

# CS4048 Data Science

Friday, November 26, 2021

## Course Instructor

Mr. Adeel Ashraf Cheema

Mr. Muhammad Usman Joyia

Serial No:

**2<sup>nd</sup> Mid Term Exam**

**Total Time: 2 Hour**

**Total Marks: 60**

\_\_\_\_\_  
Signature of Invigilator

**F178002**

**7C**

**Roll No**

**Section**

**Signature**

**DO NOT OPEN THE QUESTION BOOK OR START UNTIL INSTRUCTED.**

### Instructions:

1. Read the questions carefully for clarity of context and understanding of meaning and make assumptions wherever required, for neither the invigilator will address your queries, nor the teacher/examiner will come to the examination hall for any assistance.
2. Fit in all your answers in the provided space. You may use extra space on the last page if required. If you do so, clearly mark question/part number on that page to avoid confusion.
3. Use only your own stationery and calculator. If you do not have your own calculator, use manual calculations.
4. Use only permanent ink-pens. Only the questions attempted with permanent ink-pens will be considered. Any part of paper done in lead pencil cannot be claimed for checking/rechecking.

	Q-1	Q-2	Q-3	Q-4	Q-5	Q-6	Total
Total Marks	10	10	10	10	10	10	<b>60</b>
Marks Obtained							

Vetted By: \_\_\_\_\_ Vetter Signature: \_\_\_\_\_

University Answer Sheet Required: No ☒ Yes ☐

## Question 1:

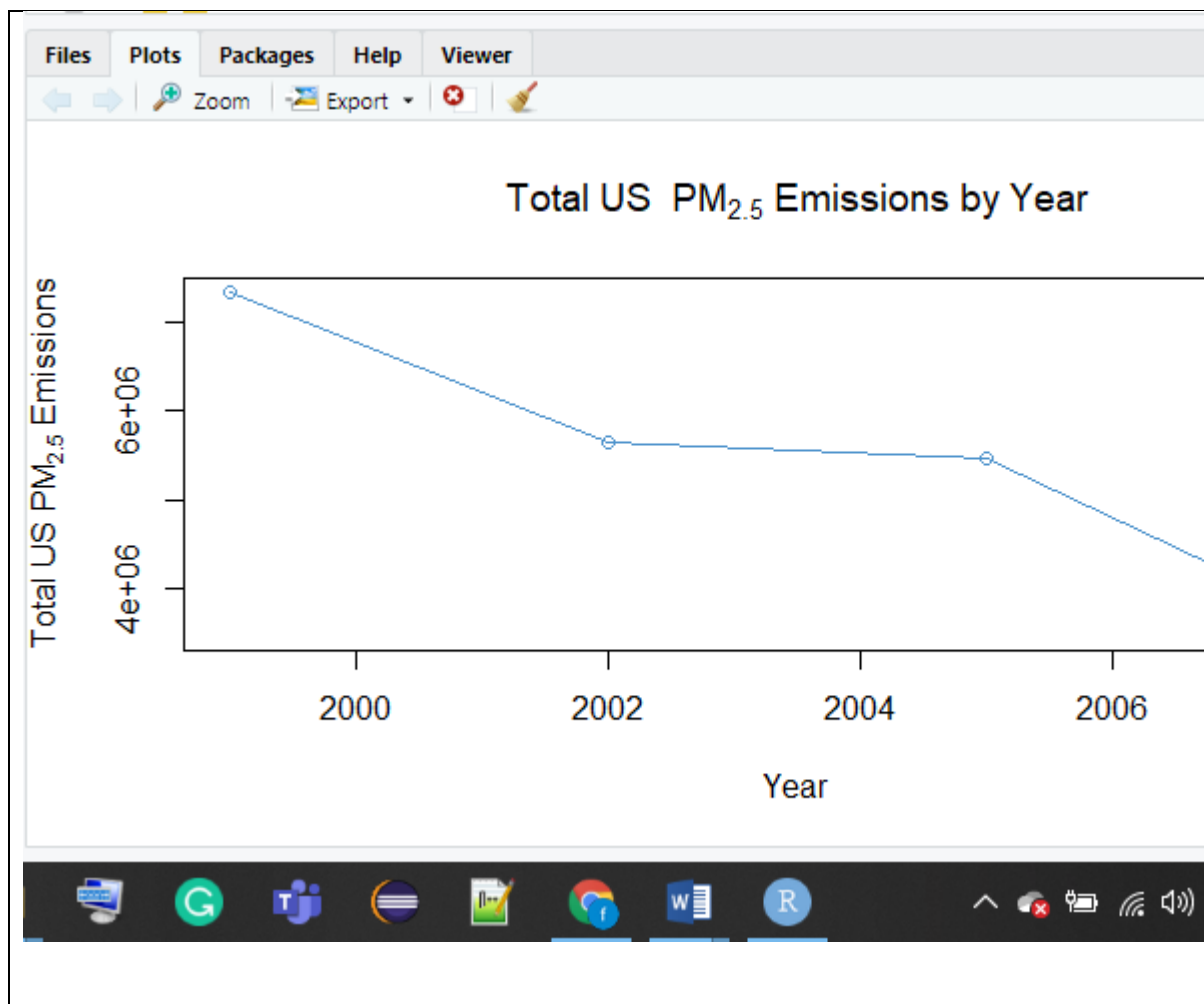
Code:

### Part 1:

```
totaldat <- aggregate(Emissions ~ year, dat, sum)

plot(totaldat$year, totaldat$Emissions, type = "o", col =
"steelblue3", main = expression("Total US "~ PM[2.5]~
"Emissions by Year"), ylab = expression("Total US "~ PM[2.5] ~
"Emissions"), xlab = "Year")
```

Output plots



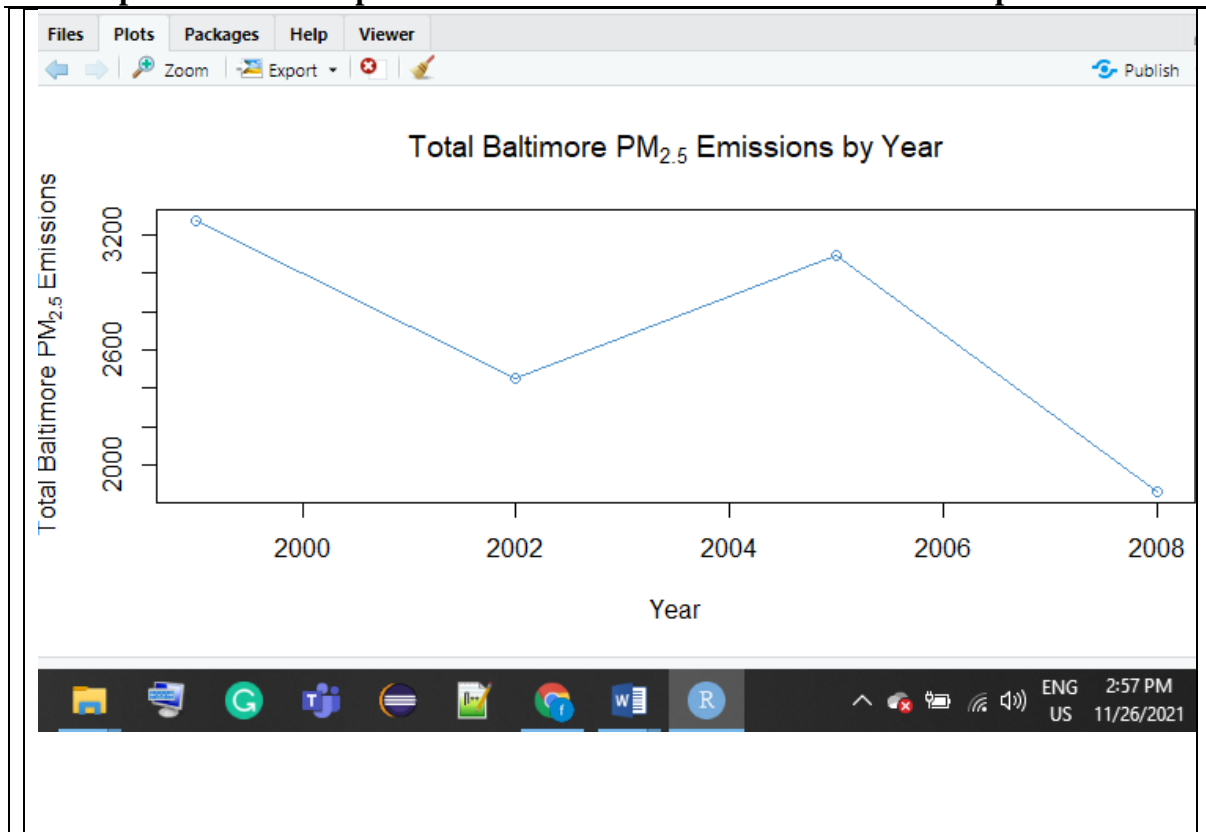
Code:

**Part 2:**

```
baltimore <- subset(dat, dat$fips == "24510")

totalBaltimore <- aggregate(Emissions ~ year, baltimore, sum)

plot(totalBaltimore$year, totalBaltimore$Emissions, type = "o",
main = expression("Total Baltimore" ~ PM[2.5] ~ "Emissions by
Year"), xlab = "Year", ylab = expression("Total Baltimore "~
PM[2.5] ~ "Emissions"), col = "steelblue3")
```

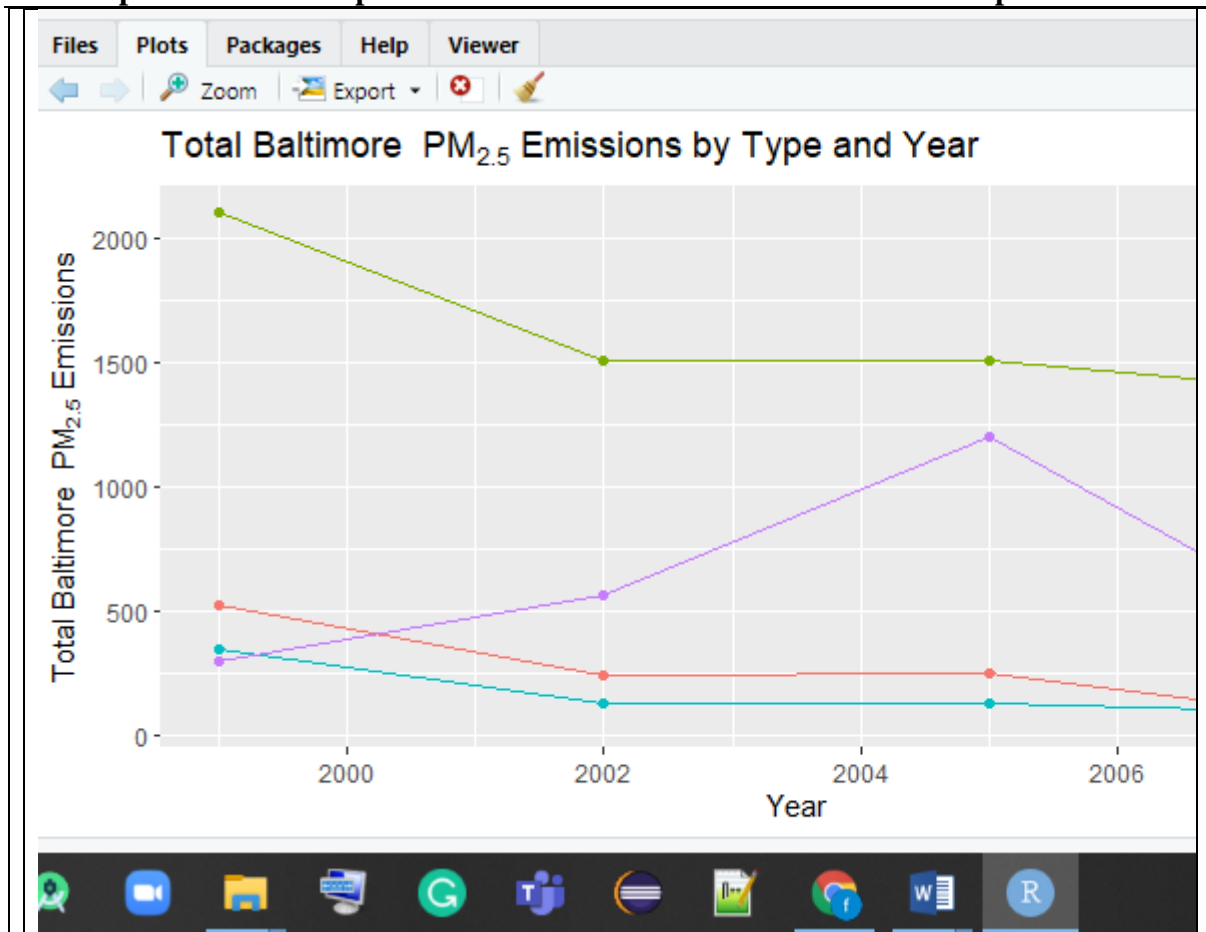


Code:

