CSEL 562: Python Programming for Data Analytics

L	T	P	C
3	0	0	3

_Pre-requisite__:_

- _Programming knowledge in any OOlanguage._

_Objectives__:_

- _Understanding the basic concepts ofPython._
- _Preparing and pre-processingdata._
- _Understanding the data aggregation and groupingconcepts._
- _Leveraging webscraping._
- _Visualizing the results of analyticseffectively._

Outcomes:

- _Ability to do a data analytics using python programs._

Module - I: Python Concepts, DataStructures,Classes   (9hrs)

Interpreter – Program Execution – Statements – Expressions – Flow Controls – Functions - Numeric Types – Sequences-Strings,Tuples,Listsand-ClassDefinition–Constructors–Inheritance–Overloading–Text&Binary   Files - Reading andWriting.

Module - II:DataWrangling   (9hrs)

Combining and Merging DataSets – Reshaping and Pivoting – Data Transformation – String Manipulation, Regular Expressions.

Module - III: Data Aggregation, Group Operations,Timeseries   (9hrs)

GoupBy Mechanics – Data Aggregation – GroupWise Operations and Transformations – Pivot Tables and Cross Tabulations – Date and Time Date Type tools – Time Series Basics – Data Ranges, Frequencies and Shifting.

Module - IV:WebScraping   (9hrs)

Data Acquisition by Scraping web applications –Submitting a form - Fetching web pages – Downloading web pages through form submission – CSS Selectors.

Module - V: VisualizationinPython   (9hrs)

Matplotlibpackage–PlottingGraphs–ControllingGraph–AddingText–MoreGraphTypes–Gettingandsetting values –Patches.

Text Book(s):

1. _Mark Lutz, Programming Python, O'Reilly Media, Fourth edition,2010._
2. _Mark Lutz, Learning Python, O'Reilly Media, Fifth Edition,2013._
3. _Tim Hall and J-P Stacey, Python 3 for Absolute Beginners, Apress, First Edition,2009._
4. _Magnus Lie Hetland, Beginning Python: From Novice to Professional, Apress, SecondEdition, 2005._
5. _Shai Vaingast, Beginning Python Visualization Crafting Visual Transformation Scripts, Apress, Second Edition,2014._

Reference Book(s):

1. _Wes Mc Kinney, Python for Data Analysis, O'Reilly Media,2012._
2. _White, Hadoop: The Definitive Guide, Third Edition - O'Reilly ,2012._
3. _BrandonRhodesandJohnGoerzen,FoundationsofPythonNetworkProgramming:TheComprehensive Guide to Building Network Applications with Python, Apress, Second Edition,2010._
4. _http://blog.matthewrathbone.com/2013/11/17/python-map-reduce-on-hadoop---a-beginnerstutorial.html_
5. _http://www.michael-noll.com/tutorials/writing-an-hadoop-mapreduce-program-in-python/_
6. _http://allthingshadoop.com/category/python/_