Exploratory Analysis

Week-4 Assignment

Badarinadh Vissapragada

Fine particulate matter (PM2.5) is an ambient air pollutant for which there is strong evidence that it is harmful to human health. In the United States, the Environmental Protection Agency (EPA) is tasked with setting national ambient air quality standards for fine PM and for tracking the emissions of this pollutant into the atmosphere. Approximatly every 3 years, the EPA releases its database on emissions of PM2.5. This database is known as the National Emissions Inventory (NEI). You can read more information about the NEI at the EPA National Emissions Inventory web site. For each year and for each type of PM source, the NEI records how many tons of PM2.5 were emitted from that source over the course of the entire year. The data that you will use **Dataless**

The data for this assignment are available from the course web site as a single zip file:

- Data for Peer Assessment [29Mb]

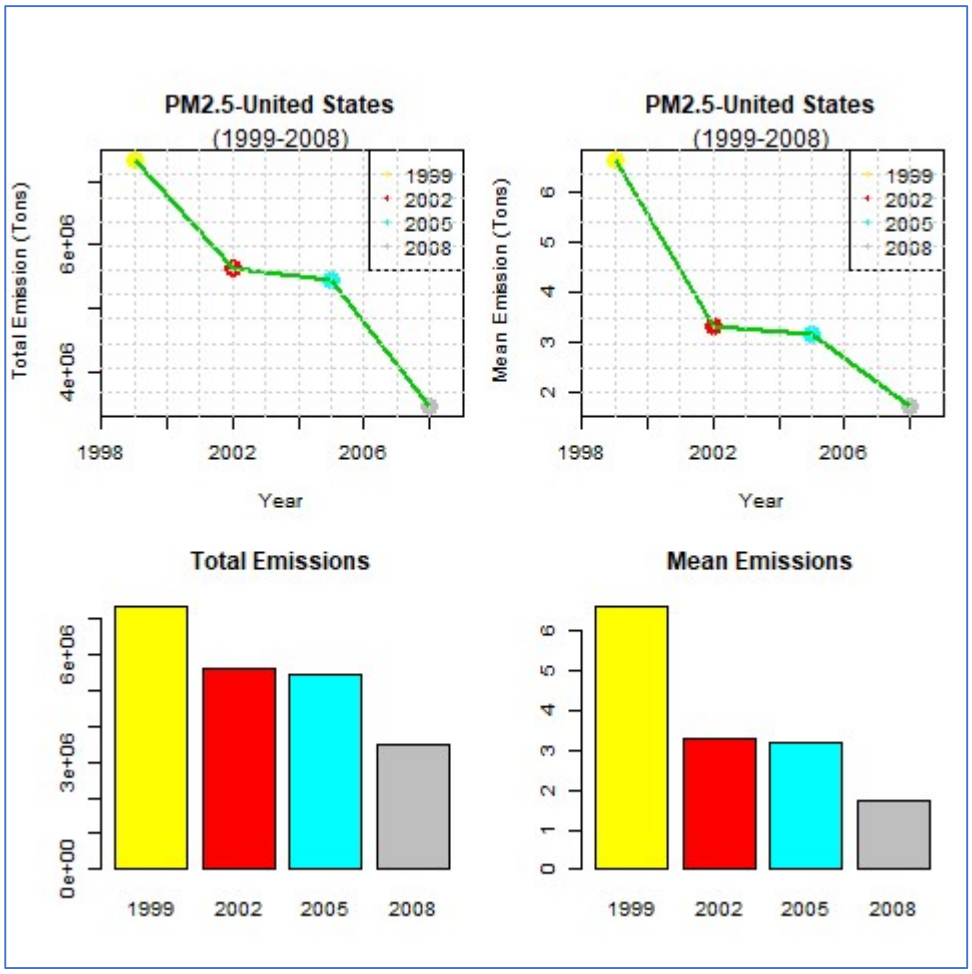
The zip file contains two files:

PM2.5 Emissions Data (summarySCC_PM25.rds): This file contains a data frame with all of the PM2.5 emissions data for 1999, 2002, 2005, and 2008. For each year, the table contains number of **tons** of PM2.5 emitted from a specific type of source for the entire year. Here are the first few rows.

