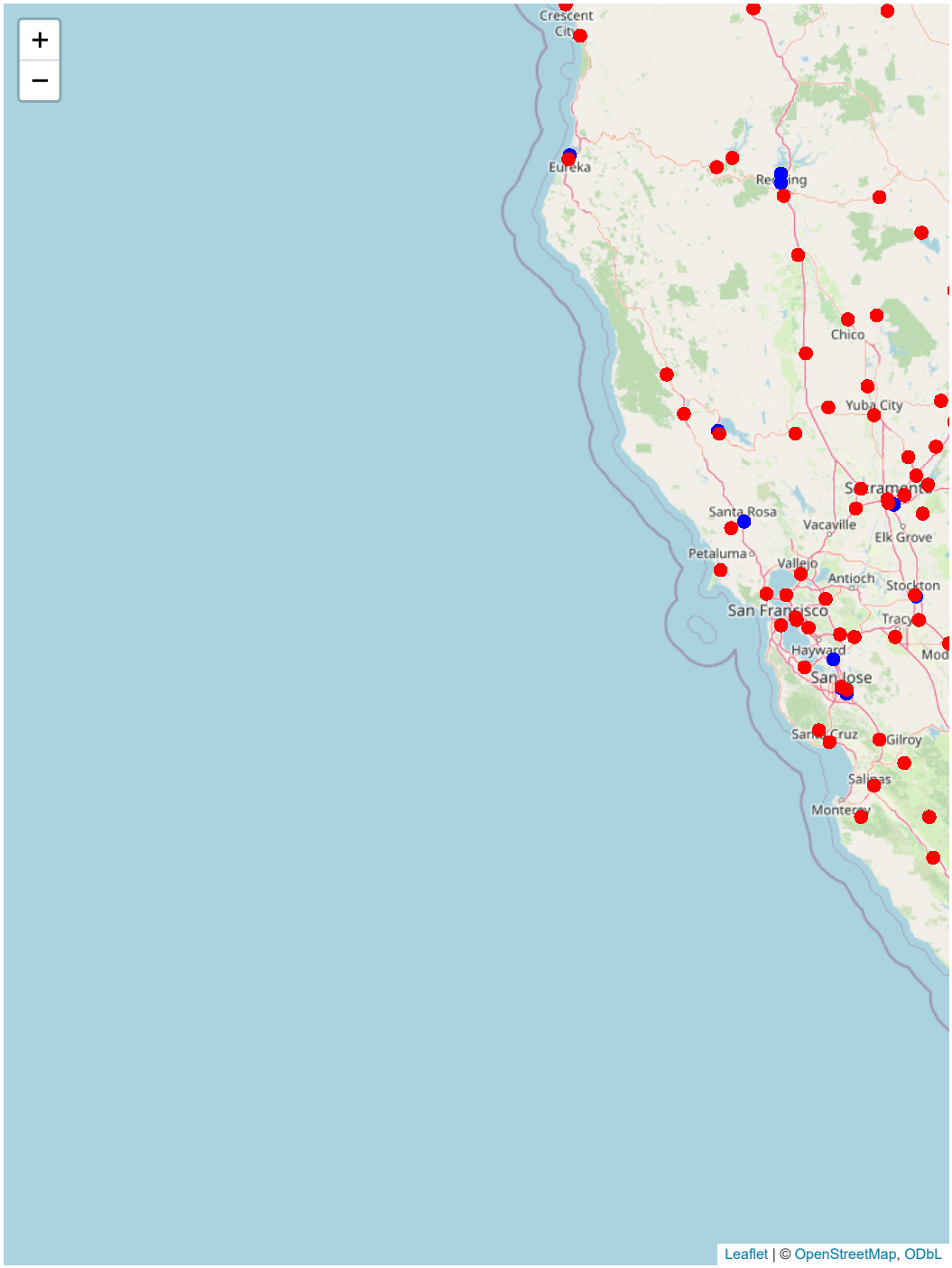
```
data1[, Year := 2002]
data2[, Year := 2022]
combined_data <- rbind(data1, data2)
setnames(combined_data, old = c( "Site Latitude", "Site Longitude"), new = c("Latitude", "Lo
```

Question 3

For the year 2002 which is represented by the blue circles, we can see that they are overtaken by the year 2022 (red) circles because of the almost 44,000 observation difference between the data sets. However, it is also evident that most of the PM2.5 concentration is along the coast.

```
library(leaflet)
map <- leaflet(data = combined_data) %>%
  addTiles() %>%
  addCircleMarkers(
    lng = ~Longitude,
    lat = ~Latitude,
```

```
    color = ~ifelse(Year == 2002, "blue", "red"),  # Color by year
    radius = 5,
    stroke = FALSE,
    fillOpacity = 0.7,
    popup = ~paste("Site ID:", `Site ID`, "<br>", "Year:", Year)
  )
map
```


Crescent City
Eureka
Chico
Yuba City
Santa Rosa
Vacaville
Elk Grove
Petaluma
Vallejo
Antioch
Stockton
San Francisco
Tracy
Hayward
San Jose
Gilroy
Salinas
Monterey
Leaflet | © OpenStreetMap, ODbL

