

Python and Data Analysis

Course info session

课程介绍公开课

2017 Copyright @ Data Application Lab
2017-04-15

Agenda

- Data Application Lab 简介
 - Python & Data Analysis 课程纲要
 - Josh老师: Python以及课程Capstone Project
 - Jason老师: Python在数据科学中的作用
 - Frank老师: Python在金融数据分析中的作用
 - Kai同学: 前学员如何实际利用Python解决问题
 - Q&A
-



Data Application Lab

Data Application Lab

Our graduates work for great companies

[24]7



BOSCH



Abbott
A Promise for Life

ACI
payment systems



Microsoft

Bloomberg



PlanGrid

ACUMEN

amazon

Booking.com

DISCOVER

OPERA
MEDIWORKS

Qumulo

Quora



Alibaba Group
阿里巴巴集团

ebay

Fanatics

GoodData

PROMETHEUS
Therapeutics & Diagnostics
A Nestlé Health Science Company

SONY



Symantec

Google

HUAWEI



irdeto



TWO SIGMA

VISA

yelp

JD.COM
京东

LinkedIn

RE/MAX
PACIFIC REALTY

MONSANTO

tableau

zest finance

Zillow

腾讯游戏
Tencent Games

FYI:

- <https://www.dataapplab.com/about-us/>
 - Partner: Data Science Association
 - <http://socaldatascience.org/> 2016.9 @ Southern California
 - <http://www.datascienceassociations.org/>
 - 2017.2 @ Dallas 2017.5 @Chicago 2017.10 @ Pasadena
 - Facebook: <https://www.facebook.com/dataapplab/>
 - Linkedin: <https://www.linkedin.com/company/7956968>
 - 微信: 大数据应用
-



Syllabus

Time: Sat & Sun 2-4 pm PST 5周课程, 3周从基础起学习Python, 后2周实战数据分析 (金融Quant分析+ Web Project)

Class 1 – 6 (Python from scratch + Mini Projects)

Topic 1: Introducing Python

- Python background
- Basic Data types
- Syntax and rules

Topic 2: Intermediate Python

- Complex Data Types
- Functions

Topic 3: Project I

- Object Oriented Programming
- API

- [Project]: Weather History

Topic 4: Advanced Python

- Data Structure
- Algorithm I

Topic 5: Python as a Pro

- Algorithm II
- Code Style
- Learn how to learn

Topic 6: Project II

- Module & Packages
- [Project] Web Scraping and Sentimental Analysis

4/22/2017

4/23

4/29

4/30

5/6

5/7

Class 5 – 8 Python in Quantitative Analysis (Python in Practice) 5/13 – 5/14

Topic 1: Monte Carlo simulation

- Stochastic process
- Variance reduction
- Option pricing – American/ European
- Risk Measurement

Topic 2: Principle Component Analysis & Bayesian Method

- Replicated index
- Bayesian Regression

Topic 3: Foreign Exchange Market Making

Topic 4: Quantitative Finance Interview

Class 9-10 Python Web Project

5/20

- Web development using Python (Flask framework)

Capstone Project

5/21

- Project: Financial Index Replication using Principle Component Analysis
- Python Visualization

