Pandas

NumPy's ndarray data structure provides essential features for the type of clean, well-organized data typically seen in numerical computing tasks. Pandas with its Series and DataFrame objects, builds on the NumPy array structure and provides efficient access to these sorts of "data munging" tasks that occupy much of a data scientist's time.

**import** **pandas** **as** **pd**

A Pandas Series is a one-dimensional array of indexed data. It can be created from a list or array. If a Series is an analog of a one-dimensional array with flexible indices, a DataFrame is an analog of a two-dimensional array with both flexible row indices and flexible column names.

To manually store data in a table, create a DataFrame. When using a Python dictionary of lists, the dictionary keys will be used as column headers and the values in each list as rows of the DataFrame.

A DataFrame is a 2-dimensional data structure that can store data of different types (including characters, integers, floating point values, categorical data and more) in columns. It is similar to a spreadsheet, a SQL table or the data.frame in R.

Selecting a single column of a pandas DataFrame, the result is a pandas Series. To select the column, use the column label in between square brackets [ ].The describe() method provides a quick overview of the numerical data in a DataFrame.