Air Pollution

Ania

18 01 2022

Introduction

The harmful effects of air pollution on human health have been confirmed by many scientific studies. Air condition monitoring mainly uses two parameters to describe air quality - PM 2.5 and PM 10. *PM* stands for *particulate matter* and number refers to the diameter of the particles (2.5 micro- and 10 micrometers, respectively). In the United States, for tracking and monitoring of the air polutions emissions the Environmental Protection Agency (*EPA*) is responsible. Approximatly every 3 years, the EPA releases its database on emissions of PM2.5 (*National Emissions Inventory - NEI*).

The records shows how many tons of PM2.5 were emitted from diferent sources anually. The data used for this analysis are for 1999, 2002, 2005, and 2008 year.

1. Data reading

```
library(ggplot2)
library(downloader)

filename <- "PM2.5 Emission Data.zip"
if(!file.exists(filename)) {
  url <- "https://d396qusza40orc.cloudfront.net/exdata%2Fdata%2FNEI_data.zip"
  download(url, filename, mode = "wb")
}
if(!file.exists("PM2.5 Emission Data")){
  unzip(filename)}

#reading files
NEI <- readRDS("summarySCC_PM25.rds")
```

2. Introductory analysis

The data frame in six columns contains information about:

- *fips*: A five-digit number (represented as a string) indicating the U.S. county
- *SCC*: The name of the source as indicated by a digit string (see source code classification table)
- *Pollutant*: A string indicating the pollutant
- *Emissions*: Amount of PM2.5 emitted, in tons
- *type*: The type of source (point, non-point, on-road, or non-road)

- *year*: The year of emissions recorded

```
dim(NEI)
```

```
## [1] 6497651       6
```

```
head(NEI)
```

```
##     fips      SCC Pollutant Emissions  type year
## 4  09001 10100401  PM25-PRI    15.714 POINT 1999
## 8  09001 10100404  PM25-PRI   234.178 POINT 1999
## 12 09001 10100501  PM25-PRI     0.128 POINT 1999
## 16 09001 10200401  PM25-PRI     2.036 POINT 1999
## 20 09001 10200504  PM25-PRI     0.388 POINT 1999
## 24 09001 10200602  PM25-PRI     1.490 POINT 1999
```

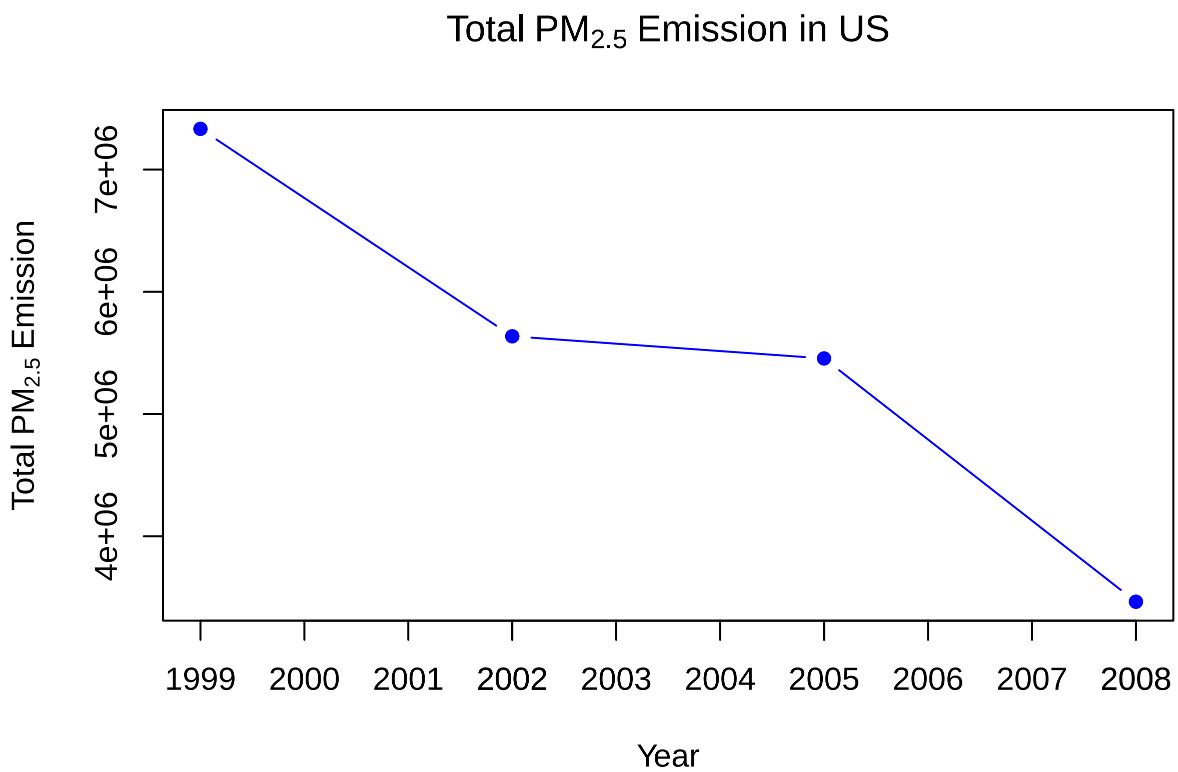
```
summary(NEI)
```

```
##      fips               SCC              Pollutant          Emissions
##  Length:6497651     Length:6497651     Length:6497651     Min.   :     0.0
##  Class :character   Class :character   Class :character   1st Qu.:     0.0
##  Mode  :character   Mode  :character   Mode  :character   Median :     0.0
##                                                           Mean   :     3.4
##                                                           3rd Qu.:     0.1
##                                                           Max.   :646952.0
##      type               year
##  Length:6497651     Min.   :1999
##  Class :character   1st Qu.:2002
##  Mode  :character   Median :2005
##                     Mean   :2004
##                     3rd Qu.:2008
##                     Max.   :2008
```

Question 1

Have total emissions from PM2.5 decreased in the United States from 1999 to 2008?

```
NEIdata <- aggregate(Emissions ~ year, NEI, sum)

plot(NEIdata$year, NEIdata$Emissions, type = "b", main = "Total"~ PM[2.5] ~"Emission in US", xlab = "Yea

axis (side = 1, at = c(1999,2001,2002,2003,2005,2005,2007,2008))
```




```r
#saving the graph as png file

dev.copy (png, "plot1.png", width=480, height=480)
```

```
## png
##   3
```

```r
dev.off()
```

```
## pdf
##   2
```

Total emission of PM 2.5 dropped between 1999 and 2008.

Question 2

Have total emissions from PM2.5 decreased in the Baltimore City, Maryland (fips = "24510") from 1999 to 2008?

To answer this question is necessary to read in another data set **Source_Classification_Code.rds** which contains digit strings for mapping in the Emissions table to the actual name of the PM2.5 source.

```r
SCC <- readRDS("Source_Classification_Code.rds")
dim(SCC)
```

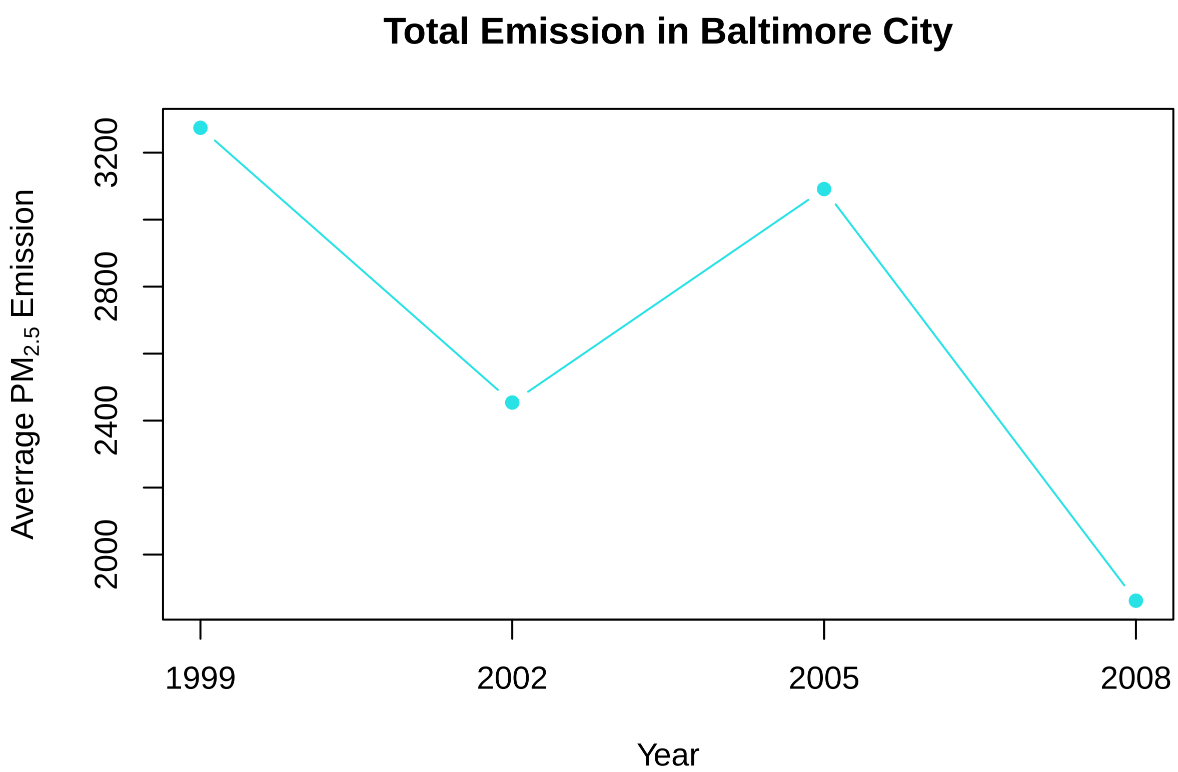
```
## [1] 11717    15
```

```r
head(SCC)
```

```
##        SCC Data.Category
## 1 10100101         Point
## 2 10100102         Point
## 3 10100201         Point
## 4 10100202         Point
## 5 10100203         Point
## 6 10100204         Point
##                                                                   Short.Name
## 1                   Ext Comb /Electric Gen /Anthracite Coal /Pulverized Coal
## 2 Ext Comb /Electric Gen /Anthracite Coal /Traveling Grate (Overfeed) Stoker
## 3       Ext Comb /Electric Gen /Bituminous Coal /Pulverized Coal: Wet Bottom
## 4       Ext Comb /Electric Gen /Bituminous Coal /Pulverized Coal: Dry Bottom
## 5                   Ext Comb /Electric Gen /Bituminous Coal /Cyclone Furnace
## 6                   Ext Comb /Electric Gen /Bituminous Coal /Spreader Stoker
##                                EI.Sector Option.Group Option.Set
## 1 Fuel Comb - Electric Generation - Coal
## 2 Fuel Comb - Electric Generation - Coal
## 3 Fuel Comb - Electric Generation - Coal
## 4 Fuel Comb - Electric Generation - Coal
## 5 Fuel Comb - Electric Generation - Coal
## 6 Fuel Comb - Electric Generation - Coal
##                SCC.Level.One       SCC.Level.Two              SCC.Level.Three
## 1 External Combustion Boilers Electric Generation              Anthracite Coal
## 2 External Combustion Boilers Electric Generation              Anthracite Coal
## 3 External Combustion Boilers Electric Generation Bituminous/Subbituminous Coal
## 4 External Combustion Boilers Electric Generation Bituminous/Subbituminous Coal
## 5 External Combustion Boilers Electric Generation Bituminous/Subbituminous Coal
## 6 External Combustion Boilers Electric Generation Bituminous/Subbituminous Coal
##                                 SCC.Level.Four Map.To Last.Inventory.Year
## 1                              Pulverized Coal     NA                  NA
## 2            Traveling Grate (Overfeed) Stoker     NA                  NA
## 3 Pulverized Coal: Wet Bottom (Bituminous Coal)     NA                  NA
## 4 Pulverized Coal: Dry Bottom (Bituminous Coal)     NA                  NA
## 5            Cyclone Furnace (Bituminous Coal)     NA                  NA
## 6            Spreader Stoker (Bituminous Coal)     NA                  NA
##   Created_Date Revised_Date Usage.Notes
## 1
## 2
## 3
## 4
## 5
## 6
```

