```
#Question 1

library(data.table)
library(arules)
library(arulesViz)
library(lubridate)
library(ggplot2)
library(knitr)
library(plyr)
library(readxl)
library(tidyverse)
library(RColorBrewer)

#a : import the coronavirus package

getwd()
setwd("/Users/Owner/Documents/Business Analytics")
install.packages("coronavirus")
library(coronavirus)

#b : show the first 100 elements of the coronavirus package

head(coronavirus, 100)

#c : describe the meaning of each column

# date = when the coronavirus data was extracted / when the cases occurred
# province = the province in a particular country that data is being extracted (similar to a state)
# country = the country where the data is being extracted / where the cases occurred
# lat = latitude
# long = longitude
# type = refers to condition of the cases (confirmed cases, cases resulting in death, and recovered cases)
# cases = the number of cases that occurred in a specific location at a specific time
```

```
#Question 2
#a : summary of total confirmed cases by country

library(dplyr)
summary_coronavirus = coronavirus %>%
  filter(type == "confirmed") %>%
  group_by(country) %>%
  summarise(total_cases = sum(cases)) %>%
  arrange(-total_cases)

head(summary_coronavirus, 20)

#b : top 5 countries in bar graph

top_countries = data.frame(head(summary_coronavirus, 5))
top_5 = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity")

#c : flip the bar graph so it is a horizontal barplot

top_5_horiz = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity") + coord_flip()

#d : add a title

print(top_5 + ggtitle("Top 5 Countries by Total Cases"))
print(top_5_horiz + ggtitle("Top 5 Countries by Total Cases"))



#Question 3
#a : create data frame that represents the confirmed # of cases sorted by dates

library(tidyr)
```

```
recent_cases = coronavirus %>%
  filter(type == "confirmed") %>%
  group_by(date) %>%
  summarise(confirmed_cases = sum(cases)) %>%
  arrange(date)

recent_cases = as.data.frame(recent_cases)

#b : show the recent_cases data in a line graph

recent_cases_graph = ggplot(recent_cases, aes(x = date, y = confirmed_cases)) + geom_line()



#Extra Credit
#1 : Change line color to red on recent_cases_graph

recent_cases_graph = ggplot(recent_cases, aes(x = date, y = confirmed_cases)) + geom_line(color = "red")

#2 : Change line type of recent_cases_graph

recent_cases_graph = ggplot(recent_cases, aes(x = date, y = confirmed_cases)) + geom_line(color = "red", linetype = 2)

#3 : Change font of recent_cases_graph

recent_cases_graph = recent_cases_graph + theme(text = element_text(size = 13, family = "Comic Sans MS"))

#4 : Change font color of recent_cases_graph to green

recent_cases_graph = recent_cases_graph + theme(text = element_text(size = 13, family = "Comic Sans MS", color =
"green"))

#5 : Add a title to recent_cases_graph
```

```
recent_cases_graph = print(recent_cases_graph + ggtitle("Confirmed Cases Over Time"))


#6 : Change bar colors of top_5 to blue


top_5 = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity", fill = "blue")


#7 : Change bar widths of top_5


top_5 = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity", fill = "blue", width = 0.5)


#8 : Change transparency of bars of top_5


top_5 = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity", fill = "blue", width = 0.5,
alpha = 0.25)


#9 : Add trim to bars of top_5


top_5 = ggplot(top_countries, aes(x = country, y = total_cases)) + geom_bar(stat = "identity", color = "black", fill =
"blue",
                                                                            width = 0.5, alpha = 0.25)
#10 : Add data labels to top_5


top_5 = top_5 + geom_text(aes(label = total_cases))
```