## Part 3:

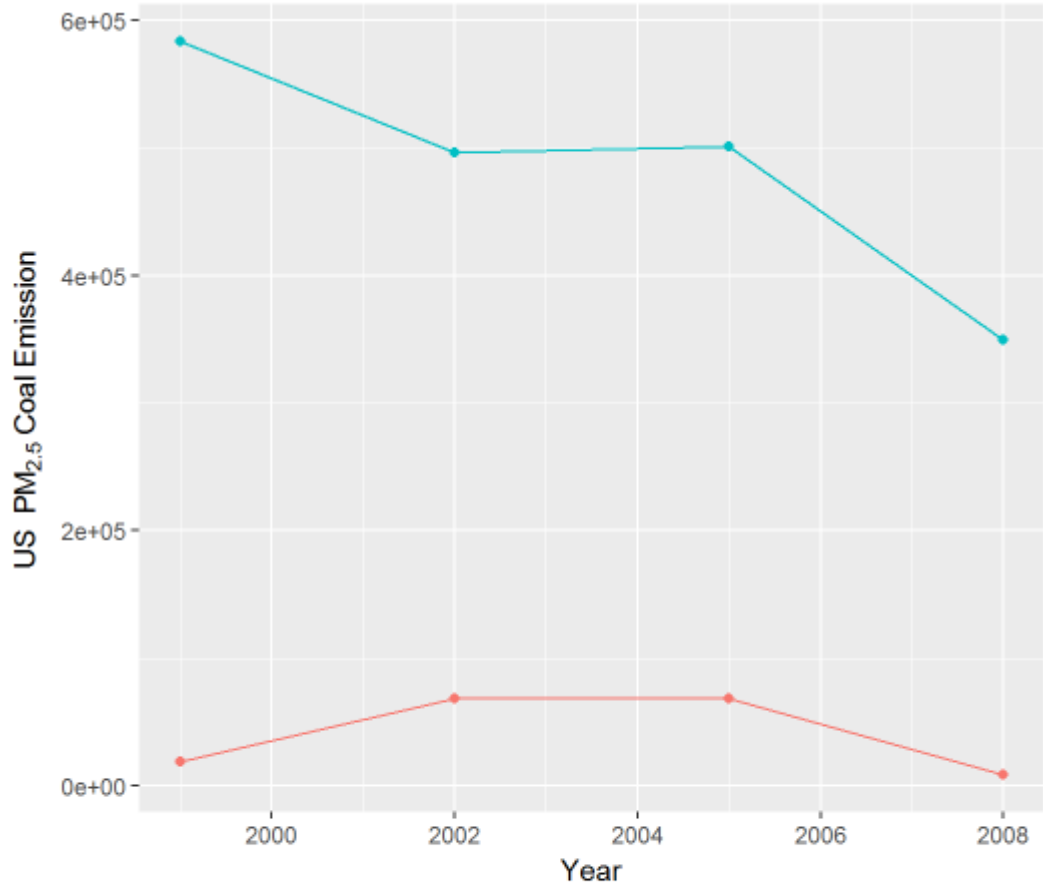
```
baltimore <- subset(dat, dat$fips == "24510")
baltimoreType <- aggregate(Emissions ~ year + type, baltimore,
sum)

ggplot(baltimoreType, aes(year, Emissions, col = type)) +
  geom_line() +
  geom_point() +
  ggtitle(expression("Total Baltimore " ~ PM[2.5] ~ "Emissions by
Type and Year")) +
  ylab(expression("Total Baltimore " ~ PM[2.5] ~ "Emissions")) +
  xlab("Year") +
  scale_colour_discrete(name = "Type of sources") +
  theme(legend.title = element_text(face = "bold"))
```



## Part 4:

```
datacoal <- data[grepl("coal", data$Short.Name, ignore.case =  
T),]  
datcoal <- dat[dat$data %in% datacoal$data,]  
totalCoal <- aggregate(Emissions ~ year + type, datcoal, sum)  
  
ggplot(totalCoal, aes(year, Emissions, col = type)) +  
  geom_line() +  
  geom_point() +  
  ggtitle(expression("Total US" ~ PM[2.5] ~ "Coal Emission by  
Type and Year")) +  
  xlab("Year") +  
  ylab(expression("US " ~ PM[2.5] ~ "Coal Emission")) +  
  scale_colour_discrete(name = "Type of sources") +  
  theme(legend.title = element_text(face = "bold"))
```

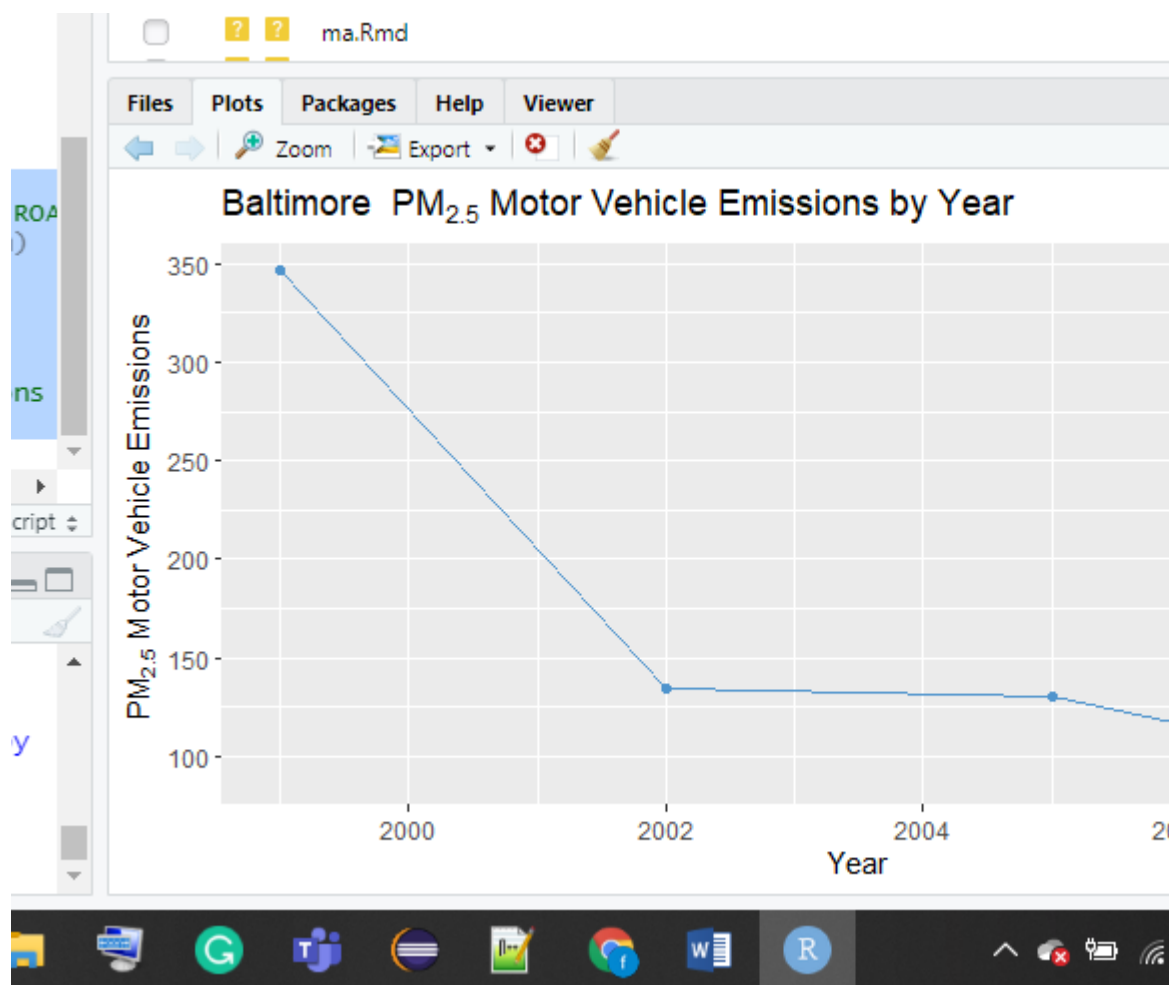


## Part 5:

```
baltimoreMotor <- subset(dat, dat$fips == "24510" & dat$type == "ON-ROAD")
```

```
baltimoreMotorAGG <- aggregate(Emissions ~ year,  
baltimoreMotor, sum)
```

```
ggplot(baltimoreMotorAGG, aes(year, Emissions)) +  
  geom_line(col = "steelblue3") +  
  geom_point(col = "steelblue3") +  
  ggtitle(expression("Baltimore " ~ PM[2.5] ~ "Motor Vehicle  
Emissions by Year")) +  
  xlab("Year") +  
  ylab(expression(~PM[2.5]~ "Motor Vehicle Emissions"))
```





