

## Course 2 - Program Description Document

<b>Course Name</b>	<b>Career Back 2 Women (CB2Women)</b>
<b>Course Name as on Certificate</b>	<b>Certified Professional in Basics of Data Science and Big Data Analytics</b>
<b>Certificate Type</b>	Certificate of Completion by IIT-MADRAS
<b>Certificate Issued by</b>	IIT MADRAS
<b>Course Objectives</b>	<p>Course is targeted at learners who are interested in learning and understanding data science and what big data is all about in sufficient depth and breadth. The course will provide an overview of how to pose meaningful data analytics problems for real-life applications. At the end of the course, participants will develop a structured thinking approach to transition from data to data science problem definitions and also learn to address challenges in handling Big Data.</p> <ul style="list-style-type: none"> <li>• Introduce the participants to Python – an easy to use tool for high level data analytics</li> <li>• Introduce the participants to a comprehensive overview of linear algebra, statistics and optimization concepts – critical concepts for the understanding of data science</li> <li>• Introduce the participants to in-depth explanation of data science algorithms – supported by hands-on work in Python from an application viewpoint</li> <li>• Introduce participants to visualization</li> <li>• Introduce the participants to the field of Big Data- background and key concepts</li> <li>• Introduce the participants to real-life applications of data science – a case study approach.</li> </ul>
<b>Eligibility</b>	<ul style="list-style-type: none"> <li>• For Indian Participants - Graduates or Diploma Holders (10+2+3) from a recognized university (UGC/AICTE/DEC/AIU/State Government) in any discipline.</li> <li>• For International Participants - Graduation or equivalent degree from any recognized University or Institution in their respective country.</li> </ul>
<b>Pre Requisites</b>	Basic understanding of technology, networks and security, while not mandatory, will be an added advantage.
<b>Target Segment</b>	This programme is aimed at the women professionals who had a job in IT industry, but had to leave the same for some reason including pregnancy or taking care of aged parents / in-laws / sick people etc. to return to IT career with a bang. FISST has database with over 4 lac women, in various domain and expressed willingness to get trained into new areas to return to work & earn.
<b>Course Content</b>	<b>See Enclosed Programme details – as Annexure 2</b>
<b>Pedagogy</b>	The primary method of instruction will be through LIVE lectures that will be delivered online via internet to participant desktops/laptops or classrooms. The lectures will be delivered by eminent academicians and practicing industry experts. The programme will be primarily taught through a combination of lectures, discussions, exercises and labs. All enrolled students will be provided access to our FISST Whizard Cloud Campus through which students may access other learning aids, reference materials, assessments and assignments as appropriate. Throughout the duration of the course, students will have the flexibility to reach out to the Professors, real time during the class or offline via the FISST Whizard Cloud Campus to raise questions and clear their doubts.
<b>Assessment</b>	There are periodic evaluation components built in as a part of the program. These maybe in the form of a quiz, assignment or other objective/subjective assessments as relevant and applicable to the program. A minimum of 70% attendance to the LIVE lectures, is a prerequisite for the successful completion of this program. Participants who satisfy the attendance criteria and successfully clear the evaluation components will be awarded a certificate of completion.

<b>Programme Faculty</b>	<p><b>Programme Director CB2Women:</b> Mr. Mohan Ram C from FISST</p> <p>Mohan has nearly 33 years of professional experience after an M.Tech from IIT-Roorkee, as IT leader specializing in Cyber Security and related physical surveillance for critical infrastructure including refinery, nuclear power plants and mission critical IT infrastructure etc. Mohan is currently pioneering Cyber Education space in India to create awareness and fill the gap in skills to tackle potential damages due to cybercrimes in partnership with leading academic institutions across India.</p> <p><b>Lead Academic Faculty Members:</b> Professor Raghunathan Rengasamy– IIT M</p> <p>And other industry experts from a pool of consultants / experts with GITAA.</p>		
<b>Duration</b>	<p>Live delivery (Virtual) by instructors with Assignments Basic course – 40 hours (10 weeks x 2 hrs per day on Sat &amp; Sun) TOTAL = 40 Hours</p>		
<b>Class Schedule</b>	<p>Twice a week on Saturdays &amp; Sundays from for 10 weeks (3 months – Basic)</p>		
<b>Programme Highlights/USPs</b>	<p><b>Course Benefits to Participants</b></p> <p>On successful completion of the programme, you will be able to</p> <ul style="list-style-type: none"> <li>The course will provide an overview of how to pose meaningful data analytics problems for real-life applications. At the end of the course, participants will develop a structured thinking approach to transition from data to data science problem definitions and also learn to address challenges in handling Big Data.</li> </ul> <p>Other benefits to participants include</p> <ul style="list-style-type: none"> <li>Opportunity to earn a Certificate from IIT Madras.</li> <li>Lectures imparted by eminent academicians and practicing industry experts.</li> <li>Get complete exposure to contemporary and most sought after skills related to Data Science &amp; Big Data Analytics</li> <li>Fully Online Course with LIVE online interactive lectures that provides a “real” classroom experience in a “virtual” environment. No isolated learning experience.</li> <li>Seamless technology that can transmit lecture videos effectively at home broadband connection of 512 kbps.</li> <li>User friendly and easy to use technology interface. No expensive and time consuming software/hardware installations required at your end.</li> <li>Virtual classrooms that allow for active interactions with other fellow students and faculty.</li> <li>Convenient weekend schedules</li> <li>In the event that students miss attending the LIVE lecture on the Virtual Classroom for some reason, students will be granted access to the recorded sessions for a specified number of days/times.</li> <li>FISST Whizard Cloud Campus – Students on our virtual social learning platform are provided access to course presentations, projects, case studies, assignments and other reference materials as applicable for specified courses. Students can raise questions and doubts either real time during the live class or offline through the Cloud Campus.</li> <li>Learn from Anywhere – No need to travel to an institute or training center. Learning continues even if you are traveling or not available at any specific location. You may also learn from the comfort of your home.</li> </ul>		
<b>Total Fees</b>		<b>Total Fees (Rs.)</b>	<b>Institute Revenue Share (Rs.)</b>
	Total Programme Fee	Rs. 35,000/- + GST	

## ANNEXURE 1

Proposed Course outline / programme / plan

### BASIC TRACK – 40 hours

#### **First module: Python programming**

- Why Python for data science?
- Opening, creating and managing Python IDE
- Basics of programming
- Variables
- Operators
- Data types
- Data structures – Lists, Tuples, Dictionary
- Control structures in Python
- Function files in Python
- Essential Python Library for data science with hands-on labs
- Numpy (Numeric Python)
- Pandas
- Matplotlib
- Object oriented programming

#### **Second module: Introduction to statistical modelling**

- Random variables, expectation
- Continuous and discrete random variables and their distributions (Poisson, Binomial, Normal and its derivatives), statistical intervals
- Bayes' theorem, independent events
- Introduction to Bayesian inference
- Hands-on session in Python through examples

#### **Third module: Data preparation**

- Data preprocessing
- Treating outliers and missing values
- Exploratory Data analysis
- Correlation
- Feature Extraction
- Hands-on session in Python through examples

#### **Fourth module: Framework for solving Data Science Problems**

- Typology of Data Science problems
- Introduction to structured approaches to solving data science problems
- Imputation example to explain the structured approach

### **Fifth module: Visualization**

- Basics of visualization
- Introduction and use of the base graphics package of Python with some examples
- Elements of the Seaborn and Matplotlib packages - aesthetic mapping, geometry, scales, stat transformations etc.
- Advanced visualization

### **Sixth module: Introduction to big data**

- Introduction to Big Data
- Characteristics of Big Data
- Challenges with Big Data
- Big Data Frameworks
- Installing and Configuring Python, Hadoop, Spark and Jupyter
- What and Why of Distributed Systems
- Distributed File System
- Distributed Programming Model
- Parallel Processing explained with WordCount
- Concept of Cloud Computing
- Big Data and Cloud Computing – Benefits
- Introduction to Hadoop
- How MapReduce works
- Parallelism in MapReduce
- Example: K means Clustering – Sequential and with MapReduce
- When does MapReduce work and Why? Comparison among Algorithms
- Implementation in Python – Regular and Spark Version of KMeans