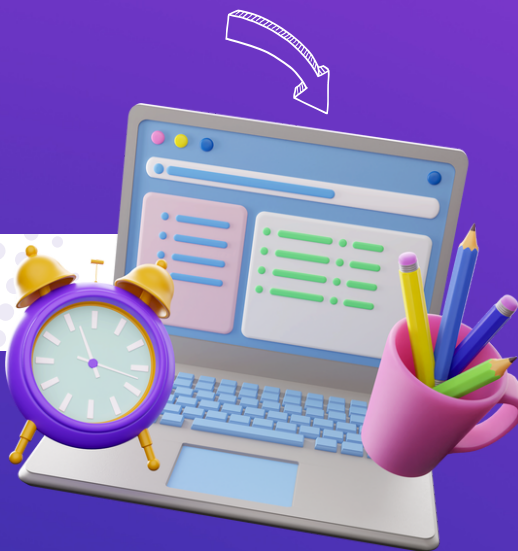


DATA  
SCIENCE  
&  
ML/AI MASTERS

THE ULTIMATE ROAD TO SUCCESS

SYLLABUS



IN ASSOCIATION WITH:



# About Us

Zep Analytics is an AI-driven EdTech Company focused on providing quality education in the form of pre-recorded & online boot camps on the latest technologies such as Data Science, Machine Learning, AI, Data Analytics, and many more. Apart from that, we also have many unique features on the platform such as a reward point system that can provide incentives to our users based on their activities.

## OUR VISION

Our vision is to bridge the gap between career expectations and reality via innovative teaching approaches.

## TEACHING BY EXPERTS

Industry experts in data science, machine learning, and deep learning with multiple years of corporate as well as teaching experience.

# Trainer

Satyajit Pattnaik has more than 14 years of industrial experience and more than a decade of experience into Data Science, Analytics & AI

- Certified Data Scientist &
- Tensorflow Certified Developer

30k+ followers    70k+ subscribers

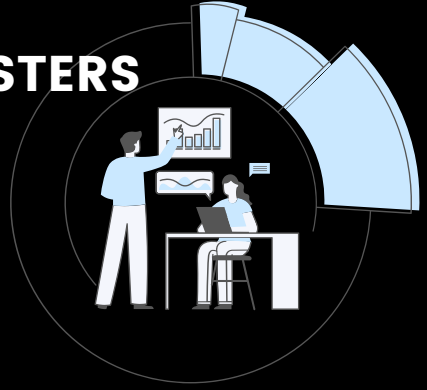


[linkedin.com/satyajitpattnaik](https://linkedin.com/satyajitpattnaik)  
[youtube.com/c/satyajitpattnaik](https://youtube.com/c/satyajitpattnaik)

The above data is as of April, 2024



# DATA SCIENCE, ML & AI MASTERS 2024



## Features:

- ✓ Pre-Recorded Contents (Gen AI Included)
- ✓ Course Completion Certificate
- ✓ 1 Month Internship Certificate\*
- ✓ Knowledge on end to end Data Analytics, Science & AI
- ✓ Guidance & Mentorship from Satyajit Pattnaik & his team.
- ✓ Interview preparation guides and e-books
- ✓ Unlimited 24\*7 Community support

The outburst of data is transforming businesses. Companies - big or small - are now expecting their business decisions to be based on data-led insight. Data specialists have a tremendous impact on business strategies and marketing tactics.

The demand for data science & analytics professionals are on the rise while the supply remains low, thus creating great job opportunities for individuals within this field.

Today, it is almost impossible to find any brand that does not have social media presence; soon, every company will need data analytics professionals. This makes it a wise career move that has a future in business.

So here is a course that will help you to become a Data Scientist, ML Engineer & AI Engineer.

The course covers all the skills required for above job role in detail.

## MODULE 1 : Introduction to the course

A general introduction to the course covering all the important aspects that you should know when you are in the Data Science industry. This module will be a basic module that will help you in framing your analytical perspective.

### Inside the Module:

1. What is Data Science?
2. Need for Data Science in today's Industry
3. Introduction to Problem Solving
4. Framing the problem
5. Analyzing the problem
6. Implementing the problem with a proper solution
7. Different problem-solving frameworks to adopt

## MODULE 2 : Python for Data Science

Python is one of the most important language when it comes to Data Science. Due to rich python libraries data analysis tasks become quite easy and interesting. So in this module we are going to start with Python for Data Science where all the important needful concepts will be covered in detail.

### Inside the Module:

1. Basics of Python
2. Data Structures in Python
3. Introduction to Problem Solving
4. Control Structures (if, if-Else, elif, nested if-else, etc.)
5. Loops in Python
6. Functions
7. Object-Oriented Programming
8. Exception Handling and Database Programming
9. Data Science Libraries: Pandas, NumPy, Matplotlib, Seaborn, Plotly etc

## MODULE 3 : Statistics

Data Science is all about statistical analysis and learning. In this module, we are going to learn about different statistical concepts which are not important for Data Science and also from an interview point of view.

### Inside the Module:

#### 3.1) Introduction to Statistics

- 1) Type of statistics
- 2) Type of Data
- 3) Different Sampling Techniques
- 4) Measures of Central Tendency (Mean, Median, Mode)
- 5) Measures of Dispersion (Variance and Standard Deviation)

#### 3.2) Intermediate Statistics

- 1) Introduction to Probability
- 2) Permutations and Combinations
- 3) Conditional Probability and Bayes Theorem



- 4) Introduction to Gaussian Distribution and properties
- 5) Central Limit Theorem
- 6) Covariance and Correlation
- 7) Pearson And Spearman Rank Correlation
- 8) Binomial Distribution
- 9) Bernoulli Distribution
- 10) Poisson Distribution

### 3.3) Advanced Statistics

- 1) Introduction to Confidence Intervals
- 2) Z-test and t-test
- 3) Hypothesis Testing
- 4) ANNOVA
- 5) AB Testing

**Note: All the concepts of the module will be covered theoretically and practically as well.**

## MODULE 4 : Data Wrangling

Once you have the core skill of programming covered – dip your feet in the nitty – gritty of working with data by learning how to wrangle and visualize them.

