In [181]:
```
import pandas as pd
```

In [182]:
```
#read csv file
df = pd.read_csv('fortune500.csv', header=0, index_col=0);
```

In [183]:
```
# get top 5 item from the frame
display = df.head(5);
display
# print("top 5 records");
```

Out[183]:

company	rank	revenues	revenue_change	profits	assets	profit_change	ceo	industry	sector	p
Walmart	1	485873	0.8	13643.0	198825	-7.2	C. Douglas McMillon	General Merchandisers	Retailing	
State Grid	2	315199	-4.4	9571.3	489838	-6.2	Kou Wei	Utilities	Energy	
Sinopec Group	3	267518	-9.1	1257.9	310726	-65.0	Wang Yupu	Petroleum Refining	Energy	
China National Petroleum	4	262573	-12.3	1867.5	585619	-73.7	Zhang Jianhua	Petroleum Refining	Energy	
Toyota Motor	5	254694	7.7	16899.3	437575	-12.3	Akio Toyoda	Motor Vehicles and Parts	Motor Vehicles & Parts	

In [184]:
```
# Display tail record
display  = df.tail()
display
# print("tail records")
```

Out[184]:

company	rank	revenues	revenue_change	profits	assets	profit_change	ceo	industry	s
Teva Pharmaceutical Industries	496	21903	11.5	329.0	92890	-79.3	Yitzhak Peterburg	Pharmaceuticals	H
New China Life Insurance	497	21796	-13.3	743.9	100609	-45.6	Wan Feng	Insurance: Life, Health (stock)	Fina
Wm. Morrison Supermarkets	498	21741	-11.3	406.4	11630	20.4	David T. Potts	Food and Drug Stores	F S
TUI	499	21655	-5.5	1151.7	16247	195.5	Friedrich Joussen	Travel Services	Bus Se
AutoNation	500	21609	3.6	430.5	10060	-2.7	Michael J. Jackson	Specialty Retailers	Re

In [185]:
```
# Display Datatype/schema infor
df.dtypes
# display
# print("Print the Datatype/schema infor")
```

Out[185]:
```
rank                          int64
revenues                      int64
revenue_change              float64
profits                     float64
assets                        int64
profit_change               float64
ceo                          object
industry                     object
sector                       object
previous_rank                 int64
country                      object
location                     object
website                      object
years_on_global_500_list      int64
employees                     int64
total_stockholder_equity      int64
dtype: object
```

In [186]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 500 entries, Walmart to AutoNation
Data columns (total 16 columns):
rank                          500 non-null int64
revenues                      500 non-null int64
revenue_change                498 non-null float64
profits                       499 non-null float64
assets                        500 non-null int64
profit_change                 436 non-null float64
ceo                           500 non-null object
industry                      500 non-null object
sector                        500 non-null object
previous_rank                 500 non-null int64
country                       500 non-null object
location                      500 non-null object
website                       500 non-null object
years_on_global_500_list      500 non-null int64
employees                     500 non-null int64
total_stockholder_equity      500 non-null int64
dtypes: float64(3), int64(7), object(6)
memory usage: 66.4+ KB
```

In [187]:
```
#Describe variation in revenues and employees (both) for min, max, mean etc
display = df.loc[:, ['revenues', 'employees']].describe();
display
# df
# get top 5 item from the frame
# display = sortrankdf.head(5);
```

Out[187]:

	revenues	employees
count	500.000000	5.000000e+02
mean	55416.358000	1.339983e+05
std	45725.478963	1.700878e+05
min	21609.000000	3.280000e+02
25%	29003.000000	4.293250e+04
50%	40236.000000	9.291050e+04
75%	63926.750000	1.689172e+05
max	485873.000000	2.300000e+06

In [195]:
```
#Fetch Top 5 companies with previous year rank and years on global 500 list
sortprevdf = df.sort_values(by='rank');
display = sortprevdf[sortprevdf['rank'].isin([1,2,3,4,5])]
display1 = display.loc[:, ['rank','previous_rank','years_on_global_500_list']]
display1.head(5)
```

Out[195]:

company	rank	previous_rank	years_on_global_500_list
Walmart	1	1	23
State Grid	2	2	17
Sinopec Group	3	4	19
China National Petroleum	4	3	17
Toyota Motor	5	8	23

In [209]:
```
#Display Count of companies by location in USA only
display = df[df['country']=='USA'].groupby('location').size();
display1 =  display.to_frame(name = 'size');
display2 = display1.sort_values(by=['size'],ascending=False);
display2
```

Out[209]:

location	size
New York, NY	15
Houston, TX	5
Chicago, IL	4
Atlanta, GA	4
Cincinnati, OH	3
San Antonio, TX	3
McLean, VA	2
Deerfield, IL	2
Falls Church, VA	2
Charlotte, NC	2
St. Louis, MO	2
Dallas, TX	2
Palo Alto, CA	2
Boston, MA	2
Washington, DC	2
Minneapolis, MN	2
San Francisco, CA	2
Peoria, IL	1
Philadelphia, PA	1
Omaha, NE	1
Armonk, NY	1
Oak Brook, IL	1
Northbrook, IL	1
North Chicago, IL	1
Newark, NJ	1
Pittsburgh, PA	1
New Brunswick, NJ	1
Nashville, TN	1
Mountain View, CA	1
Morris Plains, NJ	1
...	...
Centennial, CO	1
Camp Hill, PA	1
Burbank, CA	1

location	size
Boise, ID	1
Bloomington, IL	1
Bloomfield, CT	1
Bethesda, MD	1
Bentonville, AR	1
Beaverton, OR	1
Farmington, CT	1
Findlay, OH	1
Fort Lauderdale, FL	1
Issaquah, WA	1
Midland, MI	1
Miami, FL	1
Memphis, TN	1
Mayfield Village, OH	1
Louisville, KY	1
Lakeland, FL	1
Kenilworth, NJ	1
Irving, TX	1
Fort Worth, TX	1
Inver Grove Heights, MN	1
Indianapolis, IN	1
Hoffman Estates, IL	1
Hartford, CT	1
Goodlettsville, TN	1
Framingham, MA	1
Foster City, CA	1
Woonsocket, RI	1

93 rows × 1 columns