14.1.2 Download the Data

Kate suggests that you use Citi Bike data that has been released to the public for your analysis. You agree, but then remind her that you must be cautious when looking at the data, as it applies specifically to New York City—Des Moines is quite different!

You're excited to find out which data will and will not apply to Des Moines, drawing on both your data expertise and critical thinking skills. You'll also need to rely on Kate's salesmanship during the investor conversations. Kate can sell almost anything, but it's up to you to make sure she's selling something viable.

Before we can download our data, we need to create a file structure on our computer. Let's start by navigating to GitHub to create a new repository.

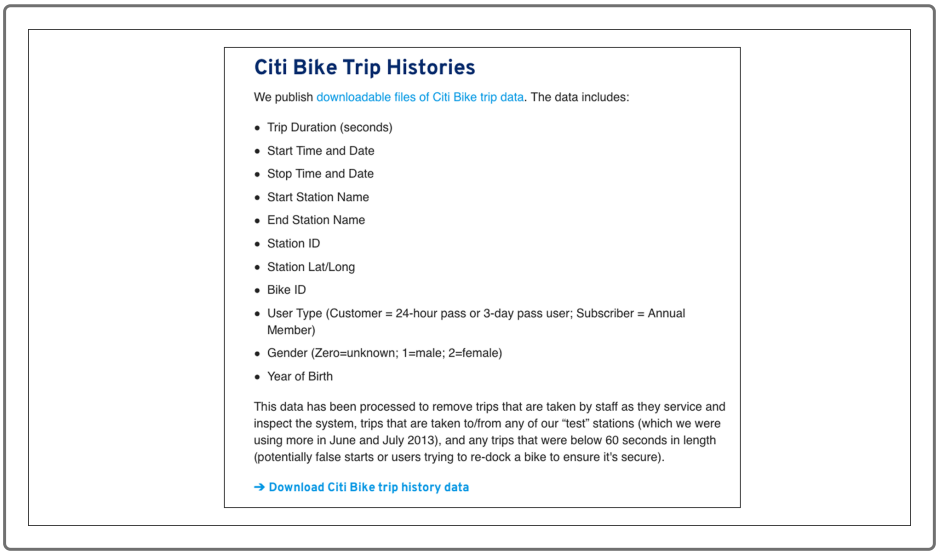
GITHUB

Create a new repository for this module named "bikesharing" and clone the empty repo into your class folder.

Download Data

For this project, we'll use data from the Citi Bike program in New York City. This data includes a variety of fields, which we'll get to shortly. The data you'll download will be contained in a flat file, a CSV.

Go to the **Citi Bike System Data page (https://www.citibikenyc.com/system-data)**. In the "Citi Bike Trip Histories" section, click the link that says "downloadable files of Citi Bike trip data," as shown in the following image:



Citi Bike Trip Histories

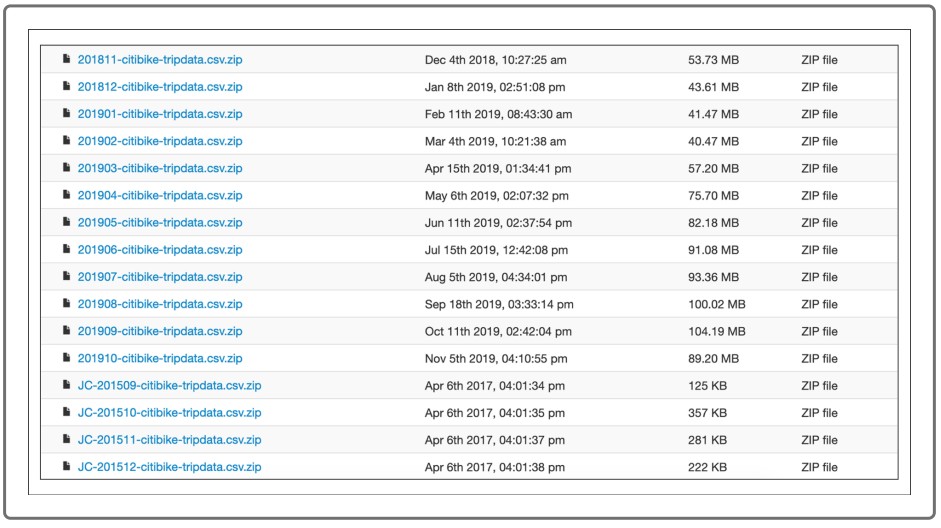
We publish downloadable files of Citi Bike trip data. The data includes:

- Trip Duration (seconds)
- Start Time and Date
- Stop Time and Date
- Start Station Name
- End Station Name
- Station ID
- Station Lat/Long
- Bike ID
- User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member)
- Gender (Zero=unknown; 1=male; 2=female)
- Year of Birth

This data has been processed to remove trips that are taken by staff as they service and inspect the system, trips that are taken to/from any of our "test" stations (which we were using more in June and July 2013), and any trips that were below 60 seconds in length (potentially false starts or users trying to re-dock a bike to ensure it's secure).

➔ Download Citi Bike trip history data

This link will take you to an index of trip data. Scroll down the list to `201908-citibike-tripdata.csv.zip`.



201811-citibike-tripdata.csv.zip	Dec 4th 2018, 10:27:25 am	53.73 MB	ZIP file
201812-citibike-tripdata.csv.zip	Jan 8th 2019, 02:51:08 pm	43.61 MB	ZIP file
201901-citibike-tripdata.csv.zip	Feb 11th 2019, 08:43:30 am	41.47 MB	ZIP file
201902-citibike-tripdata.csv.zip	Mar 4th 2019, 10:21:38 am	40.47 MB	ZIP file
201903-citibike-tripdata.csv.zip	Apr 15th 2019, 01:34:41 pm	57.20 MB	ZIP file
201904-citibike-tripdata.csv.zip	May 6th 2019, 02:07:32 pm	75.70 MB	ZIP file
201905-citibike-tripdata.csv.zip	Jun 11th 2019, 02:37:54 pm	82.18 MB	ZIP file
201906-citibike-tripdata.csv.zip	Jul 15th 2019, 12:42:08 pm	91.08 MB	ZIP file
201907-citibike-tripdata.csv.zip	Aug 5th 2019, 04:34:01 pm	93.36 MB	ZIP file
201908-citibike-tripdata.csv.zip	Sep 18th 2019, 03:33:14 pm	100.02 MB	ZIP file
201909-citibike-tripdata.csv.zip	Oct 11th 2019, 02:42:04 pm	104.19 MB	ZIP file
201910-citibike-tripdata.csv.zip	Nov 5th 2019, 04:10:55 pm	89.20 MB	ZIP file
JC-201509-citibike-tripdata.csv.zip	Apr 6th 2017, 04:01:34 pm	125 KB	ZIP file
JC-201510-citibike-tripdata.csv.zip	Apr 6th 2017, 04:01:35 pm	357 KB	ZIP file
JC-201511-citibike-tripdata.csv.zip	Apr 6th 2017, 04:01:37 pm	281 KB	ZIP file
JC-201512-citibike-tripdata.csv.zip	Apr 6th 2017, 04:01:38 pm	222 KB	ZIP file

This zip file contains all the August 2019 data. We'll use data from August because there is likely more traffic during the summer months. You will need to scroll down and find the file named `201908-citibike-tripdata.csv.zip`. Save this file to your "bikesharing" folder.

When the file is downloaded and saved on your computer, you can move on to setting up your Tableau environment.