

session-1

August 24, 2024

1 Pandas

```
[3]: import pandas as pd
import csv
```

```
[5]: dataset = pd.read_csv('companies.csv')
dataset
```

```
[5]:
```

		company	rank	revenues	revenue_change	profits	\
0		Walmart	1	485873	0.8	13643.0	
1		State Grid	2	315199	-4.4	9571.3	
2		Sinopec Group	3	267518	-9.1	1257.9	
3		China National Petroleum	4	262573	-12.3	1867.5	
4		Toyota Motor	5	254694	7.7	16899.3	
..		
495	Teva Pharmaceutical Industries		496	21903	11.5	329.0	
496	New China Life Insurance		497	21796	-13.3	743.9	
497	Wm. Morrison Supermarkets		498	21741	-11.3	406.4	
498		TUI	499	21655	-5.5	1151.7	
499		AutoNation	500	21609	3.6	430.5	

	assets	profit_change		ceo	\
0	198825	-7.2	C. Douglas	McMillon	
1	489838	-6.2		Kou Wei	
2	310726	-65.0		Wang Yupu	
3	585619	-73.7		Zhang Jianhua	
4	437575	-12.3		Akio Toyoda	
..	
495	92890	-79.3	Yitzhak	Peterburg	
496	100609	-45.6		Wan Feng	
497	11630	20.4	David T.	Potts	
498	16247	195.5	Friedrich	Joussen	
499	10060	-2.7	Michael J.	Jackson	

		industry		sector	previous_rank	\
0		General Merchandisers		Retailing	1	
1		Utilities		Energy	2	

2	Petroleum Refining	Energy	4
3	Petroleum Refining	Energy	3
4	Motor Vehicles and Parts	Motor Vehicles & Parts	8
..
495	Pharmaceuticals	Health Care	0
496	Insurance: Life, Health (stock)	Financials	427
497	Food and Drug Stores	Food & Drug Stores	437
498	Travel Services	Business Services	467
499	Specialty Retailers	Retailing	0

	country	hq_location	website \
0	USA	Bentonville, AR	http://www.walmart.com
1	China	Beijing, China	http://www.sgcc.com.cn
2	China	Beijing, China	http://www.sinopec.com
3	China	Beijing, China	http://www.cnpc.com.cn
4	Japan	Toyota, Japan	http://www.toyota-global.com
..
495	Israel	Petach Tikva, Israel	http://www.tevapharm.com
496	China	Beijing, China	http://www.newchinalife.com
497	Britain	Bradford, Britain	http://www.morrisons.com
498	Germany	Hanover, Germany	http://www.tuigroup.com
499	USA	Fort Lauderdale, FL	http://www.autonation.com

	years_on_global_500_list	employees	total_stockholder_equity
0	23	2300000	77798
1	17	926067	209456
2	19	713288	106523
3	17	1512048	301893
4	23	364445	157210
..
495	1	56960	33337
496	2	54378	8507
497	13	77210	5111
498	23	66779	3006
499	12	26000	2310

[500 rows x 17 columns]

2 Attributes vs Methods/Functions

- An attributes only gives information
- A function stored in an instance is called a method
- An attributes does not have round brackets

3 Shape

```
[6]: dataset.shape
```

```
[6]: (500, 17)
```

4 dtypes

```
[7]: type(dataset)
```

```
[7]: pandas.core.frame.DataFrame
```

```
[8]: dataset.dtypes
```

```
[8]: company          object
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
hq_location          object
website              object
years_on_global_500_list  int64
employees            int64
total_stockholder_equity  int64
dtype: object
```

5 Columns

```
[10]: dataset.columns
```

```
[10]: Index(['company', 'rank', 'revenues', 'revenue_change', 'profits', 'assets',
        'profit_change', 'ceo', 'industry', 'sector', 'previous_rank',
        'country', 'hq_location', 'website', 'years_on_global_500_list',
        'employees', 'total_stockholder_equity'],
        dtype='object')
```

6 Info()

```
[13]: dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   company               500 non-null    object
1   rank                  500 non-null    int64
2   revenues              500 non-null    int64
3   revenue_change        498 non-null    float64
4   profits               499 non-null    float64
5   assets               500 non-null    int64
6   profit_change         436 non-null    float64
7   ceo                   500 non-null    object
8   industry              500 non-null    object
9   sector                500 non-null    object
10  previous_rank          500 non-null    int64
11  country                500 non-null    object
12  hq_location            500 non-null    object
13  website                500 non-null    object
14  years_on_global_500_list 500 non-null    int64
15  employees              500 non-null    int64
16  total_stockholder_equity 500 non-null    int64
dtypes: float64(3), int64(7), object(7)
memory usage: 66.5+ KB
```

```
[14]: dataset.info(show_counts = False)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 17 columns):
#   Column                Dtype
---  -
0   company               object
1   rank                  int64
2   revenues              int64
3   revenue_change        float64
4   profits               float64
5   assets               int64
6   profit_change         float64
7   ceo                   object
8   industry              object
9   sector                object
10  previous_rank          int64
11  country                object
```

```

12 hq_location          object
13 website              object
14 years_on_global_500_list  int64
15 employees            int64
16 total_stockholder_equity int64
dtypes: float64(3), int64(7), object(7)
memory usage: 66.5+ KB

