1) Creating dataframe

2) Dealing with rows and colums

3) Operations: min, max, std, describe

4) Conditional selection

5) set_index

In [2]:
```
import pandas as pd
df = pd.read_csv("C:/Users/prasa/Desktop/ds projects/panda/weather_data.csv")
df
```

Out[2]:

	day	temperature	windspeed	event
0	01-01-2017	32	6	Rain
1	01-02-2017	35	7	Sunny
2	01-03-2017	28	2	Snow
3	01-04-2017	24	7	Snow
4	01-05-2017	32	4	Rain
5	01-06-2017	31	2	Sunny

In [4]:
```
import pandas as pd
weather_data = {
    'day': ['1/1/2017','1/2/2017','1/3/2017','1/4/2017','1/5/2017','1/6/2017'],
    'temperature': [32,35,28,24,32,31],
    'windspeed': [6,7,2,7,4,2],
    'event': ['Rain', 'Sunny', 'Snow','Snow','Rain', 'Sunny']
}
```

```
df = pd.DataFrame(weather_data)
df
```

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
5	1/6/2017	31	2	Sunny

In [16]:
```
rows,colums=df.shape
```

In [21]:
```
rows
```

Out[21]: 6

In [15]:
```
df.head(3)
```

Out[15]:

	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

In [11]:
```
df.tail(1)
```

Out[11]:

	day	temperature	windspeed	event
5	1/6/2017	31	2	Sunny

In [22]:
```
df[2:5]
```

Out[22]:

	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 [23]:
```
df[:]
```

Out[23]:

	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 [24]:
```
df.columns
```

Out[24]:
```
Index(['day', 'temperature', 'windspeed', 'event'], dtype='object')
```

In [25]:
```
df.day
```

Out[25]:
```
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 [26]:
```
df.event # or [event]
```

Out[26]:
```
0     Rain
1    Sunny
2     Snow
3     Snow
4     Rain
5    Sunny
Name: event, dtype: object
```

In [27]:
```
type(df['event'])
```

Out[27]:
```
pandas.core.series.Series
```

In [28]:
```
df[['event','day','temperature']]
```

Out[28]:

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

In [29]:
```
df
```

Out[29]:

	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

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

In [30]:
```
df['temperature'].max()
```

Out[30]: 35

In [31]:
```
df['temperature'].mean()
```

Out[31]: 30.333333333333332

In [32]:
```
df['temperature'].min()
```

Out[32]: 24

In [33]:
```
df['temperature'].std()
```

Out[33]: 3.8297084310253524

In [34]:
```
df.describe()
```

Out[34]:

	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 [35]:
```
df[df.temperature>=32]
```

Out[35]:

	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

In [36]:
```
df[df.temperature==df.temperature.max()] # or df['temperature'].max()]
```

Out[36]:

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

In [38]:
```
df[['day','temperature']][df.temperature==df.temperature.max()]
```

Out[38]:

	day	temperature
1	1/2/2017	35

Google pandas series operations

In [20]:
```
df
```

Out[20]:

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

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

In [40]:
```
df.index
```

Out[40]:
```
RangeIndex(start=0, stop=6, step=1)
```

In [ ]:
```
df.set_index("day",inplace=True)
```

In [21]:
```
df
```

Out[21]:

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

In [22]:
```
df.loc['1/3/2017']
```

Out[22]:
```
temperature      28
windspeed         2
event          Snow
Name: 1/3/2017, dtype: object
```

In [ ]:

In [33]:
```
df.reset_index(inplace=True)
df
```

Out[33]:

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

In [34]:
```
df.set_index('event',inplace=True)
df
```

Out[34]:

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

In [38]:
```
df
```

Out[38]:

	index	day	temperature	windspeed
event				
Rain	0	1/1/2017	32	6

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

In [40]:
```
df.loc['Snow']
```

Out[40]:

event	index	day	temperature	windspeed
Snow	2	1/3/2017	28	2
Snow	3	1/4/2017	24	7