

```
In [1]: !pip install kaggle
```

```
Requirement already satisfied: kaggle in c:\users\shiron salim\anaconda3\lib\site-packages (1.5.16)
Requirement already satisfied: six>=1.10 in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (1.16.0)
Requirement already satisfied: certifi in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (2023.7.22)
Requirement already satisfied: python-dateutil in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (2.8.2)
Requirement already satisfied: requests in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (2.31.0)
Requirement already satisfied: tqdm in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (4.65.0)
Requirement already satisfied: python-slugify in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (5.0.2)
Requirement already satisfied: urllib3 in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (1.26.16)
Requirement already satisfied: bleach in c:\users\shiron salim\anaconda3\lib\site-packages (from kaggle) (4.1.0)
Requirement already satisfied: packaging in c:\users\shiron salim\anaconda3\lib\site-packages (from bleach->kaggle) (23.1)
Requirement already satisfied: webencodings in c:\users\shiron salim\anaconda3\lib\site-packages (from bleach->kaggle) (0.5.1)
Requirement already satisfied: text-unidecode>=1.3 in c:\users\shiron salim\anaconda3\lib\site-packages (from python-slugify->kaggle) (1.3)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\shiron salim\anaconda3\lib\site-packages (from requests->kaggle) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\shiron salim\anaconda3\lib\site-packages (from requests->kaggle) (3.4)
Requirement already satisfied: colorama in c:\users\shiron salim\anaconda3\lib\site-packages (from tqdm->kaggle) (0.4.6)
```

```
In [2]: import kaggle
```

```
In [3]: import zipfile
```

```
In [4]: !kaggle datasets download -d hnavrodiev/london-bike-sharing-dataset
```

```
london-bike-sharing-dataset.zip: Skipping, found more recently modified local copy (use --force to force download)
```

```
In [5]: import pandas as pd
bikes=pd.read_csv(r"C:\Users\SHIRON SALIM\Downloads\london_merged.csv")
```

```
In [6]: bikes.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17414 entries, 0 to 17413
Data columns (total 10 columns):
 #   Column          Non-Null Count  Dtype  
---  -
 0   timestamp      17414 non-null  object 
 1   cnt             17414 non-null  int64  
 2   t1              17414 non-null  float64
 3   t2              17414 non-null  float64
 4   hum             17414 non-null  float64
 5   wind_speed     17414 non-null  float64
 6   weather_code    17414 non-null  float64
 7   is_holiday     17414 non-null  float64
 8   is_weekend     17414 non-null  float64
 9   season         17414 non-null  float64
```

```
dtypes: float64(8), int64(1), object(1)
memory usage: 1.3+ MB
```

```
In [7]: bikes.shape
```

```
Out[7]: (17414, 10)
```

```
In [8]: bikes
```

```
Out[8]:
```

	timestamp	cnt	t1	t2	hum	wind_speed	weather_code	is_holiday	is_weekend	season
0	2015-01-04 00:00:00	182	3.0	2.0	93.0	6.0	3.0	0.0	1.0	3.0
1	2015-01-04 01:00:00	138	3.0	2.5	93.0	5.0	1.0	0.0	1.0	3.0
2	2015-01-04 02:00:00	134	2.5	2.5	96.5	0.0	1.0	0.0	1.0	3.0
3	2015-01-04 03:00:00	72	2.0	2.0	100.0	0.0	1.0	0.0	1.0	3.0
4	2015-01-04 04:00:00	47	2.0	0.0	93.0	6.5	1.0	0.0	1.0	3.0
...
17409	2017-01-03 19:00:00	1042	5.0	1.0	81.0	19.0	3.0	0.0	0.0	3.0
17410	2017-01-03 20:00:00	541	5.0	1.0	81.0	21.0	4.0	0.0	0.0	3.0
17411	2017-01-03 21:00:00	337	5.5	1.5	78.5	24.0	4.0	0.0	0.0	3.0
17412	2017-01-03 22:00:00	224	5.5	1.5	76.0	23.0	4.0	0.0	0.0	3.0
17413	2017-01-03 23:00:00	139	5.0	1.0	76.0	22.0	2.0	0.0	0.0	3.0

