

3-MONTH ONLINE PROGRAM

Data Science: From Fundamentals to Industry-Ready

A structured, mentor-guided 12-week program designed for early professionals, graduate students, and career transitioners. Build the technical depth, practical fluency, and industry confidence to land and thrive in a data science role — starting from day one.

12 Weeks Structured learning with weekly milestones	3 Phases Foundation → Intermediate → Advanced
Live Mentoring Weekly 1:1 or small-group sessions	Industry-Aligned Real tools, real workflows, real outcomes

Why This Course Matters Today

Data science is no longer a niche specialty — it is a core business function. Organizations across every sector are making critical decisions driven by data, and the demand for skilled practitioners far outpaces supply. This course bridges the gap between academic knowledge and what industry actually expects on day one.

The Market Reality

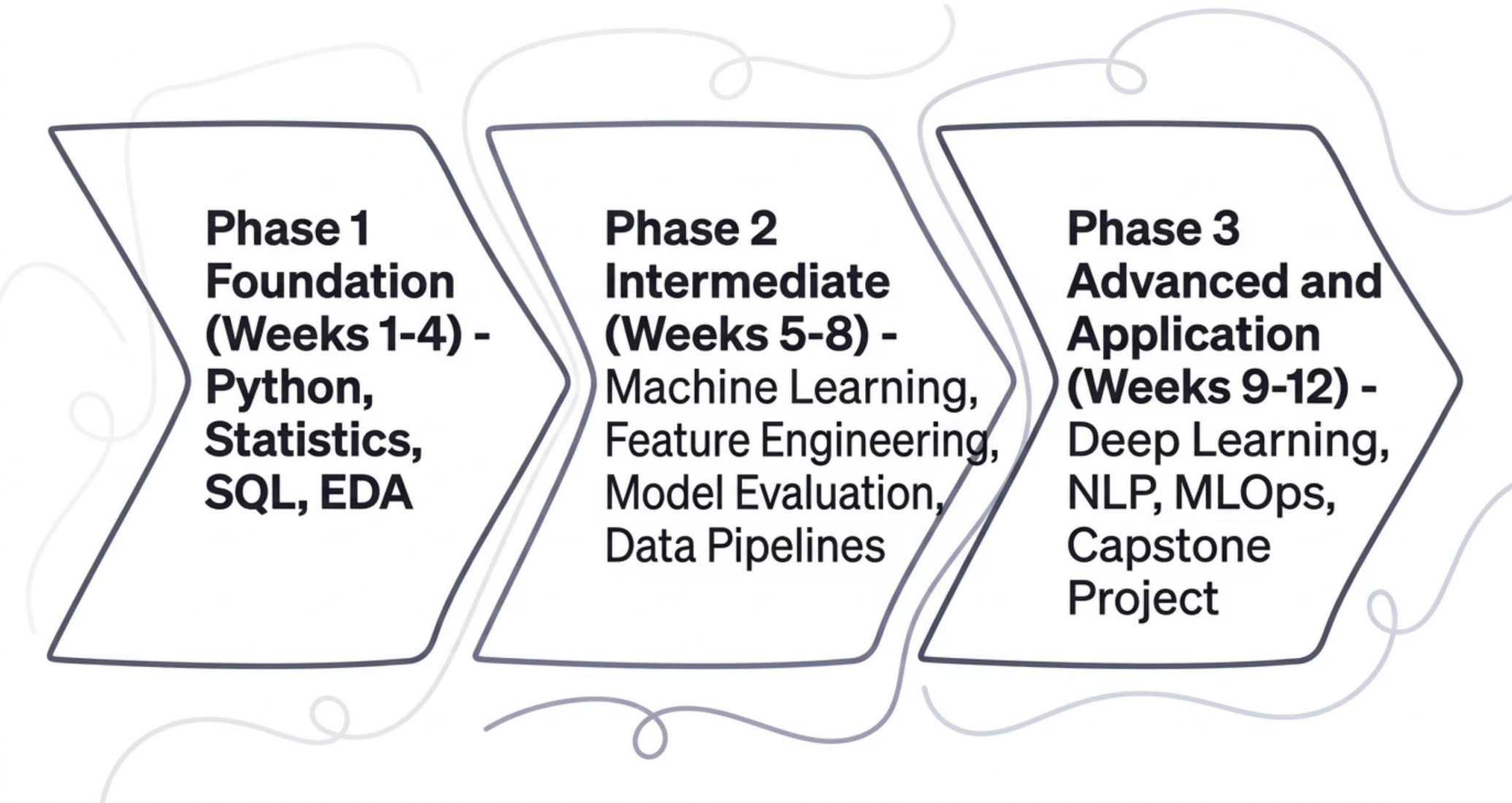
- Data science roles are projected to grow 36% through 2031 (BLS)
- Median salary exceeds \$120K in the US market
- 90% of Fortune 500 companies now have dedicated data teams
- Most entry-level job postings require hands-on project experience

What This Course Delivers

- End-to-end data science workflow mastery
- Proficiency in Python, SQL, and ML libraries
- Portfolio-ready projects and case studies
- Interview preparation and career positioning
- Direct mentorship from an experienced industry practitioner

Program Structure at a Glance

The 12-week curriculum is divided into three progressive phases, each building on the last. Every week combines a recorded concept session with a live mentoring touchpoint to ensure accountability, clarity, and application.



