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

In [1]:	<pre>import pandas as pd df = pd.read_csv('T1.csv',pa df.head()</pre>
---------	-----------------------------------------------------------------------

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

Rename the coloumn.

Out[3]:	LV Wind 7 Date ActivePower Speed
In [3]:	df.head()
In [2]:	<pre>df.rename(columns = {'Date/T</pre>

	Date	(kW)	(m/s)	
0	2018- 01-01 00:00:00	380.047791	5.311336	9
ž.	2018-	.======		

<pre>df.head()</pre>		1	201 01-0 00:10:0	01	453.7691	96	5.6721	67
3 01-01 419.645905 5.659674 00:30:00 2018- 4 01-01 380.650696 5.577941 00:40:00 In [4]: type('Date[0]') Out[4]: str Set index as Date coloumn. In [5]: df.set_index('Date',inplace=df.head()) Out[5]:		2	01-0	01	306.3765	87	5.2160	37
4 01-01 380.650696 5.577941 00:40:00 In [4]: type('Date[0]') Out[4]: str Set index as Date coloumn. In [5]: df.set_index('Date',inplace=df.head()) Out[5]:		3	01-0	01	419.6459	05	5.6596	74
Out[4]: str Set index as Date coloumn. In [5]: df.set_index('Date',inplace=df.head() Out[5]:		4	01-0	01	380.6506	96	5.5779	41
Out[4]: str Set index as Date coloumn. In [5]: df.set_index('Date',inplace=df.head() Out[5]:								
Set index as Date coloumn. In [5]: df.set_index('Date',inplace=df.head() Out[5]:	In [4]:	t	ype <mark>('l</mark>	Date	e[0]')			
<pre>In [5]: df.set_index('Date',inplace= df.head() Out[5]:</pre>	Out[4]:	st	r					
Out[5]: LV Wind Speed (kW) (m/s) Date 2018- 01-01 380.047791 5.311336 00:00:00 2018- 01-01 453.769196 5.672167		Se	t index	as i	Date colo	um	n.	
ActivePower (kW) Speed (m/s) Date 2018- 01-01 380.047791 5.311336 00:00:00 2018- 01-01 453.769196 5.672167	In [5]:	<pre>df.set_index('Date',inplace= df.head()</pre>						
2018- 01-01 380.047791 5.311336 00:00:00 2018- 01-01 453.769196 5.672167			i ilica	u()				
01-01 380.047791 5.311336 00:00:00 2018- 01-01 453.769196 5.672167	Out[5]:		Tineat		ivePower		Speed	The
01-01 453.769196 5.672167	Out[5]:				ivePower		Speed	The
	Out[5]:		Date 2018- 01-01	Act	ivePower (kW)		Speed (m/s)	The
2018- 01-01 306.376587 5.216037 00:20:00	Out[5]:	00	Date 2018- 01-01 :00:00 2018- 01-01	Act	ivePower (kW) 0.047791	5.3	Speed (m/s)	The
2018- 01-01 419.645905 5.659674 00:30:00	Out[5]:	00	Date 2018- 01-01 :00:00 2018- 01-01 :10:00	Act 38	ivePower (kW) 0.047791 3.769196	5.6	Speed (m/s) 311336 572167	The
2018- 01-01 380.650696 5.577941 00:40:00	Out[5]:	00	Date 2018- 01-01 :00:00 2018- 01-01 :10:00 2018- 01-01 :20:00	38 45	ivePower (kW) 0.047791 3.769196 6.376587	5.3 5.6 5.2	Speed (m/s) 311336 572167 216037	The

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]:	<pre>dt=pd.date_range("01-01-2018 idx = pd.DatetimeIndex(dt) df = df.reindex(idx) df.head()</pre>

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)	Th€
	2018- 12-31 23:20:00	1684.353027	7.332648	
	2018- 12-31 23:30:00	2201.106934	8.435358	
	2018- 12-31	2515.694092	9.421366	

23:40:00

	2018- 12-31 23:50:00	2820.466064	9.979332		
	2019- 01-01 00:00:00	NaN	I NaN		
In [8]:	1.5				
	dt.des	scribe()			
Out[8]:		LV ActivePower (kW)	Wind Speed (m/s)		
	count 5	0530.000000	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 v coloumr	values with m	ean of		
In [9]:	<pre>new_df1=df.fillna({'LV Activ</pre>				
	Remove the unnessecary coloumns.				
In [10]:	new_df1.drop('Theoretical_Po				
In [11]:	new_df1.drop('Wind Direction				
In [12]:	new_df	f1.head()			
Out[12]:		L	/ 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 []: