

Double-click (or enter) to edit

```
import pandas as pd
url = 'https://raw.githubusercontent.com/makhan010385/DataSet/main/data.csv'
df = pd.read_csv(url, index_col=0, parse_dates=[0])
df
```

	Pulse	Maxpulse	Calories
Duration			
60	110	130	409.1
60	117	145	479.0
60	103	135	340.0
45	109	175	282.4
45	117	148	406.0
...	...	...	...
60	105	140	290.8
60	110	145	300.0
60	115	145	310.2
75	120	150	320.4
75	125	150	330.4

169 rows × 3 columns

## ▼ Show the relationship between the columns:

```
df.corr()
```

	Pulse	Maxpulse	Calories
Pulse	1.000000	0.786535	0.025121
Maxpulse	0.786535	1.000000	0.203813
Calories	0.025121	0.203813	1.000000



Note: The `corr()` method ignores "not numeric" columns.

## Result Explained

1. The Result of the `corr()` method is a table with a lot of numbers that represents how well the relationship is between two columns.
2. The number varies from -1 to 1.
3. means that there is a 1 to 1 relationship (a perfect correlation), and for this data set, each time a value went up in the first column, the other one went up as well.
4. 0.9 is also a good relationship, and if you increase one value, the other will probably increase as well.
5. -0.9 would be just as good relationship as 0.9, but if you increase one value, the other will probably go down.
6. 0.2 means NOT a good relationship, meaning that if one value goes up does not mean that the other will.

# What is a good correlation?

It depends on the use, but I think it is safe to say you have to have at least 0.6 (or -0.6) to call it a good correlation.

## ▼ Perfect Correlation:

We can see that "Duration" and "Duration" got the number 1.000000, which makes sense, each column always has a perfect relationship with itself.

## Good Correlation:

"Duration" and "Calories" got a 0.922721 correlation, which is a very good correlation, and we can predict that the longer you work out, the more calories you burn, and the other way around: if you burned a lot of calories, you probably had a long work out.

## Bad Correlation:

"Duration" and "Maxpulse" got a 0.009403 correlation, which is a very bad correlation, meaning that we can not predict the max pulse by just looking at the duration of the work out, and vice versa.

[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 8:22 AM