Fips number for the Baltimore City in Maryland is **24510** so befor further analysis NEI data set was subset to obtain results onli for the Baltimore City.

```r
NEI_BC <- subset(NEI, fips == "24510")

NEIdata <- aggregate(Emissions ~ year, NEI_BC, sum)

plot(NEIdata$year, NEIdata$Emissions, type = "b", main = "Total Emission in Baltimore City", xlab = "Yea
axis (side = 1, at = c(1999,2002,2005,2005,2008))
```




```r
#saving the graph as png file

dev.copy (png, "plot2.png", width=480, height=480)
```

```
## png
##   3
```

```r
dev.off()
```

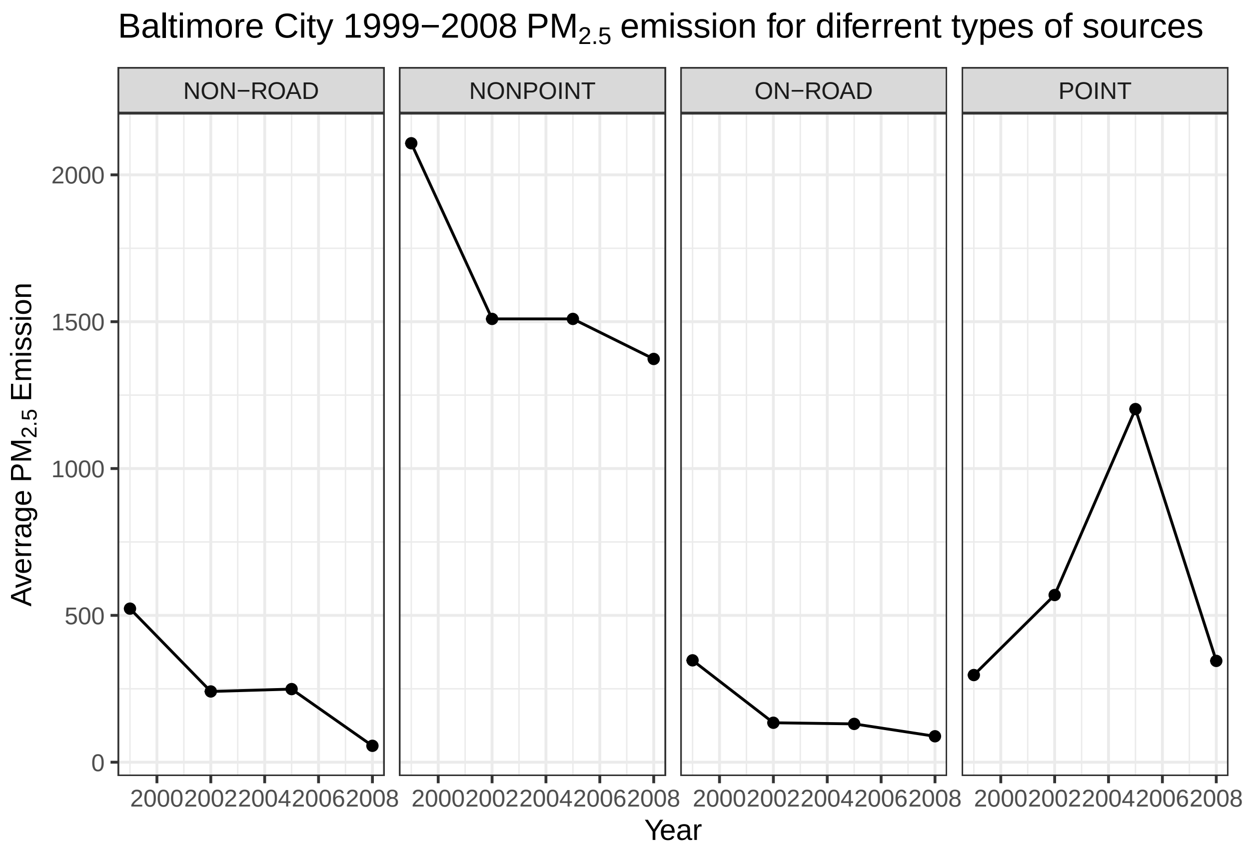
```
## pdf
##   2
```

Between 1999 and 2008 average PM 2.5 emission decreased in Baltimore City. A temporary increase of emissions can be observed in 2005, which decreased in the next measuring point.

Question 3

Of the four types of sources indicated by the type (point, nonpoint, onroad, nonroad) variable, which of these four sources have seen decreases in emissions from 1999–2008 for Baltimore City? Which have seen increases in emissions from 1999–2008?

```
NEIdata <- aggregate(Emissions ~ year + type, NEI_BC, sum)

graph_1 <- ggplot(NEIdata, aes(year, Emissions))

graph_1 + geom_point() + geom_line() + facet_grid(.~as.factor(NEIdata$type)) + theme(axis.text.x = eleme
```




```
#saving the graph as png file

dev.copy (png, "plot3.png", width=480, height=480)
```

```
## png
##   3
```

```
dev.off()
```

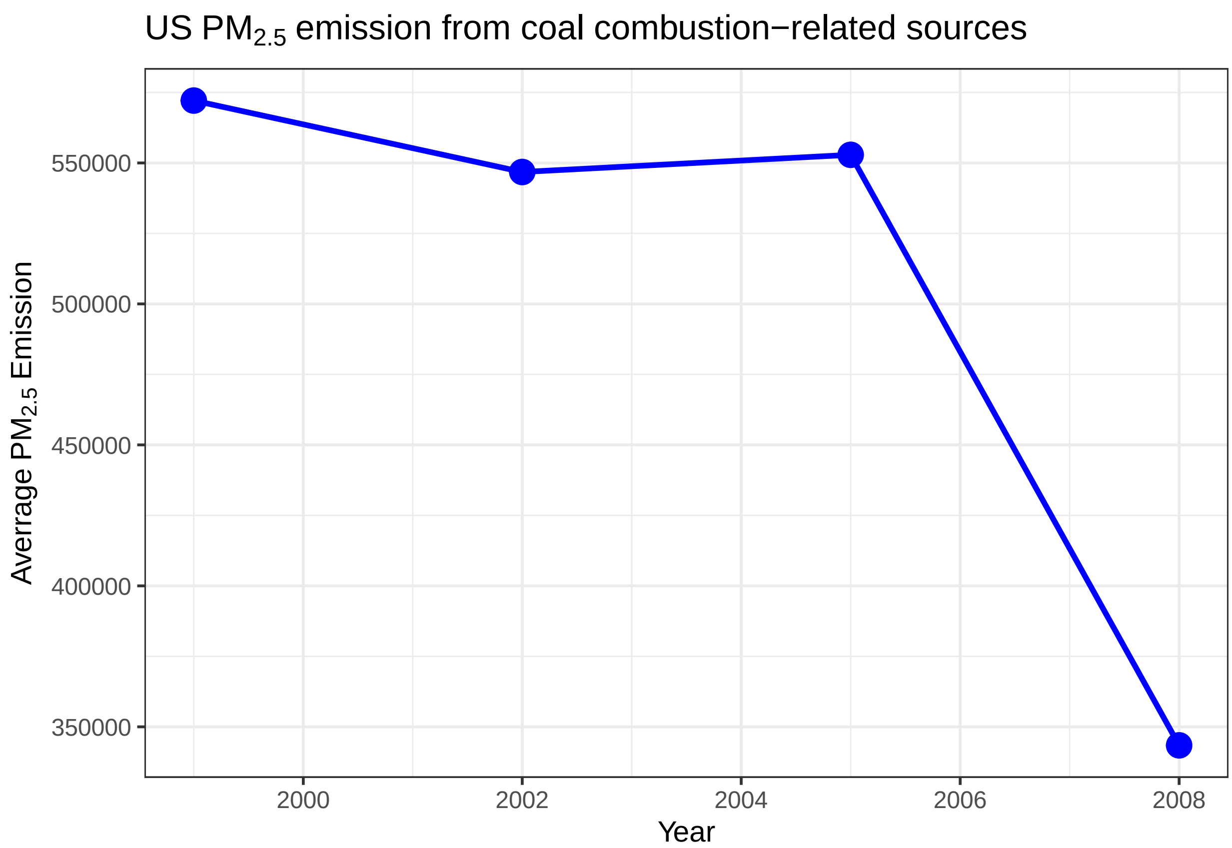
```
## pdf
##   2
```

During analyzed period for almost all sources the average emission of PM 2.5 decreased. Only exceptions are point sources, where increase in emission from 1999 to 2005 can be observed fallowed by decrease in 2008.

Question 4

Across the United States, how have emissions from coal combustion-related sources changed from 1999–2008?

```
SCC_coal <- SCC[grepl("coal",SCC$EI.Sector, ignore.case = TRUE),]
NEI_coal <- NEI[NEI$SCC %in% SCC_coal$SCC,]
NEI_coal <- aggregate(Emissions ~ year, NEI_coal, sum)

#plotting

graph_2 <- ggplot(NEI_coal, aes(year, Emissions))

graph_2 + geom_point(col = "blue", size = 4) + geom_line(col = "blue", lwd = 1) + theme(axis.text.x = el
```




```
#saving the graph as png file

dev.copy (png, "plot4.png", width=480, height=480)
```

```
## png
##   3
```

```
dev.off()
```