<http://dataapplab.com/python>

Introducing Python

Your First Programming Language

Josh

Python Popularity

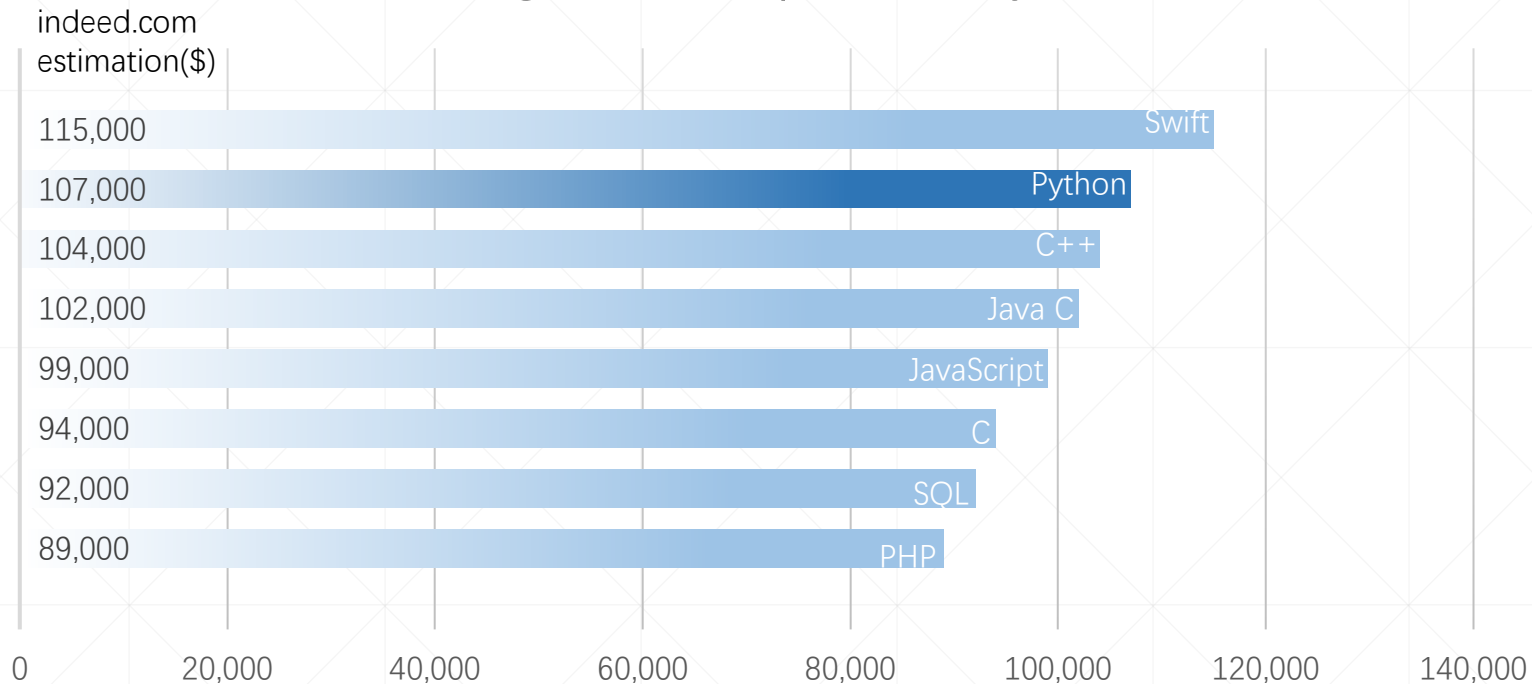
Your first programming language

- ✓ Python is NOW the most popular introductory language at TOP U.S University
- ✓ Eight of the top 10 CS departments (80%), and 27 of the top 39 (69%), teach Python in introductory CS0 or CS1 courses.
- ✓ The largest three MOOC providers (edX, Coursera, and Udacity) all offer introductory programming courses in Python.

Python Popularity

Job opportunities

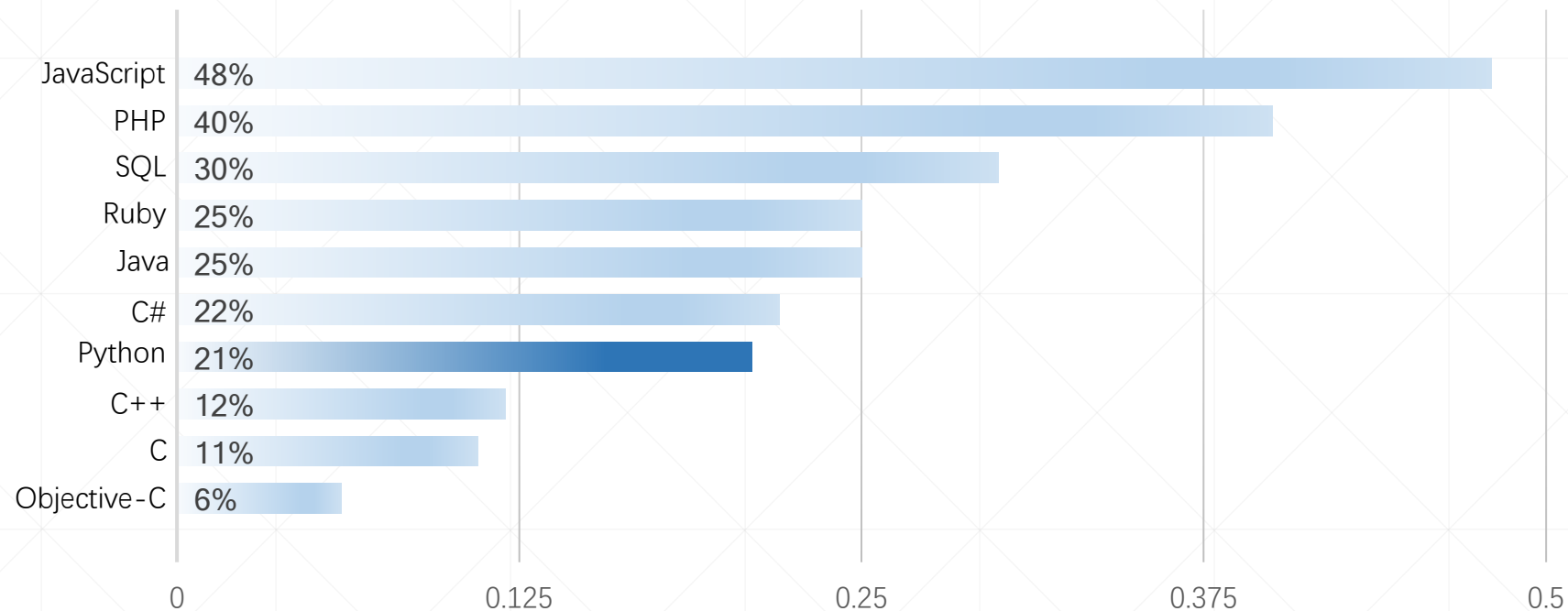
2016 Average Developer Salary in the U.S.



Python Popularity

The developer supply market

Codementors Expertise Distribution



Python Popularity

Some famous platform / software that relies on Python



Python Popularity

Python is good for data analysis

- ✓ TensorFlow:
 - Python
 - C++
 - Go (Experiment)
- ✓ Other Platform:
 - Caffe: Python
 - Paddle Paddle: Python C++
 - Java: Almost the only choice is dl4j



Caffe



Python Web

Project description

Josh

Why Web Development (offline vs online)

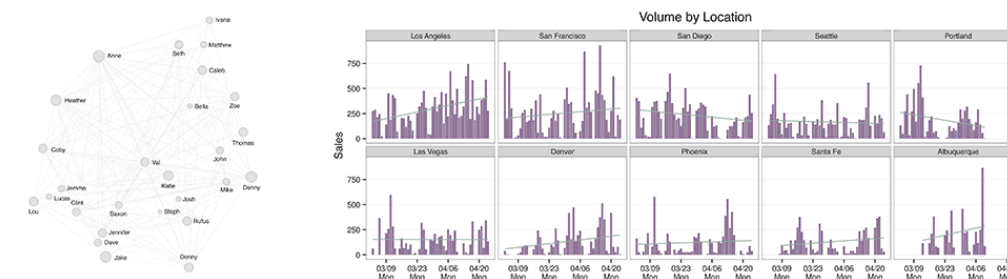
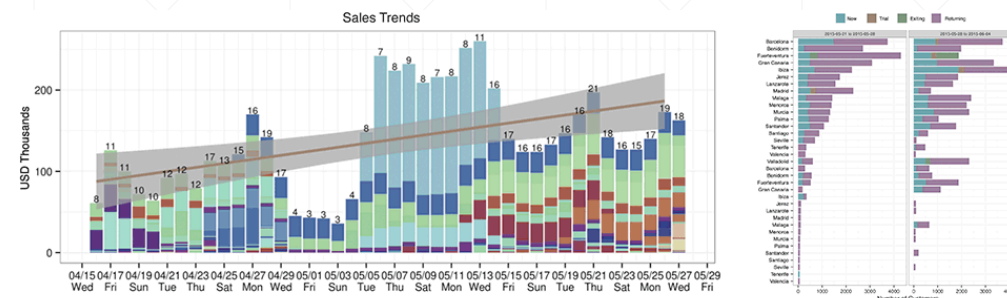
Additional items to explore [See more](#)



