



Full-Stack Data Science Course Content

Comprehensive Curriculum Integrating Python, R, SQL, Hadoop, Azure AI, AI/ML/DL, Statistics, Analytics, Power BI, Generative AI, and Prompt Engineering

Module 1: Foundations of Data Science

Introduction to Full-Stack Data Science

- Overview of the Data Science Lifecycle.
- Role of a Full-Stack Data Scientist.
- Tools and Technologies in Full-Stack Data Science.

Mathematics and Statistics for Data Science

- Linear Algebra, Calculus, and Probability Basics.
- Descriptive and Inferential Statistics.
- Hypothesis Testing and Confidence Intervals.
- Statistical Distributions and Applications.

Programming for Data Science

- **Python:**
 - Basics of Python Programming.
 - Advanced Python (OOPs, File Handling, Exception Handling).
 - Libraries: NumPy, Pandas, Matplotlib, Seaborn, SciPy, and Scikit-learn.
- **R Language:**
 - Basics of R Programming.
 - Data Manipulation with dplyr and tidyr.
 - Visualization with ggplot2.
 - Statistical Modeling in R.



Module 2: Database Management and Big Data

SQL and Relational Databases

- Database Fundamentals and SQL Basics.
- Advanced SQL Queries: Joins, Subqueries, Views, and Indexing.
- Data Manipulation and Transaction Management.
- Integration with Python and R.

Big Data and Hadoop

- Introduction to Big Data Concepts.
- Hadoop Architecture and Ecosystem (HDFS, MapReduce, YARN).
- Working with Hive and Pig for Data Processing.
- Basics of Apache Spark for Big Data Analytics.

Module 3: Data Preprocessing and Feature Engineering

- Data Cleaning and Transformation Techniques.
- Handling Missing Data, Outliers, and Data Imbalance.
- Feature Scaling, Encoding, and Selection.
- Dimensionality Reduction Techniques: PCA, t-SNE, LDA.

Module 4: Data Visualization and Business Analytics

Data Visualization Tools

- Creating Visualizations with Matplotlib, Seaborn, Plotly, and Dash.
- Interactive Dashboards with Power BI.
- Tableau Basics: Connecting, Building Dashboards, and Storytelling.



Business Analytics

- Fundamentals of Business Analytics.
- KPI Definition and Business Intelligence.
- Case Studies: Sales, Marketing, and Operational Analytics.

Module 5: Machine Learning (ML) and Deep Learning (DL)

Machine Learning Basics

- Types of Machine Learning: Supervised, Unsupervised, Reinforcement.
- Algorithms: Linear Regression, Logistic Regression, Decision Trees, Random Forest, SVM, and k-NN.
- Model Evaluation Metrics: Accuracy, Precision, Recall, F1-Score, and ROC-AUC.

Deep Learning Basics

- Neural Networks: Architecture and Training.
- Deep Learning Frameworks: TensorFlow, PyTorch, and Keras.
- Advanced Topics: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).

Module 6: Artificial Intelligence and Azure AI

Artificial Intelligence (AI)

- Introduction to AI: Applications and Use Cases.
- Fundamentals of Knowledge Representation and Search Algorithms.

Azure AI

- Introduction to Azure AI Services.
- Azure Machine Learning Studio: Building and Deploying Models.
- Cognitive Services: Computer Vision, NLP, and Speech Recognition.



Module 7: Advanced Topics

Generative AI

- Introduction to Generative AI Concepts.
- Applications of GANs and Autoencoders.
- Working with OpenAI APIs and Generative Models.

Prompt Engineering

- Basics of Prompt Engineering for Generative AI.
- Techniques for Crafting Effective Prompts.
- Case Studies: Chatbots, Summarization, and Text-to-Image Applications.

Module 8: Cloud Computing and Deployment

Cloud Platforms

- Basics of Cloud Computing and Its Importance in Data Science.
- Deploying Machine Learning Models on Azure and AWS.
- Working with Cloud Databases and Storage Solutions.

Model Deployment

- Building APIs using Flask and FastAPI.
- Deploying Models with Docker and Kubernetes.
- MLOps: Monitoring and Maintaining Deployed Models.

Module 9: Capstone Projects and Case Studies

1. **Retail Sales Prediction:** Using ML and Power BI for actionable insights.
2. **Healthcare Chatbot:** Leveraging Generative AI and Prompt Engineering.



3. **Big Data Processing:** Using Hadoop and Spark for large-scale analytics.
4. **Real-Time Fraud Detection:** Deep learning with TensorFlow.
5. **IoT Data Analysis:** Cloud-integrated dashboard with Power BI.

Module 10: Career Development

- Resume Writing and Portfolio Building for Data Science Roles.
- Preparing for Technical Interviews and Mock Sessions.
- Industry Projects to Strengthen Real-World Skills.

Course Outcomes

By the end of the course, learners will:

1. Master Python, R, SQL, Hadoop, and Azure AI for data science.
2. Build and deploy machine learning and deep learning models.
3. Gain expertise in Generative AI and Prompt Engineering.
4. Analyze data and create impactful visualizations using Power BI.
5. Complete hands-on projects to showcase real-world applications.