Import pandas library and read csv file and assing it into dataframe.

In [1]:
```
import pandas as pd
df = pd.read_csv('T1.csv',pa
df.head()
```

Out[1]:

	Date/Time	LV ActivePower (kW)	Wind Speed (m/s)
0	2018-01-01 00:00:00	380.047791	5.311336
1	2018-01-01 00:10:00	453.769196	5.672167
2	2018-01-01 00:20:00	306.376587	5.216037
3	2018-01-01 00:30:00	419.645905	5.659674
4	2018-01-01 00:40:00	380.650696	5.577941

Rename the coloumn.

In [2]:
```
df.rename(columns = {'Date/T
```

In [3]:
```
df.head()
```

Out[3]:

	Date	LV ActivePower (kW)	Wind Speed (m/s)
0	2018-01-01 00:00:00	380.047791	5.311336
1	2018-		

1	2018-01-01 00:10:00	453.769196	5.672167
2	2018-01-01 00:20:00	306.376587	5.216037
3	2018-01-01 00:30:00	419.645905	5.659674
4	2018-01-01 00:40:00	380.650696	5.577941

In [4]:
```
type('Date[0]')
```

Out[4]:
```
str
```

Set index as Date coloumn.

In [5]:
```
df.set_index('Date',inplace=
df.head()
```

Out[5]:

Date	LV ActivePower (kW)	Wind Speed (m/s)	The
2018-01-01 00:00:00	380.047791	5.311336	
2018-01-01 00:10:00	453.769196	5.672167	
2018-01-01 00:20:00	306.376587	5.216037	
2018-01-01 00:30:00	419.645905	5.659674	
2018-01-01 00:40:00	380.650696	5.577941	

To fill the data make full list of

To fill the data make full list of day in year of 2018 at frequency of 10 min and reindex the dataframe.

In [6]:
```
dt=pd.date_range("01-01-2018
idx = pd.DatetimeIndex(dt)
df = df.reindex(idx)
df.head()
```

Out[6]:

	LV ActivePower (kW)	Wind Speed (m/s)	The
2018-01-01 00:00:00	380.047791	5.311336	
2018-01-01 00:10:00	453.769196	5.672167	
2018-01-01 00:20:00	306.376587	5.216037	
2018-01-01 00:30:00	419.645905	5.659674	
2018-01-01 00:40:00	380.650696	5.577941	

In [7]:
```
df.tail()
```

Out[7]:

	LV ActivePower (kW)	Wind Speed (m/s)	The
2018-12-31 23:20:00	1684.353027	7.332648	
2018-12-31 23:30:00	2201.106934	8.435358	
2018-12-31 23:40:00	2515.694092	9.421366	

2018-12-31 23:50:00	2820.466064	9.979332
2019-01-01 00:00:00	NaN	NaN

In [8]:
```
df.describe()
```

Out[8]:

	LV ActivePower (kW)	Wind Speed (m/s)
count	50530.000000	50530.000000
mean	1307.684332	7.557952
std	1312.459242	4.227166
min	-2.471405	0.000000
25%	50.677890	4.201395
50%	825.838074	7.104594
75%	2482.507568	10.300020
max	3618.732910	25.206011

Fill nan values with mean of coloumns.

In [9]:
```
new_df1=df.fillna({'LV Activ
```

Remove the unnessecary coloumns.

In [10]:
```
new_df1.drop('Theoretical_Po
```

In [11]:
```
new_df1.drop('Wind Direction
```

In [12]:
```
new_df1.head()
```

Out[12]:

	LV	Wind

Out[12]:

	LV ActivePower (kW)	Wind Speed (m/s)
2018-01-01 00:00:00	380.047791	5.311336
2018-01-01 00:10:00	453.769196	5.672167
2018-01-01 00:20:00	306.376587	5.216037
2018-01-01 00:30:00	419.645905	5.659674
2018-01-01 00:40:00	380.650696	5.577941

In [13]:
```
new_df1.describe()
```

Out[13]:

	LV ActivePower (kW)	Wind Speed (m/s)
count	52561.000000	52561.000000
mean	1307.684332	7.557952
std	1286.851695	4.144689
min	-2.471405	0.000000
25%	71.586853	4.322062
50%	915.141724	7.320701
75%	2399.668945	10.143390
max	3618.732910	25.206011

Save the data frame into csv file.

In [15]:
```
new_df1.to_csv('file6.csv')
```

In [ ]:
```

```