

In [1]:

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

In [2]:

```
# read data set
df = pd.read_csv("sets/data.csv")
df
```

Out[2]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [3]:

```
# DROP-NA => drop rows or column based axis where it find NaN value
# By Deafult df.dropna(axis=0) which means based on rows

df.dropna()
```

Out[3]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [4]:

```
df
```

Out[4]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [5]:

```
df.dropna(axis=1)
```

Out[5]:

	ID	Name	Revenue	Profit
0	1	Lamtone	\$11,757,018	5274553
1	2	Stripfind	\$12,329,371	11412916
2	3	Canecorporation	\$10,597,009	3005820
3	4	Mattouch	\$14,026,934	6597557
4	5	Techdrill	\$10,573,990	3138627
5	6	Techline	\$13,898,119	8427816
6	7	Cityace	\$9,254,614	3005116
7	8	Kayelectronics	\$9,451,943	5573830
8	9	Ganzlax	\$14,001,180	11901180
9	10	Trantraxlax	\$11,088,336	5453060

In [ ]:

In [6]:

```
# how='any' or how='all'
df.dropna(how='any')
```

Out[6]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [7]:

```
df.dropna()
```

Out[7]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [8]:

```
# removed based on column
df.dropna(axis=1,how='any')
```

Out[8]:

	ID	Name	Revenue	Profit
0	1	Lamtone	\$11,757,018	5274553
1	2	Stripfind	\$12,329,371	11412916
2	3	Canecorporation	\$10,597,009	3005820
3	4	Mattouch	\$14,026,934	6597557
4	5	Techdrill	\$10,573,990	3138627
5	6	Techline	\$13,898,119	8427816
6	7	Cityace	\$9,254,614	3005116
7	8	Kayelectronics	\$9,451,943	5573830
8	9	Ganzlax	\$14,001,180	11901180
9	10	Trantraxlax	\$11,088,336	5453060

In [ ]:

In [9]:

```
# All => if all rows contain null or NAN value then only it will drop
df.dropna(how='all')
```

Out[9]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [10]:

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

Out[10]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1.0	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2.0	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3.0	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4.0	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5.0	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6.0	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7.0	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8.0	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	10.0	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [11]:

```
df.dropna(how='all')
```

Out[11]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1.0	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2.0	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3.0	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4.0	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5.0	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6.0	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7.0	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8.0	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
9	10.0	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [ ]:

In [12]:

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

Out[12]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%



In [13]:

```
df.dropna(how='all')
```

Out[13]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [ ]:

In [ ]:

In [14]:

```
# read Data Set
df = pd.read_csv("sets/data.csv")
df
```

Out[14]:

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

In [15]:

```
df.dropna(how='all')
```

Out[15]:

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

In [16]:

```
df.dropna(axis=1,how='all')
```

Out[16]:

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

In [ ]:

In [ ]:

In [ ]:

In [17]:

```
# read Data Set
df = pd.read_csv("sets/data.csv")
df
```

Out[17]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	NaN	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	NaN	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	NaN	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	NaN	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	NaN	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	NaN	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	NaN	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	NaN	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	NaN	7%

In [18]:

```
df.dropna(axis=1,how='all')
```

Out[18]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	NaN	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	NaN	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	NaN	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	NaN	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	NaN	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	NaN	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	NaN	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	NaN	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	NaN	7%

In [ ]:

In [ ]:

In [19]:

```
# read Data Set
df = pd.read_csv("sets/data.csv")
df
```

Out[19]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [20]:

```
# thresh=None
df.dropna()
```

Out[20]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [21]:

```
# thresh=None
# if any row contains atleast one not null values then it will not drop
df.dropna(thresh=1)
```

Out[21]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [ ]:

In [22]:

```
# read Data Set
df = pd.read_csv("sets/data.csv")
df
```

Out[22]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9	Ganzlax	NaN	NaN	NaN	NaN	NaN	NaN
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%



In [23]:

```
# thresh=None
# if any row contains atleast one not null values then it will not drop
df.dropna(thresh=1)
```

Out[23]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9	Ganzlax	NaN	NaN	NaN	NaN	NaN	NaN
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [28]:

```
# thresh=None  
# if any row contains atleast two not null values then it will not drop  
df.dropna(axis=0,thresh=2)
```

Out[28]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
8	9	Ganzlax	NaN	NaN	NaN	NaN	NaN	NaN
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [29]:

```
# thresh=None
# if any row contains atleast two not null values then it will not drop
df.dropna(axis=0,thresh=3)
```

Out[29]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
3	4	Mattouch	IT Services	2013.0	\$14,026,934	7,429,377 Dollars	6597557.0	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627.0	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
6	7	Cityace	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116.0	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830.0	4%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [30]:

```
# thresh=None
# if any row contains atleast two not null values then it will not drop
df.dropna(axis=0,thresh=8)
```

Out[30]:

	ID	Name	Industry	Inception	Revenue	Expenses	Profit	Growth
0	1	Lamtone	IT Services	2009.0	\$11,757,018	6,482,465 Dollars	5274553.0	30%
1	2	Stripfind	Financial Services	2010.0	\$12,329,371	916,455 Dollars	11412916.0	20%
2	3	Canecorporation	Health	2012.0	\$10,597,009	7,591,189 Dollars	3005820.0	7%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816.0	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060.0	7%

In [ ]:

In [ ]:

In [31]:

```
# read Data Set
df = pd.read_csv("sets/data.csv")
df
```

Out[31]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [32]:

```
# subset=None
# SubSEt is used for particular column
df.dropna(subset=['Industry'])
```

Out[32]:

	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	NaN
4	5	Techdrill	Health	NaN	\$10,573,990	7,435,363 Dollars	3138627	8%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [33]:

```
# subset=None
# SubSEt is used for particular column
df.dropna(subset=['Expenses'])
```

Out[33]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [34]:

```
df
```

Out[34]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [35]:

```
df.dropna()
```

Out[35]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [36]:

df

Out[36]:

	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	NaN
4	5	Techdrill	Health	NaN	\$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	NaN	NaN	\$9,254,614	6,249,498 Dollars	3005116	6%
7	8	Kayelectronics	NaN	2009.0	\$9,451,943	3,878,113 Dollars	5573830	4%
8	9	Ganzlax	IT Services	2011.0	\$14,001,180	NaN	11901180	18%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]:

In [37]:

```
# inplace=False
df.dropna(inplace=True)
```



In [38]:

df

Out[38]:

	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%
5	6	Techline	Health	2006.0	\$13,898,119	5,470,303 Dollars	8427816	23%
9	10	Trantraxlax	Government Services	2011.0	\$11,088,336	5,635,276 Dollars	5453060	7%

In [ ]: