

Week 4 Programming Assignment

Ken Wood

8/2/2020

```
# Load the raw data files.  
# These lines of code will take a little time to execute, so please be patient!  
  
NEI <- readRDS("exdata-data-NEI_data/summarySCC_PM25.rds")  
SCC <- readRDS("exdata-data-NEI_data/Source_Classification_Code.rds")  
merged_df <- merge(NEI,SCC,by="SCC")
```

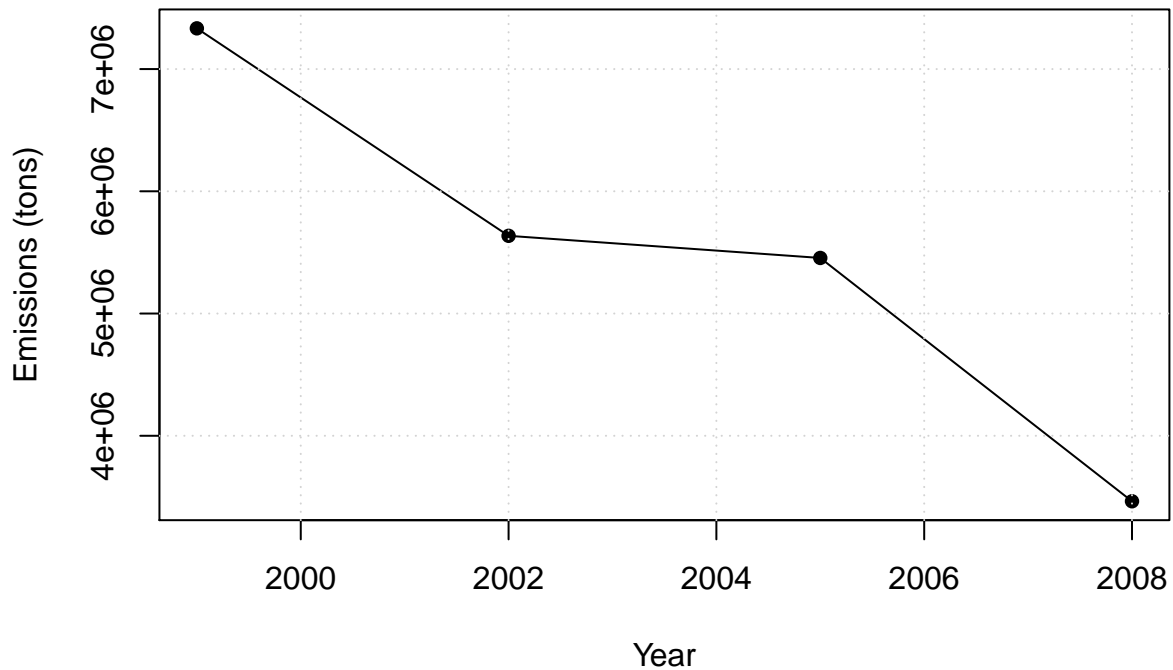
Questions

We will address the following questions and tasks in our exploratory analysis. For each question/task we will need to make a single plot. Unless specified, we can use any plotting system in R to make our plot.

1. 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.

```
total_emissions <- aggregate(NEI$Emission, by=list(NEI$year), sum)  
plot(total_emissions,pch=16,xlab="Year",ylab="Emissions (tons)",main="Total Emissions by Year")  
lines(total_emissions$Group.1,total_emissions$x)  
grid(lty="dotted")
```

