

# Prospectus

# Certified Data Science Course





## About Us

**ExcelR Is An Ed-Tech Firm**

**Headquartered In Houston USA.**

**We have been the industry leaders**

**since the last 8 years and have**

**provided the best to the students**

**by collaborating with the best**

**companies and organisations.**

## Accolades

**Deloitte.**

**50** | Technology Fast 50  
2018 INDIA WINNER

**Google**

4.8   
Great Feedback Review 4.8

**siliconindia**  
TRAINING INSTITUTE  
OF THE YEAR 2018

**siliconindia**  
TRAINING INSTITUTE  
OF THE YEAR 2015

**CIO** Review 20 MOST PROMISING  
DATA ANALYTICS  
SOLUTION PROVIDERS - 2018

**Analytics India**  
MAGAZINE

10 Leading Courses for  
Block Chain 2019



## Accreditations & Partners

Steinbeis  
Global  
Institute  
Tübingen



Steinbeis-  
University  
Berlin  
SHB



mongoDB



PEOPLECERT



## Associations , Membership & Associates



NASSCOM®



ex



## Clientele



J.P.Morgan



ORACLE



Mindtree  
Welcome to possible

ex

# Why One Should Pursue Data Science?



HARVARD  
UNIVERSITY

The Sexiest Job of The 21<sup>st</sup> Century

More than 97,000 analytics positions remain vacant in India due to the shortage of talent.



**NASSCOM**

1,40,000 Data Science and Machine Learning jobs in India alone by 2021

The demand for analytics professionals rises by 2 Lakh Shortage

**NDTV**



# Curriculum

## Module 1: Data Science Project Lifecycle

- Demo: Introduction to Types of Analytics, Project Life Cycle, LMS walk through

## Module 2: Basic Stat

- Data Types
- Measure Of central tendency
- Measures of Dispersion
- Graphical Techniques
- Skewness & Kurtosis
- Box Plot

## Module 3: R Language

- R & R Studio
- Descriptive Stats in R

## Module 4: Python

- Python (Installation and basic commands) and Libraries
- Jupyter note book
- Set up GitHub
- Descriptive Stats in Python
- Pandas and Matplotlib

## Module 5: Basic Statistics- 2

- Random Variable
- Probability
- Probability Distribution
- Normal Distribution
- SND
- Expected Value
- Sampling Funnel
- Sampling Variation
- Central Limit Theorem
- Confidence interval
- Assignments Session-1 (1 hr)



# Curriculum

## Module 6: Hypothesis Testing

- Introduction to Hypothesis Testing
- Hypothesis Testing ( 2 proportion test, 2 t sample t test)
- Anova
- Chisquare

## Module 7: EDA

- Exploratory data analysis-1(Data Cleaning, Imputation Techniques, Data analysis)
- Visualization(Scatter Diagram, Correlation Analysis, Transformations)

## Module 8: Linear Regression

- Principles of Regression
- Intro to Simple Linear Regression
- Multiple Linear Regression

## Module 9: Logistic Regression

## Module 10: Deployment Methods

- Model deployments using R and Python

## Module 11: Assignments

- Assignments Session-2(1 hr)
- Clustering introduction
- Hierarchical clustering



## Module 12: Data Mining

- Unsupervised ML Algorithms
- Kmeans
- DBSCAN

## Module 13: Dimensional Reduction Techniques

- PCA
- tSNE

# Curriculum

## Module 14: Market Basket Analysis

- Association Rules

## Module 15: Recommendation System and Assignment

- Recommender System
- Assignments Session-3 (1 hr)

## Module 16: Supervised Machine Learning

- Supervised Machine Learning Concept(Regression Tasks/ Classification Tasks)

## Module 17: Decision Tree

- Decision Tree(C5.0)

## Module 18: EDA-2

- EDA -2 ( Encoding Methods - OHE, Label Encoders, Outlier detection-Isolation Forest)
- Calculating the Predictive Power Score (PPS)

## Module 19: Feature Engineering

- Feature Engineering (Tree based methods, RFE,PCA)

## Module 20: Modal Validation Techniques

- Model Validation Methods (train-test, CV, Shuffle CV, and Accuracy methods)

## Module 21: Ensembled Techniques

- Bagging
- Random Forest
- Boosting
- XGBM
- LGBM



# Curriculum

## Module 22: Classifiers

- KNN
- Support Vector Machines

## Module 23: Regularization Techniques

- Lasso
- Ridge Regressions

## Module 24: Neural Network

- ANN
- Optimization Algorithm(Gradient descent)
- Stochastic gradient descent(intro)
- Back Propagation method
- Introduction to CNN
- Assignments Session-4 (1 hr)

## Module 25: Text Mining

- Introduction to Text Mining
- VSM
- Intro to word embedding's
- Word clouds and Document Similarity using cosine similarity
- Named Entity Recognition

## Module 26: Naïve Bayes

- Text classification using Naïve Bayes
- Emotion Mining

## Module 27: Time Series

- Introduction to Time series
- Level, Trend and Seasonality Strategy
- Scatter plot
- Lag plot



# Curriculum

- ACF
- Principles of Visualization
- Naïve forecasts

## **Module 28: Forecasting**

- Forecasting Error and its metrics
- Model Based Approaches
- AR Model for errors
- AR Model for errors
- Data driven approaches
- MA
- Exponential Smoothing
- ARIMA
- Survival Analysis

## **Module 29: Project Discussion**

- Hands on using R and Python Projects description with deployment



# Value Added Courses- AI

## Intro to Neural Network & Deep Learning

- Intro
- Deep Learning Importance [Strength & Limitations]
- SP | MLP
- Neural Network Overview
- Neural Network Representation
- Activation Function
- Loss Function
- Importance of Non-linear Activation Function
- Gradient Descent for Neural Network

## Parameter & Hyper parameter

- Train, Test & Validation Set
- Vanishing & Exploding Gradient
- Dropout
- Regularization
- Optimization algo
- Learning Rate
- Tuning
- Softmax

## CNN

- CNN
- Deep Convolution Model
- Detection Algorithm
- Face Recognition

## RNN

- RNN
- LSTM
- Bi Directional LSTM



# Value Added Courses

## Hadoop and Spark

- Introduction to Big Data
- Challenges in Big Data and Workarounds
- Introduction to Hadoop and its Components
- Hadoop components and Hands-on
- Understand the MapReduce (Distributed Computation Framework) and its Drawback
- Introduction to Spark
- Spark Components
- Spark MLlib and Hands-on (one ML model in spark)

## R

- Introduction to R Programming
- How To Install R & R Studio
- Data Structures in R
- Programming Statistical
- How to Import Dataset in R
- R-Packages
- How to Integrate R and SQL
- How to Get Data From SQL to R



## Python

- Introduction to Python
- Installation of Anaconda Python
- Difference between Python2 and Python3
- Python Environment
- Operators
- Identifiers
- Exception Handling (Error Handling)

# Value Added Courses

## Azure

- Introduction to Cloud Computing
- Difference between On Premise and Cloud
- Types of Service Models
- Advantages of Cloud Computing
- Azure Global Infrastructure
- Creation of Free tire account inside Azure
- Brief introduction to Machine Learning Services on Cloud and more

## MYSQL

- Introduction to What is DataBase
- Difference between SQL and NOSQL DB
- How to Install MYSQL and Workbench
- Connecting to DB
- Creating to DB
- What are the Languages inside SQL How to Create Tables inside DB and Inserting the Records
- Select statement and using Queries for seeing your data
- Joining 2 tables
- Where clause usage
- Indexes and views
- Different operations in SQL
- How to Connect to your applications from MYSQL includes R and Python

## Tableau

- Introduction to Data Visualization
- Tableau – Data Visualization Tool
- Tableau User Interface
- Basic Chart types
- Intermediate Chart



# Value Added Courses

- Advanced Charts
- Maps in Tableau
- Adding Background Image
- Data Connectivity in-depth understanding
- Creating Calculated Fields
- Responsive Tool Tips
- Connecting Tableau with Tableau Server
- Connecting Tableau with R



# Key Highlights

1

Dedicated  
Support  
Team

2

3+  
Capstones  
Projects  
from  
Various  
Domains

3

100%  
Placement  
Assistance

4

Full time  
Weekday/  
Weekend  
Sessions

# 1

Dedicated  
Support  
Team



## SUPPORT TEAM

**Support Is The Most Important Aspect When It Comes to learning technical courses. This is where we are so highly rated. We have a dedicated assignments team that helps you in every step when it comes to solving your doubts ( No matter how big or how small your doubt is). Shoot those doubts to us and we will help you solve them.**



# PROJECTS

## Projects From Various Domains

### (3+ Capstones Projects)

ExcelR specialises in the projects with access to projects from over 18 domains. We recommend a minimum of 2 projects to be completed from a variety of domains and there is no restriction on the maximum number of projects that have to be completed.

Get yourself immersed in practical learning by completing at least 2-3 projects.



# SAMPLE PROJECTS

## (REAL LIFE DATA PROJECTS)

### 1. "Daily" Twitter Data Analysis for a Product

Sentiment Analysis with the help of Natural Language Processing technique for identifying the sentiments of a product or service. **Domain: Social Media**

### 2. Natural Language Processing

The project is to build information retrieval system from Amazon products data based on NLP techniques. Top 5 relevant answers to be retrieved based on input question.

**Domain: E commerce**

### 3. Predicting Loan defaulters

Reducing the risk of fraudulent loans by carefully evaluating the risk & at the same time increasing profits by rejecting only those loans, which have the potential of defaulting. **Domain: Banking**

### 4. Warranty Cost prediction

The objective of the analysis to predict an item when sold, what is the probability that customer would file for warranty and to understand important factors associated with them. **Domain: FMCG**

### 5. Predict flight delays

Predict which flights would be delayed and by how long?

**Domain: Aviation**





# PLACEMENT ASSISTANCE

## **100% Placement Assistance**

Once the assignments and projects are completed, we have a placement incubation cell that will help you with the following :

- ◆ Resume Prep Sessions and Guidance
- ◆ Mock Interviews Of The Student By Senior Data Scientists
- ◆ Detailed Analysis of where the student has to improve  
( SWOT Analysis )
- ◆ Facilitate interviews from a network of 150+ Company Network
- ◆ We have a placement record\* of 91.2% over the last one year.

**\*Please note that only those participants that finish their assignments and projects will be eligible for placement assistance**





# FULL TIME/ WEEKDAY/ WEEKEND SESSIONS

## ExcelR Provides The Best Training In Two Modes :

- 1. Classroom Training**
- 2. Online Training**

### Classroom Training :

- Weekdays – Monday to Friday – Two Hours A Day**  
**( Batches Available Through The Day )**
- Weekends – Saturday and Sunday – 4 hrs a Day**  
**( 9 AM to 1 PM & 2 PM to 6PM)**

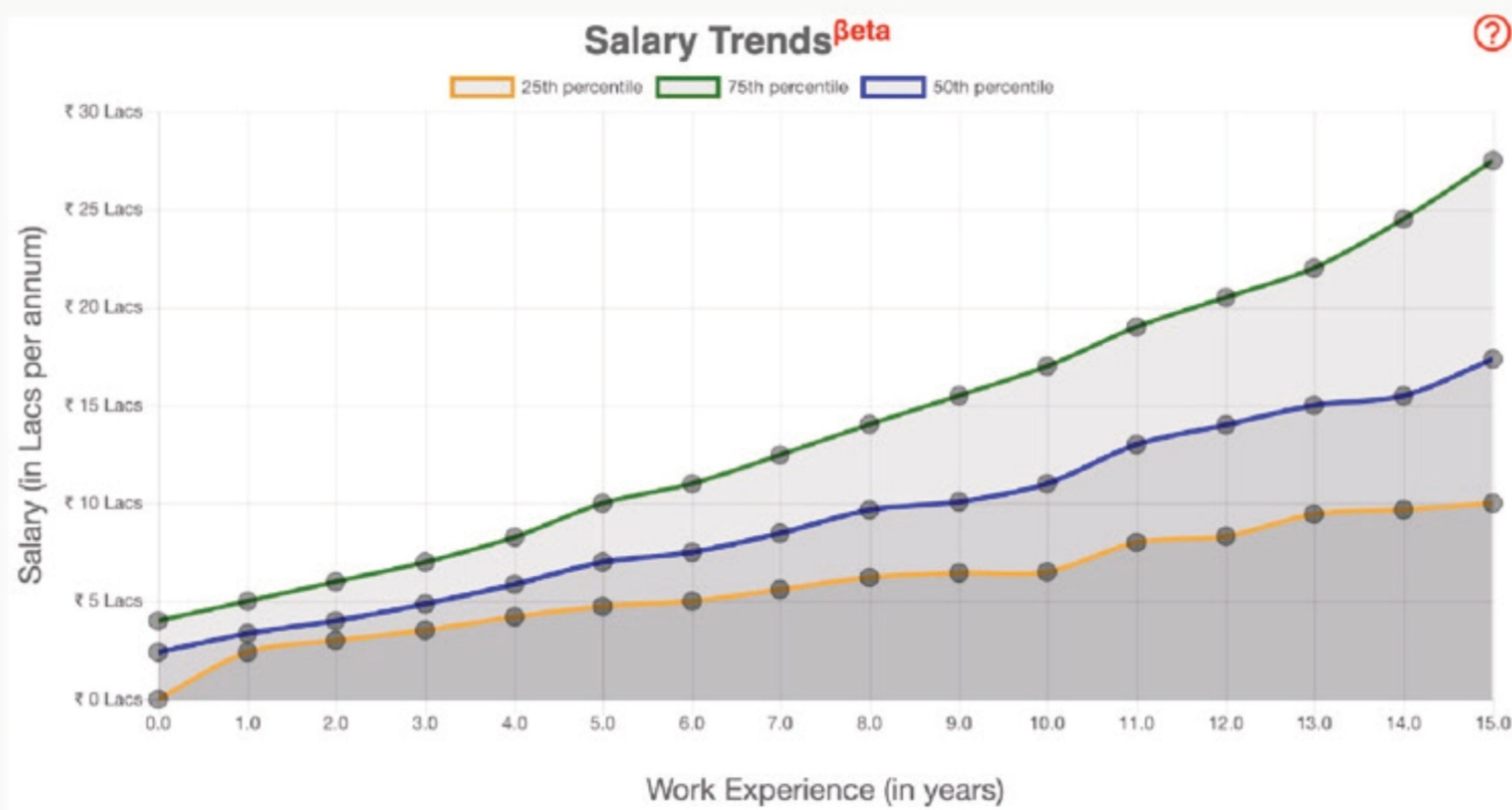
### Online Training :

- 1. Live Virtual Classes With**  
**Live Interaction With The**  
**Trainer ( Batches**  
**Available Through**  
**The Day )**



# SALARY INSIGHTS

LEVEL	EXPERIENCE	PACKAGE
Entry - Level	0 years	Min. 3 Lakhs per Annum
Early - Level	1-4 years	4-7 Lakhs per Annum
Mid - Level	5-9 years	6-15 Lakhs per Annum
High - Level	Above 10 years	Max. 25 Lakhs per Annum



Source : Naukri.com



# SUMMARY OF THE COURSE

01



**Data Science  
Certification from**

02



**Post Training Support is  
provided to help Build  
up your qualifications**

03



**Gain highly valued skills  
from the best qualified  
Faculty**

04



**Placement Assistance,  
The team will guide  
them to prepare for  
their future career  
trajectory**

# CERTIFICATE



## ExcelR Solutions

Hereby confers upon

*Mr. Name*

The

### Data Science Certification

after the successful completion of the course that was held between

March 05<sup>th</sup>, 2020 and July 09<sup>th</sup>, 2020



Reg/Cert No:103/ EXCELR/14092020

A handwritten signature in black ink.

Ram Tavva  
Director  
EXCELR SOLUTIONS

#### EXCELR SOLUTIONS

A Division of  
EXCELR SOLUTIONS (USA)  
201 S High Point Road APT 213  
Madison, Wisconsin - 53717 USA.





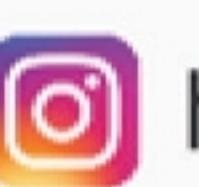
For further information on the program we offer  
please visit our website or contact us at:

 [www.excelr.com](http://www.excelr.com)

 **1800-212-2121**

**Follow us on :**

 <https://www.facebook.com/ExcelR/>

 <https://www.instagram/excelrsolutions/>

 <https://twitter.com/ExcelrS>

 <https://www.youtube.com/c/ExcelRSolutionsRaisingExcellence>

 <https://www.linkedin.com/company/excelr-solutions/>

