Pandas

By Athresh Kumar Labde

Python programming language for data manipulation and analysis.

In [1]:
```
import pandas as pd
import numpy as np
```

In [2]:
```
df = pd.DataFrame()
```

In [3]:
```
df
```

Out[3]:
—

In [6]:
```
df1 = pd.DataFrame([[2,3,1,3],[2,5,3,2],[6,7,4,7],[6,3,5,8]])
```

In [7]:
```
df1
```

Out[7]:

	0	1	2	3
0	2	3	1	3
1	2	5	3	2
2	6	7	4	7
3	6	3	5	8

In [8]:
```
df1.head(2)
```

Out[8]:

	0	1	2	3
0	2	3	1	3
1	2	5	3	2

In [9]:
```
df1.tail(2)
```

Out[9]:

	0	1	2	3
2	6	7	4	7
3	6	3	5	8

In [10]:
```
df1.shape
```

Out[10]:
```
(4, 4)
```

In [11]:
```
df1.iloc[0:2,0:2]
```

Out[11]:

	0	1
0	2	3
1	2	5

In [14]:
```
df1 = pd.DataFrame([[2,3,1,3],[2,5,3,2],[6,7,4,7],[6,3,5,8]],columns = ['A','B','C','D'])
df1
```

Out[14]:

	A	B	C	D
0	2	3	1	3
1	2	5	3	2
2	6	7	4	7
3	6	3	5	8

In [15]:
```
file = pd.read_csv('data.csv')
```

In [16]:
```
file
```

Out[16]:

	Symbol	Name	Sector
0	MMM	3M	Industrials
1	AOS	A. O. Smith	Industrials
2	ABT	Abbott Laboratories	Health Care
3	ABBV	AbbVie	Health Care
4	ABMD	Abiomed	Health Care
...	...	...	...
500	YUM	Yum! Brands	Consumer Discretionary
501	ZBRA	Zebra Technologies	Information Technology
502	ZBH	Zimmer Biomet	Health Care
503	ZION	Zions Bancorp	Financials
504	ZTS	Zoetis	Health Care

505 rows × 3 columns

In [18]:
```
file.head()
```

Out[18]:

	Symbol	Name	Sector
0	MMM	3M	Industrials
1	AOS	A. O. Smith	Industrials
2	ABT	Abbott Laboratories	Health Care
3	ABBV	AbbVie	Health Care
4	ABMD	Abiomed	Health Care

In [25]:
```
print(file['Symbol'])
```

```
0       MMM
1       AOS
2       ABT
3      ABBV
4      ABMD
       ...
500     YUM
501    ZBRA
502     ZBH
503    ZION
504     ZTS
Name: Symbol, Length: 505, dtype: object
```

Export csv

In [26]:
```
data = pd.DataFrame([[2,3,1,3],[2,5,3,2],[6,7,4,7],[6,3,5,8]],columns = ['A','B','C','D'])
```

In [27]:
```
data.head
```

Out[27]:
```
<bound method NDFrame.head of    A  B  C  D
0  2  3  1  3
1  2  5  3  2
2  6  7  4  7
3  6  3  5  8>
```

In [29]:
```
data.to_csv('export.csv')
```

Dataframe Merging

In [30]:
```
a = pd.DataFrame([[2,3,1,3],[2,5,3,2],[6,7,4,7],[6,3,5,8]],columns = ['A','B','C','D'])
```

In [31]:
```
a
```

Out[31]:

	A	B	C	D
0	2	3	1	3
1	2	5	3	2
2	6	7	4	7
3	6	3	5	8

In [32]:
```
b = pd.DataFrame([[32,3,11,3],[26,75,93,2],[6,27,4,97],[63,32,54,88]],columns = ['M','N','O','P'])
```

In [33]:
```
b
```

Out[33]:

	M	N	O	P
0	32	3	11	3
1	26	75	93	2
2	6	27	4	97
3	63	32	54	88

In [38]:
```
c = pd.merge(a,b,right_on='N',left_on='B')
```

In [39]:
```
c
```

Out[39]:

	A	B	C	D	M	N	O	P
0	2	3	1	3	32	3	11	3
1	6	3	5	8	32	3	11	3