In [1]:

```
# Pandas
import numpy as np
import pandas as pd
```

In [3]:

```
# import data set

df = pd.read_csv('sets/data.csv')
df
```

Out[3]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [4]:

```
# pandas methods
# head()
df.head()   # default value is 5, it will read first 5 records
```

Out[4]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%

In [5]:

```
# pandas methods
# head()
df.head(2)   # it will read  2  rows
```

Out[5]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%

In [6]:

```
# pandas methods
# head()
df.head(7)   # it will read  7  rows
```

Out[6]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%

In [ ]:

```

```

In [7]:

```
# pandas methods
# tail()
df.tail()   # last 5 records
```

Out[7]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
5	6	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

```

```

In [8]:

```
# pandas methods
# tail()
df.tail(3)   # last 3 records
```

Out[8]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
7	8	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

```

```

In [9]:

```
# pandas methods
# dtype()   to change any column data type
df = pd.read_csv('sets/data.csv')
df
```

Out[9]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [10]:

```
df = pd.read_csv('sets/data.csv',dtype={'ID':'float64'})
df
```

Out[10]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1.0	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2.0	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3.0	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4.0	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5.0	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6.0	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7.0	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8.0	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9.0	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10.0	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

```

```

In [11]:

```
df = pd.read_csv('sets/data.csv',dtype={'Inception':'float64'})
df
```

Out[11]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010.0	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012.0	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013.0	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009.0	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006.0	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010.0	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	Health	2009.0	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

```

```

In [14]:

```
df = pd.read_csv('sets/data.csv')
df
```

Out[14]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [12]:

```
df = pd.read_csv('sets/data.csv',dtype={'ID':'float64','Profit':'float64'})
df
```

Out[12]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1.0	Lamtone	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2.0	Stripfind	Financial Services	2010	$12,329,371	916,455 Dollars	11412916.0	20%
2	3.0	Canecorporation	Health	2012	$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4.0	Mattouch	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557.0	26%
4	5.0	Techdrill	Health	2009	$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6.0	Techline	Health	2006	$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7.0	Cityace	Health	2010	$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8.0	Kayelectronics	Health	2009	$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9.0	Ganzlax	IT Services	2011	$14,001,180	916,455 Dollars	11901180.0	18%
9	10.0	Trantraxlax	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060.0	7%

In [ ]:

```

```

In [15]:

```
# Pandas Method => True Values
df = pd.read_csv('sets/data.csv')
df
```

Out[15]:

	ID	Name	Status	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	yes	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	yes	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	yes	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	yes	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	yes	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	yes	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	yes	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	yes	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	yes	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	yes	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [16]:

```
# convert yes into true
df = pd.read_csv('sets/data.csv',true_values=['yes'])
df
```

Out[16]:

	ID	Name	Status	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	True	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553	30%
1	2	Stripfind	True	Financial Services	2010	$12,329,371	916,455 Dollars	11412916	20%
2	3	Canecorporation	True	Health	2012	$10,597,009	7,591,189 Dollars	3005820	7%
3	4	Mattouch	True	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557	26%
4	5	Techdrill	True	Health	2009	$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	True	Health	2006	$13,898,119	5,470,303 Dollars	8427816	23%
6	7	Cityace	True	Health	2010	$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	True	Health	2009	$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	True	IT Services	2011	$14,001,180	916,455 Dollars	11901180	18%
9	10	Trantraxlax	True	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [17]:

```
# Pandas Method => False Values
df = pd.read_csv('sets/data.csv')
df
```

Out[17]:

	ID	Name	Smoke	Status	Industry	Inception	Revenue	Expenses	Profit
0	1	Lamtone	no	yes	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553
1	2	Stripfind	no	yes	Financial Services	2010	$12,329,371	916,455 Dollars	11412916
2	3	Canecorporation	no	yes	Health	2012	$10,597,009	7,591,189 Dollars	3005820
3	4	Mattouch	no	yes	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557
4	5	Techdrill	no	yes	Health	2009	$10,573,990	7,435,363 Dollars	3138627
5	6	Techline	no	yes	Health	2006	$13,898,119	5,470,303 Dollars	8427816
6	7	Cityace	no	yes	Health	2010	$9,254,614	6,249,498 Dollars	3005116
7	8	Kayelectronics	no	yes	Health	2009	$9,451,943	3,878,113 Dollars	5573830
8	9	Ganzlax	no	yes	IT Services	2011	$14,001,180	916,455 Dollars	11901180
9	10	Trantraxlax	no	yes	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060

In [18]:

```
# Pandas Method => False Values ,True Values
df = pd.read_csv('sets/data.csv',true_values=['yes'],false_values=['no'])
df
```

Out[18]:

	ID	Name	Smoke	Status	Industry	Inception	Revenue	Expenses	Profit
0	1	Lamtone	False	True	IT Services	2009	$11,757,018	6,482,465 Dollars	5274553
1	2	Stripfind	False	True	Financial Services	2010	$12,329,371	916,455 Dollars	11412916
2	3	Canecorporation	False	True	Health	2012	$10,597,009	7,591,189 Dollars	3005820
3	4	Mattouch	False	True	IT Services	2013	$14,026,934	7,429,377 Dollars	6597557
4	5	Techdrill	False	True	Health	2009	$10,573,990	7,435,363 Dollars	3138627
5	6	Techline	False	True	Health	2006	$13,898,119	5,470,303 Dollars	8427816
6	7	Cityace	False	True	Health	2010	$9,254,614	6,249,498 Dollars	3005116
7	8	Kayelectronics	False	True	Health	2009	$9,451,943	3,878,113 Dollars	5573830
8	9	Ganzlax	False	True	IT Services	2011	$14,001,180	916,455 Dollars	11901180
9	10	Trantraxlax	False	True	Government Services	2011	$11,088,336	5,635,276 Dollars	5453060

In [ ]:

```

```

In [20]:

```
df.info
```

. . .

In [21]:

```
df.describe
```

. . .