```

7 describe()

```
[15]: dataset.describe()
```

```

[15]:
count      rank      revenues  revenue_change      profits      assets \
count  500.000000      500.000000      498.000000      499.000000  5.000000e+02
mean    250.500000     55416.358000        4.538353     3055.203206  2.436323e+05
std     144.481833     45725.478963      28.549067     5171.981071  4.851937e+05
min        1.000000     21609.000000     -67.300000    -13038.000000  3.717000e+03
25%     125.750000     29003.000000      -5.900000      556.950000  3.658850e+04
50%     250.500000     40236.000000       0.550000     1761.600000  7.326150e+04
75%     375.250000     63926.750000       6.975000     3954.000000  1.805640e+05
max     500.000000     485873.000000      442.300000     45687.000000  3.473238e+06

      profit_change  previous_rank  years_on_global_500_list      employees \
count      436.000000      500.000000      500.000000  5.000000e+02
mean       24.152752      222.134000        15.036000  1.339983e+05
std       437.509566      146.941961       7.932752   1.700878e+05
min      -793.700000       0.000000       1.000000   3.280000e+02
25%      -22.775000      92.750000       7.000000   4.293250e+04
50%       -0.350000     219.500000      17.000000   9.291050e+04
75%       17.700000     347.250000      23.000000   1.689172e+05
max      8909.500000     500.000000      23.000000   2.300000e+06

      total_stockholder_equity
count          500.000000
mean          30628.076000
std          43642.576833
min         -59909.000000
25%          7553.750000
50%         15809.500000
75%          37828.500000
max          301893.000000

```

8 head()

```
[19]: dataset.head()
```

```
[19]:
```

	company	rank	revenues	revenue_change	profits	assets \
0	Walmart	1	485873	0.8	13643.0	198825
1	State Grid	2	315199	-4.4	9571.3	489838
2	Sinopec Group	3	267518	-9.1	1257.9	310726
3	China National Petroleum	4	262573	-12.3	1867.5	585619
4	Toyota Motor	5	254694	7.7	16899.3	437575

	profit_change	ceo	industry \
0	-7.2	C. Douglas McMillon	General Merchandisers
1	-6.2	Kou Wei	Utilities
2	-65.0	Wang Yupu	Petroleum Refining
3	-73.7	Zhang Jianhua	Petroleum Refining
4	-12.3	Akio Toyoda	Motor Vehicles and Parts

	sector	previous_rank	country	hq_location \
0	Retailing	1	USA	Bentonville, AR
1	Energy	2	China	Beijing, China
2	Energy	4	China	Beijing, China
3	Energy	3	China	Beijing, China
4	Motor Vehicles & Parts	8	Japan	Toyota, Japan

	website	years_on_global_500_list	employees \
0	http://www.walmart.com	23	2300000
1	http://www.sgcc.com.cn	17	926067
2	http://www.sinopec.com	19	713288
3	http://www.cnpc.com.cn	17	1512048
4	http://www.toyota-global.com	23	364445

	total_stockholder_equity
0	77798
1	209456
2	106523
3	301893
4	157210

```
[20]: dataset.head(50)
```

```
[20]:
```

	company	rank	revenues	revenue_change \
0	Walmart	1	485873	0.8
1	State Grid	2	315199	-4.4
2	Sinopec Group	3	267518	-9.1
3	China National Petroleum	4	262573	-12.3
4	Toyota Motor	5	254694	7.7

5	Volkswagen	6	240264	1.5
6	Royal Dutch Shell	7	240033	-11.8
7	Berkshire Hathaway	8	223604	6.1
8	Apple	9	215639	-7.7
9	Exxon Mobil	10	205004	-16.7
10	McKesson	11	198533	3.1
11	BP	12	186606	-17.4
12	UnitedHealth Group	13	184840	17.7
13	CVS Health	14	177526	15.8
14	Samsung Electronics	15	173957	-2.0
15	Glencore	16	173883	2.0
16	Daimler	17	169483	2.2
17	General Motors	18	166380	9.2
18	AT&T	19	163786	11.6
19	EXOR Group	20	154894	1.5
20	Ford Motor	21	151800	1.5
21	Industrial & Commercial Bank of China	22	147675	-11.7
22	AmerisourceBergen	23	146850	8.0
23	China State Construction Engineering	24	144505	3.1
24	AXA	25	143722	11.2
25	Amazon.com	26	135987	27.1
26	Hon Hai Precision Industry	27	135129	-4.3
27	China Construction Bank	28	135093	-8.7
28	Honda Motor	29	129198	6.2
29	Total	30	127925	-10.8
30	General Electric	31	126661	-9.8
31	Verizon	32	125980	-4.3
32	Japan Post Holdings	33	122990	3.6
33	Allianz	34	122196	-0.6
34	Cardinal Health	35	121546	18.5
35	Costco	36	118719	2.2
36	Walgreens Boots Alliance	37	117351	13.4
37	Agricultural Bank of China	38	117275	-12.1
38	Ping An Insurance	39	116581	5.7
39	Kroger	40	115337	5.0
40	SAIC Motor	41	113861	6.7
41	Bank of China	42	113708	-7.1
42	BNP Paribas	43	109026	-2.2
43	Nissan Motor	44	108164	6.5
44	Chevron	45	107567	-18.0
45	Fannie Mae	46	107162	-2.9
46	China Mobile Communications	47	107117	0.3
47	J.P. Morgan Chase	48	105486	4.4
48	Legal & General Group	49	105235	442.3
49	Nippon Telegraph & Telephone	50	105128	9.4

