



**ANALYTIX** LABS

## **Data Science 360**

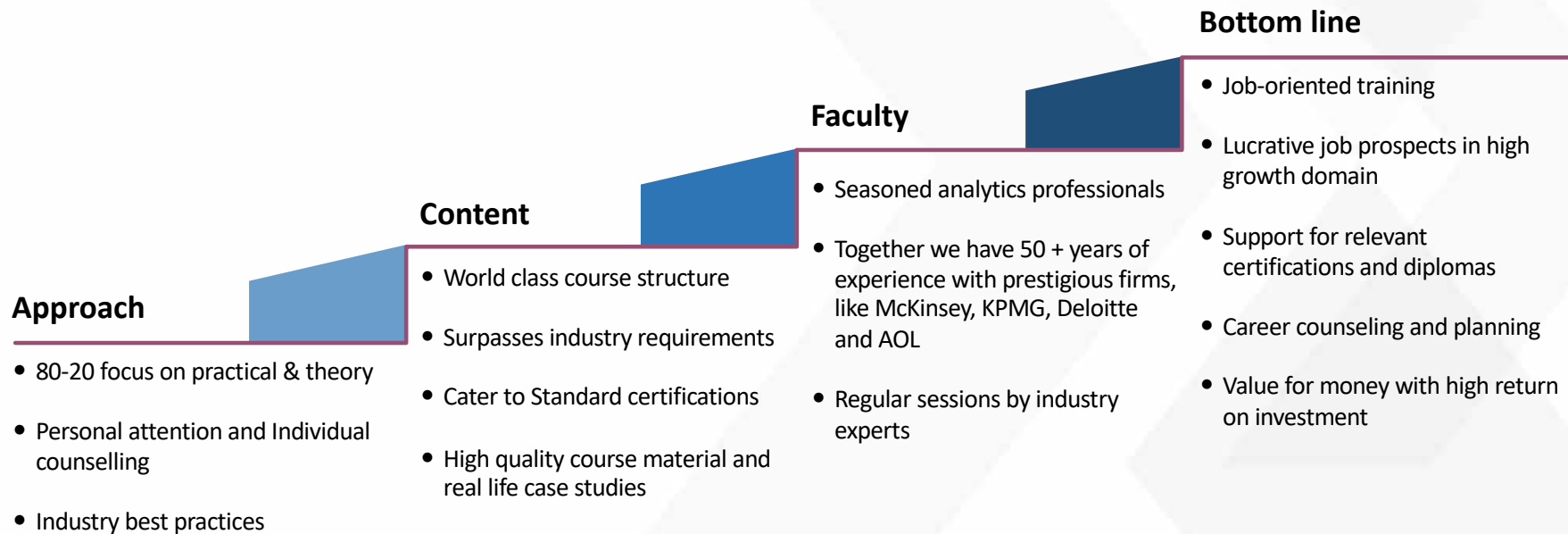
**Job oriented global certification  
program crafted by experts**

Disclaimer: This material is protected under copyright act AnalytixLabs ©, 2011-2020. Unauthorized use and/ or duplication of this material or any part of this material including data, in any form without explicit and written permission from AnalytixLabs is strictly prohibited. Any violation of this copyright will attract legal actions

***Learn to Evolve***

# About AnalytixLabs

AnalytixLabs is a capability building and training solutions firm led by McKinsey, IIM, ISB and IIT alumni with deep industry experience and a flair for coaching. We are focused at helping our clients develop skills in basic and advanced analytics to enable them to emerge as “Industry Ready” professionals and enhance career opportunities. AnalytixLabs has been also featured as top institutes by prestigious publications like Analytics India Magazine and Higher Education Review, since 2013.