Inspired by your shopping trends



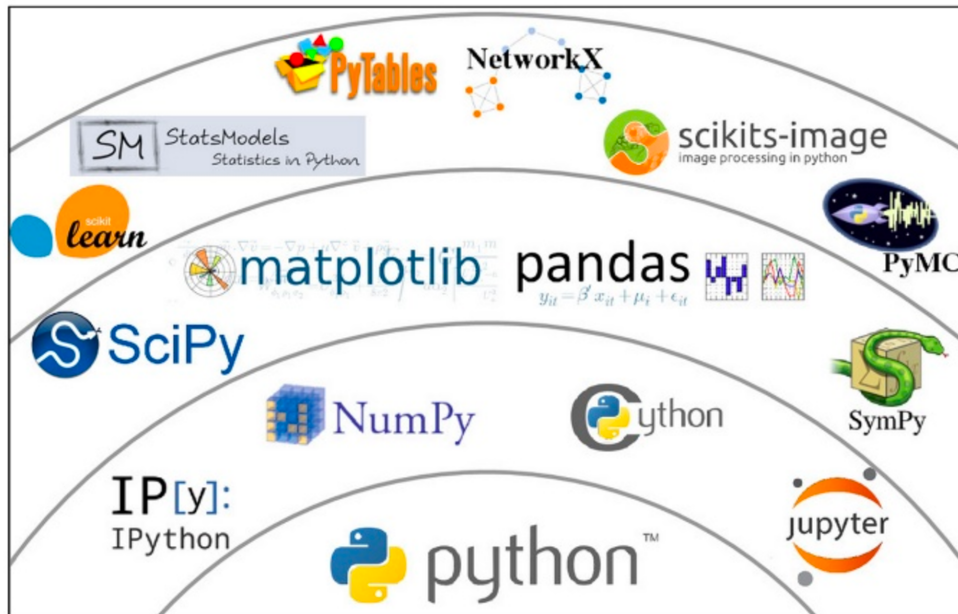
Build Better Products



Make Better Decisions

Python Ecosystem

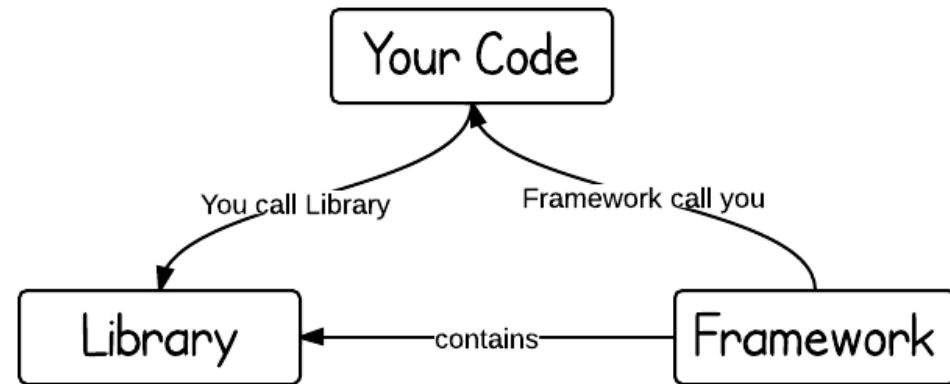
Data Analysis & Visualization



Web Development