profits assets profit_change

ceo \

0	13643.0	198825	-7.2	C. Douglas McMillon
1	9571.3	489838	-6.2	Kou Wei
2	1257.9	310726	-65.0	Wang Yupu
3	1867.5	585619	-73.7	Zhang Jianhua
4	16899.3	437575	-12.3	Akio Toyoda
5	5937.3	432116	NaN	Matthias Muller
6	4575.0	411275	135.9	Ben van Beurden
7	24074.0	620854	NaN	Warren E. Buffett
8	45687.0	321686	-14.4	Timothy D. Cook
9	7840.0	330314	-51.5	Darren W. Woods
10	5070.0	60969	124.5	John H. Hammergren
11	115.0	263316	NaN	Robert W. Dudley
12	7017.0	122810	20.7	Stephen J. Hemsley
13	5317.0	94462	1.5	Larry J. Merlo
14	19316.5	217104	16.8	Oh-Hyun Kwon
15	1379.0	124600	NaN	Ivan Glasenberg
16	9428.4	256262	0.9	Dieter Zetsche
17	9427.0	221690	-2.7	Mary T. Barra
18	12976.0	403821	-2.8	Randall L. Stephenson
19	651.3	186172	-21.1	John Elkann
20	4596.0	237951	-37.7	James P. Hackett
21	41883.9	3473238	-5.0	Gu Shu
22	1427.9	33656	NaN	Steven H. Collis
23	2492.9	201269	10.7	Guan Qing
24	6446.0	941556	3.5	Thomas Buberl
25	2371.0	83402	297.8	Jeffrey P. Bezos
26	4608.8	80436	-0.4	Terry Gou
27	34840.9	3016578	-4.0	Wang Hongzhang
28	5690.3	170165	98.3	Takahiro Hachigo
29	6196.0	230978	21.8	Patrick Pouyanne
30	8831.0	365183	NaN	Jeffrey R. Immelt
31	13127.0	244180	-26.6	Lowell C. McAdam
32	-267.4	2631385	-107.5	Masatsugu Nagato
33	7611.5	932091	3.7	Oliver Bate
34	1427.0	34122	17.4	George S. Barrett
35	2350.0	33163	-1.1	W. Craig Jelinek
36	4173.0	72688	-1.1	Stefano Pessina
37	27687.8	2816039	-3.6	Zhao Huan
38	9392.0	802490	8.9	Ma Mingzhe
39	1975.0	36505	-3.1	W. Rodney McMullen
40	4818.2	84989	1.6	Chen Zhixin
41	24773.4	2611539	-8.9	Chen Siqing
42	8517.2	2190423	14.7	Jean-Laurent Bonnafe
43	6123.4	165344	40.3	Hiroto Saikawa
44	-497.0	260078	-110.8	John S. Watson
45	12313.0	3287968	12.4	Timothy J. Mayopoulos
46	9614.3	246446	-5.2	Li Yue

47	24733.0	2490972	1.2	James Dimon
48	1697.9	577954	3.4	Nigel Wilson
49	7384.4	190740	20.2	Hiroo Unoura

	industry	sector \
0	General Merchandisers	Retailing
1	Utilities	Energy
2	Petroleum Refining	Energy
3	Petroleum Refining	Energy
4	Motor Vehicles and Parts	Motor Vehicles & Parts
5	Motor Vehicles and Parts	Motor Vehicles & Parts
6	Petroleum Refining	Energy
7	Insurance: Property and Casualty (Stock)	Financials
8	Computers, Office Equipment	Technology
9	Petroleum Refining	Energy
10	Wholesalers: Health Care	Wholesalers
11	Petroleum Refining	Energy
12	Health Care: Insurance and Managed Care	Health Care
13	Health Care: Pharmacy and Other Services	Health Care
14	Electronics, Electrical Equip.	Technology
15	Mining, Crude-Oil Production	Energy
16	Motor Vehicles and Parts	Motor Vehicles & Parts
17	Motor Vehicles and Parts	Motor Vehicles & Parts
18	Telecommunications	Telecommunications
19	Diversified Financials	Financials
20	Motor Vehicles and Parts	Motor Vehicles & Parts
21	Banks: Commercial and Savings	Financials
22	Wholesalers: Health Care	Wholesalers
23	Engineering, Construction	Engineering & Construction
24	Insurance: Life, Health (stock)	Financials
25	Internet Services and Retailing	Technology
26	Electronics, Electrical Equip.	Technology
27	Banks: Commercial and Savings	Financials
28	Motor Vehicles and Parts	Motor Vehicles & Parts
29	Petroleum Refining	Energy
30	Industrial Machinery	Industrials
31	Telecommunications	Telecommunications
32	Insurance: Life, Health (stock)	Financials
33	Insurance: Property and Casualty (Stock)	Financials
34	Wholesalers: Health Care	Wholesalers
35	General Merchandisers	Retailing
36	Food and Drug Stores	Food & Drug Stores
37	Banks: Commercial and Savings	Financials
38	Insurance: Life, Health (stock)	Financials
39	Food and Drug Stores	Food & Drug Stores
40	Motor Vehicles and Parts	Motor Vehicles & Parts
41	Banks: Commercial and Savings	Financials

42	Banks: Commercial and Savings	Financials
43	Motor Vehicles and Parts	Motor Vehicles & Parts
44	Petroleum Refining	Energy
45	Diversified Financials	Financials
46	Telecommunications	Telecommunications
47	Banks: Commercial and Savings	Financials
48	Insurance: Life, Health (stock)	Financials
49	Telecommunications	Telecommunications

	previous_rank	country	hq_location \
0	1	USA	Bentonville, AR
1	2	China	Beijing, China
2	4	China	Beijing, China
3	3	China	Beijing, China
4	8	Japan	Toyota, Japan
5	7	Germany	Wolfsburg, Germany
6	5	Netherlands	The Hague, Netherlands
7	11	USA	Omaha, NE
8	9	USA	Cupertino, CA
9	6	USA	Irving, TX
10	12	USA	San Francisco, CA
11	10	Britain	London, Britain
12	17	USA	Minnetonka, MN
13	18	USA	Woonsocket, RI
14	13	South Korea	Suwon, South Korea
15	14	Switzerland	Baar, Switzerland
16	16	Germany	Stuttgart, Germany
17	20	USA	Detroit, MI
18	23	USA	Dallas, TX
19	19	Netherlands	Amsterdam, Netherlands
20	21	USA	Dearborn, MI
21	15	China	Beijing, China
22	28	USA	Chesterbrook, PA
23	27	China	Beijing, China
24	33	France	Paris, France
25	44	USA	Seattle, WA
26	25	Taiwan	New Taipei City, Taiwan
27	22	China	Beijing, China
28	36	Japan	Tokyo, Japan
29	24	France	Courbevoie, France
30	26	USA	Boston, MA
31	30	USA	New York, NY
32	37	Japan	Tokyo, Japan
33	34	Germany	Munich, Germany
34	50	USA	Dublin, OH
35	38	USA	Issaquah, WA
36	47	USA	Deerfield, IL

37	29	China	Beijing, China
38	41	China	Shenzhen, China
39	42	USA	Cincinnati, OH
40	46	China	Shanghai, China
41	35	China	Beijing, China
42	39	France	Paris, France
43	53	Japan	Yokohama, Japan
44	31	USA	San Ramon, CA
45	40	USA	Washington, DC
46	45	China	Beijing, China
47	55	USA	New York, NY
48	0	Britain	London, Britain
49	60	Japan	Tokyo, Japan

	website	years_on_global_500_list	\
0	http://www.walmart.com	23	
1	http://www.sgcc.com.cn	17	
2	http://www.sinopec.com	19	
3	http://www.cnpc.com.cn	17	
4	http://www.toyota-global.com	23	
5	http://www.volkswagen.com	23	
6	http://www.shell.com	23	
7	http://www.berkshirehathaway.com	21	
8	http://www.apple.com	15	
9	http://www.exxonmobil.com	23	
10	http://www.mckesson.com	23	
11	http://www.bp.com	23	
12	http://www.unitedhealthgroup.com	21	
13	http://www.cvshealth.com	22	
14	http://www.samsung.com	23	
15	http://www.glencore.com	7	
16	http://www.daimler.com	23	
17	http://www.gm.com	23	
18	http://www.att.com	23	
19	http://www.exor.com	7	
20	http://www.ford.com	23	
21	http://www.icbc-ltd.com	19	
22	http://www.amerisourcebergen.com	18	
23	http://www.cscec.com	6	
24	http://www.axa.com	23	
25	http://www.amazon.com	9	
26	http://www.foxconn.com	13	
27	http://www.ccb.com	18	
28	http://www.honda.com	23	
29	http://www.total.com	23	
30	http://www.ge.com	23	
31	http://www.verizon.com	23	

32	http://www.japanpost.jp	21
33	http://www.allianz.com	23
34	http://www.cardinalhealth.com	20
35	http://www.costco.com	23
36	http://www.walgreensbootsalliance.com	23
37	http://www.abchina.com	18
38	http://www.pingan.com	8
39	http://www.thekrogerco.com	23
40	http://www.saicmotor.com	6
41	http://www.boc.cn	23
42	http://www.bnpparibas.com	23
43	http://www.nissan-global.com	23
44	http://www.chevron.com	23
45	http://www.fanniemae.com	20
46	http://www.10086.cn	17
47	http://www.jpmorganchase.com	23
48	http://www.legalandgeneralgroup.com	17
49	http://www.ntt.co.jp	23

	employees	total_stockholder_equity
0	2300000	77798
1	926067	209456
2	713288	106523
3	1512048	301893
4	364445	157210
5	626715	97753
6	89000	186646
7	367700	283001
8	116000	128249
9	72700	167325
10	64500	11095
11	74500	95286
12	230000	38274
13	204000	36830
14	325000	154376
15	93123	44243
16	282488	61116
17	225000	43836
18	268540	123135
19	302562	11582
20	201000	29170
21	461749	283438
22	18500	2129
23	263915	15344
24	97707	74454
25	341400	19285
26	726772	33476

27	362482	226851
28	211915	65482
29	102168	98680
30	295000	75828
31	160900	22524
32	248384	91532
33	140253	71020
34	37300	6554
35	172000	12079
36	300000	29880
37	501368	189682
38	318588	55177
39	443000	6698
40	97582	27617
41	308900	203134
42	184839	106164
43	137250	50550
44	55200	145556
45	7000	6071
46	463712	130459
47	243355	254190
48	8939	8579
49	274844	81254

9 tail()

```
[22]: dataset.tail()
```

```
[22]:
```

	company	rank	revenues	revenue_change	profits	\
495	Teva Pharmaceutical Industries	496	21903	11.5	329.0	
496	New China Life Insurance	497	21796	-13.3	743.9	
497	Wm. Morrison Supermarkets	498	21741	-11.3	406.4	
498	TUI	499	21655	-5.5	1151.7	
499	AutoNation	500	21609	3.6	430.5	

	assets	profit_change	ceo	\
495	92890	-79.3	Yitzhak Peterburg	
496	100609	-45.6	Wan Feng	
497	11630	20.4	David T. Potts	
498	16247	195.5	Friedrich Joussen	
499	10060	-2.7	Michael J. Jackson	

	industry	sector	previous_rank	\
495	Pharmaceuticals	Health Care	0	
496	Insurance: Life, Health (stock)	Financials	427	
497	Food and Drug Stores	Food & Drug Stores	437	

498	Travel Services	Business Services	467
499	Specialty Retailers	Retailing	0

	country	hq_location	website \
495	Israel	Petach Tikva, Israel	http://www.tevapharm.com
496	China	Beijing, China	http://www.newchinalife.com
497	Britain	Bradford, Britain	http://www.morrisons.com
498	Germany	Hanover, Germany	http://www.tuigroup.com
499	USA	Fort Lauderdale, FL	http://www.autonation.com

	years_on_global_500_list	employees	total_stockholder_equity
495	1	56960	33337
496	2	54378	8507
497	13	77210	5111
498	23	66779	3006
499	12	26000	2310

```
[23]: dataset.tail(10)
```

```
[23]:
```

	company	rank	revenues	revenue_change \
490	National Grid	491	22036	-3.2
491	Dollar General	492	21987	7.9
492	Telecom Italia	493	21941	-17.4
493	Xiamen ITG Holding Group	494	21930	34.3
494	Xinjiang Guanghui Industry Investment	495	21919	31.1
495	Teva Pharmaceutical Industries	496	21903	11.5
496	New China Life Insurance	497	21796	-13.3
497	Wm. Morrison Supermarkets	498	21741	-11.3
498	TUI	499	21655	-5.5
499	AutoNation	500	21609	3.6

	profits	assets	profit_change	ceo \
490	10150.6	82310	160.2	John Pettigrew
491	1251.1	11672	7.4	Todd J. Vasos
492	1999.4	74295	NaN	Flavio Cattaneo
493	35.6	12161	-25.1	Xu Xiaoxi
494	251.8	31957	49.9	Shang Jiqiang
495	329.0	92890	-79.3	Yitzhak Peterburg
496	743.9	100609	-45.6	Wan Feng
497	406.4	11630	20.4	David T. Potts
498	1151.7	16247	195.5	Friedrich Joussen
499	430.5	10060	-2.7	Michael J. Jackson

	industry	sector	previous_rank \
490	Utilities	Energy	471
491	Specialty Retailers	Retailing	0
492	Telecommunications	Telecommunications	404

493		Trading	Wholesalers	0
494		Trading	Wholesalers	0
495		Pharmaceuticals	Health Care	0
496	Insurance: Life, Health (stock)		Financials	427
497		Food and Drug Stores	Food & Drug Stores	437
498		Travel Services	Business Services	467
499		Specialty Retailers	Retailing	0

	country	hq_location	website \
490	Britain	London, Britain	http://www.nationalgrid.com
491	USA	Goodlettsville, TN	http://www.dollargeneral.com
492	Italy	Milan, Italy	http://www.telecomitalia.com
493	China	Xiamen, China	http://www.itgholding.com.cn
494	China	Urumqi, China	http://www.guanghui.com
495	Israel	Petach Tikva, Israel	http://www.tevapharm.com
496	China	Beijing, China	http://www.newchinalife.com
497	Britain	Bradford, Britain	http://www.morrisons.com
498	Germany	Hanover, Germany	http://www.tuigroup.com
499	USA	Fort Lauderdale, FL	http://www.autonation.com

	years_on_global_500_list	employees	total_stockholder_equity
490	12	22132	25463
491	1	121000	5406
492	18	61227	22366
493	1	18454	1066
494	1	65616	4563
495	1	56960	33337
496	2	54378	8507
497	13	77210	5111
498	23	66779	3006
499	12	26000	2310

10 Series

- A pandas Series is like a column in a table
- one-dimensional array holding data of any type

11 Accessing a column

```
[24]: dataset['company']
```

```
[24]: 0          Walmart
      1      State Grid
      2      Sinopec Group
      3  China National Petroleum
      4      Toyota Motor
```

```

...
495    Teva Pharmaceutical Industries
496          New China Life Insurance
497    Wm. Morrison Supermarkets
498                                TUI
499          AutoNation
Name: company, Length: 500, dtype: object

```

```
[25]: dataset['profits']
```

```

[25]: 0      13643.0
      1      9571.3
      2      1257.9
      3      1867.5
      4     16899.3
...
495      329.0
496      743.9
497      406.4
498     1151.7
499      430.5
Name: profits, Length: 500, dtype: float64

```

```
[27]: #dataset.company -> but not a good command
```

```
[28]: dataset['company'].shape
```

```
[28]: (500,)
```

12 Accessing multiple columns

```
[32]: dataset[['company', 'rank', 'ceo']]
```

```

[32]:
      company  rank  ceo
0      Walmart    1  C. Douglas McMillon
1    State Grid    2      Kou Wei
2  Sinopec Group    3    Wang Yupu
3  China National Petroleum  4  Zhang Jianhua
4    Toyota Motor    5    Akio Toyoda
...
495  Teva Pharmaceutical Industries  496  Yitzhak Peterburg
496    New China Life Insurance  497      Wan Feng
497    Wm. Morrison Supermarkets  498  David T. Potts
498                                TUI  499  Friedrich Joussen
499          AutoNation    500  Michael J. Jackson

```


[500 rows x 3 columns]

```
[33]: dataset[['company', 'rank', 'ceo']].shape
```

```
[33]: (500, 3)
```

13 Selecting rows n columns using indexing

14 iloc

- iloc['rows', 'column']

```
[36]: dataset.iloc[4,2]
```

```
[36]: 254694
```

```
[39]: dataset.iloc[4:7]
```

```
[39]:
```

	company	rank	revenues	revenue_change	profits	assets \
4	Toyota Motor	5	254694	7.7	16899.3	437575
5	Volkswagen	6	240264	1.5	5937.3	432116
6	Royal Dutch Shell	7	240033	-11.8	4575.0	411275

	profit_change	ceo	industry \
4	-12.3	Akio Toyoda	Motor Vehicles and Parts
5	NaN	Matthias Muller	Motor Vehicles and Parts
6	135.9	Ben van Beurden	Petroleum Refining

	sector	previous_rank	country	hq_location \
4	Motor Vehicles & Parts	8	Japan	Toyota, Japan
5	Motor Vehicles & Parts	7	Germany	Wolfsburg, Germany
6	Energy	5	Netherlands	The Hague, Netherlands

	website	years_on_global_500_list	employees \
4	http://www.toyota-global.com	23	364445
5	http://www.volkswagen.com	23	626715
6	http://www.shell.com	23	89000

