1. From R:
    1. All code from your R script (Code should be presented single-spaced in a
       <u>fixed-width font</u>. Adjust the font size so that no lines of code extend to the
       next line in the document)

```
# YUEH-TING WU
# MIS 545 Section 02
# Lab10WuY.R
# In this R programming, using read.transactions() function to
read a csv file
# into an object, examine the frequencies, and generate the
association rules.

# Install tidyverse and arules packages
# install.packages("tidyverse")
# install.packages("arules")

# Load the tidyverse and arules libraries
library(tidyverse)
library(arules)

# Set the working directory
setwd("~/MIS 545/Lab10")

# Read InstacartTransactions.csv into an object called
instacartTransactions
instacartTransactions <- read.transactions(file =
"InstacartTransactions.csv",
                                            format = "single",
                                            header = TRUE,
                                            sep = ",",
                                            cols = c("OrderID",
"ItemID"))

# Display a summary of instacartTransactions
summary(instacartTransactions)

# Display the first free transactions from instacartTransactions
inspect(instacartTransactions[1:3])

# Examine the frequency of a single item: 24852 (bananas)
itemFrequency(instacartTransactions[, "24852"])

# Convert the frequency values in instacartTransactions into a
tibble called
# instacartTransactionsFrequency
instacartTransactionsFrequency <-
```

```
  tibble(items = names(itemFrequency(instacartTransactions)),
                        Frequency =
itemFrequency(instacartTransactions))

# Display the item frequencies in instacartTransactionsFrequency
on the console
print(instacartTransactionsFrequency)

# Display the 10 most frequently purchased items on the console
instacartTransactionsFrequency %>%
  arrange(desc(Frequency)) %>%
  slice(1:10)

# Generate the association rules model in an object called
# instacartTransactionRules
instacartTransactionsRules <-
  apriori(instacartTransactions,
          parameter = list(
            support = 0.005,
            confidence = 0.2,
            minlen = 2))

# Display a summary of instacartTransactionsRules
summary(instacartTransactionsRules)

# Display the first 10 association rules
inspect(instacartTransactionsRules[1:10])

# Sort the association rules by lift and view the top 10
instacartTransactionsRules %>%
  sort(by = "lift") %>%
  head(n = 10) %>%
  inspect()
```

2. What are the 5 most frequently purchased items? Ensure you look up the names of the products so you can refer to them by name instead of by itemID.

The 5 most frequently purchased items are Bag of Organic Bananas, Organic Strawberries, Organic Baby Spinach, and Large Lemon.

3. Answer the following question in a few sentence: Interpret the association rule with the highest lift number. Do so assuming your reader has no knowledge of association rule methodology. Ensure you look up the names of the products so you can refer to them by name instead of by itemID.

The lift number in an association rule means that the increased or decreased likelihood of an itemset being sold together compared to an item in the dataset being sold along. In this association rule, the highest lift number is in Organic Garlic and Organic Yellow

Onion. Thus, it demonstrates that these two items are more likely to be sold together better than being sold alone.

4. Answer the following question in a sentence: Interpret the association rule with the 5th highest lift number. Do so assuming your reader has no knowledge of association rule methodology. How can Instacart take advantage of this information to maximize their revenues? Ensure you look up the names of the products so you can refer to them by name instead of by itemID.

The 5th highest lift number is in { Bag of Organic Bananas, Organic Strawberries, Organic Hass Avocado}. Instacart can provide some recipes near these three items, such as avocado toast, banana peanut butter toast, or fresh fruit salad. Customers who go to buy these products may take a look at the recipes and tend to buy more products which they like on the recipes.

5. Interpret the support, confidence, and lift values of the association rule with the 10th highest lift. Ensure you look up the names of the products so you can refer to them by name instead of by itemID.

The support value is 0.012, which means that 1.2% of all transactions have Organic Raspberries and Organic Strawberries in them sold together.
The confidence value is 0.3, which means that 30% of the time when Organic Raspberries are sold, Organic Strawberries are also sold in the same transaction.
The lift value is 3.63, which means that Organic Strawberries is 3.63 times more likely to be sold if we also have Organic Raspberries in the transactions.