Question 4

There are not any missing values for PM2.5 in the combined data sets. However, checking for implausible values, there were 215 total negative observations for PM2.5. This was only recorded for the year 2022 which would explain why it has so many more observations compared to 2002. Most of these observations occurred in Willows-Colusa Street during January through July and in Lebec from January to December.

```
mean(is.na(combined_data$`Daily Mean PM2.5 Concentration`))
```

```
[1] 0
```

```
implausible_PM2.5 <- combined_data[`Daily Mean PM2.5 Concentration` < 0, .(Date, Year, `Local
print(implausible_PM2.5)
```

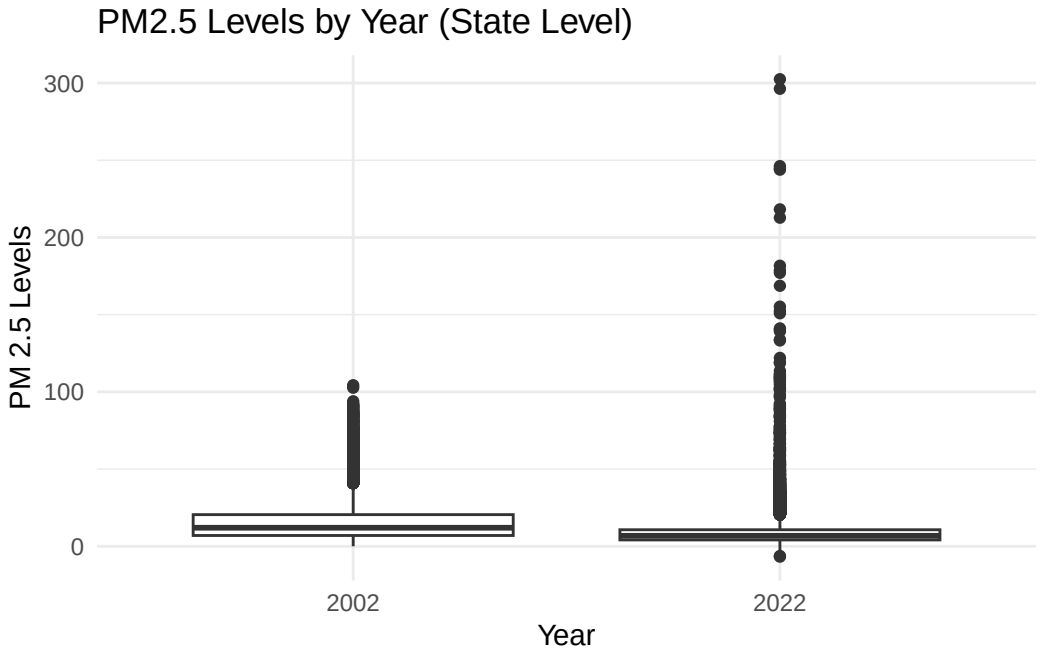
```
          Date  Year    Local Site Name Daily Mean PM2.5 Concentration
        <char> <num>             <char>                          <num>
  1: 07/06/2022  2022       Oakland West                           -0.7
  2: 07/30/2022  2022       Oakland West                           -0.1
  3: 08/26/2022  2022       Oakland West                           -0.5
  4: 02/01/2022  2022 Paradise - Theater                           -0.3
  5: 02/06/2022  2022 Paradise - Theater                           -0.1
 ---
211: 06/11/2022  2022   Davis-UCD Campus                           -0.8
212: 06/12/2022  2022   Davis-UCD Campus                           -0.4
213: 07/06/2022  2022   Davis-UCD Campus                           -0.6
214: 11/02/2022  2022   Davis-UCD Campus                           -0.1
215: 11/03/2022  2022   Davis-UCD Campus                           -0.1
```

Question 5

State Level

From the summary statistics, we can see that the max PM2.5 level is 302.50 ug/m3. From the box plot we can see that this reading belongs from the year 2022. There is also a difference seen in the increase of PM2.5 from the year 2002 to 2022 from about 100 ug/m3 to 302.50 ug/m3.

```
library(ggplot2)
ggplot(combined_data, aes(x = factor(Year), y = `Daily Mean PM2.5 Concentration`)) +
  geom_boxplot() +
  labs(title = "PM2.5 Levels by Year (State Level)", x = "Year", y = "PM 2.5 Levels") +
  theme_minimal()
```




```
summary(combined_data$`Daily Mean PM2.5 Concentration`)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -6.70    4.50    7.60   10.05   12.20  302.50
```

For County Level

After grouping by county level, we can see that Kern County has the highest PM2.5 concentration of 15.60 ug/m3 and El Dorado has the lowest with 4.47 ug/m3. This is also reflected in the histogram although it is clearer to see in the summary statistics of the counties.

```
library(dplyr)
```

