<u>Peer Assessments (https://class.coursera.org/exdata-030/human\_grading/)</u> / Course Project 2 Help Center (https://accounts.coursera.org/i/zendesk/courserahelp?return\_to=https://learner.coursera.help/hc)

due in 2wk 6d
Submission Phase
1. Do assignment ☐ (/exdata-030/human_grading/view/courses/975125/assessments/4/submissions)
Evaluation Phase
2. Evaluate peers
Results Phase
3. See results
In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.  Save draft  Submit for grading

# Introduction

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 (http://www.epa.gov/ttn/chief/eiinformation.html).

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 for this assignment are for 1999, 2002, 2005, and 2008.

## **Data**

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

<u>Data for Peer Assessment (https://d396qusza40orc.cloudfront.net/exdata%2Fdata%2FNEI\_data.zip)</u>
 [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 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
```

- 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".

You can read each of the two files using the readRDS() function in R. For example, reading in each file can be done with the following code:

```
## This first line will likely take a few seconds. Be patient!
NEI <- readRDS("summarySCC_PM25.rds")
SCC <- readRDS("Source_Classification_Code.rds")</pre>
```

as long as each of those files is in your current working directory (check by calling dir() and see if those files are in the listing).

# **Assignment**

The overall goal of this assignment is to explore the National Emissions Inventory database and see what it say about fine particulate matter pollution in the United states over the 10-year period 1999–2008. You may use any R package you want to support your analysis.

#### Questions

You must address the following questions and tasks in your exploratory analysis. For each question/task you will need to make a single plot. Unless specified, you can use any plotting system in R to make your 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.
- 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.
- 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.
- 4. Across the United States, how have emissions from coal combustion-related sources changed from 1999–2008?
- 5. How have emissions from motor vehicle sources changed from 1999–2008 in **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?

## **Making and Submitting Plots**

For each plot you should

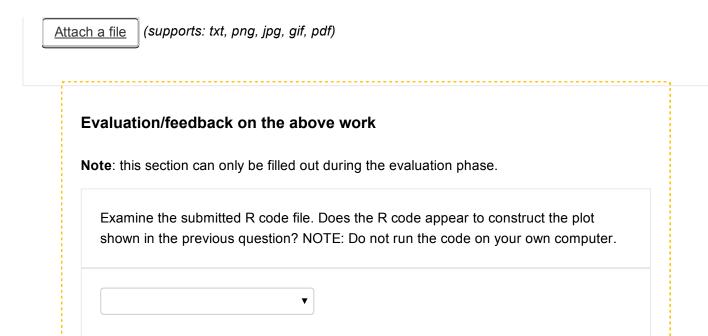
- · Construct the plot and save it to a PNG file.
- Create a separate R code file (plot1.R), plot2.R, etc.) that constructs the corresponding plot, i.e. code in plot1.R constructs the plot1.png plot. Your code file should include code for reading the data so that the plot can be fully reproduced. You must also include the code that creates the PNG file. Only include the code for a single plot (i.e. plot1.R) should only include code for producing plot1.png)
- Upload the PNG file on the Assignment submission page
- Copy and paste the R code from the corresponding R file into the text box at the appropriate point in the peer assessment.

Have total emissions from  $PM_{2.5}$  decreased in the United States from 1999 to 2008? Using the **base** plotting system, make a plot showing the *total*  $PM_{2.5}$  emission from all sources for each of the years 1999, 2002, 2005, and 2008.

Upload a PNG file containing your plot addressing this question.

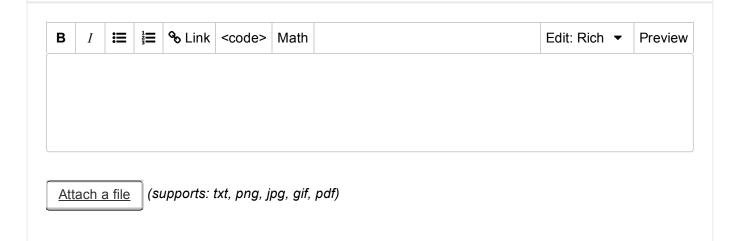
Evaluation/feedback on the above work			
Note: this section can only be filled out during the evaluati	on phase.		
Please view the plot for this question. Does the plot apple being asked? In other words, can you answer the questions shown in the plot?		-	
▼			
