Data Science Workshop Session 2

Ikenna Ivenso

20 June 2020

www.tqstem.org

Objectives

1. Why Python ?
2. Installing Anaconda and Jupyter
3. Useful Python Libraries
4. Some Python Resources

Why Python ?

Python is the most popular programming language for Data Science (see this survey)

- Versatile and easy to use
- Open source (high quality support for free!!!)
- Strong support for Data Science packages
- Supports procedural, functional and object oriented paradigms (best of all worlds)
- Easy to embed in applications
- Python has a vibrant global community
- Incorporates very high quality visualization capabilities
- Compatible with most database architectures (Hadoop, RDBMS, etc.)
- …many other reasons

Installing Anaconda and Jupyter

Install Anaconda:

- Go to Anaconda download page (here)
- Follow download link for your platform (Windows | macOS | Linux)
- Download Python 3.7 version (Python 2.x support has ended)

Run Jupyter:

- Windows: Start Menu | cmd | Anaconda Prompt
- macOS/Linux: run *jupyter notebook* from terminal or command line

Install Packages/Modules:

- To install a package that does not come pre-installed: *pip install [package-name]*
- To upgrade an installed package to a new version: *pip install –upgrade [package-name]*
- To install a specific version of a package: *pip install [package-name==version-number]*

Useful Libraries: *Pandas*

- Pandas is mainly used for data analysis and manipulation
- It is particularly useful for tabular and time-series data
- Quick demo…
- See this tutorial

Useful Libraries: *NumPy*

- NumPy is mainly used for working with arrays and performing high level math operations
- It is particularly useful for manipulating large arrays of multi-dimensional data
- Quick demo…
- See this tutorial

Useful Libraries: *Matplotlib*

- Matplotlib is mainly used for high quality plotting and data visualization
- It is particularly useful for representing data visually using different plot types
- Quick demo…
- See this tutorial

Getting Started in Python

It would take a very long time to cover all the important topics in Python

Here are some resources to help get started with Python:

- YouTube playlist: Python for Beginners
- Interactive Python tutorials (navigation links on the left pane)

THANK YOU!
Any Questions

www.tqstem.org