Basic Pandas

```
import pandas as pd
```

```
# 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

```
df["city"] = ["London", "London", "London", "Manchester", "Liverpool"]
```

```
df
```

	name	age	gender	city

```
df.shape
```

```
(5, 4)
```

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

```
df
```

	name	age	gender

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

```
df
```

	name	age	gender

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

```
df
```

	name	age	gender

```
# Column names
list(df.columns)
```

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

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

	nickname	age	sex

```
type(df["nickname"])
```

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

```
#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'>
```

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

	nickname	age	sex

```
s2 = pd.Series(["London", "London", "London", "Manchester", "Liverpool"])
df["city"] = s2
df
```

	nickname	age	sex	city

```
# write CSV file
df.to_csv("mydata.csv")
```

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

	id	name	city

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

	id	name	city

```
# import jsom
df4 = pd.read_json("data/data.json")
df4
```

	ebook	language	amazonRating	myFavorite

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

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