DATASCIENCE PROFESSIONAL PROGRAM – 2025 EDITION

"Transforming data into actionable intelligence"

Market Demand Note

Data Science continues to be one of the most in-demand careers, with businesses seeking talent who can not only clean and analyze data but also deliver actionable insights. This professional program bridges the gap between theory and practice, making learners industry ready.

Duration: 12 Weeks | Mode: Online/Offline

📘 Data Science Professional Program — 2025 Edition

Key Tools & Technologies:
Python, Pandas, NumPy, Matplotlib, Seaborn, SQL, Power BI / Tableau, Scikit-learn, Jupyter Notebook, Git, GitHub, Kaggle Datasets.

Learning Objectives

By the end of this program, learners will be able to:

1. Perform complex data wrangling and cleaning operations.

2. Conduct exploratory data analysis (EDA) with advanced visualization.

3. Build predictive models using machine learning.

4. Communicate insights using dashboards and storytelling techniques.

5. Apply data science workflows to solve real-world problems.

Table of Contents

Detailed Content

Week 1: Advanced Python for Data Science

- Python Refresher:
  Variables, loops, functions, list/dict/set comprehensions, generators, lambda functions.

- Writing Production-Ready Code:
  Error handling, logging, modular functions, docstrings.

Data Science Professional Program @ https://udaan.x-fuzion.com/

- Efficient Data Structures:
  Choosing the right structure for speed and memory.

- Hands-on:

  - Solve complex data problems using advanced Python.

  - Write reusable, documented functions for data tasks.

Current Trends:
Python remains the lingua franca for data science. Focus on clean, modular, and efficient code is a core expectation in industry roles.

Week 2: Data Manipulation with Pandas & NumPy

- Pandas Mastery:
  Advanced indexing, multi-index DataFrames, merging/joining, groupby operations, handling missing data, datetime manipulation.

- NumPy for Speed:
  Vectorized operations, broadcasting, memory-efficient arrays.

- Handling Outliers:
  Statistical and visual methods for detection and treatment.

- Hands-on:

  - Clean and transform messy real-world datasets (e.g., sales, IoT sensor data).

  - Benchmark Pandas vs. NumPy for large-scale operations.

Current Tools:
Pandas 2.x (with PyArrow backend for speed), Polars (for big data), cuDF (GPU-accelerated).

Week 3: Exploratory Data Analysis (EDA)

- Descriptive Stats:
  Mean, median, skewness, kurtosis, quantiles.

- Correlation & Covariance:
  Detect relationships, multicollinearity.

- Visual EDA:
  Pair plots, heatmaps, boxplots, violin plots, distplots.

Data Science Professional Program @ https://udaan.x-fuzion.com/

- Hands-on:
    - Perform EDA on a Kaggle dataset.
    - Identify data quality issues, anomalies, and insights.

Industry Relevance:
EDA is the foundation of every data science project—expected in interviews and daily work.

Week 4: Data Visualization Mastery

- Matplotlib/Seaborn Advanced:
  Custom themes, subplots, annotations, interactive elements.

- Plotly/Dash:
  Interactive dashboards, real-time updates.

- Storytelling:
  Design visuals that communicate insights clearly.

- Hands-on:
    - Build a dashboard that tells a data story.
    - Compare static vs. interactive visualization tools.

Current Tools:
Plotly, Altair, Tableau/Power BI integration, Streamlit for Python apps.

Week 5: SQL for Data Science

- Advanced SQL:
  Complex joins, subqueries, CTEs, window functions (ROW_NUMBER, RANK, LAG/LEAD).

- Aggregation & Transformation:
  Pivoting, rolling aggregates, case statements.

- Python + SQL:
  Connect to databases, execute queries, fetch results.

- Hands-on:
    - Write and optimize SQL queries on a large database.
    - Integrate SQL with Python for automated ETL.

Industry Demand:
SQL is mandatory for data roles. Window functions and query optimization are high-value skills.

Week 6: Introduction to Machine Learning

- Supervised Learning:
  Regression (linear, ridge, lasso), classification (logistic, decision trees, random forest).

- Model Evaluation:
  Accuracy, precision, recall, F1, ROC-AUC, confusion matrix.

- Feature Engineering Basics:
  Creating new features, handling categorical variables.

