PartⅠ：

Python Pandas file in/out：

The module built in a few functions to read data from the hardware,

read\_csv; read\_table; read\_fwf; read\_clipboard;

All those functions have some common keywords: 1.index, 2.type inference and data transformation; 3.date analyses: include combination; 4.literation: read some big file block by block;5.unformat data problem: jump across some lines, notes or other unimportant staffs,

Tips:

Keyword for read\_csv; read\_table:

Path: file location

Sep: string or regular expression

Header: name of the column, if absent, none should be set

Indexical:

Skiprow:as the name;

Na\_value: set the data to replact the NAN

Comment: used to depart the comment data;

Nrows: set for lines need to read

Iterator: get no idea on how this used:

JSON：often used in HTTP and Web browser also some communication between application.

Before use:

Import Json – this line is necessary

Two important functions in module: json.load / json.dumps

Also this is a good way to build Dataframe.

Web information

To read data from the websites, module lxml is given to this occasion,

decide more practice needs on its way.

Chapter 6 mainly concentrate on those useful way to load data from different sources. Among all those, the web source i think is the most valuable one.

Practice for today. Let’s do some practice to get data from finance YaHoo.

Chapter 7 Data manipulate.

If you may would like to make calculation of all the elements in the Dataframe. Then you may need to use the method **DataFrame.map()** this function is similar to the **apply() [**chapter 5, page 138 in python for data analysis]

Yet the **map()** method seems to be more powerful and can dealing with not only functions but also the list and dictionary.

Dummy matrix: the target to import dummy variable is to quantify those variable which can not be quantified before, like the influence of sexual in career choice, war and nature disaster to GDP, according to the attribute of those element, build the man-made variable exist only 0/1