y and paste the R code file for the plot uploaded in the previ	ous question.		
by and paste the R code file for the plot uploaded in the previous I I I I I I I I I I I I I I I I I I I		dit: Rich ▼	Prev
by and paste the R code file for the plot uploaded in the previous I		dit: Rich ▼	Prev
		dit: Rich ▼	Prev
I I I I I I I I I I I I I I I I I I I		dit: Rich ▼	Prev
I		dit: Rich ▼	Pre

Have total emissions from  $PM_{2.5}$  decreased in the **Baltimore City**, Maryland ( fips == 24510 ) from 1999 to 2008? Use the **base** plotting system to make a plot answering this question. Upload a PNG file containing your plot addressing this question. % Link | <code> | Math В Edit: Rich ▼ Preview Attach a file (supports: txt, png, jpg, gif, pdf) Evaluation/feedback on the above work **Note**: this section can only be filled out during the evaluation phase. Please view the plot for this question. Does the plot appear to address the question being asked? In other words, can you answer the question using the information shown in the plot? Copy and paste the R code file for the plot uploaded in the previous question. Edit: Rich ▼ Preview



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.

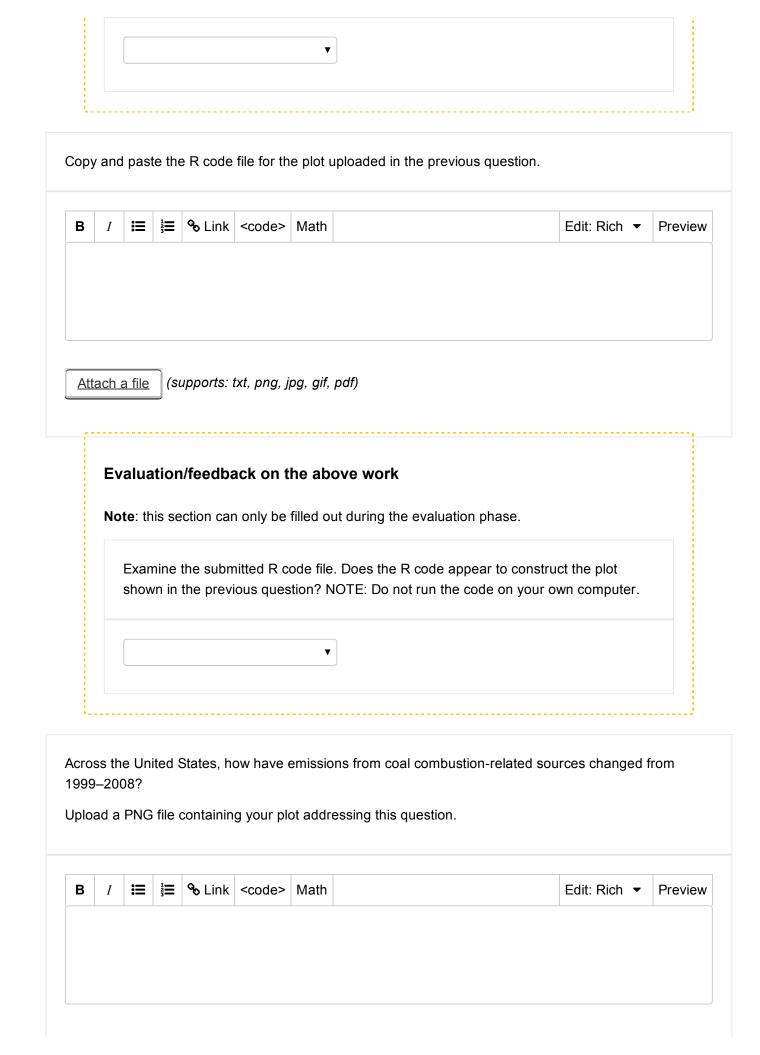
Upload a PNG file containing your plot addressing this question.



### Evaluation/feedback on the above work

**Note**: this section can only be filled out during the evaluation phase.

Please view the plot for this question. Does the plot appear to address the question being asked? In other words, can you answer the question using the information shown in the plot?



s the question information  Edit: Rich	Previ
information	Previ
Edit: Rich ▼	Previ
ruct the plot	

В	I	≔	1 2 3	<b>%</b> Link	<code></code>	Math				Edit: Rich	า ▼	Previe
Att	ach	a file	(sı	upports: 1	txt, png, j <sub>l</sub>	pg, gif,	odf)					
							ve work t during the e	valuatior	n phase.			
		DI										
		being	g ask	-	ther word	-	on. Does the prou answer th				on	
		being	g ask	ed? In o	ther word	-					on	
		being	g ask	ed? In o	ther word	s, can					on	
opy	/ and	being	g ask /n in	the plot?	ther word	s, can		e questic	on using the	e information	on	
Сору	/ and	being	g ask /n in	ed? In of the plot?	ther word	e plot u	ou answer th	e questic	on using the	e information		Previe
		being show	g ask vn in	ed? In of the plot?	file for th	e plot u	ou answer th	e questic	on using the	e information		Previe

								e appear t		•		
	shov	wn in t	he previ	ous ques	stion? N	IOTE: D	o not run	the code	on your o	wn compute	er.	
•								•		from motor en greater o		
			e emissi ontainin	ions? g your plo	ot addr	essing th	his questi	on.				
3	<i>I</i>	1-	0 1:-1-								_	Previe
		1 =	*S LINK	<code></code>	Math					Edit: Rich	•	1.000
Attac	ch a file			<code></code>		pdf)				Edit: Rich		
	Evalu	(su <sub>l</sub>	pports: i	txt, png, j <sub>l</sub>	pg, gif, he ab	ove wo		uation pha	ase.	Edit: Rich		
	<b>Evalu</b> a <b>Note</b> : t Plea bein	ation/ his sec	pports: i	ack on to a contract on the contract only be the contract of t	pg, gif, he ab	ove would during on. Doe	g the eval		address	the questio		

В	I	<b>≔</b>	1 2 3	& Link	<code></code>	Math				Edit: Rich	•	Prev
Att		a file				pg, gif, pdf) he above	work					
	No	Exar	nine	the subn	nitted R c	filled out du	es the R co	ode appear	to constru	•		
		SHOV	VIII III	the prev	ious ques	vtion? NOTE	Do not ro	in the code	on your o	wir comput	σι.	
	Over	all e	valu	uation/f	eedback	·						
N	lote:	this	secti	on can o	nly be fille	ed out during	g the evalu	ation phas	e.			
	the	wor	k. Po	int out th	ne submis	provide consion's strençolain your gr	gths as we	ll as areas				

$\hfill \square$ In accordance with the Honor Code, I certify that my answers here are my own work, and that
have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft

Submit for grading