Numpy and Pandas Fundamentals

MAYANK MISHRA

Numpy and Pandas Fundamentals

1. Introduction to Numpy

2. Coding with Numpy

3. Introduction to Pandas

4. Coding with Pandas

Introduction to Numpy

Numpy is Python's 'C' extension library for array oriented computing.
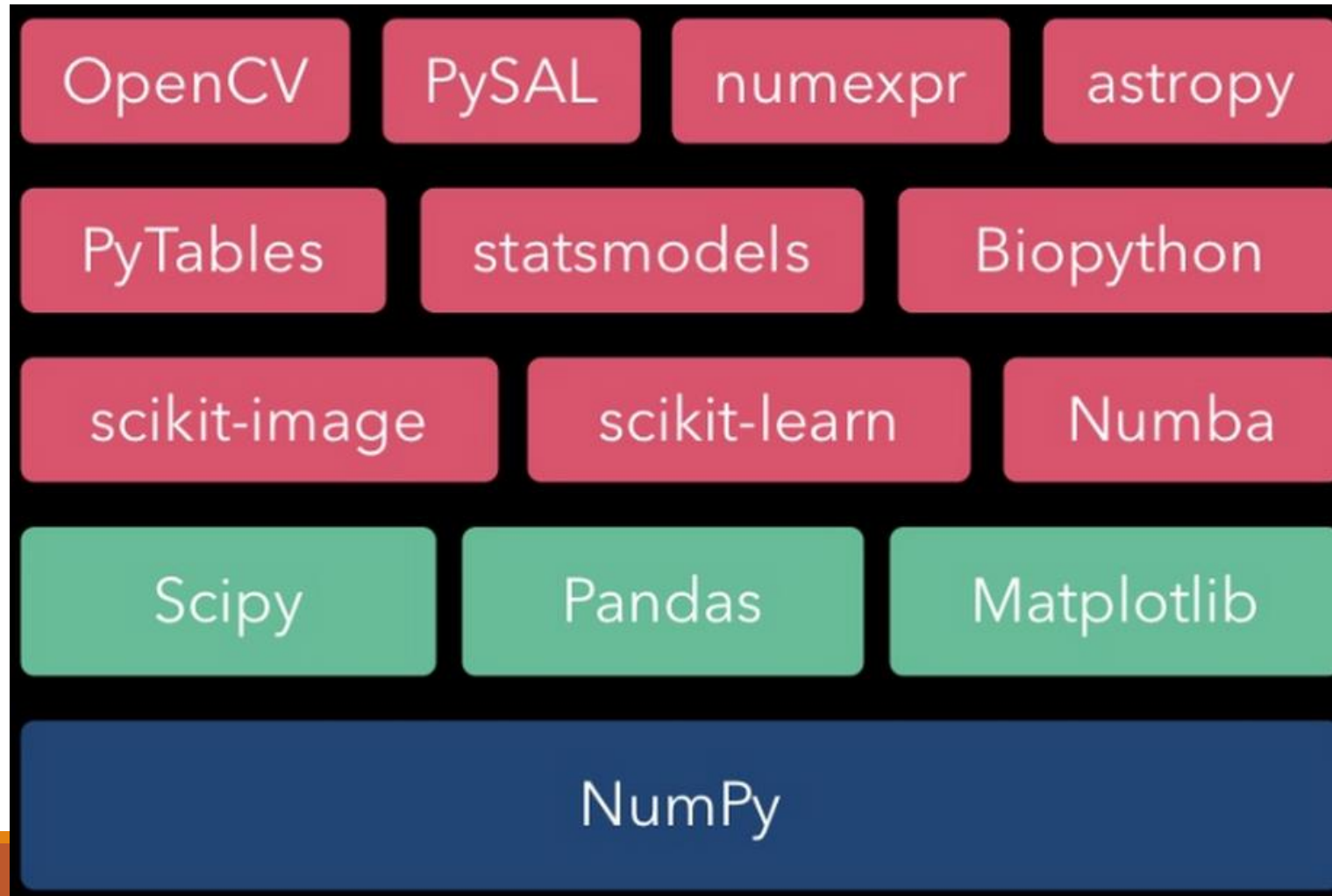
- ◦ Efficient

- ◦ In memory

- ◦ Contiguous (or strides)

- ◦ Homogeneous

Numpy is suited to many application

- ◦ Image Processing

- ◦ Signal Processing

- ◦ Linear Algebra

- ◦ A plethora of others

Numpy is the foundation of the python scientific stack.

Numpy Ecosystem

Let's Code

Introduction to Pandas

Rich relational data tool build on top of Numpy

◦ Excellent Performance

◦ Easy to use

◦ A foundation for data analysis in Python

Heavy production used in finance industry and others

Hope: Basis for the next generation statistical computing and analysis environment.

Introduction to Pandas Cont...

Simplifying Data Wrangling

◦ Data munging / preparation/ cleaning / integration is slow error prone and time consuming.

◦ Every one already love python <3 for the above tasks and pandas is icing on the cake.

◦ Data munging or Data wrangling is the process of cleaning the messy data.

DataFrame

columns	foo	bar	baz	qux
index				
A	0	x	2.7	True
B	4	y	6	True
C	8	z	10	False
D	-12	w	NA	False
E	16	a	18	False

- NumPy array-like
- Each column can have a different type
- Row and column index
- Size mutable: insert and delete columns

Let's Code

Thank You!