



# MASTERS IN DATA SCIENCE WITH POWER BI



[ConsoleFlare.com](https://ConsoleFlare.com)

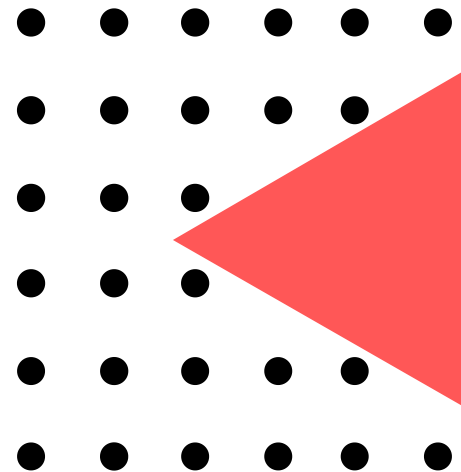


# DATA SCIENCE

## Why Data Science?

Without the expertise of professionals who turn cutting-edge technology into actionable insights, Big Data is nothing. Today, more and more organizations are opening up their doors to big data and unlocking its power—increasing the value of a data scientist who knows how to tease actionable insights out of gigabytes of data.

Most data scientists in the industry have advanced and training in statistics, math, and computer science. Their experience is a vast horizon that also extends to data visualization, data mining, and information management. It is fairly common for them to have previous experience in infrastructure design, cloud computing, and data warehousing.



## Job Opportunities

Data Analyst, Data Scientist, Product Analyst, Machine Learning Engineer, Business Analyst

## Estimated Base Salary

\$118,000/year





# BIGDATA AND BEYOND

## WHY BIGDATA?

Large content repositories house unstructured data such as documents, and companies like **Google**, **Oracle**, **SAP** and **NetSuite** and others are utilizing, sharing and storing a great deal of structured and unstructured information.

Processing this huge amount of data is not possible with the traditional means of computing.

### THE SOLUTION IS BIGDATA

This rapid growth is expected to create thousands of Big Data Engineering jobs in the next 3-5 years, presenting significant opportunities for individuals to create their careers and grow.

As per recent hiring trends, skills required in building Big Data applications and platforms (For e.g., technologies like Hadoop & Spark) are sought after in the technology world today. Big Data finds applications across industries - ranging from IT/BPO to E-commerce/Internet to Public Sector. This fast evolving technology presents a promising career path with excellent long term prospects.

For many entertainment, technology and media organizations, Big Data analytics is the key to retaining subscribers, securing advertising revenues, and understanding customer needs.

### Job Opportunities

Big Data Engineer, Big Data Admin, Big Data Architect, Big Data Analyst

### Average Salary

Big Data Engineer - \$ 140,000 per year  
Big Data Admin - \$100,000 per year  
Big Data Architect - \$150,000 per year  
Big Data Analyst - \$120,000 per year



**BIGDATA  
ANALYTICS INDUSTRY  
IN INDIA  
IS TO BE WORTH  
US \$ 16 BN  
IN 2025**

# MACHINE LEARNING & AI

## Why Machine Learning & Artificial Intelligence?

We live in a world with humans and machines. Humans learn from their experiences and machines follow instructions. What if, machines could learn from experiences too and perform tasks more efficiently without being explicitly programmed to do so.

This is where machine Learning and artificial intelligence comes into play.

Machine learning is a core sub-area of artificial intelligence; it enables computers to get into a mode of self-learning without being explicitly programmed. And artificial intelligence is to create technology that allows computers and machines to function in an intelligent manner.

ANY SUFFICIENTLY  
ADVANCED  
TECHNOLOGY IS  
INDISTINGUISHABLE  
FROM MAGIC

- ARTHUR C. CLARKE

It is growing very vast:

The self-driving Google car, cyber fraud detection, online recommendation engines—like friend suggestions on Facebook, Netflix showcasing the movies and shows you might like, and “more items to consider” and “get yourself a little something” on Amazon—are all examples of applied machine learning.

And who knows you might make your own Jarvis.

A study conducted by Analytics India Magazine reveals that there are more than 78,000 jobs in the Machine Learning and AI are lying vacant in India.

