Data Analysis with Python

This course aims to familiarize the students with the most popular scripting language – Python and data science must know tools – Numpy,Pandas, Matplotlib and Seaborn. The course will start with learning the syntax of functional aspect of python and how you can develop scripts rapidly. It then goes to the data science must know tools and augmenting your python skills with shell scripting. This course aims to teach the fundamentals of data science. As a data scientist, you are not expected to write machine learning code everyday. A lot of work focus around gathering data, cleaning data and coming up with useful analysis and visualizing them. This course will prove to a strong base for students who want to take advance courses in data science.

Syllabus :

Lecture	Topic
1	Introduction to Python: • How to move away from OOP and develop rapid scripts. • Numeric Type • String Fundamentals • Understanding Modules • Understanding the tools for development
2	Python Data Structures: • List, Dictionary and Tuples
3	Statement and Syntax: • Python Statements and assignments • Conditionals Constructs • Iterators
4	Functions and Generators • Basics of Functions • Variables Scopes and Arguments • Advance Function Topics
5	• Comprehensions and Generators • Exception Handling
6	Numpy
7	**Midterm**
8	Pandas

	• Introduction to pandas data structure • Data loading and storing
9	Advance Pandas concept - I • Cleaning, Transform and merging data. • Data Aggregation
10	Advance Pandas concept – II • Data Grouping • Working with time series data
11	Introduction to Matplotlib
12	Introduction to Seaborn
13	Shell Scripting : • ZSH shell • AWK,SED commands • Cron Jobs.
14	Analyzing example data sets.
15	**Final Exam**

Reference Books :

1. Learning Python, 5th Edition
   By Mark Lutz
   Publisher: O'Reilly Media
2. Python for Data Analysis
   By Wes McKinney
   Publisher: O'Reilly Media