- Hands-on:

  - Train and evaluate models on a real dataset.

  - Compare performance across algorithms.

Current Tools:
scikit-learn, XGBoost, LightGBM, CatBoost.

Week 7: Unsupervised Learning & Dimensionality Reduction

- Clustering:
  K-Means, DBSCAN, hierarchical.

- Dimensionality Reduction:
  PCA, t-SNE, UMAP.

- Hands-on:

  - Segment customers using clustering.

  - Visualize high-dimensional data in 2D/3D.

Industry Applications:
Customer segmentation, anomaly detection, feature engineering for supervised models.

Week 8: Feature Engineering & Model Optimization

- Encoding Categoricals:
  One-hot, target, embeddings.

- Scaling & Normalization:
  StandardScaler, MinMaxScaler, RobustScaler.

- Cross-Validation:
  k-fold, stratified, time series splits.

- Hyperparameter Tuning:
  Grid search, random search, Bayesian optimization.

- Hands-on:

    - Optimize a model pipeline end-to-end.

    - Deploy the best model.

Current Trends:
AutoML tools (Auto-sklearn, H2O), feature stores, MLOps integration.

Week 9: Time Series Analysis

- Time Series Components:
  Trend, seasonality, noise.

- Forecasting:
  ARIMA, SARIMA, Prophet, LSTM (intro).

- Hands-on:

    - Forecast sales or stock prices.

    - Evaluate models on time series data.

Industry Tools:
statsmodels, Prophet, PyCaret, Darts.

Week 10: Storytelling with Data & Dashboard Development

- Data Storytelling:
  Structuring narratives, highlighting insights, avoiding misrepresentation.

- Dashboard Tools:
  Power BI, Tableau, Streamlit, Dash.

- Hands-on:

Data Science Professional Program @ https://udaan.x-fuzion.com/

- Build an interactive dashboard with business KPIs.

- Present findings to a simulated stakeholder.

Job Market:
Dashboard and storytelling skills are critical for senior data roles and consulting.

Week 11: Project Work & Real-World Data

- Kaggle/Industry Datasets:
  Work on a complex, real-world problem (e.g., churn prediction, credit risk).

- Full Pipeline:
  Data ingestion, cleaning, EDA, feature engineering, modeling, evaluation, deployment.

- Hands-on:

  - Collaborate in teams.

  - Document and version control your work (Git).

Portfolio Value:
Real projects are the #1 differentiator for job seekers.

Week 12: Capstone Project Presentation

- Project Showcase:
  Present your end-to-end data solution to peers and instructors.

- Peer Review:
  Give and receive constructive feedback.

- Portfolio Submission:
  GitHub repo with code, README, visuals, and presentation.

Capstone Ideas:
Predict customer churn, forecast COVID-19 trends, analyze credit risk—all using real data.

Capstone Project Examples

- Predicting Customer Churn: Use historical usage data to identify at-risk customers (common in telecom, SaaS).

Data Science Professional Program @ https://udaan.x-fuzion.com/

- COVID-19 Dashboard: Track cases, deaths, vaccinations; forecast trends; visualize geospatial patterns.

- Credit Risk Analysis: Build a model to assess loan applicant risk for a bank.

Current Tools & Industry Trends (2025)

- Python Ecosystem: Pandas, NumPy, scikit-learn, XGBoost, LightGBM, CatBoost, Polars, cuDF.

- SQL: PostgreSQL, MySQL, BigQuery, Redshift, Synapse—windowing and optimization are key.

- Visualization: Plotly, Altair, Streamlit, Dash, Tableau, Power BI.

- MLOps: MLflow, DVC, Airflow, Docker, Kubernetes—deployment and automation are expected.

- Cloud: AWS, GCP, Azure—data engineering is increasingly cloud-native.

- Big Data: Spark, Dask—scalable processing for large datasets.

- Dashboarding: Interactive, self-service dashboards are the norm in enterprises.

Key Differentiators for Job Seekers

- End-to-end project experience (from raw data to deployed model/dashboard).

- Clean, modular, documented code (GitHub portfolio).

- SQL mastery (complex queries, optimization).

- Storytelling and visualization (clear communication of insights).

- Exposure to MLOps and cloud tools (deployment, automation, scalability).