

Setting up the Environment



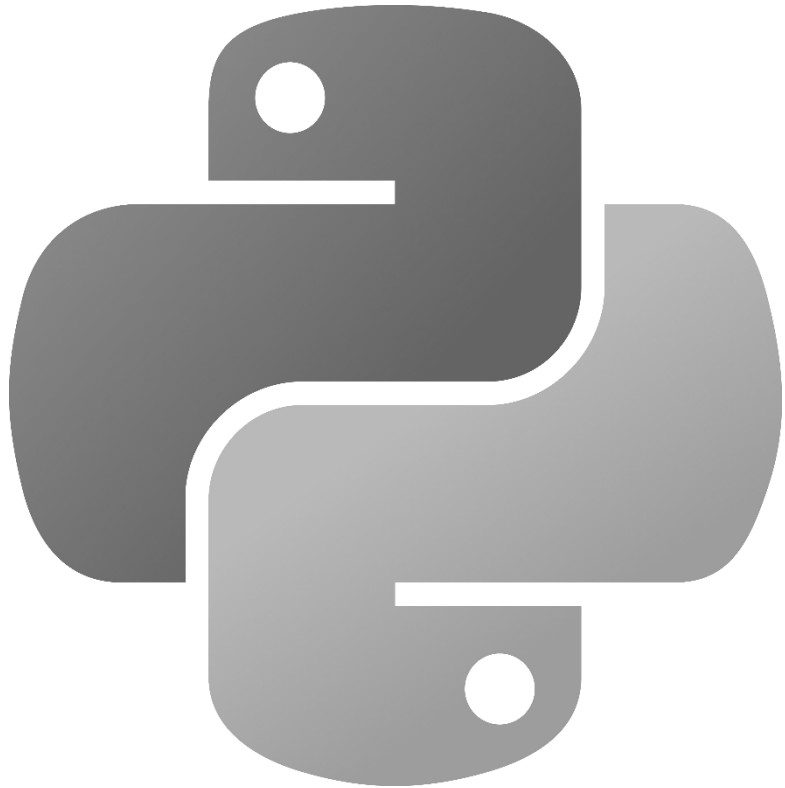
Abhishek Kumar

AUTHOR

@meabhishekkumar



Overview



Python distributions

Jupyter Notebook

Notebook basics

Data science project template

Versioning system

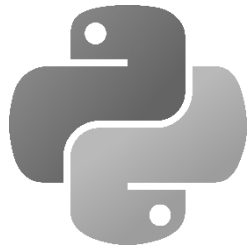
Python Distributions for Data Science



Option 1

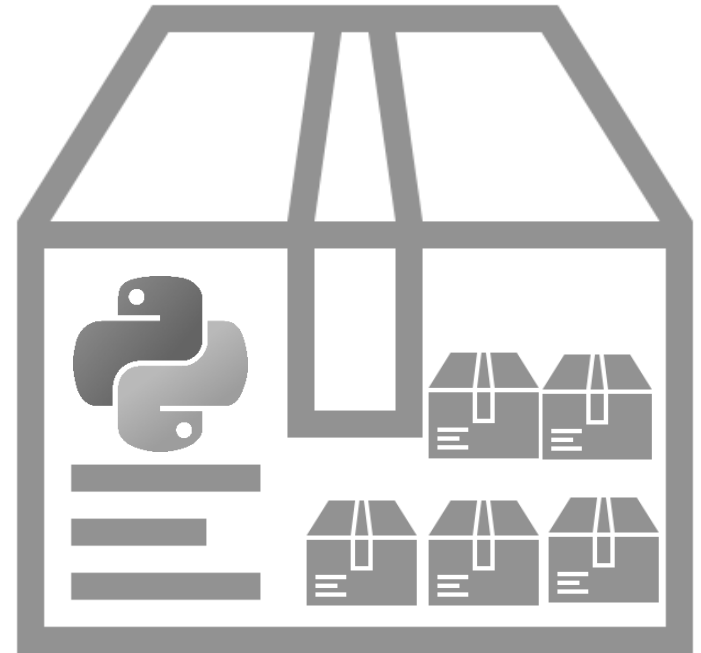


Packages



Base
Python

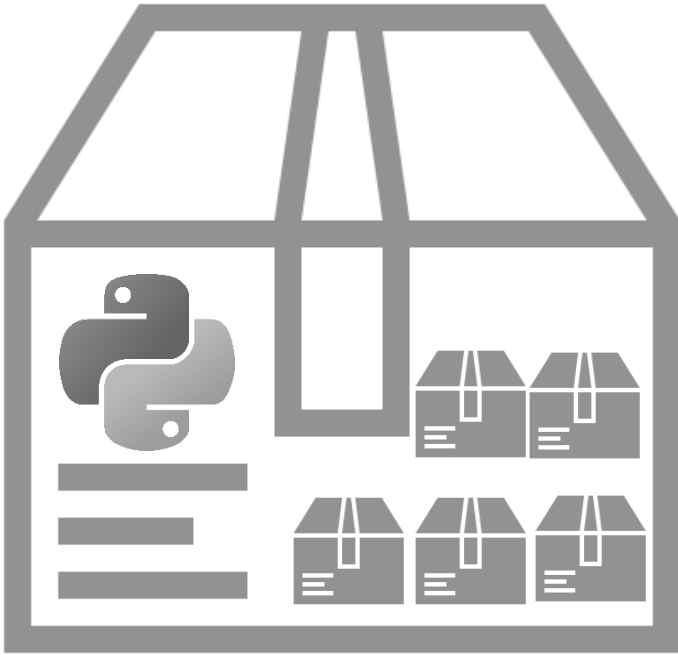
Option 2



Specialized Python
Distributions



Python Distributions for Data Science



Anaconda

Enthought canopy

Anaconda Distribution

Free version

Python 2.x or 3.x?



