Topics : 1. Creating Dataframe 2. Dealing with rows and coloumns 3. Operations : min,max,std,describe 4. Conditional Selection 5. set_index

In [3]:

```
import pandas as pd
df = pd.read_csv("weather_data.csv")
df
```

Out[3]:

	day	temperature	windspeed	event
0	1/1/2017	32	6	Rain
1	1/2/2017	35	7	Sunny
2	1/3/2017	28	2	Snow
3	1/4/2017	24	7	Snow
4	1/5/2017	32	4	Rain
5	1/6/2017	31	2	Sunny

In [4]:

```
df.head() # df.head(3)
```

Out[4]:

	day	temperature	windspeed	event
0	1/1/2017	32	6	Rain
1	1/2/2017	35	7	Sunny
2	1/3/2017	28	2	Snow
3	1/4/2017	24	7	Snow
4	1/5/2017	32	4	Rain

In [5]:

```
df.tail() # df.tail(2)
```

Out[5]:

	day	temperature	windspeed	event
1	1/2/2017	35	7	Sunny
2	1/3/2017	28	2	Snow
3	1/4/2017	24	7	Snow
4	1/5/2017	32	4	Rain
5	1/6/2017	31	2	Sunny

In [7]:

```
df.shape
```

Out[7]:

```
(6, 4)
```

In [11]:

```
rows , coloumns = df.shape
```

In [9]:

```
rows
```

Out[9]:

```
6
```

In [13]:

```
print(coloumns)
```

```
4
```

In [14]:

```
# Slicing
df[2:5] # Includes row 2 and exlcudes 4
```

Out[14]:

	day	temperature	windspeed	event
2	1/3/2017	28	2	Snow
3	1/4/2017	24	7	Snow
4	1/5/2017	32	4	Rain

In [15]:

```
df.columns # To print all the Columns details
```

Out[15]:

```
Index(['day', 'temperature', 'windspeed', 'event'], dtype='object')
```

In [17]:

```
df.day # To print any of the column data
```

Out[17]:

```
0    1/1/2017
1    1/2/2017
2    1/3/2017
3    1/4/2017
4    1/5/2017
5    1/6/2017
Name: day, dtype: object
```

In [18]:

```
# To print only particular columns

df[['event','day']]
```

Out[18]:

	event	day
0	Rain	1/1/2017
1	Sunny	1/2/2017
2	Snow	1/3/2017
3	Snow	1/4/2017
4	Rain	1/5/2017
5	Sunny	1/6/2017

In [19]:

```
df['temperature'].max()
```

Out[19]:

```
35
```

In [21]:

```
df['windspeed'].mean()
```

Out[21]:

```
4.666666666666667
```

In [24]:

```
df.describe()
```

Out[24]:

	temperature	windspeed
count	6.000000	6.000000
mean	30.333333	4.666667
std	3.829708	2.338090
min	24.000000	2.000000
25%	28.750000	2.500000
50%	31.500000	5.000000
75%	32.000000	6.750000
max	35.000000	7.000000

In [27]:

```
# Conditional Selection

df[df.temperature > 28]
```

Out[27]:

	day	temperature	windspeed	event
0	1/1/2017	32	6	Rain
1	1/2/2017	35	7	Sunny
4	1/5/2017	32	4	Rain
5	1/6/2017	31	2	Sunny

In [29]:

```
df[df.temperature == df.temperature.max()]
```

Out[29]:

	day	temperature	windspeed	event
1	1/2/2017	35	7	Sunny

In [31]:

```
# If you don't need the total columns
df[['day','temperature']][df.temperature == df.temperature.max()]
```

Out[31]:

	day	temperature
1	1/2/2017	35

In [33]:

```
# Setting index
df.index
```

Out[33]:

```
RangeIndex(start=0, stop=6, step=1)
```

In [37]:

```
df.set_index('day', inplace=True)
```

In [40]:

```
# To get the particular information of an row
df.loc['1/2/2017']
```

Out[40]:

```
temperature       35
windspeed          7
event          Sunny
Name: 1/2/2017, dtype: object
```

In [41]:

```
# Reset index
df.reset_index(inplace=True)
```

In [42]:

```
df
```

Out[42]:

	day	temperature	windspeed	event
0	1/1/2017	32	6	Rain
1	1/2/2017	35	7	Sunny
2	1/3/2017	28	2	Snow
3	1/4/2017	24	7	Snow
4	1/5/2017	32	4	Rain
5	1/6/2017	31	2	Sunny