|  |
| --- |
| Serial No: |
| **2nd Mid Term Exam** |
| **Total Time:2 Hour** |
| **Total Marks: 60** |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Signature of Invigilator |

|  |
| --- |
| **CS4048 Data Science** |
| Friday, November 26, 2021 |
| **Course Instructor** |
| Mr. Adeel Ashraf Cheema  Mr. Muhammad Usman Joyia |

|  |
| --- |
| **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ . .**  **Roll No Section Signature** |

Name: Usama Khalid

Roll no: F17-8013

Sec: 7A

## 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 | **60** |
| **Marks Obtained** |  |  |  |  |  |  |  |

**Vetted By: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Vetter Signature: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **University Answer Sheet Required:** | **No** |  |  |  |  | **Yes** |  |

**Question 1:**

Code:

|  |
| --- |
| NEI <- readRDS("summarySCC\_PM25.rds")  SCC <- readRDS("Source\_Classification\_Code.rds")  totalNEI <- aggregate(Emissions ~ year, NEI, sum)  plot(totalNEI$year, totalNEI$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

|  |
| --- |
| Chart, line chart  Description automatically generated |

**Figure 1Caption Me**

**Question 2:**

Code:

|  |
| --- |
| baltimoredata <- subset(NEI, NEI$fips == "24510")  totalBaltimoredata <- aggregate(Emissions ~ year, baltimoredata, sum)  plot(totalBaltimoredata$year, totalBaltimoredata$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") |

Output plots

|  |
| --- |
| **Chart, line chart  Description automatically generated**  **Chart, line chart  Description automatically generated** |

**Question 3:**

Code:

|  |
| --- |
| baltimoredata <- subset(NEI, NEI$fips == "24510")  baltimoreType <- aggregate(Emissions ~ year + type, baltimoredata, sum)  install.packages("ggplot2")  library(ggplot2)  ggplot(baltimoreType, aes(year, Emissions, col = type)) +  geom\_line() +  geom\_point() +  ggtitle(expression("Total Baltimore info" ~ PM[2.5] ~ "Emission Type and Year")) +  ylab(expression("Total Baltimore info " ~ PM[2.5] ~ "Emissions")) +  xlab("Year") +  scale\_colour\_discrete(name = "Type of sources") +  theme(legend.title = element\_text(face = "bold")) |

Output plots

|  |
| --- |
| **Chart, line chart  Description automatically generated** |

**Question 4:**

**Code**

|  |
| --- |
| SCCcoal <- SCC[grepl("coal", SCC$Short.Name, ignore.case = T),]  NEIcoal <- NEI[NEI$SCC %in% SCCcoal$SCC,]  totalCoal <- aggregate(Emissions ~ year + type, NEIcoal, 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 info")) +  scale\_colour\_discrete(name = "Type of sources") +  theme(legend.title = element\_text(face = "bold")) |

Output plots

|  |
| --- |
| Chart, line chart  Description automatically generated |

**Question 5:**

**Code**

|  |
| --- |
| baltimoreMotordata <- subset(NEI, NEI$fips == "24510" & NEI$type == "ON-ROAD")  baltimoreMotorAGGdata <- aggregate(Emissions ~ year, baltimoreMotordata, sum)  ggplot(baltimoreMotorAGGdata, aes(year, Emissions)) +  geom\_line(col = "steelblue3") +  geom\_point(col = "steelblue3") +  ggtitle(expression("Baltimore info " ~ PM[2.5] ~ "Motor Vehicle Emissions by Year info")) +  xlab("Year") +  ylab(expression(~PM[2.5]~ "Motor Vehicle Emissions")) |

Output plots

|  |
| --- |
| Chart, line chart  Description automatically generated |

**Question 6:**

**Code**

|  |
| --- |
| baltLosAngelesMotorsdata <- subset(NEI, NEI$fips %in% c("24510","06037") & NEI$type == "ON-ROAD")  baltLosAngelesMotorsAGGdata <- aggregate(Emissions ~ year + fips, baltLosAngelesMotorsdata, sum)  ggplot(baltLosAngelesMotorsAGGdata, aes(year, Emissions, col = fips)) +  geom\_line() +  geom\_point() +  ggtitle(expression("Baltimore and Los Angeles info" ~ PM[2.5] ~ "Motor Vehicle Emissions by Year info")) +  labs(x = "Year", y = expression(~PM[2.5]~ "Motor Vehicle Emissions info") ) +  scale\_colour\_discrete(name = "City", labels = c("Los Angeles", "Baltimore")) +  theme(legend.title = element\_text(face = "bold")) |

Output plots

|  |
| --- |
| Chart, line chart  Description automatically generated |