Machine Learning jobs for freshers may vary between ₹ 699,807- 891,326. per year

### Job Opportunities

Machine Learning Engineer, Research Scientist, Business Intelligence Developer, Algorithm specialists. And many more

### Average Salary

\$134,135 per year - \$169,930 per year





# Course Modules

Python Programming



Data Manipulation



Wielding Bigdata through Pyspark



Importing & Cleaning Data



Data Visualization



Statistical exploration in Python



Machine Learning



Live Data Streaming



Reporting



Project



# 1 Python Programming

- Course Introduction
- Environment Setup
- Python Part 1 | Fundamentals of Programming
- Python Part 2 | Working with in-built Functions & Modules
- Milestone Project 1
- Python Part 3 | String & String Formatting
- Python Part 4 | Python Data Structures
- Python Part 5 | Conditional Statements
- Python Part 6 | Loop and Iterator
- Milestone Project 2
- Python Part 7 | Exception handling
- Python Part 8 | User defined Functions and Modules
- Python Part 9 | File Handling

# 2 Data Manipulation

## Numpy

Learning about NumPy which is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays

- NumPy Arrays
- Array Constructor
- Introduction To Arrays
- NumPy. Range ()
- LINPAC
- Create 2-D Array
- Indexing & Slicing Arrays
- Indexing In 1-D Arrays
- Indexing In 2-D Arrays
- Slicing In 1-D Arrays
- Slicing In 2-D Arrays
- Indexing
- Slicing
- Vectorization & Broadcasting In Arrays
- Scalar Vectorization
- Array Comparison
- New Axis

# Pandas

This section is about Pandas which is an open-source library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language

- Pandas And Data Manipulation
- Dropping/Deleting Columns
- Data Frame Indexing
- Series
- Series. Apply ()
- Creating Series
- Applying Method To A Series
- Data Frame & Basic Functionality
- Printing Columns
- DataFrame. Head ()
- Creating DataFrame
- Descriptive Statistical Function
- Data Frame Manipulation
- Mean()
- Correlation
- Indexing & Missing Values
- Grouping & Reshaping

## 3 Wielding Big Data through Pyspark

- How big is Big Data?
- Cluster Computing
- Hadoop Architecture
- In-memory Computation
- Apache Spark Architecture
- Hadoop vs Spark
- What Is Spark
- Why Pyspark
- Databricks setup and forming cloud cluster
- Introduction To Data Frames
- Data Frames Basics
- Dataframe Operations
- Group By And Aggregate Operations
- Handling Missing Data
- Triggering Sql query in pyspark
- Dates And Timestamp
- MileStone Project



## 4 Importing & Cleaning Data

Learning about database performing CRUD operations and importing data from a database and store in DataFrame using Pandas Package

- Hosting A Database On Cloud/Local System
- Python Sql Connectivity
- CRUD Operation On Database Tables Through Python
- Importing Data Into Dataframe Of Pandas
- Processing & Cleaning Data Through Pandas Methods
- Dealing with missing values

## 5 Data Visualization

Data Visualization is the presentation of data in a graphical format

- Introduction to Data Visualization
- Data Visualization with Python

### Matplotlib

- Matplotlib is a plotting library for the Python programming language
- Using the matplotlib.pyplot interface
- Types of Charts
- Bar Chart using Matplotlib
- Stacked Bar Chart using Matplotlib
- Histogram Chart using Matplotlib
- Scatter plots using Matplotlib
- Adding data to an Axes object
- Customizing your plots
- Customizing data appearance
- Creating a grid of subplots
- Area Plot for Indexed Data

### Seaborn

- Seaborn is a library for making statistical graphics in Python. It is built on top of matplotlib and closely integrated with pandas data structures
- Seaborn vs Matplotlib
- How To Show Seaborn Plots
- How To Use Seaborn With Matplotlib Defaults
- How To Set xlim or ylim in Seaborn
- Box and whiskers

## 6 Statistical exploration in Python

- Graphical exploratory data analysis
- Plotting and computing simple summary statistics
- Quantitative exploratory data analysis
- Describe salient features of a dataset with numbers.
- Probabilistically-- Discrete variables
- Statistical inference rests upon probability
- Thinking probabilistically-- Continuous variables

## 7 Introduction to Machine Learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed

- Machine Learning Introduction
- History of Machine Learning
- Examples of Machine learning
- Difference between Artificial intelligence and Machine learning and Deep Learning
- Machine learning Life cycle
- Types of Machine Learning
- What are Features and labels

### Supervised Learning

Supervised machine learning algorithms are designed to learn by example.

- Introduction to Supervised Learning
- Different Types of Supervised Learning
- Difference between Regression and Classification

# Regression

Regression analysis is a form of predictive modelling technique, which investigates the relationship between a dependent and independent variable

- Introduction to Regression
- Introduction of Simple Linear Regression
- Simple Linear Regression Understanding Algorithm
- Simple Linear Regression Example with Python
- Introduction of Multiple Linear Regression
- Multiple Linear Regression Understanding Algorithm
- Multiple Linear Regression Example with Python
- Introduction of Polynomial Regression
- Polynomial Regression Understanding Algorithm
- Polynomial Regression Example with Python
- Introduction of Backward Elimination
- Backward Elimination Understanding Algorithm
- Backward Elimination Example with Python

# Classification

Classification separates observations into groups based on their characteristics.

