MATH 216: Assignment 5

Your Name Here

Instructions

- submit the .html file to Canvas
- you are encouraged to work together and ask your peers questions. Each person should submit their own work.
- You may share parts of your code to ask or answer questions on Slack. You should avoid sharing (copying and pasting) the entirety of your answers.
- make sure you include at least one acknowledgement
- The assignment is worth 15 points total. The specific number of points for each part is listed below.

Loading the data

You have decided to open an online store selling things to keep people busy and creative.

```
products <-read.csv("https://ebmwhite.github.io/MATH0216/assignments/products.csv")
orders <-read.csv("https://ebmwhite.github.io/MATH0216/assignments/orders.csv")
customers <-read.csv("https://ebmwhite.github.io/MATH0216/assignments/customers.csv")
```

A list of the products you sell is available in `products`. A list of your current orders is available in `orders`. A list of your customers is available in `customers`.

Please click on each dataset to get familar with what each row represents and what columns are available.

Question 1: Print out a list of all the products you sell which cost under $10 (1 point)

```
products %>%
  filter(price <= 10)
```

```
##   product_id    description price
## 1          2  coloring-book     5
## 2          3 pencil-crayons     8
## 3          4        markers     7
```

Question 2: How many puzzles have you sold? (1 point)

```
left_join(orders, products, by="product_id") %>%
  filter(description == "puzzle") %>%
  summarize(revenue = sum(quantity))
```

```
##   revenue
## 1       1
```

Some students may have used slightly different logic. As long as they get the answer 1, give them the point.

Question 3: Consider the order with `order_id=4`. What type of product did they order? (1 point)

Note: you can answer this by manually looking at the tables, or you can use the appropriate join function.

```
left_join(orders, products, by="product_id") %>%
  filter(order_id == 4) %>%
  select(description)
```

```
##   description
## 1     markers
```

Some students may have used slightly different logic. As long as they get the answer "markers", give them the point.

Question 4: Consider the order with `order_id=7`. What is the email address associated with this order? (1 point)

Note: you can answer this by manually looking at the tables, or you can use the appropriate join function.

```
left_join(orders, customers, by="customer_id") %>%
  filter(order_id == 7) %>%
  select(email)
```

```
##                     email
## 1 jasmine077@boxsquat.com
```

Some students may have used slightly different logic. As long as they get the answer "jasmine007@boxsquat.com", give them the point.

Question 5 (6 points, 2 points for each part)

(a) Return a dataframe which has all the columns and rows from `orders` but also contains an additional column with the `price` of the item listed in the `products` dataset. For example, if `product_id=3`, the column `price` should contain 8.

```
left_join(orders, products, by="product_id")
```

```
##    order_id customer_id product_id quantity time_stamp    description
## 1         1           1          3        2    3/11/20 pencil-crayons
## 2         2           2          2        1    3/11/20  coloring-book
## 3         3           2          3        1    3/11/20 pencil-crayons
## 4         4           3          4        2    3/12/20        markers
## 5         5           4          1        1    3/12/20         puzzle
## 6         6           4          5        1    3/12/20     board-game
## 7         7           5          4        5    3/12/20        markers
## 8         8           1          2        1    3/13/20  coloring-book
## 9         9           6          3        1    3/13/20 pencil-crayons
## 10       10           7          5        2    3/13/20     board-game
##    price
## 1      8
## 2      5
## 3      8
## 4      7
## 5     20
## 6     25
## 7      7
```

```
## 8      5
## 9      8
## 10    25
```

(b) In the `orders` dataset, add a new column called `total` which indicates the total cost of the order to the customer. You may want to use your answer from (a) to help you.

```
left_join(orders, products, by="product_id")  %>%
  mutate(total = price*quantity)
```

```
##    order_id customer_id product_id quantity time_stamp    description
## 1         1           1          3        2    3/11/20 pencil-crayons
## 2         2           2          2        1    3/11/20  coloring-book
## 3         3           2          3        1    3/11/20 pencil-crayons
## 4         4           3          4        2    3/12/20        markers
## 5         5           4          1        1    3/12/20         puzzle
## 6         6           4          5        1    3/12/20     board-game
## 7         7           5          4        5    3/12/20        markers
## 8         8           1          2        1    3/13/20  coloring-book
## 9         9           6          3        1    3/13/20 pencil-crayons
## 10       10           7          5        2    3/13/20     board-game
##    price total
## 1      8    16
## 2      5     5
## 3      8     8
## 4      7    14
## 5     20    20
## 6     25    25
## 7      7    35
## 8      5     5
## 9      8     8
## 10    25    50
```

(c) What is the total amount of revenue you have brought in so far? You may want to use your answer from (b) to help you.

```
left_join(orders, products, by="product_id")  %>%
  mutate(total = price*quantity) %>%
  summarize(revenue = sum(total))
```

```
##   revenue
## 1     186
```

Question 6: List the names and addresses (including `address`, `city`, and `state`) of the people who placed orders on March 13, 2020. (2 points).

```
left_join(orders, customers, by="customer_id") %>%
  filter(time_stamp =="3/13/20") %>%
  select(customer_name, address, City, State)
```

```
##   customer_name                 address         City State
## 1   Boris Bower 1328  Village View Drive    Frederick    MD
```

```
## 2  Brian McLean     914 Twin Willow Lane Fayetteville    NC
## 3 DeSean Mackay        885 Cecil Street     Wheeling     IL
```

Question 7: Find a dataset on Wikipedia that you think is interesting. Scrape the data from the web. Please include a link to the wikipedia pages and print out the first few rows of the dataset using the `head` command.

You can choose any dataset you like. https://en.wikipedia.org/wiki/Wikipedia:Featured_lists might be a good place to start looking.

Deduct points as you see appropriate... - 1 point deduction (2/3 marks) for not having the headers formatted appropriately - 1 point deduction (2/3 marks) for not listing the source of their dataset somewhere