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

In [2]:
```python
# 1. Create a DataFrame for the 201908-citibike-tripdata data.
citibike_data = "201908-citibike-tripdata.csv"
citibike_df = pd.read_csv(citibike_data)
```

In [4]: `citibike_df.info`

Out[4]: `<bound method DataFrame.info of`   393   2019-08-01 00:00:01.4680
        2019-08-01 00:06:35.3780   531.0 \
        0   627   2019-08-01 00:00:01.9290   2019-08-01 00:10:29.7840   274.0
        1   1132   2019-08-01 00:00:04.0480   2019-08-01 00:18:56.1650   2000.0
        2   1780   2019-08-01 00:00:04.1630   2019-08-01 00:29:44.7940   479.0
        3   1517   2019-08-01 00:00:05.4580   2019-08-01 00:25:23.4550   3312.0
        4   632   2019-08-01 00:00:11.8060   2019-08-01 00:10:43.9590   173.0
        ...   .
        2344218   216   2019-08-31 23:59:46.2930   2019-09-01 00:03:23.0360   116.0
        2344219   117   2019-08-31 23:59:47.7970   2019-09-01 00:01:45.7440   3429.0
        2344220   1614   2019-08-31 23:59:48.1560   2019-09-01 00:26:42.7760   387.0
        2344221   1301   2019-08-31 23:59:58.3620   2019-09-01 00:21:39.7040   3168.0
        2344222   419   2019-08-31 23:59:59.4520   2019-09-01 00:06:59.0210   447.0

                              Forsyth St & Broome St   40.71893904   -73.99266288   408.
        0   \
        0   Lafayette Ave & Fort Greene Pl   40.686919   -73.976682   3409.
        0
        1   Front St & Washington St   40.702551   -73.989402   3388.
        0
        2   9 Ave & W 45 St   40.760193   -73.991255   473.
        0
        3   1 Ave & E 94 St   40.781721   -73.945940   3312.
        0
        4   Broadway & W 49 St   40.760683   -73.984527   3707.
        0
        ...   .
        ...
        2344218   W 17 St & 8 Ave   40.741776   -74.001497   509.
        0
        2344219   Hanson Pl & Ashland Pl   40.685068   -73.977908   353.
        0
        2344220   Centre St & Chambers St   40.712733   -74.004607   3440.
        0
        2344221   Central Park West & W 85 St   40.784727   -73.969617   423.
        0
        2344222   8 Ave & W 52 St   40.763707   -73.985162   529.
        0

                              Market St & Cherry St   40.71076228   -73.99400398   35305 \
        0   Bergen St & Smith St   40.686744   -73.990632   38822
        1   President St & Henry St   40.682800   -73.999904   18373
        2   Rivington St & Chrystie St   40.721101   -73.991925   25002
        3   1 Ave & E 94 St   40.781721   -73.945940   31198
        4   Lexington Ave & E 26 St   40.741459   -73.983293   21628
        ...   .
        2344218   9 Ave & W 22 St   40.745497   -74.001971   38655
        2344219   S Portland Ave & Hanson Pl   40.685396   -73.974315   25264
        2344220   Fulton St & Adams St   40.692418   -73.989495   28485
        2344221   W 54 St & 9 Ave   40.765849   -73.986905   38664
        2344222   W 42 St & 8 Ave   40.757570   -73.990985   35210

                      Subscriber   1996   2
        0   Subscriber   1998   2
        1   Subscriber   1988   1

```
2          Subscriber   1988  1
3          Subscriber   1965  2
4          Subscriber   1998  1
...                ...   ...  ..
2344218    Subscriber   1972  1
2344219    Subscriber   1982  2
2344220      Customer   1969  0
2344221      Customer   1969  0
2344222    Subscriber   1994  1

[2344223 rows x 15 columns]>
```

In [5]: `# 2. Check the datatypes of your columns.`
`citibike_df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2344223 entries, 0 to 2344222
Data columns (total 15 columns):
 #   Column                     Dtype
---  ------                     -----
 0   393                        int64
 1   2019-08-01 00:00:01.4680   object
 2   2019-08-01 00:06:35.3780   object
 3   531.0                      float64
 4   Forsyth St & Broome St     object
 5   40.71893904                float64
 6   -73.99266288               float64
 7   408.0                      float64
 8   Market St & Cherry St      object
 9   40.71076228                float64
 10  -73.99400398               float64
 11  35305                      int64
 12  Subscriber                 object
 13  1996                       int64
 14  2                          int64
dtypes: float64(6), int64(4), object(5)
memory usage: 268.3+ MB
```

```
In [6]:  # 3. Convert the 'tripduration' column to datetime datatype.
         citibike_df['393_orig'] = citibike_df['393']
         citibike_df['393'] = pd.to_datetime(citibike_df['393'], unit='m')
         citibike_df.head()
```

Out[6]:

	393	2019-08-01 00:00:01.4680	2019-08-01 00:06:35.3780	531.0	Forsyth St & Broome St	40.71893904	-73.99266288	408.0	
0	1970-01-01 10:27:00	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	S
1	1970-01-01 18:52:00	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	Pr H
2	1970-01-02 05:40:00	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	Ri C
3	1970-01-02 01:17:00	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 A
4	1970-01-01 10:32:00	2019-08-01 00:00:11.8060	2019-08-01 00:10:43.9590	173.0	Broadway & W 49 St	40.760683	-73.984527	3707.0	Lex A

In [7]: *# 4. Check the datatypes of your columns.*
```python
citibike_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2344223 entries, 0 to 2344222
Data columns (total 16 columns):
 #   Column                    Dtype
---  ------                    -----
 0   393                       datetime64[ns]
 1   2019-08-01 00:00:01.4680  object
 2   2019-08-01 00:06:35.3780  object
 3   531.0                     float64
 4   Forsyth St & Broome St    object
 5   40.71893904               float64
 6   -73.99266288              float64
 7   408.0                     float64
 8   Market St & Cherry St     object
 9   40.71076228               float64
 10  -73.99400398              float64
 11  35305                     int64
 12  Subscriber                object
 13  1996                      int64
 14  2                         int64
 15  393_orig                  int64
dtypes: datetime64[ns](1), float64(6), int64(4), object(5)
memory usage: 286.2+ MB
```

In [8]: *# 5. Export the Dataframe as a new CSV file without the index.*
```python
citibike_df.to_csv('citibike_201908_updt.csv', index=False)
```

In [ ]: