ExploratoryData

1

This program assumes the required files are downloaded in working directory

and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

in the United States

from 1999 to 2008 as PNG in working directory

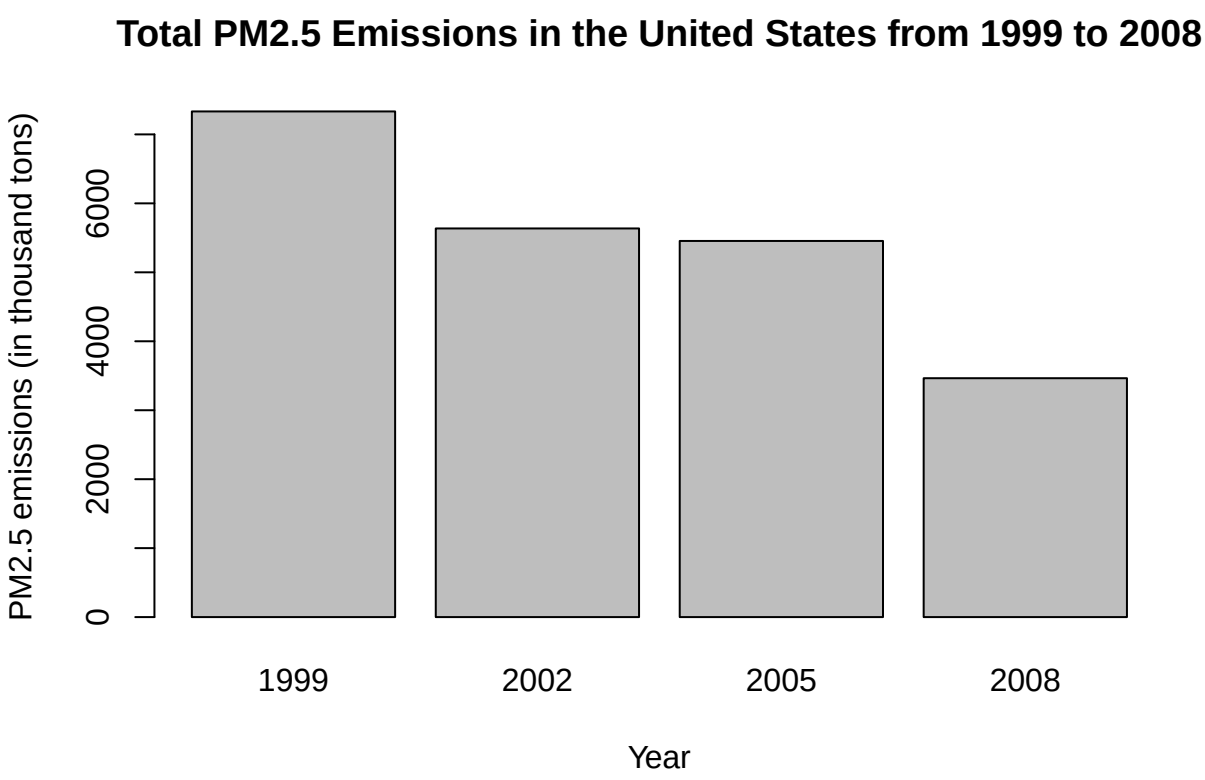
```
## Read the data
NEI <- readRDS("summarySCC_PM25.rds")
SCC <- readRDS("Source_Classification_Code.rds")
```

Q1

Have total emissions from PM2.5 decreased in the United States from 1999 to 2008? Using the base plotting system, make a plot showing the total PM2.5 emission from all sources for each of the years 1999, 2002, 2005, and 2008.

```
## organise the figures by year and scale them down
ds1 <- tapply(NEI$Emissions/1000, NEI$year, sum)
ds1.factor <- data.frame(year = names(ds1),
    Emissions = as.numeric(ds1))

## make the plot
barplot(ds1.factor$Emissions, names.arg = ds1.factor$year,
    ylab = "PM2.5 emissions (in thousand tons)", xlab = "Year",
    main = "Total PM2.5 Emissions in the United States from 1999 to 2008")
```



The total emissions of PM2.5 in the US from 1999 to 2008 have steadily decreased.

2

This program assumes the required files are downloaded in working directory and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

in the Baltimore City, Maryland

from 1999 to 2008 as PNG in working directory

Q2

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

```
## get the emissions figures
ds2 <- NEI[which(NEI$fips == "24510"), c("Emissions","year")]
```

```
## organise the figures by year and scale them down
ds2.t <- tapply(ds2$Emissions/1000, ds2$year, sum)
ds2.factor <- data.frame(year = names(ds2.t),
    Emissions = as.numeric(ds2.t))
```

from 1999 to 2008? Use the base plotting system to make a plot answering this question.

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

The total emissions of PM2.5 in the Baltimore City, Maryland from 1999 to 2008 showed a decreasing trend with vigorous fluctuation.

3

This program assumes the required files are downloaded in working directory

and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

by source types in the United States

from 1999 to 2008 as PNG in working directory

Q3

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?

```
## get the data
ds3 <- NEI[which(NEI$fips == "24510"), c("Emissions","year","type")]

## get the total numbers
ds3.agg <- aggregate(Emissions ~ type + year, data = ds3, FUN = sum)
ds3.agg$Emissions<- as.factor(ds3.agg$Emissions)
```

Which have seen increases in emissions from 1999–2008?

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

From 1999–2008 for Baltimore City, the 4 types of PM2.5 emissions showed a descreasing tendency.

PM2.5 emissions from NONPOINT sources slightly decreased and remained the greatest emissions over the period.

PM2.5 emissions from ON-ROAD sources steadily decreased.

PM2.5 emissions from POINT sources increased with fluctuations.

PM2.5 emissions from NON-ROAD sources decreased with fluctuations and became the least emissions sources in 2008.

4

This program assumes the required files are downloaded in working directory

and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

from coal combustion-related sources

in the United States from 1999 to 2008 as PNG in working directory

Q4

```
## get the SCC code for coal consumption related records
coalcode <- SCC[grep(".*[cC]oal.*", SCC$EI.Sector), c("SCC")]

## get the data according to the SCC codes
ds4 <- NEI[NEI$SCC %in% coalcode, c("Emissions","year")]

## organise the figures by year and scale them down
ds4.t <- tapply(ds4$Emissions/1000, ds4$year, sum)
ds4.factor <- data.frame(year = names(ds4.t),
    Emissions = as.numeric(ds4.t))
```

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

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

The PM2.5 emissions showed a decreasing trend throughout the period

and a remarkable drop from 2005 to 2008, which was approximately one-third of the emissions in 1999.

5

This program assumes the required files are downloaded in working directory

and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

from motor vehicle sources in Baltimore City

from 1999 to 2008 as PNG in working directory.

Q5

```
## get the SCC code for motor vehicle sources records
motorcode <- SCC[grep(".*[mM]otor.*", SCC$EI.Sector), c("SCC")]
vehiclecode <- SCC[grep(".*[vV]ehicle.*", SCC$EI.Sector), c("SCC")]

## get the data according to the SCC codes
ds5 <- NEI[which(NEI$fips == "24510"), c("Emissions","year","SCC")]
ds5 <- ds5[which(ds5$SCC %in% motorcode | ds5$SCC %in% vehiclecode),]

## organise the figures by year
ds5.t <- tapply(ds5$Emissions, ds5$year, sum)
ds5.factor <- data.frame(year = names(ds5.t),
    Emissions = as.numeric(ds5.t))
```

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

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

The PM2.5 emissions showed a promising decreasing trend.

The figure dropped more than halved from 1999 to 2002

and then moderately decreased from 2002 to 2008.

6

This program assumes the required files are downloaded in working directory

and the libraries are present.

It creates a bar chart showing total PM2.5 emissions

from motor vehicle sources in Baltimore City and Los Angeles County

from 1999 to 2008 as PNG in working directory.

Q6

Compare emissions from motor vehicle sources in Baltimore City

with emissions from motor vehicle sources in Los Angeles County, California (fips == "06037").

```
## get the SCC code for motor vehicle sources records
motorcode <- SCC[grep(".*[mM]otor.*", SCC$EI.Sector), c("SCC")]
vehiclecode <- SCC[grep(".*[vV]ehicle.*", SCC$EI.Sector), c("SCC")]

## get the data according to the fipds and SCC codes
ds6 <- NEI[which(NEI$fips == "24510" | NEI$fips == "06037"), c("Emissions","year","SCC","fips")]
ds6 <- ds6[which(ds6$SCC %in% motorcode | ds6$SCC %in% vehiclecode),]
ds6[which(ds6$fips == "24510"), c("fips")] <- "Baltimore"
ds6[which(ds6$fips == "06037"), c("fips")] <- "Los Angeles County"

library(ggplot2)

## get the total numbers
ds6.agg <- aggregate(Emissions ~ fips + year, data = ds6, FUN = sum)
ds6.agg$Emissions<- as.factor(ds6.agg$Emissions)
```

Which city has seen greater changes over time in motor vehicle emissions?

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

The PM2.5 emissions in Baltimore City steadily decreased over the 10 years.

Meanwhile the emissions in L. A. County rose significantly from 1999 to 2005

although a slight decline followed until 2008.

L. A. County always emitted more than Baltimore City did by 10-50 times during the period.

The difference was enlarging by the drop in Baltimore and the increase in L. A..