Ques 02: How do I select a `Series / Column` from a `Dataframe` ?

`DataFrames` and `Series` are the two main `data structures` in pandas for data storage:

> A **DataFrame** is like a table, and each column of the table is called a **Series**. We will often select a Series in order to analyze or manipulate it.

In [2]:
```
import pandas as pd
```

In [3]:
```
# UFO dataset
# read_csv( ) is used for 'comma' seperated file

ufo_ds = pd.read_csv('http://bit.ly/uforeports')  # ufo = pd.read_table('http://bit.ly/uforeports', sep = ',')
ufo_ds.head()
```

Out[3]:

	City	Colors Reported	Shape Reported	State	Time
0	Ithaca	NaN	TRIANGLE	NY	6/1/1930 22:00
1	Willingboro	NaN	OTHER	NJ	6/30/1930 20:00
2	Holyoke	NaN	OVAL	CO	2/15/1931 14:00
3	Abilene	NaN	DISK	KS	6/1/1931 13:00
4	New York Worlds Fair	NaN	LIGHT	NY	4/18/1933 19:00

In [4]:
```
type(ufo_ds)
```

Out[4]:
```
pandas.core.frame.DataFrame
```

In [5]:
```
# Selecting 'City' column (a series)
# Column name is case-sensitive

ufo_ds['City'].head()
```

Out[5]:
```
0                 Ithaca
1            Willingboro
2                Holyoke
3                Abilene
4    New York Worlds Fair
Name: City, dtype: object
```

In [6]:
```
type(ufo_ds['City'])
```

Out[6]:
```
pandas.core.series.Series
```

In [7]:
```
# Each name becomes an attribute of the dataframe.
# So, we can use dot notation which is more easy but there is a limitation.
# The column name must not be a "keyword", and there should be "no space" in column name.

ufo_ds.City.head()
```

Out[7]:
```
0                 Ithaca
1            Willingboro
2                Holyoke
3                Abilene
4    New York Worlds Fair
Name: City, dtype: object
```

`Dot notation` does not always work but `Bracket notation` always works. If we want to stick to 'Dot' notation, then we have to rename all the columns so that there is no space in the columns name and no name is a keyword (built-in method or attribute) of python.

We can add two series (columns) if they are strings.

Creating a new column by combining two columns.

In [8]:
```
# We can create a new series/colum

ufo_ds['Location'] = ufo_ds.City + ' , ' + ufo_ds.State   # Need Bracket notation for the new column

ufo_ds.head()
```

Out[8]:

	City	Colors Reported	Shape Reported	State	Time	Location
0	Ithaca	NaN	TRIANGLE	NY	6/1/1930 22:00	Ithaca , NY
1	Willingboro	NaN	OTHER	NJ	6/30/1930 20:00	Willingboro , NJ
2	Holyoke	NaN	OVAL	CO	2/15/1931 14:00	Holyoke , CO
3	Abilene	NaN	DISK	KS	6/1/1931 13:00	Abilene , KS
4	New York Worlds Fair	NaN	LIGHT	NY	4/18/1933 19:00	New York Worlds Fair , NY