Weekly Rhythm

- **Weekend Session:** Recorded concept-focused lecture (2–3 hrs)
- **Live Mentoring:** 1:1 or small group doubt clearing (1 hr)
- **Assignments:** Practical exercises due before next session

Learning Ratio

- **70–80%:** Core concepts and strong fundamentals
- **10–15%:** Industry use cases and real applications
- **10%:** Projects, reviews, and career prep

PHASE 1 • WEEKS 1–4

Foundation Phase

Build the unshakeable bedrock every data scientist needs. This phase ensures you speak the language of data fluently — from Python scripting and SQL querying to statistical reasoning and exploratory analysis. No shortcuts; strong fundamentals compound into superior outcomes later.

1 Week 1 Python for Data Science - Python environment setup (Anaconda, Jupyter) - Data types, loops, functions, and OOP basics - NumPy arrays and vectorized operations - Pandas DataFrames: load, filter, transform - File I/O: CSV, JSON, Excel handling - Mini-project: Clean and summarize a public dataset	2 Week 2 Statistics & Probability Foundations - Descriptive statistics: mean, median, variance, skewness - Probability distributions: Normal, Binomial, Poisson - Hypothesis testing: t-test, chi-square, p-values - Correlation vs. causation — a critical distinction - A/B testing fundamentals and business framing - Exercise: Run a hypothesis test on e-commerce data
3 Week 3 SQL & Data Wrangling - SQL fundamentals: SELECT, WHERE, GROUP BY, JOIN - Subqueries, CTEs, and window functions - Working with databases: PostgreSQL and SQLite - Python + SQL integration via SQLAlchemy - Data cleaning: missing values, duplicates, outliers - Assignment: Build a customer analytics query report	4 Week 4 Exploratory Data Analysis & Visualization - EDA framework: understand before modeling - Matplotlib and Seaborn for visual storytelling - Plotly for interactive dashboards - Identifying trends, distributions, and anomalies - Feature correlation heatmaps and pair plots - Mini-project: Full EDA report on a retail dataset

PHASE 1 • OUTCOMES

Foundation Phase: What You'll Be Able to Do



Write Production-Quality Python

Confidently manipulate datasets with Pandas and NumPy, write reusable functions, and handle real-world data formats without assistance.



Conduct Full EDA

Systematically explore any dataset, identify patterns, flag data quality issues, and present findings through compelling visualizations.



Query Any Database

Write complex SQL queries including window functions and CTEs to extract insights directly from relational databases used in industry.



Apply Statistical Reasoning

Frame business questions statistically, run hypothesis tests, and interpret results with confidence — avoiding common pitfalls like p-hacking.

Intermediate Phase: Machine Learning in Practice

With a solid foundation in place, you now enter the heart of data science — machine learning. This phase emphasizes not just knowing algorithms, but knowing *when*, *why*, and *how* to apply them in real business contexts. Feature engineering and model evaluation receive dedicated focus because they determine whether a model actually works in production.

1

Week 5

Supervised Learning — Regression

- ML workflow: train/test split, cross-validation
- Linear and polynomial regression deep dive
- Ridge, Lasso, and ElasticNet regularization
- Gradient descent intuition and implementation
- Evaluation: RMSE, MAE, R^2 — picking the right metric
- Project: Predict housing prices with Scikit-learn

2

Week 6

Supervised Learning — Classification

- Logistic regression and decision boundaries
- Decision Trees, Random Forest, Gradient Boosting (XGBoost)
- Support Vector Machines and k-NN
- Class imbalance: SMOTE, class weights, threshold tuning
- Evaluation: Confusion matrix, ROC-AUC, F1-score
- Assignment: Customer churn prediction model

3

Week 7

Unsupervised Learning & Feature Engineering

- K-Means, DBSCAN, Hierarchical Clustering
- PCA and dimensionality reduction techniques
- Feature creation: binning, encoding, polynomial features
- Feature selection: importance scores, Recursive Feature Elimination
- Anomaly detection fundamentals (Isolation Forest)
- Mini-project: Customer segmentation for a retail brand

4

Week 8

Model Evaluation, Tuning & Pipelines

- Hyperparameter tuning: GridSearchCV, RandomizedSearchCV, Optuna
- Bias-variance tradeoff: diagnosing under/overfitting
- Scikit-learn Pipelines for reproducible workflows
- Model persistence: saving and loading models (joblib, pickle)
- Experiment tracking basics with MLflow
- Assignment: Build an end-to-end ML pipeline with tuning

PHASE 3 · WEEKS 9–12

Advanced & Application Phase

The final phase elevates your skillset into the territory that top employers value most: deploying models, working with unstructured text, and building systems that work at scale. Week 12 is dedicated entirely to the capstone project — your flagship portfolio piece and the culmination of everything learned.

