Basic Pandas

Create a new dataframe

```python
import pandas as pd
```

```python
# create dataframe from scratch
raw_data = {
    "name": ["Toy", "Joe", "Mary", "John", "Anna"],
    "age": [33, 25, 20, 22, 31],
    "gender": ["M", "M", "F", "M", "F"]
}

df = pd.DataFrame(raw_data)

df
```

	name	age	gender
0	Toy	33	M
1	Joe	25	M
2	Mary	20	F
3	John	22	M
4	Anna	31	F

Add a new column

```python
# add new column
df['city'] = ['London', 'London', 'London', 'Manchester', 'Liverpool']
```

```python
df
```

	name	age	gender	city
0	Toy	33	M	London
1	Joe	25	M	London
2	Mary	20	F	London
3	John	22	M	Manchester
4	Anna	31	F	Liverpool

```python
df.shape
```

```
(5, 4)
```

Drop a column

```python
# drop column city
df = df.drop('city', axis=1)
```

```python
df
```

	name	age	gender
0	Toy	33	M
1	Joe	25	M
2	Mary	20	F
3	John	22	M
4	Anna	31	F

```python
# remove index = 2
df = df.drop(2, axis=0)
```

```python
df
```

	name	age	gender
0	Toy	33	M
1	Joe	25	M
3	John	22	M
4	Anna	31	F

```python
# reset index
df = df.reset_index(drop = True)
df
```

	name	age	gender
0	Toy	33	M
1	Joe	25	M
2	John	22	M
3	Anna	31	F

Rename columns

```python
# column names
list(df.columns)
```

```
['name', 'age', 'gender']
```

```python
# rename columns
df.columns = ['nickname', 'age', 'sex']
df
```

	nickname	age	sex
0	Toy	33	M
1	Joe	25	M
2	John	22	M
3	Anna	31	F

Series vs. DataFrame

```python
type(df['nickname'])
```

```
pandas.core.series.Series
```

```python
type(df)
```

```
pandas.core.frame.DataFrame
```

```python
# create a new series
s1 = pd.Series(['Mary', 20, 'F'], index=['nickname', 'age', 'sex'])
print(s1)
print(type(s1))
```

```
nickname    Mary
age           20
sex            F
dtype: object
<class 'pandas.core.series.Series'>
```

```python
# append s1 to df
df = df.append(s1, ignore_index=True)
df
```

	nickname	age	sex
0	Toy	33	M
1	Joe	25	M
2	John	22	M
3	Anna	31	F
4	Mary	20	F

```python
# create a new column using s2
s2 = pd.Series(['London', 'London', 'London', 'Manchester', 'Liverpool'])
df['city'] = s2
df
```

	nickname	age	sex	city
0	Toy	33	M	London
1	Joe	25	M	London
2	John	22	M	London
3	Anna	31	F	Manchester
4	Mary	20	F	Liverpool

Write csv file

```python
# write csv file
df.to_csv('mydata.csv')
```

Import csv file

```python
# import csv file
df2 = pd.read_csv('data/data.csv')
df2
```

	id	name	city
0	1	John	London
1	2	Joe	Liverpool
2	3	Mary	Manchester
3	4	Anna	Swansea
4	5	David	London

Import Excel

```python
# import excel file
import pandas as pd
df3 = pd.read_excel('data/data.xlsx')
```

df3

	id	name	city
0	1	John	London
1	2	Joe	Liverpool
2	3	Mary	Manchester
3	4	Anna	Swansea
4	5	David	London

Import JSON file

```python
# import json
df4 = pd.read_json('data/data.json')
df4
```

	ebook	language	amazonRating	myFavorite
0	Getting started with Python	python	4.89	True
1	Introduction to R	r	4.88	False
2	SQL for Beginners	sql	4.75	True

```python
df4['myFavorite'].dtype
```

```
dtype('bool')
```