

DATA SCIENCE PRODEGREE



Knowledge Partner:



Global Leader in Digitally-Powered
Business Process Management & Services



TOP4
ANALYTICS TRAINING
INSTITUTE IN INDIA

Leading Institute in
SAS & Python
Analytics Vidhya

BIG DATA  2016
TRAINING INSTITUTE

INDUSTRY LANDSCAPE



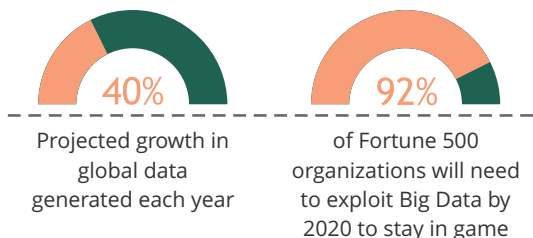
WHAT IS DATA ANALYTICS?

The scientific process of transforming data into insights for making better decisions and offering new opportunities for a competitive advantage

Why is Data Analytics Important?

It helps organizations harness their data and use it to identify new opportunities, leading to smarter business moves, more efficient operations, higher profits and happier customers.

THE SKILLS GAP

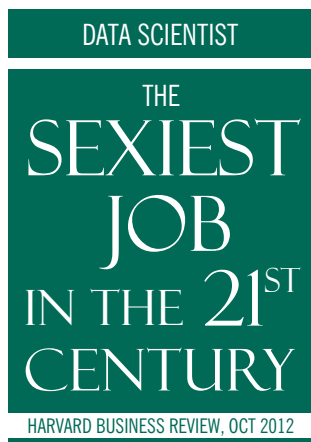
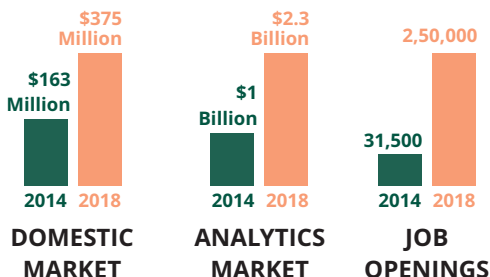


IN DEMAND SKILL SETS

- In Demand Skill Sets
- Predictive Analytics
- Data Analysis & Management
- Data Visualization
- Business Intelligence
- SAS Programming
- New Tools like R, Python

GROWING DEMAND

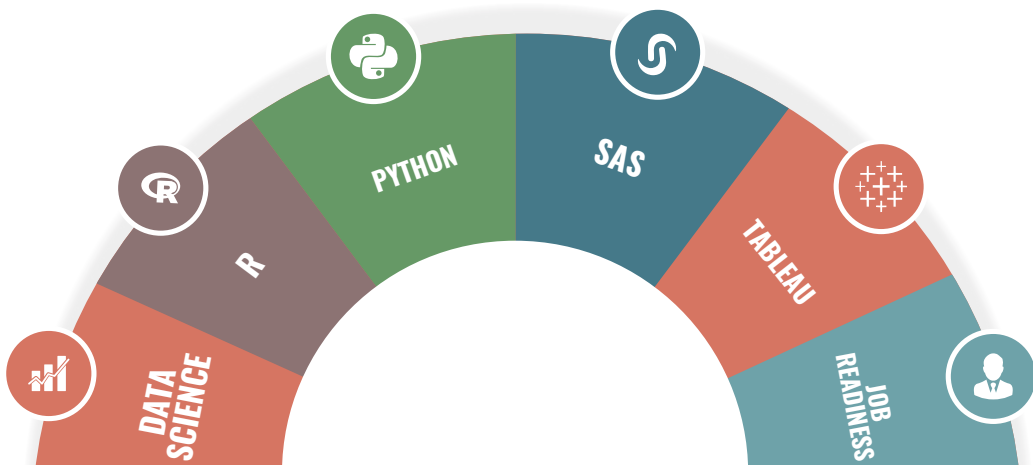
Projected Demand for Analytics Professionals



EMPLOYMENT LANDSCAPE

- Genpact
- Infosys
- Target
- HSBC
- Cap Gemini
- Accenture
- Wipro Analytics
- Cognizant
- Fractal
- Citi Bank Analytics
- EXL
- Mu Sigma
- HCL
- Mindtree
- Latent View
- IBM

OVERVIEW OF PROGRAM



PROGRAM AVAILABLE IN CLASSROOM AND ONLINE DELIVERY FORMAT

CURRICULUM

INTRODUCTION

BATCH LAUNCH	Intro to Program Curriculum Overview Learning Methodology Guest Lecture
ALL ABOUT DATA	Data Variables Data Types Measures of Central Tendency in Data Understanding Skewness in Data Measures of Dispersion Data Distribution
ANOVA/ REGRESSION ANALYSIS	Analysis of Variance and Covariance One way analysis of variance Assumption of ANOVA Statistics associated with one way analysis of variance Interpreting the ANOVA Results Two way analysis of variance Interpreting the ANOVA Results Analysis of Covariance Examine Regression Results What is Regression Analysis Linear and Logistic Regression Statistics Associated with Regression
PREDICTIVE MODELLING	Decision Trees and Neural Networks Introduction to Predictive Modelling with Decision Trees Assumptions Formulate the Model Estimate the Parameters Check the Prediction Accuracy
TREE AND BAYESIAN NETWORK MODELS	Decision Trees, Bagging Random Forests, Boosted Trees Bayesian Classification Models
NEURAL NETWORKS	Perceptron, MLP, Back Propagation Revision of Key Concepts

R

R BASICS	R Base Software Understanding CRAN RStudio The IDE Basic Building Blocks in R Sequence of Numbers in R Understanding Vectors in R Basic Operations Operators and Types
R FUNCTIONS	Handling Missing Values in R Subsetting Vectors in R Matrices and Data Frames in R Logical Statements in R Lapply, sapply, vapply and tapply Functions
LINEAR REGRESSION THEORY - R	Covariance and Correlation Multivariate Analysis Assumptions of Linearity Hypothesis Testing Limitations of Regression
BUSINESS CASE: MANAGING CREDIT RISK	Business Case : Managing Credit Risk Meaning of Credit Risk Impact of Credit Default Sources of Data for Managing Risk Understanding Loss Given Default Understanding Default
LOSS GIVEN DEFAULT LINEAR REGRESSION R	Loss Given Default Linear Regression R Extract Data in R Univariate Analysis of Data Apply Data Transformations Bivariate Analysis of Data Identify Multicollinearity in Data Treatment on Data Identify Heteroscedasticity Discuss what could be the Reason for Heteroscedasticity Modelling of Data Variable Significance Identification Model Significance Test Predict using Testing Data Set Validate the Model Performance
LOGISTIC REGRESSION THEORY - R	Reason for Logistic Regression The Logistic Transform Logistic Regression Modelling Model Optimisation Understanding ROC Curve
PROJECT 1	Project 1 - Default Modelling using Logistic Regression in R
SUPPORT VECTOR MACHINES (THEORY)	Introduction to SVM Classification as a Hyper Plane Location Problem Motivation for Linear Support Vectors SVM as Quadratic Optimization Problem Non Linear SVM Introduction to Kernel Functions
PROJECT 2	Project 2 - Default Modelling using SVM in R
DECISION TREES	Introduction to Decision Trees Theory of Entropy & Information Gain Stopping Rules Overfitting Problem Cross Validations for Overfitting Problem Pruning as a Solution for Overfitting Ensemble Learning Notion Concept of Bootstrap Aggregation Concept of Random Forest
BUSINESS CASE	Business Case : Intrusion Detection in IT Network Meaning of Intrusion in IT Cost of Intrusion Meaning of Intrusion Detection System
PROJECT 3	Project 3 - Network Intrusion Detection using Decision Tree & Ensemble Learning in R
GUEST LECTURE	Industry View from Expert Refresher on R Open House

PYTHON

PYTHON BASICS	What is Python? Installing Anaconda Understanding the Spyder Integrated Development Environment (IDE) Lists, tuples, dictionaries, variables
DATA STRUCTURES IN PYTHON USED FOR DATA ANALYSIS	Intro to Numpy Arrays Creating ndarrays Indexing Data Processing using Arrays File Input and Output Getting Started with Pandas

CURRICULUM

DATA FRAME MANIPULATION	Data Acquisition(Import & Export) Indexing Selection and Filtering Sorting & Summarizing Descriptive Statistics Combining and Merging Data Frames Removing Duplicates Discretization and Binning String Manipulation PLUS: Project Work on Python
OTHER PREDICTIVE MODELLING TOOLS	Intro to Machine Learning Random Forests Sklearn Library & Statsmodels
PROJECT 4	Project 4 - Default Modelling using Logistic Regression in Python
PROJECT 5	Project 5 - Credit Risk Analytics using SVM in Python
PROJECT 6	Project 6 - Intrusion Detection using Decision Trees & Ensemble Learning in Python

SAS

INTRODUCTION TO SAS AND SAS PROGRAMS	What is SAS? Key Features Submitting a SAS Program SAS Program Syntax Examining SAS Datasets Accessing SAS Libraries Sorting and Grouping Reporting Data Using SAS Formats
READING AND MANIPULATING DATA	Reading SAS Datasets Reading Excel Data Reading Raw Files Reading Database Data Creating Summary Reports Combining Datasets
DATA TRANSFORMATIONS	Writing Observations Writing to Multiple Datasets Accumulating Total Creating Accumulating Total for a Group of Data Data Transformations
MACROS	Introduction to Macro Variables Automatic Macro Variables User Defined Macro Variables Macro Variable Reference Defining and Calling Macros Macro Parameters Global and Local Symbol Table Creating Macro Variables in the Data Step
SQL	Introduction to SQL How Does RDBMS Work? SQL Procedures Specifying Columns Specifying Rows Presenting Data Summarizing Data Writing Join Queries using SQL Working with Subqueries, Indexes and Views Set Operators Creating Tables and Views using Proc SQL
PROJECT 7	Project 7 - Store Data Analytics in SAS








TABLEAU

TABLEAU BASIC	Introduction to Visualization Working with Tableau Visualization in Depth Data Organisation Advanced Visualization Mapping Enterprise Dashboards Data Presentation
BEST PRACTICES FOR DASHBOARDING AND REPORTING AND CASE STUDY	Have a Methodology Know Your Audience Define Resulting Actions Classify Your Dashboard Profile Your Data Use Visual Features Properly Design Iteratively
INTRODUCTION TO THE GROUP PROJECT	Choice of three projects on various domains

JOB READINESS

RESUME BUILDING AND INTERVIEW PREP	Resume Building Personal Branding Interview Skills
1:1 MOCK INTERVIEWS	1:1 Mock Interviews with Industry Veterans to Clear the Technical Round of Interviews to Give You Confidence to Face Real World Scenarios

HANDS-ON PROJECTS

 DEFAULT MODELLING USING LOGISTIC REGRESSION IN R	 DEFAULT MODELLING USING SVM IN R	 NETWORK INTRUSION DETECTION USING DECISION TREE & ENSEMBLE LEARNING IN R
 DEFAULT MODELLING USING LOGISTIC REGRESSION IN PYTHON	 CREDIT RISK ANALYTICS USING SVM IN PYTHON	 INTRUSION DETECTION USING DECISION TREES & ENSEMBLE LEARNING IN PYTHON
 STORE DATA ANALYTICS IN SAS		

KEY HIGHLIGHTS



COMPREHENSIVE COVERAGE

The 180-hour training program provides comprehensive knowledge of Data Analysis and Statistics, along with business perspectives and cutting-edge practices using SAS, R, Python and Tableau to ensure you enter the work force as well-rounded professionals.

ENDORSED BY GENPACT

The program is co-created with Genpact as the Knowledge Partner and comes with a cutting edge industry-aligned curriculum that is aligned as per Genpact's exacting requirements.

GENPACT



EXPERIENTIAL LEARNING

We believe in 'Learning by Doing' and place utmost importance to practical understanding of the subject matter. You will spend 50 hours of this program getting hands-on with industry projects and build a portfolio of demonstrable work.

MENTORSHIP

Industry experts from leading companies advise and mentor students in their journey towards job-readiness. We have a dedicated mentor assigned to each student who you can approach at any time to clear any doubts about the industry or your career prospects.



24/7 LEARNING

Our state of the art online portal provides 24/7 access to your study material, learning aids and tests. Stay in touch with students and faculty for continued learning and support.

DELIVERY MODES TO CHOOSE FROM



ONLINE DELIVERY

Live Instructor-led Virtual Classes with expert faculty for real-time learning as per your convenience



SELF PACED INSTRUCTOR VIDEOS

Active, self-paced, data-driven learning through HD videos

FACULTY

ARUN UPADHYAY

Arun has over 14 years experience in IT and has conducted SAS training for Infosys, Wipro, IBM, Genpact, ICICI Bank, Reliance Mutual Fund. He is a certified, accredited IT professional who has successfully trained more than 10,000 students in different technologies like SAS and R. He has cleared many Microsoft international certifications such as MCAD, MCPD, MCTS etc. and is a Microsoft-certified trainer.

YOGESH PARTE

Yogesh is a research engineer with over 14 years of experience in algorithmic development and PoC demonstration using C/C++, Python and R. He is the Founder of Y P Consulting Services, which specializes in innovation engineering and technology applications. He holds a PhD. in Applied Mathematics from University of Paul Sabatier, France and has won over 30 awards for academic excellence.

MOHAN RAI

Mohan has 10+ years of experience in Core Analytics (Sales & IT). Mohan is a Director for S & R Analytics involved in Delivery of Analytics Consulting /Training and SIP Partners of TCS. He is also a visiting Faculty for Analytics at various colleges and institutes. Mohan holds degrees in Business Analytics and Intelligence from IIM-Bangalore, MBA in Marketing and BSC in Statistics.

PLACEMENT ASSISTANCE

The Career Assistance team at Imarticus provides 100% support throughout the program to guide and help navigate ample career options.

1 RESUME BUILDING



We help you refine and polish your resume with tips to help you land your coveted job

2 INTERVIEW PREP



We prepare you to ace the Technical interview rounds with model interview Q&A and extensive mock interviews

COLLABORATION WITH GENPACT

Genpact is a global leader in digitally-powered business process management and services and works with over 1/5th of the Fortune Global 500 companies across technology and analytics with revenues of \$2.46 billion and 70,000 employees spread across 25 countries.



**Project
Evaluation**



**Guest Lecture &
Mentorship**



**Industry-Approved
Curriculum**



**FOR ONLINE DELIVERY INQUIRIES:
1800-267-7679**



WWW.IMARTICUS.ORG



INFO@IMARTICUS.COM