



# SYMBIOSIS SKILLS AND PROFESSIONAL UNIVERSITY

(Established under Govt. of Maharashtra Act No. XXXVII 2017 dated 3rd May 2017)

Kiwale, Adjoining Mumbai - Pune Express Highway, Pune 412 101,

State - Maharashtra, INDIA. | <http://www.ssou.ac.in>

<b>School Name</b>	School of Data Science
<b>Program Name</b>	Certificate Course in Data Associate
<b>Duration</b>	3 Months (300 Hours)
<b>Occupation &amp; Description of Role</b>	Python Developer, MySQL Developer, Data Engineer, Data Analyst, Data Visualizer, Data Associate.
<b>Eligibility (Educational)</b>	BE/B.Tech (Engg, CSE), MS (Stat, Math), BSc (Stat, Math). Master's Degree in Science/ Technology/ Mathematics/ Statistics <b>and</b> Mathematics subjects studied in 12th Standard is a compulsory. Candidate need to qualify the Entrance Test conducted by the Competent Authority.
<b>Pre-requisite</b>	<ul style="list-style-type: none"><li>• Understanding of Math / Stats at 12th standard level.</li><li>• Knowledge of basic computer science principles and skills.</li><li>• Basic programming knowledge – writing and executing code in any language</li><li>• Laptop with minimum 8 GB RAM.</li></ul>
<b>Skills Students Acquire at end of the course</b>	<i>Generic:</i> Communication skills (spoken, basic), Presentation Skills <i>Technical:</i> Collate and Analyze Data, Use Python for Programming, Use Tableau, MySQL as part of data analysis, Hadoop Eco System, HIVE, Apache Spark. <i>Professional:</i> Team Work
<b>Course Objective</b>	<ul style="list-style-type: none"><li>• To develops skills of Python programming.</li><li>• To implement a different types of wide variety of python modules in various applications.</li><li>• To develop data handling and data analysing skills.</li><li>• To handle big data.</li></ul>

	<ul style="list-style-type: none"> <li>● To give each student a realistic perspective of work and work expectations, to help formulate problem solving skills, to guide students in making appropriate and responsible decisions.</li> <li>● To create a desire to fulfil individual goals, and to educate students about unproductive thinking, self-defeating emotional impulses, and self-defeating behaviour.</li> </ul>
<b>Course Learning Outcomes</b>	<p>After completing this programme, participants will be able to:</p> <ul style="list-style-type: none"> <li>● Work with Python scripting elements such as variables, flow control structures, functions, file handling.</li> <li>● Determine the methods to create and manipulate Python programs by utilizing the data structures like lists, dictionaries, tuples and sets.</li> <li>● Apply fundamentals of some of the most widely used Python packages; including Numpy, Pandas and Matplotlib, then apply them to Data Analysis and Data Visualization projects.</li> <li>● Execute various basic and advance SQL queries.</li> <li>● Apply various Normalization techniques, MySQL programming using concept.</li> <li>● Creating Crosstabs, Charts, Maps and Dashboards using Tableau.</li> <li>● Develop Big Data Solutions using Hadoop Eco System.</li> <li>● Querying big data with Hive.</li> <li>● Managing big data with Apache Spark.</li> <li>● Understanding the significance Personality Development and Presentation skills</li> <li>● Able to build right Attitude</li> <li>● Adopt the Interview skills and self-introduction</li> </ul>
<b>Module list</b>	<ul style="list-style-type: none"> <li>● Python for Data Analysis</li> <li>● Managing with Data</li> <li>● Analyzing Data from Disparate Sources</li> <li>● IDSC</li> <li>● Project</li> </ul>

## Suggested learning Resources (but not limited)

Sr. No	Title of the Book / Link	Author / Webiste	Edition / volume	Text (T) Reference (R)
1.	Hadoop Illuminated	Mark Kerzner & Sujee Maniyam	2014	T1
2	Data Mining: Practical Machine Learning Tools and Techniques	Ian H. Witten & Eibe Frank	2005	T2
3	Big Data Now	O'Reilly Media	2012	R1
4	Head-First Python	Paul Barry	2 <sup>nd</sup> Edition	R2
5	Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale	<i>Tom White</i>	4 <sup>th</sup> Edition	R3
6	Big Data Analytics with Hadoop 3.0	<i>Sridhar Alla</i>		R4
7	CASE STUDY OF HIVE USING HADOOP	Sai Prasad Potharaju, Shanmuk Srinivas A, Ravi Kumar Tirandasu	<a href="https://www.researchgate.net/publication/314724458_CASE_STUDY_OF_HIVE_USING_HADOOP/link/58c5114192851c0ccbf7fb57/download">https://www.researchgate.net/publication/314724458_CASE_STUDY_OF_HIVE_USING_HADOOP/link/58c5114192851c0ccbf7fb57/download</a>	Case Study 1
8	<a href="https://hadoopbaseblog.wordpress.com/2017/09/04/hive-ddl-loading-data-into-hive-tables/">https://hadoopbaseblog.wordpress.com/2017/09/04/hive-ddl-loading-data-into-hive-tables/</a>			Online