



DATA SCIENCE PRODEGREE

Knowledge Partner:



Global Leader in Digitally-Powered
Business Process Management & Services

In Collaboration with Genpact, a Global Leader in Analytics

180+ Hours of Learning Delivered in Classroom and Online Format

Course Covering Multiple Analytics Tools such as R, Python, SAS and Tableau

Hands-on Learning with 14 Industry Projects Across Diverse Industries

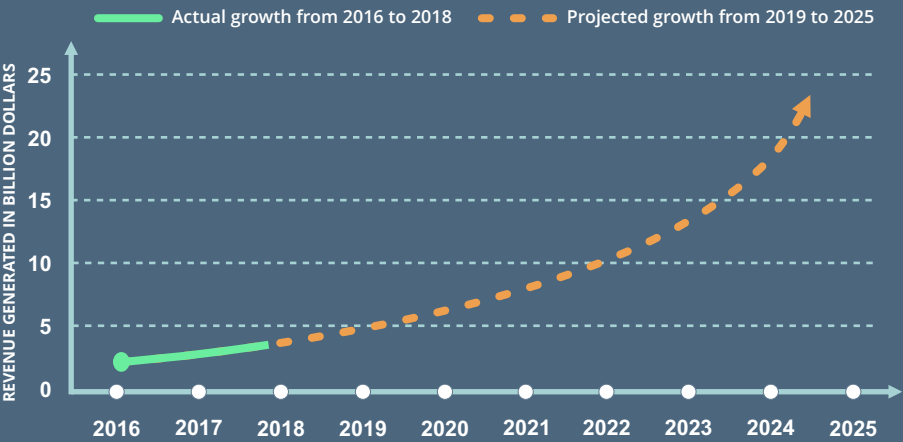
Learn to Build Complex Data Science Models and Solve Real-world Business Problems



INDUSTRY LANDSCAPE

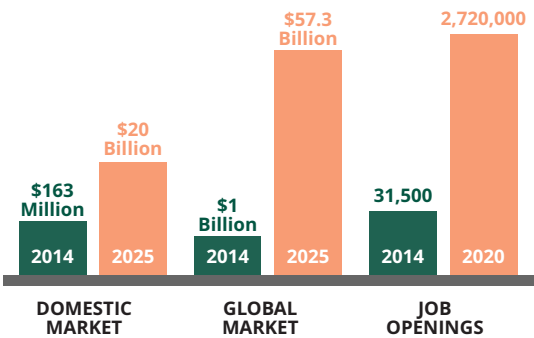
Analytics, data science and big data industry in India is currently estimated to be **\$2.71** billion annually in revenues, growing at a healthy rate of **33.5% CAGR**.

ANALYTICS REVENUE OVER YEARS



GROWING DEMAND

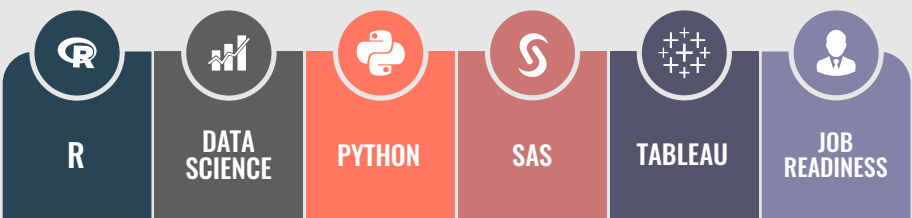
Projected Demand for Analytics Professionals in India



EMPLOYMENT LANDSCAPE



OVERVIEW OF PROGRAM



180-HOUR PROGRAM AVAILABLE IN CLASSROOM & ONLINE DELIVERY FORMAT

CURRICULUM

STATISTICS FUNDAMENTALS & R

R FOR DATA SCIENCE	R Installation, R Studio, Understanding Data Structures in R - Lists, Matrices, Vectors R Studio the IDE Basic Building Blocks in R Understanding Vectors in R Basic Operations Operators and Types 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
DATA VISUALIZATION USING R	Grammar of Graphics Bar Charts Histograms Pie Charts Scatter Plots Line Plots and Regression Word Clouds Box Plots GGLOT2
STATISTICAL LEARNING	Measures of Central Tendency in Data Measures of Dispersion Understanding Skewness in Data Probability Theory Bayes Theorem Probability Distributions Hypothesis Testing
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

DATA SCIENCE WITH R

EXPLORATORY DATA ANALYSIS WITH R	Merge, Rollup, Transpose and Append Missing Analysis and Treatment Outlier Analysis and Treatment Summarizing and Visualizing the Important Characteristics of Data Univariate, Bivariate Analysis Crosstabs, Correlation
LINEAR REGRESSION	What is Regression Analysis Limitations of Regression Covariance and Correlation Multivariate Analysis Assumptions of Linearity Hypothesis Testing Limitations of Regression Implementing Simple & Multiple Linear Regression Making Sense of Result Parameters Model Validation Handling Other Issues/Assumptions in Linear Regression Handling Outliers, Categorical Variables, Autocorrelation, Multicollinearity, Heteroskedasticity Prediction and Confidence Intervals
LOGISTIC REGRESSION	Implementing Logistic Regression Making Sense of Result Parameters: Wald Test, Likelihood Ratio Test Statistic, Chi-Square Test Goodness of Fit Measures Model Validation: Cross Validation, ROC Curve, Confusion Matrix
DECISION TREES	Introduction to Predictive Modelling with Decision Trees Entropy & Information Gain Standard Deviation Reduction (SDR) Overfitting Problem Cross Validation for Overfitting Problem Running as a Solution for Overfitting
LINEAR DISCRIMINANT ANALYSIS	Multi-class classification

DATA SCIENCE WITH PYTHON

BASICS OF PYTHON FOR DATA SCIENCE	Python Basics Data Structures in Python Control & Loop Statements in Python Functions & Classes in Python Working with Data
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
EXPLORATION DATA ANALYSIS	Data Visualization & EDA
TIME SERIES FORECASTING	Understand Time Series Data Visualizing Time Series Components Exponential Smoothing Holt's Model Holt-Winter's Model ARIMA ARCH & GARCH
UNSUPERVISED LEARNING	K-Means Clustering
DIMENSIONALITY REDUCTION	Principal Component Analysis (PCA) Scree Plot One-Eigen Value Criterion Factor Analysis
INTRODUCTION TO MACHINE LEARNING	Machine Learning Modelling Flow How to Treat Data in ML Parametric & Non-Parametric ML Algorithm Types of Machine Learning Performance Measures Bias-Variance Trade-Off Overfitting & Underfitting Optimization Techniques Scikit-Learn Library
SUPERVISED LEARNING	Linear Regression Linear Regression with Stochastic Gradient Descent, Batch GD Optimizing Learning Rate Momentum
LOGISTIC REGRESSION	Logistic Regression with Stochastic Gradient Descent, Batch GD Optimizing Learning Rate Momentum

CURRICULUM

K NEAREST NEIGHBOR

Understanding KNN | Voronoi Tessellation | Choosing K | Distance Metrics - Euclidean, Manhattan, Chebyshev

DECISION TREE & RANDOM FOREST

Fundamental Concepts of Ensemble | Hyper-Parameters

SUPPORT VECTOR MACHINES

What is SVM? | When to Use SVM? | What is Support Vector? | Understanding Hyperplane | Understanding Lagrangian Multiplier, Karush Kuhn Tucker Conditions | SVM Kernels - Radial Basis Function, Gaussian Kernel, Linear Kernel | Optimizing the C Parameter | Regularization

SAS PROGRAMMING

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 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

DATA VISUALIZATION WITH 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

Have a Methodology | Know Your Audience | Define Resulting Actions Classify Your Dashboard | Profile Your Data | Use Visual Features Properly | Design Iteratively

PROJECTS



Property Price Prediction using Linear Regression in R



Real Estate Price Prediction using Linear Regression



Bank Credit Card Default Prediction using Logistic Regression in R



Identifying Good and Bad Customers for Granting Credit



Predict Wine Quality with Decision Tree (Regression Trees * Classification Trees)



Breast Cancer Prediction - KNN Classifier & How to Choose the K Value



Multi-Class Classification with Linear Discriminant Analysis



Bank Marketing Analytics - Decision Tree & Random Forest Classifier



Forecasting and Predicting the Furniture Sales using ARIMA



Default Prediction of Credit Card Clients - SVM Classifier using Different Kernels



Reduce Data Dimensionality for a House Attribute Dataset using PCA



Store Data Analytics in SAS



Use K-means Clustering to Group Teen Students into Segments for Targeted Marketing Campaigns



Building Tableau Dashboard

KEY HIGHLIGHTS



GENPACT-ENDORSED CURRICULUM

Cutting-edge curriculum and best in-class content covering data analysis and statistics using SAS, R, Python & Tableau, developed in collaboration with Genpact.



PROJECT-BASED LEARNING: 14 INDUSTRY PROJECTS

The Prodegree provides an edge through our unique project-based methodology, focusing on 14 real-world projects. Students will gain hands-on experience of in-demand data science tools and build a portfolio of demonstrable work.



CLOUD LAB

24/7 access to your very own cloud-based lab that's agnostic of machine configuration, to hone your programming skills and build your projects.



DEDICATED CAREER MENTORSHIP

A dedicated industry mentor with over a decade of experience to guide you on the most suitable career path based on your skills and interests and resolve your career-related queries.



CAREER ASSISTANCE

The Imarticus Career Assistance Services team prepares you to be job-ready through extensive interview prep, resume building & mock interviews.

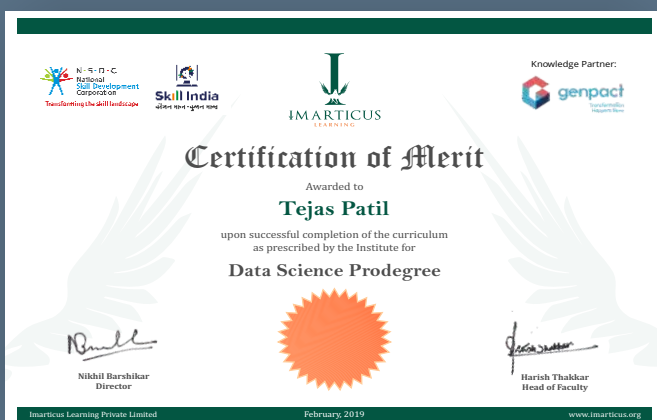


PROJECT PORTFOLIO

Build a demonstrable project portfolio on Github and showcase it to potential employers.

CERTIFICATION

On completion of the Data Science Prodegree, aspirants will receive an industry endorsed Certificate of Achievement, which is co-branded by Genpact and Imarticus Learning.



FACULTY

VINAY BORHADE

Vinay's tech expertise includes AI – Machine Learning, Python, PL-SQL, and Big Data – Netezza, Java/J2EE. Having served more than 10 years with Bank of America (Merrill Lynch), he has worked on projects like Finance, Liquidity and Capital Risk (Regulatory Reporting) and has won repeat business from clients for BOA using technologies like Machine Learning, Capitalize: Data Analytics, Quartz, Python, IBM Netezza, Oracle (Hexadata).

Y LAXMI PRASAD

With close to eight years of experience as the director of Clinitral, YL Prasad is an expert in SAS programming. Based in Hyderabad, he has worked extensively in the SAS programming and consulting sector. Educated at JNTU, YL Prasad started off his career as a consultant with Cognizant before working as a SAS Consultant with Siro Clinipharm.

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.

CAREER 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

3 PLACEMENT PORTAL



We give you unlimited access to our private and public leads and references on our placement portal

COLLABORATION WITH GENPACT

The Data Science Prodegree is co-created with Genpact as the Knowledge Partner and comes with a cutting-edge industry aligned curriculum and learning methodology. You will benefit in terms of:



Certification



Industry-Approved Curriculum

CONTACT US FOR A PROFILE REVIEW

AHMEDABAD

 +91 8980682998

DELHI

 +91 8448584775

MUMBAI

 +91 8929129624

BANGALORE

 +91 8971729953

GURUGRAM

 +91 8448984141

PUNE

 +91 8433905205

CHENNAI

 +91 9789879741

HYDERABAD

 +91 40 4712 3308/09

THANE

 +91 9319746164

FOR ONLINE DELIVERY INQUIRIES: **1-800-267-7679**



WWW.IMARTICUS.ORG



INFO@IMARTICUS.COM