

Project 2: Fine particulate matter (PM2.5)

```
library(tidyverse)

## Warning: package 'tidyverse' was built under R version 3.6.2
## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.2.1      v purrr   0.3.3
## v tibble  2.1.3      v dplyr  0.8.3
## v tidyr   1.0.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## Warning: package 'ggplot2' was built under R version 3.6.2
## Warning: package 'tidyr' was built under R version 3.6.2
## Warning: package 'readr' was built under R version 3.6.2
## Warning: package 'purrr' was built under R version 3.6.2
## Warning: package 'forcats' was built under R version 3.6.2

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(scales)

##
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':
##
##   discard

## The following object is masked from 'package:readr':
##
##   col_factor
```

0 Introduction

0.1 PM2.5 Emissions Data (NEI):

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.

Variables:

- 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

```
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

0.2 Source Classification Code Table (SCC):

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.

```
head(SCC)
```

##	SCC	Data.Category	Short.Name
## 1	10100101	Point	
## 2	10100102	Point	
## 3	10100201	Point	
## 4	10100202	Point	
## 5	10100203	Point	
## 6	10100204	Point	
## 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	
## 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	
## 1		External Combustion Boilers Electric Generation	
## 2		External Combustion Boilers Electric Generation	
## 3		External Combustion Boilers Electric Generation	
## 4		External Combustion Boilers Electric Generation	
## 5		External Combustion Boilers Electric Generation	
## 6		External Combustion Boilers Electric Generation	
## 1		Anthracite Coal	
## 2		Anthracite Coal	
## 3		Bituminous/Subbituminous Coal	
## 4		Bituminous/Subbituminous Coal	
## 5		Bituminous/Subbituminous Coal	
## 6		Bituminous/Subbituminous Coal	
## 1		Pulverized Coal	NA
## 2		Traveling Grate (Overfeed) Stoker	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
```

1

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

Yes, total emissions have decreased.

```
year=c(1999,2002,2005,2008)
total_emissions=c()

for (i in 1:4){
  NEI_year=NEI[NEI$year==year[i],]
  total_emissions[i]=sum(NEI_year$Emissions)
  assign(paste0("NEI_year_",year[i]),NEI_year)
}

png(filename="plot1.png")
barplot(total_emissions,names = year, xlab = "Years", ylab = "Tons",
        main = "Total PM2.5 emissions in the United States")
dev.off()
```

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

2

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

Yes, total emissions in Baltimore have decreased.

```
total_emissions_Baltimore=c()

for (i in 1:4){
  NEI_year=get(paste0("NEI_year_",year[i]))
  NEI_year_Baltimore=NEI_year[NEI_year$fips=="24510",]
  total_emissions_Baltimore[i]=sum(NEI_year_Baltimore$Emissions)
}

png(filename="plot2.png")
barplot(total_emissions_Baltimore,names = year, xlab = "Years", ylab = "Tons",
        , main = "Total PM2.5 emissions in Baltimore City, 1999-2008")
dev.off()
```

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

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?

Onroad, Non-road and Nonpoint have seen decreases in emissions. Point has seen decreases in emissions.

```
png(filename="plot3.png")
NEI%>%filter(fips=="24510")%>%
  select(Emissions,type,year)%>%
  mutate(type=factor(type,levels = c("ON-ROAD","NON-ROAD","POINT","NONPOINT")),
         year=factor(year))%>%
  group_by(type,year)%>%
  summarise(total=sum(Emissions))%>%#ungroup()%>%
  ggplot(aes(x=year,y=total))+
  geom_bar(stat="identity")+
  facet_wrap(~type,ncol=4)+
  labs(title = "Total PM2.5 emissions by types in Baltimore City",
       y="Tons")+
  theme(plot.title = element_text(hjust = 0.5))
dev.off()
```

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

4

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

Emissions from coal combustion-related sources have decreased.

```
SCC_Comb.Coal=SCC%>%
  filter(str_detect(Short.Name,"Comb|Coal"))%>%
  pull(SCC)

png(filename="plot4.png")
NEI%>%filter(SCC %in% SCC_Comb.Coal)%>%
  select(Emissions,year)%>%
  mutate(year=factor(year))%>%
  group_by(year)%>%
  summarise(total=sum(Emissions))%>%
  ggplot(aes(x=year,y=total))+
  geom_bar(stat="identity")+
  labs(title = "Total PM2.5 emissions from coal combustion-related sources",
       y="Tons")+
  theme(plot.title = element_text(hjust = 0.5))
dev.off()
```

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

5

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

Emissions from motor vehicle sources have decreased.

```
SCC_vehicle=SCC%>%
  filter(str_detect(SCC.Level.Two,"Vehicle"))%>%
  pull(SCC)

png(filename="plot5.png")
NEI%>%filter(fips=="24510" & SCC %in% SCC_vehicle)%>%
  select(Emissions,year)%>%
  mutate(year=factor(year))%>%
  group_by(year)%>%
  summarise(total=sum(Emissions))%>%
  ggplot(aes(x=year,y=total))+
  geom_bar(stat="identity")+
  labs(title = "Total PM2.5 emissions from motor vehicle sources in Baltimore City",
       y="Tons")+
  theme(plot.title = element_text(hjust = 0.5))
dev.off()
```

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

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?

Baltimore City has seen greater changes in motor vehicle emissions, we can tell from dividing emissions in later years by emissions in 1999.

```
(NEI_vehicle_city=NEI%>%
  filter(fips %in% c("24510","06037") & SCC %in% SCC_vehicle)%>%
  dplyr::select(Emissions,fips,year)%>%
  mutate(fips=factor(fips,levels=c("24510","06037"),labels=c("Baltimore City","Los Angeles")),
        year=factor(year))%>%
  group_by(fips,year)%>%
  summarise(total=sum(Emissions))%>%
  ungroup()%>%group_by(fips)%>%
  mutate(percent_year1999=percent(total/total[1])))
```

```
## # A tibble: 8 x 4
## # Groups:   fips [2]
##   fips      year  total percent_year1999
##   <fct>    <fct> <dbl> <chr>
## 1 Baltimore City 1999  404. 100.0%
## 2 Baltimore City 2002  192. 47.6%
## 3 Baltimore City 2005  185. 45.9%
## 4 Baltimore City 2008  138. 34.2%
## 5 Los Angeles   1999 6110. 100.0%
## 6 Los Angeles   2002 7189. 117.7%
## 7 Los Angeles   2005 7304. 119.5%
```

```
## 8 Los Angeles      2008  6421. 105.1%
png(filename="plot6.png")
NEI_vehicle_city%>%ggplot(aes(x=year,y=total))+
  geom_bar(stat="identity")+
  facet_wrap(~fips,ncol=2)+
  labs(title = "Total PM2.5 emissions from motor vehicle sources in Baltimore and Los Angeles",
        y="Tons")+
  theme(plot.title = element_text(hjust = 0.5))
dev.off()

## pdf
## 2
```