Data Science with Python

Why Python

1. Easy to learn
2. Not just statistics language but data acquisition, cleaning, databases, high performance computing and modeling
3. Strong data science libraries eg Scipy ecosystem

Python pandas toolkit

Advanced querying and manipulation. Basic statistical analysis with numpy and scipy

Good data science brings skepticism, experimentation, simulation and replication.

Python is high level language, interpreted directly compiled into machine code. Good for where exploration is better, dynamically typed language. Set and modify the variables. Python interpreter is stateful.

Static typing in python – strings

Tuples – immutable, mixed types, unpacking in python

Lists – mutable, mixed types, + sign concatenates lists, \* operation repeats elements in the lists, all strings are lists of characters, Slicing in python

Dictionaries – iterate with items function

%precision 2

With open(‘file.csv’) as csvfile:

List(csv.DictReader(csvfile))