- Introduction to Classification and Algorithm
- Introduction Of Naive Bayes
- Naive Bayes Understanding The Algorithm
- Naive Bayes Example with Python
- Introduction Of K-Nearest Neighbours (K-NN)
- K-Nearest Neighbours (K-NN) Understanding The Algorithm
- K-Nearest Neighbours (K-NN)Example with Python
- Introduction Of Support Vector Machine (SVM)
- Support Vector Machine (SVM) Understanding The Algorithm
- Support Vector Machine (SVM)Example with Python
- Introduction Of Decision Tree Classification
- Decision Tree Classification Understanding The Algorithm
- Decision Tree Classification Example with Python
- Introduction Of Random Forest Classification
- Random Forest Classification Understanding The Algorithm
- Random Forest Classification Example with Python
- Introduction Of Logistic Regression
- Logistic Regression Understanding The Algorithm
- Logistic Regression Example with Python
- Classification Model Selection in Python



# Unsupervised Learning

Unsupervised learning is the training of machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance

- Introduction to Unsupervised Learning
- Introduction Of K-Means Clustering
- K-Means Clustering Understanding The Algorithm
- K-Means Clustering Example with Python
- Introduction Of Hierarchical Clustering
- Hierarchical Clustering Understanding The Algorithm
- Hierarchical Clustering Example with Python
- Introduction Of Fuzzy K-means
- Fuzzy K-means Understanding The Algorithm
- Fuzzy K-means Example with Python

## 8 Live Data Streaming

Streaming data is data that is continuously generated by different sources. Such data should be processed incrementally using Stream Processing techniques without having access to all of the data.

- Spark Streaming With Python
- Data in Motion
- Documentation Example
- Converting Your Terminal Into A Stream For Live Data And Reading From That Live Stream
- Spark Stream Twitter Project (Create A Live Stream For Tweets Coming In And Creating Live Dashboard With Data Analysis And Visualization)
- Quizzes And Doubt Clearing Sessions

# 9 Reporting With Microsoft Power BI

## Introduction

- What is Power BI?
- Environment setup
- Understanding the Power BI Desktop Workflow
- Exploring the Interface of the Data Model
- Understanding the Query Editor Interface

## Working in the Query Editor

- Introduction
- Connecting Power BI Desktop to Source Files
- Keeping & Removing Rows
- Working with Filters
- Removing Empty Rows
- Appending Queries
- Formatting Data & Handling Formatting Errors
- Pivoting & Unpivoting Data
- Splitting Columns
- Data Model - The Star Schema
- Query Duplicates vs References
- Removing Duplicates
- Understanding "Join Kind"
- Understanding "Extract"
- Creating & understanding the concept of the FACT-Table
- Performance Optimization
- Assignment 1 (Loading your project data efficiently)

# Understanding the Data Model: Data & Relationship View

- Introduction
- Understanding Relationships
- M-Language vs DAX (Data Analysis Expressions)
- Understanding the DAX Basics
- Diving Into Operators
- DAX Functions - An Overview
- Applying the DAX Basics
- Calculated Columns vs Measures
- Creating Measures with Measures
- Assignment 2 (Manipulating & processing the above loaded data)

# Creating Visuals in the Report View

- Introduction
- Understanding Basic Visual Concepts
- Creating our First Visuals
- Understanding Tooltips & Interactions
- Diving Intro Hierarchies & Drill Mode
- Data Colors & Conditional Formatting
- Formatting Report Pages
- Using the Slicer
- Understanding Default Summarization & Sorting
- Syncing Slicers
- Understanding Filter Types (Visual, Page & Report)
- Creating Combined Visuals & Waterfalls
- Using Custom Visuals
- Assignment 3 (Visualizing the above processed data)

# Deploying Project to the Cloud with Power BI Pro (Service)

- Introduction
- Understanding Workspaces
- Working with Reports
- Creating Dashboards
- Refreshing Data with Gateways
- Sharing Data From "My Workspace"
- Publishing an App



- Student Project Assistance
- Revisiting the course timeline
- Future enhancements and Tips
- Bonus: Selected Features by Example & Refresher Project

○—————○

## Congratulations on becoming a Data Scientist

### Our Students are at.



**Deloitte.** infoedge



**MAQ**  
Software



Infosys®

ggk tech  
An ACS Group Company

**nagarro**

GlobalLogic®



**Cognizant**

**accenture**



**Red Hat**



**IBM**

**CONDUENT**

# Get in touch.

---

**Website | [Consoleflare.com](https://consoleflare.com)**

By Arbre Creations Pvt. Ltd.

**Meet @ | A-93, Sector 4 Near Sector 16 metro station, Noida, Uttar Pradesh 201301**

**Call @ | +91-7376114845,  
+91-7428334865  
+91-8810480042**

**Mail @ | [info@consoleflare.com](mailto:info@consoleflare.com)**