Groupby and Group Manipulation

In this play, you will learn how to do analysis on groups(individual categories) in a pandas dataframe of a categorical feature. The concept of split-apply-combine, manipulate groups separately based on different rules.

When you are working on a dataset, be it data preparation phase, EDA phase, you will often come upon a situation where you would like to get some statistical inference from the data within each individual category.
Too much to process what I just said? Let's think with an example.

For example,
Let's say you have a dataset where you have records of your friends and the duration of talk you had with that friend each time you met him/her.

S.no	Friend	Duration of Meet
1.	A	10 mins
2.	A	13 mins
3.	B	6 mins
4.	B	2 mins
5.	B	3 mins
6.	C	40 mins
7.	D	1 min

In real life let's assume you had about 1000s to 10,000s of such records, now some day you wanted to just see whom you are more in touch with whom you talked the most etc. etc.

The result we want to generate would be something like this

Friend	Average talk time	Frequency of meet	Longest Talk	Shortest talk
A	11.5 mins	2	13 mins	10 mins
B	3.67 mins	3	6 mins	2 mins
C	40 mins	1	40 mins	40 mins
D	1 min	1	1 mins	1 mins

So as you can see we just checked some statistical analysis on group level (your each friend).