Library, Framework



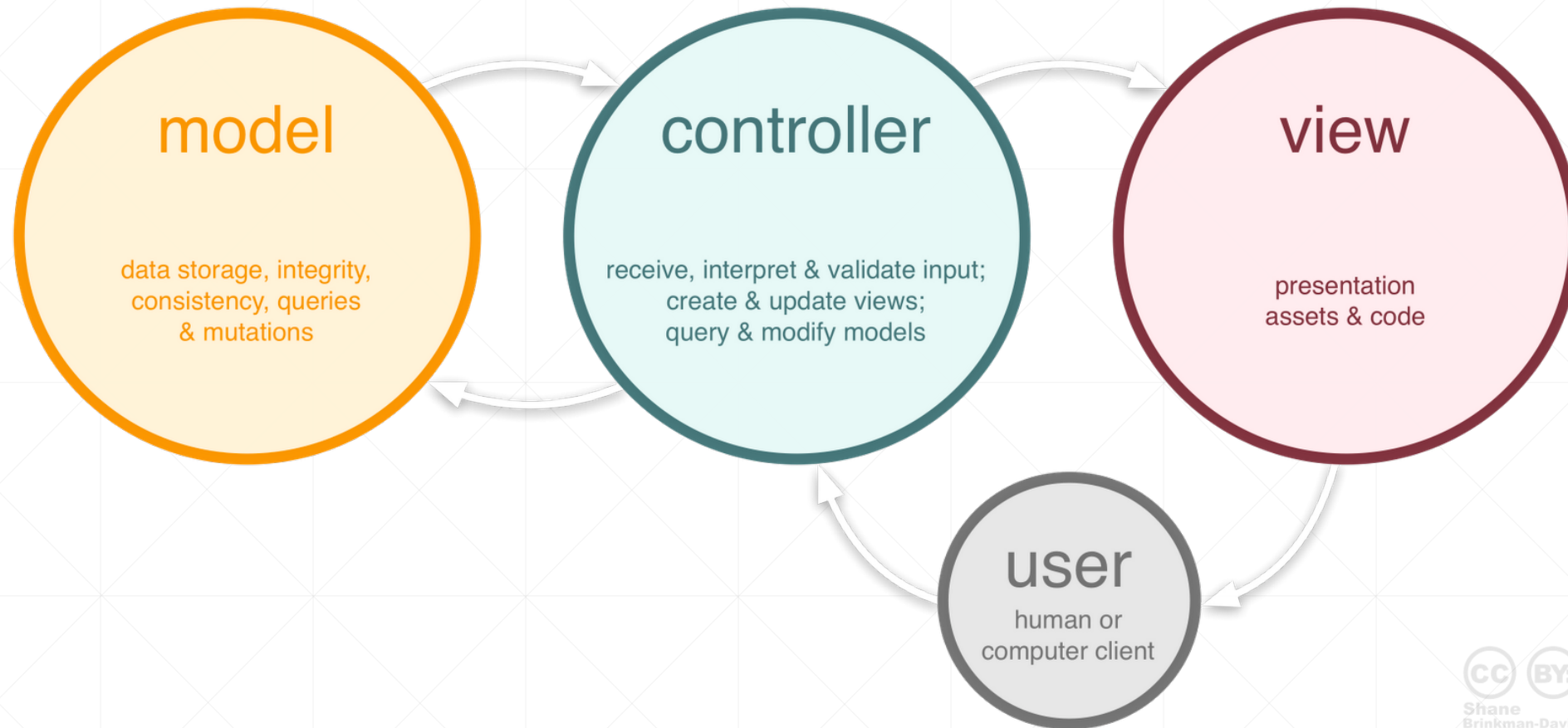


Web Development Full Stack

Front End	HTML5, Javascript, CSS
Back End	Python, NodeJs, Ruby on Rail
Data Sets	Mysql, MongoDB
Infrastructure	Servers, Cloud (AWS, Google, Azure), Container



We will learn MVC Architecture



Course Purpose

- Understand Web Basic Knowledge
 - Learn Flask Framework
 - Understand How Data Analysis Model being applied on Web Application
-

Project Demo Example:

European Option Simulation

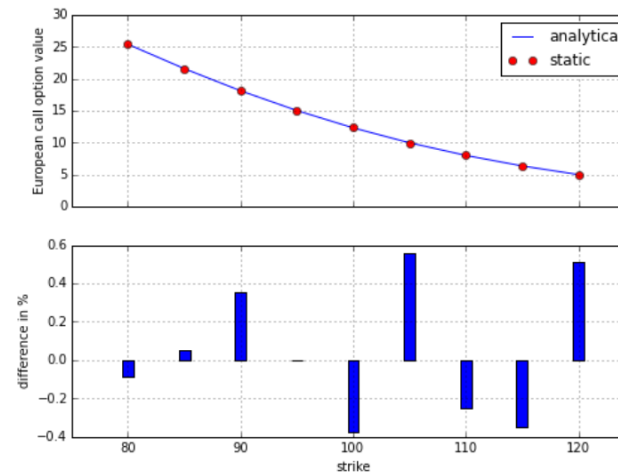
Valuation of European call option in Black-Scholes-Merton by Monte Carlo simulation (of index level at maturity)



S_0 _____
 r _____
 σ _____
 T _____
 I _____

Batch Upload

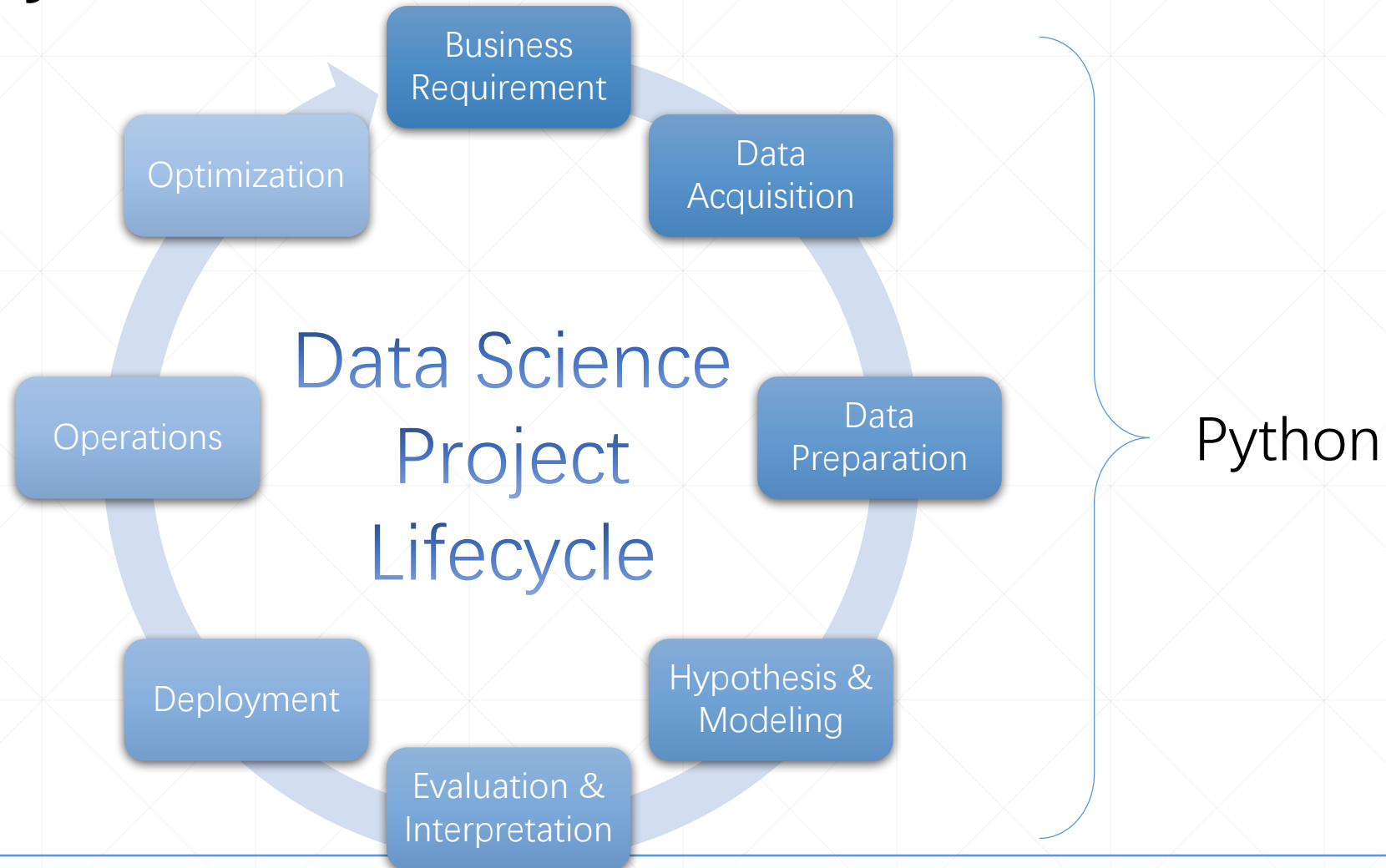
Submit



Python in Data Science

Jason

Why Data Science?



Why use python?

Everywhere in industry for all kinds of jobs:

- National Labs
- Finance
- Oil & Gas
- Consumer Products
- Aerospace/Defense
- Data Analytics
- School teaching
- Mini Projects