- 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

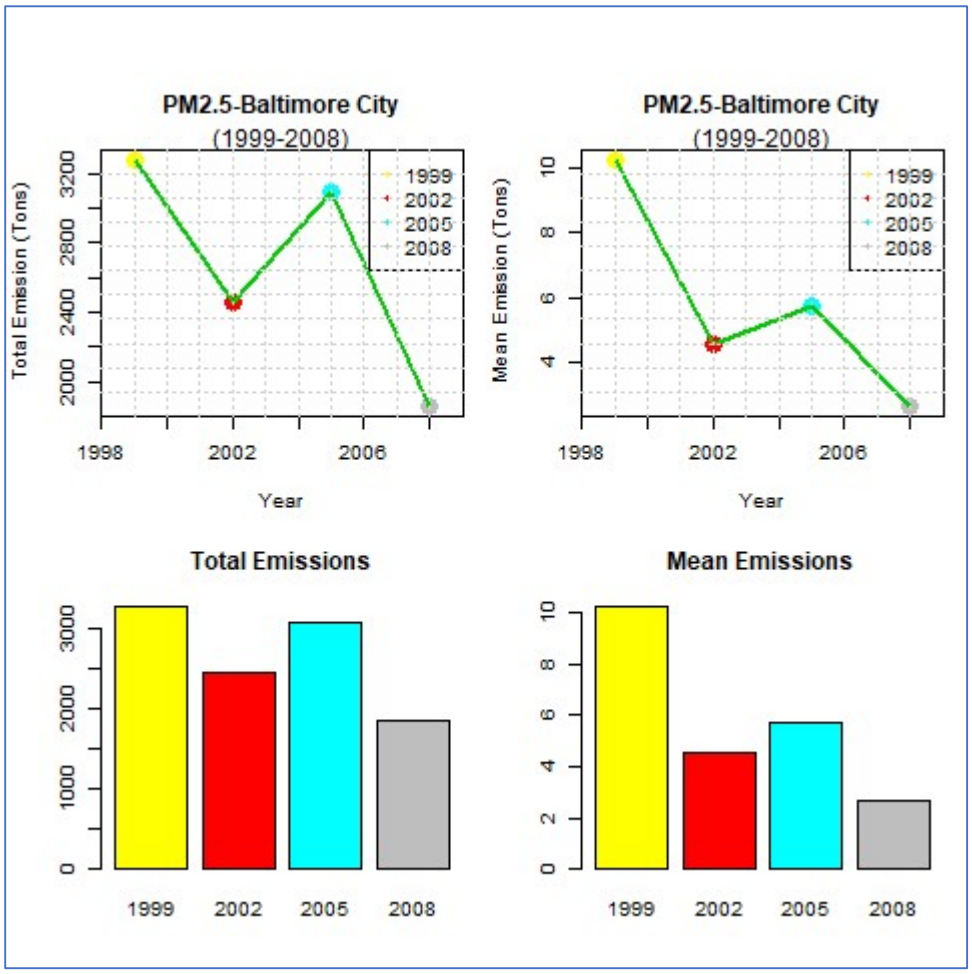
Source Classification Code Table Source_Classification_Code.rds): This table provides a mapping from the SCC digit strings in the Emissions table to the actual name of the PM2.5 source. The sources are categorized in a few different ways from more general to more specific and you may choose to explore whatever categories you think are most useful. For example, source “10100101” is known as “Ext Comb /Electric Gen /Anthracite Coal /Pulverized Coal”.