1

Week 9

Deep Learning Foundations

- Neural network architecture: layers, activations, backprop
- Building models with TensorFlow and Keras
- CNNs for image data: convolution, pooling, feature maps
- Overfitting prevention: Dropout, Batch Normalization
- Transfer learning with pre-trained models (ResNet, MobileNet)
- Mini-project: Image classifier using transfer learning

2

Week 10

Natural Language Processing (NLP)

- Text preprocessing: tokenization, stemming, lemmatization
- TF-IDF, Word2Vec, and GloVe embeddings
- Sentiment analysis and text classification
- Named Entity Recognition (NER) with spaCy
- Introduction to Transformer models and Hugging Face
- Assignment: Build a product review sentiment analyzer

3

Week 11

MLOps, Deployment & Cloud Basics

- Model deployment with Flask and FastAPI REST endpoints
- Containerization with Docker: building and running containers
- Cloud platforms: AWS SageMaker, GCP Vertex AI overview
- CI/CD pipelines for ML: GitHub Actions basics
- Monitoring model drift and data quality in production
- Mini-project: Deploy a classification model as an API

4

Week 12

Capstone Project & Career Prep

- End-to-end project: define problem → model → deploy
- GitHub portfolio cleanup and README best practices
- Resume and LinkedIn optimization for data science roles
- Mock technical interview: ML concepts + case study
- Capstone demo presentation with mentor feedback
- Final deliverable: Deployed project + presentation deck

Tools & Technology Stack

Every tool in this program is actively used in industry right now. You won't just learn what these tools do — you'll build muscle memory through repeated, contextual use across projects and assignments throughout all 12 weeks.



Core Languages

- Python 3.x (primary)
- SQL (PostgreSQL, SQLite)
- Bash scripting basics



Data & ML Libraries

- Pandas, NumPy, Scikit-learn
- XGBoost, LightGBM
- TensorFlow / Keras, spaCy



Visualization

- Matplotlib, Seaborn
- Plotly, Streamlit
- Power BI / Tableau basics



MLOps & Cloud

- MLflow, FastAPI, Docker
- GitHub Actions (CI/CD)
- AWS SageMaker / GCP basics



Development Env.

- Jupyter Notebooks / JupyterLab
- VS Code with Python extensions
- Google Colab (cloud notebooks)



Collaboration

- Git & GitHub (version control)
- Notion / Confluence (docs)
- Slack / Teams (async comms)

Predicting Customer Churn at a Telecom Company

The Business Problem

A major US telecom provider was losing 8–10% of its subscribers quarterly. Retaining an existing customer costs 5× less than acquiring a new one, but the retention team was using a blanket outreach strategy with low conversion and wasted spend.

Role of Early Professionals

Junior data scientists cleaned and processed 3 years of billing, usage, and support ticket data. They built and benchmarked three classification models and presented results to the product team.

How Data Science Was Applied

- Built a binary classification model (XGBoost) to predict churn probability per customer
- Features: average monthly spend, call drops, contract type, customer tenure, support contacts
- Used SHAP values to explain model decisions to non-technical stakeholders
- Integrated predictions into CRM system for automated targeted outreach

Measurable Business Impact

- Churn rate reduced from 9.2% to 5.8% within two quarters
- Retention campaign ROI improved by 34%
- Tools: Python, Scikit-learn, XGBoost, SHAP, Tableau, Salesforce CRM

INDUSTRY EXAMPLE 2

Demand Forecasting for a Retail Supply Chain

The Business Problem

A national grocery chain with 400+ stores was experiencing 15% inventory waste due to overstocking perishables, alongside frequent stockouts of high-velocity SKUs during weekends and holidays. Manual Excel-based forecasting couldn't keep pace with demand variability.

Role of Early Professionals

Analysts on the data team took ownership of cleaning historical POS data, building baseline models, and running weekly backtests. They also created automated reporting dashboards for the supply chain team.

How Data Science Was Applied

- Time-series forecasting using Facebook Prophet and LightGBM with lag features
- External signals incorporated: weather data, public holidays, local events
- Store-level models trained for top 200 SKUs; generic model for the tail
- Forecasts delivered weekly via automated pipeline to the procurement system

Measurable Business Impact

- Perishable waste reduced by 22% in six months
- Stockout incidents dropped 18% — improving in-store customer satisfaction
- Tools: Python, Prophet, LightGBM, Apache Airflow, AWS S3, Power BI

INDUSTRY EXAMPLE 3

NLP-Powered Review Intelligence at an E-Commerce Platform

The Business Problem

A mid-sized e-commerce marketplace was receiving 50,000+ product reviews monthly but had no scalable system to extract actionable insights. The customer experience team was manually reading samples — missing critical product quality signals and seller compliance issues.

Role of Early Professionals

Data science trainees and junior NLP engineers built the text preprocessing pipeline, fine-tuned the classification model, and created the Streamlit dashboard — presenting results in a sprint review to the product and CX teams.

How Data Science Was Applied

- Fine-tuned a BERT-based model (Hugging Face) for multi-label sentiment classification
- Topic modeling with LDA to identify top complaint and praise themes by category
- Named Entity Recognition (NER) to flag product defects and brand mentions
- Built a Streamlit dashboard for real-time review monitoring by category managers

Measurable Business Impact

- Product return rate decreased 11% after proactive defect identification
- CX team response time to negative trends cut from 2 weeks to 48 hours
- Tools: Python, spaCy, Hugging Face Transformers, Streamlit, PostgreSQL

Skills Progression Map

This course is architected around a deliberate three-stage skill progression. Each stage unlocks the next, ensuring you're never memorizing tools in isolation — you're building compounding, transferable competence that employers recognize immediately.



Beginner (Weeks 1–4)

Python scripting · SQL queries · Descriptive stats · EDA · Data cleaning · Basic visualization



Intermediate (Weeks 5–8)

Supervised & unsupervised ML · Feature engineering · Model tuning · Scikit-learn pipelines · MLflow tracking



Job-Ready (Weeks 9–12)

Deep learning · NLP & transformers · Model deployment · Docker & cloud basics · Capstone portfolio project

 **Benchmark:** By the end of Week 12, you should be able to take a raw business problem from data → model → deployed solution — and explain every decision along the way. That's the standard interviewers are measuring you against.

Career & Job Role Alignment

This curriculum is deliberately mapped to the roles most accessible to candidates with 0–2 years of experience post-program completion. The skills you develop don't fit just one job title — they open multiple doors across industries.

Data Analyst

Core skills used: SQL, Python, EDA, Pandas, Visualization

Interpret data, build dashboards, and report insights to business stakeholders. High demand across all industries.

Avg. US Salary: \$75K–\$95K

Junior Data Scientist

Core skills used: ML algorithms, feature engineering, model evaluation, Python pipelines

Build and validate predictive models to solve specific business problems in product, finance, or operations.

Avg. US Salary: \$95K–\$120K

ML Engineer

Core skills used: MLOps, Docker, FastAPI, cloud deployment, pipelines

Deploy, monitor, and maintain machine learning systems in production environments at scale.

Avg. US Salary: \$110K–\$140K

NLP / AI Engineer

Core skills used: NLP, transformers, Hugging Face, text classification, embeddings

Build language-based AI applications — from chatbots to review analyzers and document intelligence tools.

Avg. US Salary: \$115K–\$145K

Business Intelligence Developer

Core skills used: SQL, EDA, visualization, Power BI / Tableau, statistical reasoning

Design and maintain BI reporting infrastructure that drives strategic decision-making across the organization.

Avg. US Salary: \$80K–\$105K

Capstone Project: Your Portfolio Flagship

Week 12 culminates in a full end-to-end capstone project that serves as the centerpiece of your professional portfolio. This is not a textbook exercise — it is a real-world simulation graded on industry standards.

Exploratory Analysis

Investigate data and engineer features.

Model Building

Train models and evaluate performance.

Problem Definition

Define goals and choose dataset.

Deployment & Presentation

Deploy solution and present results.

Deliverables

- A public GitHub repository with clean, documented code
- A deployed model accessible via a web endpoint or Streamlit app
- A 10-slide presentation deck summarizing the business problem and results
- A written README that reads like a professional project brief

Evaluation Criteria

- Problem framing and business relevance
- Data quality, feature engineering decisions
- Model selection rationale and performance metrics
- Code quality, reproducibility, and documentation
- Clarity of the final presentation and demo

Your Roadmap After Completing the Course

Completing the program is not the finish line — it is the launch ramp. Here is the recommended pathway to move from "course complete" to "offer accepted" within 60–90 days of graduation.



 **Pro Tip:** Start applying during Week 11 — before the capstone is complete. Use the capstone as a talking point in interviews. "I'm currently building X to solve Y business problem" is a powerful conversation opener.