

# Reading in Files

In [6]: `import pandas as pd`

In [9]: `df = pd.read_csv(r"D:\COURSES\YOUTUBE\ALEX THE ANALYST\PYTHON\Web scraping comp  
df`

	Rank	Name	Industry	(USD millions)	Revenue growth	Employees	Headquarters
0	1	Walmart	Retail	611,289	6.7%	2,100,000	Bentonville, Arkansas
1	2	Amazon	Retail and cloud computing	513,983	9.4%	1,540,000	Seattle, Washington
2	3	ExxonMobil	Petroleum industry	413,680	44.8%	62,000	Spring, Texas
3	4	Apple	Electronics industry	394,328	7.8%	164,000	Cupertino, California
4	5	UnitedHealth Group	Healthcare	324,162	12.7%	400,000	Minnetonka, Minnesota
...	...	...	...	...	...	...	...
95	96	Best Buy	Retail	46,298	10.6%	71,100	Richfield, Minnesota
96	97	Bristol-Myers Squibb	Pharmaceutical industry	46,159	0.5%	34,300	New York City, New York

In [ ]: `#df = pd.read_table(r"D:\COURSES\YOUTUBE\ALEX THE ANALYST\PYTHON\Web scraping c  
# for text files  
# pd.read_excel for excel files and for multiple sheets add a comma after the p  
#otherwise its gonna read the first one by default.  
  
#To read all the rows/columns  
#pd.set_option('display.max.rows/columns', 235/38)`

In [10]: *#to get details about the data*

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Rank                                  100 non-null    int64
1   Name                                  100 non-null    object
2   Industry                              100 non-null    object
3   Revenue (USD millions)                100 non-null    object
4   Revenue growth                        100 non-null    object
5   Employees                             100 non-null    object
6   Headquarters                          100 non-null    object
dtypes: int64(1), object(6)
memory usage: 5.6+ KB
```

In [11]: `df.shape`

Out[11]: (100, 7)

In [13]: `df.head(10)`

Out[13]:

	Rank	Name	Industry	Revenue (USD millions)	Revenue growth	Employees	Headquarters
0	1	Walmart	Retail	611,289	6.7%	2,100,000	Bentonville, Arkansas
1	2	Amazon	Retail and cloud computing	513,983	9.4%	1,540,000	Seattle, Washington
2	3	ExxonMobil	Petroleum industry	413,680	44.8%	62,000	Spring, Texas
3	4	Apple	Electronics industry	394,328	7.8%	164,000	Cupertino, California
4	5	UnitedHealth Group	Healthcare	324,162	12.7%	400,000	Minnetonka, Minnesota
5	6	CVS Health	Healthcare	322,467	10.4%	259,500	Woonsocket, Rhode Island
6	7	Berkshire Hathaway	Conglomerate	302,089	9.4%	383,000	Omaha, Nebraska
7	8	Alphabet	Technology and cloud computing	282,836	9.8%	156,000	Mountain View, California
8	9	McKesson Corporation	Health	276,711	4.8%	48,500	Irving, Texas
9	10	Chevron Corporation	Petroleum industry	246,252	51.6%	43,846	San Ramon, California

In [14]: `df.tail(10)`

Out[14]:

	Rank	Name	Industry	Revenue (USD millions)	Revenue growth	Employees	Headquarters
90	91	American Airlines	Airline	48,971	63.9%	129,700	Fort Worth, Texas
91	92	CHS	Agriculture cooperative	47,194	24.3%	10,014	Inver Grove Heights, Minnesota
92	93	Performance Food Group	Food processing	47,194	61.6%	34,825	Richmond, Virginia
93	94	PBF Energy	Petroleum industry	46,830	71.8%	3,616	Parsippany–Troy Hills, New Jersey
94	95	Nike	Apparel	46,710	4.9%	79,100	Beaverton, Oregon
95	96	Best Buy	Retail	46,298	10.6%	71,100	Richfield, Minnesota
96	97	Bristol-Myers Squibb	Pharmaceutical industry	46,159	0.5%	34,300	New York City, New York
97	98	United Airlines	Airline	44,955	82.5%	92,795	Chicago, Illinois
98	99	Thermo Fisher Scientific	Laboratory instruments	44,915	14.5%	130,000	Waltham, Massachusetts
99	100	Qualcomm	Technology	44,200	31.7%	51,000	San Diego, California

In [15]: `df['Rank']`

Out[15]:

```

0      1
1      2
2      3
3      4
4      5
...
95     96
96     97
97     98
98     99
99    100
Name: Rank, Length: 100, dtype: int64

```

In [17]: `df.loc[95]`

Out[17]:

```

Rank      96
Name      Best Buy
Industry   Retail
Revenue (USD millions)  46,298
Revenue growth      10.6%
Employees      71,100
Headquarters      Richfield, Minnesota
Name: 95, dtype: object

```

```
In [19]: df.iloc[95]
```

```
Out[19]: Rank          96  
Name          Best Buy  
Industry       Retail  
Revenue (USD millions)  46,298  
Revenue growth    10.6%  
Employees        71,100  
Headquarters    Richfield, Minnesota  
Name: 95, dtype: object
```

```
In [ ]:
```