Python 3.x vs Python 2.x

Python 3.x

Clean & faster

Future

Python 2.x

Stable third-party packages

Better community support

Backward compatibility



Demo



Install Anaconda on local windows machine



Jupyter Notebook



Jupyter Notebook

Formerly known as IPython Notebook

Combine code blocks, human-friendly text, images, videos in single document

Run in web browsers

Support different kernels such as Python, R, Julia, Scala

Viewed with nbviewer (also in GitHub)

Export to various formats such as PDF



Demo



Setup Jupyter Notebook on local machine



Demo



Jupyter Notebook - basics



Demo



Jupyter notebook – magic functions



Data Science Project Template



Data Science Project Template

Raw Data

Processed Data

Documentation

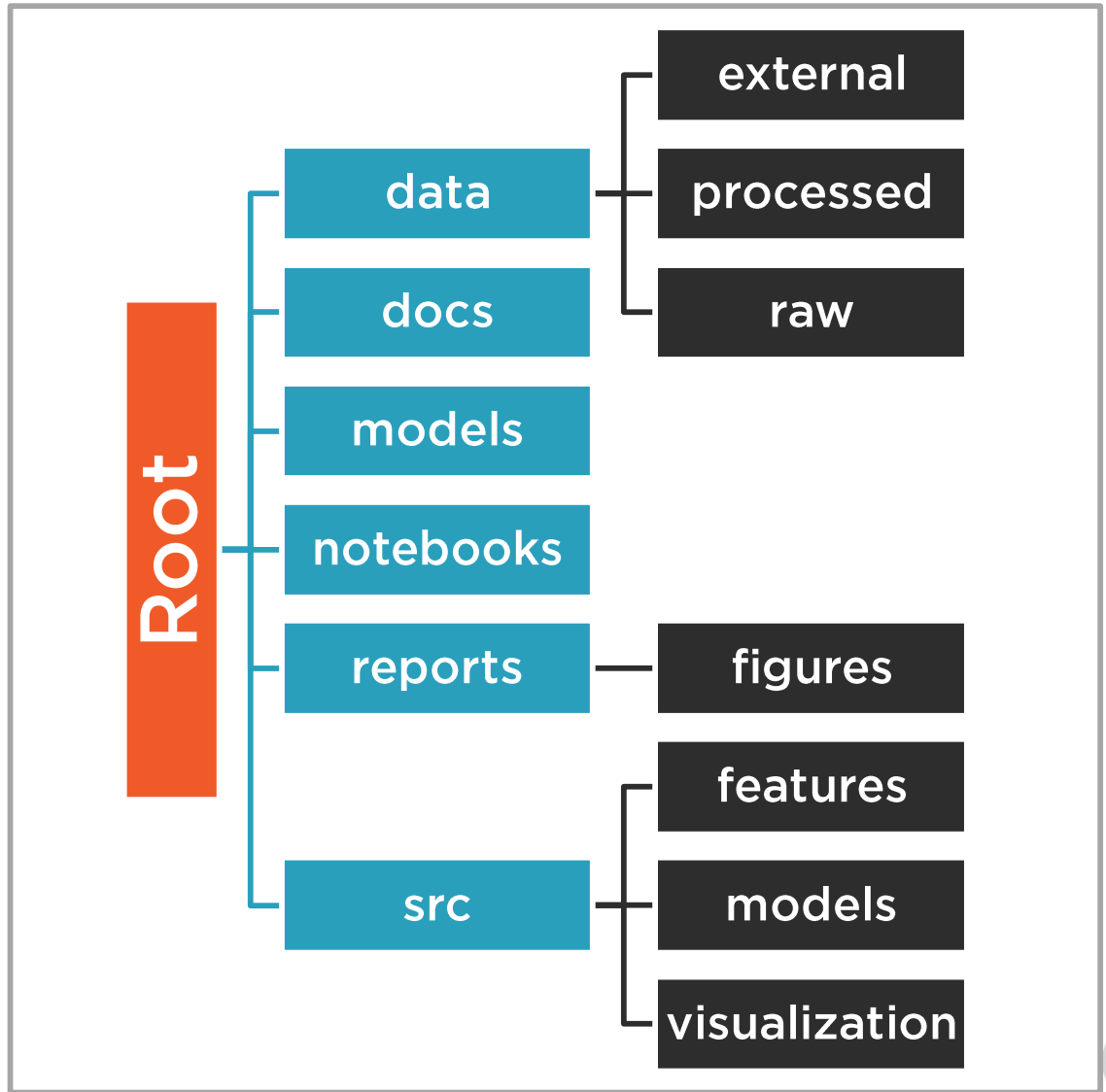
Models

Notebooks

Reports

Visualizations

References



Why Use a Template?

Consistency

Standardization

Collaboration

Reusability

Code Quality



Which Project Template to use?

Cookiecutter data science - Driven Data

Link : <https://github.com/drivendata/cookiecutter-data-science>



Demo



Setting up the cookiecutter data science project template



Versioning for Data Science Projects



Why Versioning?

Common
repository

Track changes

Suitable for
individuals too



Which Versioning System to Use?

Local

Git

Cloud

GitHub

Bitbucket



Demo



Add the project to git



Summary



Python distributions for data science

Python 2.x vs 3.x

Ananconda

Jupyter notebook

Data science project template

Versioning system