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.*



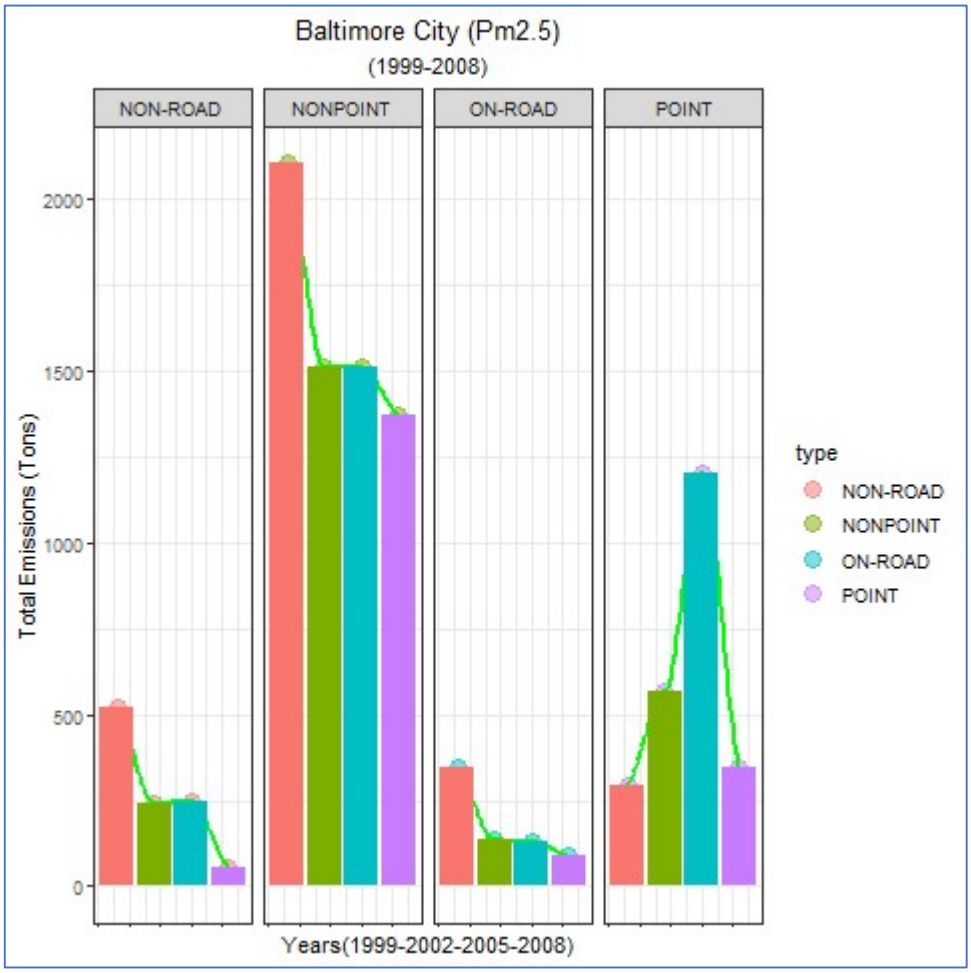
Interpretation: The total emissions of PM2.5 in the United States decreased significantly from the year 1999 t0 2002. During 2002-2005 the rate is slowed down. Between 2005-2008 again there is a good drop down of the PM2.5 emissions. The mean and total emissions are following the same trend indicating that the total decrease not just due to monitoring but overall impact of reduced emissions.

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.*



Interpretation: The total emissions of PM2.5 in the Baltimore City decreased significantly from the year 1999 t0 2002. During 2002-2005 the total emissions increased. Between 2005-2008 again there is a good drop down of the PM2.5 emissions. The mean emissions show the same trend but the increase between 2002-2005 is less drastic. Need to investigate why there is an increase in PM2.5 emissions during the 2002-2005.

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.*



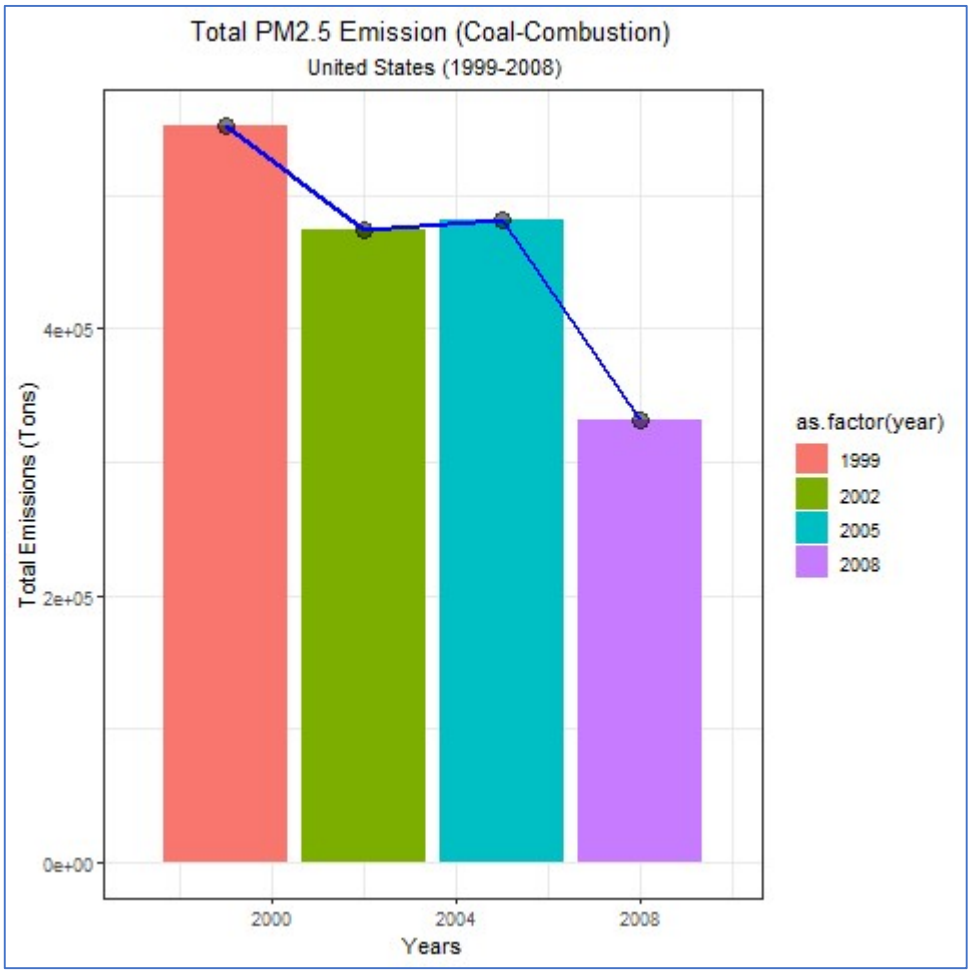
Interpretation:

