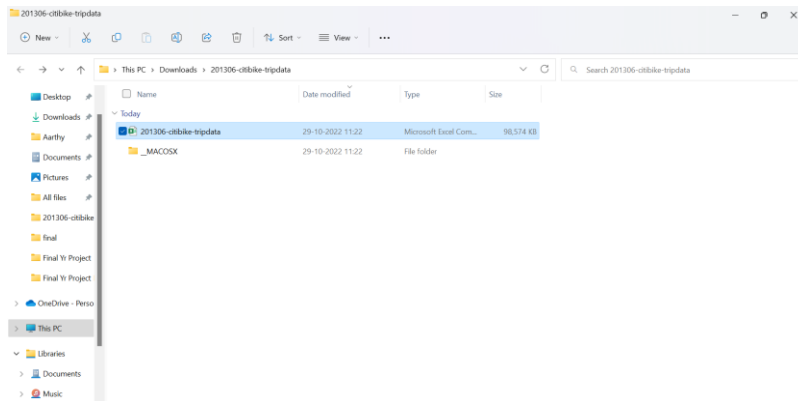


WORKING WITH THE DATASET

Downloading and Understanding the Dataset

Download the zip file from the dataset link and extract



Understand the dataset

	tripduration	starttime	stoptime	start station	start latitude	start longitude	end station	end latitude	end longitude	bikeid	usertype	birth year	gender
1	605	01-06-2013 00:00:01	01-06-2013 00:11:36	444 Broadway	40.74235	-73.9892	434 9 Ave & W	40.74317	-74.0037	19678	Subscriber	1983	1
2	603	01-06-2013 00:00:01	01-06-2013 00:11:36	444 Broadway	40.74235	-73.9892	434 9 Ave & W	40.74317	-74.0037	16649	Subscriber	1984	1
3	2059	01-06-2013 00:00:01	01-06-2013 00:11:36	406 Hicks St &	40.69513	-73.996	406 Hicks St &	40.69513	-73.996	19599	Customer	NULL	0
4	123	01-06-2013 00:00:01	01-06-2013 00:11:36	475 E 15 St & I	40.73524	-73.9876	262 Washingtr	40.69178	-73.9737	16352	Subscriber	1960	1
5	1521	01-06-2013 00:00:01	01-06-2013 00:11:36	2008 Little West	40.70569	-74.0168	310 State St &	40.68927	-73.9891	15567	Subscriber	1983	1
6	2028	01-06-2013 00:00:01	01-06-2013 00:11:36	485 W 37 St &	40.75038	-73.9834	406 Hicks St &	40.69513	-73.996	18445	Customer	NULL	0
7	2057	01-06-2013 00:00:01	01-06-2013 00:11:36	285 Broadway	40.73455	-73.9907	532 S 5 St & S 1	40.71045	-73.9609	15693	Subscriber	1991	1
8	369	01-06-2013 00:00:01	01-06-2013 00:11:36	509 9 Ave & W	40.7455	-74.002	521 8 Ave & W	40.75097	-73.9944	16100	Subscriber	1981	1
9	1829	01-06-2013 00:00:01	01-06-2013 00:11:36	265 Stanton St	40.72229	-73.9915	436 Hancock S	40.68217	-73.954	15234	Subscriber	1984	1
10	829	01-06-2013 00:00:01	01-06-2013 00:11:36	404 9 Ave & W	40.74058	-74.0055	303 Mercer St	40.72363	-73.9995	16400	Subscriber	1987	1
11	1316	01-06-2013 00:00:01	01-06-2013 00:11:36	423 W 54 St &	40.76585	-73.9869	314 Cadman P	40.69383	-73.9905	19781	Subscriber	1960	1
12	1456	01-06-2013 00:00:01	01-06-2013 00:11:36	502 Henry St &	40.71422	-73.9813	532 S 5 St & S 1	40.71045	-73.9609	18886	Customer	NULL	0
13	386	01-06-2013 00:00:01	01-06-2013 00:11:36	241 Dekalb Av	40.68981	-73.9749	365 Fulton St &	40.68223	-73.9615	19039	Subscriber	1981	1
14	924	01-06-2013 00:00:01	01-06-2013 00:11:36	486 Broadway	40.7462	-73.9886	521 8 Ave & W	40.75097	-73.9944	16608	Customer	NULL	0
15	1213	01-06-2013 00:00:01	01-06-2013 00:11:36	537 E 33 St & 2	40.74402	-73.9761	296 Division St	40.71413	-73.987	14761	Subscriber	1987	1
16	512	01-06-2013 00:00:01	01-06-2013 00:11:36	309 Murray St	40.71498	-74.013	300 Shevcheni	40.72815	-73.9902	19080	Subscriber	1979	2
17	505	01-06-2013 00:00:01	01-06-2013 00:11:36	309 Murray St	40.71498	-74.013	347 Greenwich	40.72885	-74.0086	16798	Subscriber	1984	1
18	833	01-06-2013 00:00:01	01-06-2013 00:11:36	503 E 20 St & I	40.73827	-73.9875	503 E 20 St & I	40.73827	-73.9875	19072	Customer	NULL	0
19	1818	01-06-2013 00:00:01	01-06-2013 00:11:36	257 Uppercad	40.71939	-74.0025	500 Broadway	40.76229	-73.9834	20349	Customer	NULL	0

1. Trip Duration: How long a trip lasted in seconds
2. Start Date and Time: EX->01-06-2013 00:00:01
3. Stop Date and Time: EX->01-06-2013 00:11:36
4. Start Station ID: Unique identifier for each station
5. Start Station Name
6. Start Station Latitude: Coordinates
7. Start Station Longitude: Coordinates
8. End Station ID: Unique identifier for each station
9. End Station Name

10. End Station Latitude

11. End Station Longitude

12. Bike ID: Unique identifier for each bike

13. User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member): Customers are usually tourists, subscribers are usually NYC residents

14. Year of Birth: Self-entered, not validated by an ID

15. Gender (Zero=unknown; 1=male; 2=female): Usually unknown for customers since they often sign up at a kiosk