Learning the DataFrame in Pandas

In [ ]:
```
# importing the pandas

import pandas as pd
```

In [ ]:
```
# Importing the dataset

df = pd.read_csv("D:\\Study\DataScience\\Learning_Datascience\\Learning_Pandas\\Codebas
df
```

Out[ ]:

	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 [ ]:
```
#Shape of the dataframe

df.shape

rows, cols  = df.shape
```

In [ ]:
```
print('Rows : ', rows, ' Cols: ', cols)
```

```
Rows :  6  Cols:  4
```

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

Out[ ]:

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

Out[ ]:

	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 [ ]:
```
df[2:5]
```

Out[ ]:

	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 [ ]:
```
#headings of the cols

df.columns
```

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

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

Out[ ]:
```
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 [ ]:
```
df['day']
```

Out[ ]:
```
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 [ ]:
```
type(df['event'])
```

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

In [ ]:
```
type(df)
```

Out[ ]:
```
pandas.core.frame.DataFrame
```

In [ ]:
```
#for printing custome columns

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

Out[ ]:

	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 [ ]:
```
# now we are doing the analysis

# Now we can find the Max/Min temperatures and also the average temperature

maximumTemperature = df['temperature'].max()
minimumTemperature = df['temperature'].min()
averageTemperature = df['temperature'].mean()

print('Max Temp: ', maximumTemperature,'\nMin Temp: ', minimumTemperature, '\nAvg Temp:
```

```
Max Temp:  35
Min Temp:  24
Avg Temp:  30.333333333333332
```

In [ ]:
```
df
```

Out[ ]:

	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 [ ]:
```
df[['day', 'temperature']][df['event'] == 'Sunny']
```

Out[ ]:

	day	temperature
1	1/2/2017	35
5	1/6/2017	31

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

Out[ ]:

	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 [ ]:
```
df
```

Out[ ]:

	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 [ ]:
```
# as we can see the we have index staring form 0 to 5 but if we want that our dataframe
# to the day column we can do it so that we can give a date as an index and recieve the

# we can see the index by

df.index
```

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

In [ ]:
```
#to set the index of the df to day

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

Out[ ]:

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

day	temperature	windspeed	event
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 [ ]:
```
df.loc['1/1/2017']
```

Out[ ]:
```
temperature      32
windspeed         6
event          Rain
Name: 1/1/2017, dtype: object
```

In [ ]:
```
# Now to reset the indexes to 0 to 5

df.reset_index(inplace = True)
df
```

Out[ ]:

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

In [ ]:
```

```