Total Emissions by Year

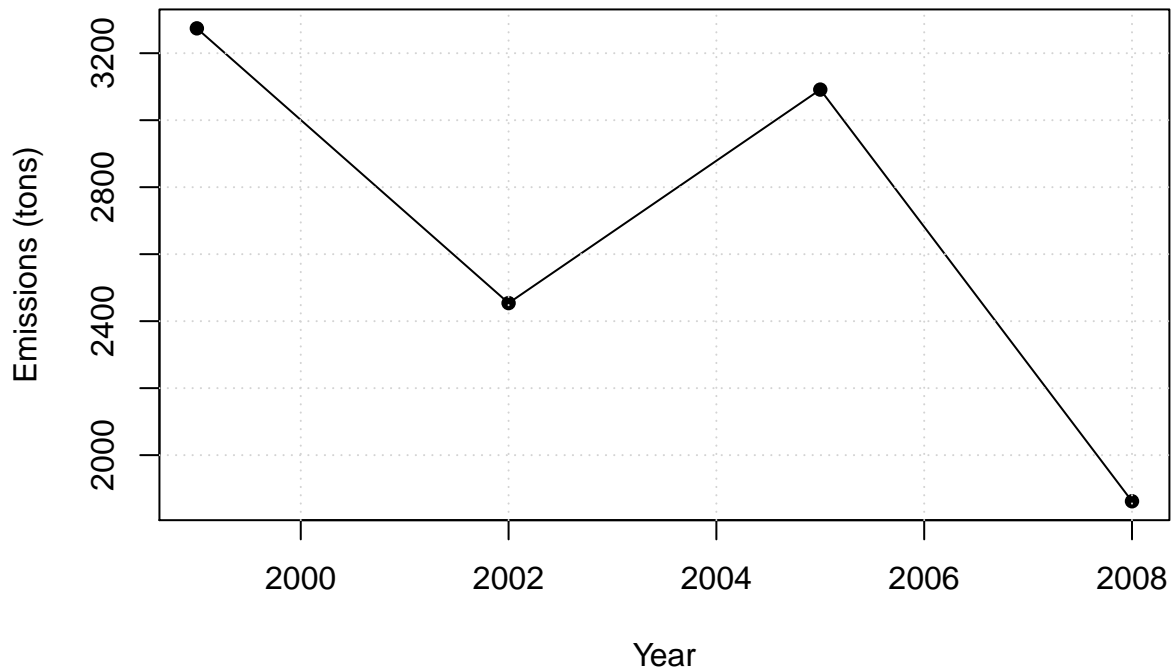


2. Have total emissions from PM2.5 decreased in the **Baltimore City, Maryland** (`fips == "24510"`) from 1999 to 2008? Use the base plotting system to make a plot answering this question.

```
# subset NEI for Baltimore City (fips=="24510")
baltimore_city <- subset(NEI,fips=="24510")
balt_emissions <- aggregate(baltimore_city$Emission, by=list(baltimore_city$year), sum)

plot(balt_emissions,pch=16,xlab="Year",ylab="Emissions (tons)",main="Total Emissions by Year: Baltimore")
lines(balt_emissions$Group.1,balt_emissions$x)
grid(lty="dotted")
```

