```
In [2]: pip install pyspark
```

```
6
    - ------------------------------------ 13.5/317.0 MB 3.1 MB/s eta 0:01:3
8
    - ------------------------------------ 13.6/317.0 MB 3.1 MB/s eta 0:01:3
8
    - ------------------------------------ 13.8/317.0 MB 3.1 MB/s eta 0:01:3
8
    - ------------------------------------ 13.9/317.0 MB 3.1 MB/s eta 0:01:3
7
    - ------------------------------------ 14.1/317.0 MB 3.1 MB/s eta 0:01:3
7
    - ------------------------------------ 14.3/317.0 MB 3.2 MB/s eta 0:01:3
5
    - ------------------------------------ 14.4/317.0 MB 3.2 MB/s eta 0:01:3
4
    - ------------------------------------ 14.5/317.0 MB 3.2 MB/s eta 0:01:3
4
    - ------------------------------------ 14.7/317.0 MB 3.2 MB/s eta 0:01:3
4
    - ------------------------------------ 14.8/317.0 MB 3.2 MB/s eta 0:01:3
```

```
In [3]: import pyspark
```

```
In [5]: from pyspark.sql import SparkSession
        spark=SparkSession.builder.appName("Practise").getOrCreate()
```

```
In [6]: spark
```

Out[6]: **SparkSession - in-memory**

SparkContext

Spark UI (http://Arnak:4040)

Version
 v3.5.1
Master
 local[*]
AppName
 Practise

```
In [8]: master=spark.read.csv(r"D:\NIT\FEBRUARY\26 feb (time series and spark)\sample.csv"
```

```
In [9]: master
```

Out[9]: DataFrame[_c0: string, _c1: string]

```
In [10]:  master.show()
```

```
+------+---+
|   _c0|_c1|
+------+---+
|  Name |Age|
|Michel| 30|
|  John| 32|
|   Max| 33|
+------+---+
```

```
In [11]:  master1=spark.read.option("header","true").csv(r"D:\NIT\FEBRUARY\26 feb (time seri
```

```
In [13]:  master1
```

```
Out[13]:  DataFrame[Name : string, Age: string]
```

```
In [12]:  master1.show()
```

```
+------+---+
| Name |Age|
+------+---+
|Michel| 30|
|  John| 32|
|   Max| 33|
+------+---+
```

```
In [15]:  type(master1)
```

```
Out[15]:  pyspark.sql.dataframe.DataFrame
```

```
In [16]:  master1.head()
```

```
Out[16]:  Row(Name ='Michel', Age='30')
```

```
In [17]:  master1.head(2)
```

```
Out[17]:  [Row(Name ='Michel', Age='30'), Row(Name ='John', Age='32')]
```

```
In [18]:  master1.head(3)
```

```
Out[18]:  [Row(Name ='Michel', Age='30'),
           Row(Name ='John', Age='32'),
           Row(Name ='Max', Age='33')]
```

```
In [19]: master1.printSchema()
```

```
root
 |-- Name : string (nullable = true)
 |-- Age: string (nullable = true)
```

```
In [ ]:
```