

```
In [1]: import pandas as pd
```

```
In [2]: bike1=pd.read_csv("201907-citibike-tripdata.csv")
bike1.head()
```

Out[2]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	897	2019-07-01 00:00:00.1320	2019-07-01 00:14:58.0040	493.0	W 45 St & 6 Ave	40.756800	-73.982912	454.0
1	267	2019-07-01 00:00:05.1780	2019-07-01 00:04:32.4500	3143.0	5 Ave & E 78 St	40.776321	-73.964274	3226.0
2	2201	2019-07-01 00:00:05.2130	2019-07-01 00:36:46.7490	317.0	E 6 St & Avenue B	40.724537	-73.981854	3469.0
3	1660	2019-07-01 00:00:08.6010	2019-07-01 00:27:48.8050	249.0	Harrison St & Hudson St	40.718710	-74.009001	369.0
4	109	2019-07-01 00:00:12.1580	2019-07-01 00:02:01.5670	3552.0	W 113 St & Broadway	40.805973	-73.964928	3538.0

```
In [3]: bike1["month"]=7
bike1.head()
```

Out[3]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	897	2019-07-01 00:00:00.1320	2019-07-01 00:14:58.0040	493.0	W 45 St & 6 Ave	40.756800	-73.982912	454.0
1	267	2019-07-01 00:00:05.1780	2019-07-01 00:04:32.4500	3143.0	5 Ave & E 78 St	40.776321	-73.964274	3226.0
2	2201	2019-07-01 00:00:05.2130	2019-07-01 00:36:46.7490	317.0	E 6 St & Avenue B	40.724537	-73.981854	3469.0
3	1660	2019-07-01 00:00:08.6010	2019-07-01 00:27:48.8050	249.0	Harrison St & Hudson St	40.718710	-74.009001	369.0
4	109	2019-07-01 00:00:12.1580	2019-07-01 00:02:01.5670	3552.0	W 113 St & Broadway	40.805973	-73.964928	3538.0

```
In [4]: bike2=pd.read_csv("201908-citibike-tripdata.csv")
bike2.head()
```

Out[4]:

	duration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	e stati nar
	393	2019-08-01 00:00:01.4680	2019-08-01 00:06:35.3780	531.0	Forsyth St & Broome St	40.718939	-73.992663	408.0	Marl S Cherry
	627	2019-08-01 00:00:01.9290	2019-08-01 00:10:29.7840	274.0	Lafayette Ave & Fort Greene Pl	40.686919	-73.976682	3409.0	Berg S Smith
	1132	2019-08-01 00:00:04.0480	2019-08-01 00:18:56.1650	2000.0	Front St & Washington St	40.702551	-73.989402	3388.0	Preside S Henry
	1780	2019-08-01 00:00:04.1630	2019-08-01 00:29:44.7940	479.0	9 Ave & W 45 St	40.760193	-73.991255	473.0	Rivingt S Chrys
	1517	2019-08-01 00:00:05.4580	2019-08-01 00:25:23.4550	3312.0	1 Ave & E 94 St	40.781721	-73.945940	3312.0	1 Ave E 94

```
In [5]: bike2["month"]=8
bike2.head()
```

Out[5]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	393	2019-08-01 00:00:01.4680	2019-08-01 00:06:35.3780	531.0	Forsyth St & Broome St	40.718939	-73.992663	408.0
1	627	2019-08-01 00:00:01.9290	2019-08-01 00:10:29.7840	274.0	Lafayette Ave & Fort Greene Pl	40.686919	-73.976682	3409.0
2	1132	2019-08-01 00:00:04.0480	2019-08-01 00:18:56.1650	2000.0	Front St & Washington St	40.702551	-73.989402	3388.0
3	1780	2019-08-01 00:00:04.1630	2019-08-01 00:29:44.7940	479.0	9 Ave & W 45 St	40.760193	-73.991255	473.0
4	1517	2019-08-01 00:00:05.4580	2019-08-01 00:25:23.4550	3312.0	1 Ave & E 94 St	40.781721	-73.945940	3312.0

```
In [6]: bike3=pd.read_csv("201909-citibike-tripdata.csv")
bike3.head()
```