Total Emissions by Year: Baltimore–City

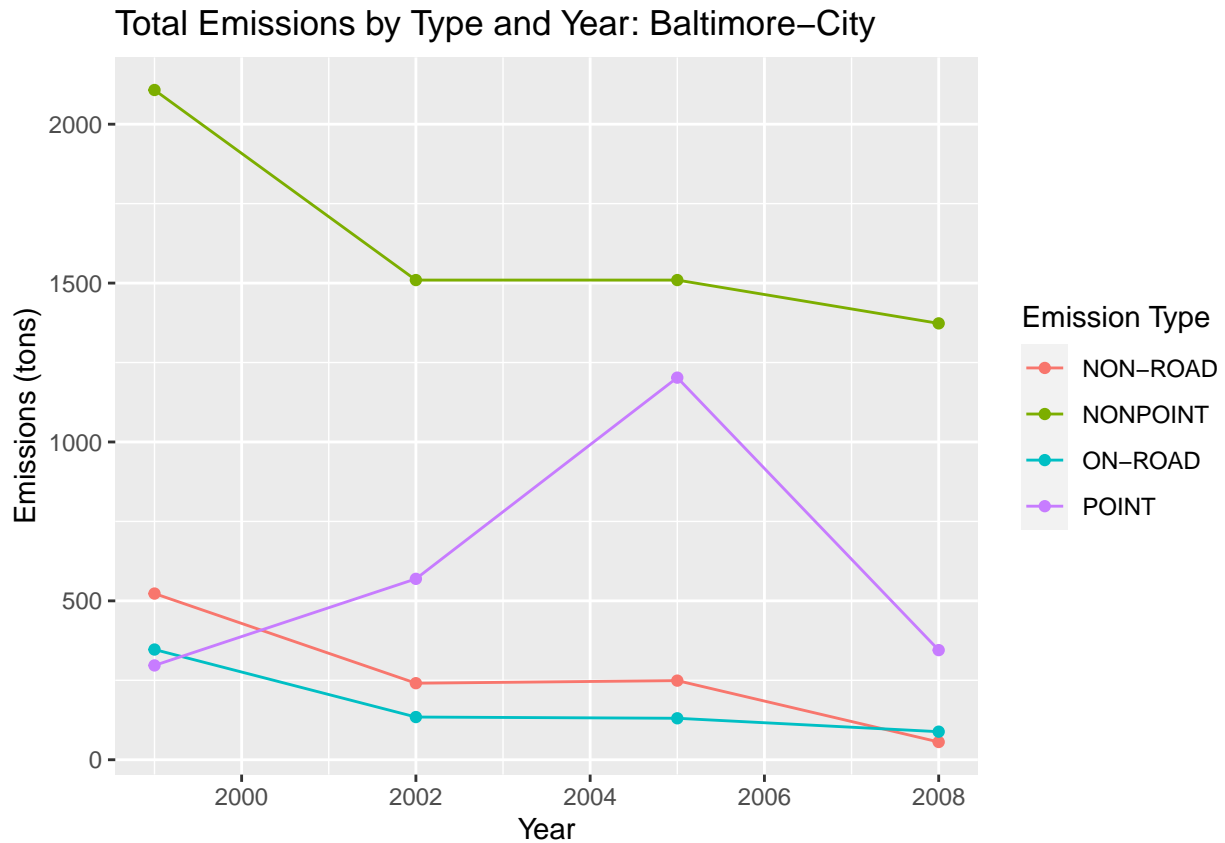


3. Of the four types of sources indicated by the `type` (point, nonpoint, onroad, nonroad) variable, which of these four sources have seen decreases in emissions from 1999–2008 for **Baltimore City**? Which have seen increases in emissions from 1999–2008? Use the **ggplot2** plotting system to make a plot answer this question.

```
library(ggplot2)
```

```
balt_emissions_type_year <- aggregate(baltimore_city$Emission, by=list(baltimore_city$type,baltimore_ci  
names(balt_emissions_type_year) <- c("type","Year","Emissions")
```

```
ggplot(balt_emissions_type_year, aes(Year,Emissions))+ geom_point(aes(color=type))+geom_line(aes(color=
```