	total_stockholder_equity
4	157210
5	97753
6	186646

```
[40]: dataset.iloc[4:7, 0]
```

```
[40]: 4      Toyota Motor
      5      Volkswagen
      6      Royal Dutch Shell
      Name: company, dtype: object
```

```
[41]: dataset.iloc[4:7, 0:3]
```

```
[41]:      company  rank  revenues
      4      Toyota Motor    5    254694
      5      Volkswagen    6    240264
      6  Royal Dutch Shell    7    240033
```

```
[42]: dataset.iloc[[4,7,9], [0,1,7]]
```

```
[42]:      company  rank      ceo
      4      Toyota Motor    5      Akio Toyoda
      7  Berkshire Hathaway    8  Warren E. Buffett
      9      Exxon Mobil    10  Darren W. Woods
```

```
[47]: dataset.iloc[0:5, [0,1,7]]
```

```
[47]:      company  rank      ceo
      0      Walmart    1  C. Douglas McMillon
      1      State Grid    2      Kou Wei
      2      Sinopec Group    3      Wang Yupu
      3  China National Petroleum    4      Zhang Jianhua
      4      Toyota Motor    5      Akio Toyoda
```

15 loc

```
[48]: dataset.head()
```

```
[48]:      company  rank  revenues  revenue_change  profits  assets \
      0      Walmart    1    485873         0.8  13643.0  198825
      1      State Grid    2    315199        -4.4   9571.3  489838
      2      Sinopec Group    3    267518        -9.1   1257.9  310726
      3  China National Petroleum    4    262573       -12.3   1867.5  585619
      4      Toyota Motor    5    254694         7.7  16899.3  437575

      profit_change      ceo      industry \
      0      -7.2  C. Douglas McMillon  General Merchandisers
      1      -6.2      Kou Wei      Utilities
      2     -65.0      Wang Yupu  Petroleum Refining
      3     -73.7  Zhang Jianhua  Petroleum Refining
      4     -12.3  Akio Toyoda  Motor Vehicles and Parts
```

	sector	previous_rank	country	hq_location	\
0	Retailing	1	USA	Bentonville, AR	
1	Energy	2	China	Beijing, China	
2	Energy	4	China	Beijing, China	
3	Energy	3	China	Beijing, China	
4	Motor Vehicles & Parts	8	Japan	Toyota, Japan	

	website	years_on_global_500_list	employees	\
0	http://www.walmart.com	23	2300000	
1	http://www.sgcc.com.cn	17	926067	
2	http://www.sinopec.com	19	713288	
3	http://www.cnpc.com.cn	17	1512048	
4	http://www.toyota-global.com	23	364445	

	total_stockholder_equity
0	77798
1	209456
2	106523
3	301893
4	157210

```
[50]: dataset.loc[0]
```

```
[50]: company          Walmart
rank                1
revenues            485873
revenue_change      0.8
profits             13643.0
assets              198825
profit_change       -7.2
ceo                 C. Douglas McMillon
industry            General Merchandisers
sector              Retailing
previous_rank        1
country              USA
hq_location          Bentonville, AR
website              http://www.walmart.com
years_on_global_500_list  23
employees            2300000
total_stockholder_equity  77798
Name: 0, dtype: object
```

```
[49]: dataset.loc[0:5, 'company']
```

```
[49]: 0          Walmart
1      State Grid
2      Sinopec Group
```

```

3    China National Petroleum
4          Toyota Motor
5          Volkswagen
Name: company, dtype: object

```

```
[51]: dataset.loc[4:7]
```

```

[51]:      company  rank  revenues  revenue_change  profits  assets \
4    Toyota Motor    5    254694          7.7  16899.3  437575
5    Volkswagen     6    240264          1.5   5937.3  432116
6  Royal Dutch Shell    7    240033         -11.8   4575.0  411275
7  Berkshire Hathaway    8    223604          6.1  24074.0  620854

      profit_change      ceo      industry \
4         -12.3    Akio Toyoda  Motor Vehicles and Parts
5          NaN  Matthias Muller  Motor Vehicles and Parts
6        135.9  Ben van Beurden  Petroleum Refining
7          NaN  Warren E. Buffett  Insurance: Property and Casualty (Stock)

      sector  previous_rank  country  hq_location \
4  Motor Vehicles & Parts      8    Japan  Toyota, Japan
5  Motor Vehicles & Parts      7  Germany  Wolfsburg, Germany
6          Energy          5  Netherlands  The Hague, Netherlands
7      Financials         11    USA      Omaha, NE

      website  years_on_global_500_list  employees \
4  http://www.toyota-global.com          23    364445
5  http://www.volkswagen.com          23    626715
6  http://www.shell.com          23    89000
7  http://www.berkshirehathaway.com        21    367700

      total_stockholder_equity
4          157210
5          97753
6         186646
7         283001

```

```
[52]: dataset.loc[4:7, 'profits']
```

```

[52]: 4    16899.3
5     5937.3
6     4575.0
7    24074.0
Name: profits, dtype: float64

```

```
[53]: dataset.loc[4:7, 'ceo':]
```

```
[53]:
```

	ceo	industry \
4	Akio Toyoda	Motor Vehicles and Parts
5	Matthias Muller	Motor Vehicles and Parts
6	Ben van Beurden	Petroleum Refining
7	Warren E. Buffett	Insurance: Property and Casualty (Stock)

	sector	previous_rank	country	hq_location \
4	Motor Vehicles & Parts	8	Japan	Toyota, Japan
5	Motor Vehicles & Parts	7	Germany	Wolfsburg, Germany
6	Energy	5	Netherlands	The Hague, Netherlands
7	Financials	11	USA	Omaha, NE

	website	years_on_global_500_list	employees \
4	http://www.toyota-global.com	23	364445
5	http://www.volkswagen.com	23	626715
6	http://www.shell.com	23	89000
7	http://www.berkshirehathaway.com	21	367700

	total_stockholder_equity
4	157210
5	97753
6	186646
7	283001

```
[54]: dataset.loc[4:7, 'ceo':'website']
```

```
[54]:
```