Out[6]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	327	2019-09-01 00:00:01.9580	2019-09-01 00:05:29.3410	3733	Avenue C & E 18 St	40.730563	-73.973984	504
1	1145	2019-09-01 00:00:04.1430	2019-09-01 00:19:09.8360	3329	Degraw St & Smith St	40.682915	-73.993182	270
2	1293	2019-09-01 00:00:07.3090	2019-09-01 00:21:40.7580	3168	Central Park West & W 85 St	40.784727	-73.969617	423
3	1753	2019-09-01 00:00:08.0640	2019-09-01 00:29:21.5040	3299	E 98 St & Park Ave	40.788130	-73.952060	3160
4	613	2019-09-01 00:00:12.8510	2019-09-01 00:10:26.1850	486	Broadway & W 29 St	40.746201	-73.988557	478

```
In [7]: bike3["month"]=9
bike3.head()
```

Out[7]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	327	2019-09-01 00:00:01.9580	2019-09-01 00:05:29.3410	3733	Avenue C & E 18 St	40.730563	-73.973984	504
1	1145	2019-09-01 00:00:04.1430	2019-09-01 00:19:09.8360	3329	Degraw St & Smith St	40.682915	-73.993182	270
2	1293	2019-09-01 00:00:07.3090	2019-09-01 00:21:40.7580	3168	Central Park West & W 85 St	40.784727	-73.969617	423
3	1753	2019-09-01 00:00:08.0640	2019-09-01 00:29:21.5040	3299	E 98 St & Park Ave	40.788130	-73.952060	3160
4	613	2019-09-01 00:00:12.8510	2019-09-01 00:10:26.1850	486	Broadway & W 29 St	40.746201	-73.988557	478

```
In [8]: bike4=pd.read_csv("201910-citibike-tripdata.csv")
bike4.head()
```

Out[8]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id
0	527	2019-10-01 00:00:05.6180	2019-10-01 00:08:52.9430	3746	6 Ave & Broome St	40.724308	-74.004730	223
1	174	2019-10-01 00:00:15.8750	2019-10-01 00:03:10.1680	3301	Columbus Ave & W 95 St	40.791956	-73.968087	3283
2	759	2019-10-01 00:00:19.8240	2019-10-01 00:12:59.7070	161	LaGuardia PI & W 3 St	40.729170	-73.998102	174
3	615	2019-10-01 00:00:21.0680	2019-10-01 00:10:36.6790	254	W 11 St & 6 Ave	40.735324	-73.998004	477
4	761	2019-10-01 00:00:26.3800	2019-10-01 00:13:08.3130	161	LaGuardia PI & W 3 St	40.729170	-73.998102	174

```
In [9]: bike4["month"]=10
bike4.head()
```

Out[9]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	end station name
0	527	2019-10-01 00:00:05.6180	2019-10-01 00:08:52.9430	3746	6 Ave & Broome St	40.724308	-74.004730		
1	174	2019-10-01 00:00:15.8750	2019-10-01 00:03:10.1680	3301	Columbus Ave & W 95 St	40.791956	-73.968087	3	
2	759	2019-10-01 00:00:19.8240	2019-10-01 00:12:59.7070	161	LaGuardia PI & W 3 St	40.729170	-73.998102		
3	615	2019-10-01 00:00:21.0680	2019-10-01 00:10:36.6790	254	W 11 St & 6 Ave	40.735324	-73.998004		
4	761	2019-10-01 00:00:26.3800	2019-10-01 00:13:08.3130	161	LaGuardia PI & W 3 St	40.729170	-73.998102		

```
In [10]: bike5=pd.read_csv("201911-citibike-tripdata.csv")
bike5.head()
```