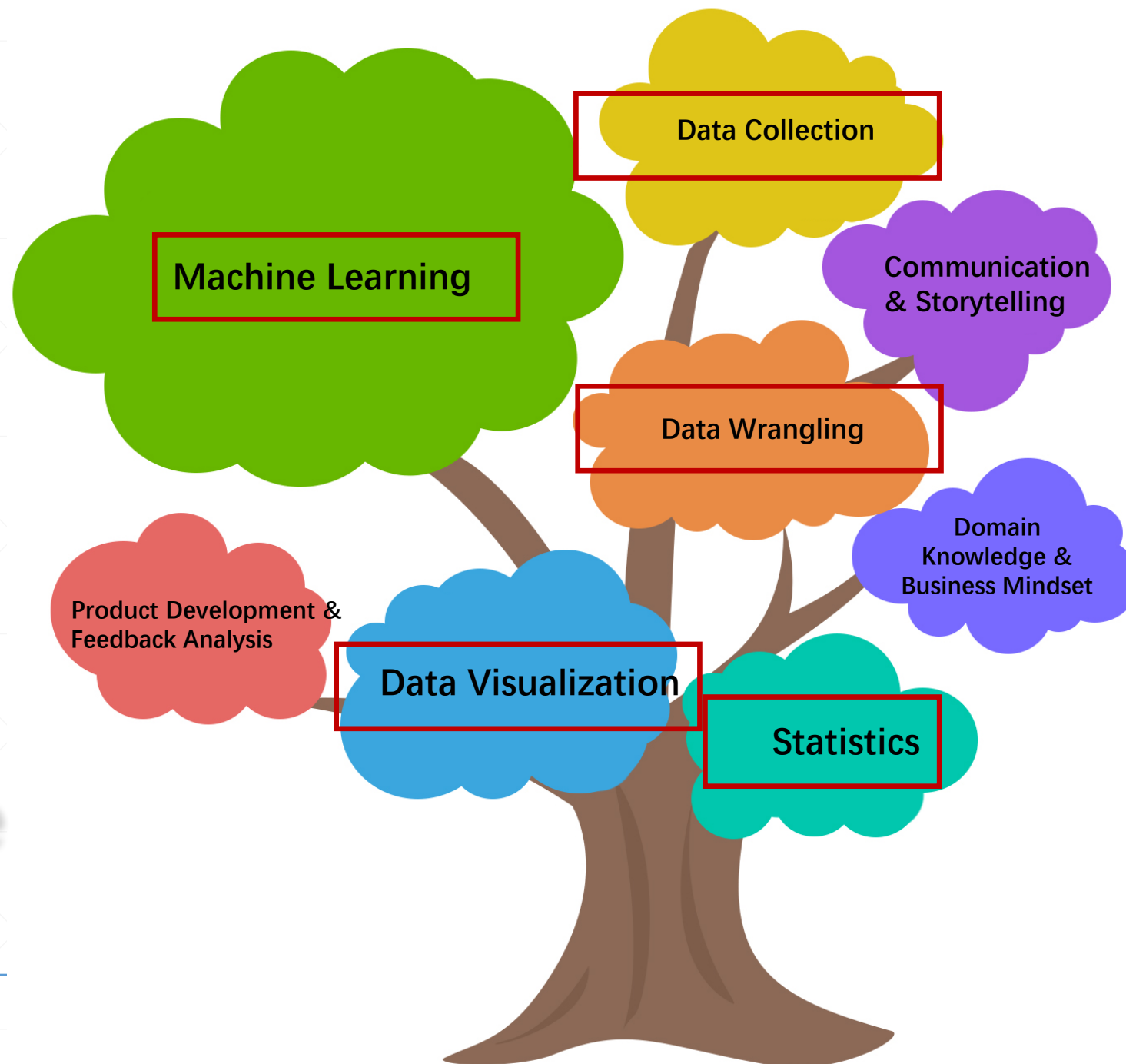


Python common toolkits in Data Science

- **numpy**: basic array manipulation
 - **scipy**: scientific computing in python, including signal processing and optimization
 - **matplotlib**: visualization and plotting
 - **IPython**: write and run python code interactively in a shell or a notebook
 - **pandas**: data manipulation
 - **scikit-learn**: machine learning
-

Need Python

Data Science Skill Tree



Data (Quant) Analysis

Brief introduction

Frank

Python in Analysis?

- HFT: AQR, KCG, Citadel
 - Asset Management: BlackRock, Vanguard
 - Banking: Goldman Sachs, JP Morgan
 - Rating Agency: S&P, Moody
 - Bloomberg
-

C++ is used everywhere, Python is widely used in financial industry.

- Goldman Sachs: Slang
- Morgan Stanley: Q
- JP Morgan: Python
- Bank of America Merrill Lynch: Python
- Jane Street: Ocaml
- Two Sigma: Java
- AQR: Python
- KCG: C++/Python





For OOP fans

- Interested in coding.
- Know how to prepare programming interview.





What else?

- Stochastic Calculus
 - Machine learning
 - Time series, prob, stats
 - Derivatives
 - Data Structures and Algo, Linux
-





Course

- Monte Carlo with variance reduction
 - European and American Option
 - VaR and CVA
 - PCA in Finance
 - Black-Litterman model
-

After course

- You can answer interview questions like:
 1. What's the difference between PCA and Linear Regression?
 2. When do we need PCA?
 3. What kind of data do we need before PCA?
 4. How to explain the result in PCA? What does **COMPONENT** mean? What does eigenvector mean?
-

After course

- Project:
 - PCA in DAX30 replication for ETF product
 - European and American option
-

Our Student Story

Python project to crawl data

Kai

Q&A?

课程报名方式:

- <http://dataapplab.com/python>
 - 网页直接报名,我们会联络您
 - 网页在线Chat
 - 邮件报名: 发送个人信息到 DataAnalyst@DataAppLab.com
 - 电话咨询 +1 (800) 485 – 7918
 - 微信公众号 “大数据应用”
 - 加微信号私聊(微信号: data-science)
-