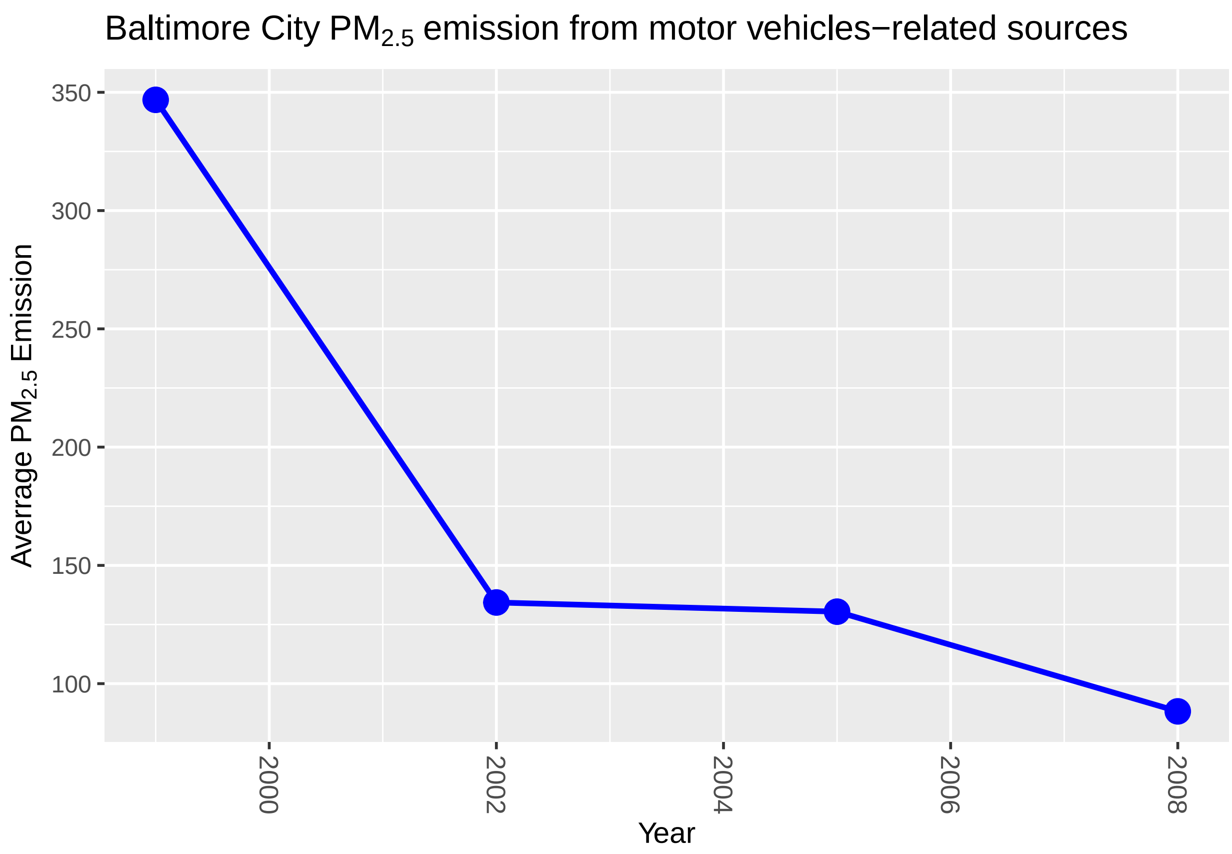
```
## pdf
##   2
```

Between 1999 and 2008 emission from combustion-related sources decreased.

Question 5

How have emissions from motor vehicle sources changed from 1999–2008 in Baltimore City?

```
SCC_vehicle <- SCC[grepl("Vehicles",SCC$EI.Sector, ignore.case = TRUE),]

NEI_vehicle <- NEI_BC[NEI_BC$SCC %in% SCC_vehicle$SCC,]
NEI_sum <- aggregate(Emissions ~ year, NEI_vehicle, sum)

#plotting

graph_3 <- ggplot(NEI_sum, aes(year, Emissions))

graph_3 + geom_point(col = "blue", size = 4) + geom_line(col = "blue", lwd = 1) + theme(axis.text.x = el
```




```
#saving the graph as png file

dev.copy (png, "plot5.png", width=480, height=480)
```

```
## png
##   3
```

```
dev.off()
```

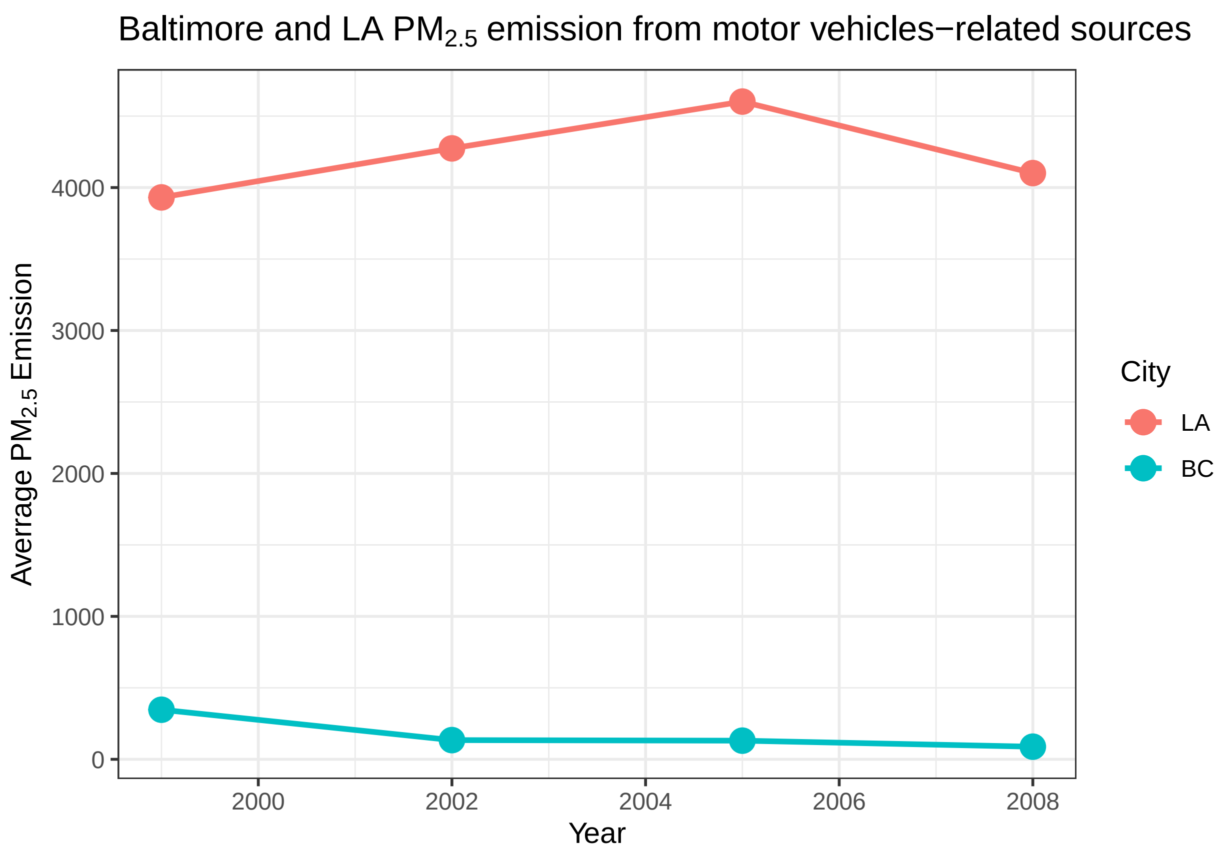
```
## pdf
##   2
```

Between 1999 and 2008 emission from motor vehicle sources decreased.

Question 6

Compare emissions from motor vehicle sources in Baltimore City with emissions from motor vehicle sources in Los Angeles County, California (fips = "06037"). Which city has seen greater changes over time in motor vehicle emissions?

```
NEI_BC_LA <- subset(NEI, fips %in% c("24510","06037"))

SCC_vehicle <- SCC[grepl("Vehicles",SCC$EI.Sector, ignore.case = TRUE),]

NEI_BC_LA_vehicle <- NEI_BC_LA[NEI_BC_LA$SCC %in% SCC_vehicle$SCC,]
NEI_BC_LA_sum <- aggregate(Emissions ~ year + fips, NEI_BC_LA_vehicle, sum)


#plotting

graph_4 <- ggplot(NEI_BC_LA_sum, aes(year, Emissions))

graph_4 + geom_point(aes(col = fips), size = 4) + geom_line(aes(col = fips), lwd = 1) + theme(axis.text
```




```r
#saving the graph as png file

dev.copy (png, "plot6.png", width=480, height=480)
```

```
## png
##   3
```

```r
dev.off()
```

```
## pdf
##   2
```

Changes in PM 2.5 emissions from otor vehicle sources in the cities analyzed have the opposite pattern. The data show that in Baltimore City, emission decreased during the investigated period of time. For Los Angles these sources increased.