greatlearning
Learning for Life

DATA SCIENCE USING PYTHON Week 1

greatlearning
Learning for Life

LET'S INTRODUCE OURSELVES!!

- Work-ex

- Company

- Motivation behind joining this course

Please switch on our web camera

LEARNING OBJECTIVE OF THIS MODULE

- Basic Working proficiency in Python

- Basic Data-Manipulation using Python

- Basic Data-Visualization using Python

LET'S SET SOME GROUND RULES

- Come prepared for these sessions by watching the videos.
    - Concepts will be covered in the videos.
    - Hands-On Application will be covered in Mentor Sessions.

- Submit all assignments on time.

- Let's be punctual & respect each other's time.

**greatlearning**
Learning for Life

TRY ANSWERING THE FOLLOWING

- Who will be the winner of upcoming Cricket World Cup T-20?
- Can you predict when an employee will resign from his/her organization?
- Can you predict how many comments a user generated post on Twitter is expected to receive in the given set of hours?
- Can you predict which client will default the loan payment based on the client's spending?
- Which stock will give >30% returns in the next 6 months?

![greatlearning - Learning for Life]
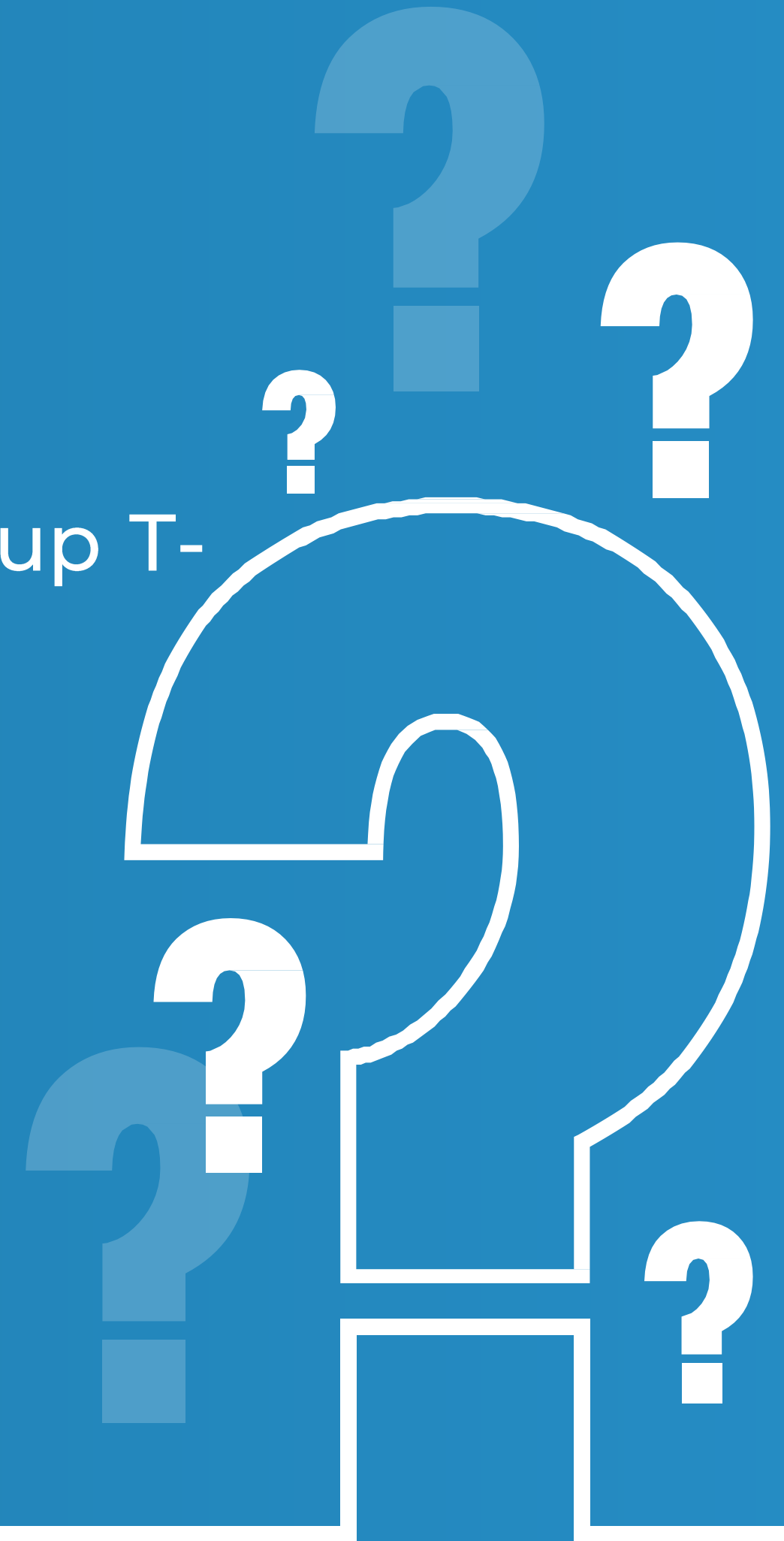
LEARNING OBJECTIVES OF THIS SESSION
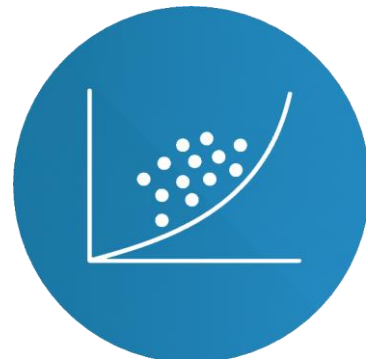
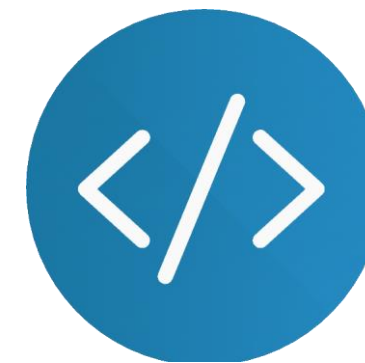- Understand the big picture of Data Science & Analytics
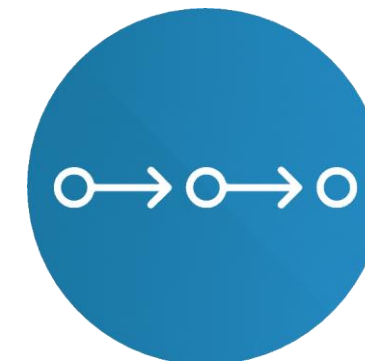
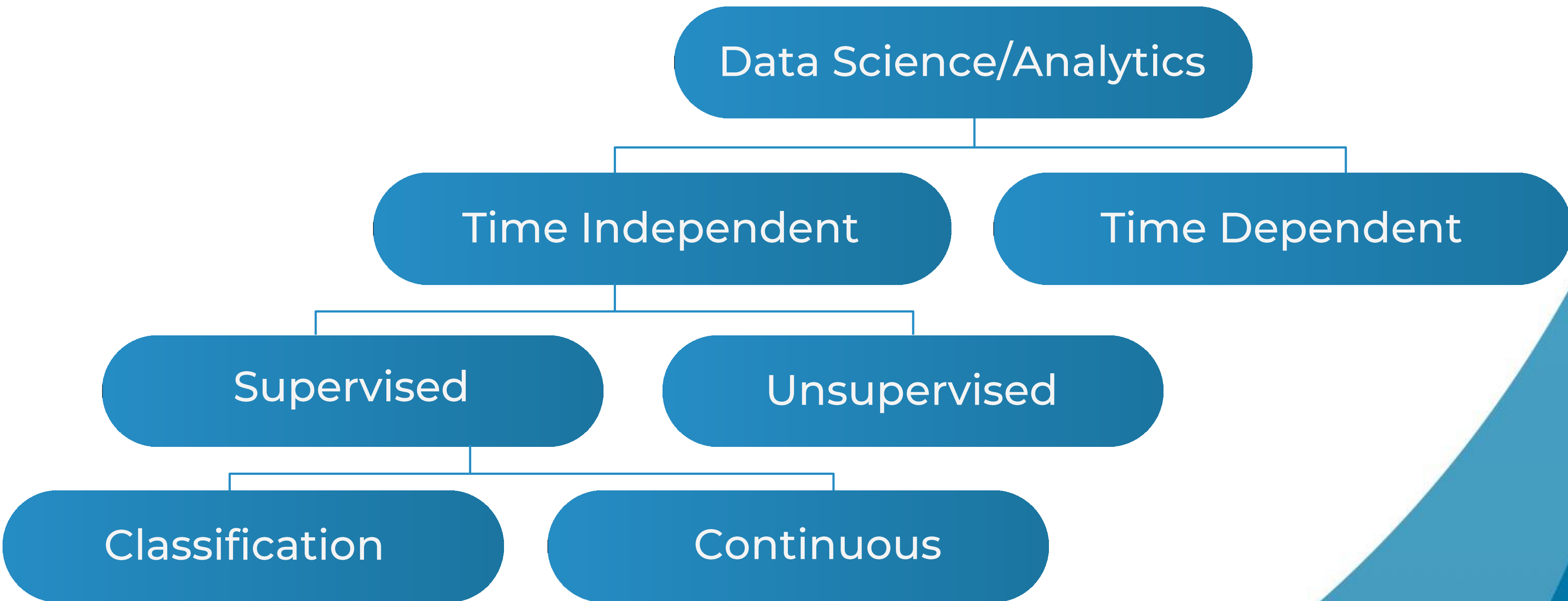- PGPDSBA Curriculum

- Introduction to Python

- Installation Steps

- Basic Operations in Python using a Case Study

- A journey of a thousand miles begins with a single step

DSBA CURRICULUM DESIGN

FOUNDATIONS

Python for Data Science(1/4)

Statistical Methods for Decision Making

CORE COURSES

Advanced Statistics

Data Mining

Predictive Modelling

Machine Learning

Time Series Forecasting

Data Visualization

SQL

DOMAIN APPLICATIONS

Financial Risk Analytics

Marketing Retail Analytics

BY THE ALUMS

> This program helped me add skills and tools to transition to analytics with 45% hike

Divya Sharma

> Make a non-techie stand as a technology specialist. Thanks to the pedagogy, content and support provided by Great Learning

Sonakshi Pattnaik

PYTHON (WHAT AND WHY ?)

- Python is the most popular programming language & choice for Data Scientist / Data Engineer across the world

- Very rich libraries & functions

- Community support

- Easy to deploy in production

- Support for all the new state of the art technologies
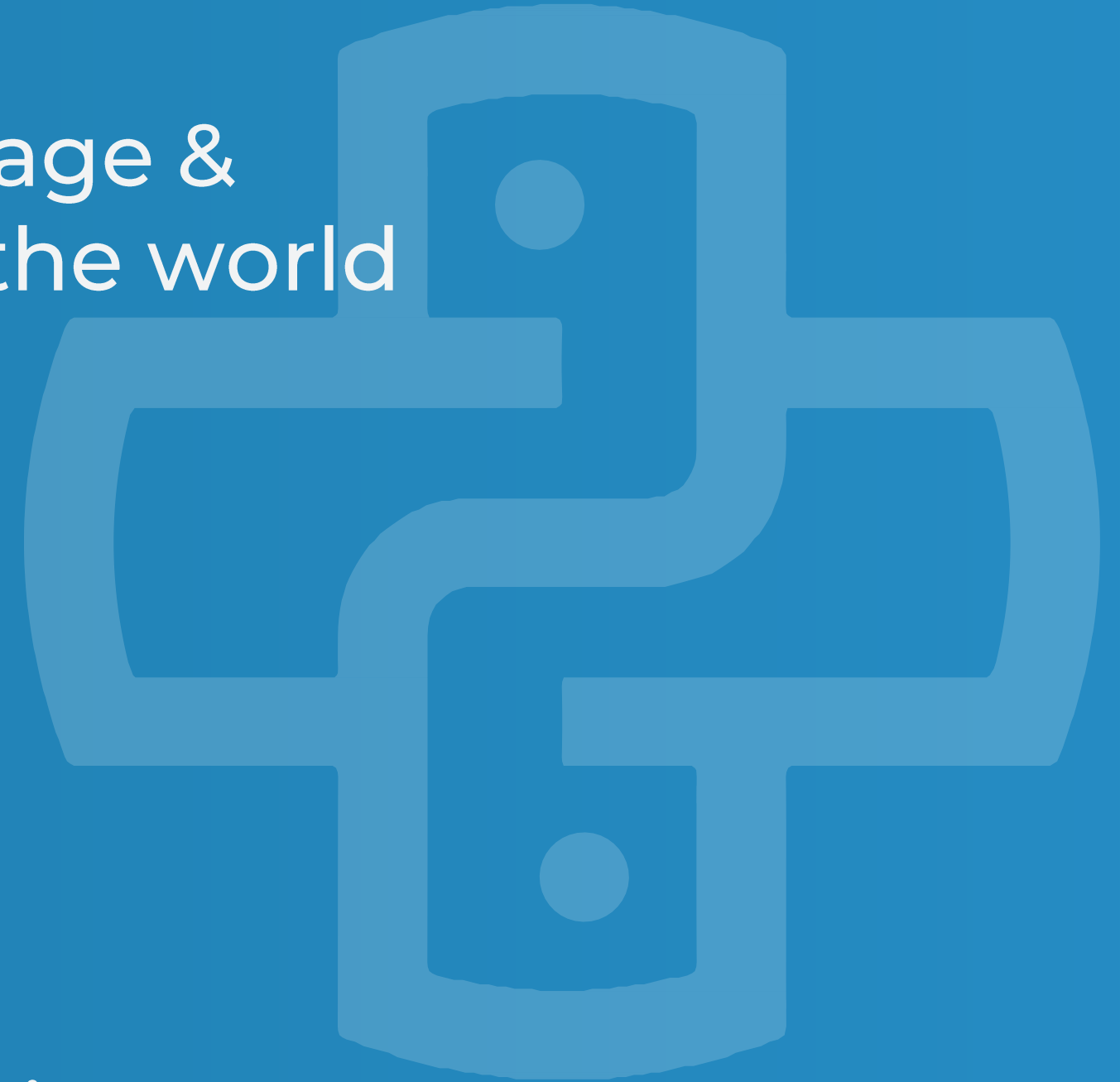
```python
import random
n = random.randint(1, 99)
guess = int(raw_input("Enter a number from 1 to 99: "))
while n != "guess":
    print
    if guess < n:
        print "guess is low"
        guess = int(raw_input("Enter a number from 1 to 99: "))
    elif guess > n:
        print "guess is high"
        guess = int(raw_input("Enter a number from 1 to 99: "))
    else:
        print "Congrats! you guessed it!"
        break
    print
```

HOW MANY OF US HAVE ALREADY INSTALLED PYTHON & JUPYTER ON THEIR SYSTEMS ?

INSTALLATION STEPS

Install using the instruction given in the below links:

1. Install Jupyter - http://jupyter.org/install
    Preferred installation method is through
    Anaconda distribution.

Install **Python 3.6 version**

2. Anaconda 5.2 For Linux Installer
    -https://www.anaconda.com/download/#linux
3. Anaconda 5.2 For macOS Installer
    - https://www.anaconda.com/download/#macos

(You need to download the version compatible with your OS)

3. Anaconda 5.2 For Windows Installer
   - https://www.anaconda.com/download/#windows

greatlearning
Learning for Life

ANACONDA NAVIGATOR

File   Help

 ANACONDA® NAVIGATOR

Sign in to Anaconda Cloud

- Home
- Environments
- Learning
- Community

Applications on   base (root)   Channels   Refresh

JupyterLab
1.0.2
An extensible environment for interactive and reproducible computing, based on the Jupyter Notebook and Architecture.

Launch

Notebook
6.0.0
Web-based, interactive computing notebook environment. Edit and run human-readable docs while describing the data analysis.
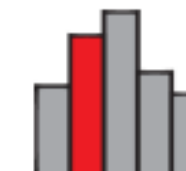
Launch

Spyder
3.3.6
Scientific PYthon Development EnviRonment. Powerful Python IDE with advanced editing, interactive testing, debugging and introspection features

Launch

Glueviz
0.13.3
Multidimensional data visualization across files. Explore relationships within and among related datasets.

Install

Orange 3
3.19.0
Component based data mining framework. Data visualization and data analysis for novice and expert. Interactive workflows

RStudio
1.1.456
A set of integrated tools designed to help you be more productive with R. Includes R essentials and notebooks.

VS Code
1.44.2
Streamlined code editor with support for development operations like debugging, task running and version control.

Documentation

Developer Blog

Let's start with Python

- Launching Jupyter Notebook

- Opening ipnyb file

- Setting Working Directory

- Changing Working Directory

- Saving ipnyb file

Basic Python Hands-on Exercise

- Data Types

- Conditional Statements and Loops

greatlearning
Learning for Life

ANY QUESTIONS

HAPPY LEARNING