a. The total emissions of PM2.5 in the Baltimore city for "Non-Road, "Non-Point", "On-Road" types show there is general trend of decrease 1999-2008 following the same trend as the total emissions from all sources in the national trend.

b. The total emissions from the "Point" type show an increase between 1999-2005 and dropped between 2005-2008. This "point" type may be responsible for the Baltimore overall increase of emissions we observed between 2005-2008.

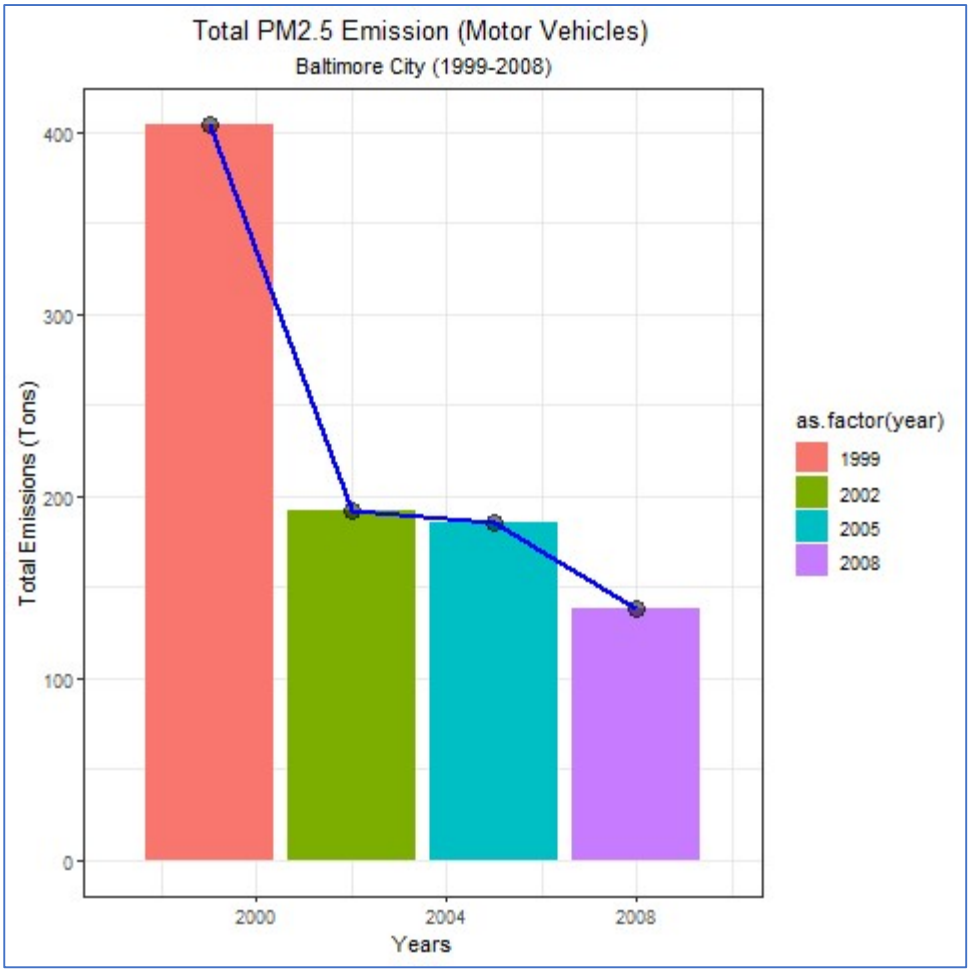
c. In general, the "Non-Point" type has highest level of emissions though it is good that there is overall decrease between 1999-2008. The "On-Road type" has lowest total emissions of PM2.5