Out[10]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	end station name
0	273	00:00.3	04:34.2	237	E 11 St & 2 Ave	40.730473	-73.986724	335	Washington PI & Broadway
1	159	00:03.6	02:43.5	3572	34 Ave & 38 St	40.756913	-73.921631	3568	34 St & 35 Ave
2	967	00:06.3	16:14.1	334	W 20 St & 7 Ave	40.742388	-73.997262	529	W 42 St & 8 Ave
3	604	00:21.8	10:25.9	280	E 10 St & 5 Ave	40.733320	-73.995101	445	E 10 St & Avenue A
4	153	00:25.2	02:59.0	3761	Cedar St & Myrtle Ave	40.697842	-73.926241	3846	Central Ave & Himrod St

```
In [11]: bike5["month"]=11
bike5.head()
```

Out[11]:

ripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	end station name	s la
273	00:00.3	04:34.2	237	E 11 St & 2 Ave	40.730473	-73.986724	335	Washington PI & Broadway	40.7
159	00:03.6	02:43.5	3572	34 Ave & 38 St	40.756913	-73.921631	3568	34 St & 35 Ave	40.7
967	00:06.3	16:14.1	334	W 20 St & 7 Ave	40.742388	-73.997262	529	W 42 St & 8 Ave	40.7
604	00:21.8	10:25.9	280	E 10 St & 5 Ave	40.733320	-73.995101	445	E 10 St & Avenue A	40.7
153	00:25.2	02:59.0	3761	Cedar St & Myrtle Ave	40.697842	-73.926241	3846	Central Ave & Himrod St	40.6

```
In [13]: bikenyc=pd.concat([bike1,bike2,bike3,bike4,bike5])
bikenyc.head()
```

Out[13]:

starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	end station name	s la
2019-07-01 00:00:00.1320	2019-07-01 00:14:58.0040	493.0	W 45 St & 6 Ave	40.756800	-73.982912	454.0	E 51 St & 1 Ave	40.7
2019-07-01 00:00:05.1780	2019-07-01 00:04:32.4500	3143.0	5 Ave & E 78 St	40.776321	-73.964274	3226.0	W 82 St & Central Park West	40.7
2019-07-01 00:00:05.2130	2019-07-01 00:36:46.7490	317.0	E 6 St & Avenue B	40.724537	-73.981854	3469.0	India St & West St	40.7
2019-07-01 00:00:08.6010	2019-07-01 00:27:48.8050	249.0	Harrison St & Hudson St	40.718710	-74.009001	369.0	Washington PI & 6 Ave	40.7
2019-07-01 00:00:12.1580	2019-07-01 00:02:01.5670	3552.0	W 113 St & Broadway	40.805973	-73.964928	3538.0	W 110 St & Amsterdam Ave	40.8

```
In [14]: bikenyc.to_csv("bike_trip_data")
```

```
In [15]: bikeall=bikenyc.sample(n=100000)
```

```
In [16]: bikeall.tail()
```

Out[16]:

	tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	s
1372345	982	2019-07-20 16:14:09.5430	2019-07-20 16:30:32.1550	3256.0	Pier 40 - Hudson River Park	40.727714	-74.011296	
949105	429	2019-09-13 16:31:52.2170	2019-09-13 16:39:02.0710	325.0	E 19 St & 3 Ave	40.736245	-73.984738	
361041	850	2019-09-06 06:02:06.4370	2019-09-06 06:16:16.4680	349.0	Rivington St & Ridge St	40.718502	-73.983299	
1321300	1009	2019-10-19 16:46:58.9860	2019-10-19 17:03:48.0530	3323.0	W 106 St & Central Park West	40.798186	-73.960591	3
770381	783	2019-09-11 12:21:59.7360	2019-09-11 12:35:03.1940	497.0	E 17 St & Broadway	40.737050	-73.990093	

```
In [17]: len(bikenyc["tripduration"])
```

Out[17]: 10111336

```
In [18]: bikeall.to_csv("bike_data")
```

```
In [19]: len(bikeall["tripduration"])
```

Out[19]: 100000

```
In [ ]:
```