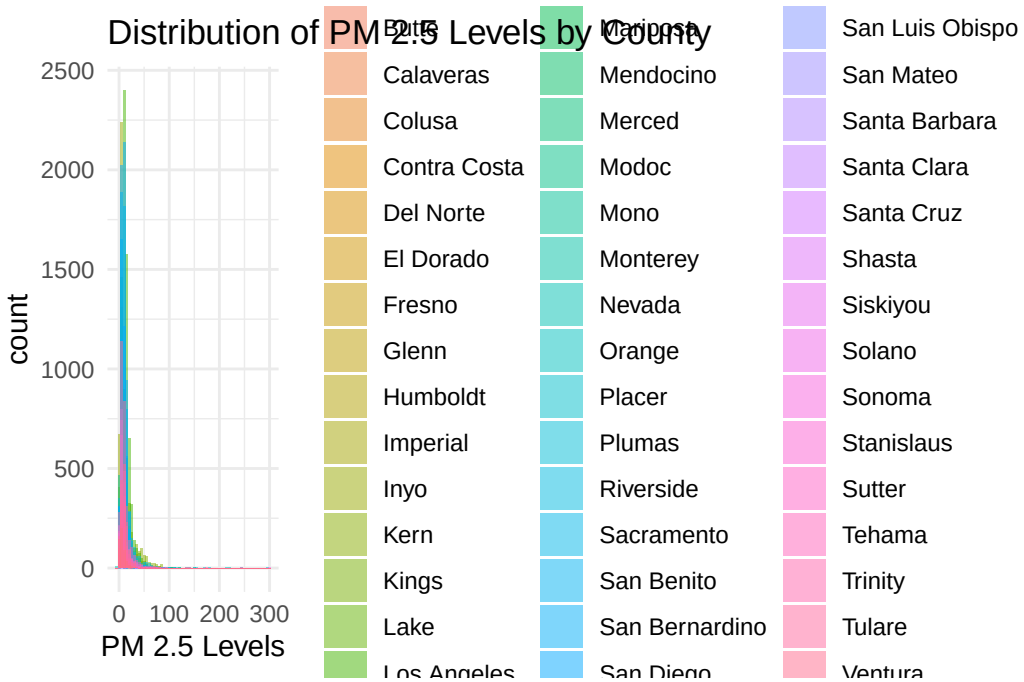
```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:data.table':

    between, first, last

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
county_summary <- combined_data %>%
  group_by(County) %>%
  summarise(
    Mean_PM2.5 = mean(`Daily Mean PM2.5 Concentration`, na.rm = TRUE),
    Median_PM2.5 = median(`Daily Mean PM2.5 Concentration`, na.rm = TRUE),
    SD_PM2.5 = sd(`Daily Mean PM2.5 Concentration`, na.rm = TRUE),
    .groups = 'drop'
  )

print(county_summary)
```

```
# A tibble: 51 x 4
   County       Mean_PM2.5 Median_PM2.5 SD_PM2.5
   <chr>             <dbl>        <dbl>    <dbl>
 1 Alameda            8.81         7.2      6.21
 2 Butte              8.73         6        8.90
 3 Calaveras          6.60         5.3      4.71
 4 Colusa             8.40         7        6.32
 5 Contra Costa       9.98         7.8      8.93
 6 Del Norte          4.75         4.05     3.43
 7 El Dorado          4.47         3.1      7.21
 8 Fresno            12.3          8.4     12.1
 9 Glenn              5.34         4.4      4.98
10 Humboldt           7.11         6        4.45
# i 41 more rows
```

```r
ggplot(combined_data, aes(x = `Daily Mean PM2.5 Concentration`, fill = factor(County))) +
  geom_histogram(binwidth = 5, position = "identity", alpha = 0.5) +
  labs(title = "Distribution of PM 2.5 Levels by County", x = "PM 2.5 Levels", fill = "County
  theme_minimal()
```



For Los Angeles Level

Filtering by only the LA site level, the mean PM2.5 level is 13.32 ug/m3 just below that of Kern County. From the line plot, it appears that the particulate matter concentrations increase as the year ends. Also, sites 60377500 and below seem to have the lowest PM2.5 concentration with sites 60372500 and above having the highest concentrations.

```
LA_Site <- combined_data %>% filter(County == "Los Angeles")
LA_summary <- LA_Site %>%
  summarise(
    Mean_PM2.5 = mean(`Daily Mean PM2.5 Concentration`, na.rm = TRUE),
    Median_PM2.5 = median(`Daily Mean PM2.5 Concentration`, na.rm = TRUE),
    SD_PM2.5 = sd(`Daily Mean PM2.5 Concentration`, na.rm = TRUE)
  )

print(LA_summary)
```

```
  Mean_PM2.5 Median_PM2.5 SD_PM2.5
1   13.31989         11.4  8.54839
```

```
ggplot(LA_Site, aes(x = Date, y = `Daily Mean PM2.5 Concentration`, group = `Site ID`, color
  geom_line() +
  labs(title = "PM 2.5 Levels Over Time at Sites in Los Angeles", x = "Date", y = "PM 2.5 Lev
```