### Inside the Module:

1. Reading CSV, JSON, XML, and HTML files using Python
2. NumPy & Pandas
3. Relational databases and data manipulation
4. Scipy libraries
5. Loading, cleaning, transforming, merging & reshaping data

## MODULE 5 : RDBMS and SQL for Data Science

Data is stored in Databases and databases can be of various types. But when we talk about Databases what first comes in our mind is SQL (Structured Query Language). Even in most of the data analytics job description SQL is always mentioned.

### Inside the Module:

#### 5.1) SQL

- 1) Introduction to Databases
- 2) What is SQL?
- 3) Introduction to Schemas and its types
- 4) Introduction to Relational and Non-Relational Schemas
- 5) Different integrity constraints in SQL
- 6) Details on different keys in SQL
- 7) Data Definition Language (DDL)
- 8) Data Manipulation Language (DML)

- 9) Data Control Language (DCL)
- 10) Joins in SQL
- 11) Data import and export
- 12) Functions in SQL
- 13) Nested Queries
- 14) Views in SQL
- 15) Stored Procedures
- 16) Window Functions
- 17) Python connectivity for SQL

## MODULE 6 : Machine Learning

Machines have increased the ability to interpret large volumes of complex data. Combine aspects of Computer science with statistics to formulate algorithms that helps machines draw insights from structured and unstructured data

### Inside the Module:

1. Introduction to Machine Learning
2. Building models using below algorithms
3. Evaluation Metrics
4. Regression Algorithms
5. Classification Algorithms
6. Clustering Algorithms
7. Association Rule Learning
8. Ensemble Techniques
9. Feature Engineering
10. Feature Scaling & Feature Selection
11. Time Series Analysis
12. Hyper Parameter Optimization

## MODULE 7 : Natural Language Processing

NLP helps resolve ambiguity in language and adds useful numeric structure to the data for many downstream applications, such as speech recognition or text analytics.

### Inside the Module:

1. Text Data Cleaning
2. Advanced Topics: Bag of words, n-Grams, Word Embedding, Embedding Layer etc
3. Relationships, Joins and Unions
4. Part of Speech Tagging
5. Neural Networks: RNNs, LSTMs, Bi-Directional RNNs, GRU etc
6. Use Case: Text Classification
7. Topic Modelling

## MODULE 8 : Deep Learning & Computer Vision

Go beyond superficial analysis of data by learning how to interpret them deeply. Use deep learning nets to uncover hidden structures in even unlabeled and unstructured data.

### Inside the Module:

1. Basics of Neural Networks
2. Linear Algebra
3. Implementation of neural networks
4. Basics of TensorFlow
5. CNNs, RNNs, ANNs
6. Generative models
7. Encoder/Decoders, Seq-to-Seq models
8. Transfer Learning
9. Pre-Trained Model
10. Introduction to Computer Vision
11. Basics of Object Detection
12. Transformers

## MODULE 9: Gen AI & LLMs

Unlock the potential of language generation with Gen AI and LLM, delving deeper to create rich, meaningful content beyond the surface

### Inside the Module:

1. Fundamentals
2. Gen AI: Text, Audio, Images, Video
3. Text & Chatbots: LLMs
4. Image Generation
5. Audio Generation
6. Prompt Engineering
7. Langchain Embeddings
8. Use Cases: Chatbots



## **MODULE 10 : Deploying ML Models using Flask (ON PREM), Streamlit (ON PREM) & CLOUD**

Go beyond the model preparation by learning how to deploy a machine learning model using Flask, Streamlit and Cloud etc.

### **Inside the Module:**

1. Concepts on Apache Flask, Waitress
2. Various cloud providers – AWS, Azure, GCP etc
3. How to prepare a model?
4. How to save a model?
5. Deploying on streamlit
6. Deploy a model using Apache Flask
7. Deploy a model using Cloud

## **MODULE 11 : DATA VISUALIZATION - Tableau & Power BI [Add-on]**

Learn the Art of Dashboarding and Data Visualization using Power BI and Tableau.

### **Inside the Module:**

1. Introduction to both the tools
2. Connecting to files & databases
3. Data filters
4. Calculations and Parameters
5. Creating Dashboards
6. Data Blending
7. Creating superimposed graphs

## Module 12: Data Engineering [BASICS] [Add-on]

Get prepared on the concepts of Data Engineering as an add-on feature.

### Inside the Module:

1. Data Engineering Basics
2. ETL vs ELT
3. Data Bricks, Data Lakes, Delta Lakes etc
4. Data Marts
5. Data Architectures

## Module 13: LIVE PROJECTS

We Provide 50+ Projects during the course , and various industry level use case for practice.

Students get one to one mentoring while solving various industry level projects too. Real world industry Projects and deployment in AWS and Azure Cloud.

# Ready to take the next steps?

Zep Analytics offers a complete data science & analytics training taught by expert instructors with a fun, interactive, and beginner-friendly approach.

The courses start with the fundamentals, cover in-demand programming languages like Python, and SQL, visualization tools like Power BI and Tableau, and finish off with advanced specialized courses, including state-of-the-art Machine and Deep Learning.

