# Masters Program in Data Science (6 Months)

Course Overview:

The Masters Program in Data Science is designed to provide students with a comprehensive understanding of the various aspects of data science, including data analysis, machine learning, big data, and data visualization. Over the course of six months, students will develop hands-on skills in the tools and techniques required to extract valuable insights from data, build predictive models, and deploy data-driven solutions to solve real-world problems.

**Month 1: Introduction to Data Science and Python Programming**

- Course Overview
  - Introduction to the field of Data Science
  - Role of Data Science in decision making and business intelligence
- Python Programming for Data Science
  - Basics of Python programming (variables, loops, conditionals, functions)
  - Working with data structures (lists, dictionaries, sets, tuples)
  - Introduction to libraries: Numpy, Pandas, Matplotlib, and Seaborn
- Data Exploration and Cleaning
  - Data wrangling and preprocessing techniques
  - Handling missing data, duplicates, and outliers
  - Exploratory Data Analysis (EDA)

**Month 2: Statistics and Probability for Data Science**

- Descriptive Statistics
  - Measures of central tendency (mean, median, mode)

  - Measures of dispersion (variance, standard deviation, range)
  - Visualizing data (box plots, histograms, bar plots)
- Inferential Statistics
  - Probability theory and distributions (normal, binomial, Poisson)
  - Hypothesis testing and confidence intervals
  - Sampling techniques and Central Limit Theorem
- Statistical Analysis in Python
  - Using libraries like Scipy and Statsmodels for statistical analysis

**Month 3: Data Visualization and Communication**

- Data Visualization Principles
  - Importance of effective data visualization
  - Types of charts and graphs: histograms, pie charts, line plots, heatmaps
- Advanced Visualization with Python
  - Using Matplotlib, Seaborn, and Plotly for dynamic plots
  - Interactive data visualizations
  - Dashboard development basics using Dash
- Storytelling with Data
  - Communicating insights through visualizations
  - Creating data-driven presentations
  - Building reports and storytelling with charts and graphs

**Month 4: Machine Learning Fundamentals**

- Introduction to Machine Learning
  - Supervised vs Unsupervised learning

  - Overview of ML algorithms: regression, classification, clustering
- Supervised Learning Algorithms
  - Linear Regression, Logistic Regression, and Decision Trees
  - Model evaluation techniques (cross-validation, train-test split)
  - Overfitting and underfitting
- Unsupervised Learning Algorithms
  - K-Means Clustering, DBSCAN, and Hierarchical Clustering
  - Principal Component Analysis (PCA) for dimensionality reduction
- Hands-on Projects
  - Implementing algorithms with scikit-learn
  - Working with real-world datasets for training and testing models

**Month 5: Deep Learning and Big Data Analytics**

- Introduction to Deep Learning
  - Neural Networks and the basics of deep learning
  - Building a neural network using TensorFlow/Keras
  - Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)
- Natural Language Processing (NLP)
  - Text preprocessing and feature extraction (TF-IDF, Word2Vec, GloVe)
  - Sentiment analysis, named entity recognition, and topic modeling
  - Working with NLP tools (NLTK, spaCy)
- Big Data Technologies
  - Introduction to Big Data and Hadoop Ecosystem
  - Distributed data processing with Spark
  - Introduction to NoSQL databases (MongoDB, Cassandra)
  - Using PySpark for big data analysis

**Month 6: Capstone Project and Industry Applications**

- Capstone Project
  - Applying knowledge to solve a real-world data science problem
  - Students will work on a large dataset, build models, and present their findings
  - Focus on end-to-end project workflow (data collection, cleaning, analysis, model building, and evaluation)
- Industry Applications of Data Science
  - Data Science in business, healthcare, finance, and marketing
  - Ethical considerations in data science (privacy, fairness, and bias)
  - Case studies of successful data science applications in various industries
- Career Preparation and Soft Skills
  - Resume building and LinkedIn profile optimization for Data Science roles
  - Mock interviews and interview preparation
  - How to present data science projects to non-technical audiences