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

```
library(dplyr)
```

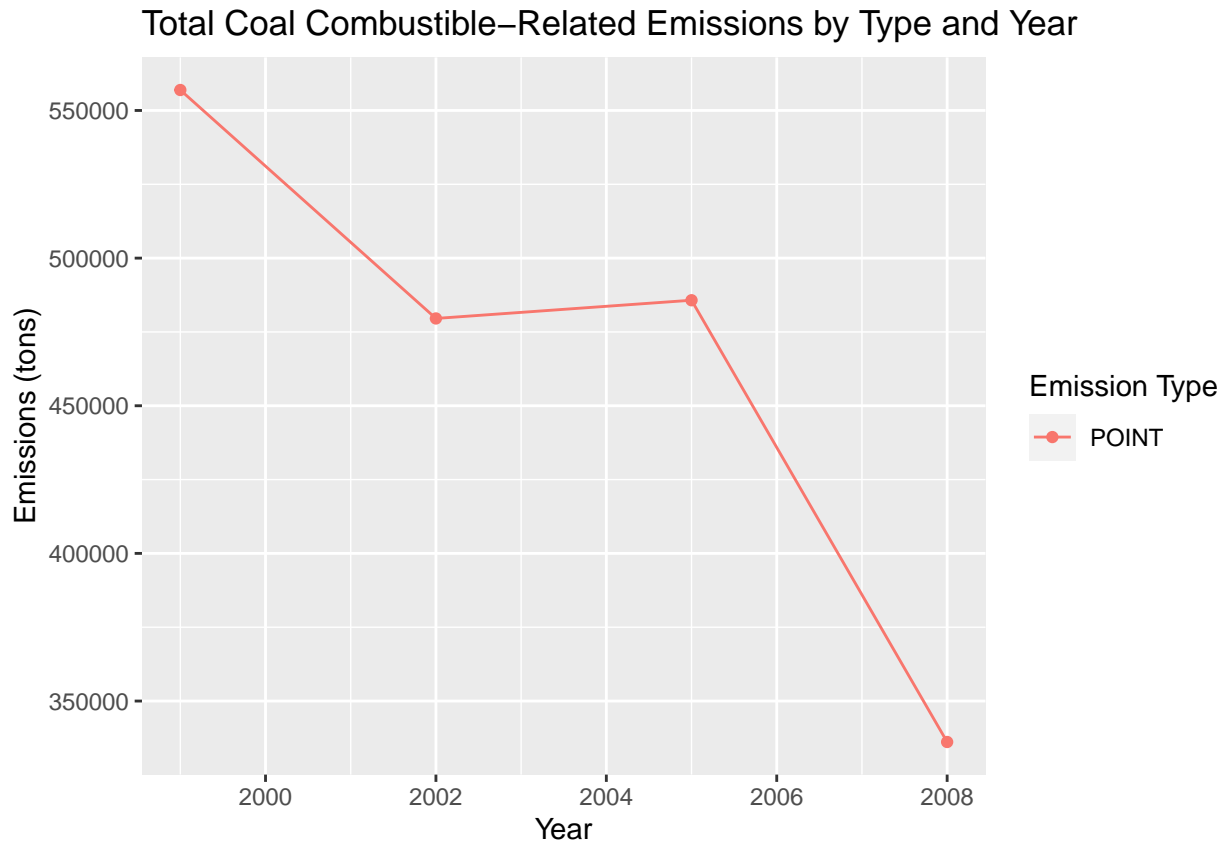
```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
# Filter the merged df for 'coal combustion-related' sources
```

```
coal_combustion <- dplyr::filter(merged_df, grepl('Coal', Short.Name) & grepl('Ext Comb', Short.Name))
emissions_type_year <- aggregate(coal_combustion$Emission, by=list(coal_combustion$type, coal_combustion$Year), FUN=sum, na.rm=TRUE)
names(emissions_type_year) <- c("type", "Year", "Emissions")
```

```
ggplot(emissions_type_year, aes(Year, Emissions)) + geom_point(aes(color=type)) + geom_line(aes(color=type))
```



