```
    #move to the correct working directory
    setwd("~/Documents/COURSERA/4.ExploratoryDataAnalysis/project2/exdata-data-
    NEI_data")

    #read the data to use.
    NEI <- readRDS("summarySCC_PM25.rds")
    head(NEI)
    str(NEI)

    #install data.table
    install.packages('data.table')
    library(data.table)

    #plot1
df <- data.table(NEI)
by_year <- df[, list(emissions=sum(Emissions)), by=year]
by_year$year = as.numeric(as.character(by_year$year))
by_year$emissions = as.numeric(as.character(by_year$emissions))
plot(by_year$year, by_year$emissions, type='l', ylab='Emissions', xlab='Year')
dev.copy(png, file="plot1_EG.png", width=480, height=480)
dev.off()

#plot2
Baltimore <- subset(df, fips == '24510')
by_year <- Baltimore[, list(emissions=sum(Emissions)), by=year]
by_year$year = as.numeric(as.character(by_year$year))
by_year$emissions = as.numeric(as.character(by_year$emissions))
#print on screen to see it
plot(by_year$year, by_year$emissions, type='l', ylab='Emissions', xlab='Year')
#print png
dev.copy(png, file="plot2_EG.png", width=480, height=480)
dev.off()

#plot3
#only change from previous plot2 : new library to use. The rest is the same
library('ggplot2')

NEI <- readRDS("summarySCC_PM25.rds")
head(NEI)
str(NEI)
df <- data.table(NEI)
Baltimore <- subset(df, fips == '24510')
by_year <- Baltimore[, list(emissions=sum(Emissions)), by=c('year', 'type')]
by_year$year = as.numeric(as.character(by_year$year))
by_year$emissions = as.numeric(as.character(by_year$emissions))

#print on screen to see it
```

```
ggplot(data=by_year, aes(x=year, y=emissions, col=type)) + geom_line() + geom_point() + ggtitle("Emissions

    #print png plot3_EG.png
    dev.copy(png, file="plot3_EG.png", width=480, height=480)
    dev.off()


    #plot4 (takes time the processing of data)
    NEI <- readRDS("summarySCC_PM25.rds")
    head(NEI)
    str(NEI)
    SCC <- readRDS("Source_Classification_Code.rds")
    head(SCC)
    str(SCC)
    merged_data <- merge(NEI, SCC, by="SCC")
    df <- data.table(merged_data)

    #filter data to find coal in Short.Name
    coal <- grepl("coal", df$Short.Name, ignore.case=TRUE)
    coal <- data.table(merged_data[coal, ])

    by_year <- coal[, list(emissions=sum(Emissions)), by=c('year')]
    by_year$year = as.numeric(as.character(by_year$year))
    by_year$emissions = as.numeric(as.character(by_year$emissions))

    #plot on screen
    ggplot(data=by_year, aes(x=year, y=emissions)) + geom_line() + geom_point() +
    ggtitle("Emissions from Coal Sources in the US")
    #plot png plot4_EG.png
    dev.copy(png, file="plot4_EG.png", width=480, height=480)
    dev.off()



    #plot5
    NEI <- readRDS("summarySCC_PM25.rds")
    head(NEI)
    str(NEI)
    df <- data.table(NEI)

    #filering data: only Baltimore with type ‘ON-ROAD’
    Baltimore <- subset(df, fips == '24510' & type == 'ON-ROAD')
    by_year <- Baltimore[, list(emissions=sum(Emissions)), by=c('year', 'type')]
    by_year$year = as.numeric(as.character(by_year$year))
    by_year$emissions = as.numeric(as.character(by_year$emissions))

    #plot on screen
```

```
ggplot(data=by_year, aes(x=year, y=emissions)) + geom_line() + geom_point() +
ggtitle("Emissions in Baltimore City from Motor Vehicles")

#plot png plot5_EG
dev.copy(png, file="plot5_EG.png", width=480, height=480)
dev.off()


#plot6
NEI <- readRDS("summarySCC_PM25.rds")
head(NEI)
str(NEI)
df <- data.table(NEI)

#filter Baltimore and Los Angeles with type ‘ON-ROAD’
data <- subset(df, fips %in% c('06037', '24510') & type == 'ON-ROAD')
by_year <- data[, list(emissions=sum(Emissions)), by=c('year', 'fips')]
by_year$year = as.numeric(as.character(by_year$year))
by_year$emissions = as.numeric(as.character(by_year$emissions))

#plot on screen to check
ggplot(data=by_year, aes(x=year, y=emissions, col=fips)) + geom_line() + geom_point()
+ ggtitle("Emissions from Motor Vehicles in Baltimore City compared to Los Angeles")

#plot png plot6_EG
dev.copy(png, file="plot6_EG.png", width=480, height=480)
dev.off()
```