**Assessment:**

**Data Pipeline & Data Ingestion**

| **Section B (70 points total)** |
| --- |

**Context: Explore US Bikeshare Data**

You are assigned the task to analyze the Bicycle-sharing data for three major cities in the United States—Chicago, New York City, and Washington.

The [datasets](https://drive.google.com/drive/folders/1L9ijKFUJ36ORSiCLYgM22iDbmDeT4Sew?usp=sharing) used for this project contain bike share data for the first six months of 2017. The csv files can be found in the same folder. The data is provided by Motivate, which is a bike share system provider for many cities in the United States. The data files for all three cities contain the same six columns:

● Start Time (e.g., 2017-01-01 00:07:57)

● End Time (e.g., 2017-01-01 00:20:53)

● Trip Duration (in seconds - e.g., 776)

● Start Station (e.g., Broadway & Barry Ave)

● End Station (e.g., Sedgwick St & North Ave)

● User Type (Subscriber or Customer)

The Chicago and New York City files also contain the following two columns:

● Gender

● Birth Year

**Tasks**:

You are required to write Python code that prints answers to the following questions about the dataset. (*10 points each*)

1. What is the most popular month for start time?
2. What is the most popular day of week (Monday, Tuesday, etc.) for start time?
3. What is the most popular hour of day for start time?
4. What is the total trip duration and average trip duration?
5. What is the most popular start station and most popular end station?
6. What is the most popular trip?
7. What are the counts of each user type?