**Part 6:**

```
baltLosAngelesMotors <- subset(dat, dat$fips %in%  
c("24510","06037") & dat$type == "ON-ROAD")  
baltLosAngelesMotorsAGG <- aggregate(Emissions ~ year + fips,  
baltLosAngelesMotors, sum)  
  
ggplot(baltLosAngelesMotorsAGG, aes(year, Emissions, col =  
fips)) +  
  geom_line() +  
  geom_point() +  
  ggtitle(expression("Baltimore and Los Angeles" ~ PM[2.5] ~  
"Motor Vehicle Emissions by Year")) +
```

```
labs(x = "Year", y = expression(~PM[2.5]~ "Motor Vehicle  
Emissions")) +  
  scale_colour_discrete(name = "City", labels = c("Los Angeles",  
"Baltimore")) +  
  theme(legend.title = element_text(face = "bold"))
```

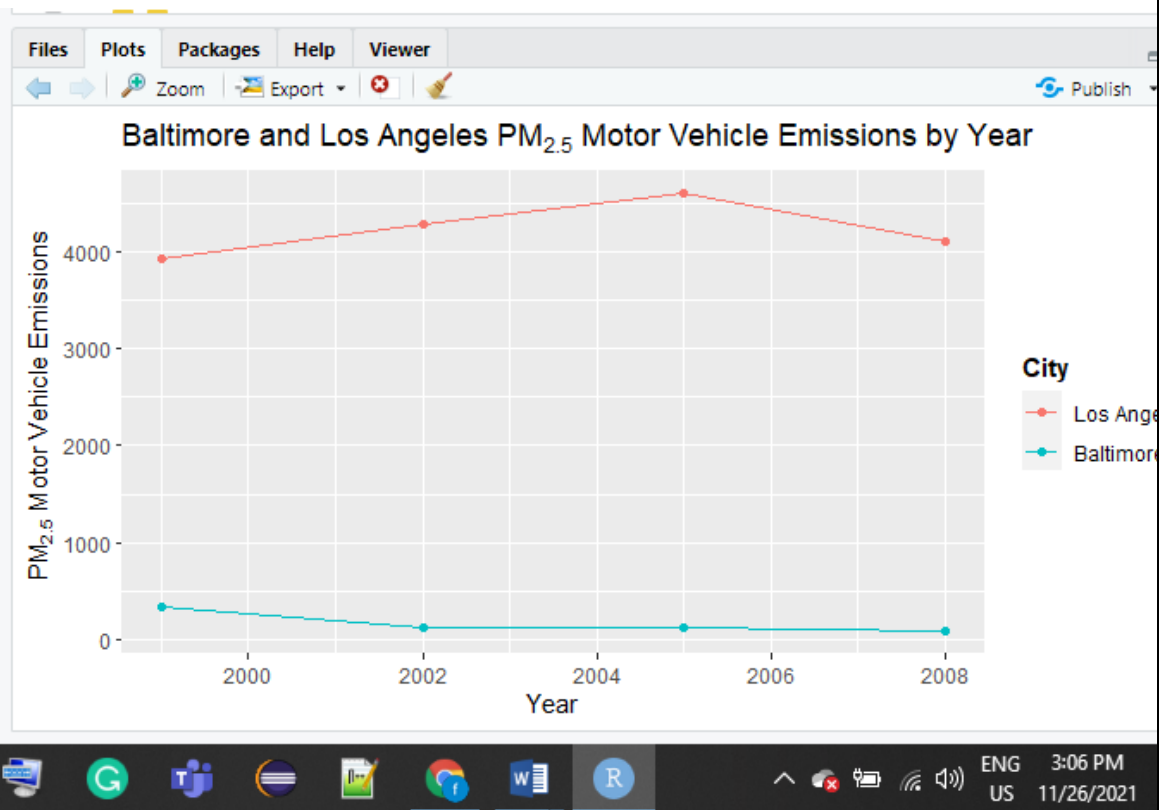


Figure 1Caption Me