**Exploratory Data Analysis Project 2:**

# **Import the data frame**

library(dplyr)

setwd("C:/Users/KHAWLA/Downloads/")

file.name<- "summarySCC\_PM25.rds"

data<- readRDS(file.name)

# **The code in R:**

## Question 1:

emission\_per\_year <- data %>% group\_by(year) %>% summarise(total=sum(Emissions))

emission\_per\_year

## Question 2:

emission\_balt <- data %>% group\_by(year) %>% filter(fips == "24510") %>% summarise(total=sum(Emissions))

emission\_balt

## Question 3:

emission\_balt\_type <- data %>% group\_by(type,year) %>% filter(fips == "24510") %>% summarise(total=sum(Emissions))

emission\_balt\_type

## Question 4:

file2 <- "Source\_Classification\_Code.rds"

SCC <- readRDS(file2)

# select coal combustion related sources from Ei.Sector:

coal\_comb <- SCC[grepl("Comb.\*Coal", SCC$EI.Sector),]

# total coal combustion emissions

coal\_scc <- unique(coal\_comb$SCC)

coal\_emission <- data[(data$SCC %in% coal\_scc),]

coal\_year <- coal\_emission %>% group\_by(year) %>% summarise(total=sum(Emissions))

## Question 5:

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

## # records related to motor vehicle sources:

## motor\_data <- SCC[grepl("Vehicle", SCC$SCC.Level.Two),]

## # total emissions for motor vehicle sources in Baltimore city:

## motor\_scc <- unique(motor\_data$SCC)

## motor\_emission <- data[(data$SCC %in% motor\_scc),]

## motor\_year <- motor\_emission %>% filter(fips == "24510") %>%

## group\_by(year) %>% summarise(total=sum(Emissions))

## Question 6:

# Compare the motor vehicle emissions in Baltimore City, LA:

motor\_balti\_la <- motor\_emission %>% filter(fips == "24510" | fips == "06037") %>%

group\_by(fips,year) %>% summarise(total=sum(Emissions))

# add the name of city for each code:

motor\_balti\_la <- mutate(motor\_balti\_la, Unit=

ifelse(fips == "24510","Baltimore city",

ifelse(fips== "06037", "Los Angeles County")))

# **Graphs:**

## Plot 1:

png(filename = "plot1.png")

plot1 <- barplot(emission\_year$total/1000, xlab = "Year", ylab = " PM 2.5 Emissions in kilotons",

main = "Total PM 2.5 Emissions", names.arg = emission\_year$year,

col = "darkgreen", ylim = c(0,8500))

text(plot1, round(emission\_year$total/1000), labels = round(emission\_year$total/1000),

pos = 3, cex = 1.2)

dev.off()

## Plot 2:

## png(filename = "plot2.png")

## plot2 <- barplot(emission\_balt$total, xlab = "Year", ylab = " PM 2.5 Emissions",

## main = "Total PM 2.5 Emissions in Baltimore City", names.arg = emission\_balt$year,

## col = "blue")

## text(plot2, round(emission\_balt$total), labels = round(emission\_balt$total),pos = 3, cex = 1.2)

## dev.off()

## Plot 3:

## library(ggplot2)

## png(filename = "plot3.png")

## plot3 <- ggplot(emission\_balt\_type, aes(x= factor(year), y= total, fill = type, label= round(total)))+

## geom\_bar(stat="identity")+ facet\_grid(.~type)+

## ggtitle("Total of PM 2.5 Emissions in Baltimore City, Maryland")+ xlab("Year")+

## ylab("PM 2.5 Emissions")+ theme\_classic()

## dev.off()

## Plot 4:

## png(filename = "plot4.png")

## ggplot(coal\_year, aes(x= factor(year), y=total/1000, label=round(total/1000)))+

## geom\_bar(stat = "identity", fill= "darkred")+ ggtitle("The Total of Coal Combustion related to PM 2.5 Emissions")+

## xlab("Year")+ ylab("PM 2.5 Emissions in kiloTons")+ theme\_classic()+

## ylim(c(0,640))+ geom\_text(size = 5, vjust = -1)

## dev.off()

## Plot 5:

## png(filename = "plot5.png")

## ggplot(motor\_year, aes(x= factor(year), y=total, label=round(total)))+

## geom\_bar(stat = "identity", fill= "darkorchid4")+ ggtitle("The Total of Motor Vehicle Emissions in Baltimore city, Maryland")+

## xlab("Year")+ ylab("PM 2.5 Emissions")+ theme\_classic()+ ylim(c(0,500))+

## geom\_text(size = 5, vjust = -1)

## dev.off()

## Plot 6:

png(filename = "plot6.png")

ggplot(motor\_balti\_la, aes(x = factor(year), y = total,

fill= Unit, label=round(total)))+

geom\_bar(stat = "identity")+facet\_grid(.~Unit)+

ggtitle("Total of Motor Vehicle Emissions")+ xlab("year")+ ylab("PM 2.5 Emissions")+

theme\_classic()+ylim(c(0,8000))+ geom\_text(size = 3, vjust=-1)

dev.off()