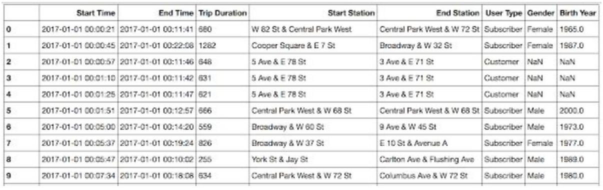
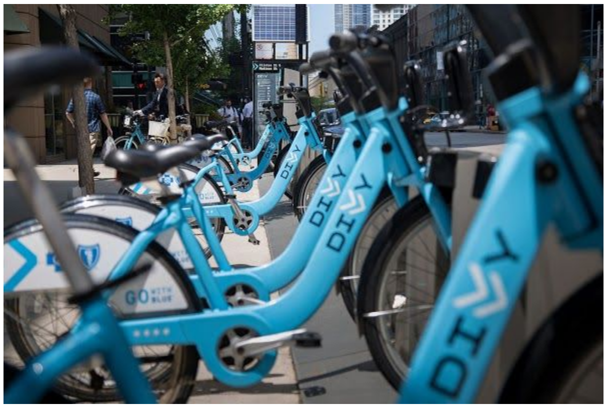


	Start Time	End Time	Trip Duration	Start Station	End Station	User Type	Gender	Birth Year
0	2017-01-01 00:00:21	2017-01-01 00:11:41	680	W 82 St & Central Park West	Central Park West & W 72 St	Subscriber	Female	1965.0
1	2017-01-01 00:00:45	2017-01-01 00:22:08	1282	Cooper Square & E 7 St	Broadway & W 32 St	Subscriber	Female	1987.0
2	2017-01-01 00:00:57	2017-01-01 00:11:46	648	5 Ave & E 78 St	3 Ave & E 71 St	Customer	NaN	NaN
3	2017-01-01 00:01:10	2017-01-01 00:11:42	631	5 Ave & E 78 St	3 Ave & E 71 St	Customer	NaN	NaN
4	2017-01-01 00:01:25	2017-01-01 00:11:47	621	5 Ave & E 78 St	3 Ave & E 71 St	Customer	NaN	NaN
5	2017-01-01 00:01:51	2017-01-01 00:12:57	666	Central Park West & W 68 St	Central Park West & W 68 St	Subscriber	Male	2000.0
6	2017-01-01 00:05:00	2017-01-01 00:14:20	559	Broadway & W 60 St	9 Ave & W 45 St	Subscriber	Male	1973.0
7	2017-01-01 00:05:37	2017-01-01 00:19:24	826	Broadway & W 37 St	E 10 St & Avenue A	Subscriber	Female	1977.0
8	2017-01-01 00:05:47	2017-01-01 00:10:02	255	York St & Jay St	Carlton Ave & Flushing Ave	Subscriber	Male	1989.0
9	2017-01-01 00:07:34	2017-01-01 00:18:08	634	Central Park West & W 72 St	Columbus Ave & W 72 St	Subscriber	Male	1980.0



Project Overview

In this project, you will make use of Python to explore data related to bike share systems for three major cities in the United States—Chicago, New York City, and Washington. You will write code to import the data and answer interesting questions about it by computing descriptive statistics. You will also write a script that takes in raw input to create an interactive experience in the terminal to present these statistics.

What Software Do I Need?
To complete this project, the following software requirements apply:

You should have Python 3, NumPy, and pandas installed using Anaconda
A text editor, like Sublime - https://www.sublimetext.com/ or Atom - https://atom.io/
A terminal application (Terminal on Mac and Linux or Cygwin on Windows).

Bike Share Data
Over the past decade, bicycle-sharing systems have been growing in number and popularity in cities across the world. Bicycle-sharing systems allow users to rent bicycles on a very short-term basis for a price. This allows people to borrow a bike from point A and return it at point B, though they can also return it to the same location if they'd like to just go for a ride. Regardless, each bike can serve several users per day.

Thanks to the rise in information technologies, it is easy for a user of the system to access a dock within the system to unlock or return bicycles. These technologies also provide a wealth of data that can be used to explore how these bike-sharing systems are used.

In this project, you will use data provided by Motivate - https://www.motivateco.com/ ...a bike share system provider for many major cities in the United States, to uncover bike share usage patterns. You will compare the system usage between three large cities: Chicago, New York City, and Washington, DC.

The Datasets
Randomly selected data for the first six months of 2017 are provided for all three cities. All three of the data files contain the same core six (6) 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 have the following two columns:

Gender
Birth Year.

A screenshot of Data, for the first 10 rides in the new_york_city.csv file is attached.