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



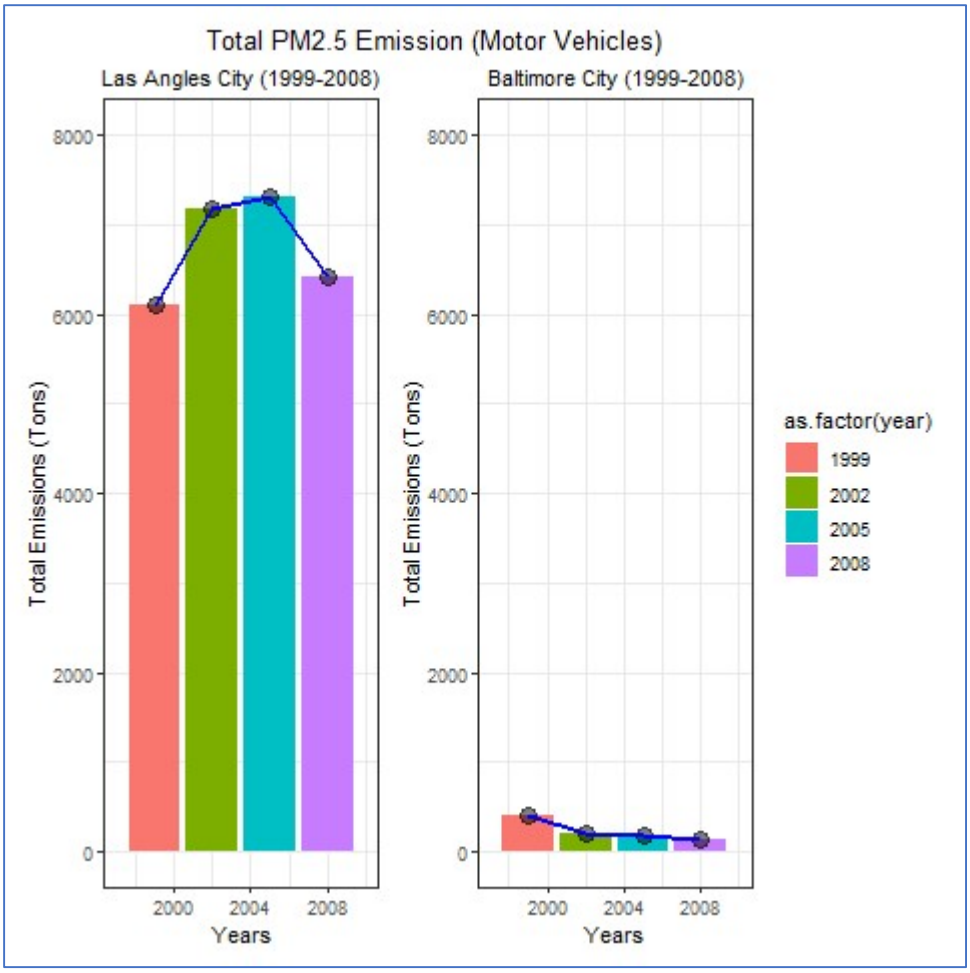
Interpretation: The total emissions of PM2.5 due to Coal-Combustion in the United States show overall decrease from the year 1999 t0 2008.

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



Interpretation: The total emissions of PM2.5 due to Motor Vehicles in Baltimore City in the show an overall decrease during 1999-2008. The trend is consistent with overall drop of PM 2.5 Total and mean emissions in the United States.

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?*



Interpretation:

a. The Las Angles city is way high in PM2.5 emissions due to Motor vehicles compared to the Baltimore city.
b. The total emissions of PM2.5 due to Motor Vehicles in the Las Angles increased between 1999-2005 in contrast to the Baltimore city, where the trend showed a significant drop.
c. PM2.5 trend shows decrease between 2005-2008
d. The PM2.5 emissions due to motor vehicles showed overall decrease in the Baltimore city .