

Project 03 Answer Key

Question 1a (3 pts)

Make a table of the number of transactions that occur in each of the four stores in the 84.51 data set called 5000_transactions.csv, found in the folder The_Complete_Journey_2_Master.

Solution

```
# Read in the data with the read.csv() function and use the table() function on
# the STORE_R column to get a frequency table of each store.
transactions = read.csv('class/datamine/data/8451/
  The_Complete_Journey_2_Master/5000_transactions.csv')
table(transactions$STORE_R)

>>>
CENTRAL EAST    SOUTH    WEST
2463343 3263360 2221500 2677350
```

Question 2a (1 pt)

Load the data about yellow taxi cab rides in New York City in June 2019. This data is already stored on Scholar.

Solution

```
# Read in the data with the read.csv() function
taxi_data = read.csv('/class/datamine/data/taxi/yellow/yellow_tripdata_2019-06.csv')
```

Question 2b (2 pts)

How many passengers (altogether) rode in yellow taxi cab rides in New York City in June 2019?

Solution

```
# Use the sum() function to get a total of all passenger counts
sum(taxi_data$passenger_count)

>>>
10878820
```

Question 3 (4 pts)

Modify the code from Project 2 to visualize the locations of all airports in the (continental) USA.

Solution

```
# Load the ggmap library
library(ggmap)

# Download and read in the airports data with read.csv()
airports = read.csv('http://stat-computing.org/dataexpo/2009/airports.csv')

# Put the latitude and longitude data into a data frame using data.frame()
usa_points = data.frame(lon=airports$long, lat=airports$lat)

# Register your API key
register_google(key='<your key here>', write=TRUE)

# Select Lebanon, Kansas as the map center. This center was determined by the
# National Geodetic Survey. Selecting 'Kansas' in general should yield pretty
# good results.
usa_center = as.numeric(geocode('Lebanon, Kansas'))

# Generate and display the map
usa_map = ggmap(get_googlemap(center=usa_center, zoom=4))
usa_map

# Add points to the map and display it
usa_map_with_points = map + geom_point(data=usa_points, size=0.07)
usa_map_with_points
```