The original files are much larger and messier, and you don't need to download them, but they can be accessed here if you'd like to see them (Chicago(https://www.divvybikes.com/system-data), New York City(https://www.citibikenyc.com/system-data), Washington(https://www.capitalbikeshare.com/system-data)). These files had more columns and they differed in format in many cases. Some data wrangling has been performed to condense these files to the above core six columns to make your analysis and the evaluation of your Python skills more straightforward. In the Data Wrangling course that comes later in the Data Analyst Nanodegree program, students learn how to wrangle the dirtiest, messiest datasets, so don't worry, you won't miss out on learning this important skill!

Statistics Computed
You will learn about bike share use in Chicago, New York City, and Washington by computing a variety of descriptive statistics. In this project, you'll write code to provide the following information:

#1 Popular times of travel (i.e., occurs most often in the start time)

most common month
most common day of week
most common hour of day

#2 Popular stations and trip

most common start station
most common end station
most common trip from start to end (i.e., most frequent combination of start station and end station)

#3 Trip duration

total travel time
average travel time

#4 User info

counts of each user type
counts of each gender (only available for NYC and Chicago)
earliest, most recent, most common year of birth (only available for NYC and Chicago)

The Files
To answer these questions using Python, you will need to write a Python script. To help guide your work in this project, a template with helper code and comments is provided in a bikeshare.py file, and you will do your scripting in there also. You will need the three city dataset files too:

- chicago.csv
- new_york_city.csv
- washington.csv

All four of these files are zipped up in the Bikeshare file, you will find in the resource folder.

You may use the template provided in bikeshare.py to complete this project. You should feel free to change the template however you'd like, as long as your code provides the statistics shown in the template, and allows a user to give input on which data they would like to see.

WATCH THE VIDEO ON "BIKESHARE PROJECT WALKTHROUGH" BEFORE READING THIS!
An Interactive Experience
The bikeshare.py file is set up as a script that takes in raw input to create an interactive experience in the terminal that answers questions about the dataset. The experience is interactive because depending on a user's input, the answers to the questions on the previous page will change! There are four questions that will change the answers:

1. Would you like to see data for Chicago, New York, or Washington?
2. Would you like to filter the data by month, day, or not at all?
(If they chose month) Which month - January, February, March, April, May, or June?
(If they chose day) Which day - Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, or Sunday?

The answers to the questions above will determine the city and timeframe on which you'll do data analysis. After filtering the dataset, users will see the statistical result of the data, and choose to start again or exit.

Remember that any time you ask users for input, there is a chance they may not enter what you expect, so your code should handle unexpected input well without failing. You need to anticipate raw input errors like using improper upper or lower case, typos, or users misunderstanding what you are expecting. Use the tips provided in the sections of the Scripting lesson in this course to make sure your code does not fail with an execution error due to unexpected raw input.

NOTE: Your script also needs to prompt the user whether they would like want to see the raw data. If the user answers 'yes,' then the script should print 5 rows of the data at a time, then ask the user if they would like to see 5 more rows of the data. The script should continue prompting and printing the next 5 rows at a time until the user chooses 'no,' they do not want any more raw data to be displayed.

Note that this bikeshare.py file is simply a template you can use, but you are not required to use it. You can change the functions however you like as long as you have an ending product that meets the project requirements. Changes to the structure of bikeshare.py (e.g., adding and/or deleting helper functions) that you think make the code more efficient or have a better style are encouraged!

How to Complete and Submit this Project.
You can work on your Python script on your local machine. To do this, you should download the necessary project files from the Resources folder. Before you submit, make sure your project meets all the necessary elements of the Project Rubric.

Project Submission
In this project, you will write Python code to import US bike share data and answer interesting questions about it by computing descriptive statistics. You will also write a script that takes in raw input to create an interactive experience in the terminal to present these statistics.

Before You Submit
Check the Rubric
Your project will be evaluated by your mentor, according to this Project Rubric. Be sure to review it thoroughly before you submit. Your project "meets specifications" only if it meets specifications in all the criteria. If you see room for improvement in any category in which you do not meet specifications, be sure to take some time to revise your work until you feel it is up to expectations. In particular, there is one section of the rubric that focuses on the quality of your code. It is important that you not only obtain the correct answers with your code, but that you have followed good coding practices to obtain your solutions.

Gather Submission Materials
All you need to submit for this project, are two files:

1. bikeshare.py: Your code
2. readme.txt: If you refer to other websites, books, and other resources to help you in solving tasks in the project, make sure that you document them in this file

There is no need for you to include any data files with your submission.

Please, accept my congratulations, on behalf of Wejapa Internships and all mentors, as you endeavour to pull through, a successful project, meeting the stipulated rubrics and deadline.

Yours sincerely,
Samuel George (mi yang).
Lead Mentor,
Data Science - Python,
Wejapa Internships.

May the force be with you! 🙏🏾