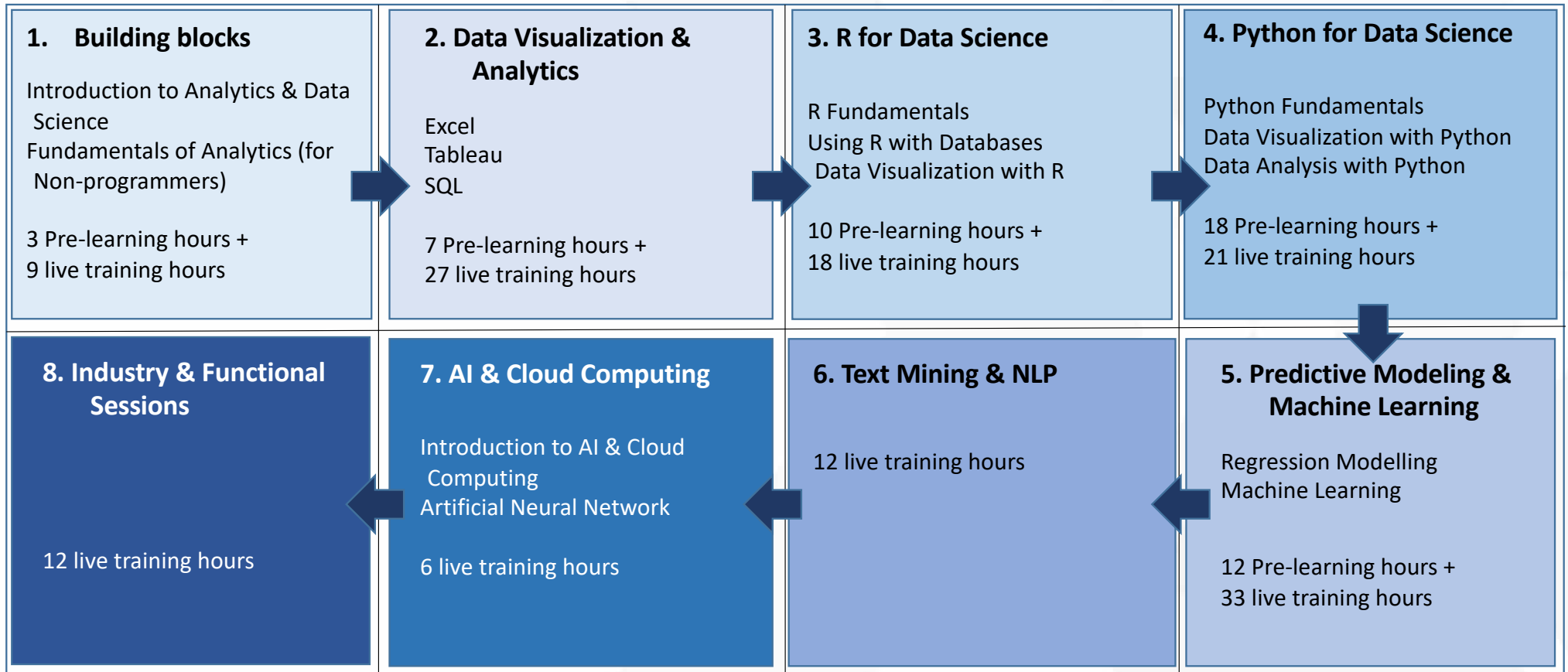
Candidates trained by us are working in leading companies across industries...



# Program Objective – Data Science 360

18 Assignments and  
Projects included

On average 8 hours of  
self-study per week



*\*Pre-learning content refers to basic foundation concepts and candidates are required to go through these using e-learning modules in self-paced mode*

# 1. Building Blocks

## Introduction to Bridge Course & Analytics Software's

### Basic Excel

- Excel Environment
- Key Terminologies
- Short Cuts
- Key Functionalities
- Copy-paste-paste special
- Formatting & conditional Formatting
- Basic Excel Functions - Types of Functions
- Relational operators
- Data Sorting, Filtering and Data Validation
- Understanding of Name Ranges
- Pivot tables - Charts
- Basics of charts

### Basic Programming Elements

- Overview of programming languages
- Basics of programming elements
  - variables, data types, data structures, loops, conditional statements, inputs, outputs, functions etc
- Understanding key terms
  - Client/server
  - Database
  - Hosting/deployment

## Introduction to Basic Statistics

- Introduction to Statistics
- Measures of central tendencies
- Measures of variance
- Measures of frequency
- Measures of Rank
- Basics of Probability, distributions
- Conditional Probability (Bayes Theorem)

## RDBMS & SQL (Basics)

- **Basic RDBMS Concepts**
  - Introduction to Relational Database management system. Why SQL?
  - A glance at the tool and its advantages and disadvantages
  - Understanding Schema, ERDs and Metadata
- Introduction to MS SQL Server
  - What is SQL – A Quick Introduction
  - Installing MS SQL Server for windows
  - Introduction to SQL Server Management Studio
  - Understanding basic database concepts
  - Getting started

## Introduction to Analytics & Data Science

- What is analytics & Data Science?
- Business Analytics vs. Data Analytics vs. Data Science
- Common Terms in Analytics
- Analytics vs. Data warehousing, OLAP, MIS Reporting
- Types of data (Structured vs. Unstructured vs. Semi Structured)
- Relevance of Analytics in industry and need of the hour
- Critical success drivers
- Overview of analytics tools & their popularity
- Analytics Methodology & problem solving framework
- Stages of Analytics

## Introduction to Mathematical foundations

- Introduction to Linear Algebra
  - Matrices Operations
- Introduction to Calculus
  - Derivatives & Integration
  - Maxima, minima
  - Area under the curve
- Theory of optimization

## 2. Data Visualization & Analytics (Excel) (1/3)

### Quick Recap of Basics of Excel

#### Data manipulation using functions

- Descriptive functions
- Logical functions: IF, and, or, not
- Date and Time functions
- Text functions
- Array functions
- Use and application of lookup functions
- Limitations of lookup functions
- Using Index, Match, Offset, reverse vlookup

#### Data analysis and reporting

- Data Analysis using Pivot Tables - use of row and column shelf, values and filters
- Difference between data layering and cross tabulation, summary reports, advantages and limitations
- Change aggregation types and summarization
- Creating groups and bins in pivot data
- Concept of calculated fields, usage and limitations
- Changing report layouts - Outline, compact and tabular forms
- Show and hide grand totals and subtotals
- Creating summary reports using pivot tables

### Data Visualization in Excel

- Overview of chart types – column/bar charts, line/area , pie, doughnut charts, scatter plots
- How to select right chart for your data
- Creating and customizing advance charts - thermometer charts, waterfall charts, population pyramids

### Overview of Dashboards

- What is dashboard & Excel dashboard
- Adding icons and images to dashboards
- Making dashboards dynamic

### Create dashboards in Excel - Using Pivot controls

- Concept of pivot cache and its use in creating interactive dashboards in excel
- Pivot table design elements - concept of slicers and timelines
- Designing sample dashboard using Pivot Controls
- Design principles for including charts in dashboards - do's and dont's

### Business Dashboard Creation

- Management Dashboard for Sales & Services
- Best practices - Tips and Tricks to enhance dashboard designing

## 2. Data Visualization & Analytics (SQL) (2/3)

### Quick Recap of RDBMS & Basic SQL

#### Data based objects creation (DDL Commands)

- Creating databases and tables. Understanding data types
- Inserting values into the table
- Altering table properties
- Introduction to Keys and constraints
- Creating, Modifying & Deleting Tables
- Create Table & Create Index statements
- Drop & Truncate statements – Uses & Differences
- DDL Statements with constraints
- Import and Export wizard to get the data in SQL server from excel files or delimited files

#### Data manipulation (DML Commands)

- Data Manipulation statements
- Insert, Update & Delete statements
- Select statement – Sub setting, Filters, Sorting. Removing Duplicates, grouping and aggregations etc
- Operators, predicates and built in functions(Top, distinct, Limit)
- Where, Group By, Order by & Having clauses
- SQL Functions – Number, Text, Date, etc
- SQL Keywords – Top, Distinct, Null, etc
- SQL Operators - Relational (single valued and multi valued), Logical (and, or, not), Use of wildcard operators and wildcard characters, etc

#### Accessing data from Multiple Tables using SELECT

- Append and Joins
- Union and Union All – Use & constraints
- Intersect and Except statements
- Table Joins - inner join, left join, right join, full join
- Cross joins/cartisian products, self joins, natural joins etc
- Inline views and sub-queries & it's types
- Optimizing your work
- Update operations with and without joins

#### Advanced SQL

- Creating table copy and database copy
- Views
- Transactions
- Stored Procedures in SQL
- Crud operations using stored procedures
- Window functions in SQL
- Miscellaneous Topics: Rollup and cube

#### Apply learning's on Business Case study

## 2. Data Visualization & Analytics (Tableau) (3/3)

### Getting Started

- What is Tableau?
- Tableau product suite
- How Does Tableau Work?
- Tableau Architecture
- Connecting to Data & data source concepts
- Understanding the Tableau workspace
- Dimensions and Measures
- Data Types & Default Properties
- Tour of Shelves & Marks Card
- Using Show Me
- Saving and Sharing your work-overview

### Data handling & summaries

- Date Aggregations and Date parts
- Cross tab & Tabular charts
- Totals & Subtotals
- Bar Charts & Stacked Bars
- Line Graphs with Date & Without Date
- Tree maps/Scatter Plots
- Individual Axes, Blended Axes, Dual Axes & Combination chart/Edit axis
- Parts of Views
- Sorting
- Trend/Reference Lines/Forecasting
- Filters/Context filters

### Data handling & summaries

- Sets (In/Out Sets/Combined Sets)
- Grouping/Bins/Histograms
- Drilling up/down – drill through
- Hierarchies
- View data
- Actions (across sheets)

### Building Advanced Reports/ Maps

- Explain latitude and longitude
- Default location/Edit locations
- Building geographical maps
- Using Map layers

### Calculated Fields

- Aggregate vs. Disaggregate data
- Explain - #Number of Rows
- Basic Functions (String/Date/Numbers etc)
- Usage of Logical conditions

### Table calculations

- Explain scope and direction
- Percent of Total, Running / Cumulative calculations
- Introduction to LOD (Level of Detail) Expressions
- User applications of Table calculations

### Parameters

- Using Parameters in calculated fields
- Bins/Reference Lines
- Filters/Sets
- Display Options (Dynamic Dimension/Measure Selection)
- Create What-If/ Scenario analysis

### Building Interactive Dashboards

- Combining multiple visualizations into a dashboard (overview)
- Making your worksheet interactive by using actions
- Filter/URL/Highlight
- Complete Interactive Dashboard for Sales & Services

### Building Stories

- Story Points
- Options in Formatting your Visualization
- Working with Labels and Annotations
- Effective Use of Titles and Captions

### Working with Data

- Multiple Table Join
- Data Blending
- Difference between joining and blending data, and when we should do each
- Toggle between to Direct Connection and Extracts

### Sharing work with others

- Sharing Workbooks
- Publish to Reader/PDF
- Publish to Tableau Server and sharing on the web



# 3. R For Data Science

## Data Importing/Exporting

- Introduction R/R-Studio - GUI
- Concept of Packages - Useful Packages (Base & Other packages)
- Data Structure & Data Types (Vectors, Matrices, factors, Data frames, and Lists)
- Importing Data from various sources
- Exporting Data to various formats
- Viewing Data (Viewing partial data and full data)
- Variable & Value Labels – Date Values

## Data Manipulation

- Creating New Variables (calculations & Binning)
- Dummy variable creation
- Applying transformations
- Handling duplicates/missing's
- Sorting and Filtering
- Sub setting (Rows/Columns)
- Appending (Row/column appending)
- Merging/Joining (Left,right,inner,full,outer)
- Data type conversions
- Renaming
- Formatting

## Data Manipulation

- Reshaping data
- Sampling
- Operators
- Control Structures (if, if else)
- Loops (Conditional, iterative loops)
- apply functions
- Arrays
- R Built-in Functions
- Text, Numeric, Date, utility
- R User Defined Functions
- Aggregation/Summarization

## Data Analysis

- Introduction exploratory data analysis
- Descriptive statistics, Frequency Tables and summarization
- Uni-variate Analysis (Distribution of data)
- Bivariate Analysis(Cross Tabs, Distributions & Relationships)

## Using R with Databases

- R and Relational Databases
- Connecting to Relational Databases using RJDBC and RODBC
- Database Design and Querying Data
- Modifying Data and Using Stored Procedures
- In-Database Analytics with R

## Data Visualization with R

- Basic Visualization Tools
  - Bar Charts/Histograms/Pie Charts
- Basic Visualization Tools Continued
  - Scatter Plots
  - Line Plots and Regression
- Specialized Visualization Tools
  - Word Clouds/ Radar Charts
  - Waffle Charts/ Box Plots
- How to create Maps
  - Creating Maps in R
- How to build interactive web pages
  - Introduction to Shiny
  - Creating and Customizing Shiny Apps
  - Additional Shiny Features

# 4. Python For Data Science (1/2)

## Python Essentials (Core)

- Overview of Python- Starting with Python
- Why Python for data science?
  - Anaconda vs. python
- Introduction to installation of Python
- Introduction to Python IDE's(Jupyter,/Ipython)
- Concept of Packages - Important packages
  - NumPy, SciPy, scikit-learn, Pandas, Matplotlib, etc
- Installing & loading Packages & Name Spaces
- Data Types & Data objects/structures (strings, Tuples, Lists, Dictionaries)
- List and Dictionary Comprehensions
- Variable & Value Labels – Date & Time Values
- Basic Operations – Mathematical/string/date
- Control flow & conditional statements
- Debugging & Code profiling
- Python Built-in Functions (Text, numeric, date, utility functions)
- User defined functions – Lambda functions
- Concept of apply functions
- Python – Objects – OOPs concepts
- How to create & call class and modules?

## Operations with NumPy (Numerical Python)

- What is NumPy?
- Overview of functions & methods in NumPy
- Data structures in NumPy
- Creating arrays and initializing
- Reading arrays from files
- Special initializing functions
- Slicing and indexing
- Reshaping arrays
- Combining arrays
- NumPy Maths

## Overview of Pandas

- What is pandas, its functions & methods
- Pandas Data Structures (Series & Data Frames)
- Creating Data Structures (Data import – reading into pandas)

## Cleansing Data with Python

- Understand the data
- Sub Setting / Filtering / Slicing Data
  - Using [] brackets
  - Using indexing or referring with column names/rows
  - Using functions
  - Dropping rows & columns

- Mutation of table (Adding/deleting columns)
- Binning data (Binning numerical variables in to categorical variables)
- Renaming columns or rows
- Sorting (by data/values, index)
  - By one column or multiple columns
  - Ascending or Descending
- Type conversions
- Setting index
- Handling duplicates /missing/Outliers
- Creating dummies from categorical data (using get\_dummies())
- Applying functions to all the variables in a data frame (broadcasting)
- Data manipulation tools(Operators, Functions, Packages, control structures, Loops, arrays etc.)

## Data Analysis using Python

- Exploratory data analysis
- Descriptive statistics, Frequency Tables and summarization
- Uni-variate Analysis (Distribution of data & Graphical Analysis)
- Bi-Variate Analysis(Cross Tabs, Distributions & Relationships, Graphical Analysis)

# 4. Python For Data Science (2/2)

## Data Visualization with Python

- Introduction to Data Visualization
- Introduction to Matplotlib
- Basic Plotting with Matplotlib
- Line Plots

## Basic Visualization Tools

- Area Plots
- Histograms
- Bar Charts
- Pie Charts
- Box Plots
- Scatter Plots
- Bubble Plots

## Advanced Visualization Tools

- Waffle Charts
- Word Clouds
- Seaborn and Regression Plots

## Visualizing Geospatial Data

- Introduction to Folium
- Maps with Markers
- Choropleth Maps

## Statistical Methods & Hypothesis Testing

- Descriptive vs. Inferential Statistics
- What is probability distribution?
- Important distributions (discrete & continuous distributions)
- Deep dive of normal distributions and properties
- Concept of sampling & types of sampling
- Concept of standard error and central limit theorem
- Concept of Hypothesis Testing
- Statistical Methods - Z/t-tests (One sample, independent, paired), ANOVA, Correlation and Chi-square

# 5. Predicting Modeling & Introduction to Machine Learning

## Introduction to Predictive Modeling

- Concept of model in analytics and how it is used?
- Common terminology used in modeling process
- Types of Business problems - Mapping of Algorithms
- Different Phases of Predictive Modeling
- Data Exploration for modeling
- Exploring the data and identifying any problems with the data (Data Audit Report)
- Identify missing/Outliers in the data
- Visualize the data trends and patterns

## Introduction to Machine Learning

- Applications of Machine Learning
- Supervised vs Unsupervised Learning
- Overall process of executing the ML project
- Stages of ML Project
- Concept of Over fitting and Under fitting (Bias-Variance Trade off) & Performance Metrics
- Concept of feature engineering
- Regularization (LASSO, Elastic net and Ridge)
- Types of Cross validation(Train & Test, K-Fold validation etc.)
- Concept of optimization - Gradient descent algorithm
- Cost & optimization functions
- Python libraries suitable for Machine Learning

## Supervised Learning: Regression problems

- Linear Regression
- Non-linear Regression
- K-Nearest Neighbor
- Decision Trees
- Ensemble Learning - Bagging, Random Forest, Adaboost, Gradient Boost, XGBoost
- Support Vector Regressor

## Supervised Learning: Classification problems

- Logistic Regression
- K-Nearest Neighbor
- Naïve Bayes Classifier
- Decision Trees
- Ensemble Learning - Bagging, Random Forest, Adaboost, Gradient Boost, XGBoost
- Support Vector Classifier

## Unsupervised Learning

- Principle Component Analysis
- K-Means Clustering
- Hierarchical Clustering
- Density-Based Clustering

## Recommender Systems

- Content-based recommender systems
- Collaborative Filtering

## Time Series Forecasting

- What is forecasting?
- Applications of forecasting
- Time Series Components and Decomposition
- Types of Seasonality
- Important terminology: lag, lead, Stationary, stationary tests, auto correlation & white noise, ACF & PACF plots, auto regression, differencing
- Classification of Time Series Techniques (Uni-variate & Multivariate)
- Time Series Modeling & Forecasting Techniques
  - Averages (Moving average, Weighted Moving Average)
  - ETS models (Holt Winter Methods)
  - Seasonal Decomposition
  - ARIMA/ARIMAX/SARIMA/SARIMAX
  - Regression
  - Evaluation of Forecasting Models

# 6. Text Mining using NLP

## Introduction to Text Mining

- Text Mining - characteristics, trends
- Text Processing using Base Python & Pandas, Regular Expressions
  - Text processing using string functions & methods
  - Understanding regular expressions
  - Identifying patterns in the text using regular expressions

## Text Processing with modules like NLTK, sklearn

- Getting Started with NLTK
- Introduction to NLP & NLTK
- Introduction to NLTK Modules (corpus, tokenize, Stem, collocations, tag, classify, cluster, tbl, chunk, Parse, ccg, sem, inference, metrics, app, chat, toolbox etc)

## Initial data processing and simple statistical tools

- Reading data from file folder/from text file, from the Internet & Web scrapping, Data Parsing
- Cleaning and normalization of data
- Sentence Tokenize and Word Tokenize, Removing insignificant words(“stop words”), Removing special symbols, removing bullet points and digits, changing letters to lowercase, stemming /lemmatization /chunking

- Creating Term-Document matrix
- Tagging text with parts of speech
- Word Sense Disambiguation
- Finding associations
- Measurement of similarity between documents and terms
- Visualization of term significance in the form of word clouds

## Advanced data processing and visualization

- Vectorization (Count, TF-IDF, Word Embedding's)
- Sentiment analysis (vocabulary approach, based on Bayesian probability methods)
- Name entity recognition (NER)
- Methods of data visualization
  - word length counts plot
  - word frequency plots
  - word clouds
  - correlation plots
  - letter frequency plot
  - Heat map
- Grouping texts using different methods
- Language Models and n-grams -- Statistical Models of Unseen Data (Smoothing)

## Text Mining – Predictive Modeling

- Semantic similarity between texts
- Text Segmentation
- Topic Mining (LDA)
- Text Classification(spam detection, sentiment analysis, Intent Analysis)

# 7. Introduction to AI & DL & Cloud Computing

## Introduction to Artificial Intelligence (AI)

- Modern era of AI
- Role of Machine learning & Deep Learning in AI
- Hardware for AI (CPU vs. GPU vs. FPGA)
- Software Frameworks for AI & Deep Learning
- Key Industry applications of AI

## Introduction to Deep Learning

- What are the Limitations of Machine Learning?
- What is Deep Learning?
- Advantage of Deep Learning over Machine learning
- Reasons to go for Deep Learning
- Real-Life use cases of Deep Learning
- Overview of important python packages for Deep Learning

## Artificial Neural Network

- Overview of Neural Networks
- Activation Functions, hidden layers, hidden units
- Illustrate & Training a Perceptron
- Important Parameters of Perceptron
- Understand limitations of A Single Layer Perceptron
- Illustrate Multi-Layer Perceptron
- Understand Backpropagation – Using Example
- Implementation of ANN in Python- Keras

## Introduction to Google Colab

## Introduction to Cloud Computing

- What is Cloud Computing? Why it matters?
- Traditional IT Infrastructure vs. Cloud Infrastructure
- Cloud Companies (Microsoft Azure, GCP, AWS ) & their Cloud Services (Compute, storage, networking, apps, cognitive etc.)
- Use Cases of Cloud computing
- Over view of Cloud Segments: IaaS, PaaS, SaaS
- Overview of Cloud Deployment Models
- Overview of Cloud Security
- AWS vs. Azure vs. GCP
- Implementation of ML/DL model in Cloud

# 8. Industrial & Functional Sessions (Domain Understanding)

## Introduction to Data Sources for Various Industries

## Introduction to Analytics Project Management

### Marketing Analytics

- Introduction to Marketing Function
- Marketing Research - Analytics
- Customer Analytics
- Campaign Analytics
- Pricing Analytics
- Marketing Return on Investment (MROI)
- Market Mix Models

### Risk Analytics

- Introduction to Risk Function
- Enterprise Risk Function
- Credit Scoring (Application & Behavioral)
- Fraud Analytics

### Operation Analytics

- Overview of Operation Analytics
- Applications Analytics in different functions
- Service Operations
- Manufacturing
- Logistics
- Business Support Functions (HR Analytics)
- Inventory Management

### Digital Analytics(Web Analytics)

### Social Network Analytics

## Banking & Financial Services, Insurance

- Applications of analytics to various functions
- Retail Banking
- Commercial Banking
- Investment Banking
- Insurance

### Retail & E-Commerce

- Customer Analytics
- Demand & Supply Chain
- Merchandizing & planning
- Price & Promotion Analysis
- Trust Analytics
- Customer Service
- A/B Testing
- Recommendation Engines

### Pharma & Health Care

- Marketing & Sales Analytics
- Provider-Payer-Patient Analytics
- Claims Analytics
- Fraud Analytics
- MROI

### Telecom & Network

- Network Analytics
- Subscriber Analytics
- Loyalty Analytics
- Revenue Leakage Analytics

# Course completion and career assistance

## Course completion & Certification criteria

---

- You shall be awarded an AnalytixLabs certificate only post the submission and evaluation of mandatory course project work. These will be provided as a part of the training.
- There is no pass/fail for these assignments and projects. Our objective is to ensure that trainees get strong hands-on experience so that they are well-prepared for job interviews along with performance at their jobs.
- In case the assignments and projects are not up-to-the-mark, trainees are welcome to take help and support for improvisation.
- While weekly schedule is shared with trainees for regular assignments, candidates get 3 months, post course completion, to submit their final assignment and projects.

## What is included in career assistance?

---

- Post successful course completion, candidates can seek assistance from AnalytixLabs for profile building. A team of seasoned professionals will help you based on your overall education background and work experience. This will be followed by interview preparation along with mock interviews (if required)
- Job referrals are based on the requirements we get from various organizations, HR consultants and large pool of AnalytixLabs' ex-students working in various companies.
- No one can truthfully provide job guarantee, particularly for good quality job profiles in Analytics. However, most of our students do get multiple interview calls and good career options based on the skills they learn during training. For this there will be continuous support from our side for as long as required.



# Time and Investment

**Training:** 50 hours Pre-learning video based training + 138 hours live training + Practice,  
INR ~~70,000~~ 52,000 + 18% GST - **with ~25% Combo discount**

**Timing:** 6 hours per weekend live training (Saturday & Sunday 3 hours each) + Practice

**Training mode:** Fully interactive live online class / Classroom (In Gurgaon, Bangalore and Noida center only)

*(In addition to the above, you will also get access to the recordings for future reference and self study)*

**Components:** Learning Management System access for courseware like class recordings - study material, Industry-relevant project work

**Certification:** Participants will be awarded a **Co-branded Global certificate** on successful completion of the stipulated requirements including an evaluation

# We provide trainings both in 'fully interactive live online' and classroom\* mode



\*Classroom only available at Gurgaon, Bangalore and Noida center

# Contact Us

Visit us on: <http://www.analytixlabs.in/>

For course registration, please visit: <http://www.analytixlabs.co.in/course-registration/>

For more information, please contact us: <http://www.analytixlabs.co.in/contact-us/>

Or email: [info@analytixlabs.co.in](mailto:info@analytixlabs.co.in)

Call us we would love to speak with you: (+91) 9555219007

Join us on:

Twitter - <http://twitter.com/#!/AnalytixLabs>

Facebook - <http://www.facebook.com/analytixlabs>

LinkedIn - <http://www.linkedin.com/in/analytixlabs>

Blog - <http://www.analytixlabs.co.in/category/blog/>

# Visit Us

## **Gurgaon Address:**

GF 382, Sector 29,  
Adjoining IFFCO Chowk Metro  
Station (Gate 2),  
Next to Magicpin Office,  
Gurgaon,  
Haryana- 122001

## **Bengaluru Address:**

Bldg 41, First floor,  
14th Main Road, Near BDA  
complex,  
(Landmark: Max store)  
Sector 7, HSR Layout, Bengaluru  
Karnataka - 560102

## **Noida Address:**

First Floor, A-78,  
A Block, Sector 2,  
(Adjacent to Sector 15 metro  
station)  
Noida,  
Uttar Pradesh - 201301