	ceo	industry \
4	Akio Toyoda	Motor Vehicles and Parts
5	Matthias Muller	Motor Vehicles and Parts
6	Ben van Beurden	Petroleum Refining
7	Warren E. Buffett	Insurance: Property and Casualty (Stock)

	sector	previous_rank	country	hq_location \
4	Motor Vehicles & Parts	8	Japan	Toyota, Japan
5	Motor Vehicles & Parts	7	Germany	Wolfsburg, Germany
6	Energy	5	Netherlands	The Hague, Netherlands
7	Financials	11	USA	Omaha, NE

	website
4	http://www.toyota-global.com
5	http://www.volkswagen.com
6	http://www.shell.com
7	http://www.berkshirehathaway.com

```
[55]: dataset.loc[4:7, ['ceo', 'website', 'company', 'rank']]
```

```
[55]:
```

	ceo	website	company \
4	Akio Toyoda	http://www.toyota-global.com	Toyota Motor
5	Matthias Muller	http://www.volkswagen.com	Volkswagen
6	Ben van Beurden	http://www.shell.com	Royal Dutch Shell
7	Warren E. Buffett	http://www.berkshirehathaway.com	Berkshire Hathaway

	rank
4	5
5	6
6	7
7	8

```
[57]: dataset.loc[[4,7], ['ceo','website','company','rank']]
```

```
[57]:
```

	ceo	website	company \
4	Akio Toyoda	http://www.toyota-global.com	Toyota Motor
7	Warren E. Buffett	http://www.berkshirehathaway.com	Berkshire Hathaway

	rank
4	5
7	8

16 Boolean Indexing

```
[59]: dataset['country']
```

```
[59]:
```

0	USA
1	China
2	China
3	China
4	Japan
	...
495	Israel
496	China
497	Britain
498	Germany
499	USA

Name: country, Length: 500, dtype: object

```
[62]: dataset['country'] == 'Pakistan'
```

```
[62]:
```

0	False
1	False
2	False
3	False
4	False

```

...
495 False
496 False
497 False
498 False
499 False
Name: country, Length: 500, dtype: bool

```

```
[71]: boolean_indexing = dataset['country'] == 'India'
```

```
[72]: dataset[boolean_indexing]
```

```
[72]:
```

	company	rank	revenues	revenue_change	profits	assets	\
167	Indian Oil	168	53562	-2.1	2960.0	42132	
202	Reliance Industries	203	46931	8.0	4458.9	108856	
216	State Bank of India	217	44533	6.8	36.0	530590	
246	Tata Motors	247	40329	-4.2	1111.6	42162	
294	Rajesh Exports	295	36114	43.1	185.8	3717	
359	Bharat Petroleum	360	30316	4.2	1300.5	16801	
383	Hindustan Petroleum	384	28166	-2.3	1228.1	12370	

	profit_change	ceo	industry	\
167	72.7	Sanjiv Singh	Petroleum Refining	
202	5.6	Mukesh D. Ambani	Petroleum Refining	
216	-98.1	Arundhati Bhattacharya	Banks: Commercial and Savings	
246	-34.0	Guenter Butschek	Motor Vehicles and Parts	
294	13.8	Rajesh J. Mehta	Trading	
359	6.7	D. Rajkumar	Petroleum Refining	
383	63.4	Mukesh Kumar Surana	Petroleum Refining	

	sector	previous_rank	country	hq_location	\
167	Energy	161	India	New Delhi, India	
202	Energy	215	India	Mumbai, India	
216	Financials	232	India	Mumbai, India	
246	Motor Vehicles & Parts	226	India	Mumbai, India	
294	Wholesalers	423	India	Bengaluru, India	
359	Energy	358	India	Mumbai, India	
383	Energy	367	India	Mumbai, India	

	website	years_on_global_500_list	employees	\
167	http://www.iocl.com	23	34999	
202	http://www.ril.com	14	140483	
216	http://www.sbi.co.in	12	278872	
246	http://www.tatamotors.com	8	79558	
294	http://www.rajeshindia.com	2	328	
359	http://www.bharatpetroleum.in	14	13395	
383	http://www.hindustanpetroleum.com	14	10422	

	total_stockholder_equity
167	15724
202	40614
216	33450
246	8942
294	868
359	4747
383	3245

```
[73]: dataset[dataset['country'] == 'India']
```

```
[73]:
```

	company	rank	revenues	revenue_change	profits	assets	\
167	Indian Oil	168	53562	-2.1	2960.0	42132	
202	Reliance Industries	203	46931	8.0	4458.9	108856	
216	State Bank of India	217	44533	6.8	36.0	530590	
246	Tata Motors	247	40329	-4.2	1111.6	42162	
294	Rajesh Exports	295	36114	43.1	185.8	3717	
359	Bharat Petroleum	360	30316	4.2	1300.5	16801	
383	Hindustan Petroleum	384	28166	-2.3	1228.1	12370	

	profit_change	ceo	industry	\
167	72.7	Sanjiv Singh	Petroleum Refining	
202	5.6	Mukesh D. Ambani	Petroleum Refining	
216	-98.1	Arundhati Bhattacharya	Banks: Commercial and Savings	
246	-34.0	Guenter Butschek	Motor Vehicles and Parts	
294	13.8	Rajesh J. Mehta	Trading	
359	6.7	D. Rajkumar	Petroleum Refining	
383	63.4	Mukesh Kumar Surana	Petroleum Refining	

	sector	previous_rank	country	hq_location	\
167	Energy	161	India	New Delhi, India	
202	Energy	215	India	Mumbai, India	
216	Financials	232	India	Mumbai, India	
246	Motor Vehicles & Parts	226	India	Mumbai, India	
294	Wholesalers	423	India	Bengaluru, India	
359	Energy	358	India	Mumbai, India	
383	Energy	367	India	Mumbai, India	

	website	years_on_global_500_list	employees	\
167	http://www.iocl.com	23	34999	
202	http://www.ril.com	14	140483	
216	http://www.sbi.co.in	12	278872	
246	http://www.tatamotors.com	8	79558	
294	http://www.rajeshindia.com	2	328	
359	http://www.bharatpetroleum.in	14	13395	
383	http://www.hindustanpetroleum.com	14	10422	

	total_stockholder_equity
167	15724
202	40614
216	33450
246	8942
294	868
359	4747
383	3245

17 Skill Test - 1

Extract Chinese Companies Name

```
[81]: dataset[dataset['country'] == 'China']['company']
```

```
[81]: 1          State Grid
      2          Sinopec Group
      3      China National Petroleum
      21  Industrial & Commercial Bank of China
      23      China State Construction Engineering
      ...
      487          Xiamen C&D
      489      China General Technology
      493          Xiamen ITG Holding Group
      494  Xinjiang Guanghui Industry Investment
      496          New China Life Insurance
      Name: company, Length: 109, dtype: object
```

```
[82]: dataset[dataset['country'] == 'China']
```

```
[82]:
```

	company	rank	revenues	revenue_change	\
1	State Grid	2	315199	-4.4	
2	Sinopec Group	3	267518	-9.1	
3	China National Petroleum	4	262573	-12.3	
21	Industrial & Commercial Bank of China	22	147675	-11.7	
23	China State Construction Engineering	24	144505	3.1	
..	
487	Xiamen C&D	488	22145	6.6	
489	China General Technology	490	22113	-20.1	
493	Xiamen ITG Holding Group	494	21930	34.3	
494	Xinjiang Guanghui Industry Investment	495	21919	31.1	
496	New China Life Insurance	497	21796	-13.3	

	profits	assets	profit_change	ceo	\
1	9571.3	489838	-6.2	Kou Wei	
2	1257.9	310726	-65.0	Wang Yupu	

3	1867.5	585619	-73.7	Zhang Jianhua
21	41883.9	3473238	-5.0	Gu Shu
23	2492.9	201269	10.7	Guan Qing
..
487	280.2	21729	15.6	Huang Wenzhou
489	413.6	20860	-20.8	Xu Xianping
493	35.6	12161	-25.1	Xu Xiaoxi
494	251.8	31957	49.9	Shang Jiqiang
496	743.9	100609	-45.6	Wan Feng

	industry	sector \
1	Utilities	Energy
2	Petroleum Refining	Energy
3	Petroleum Refining	Energy
21	Banks: Commercial and Savings	Financials
23	Engineering, Construction	Engineering & Construction
..
487	Trading	Wholesalers
489	Engineering, Construction	Engineering & Construction
493	Trading	Wholesalers
494	Trading	Wholesalers
496	Insurance: Life, Health (stock)	Financials

	previous_rank	country	hq_location	website \
1	2	China	Beijing, China	http://www.sgcc.com.cn
2	4	China	Beijing, China	http://www.sinopec.com
3	3	China	Beijing, China	http://www.cnpc.com.cn
21	15	China	Beijing, China	http://www.icbc-ltd.com
23	27	China	Beijing, China	http://www.cscec.com
..
487	0	China	Xiamen, China	http://www.chinacdc.com
489	383	China	Beijing, China	http://www.genertec.com.cn
493	0	China	Xiamen, China	http://www.itgholding.com.cn
494	0	China	Urumqi, China	http://www.guanghui.com
496	427	China	Beijing, China	http://www.newchinalife.com

	years_on_global_500_list	employees	total_stockholder_equity
1	17	926067	209456
2	19	713288	106523
3	17	1512048	301893
21	19	461749	283438
23	6	263915	15344
..
487	1	18381	3985
489	4	38589	5114
493	1	18454	1066
494	1	65616	4563

496

2

54378

8507

[109 rows x 17 columns]

```
[88]: dataset[dataset['country'] == 'China'][['company', 'rank']]
```

```
[88]:
```

	company	rank
1	State Grid	2
2	Sinopec Group	3
3	China National Petroleum	4
21	Industrial & Commercial Bank of China	22
23	China State Construction Engineering	24
..
487	Xiamen C&D	488
489	China General Technology	490
493	Xiamen ITG Holding Group	494
494	Xinjiang Guanghui Industry Investment	495
496	New China Life Insurance	497

[109 rows x 2 columns]

```
[90]: #dataset[dataset['country'] == 'China'][['company', 'rank']] -> run this
      ↪ command using loc
dataset.loc[dataset['country'] == 'China', ['company', 'rank']]
```

```
[90]:
```

	company	rank
1	State Grid	2
2	Sinopec Group	3
3	China National Petroleum	4
21	Industrial & Commercial Bank of China	22
23	China State Construction Engineering	24
..
487	Xiamen C&D	488
489	China General Technology	490
493	Xiamen ITG Holding Group	494
494	Xinjiang Guanghui Industry Investment	495
496	New China Life Insurance	497

[109 rows x 2 columns]

18 value_counts()

```
[91]: dataset['country'].value_counts()
```

```
[91]: USA          132
      China        109
```

Japan	51
Germany	29
France	29
Britain	24
South Korea	15
Netherlands	14
Switzerland	14
Canada	11
Spain	9
Australia	7
Brazil	7
India	7
Italy	7
Taiwan	6
Russia	4
Ireland	4
Singapore	3
Sweden	3
Mexico	2
Malaysia	1
Thailand	1
Belgium	1
Norway	1
Luxembourg	1
Indonesia	1
Denmark	1
Saudi Arabia	1
Finland	1
Venezuela	1
Turkey	1
U.A.E	1
Israel	1

Name: country, dtype: int64

[]: