pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$
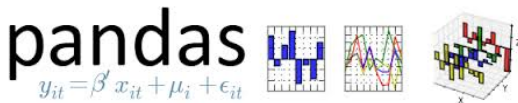
- ▶ pandas is a high-performance module that provides a comprehensive set of structures for working with data.
- ▶ pandas excels at handling structured data, such as data sets containing many variables, working with missing values and merging across multiple data sets.

Data Wrangling with Pandas, NumPy, and IPython

Python for Data Analysis

pandas

- While extremely useful, pandas is not an essential component of the Python scientific stack unlike NumPy, SciPy or matplotlib, and so while pandas doesnt make data analysis possible in Python, it makes it much easier.
- pandas also provides high-performance, robust methods for importing from and exporting to a wide range of formats.

pandas

Data Structures

pandas provides a set of data structures which include Series, DataFrames and Panels.

- ▶ **Series** are 1-dimensional arrays.
- ▶ **DataFrames** are collections of Series and so are 2-dimensional,
- ▶ **Panels** are collections ofDataFrames, and so are 3-dimensional.