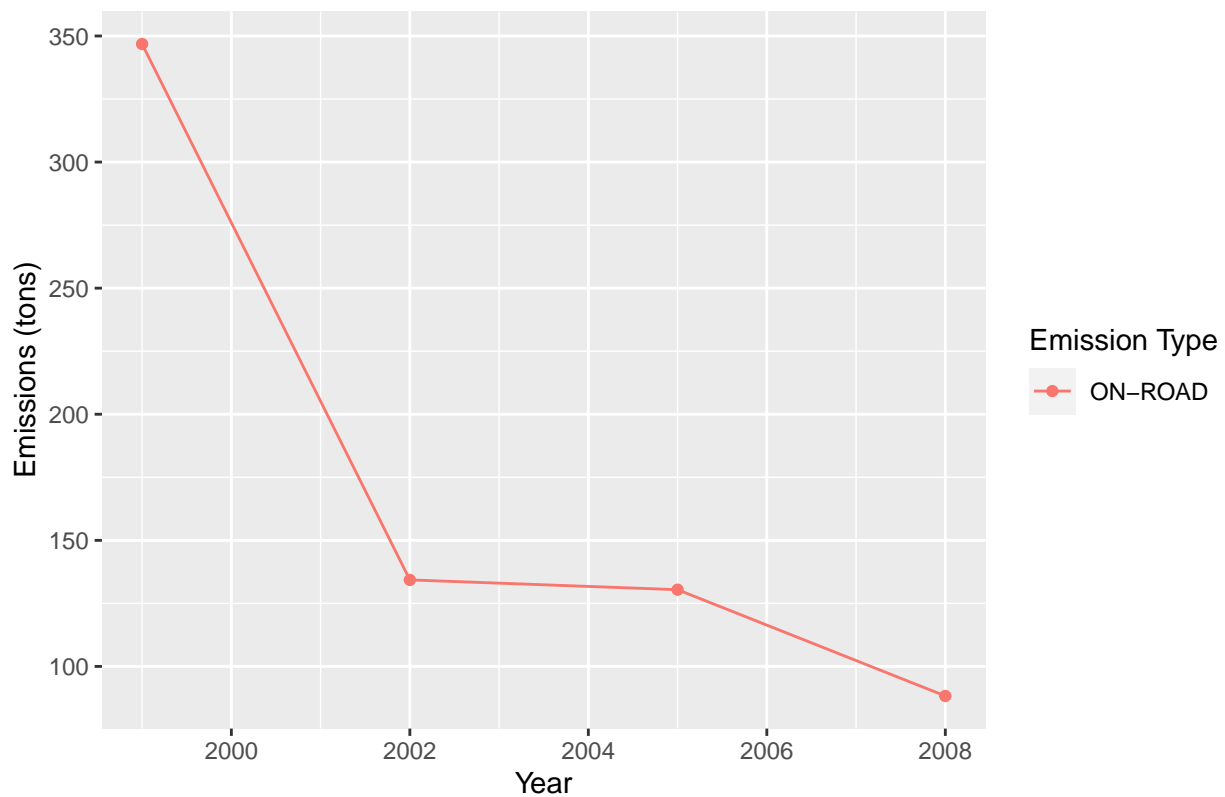
5. How have emissions from motor vehicle sources changed from 1999–2008 in **Baltimore City**?

```
library(dplyr)
library(ggplot2)

# Filter the merged_df for 'Baltimore City' and 'Highway Veh'
balt_cars <- dplyr::filter(merged_df, grepl("24510",fips) & grepl('Highway Veh',Short.Name))
balt_car_emissions <- aggregate(balt_cars$Emission, by=list(balt_cars$type,balt_cars$year), sum)
names(balt_car_emissions) <- c("type","Year","Emissions")

ggplot(balt_car_emissions, aes(Year,Emissions))+ geom_point(aes(color=type))+geom_line(aes(color=type)).
```

Total Motor Vehicle Emissions by Type and Year: Baltimore–City



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?

```
library(dplyr)
library(ggplot2)

# Filter the merged_df for fips==24510/fips==06037 and 'Highway Veh'
balt_la_cars <- dplyr::filter(merged_df, grepl("24510|06037",fips) & grepl('Highway Veh',Short.Name))
balt_la_car_emissions <- aggregate(balt_la_cars$Emission, by=list(balt_la_cars$fips,balt_la_cars$year),
names(balt_la_car_emissions) <- c("fips","Year","Emissions")

ggplot(balt_la_car_emissions, aes(Year,Emissions))+ geom_point(aes(color=fips))+geom_line(aes(color=fips))
```

Total Motor Vehicle Emissions by FIPS Location and Year