17414 rows × 10 columns

```
In [9]: bikes.weather_code.value_counts()
```

```
Out[9]: weather_code
1.0      6150
2.0      4034
3.0      3551
7.0      2141
4.0      1464
26.0        60
10.0        14
Name: count, dtype: int64
```

```
In [10]: bikes.season.value_counts()
```

```
Out[10]: season
0.0      4394
1.0      4387
3.0      4330
2.0      4303
Name: count, dtype: int64
```

```
In [11]: new_cols_dict={
    'timestamp': 'time',
    'cnt': 'count',
    't1': 'temp_real_c',
    't2': 'temp_feels_like_c',
    'hum': 'humidity_percent',
    'wind_speed': 'wind_speed_kph',
    'weather_code': 'weather',
    'is_holiday': 'is_holiday',
    'is_weekend': 'is_weekend',
    'season': 'season'
```

```
}

bikes.rename(new_cols_dict,axis=1,inplace=True)

bikes.head()
```

Out[11]:

	time	count	temp_real_c	temp_feels_like_c	humidity_percent	wind_speed_kph	weather	is_holiday	is_we
0	2015-01-04 00:00:00	182	3.0	2.0	93.0	6.0	3.0	0.0	
1	2015-01-04 01:00:00	138	3.0	2.5	93.0	5.0	1.0	0.0	
2	2015-01-04 02:00:00	134	2.5	2.5	96.5	0.0	1.0	0.0	
3	2015-01-04 03:00:00	72	2.0	2.0	100.0	0.0	1.0	0.0	
4	2015-01-04 04:00:00	47	2.0	0.0	93.0	6.5	1.0	0.0	

In [12]:

```
bikes.humidity_percent=bikes.humidity_percent / 100
bikes.head()
```

Out[12]:

	time	count	temp_real_c	temp_feels_like_c	humidity_percent	wind_speed_kph	weather	is_holiday	is_we
0	2015-01-04 00:00:00	182	3.0	2.0	0.930	6.0	3.0	0.0	
1	2015-01-04 01:00:00	138	3.0	2.5	0.930	5.0	1.0	0.0	
2	2015-01-04 02:00:00	134	2.5	2.5	0.965	0.0	1.0	0.0	
3	2015-01-04 03:00:00	72	2.0	2.0	1.000	0.0	1.0	0.0	
4	2015-01-04 04:00:00	47	2.0	0.0	0.930	6.5	1.0	0.0	

In [13]:

```
season_dict={
    '0.0':'spring',
    '1.0':'summer',
    '2.0':'autumn',
    '3.0':'winter'
}

weather_dict={
    '1.0':'clear',
    '2.0':'Scattered clouds',
```

```

        '3.0':'Broken clouds',
        '4.0':'cloudy',
        '7.0':'Rain',
        '10.0':'Rain with thunderstorm',
        '26.0':'snowfall'
    }

    bikes.season=bikes.season.astype('str')
    bikes.season=bikes.season.map(season_dict)

    bikes.weather=bikes.weather.astype('str')

    bikes.weather=bikes.weather.map(weather_dict)

```

In [14]: `bikes.head()`

Out[14]:

	time	count	temp_real_c	temp_feels_like_c	humidity_percent	wind_speed_kph	weather	is_holiday	is_we
0	2015-01-04 00:00:00	182	3.0	2.0	0.930	6.0	Broken clouds	0.0	
1	2015-01-04 01:00:00	138	3.0	2.5	0.930	5.0	clear	0.0	
2	2015-01-04 02:00:00	134	2.5	2.5	0.965	0.0	clear	0.0	
3	2015-01-04 03:00:00	72	2.0	2.0	1.000	0.0	clear	0.0	
4	2015-01-04 04:00:00	47	2.0	0.0	0.930	6.5	clear	0.0	

In [16]: `bikes.to_excel('london_bikes_final.xlsx', sheet_name='Data')`

In []: