

# 1

## Building Blocks for Data Visualization & Analytics (Pre – Learning)

### 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

### 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

# 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 lookup

### 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
- DD 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/ cartesian 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 (PowerBI)

## (3/3)

### Introduction

- Introduction to Power BI
- Installing Power BI Desktop (Signup for PowerBI)
- Various Options in Power BI Desktop
- Views in Power BI Desktop
- Template Apps
- Task pipeline when your working on a project

### Data Preparation and Modelling

- Connect and Retrieve data from different sources (csv, excel etc.)
- Query editor in Power BI
- Power Query for cleaning the data
- Power Query Functions – Text, Date, Numeric
- Power Query Conditional Columns
- Clean & transform data with Query Editor
- Define data granularity
- Combining data – Merging & Appending
- Fill Down in Power BI, Grouping, Transpose, Unpivot, Data Types, Replace errors and values, Keep and Remove rows, Add Remove and Go To Columns
- Work with relationships and cardinality
- Types of Relationships (1:1, 1: Many, Many:1)
- Optimizing for performance
- PBIDS Files

### Data Analysis Expressions (DAX)

- Introduction to DAX
- Calculated tables, Columns & Measures
- Time Intelligence in DAX
- Frequently Used DAX functions in Real time (Calendar Functions, Filter Functions, Information functions, Text Functions, Logical Functions, Math functions, Parent & Child functions etc..)

### Reports Development (Visuals in Power BI)

- Introduction to work with Power BI visuals
- Reports Development in Power BI
- Working with Different Visuals /Charts
- Formatting Options in Reports
- Use a slicer to filter visualizations
- Working with Filters (Page Level, Include/Exclude, Report Level, Cross report Filter)
- Download & use Custom Visuals from the galaxy
- Add an R or Python visual
- Work with key performance indicators
- Project to Implement the learning's

# 2 Data Visualization & Analytics (PowerBI)

## (3/3)

### Data Driven Story Reports:

- Introduction to create a data-driven story
- Design a report layout
- Add buttons, bookmarks, and selections
- Creating Interactive reports with bookmarks
- Design report navigation
- Use interactions and drill through
- Comment on reports
- Tune report performance
- Optimize reports for mobile use

### Dashboards:

- Introduction to dashboards
- Configure data alerts
- Explore data by asking questions
- Add a dashboard theme
- Pin a live report page to a dashboard
- Configure a real-time dashboard
- Configure data classification
- Set mobile view

### Advanced / Other Power BI Concepts

- Row level Security (Static Row Level Security, User login based row level security, Organizational level security)
- Dynamic Measures, Filters, Axis in Charts
- Power BI Template file
- Wallpapers, Themes (create custom themes)

### Power BI Analytics:

- Explore statistical summary
- Identify outliers with Power BI visuals
- Group and bin data for analysis
- Use the Analyze feature
- Use advanced analytics custom visuals
- Review Quick insights

### Publishing workbooks and Workspace

- Publishing the Reports, Dashboards, APP
- Share data with Colleagues and Others
- Publish & manage report to the web
- Create an app workspace and add users
- Create a QR code to share a tile
- Embed a report in share point Online

# 3 Building Blocks for Python and ML (Pre-Learning)

## Programming Basics

- Introduction to programming
- Computer programs and business use
- Database and its requirement in the software applications.
- What is an IDE - Integrated development environment.
- Different programming languages, High level vs Low level languages,
- Language translators - Compiler and Interpreter, Why syntax rules?
- Programming basics: variables, INC rules: Identifier Naming Conventions, Datatypes, Operators.
- Control flow statements: Conditional statements and Loops.
- Functions and UDFS.
- Logic building and Pseudo codes.

## 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 (Baes Theorem)

## Introduction to Mathematical foundations

- Sets & Functions
- Introduction to Linear Algebra
  - Matrices Operations
- Introduction to Calculus
  - Derivatives & Integration
  - Maxima, minima
  - Area under the curve

# 3 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,/python)
- Concept of Packages - Important packages
- 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)

# 3 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/Density plots
- Bar Charts/Stacked charts
- Pie Charts
- Box Plots
- Scatter Plots
- Sub Plots

## 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
- Hypothesis Testing & Applications



# 4 Predictive Modeling & 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 vs. Reinforcement 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
- Nonlinear 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
- Naive Bayes Classifier
- Decision Trees
- Ensemble Learning - Bagging, Random
- Forest, Adaboost, Gradient Boost, XGBoost
- Support Vector Classifier

## Unsupervised Learning

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

## Recommender Systems

- Market Basket Analysis (MBA)
- 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

# 5 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)

# 6 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 MultiLayer Perceptron
- Understand Backpropagation - Using Example
- Implementation of ANN in Python- Keras

## Introduction to Google Colab/Kaggle workbooks

## 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

# 7 Introduction to ML-Ops & Model Deployment (self-paced)

## Introduction to MLOps

- What is MLOps
- MLOps vs. DevOps vs. Data Engineering
- Why MLOps is important?
- ML Engineering Pipeline
- How to implement MLOps?
- Understand end to end MLOps solution

## Deployment of ML Model in the cloud

- What is model deployment?
- Ways of deployment of models
- Introduction to Flask
- How to create simple app?
- Deployment of ML model in the cloud

# 8 Industrial & Functional Sessions (Domain Understanding)

## Business problem solving

- Introduction to Business requirement understanding
- BRD VS. FRD
- Creating analytics project documentation
- Creating visualizations based on the requirement
- How to define KPI's & understand design elements?
- Define business outcomes and create visualizations from results of the analysis
- How to prepare for client discussions?
- Introduction to Analytics Project Management

## Introduction to Industry & Functional Sessions

- Understand to Data Sources for Various Industries
- Overview of different industries & applications
  - Retail/E-Commerce
  - Banking & Financial Services
  - Telecom & Networking
  - Digital & Social Media
  - Travel, Transportation and Hospitality
  - Utilities
  - Manufacturing & FMCG

## Marketing Analytics

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

## Risk Analytics

- Introduction to Risk Function
- Enterprise Risk Function
- Credit Scoring (Application & Behavioral)
- Fraud Analytics
- Applications in across industries

## Operation Analytics

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

## Digital Analytics(Web Analytics)