🚀12-Week Data Science & GenAI Roadmap (4 Hours/Day)

This roadmap is designed for a **3-month intensive prep** (4 hours/day, Mon–Fri) for Data Science, Machine Learning, and Generative AI roles in 2025, especially in Dubai's tech job market.

It combines **Python, ML, Cloud, Big Data, Generative AI, and interview prep**, with portfolio-ready projects every step of the way.

🚃Month 1 — Foundations for Data Science

Goal: Build strong fundamentals in Python, data handling, statistics, and SQL.

Week 1: Python Foundations

- **Mon:** Python basics → notebook `python_basics.ipynb`
- **Tue:** Control flow (loops, conditionals)
- **Wed:** Functions & modules
- **Thu:** OOP (classes, objects)
- **Fri:** File I/O & error handling

Week 2: Data Handling with Python

- **Mon:** NumPy arrays & vectorization
- **Tue:** Pandas basics
- **Wed:** Pandas joins & pivot tables
- **Thu:** Data cleaning & missing values
- **Fri:** Visualization with Matplotlib & Seaborn

Week 3: Statistics & Probability

- **Mon:** Descriptive stats
- **Tue:** Probability & Bayes theorem
- **Wed:** Hypothesis testing (t-test, chi-square)
- **Thu:** Correlation & causation
- **Fri:** Linear algebra refresher (vectors, matrices)

Week 4: SQL & Databases

- **Mon:** SQL basics (SELECT, WHERE)
- **Tue:** Joins (INNER, LEFT, RIGHT)
- **Wed:** Aggregations (GROUP BY, HAVING)
- **Thu:** Window functions (RANK, ROW_NUMBER)

- **Fri:** Subqueries, CTEs, optimization

🚃Month 2 — Applied ML, Cloud & MLOps

Goal: Train, tune, and deploy ML models. Learn cloud/data tools.

Week 5: Supervised ML Models

- **Mon:** Regression models
- **Tue:** Decision Trees & Random Forests
- **Wed:** Gradient Boosting (XGBoost, LightGBM)
- **Thu:** Evaluation metrics
- **Fri:** Imbalanced data handling (SMOTE)

Week 6: Advanced ML & Interpretability

- **Mon:** Feature engineering (PCA, encoding)
- **Tue:** Explainability (SHAP, LIME)
- **Wed:** Time Series OR Recommender Systems
- **Thu:** Hyperparameter tuning
- **Fri:** Ensembles & stacking

Week 7: MLOps & Deployment

- **Mon:** Pipelines & reproducibility
- **Tue:** REST APIs with FastAPI/Flask
- **Wed:** Experiment tracking (MLflow)
- **Thu:** Docker basics
- **Fri:** CI/CD basics (GitHub Actions)

Week 8: Big Data & Cloud Tools

- **Mon:** SQL on BigQuery/Snowflake
- **Tue:** PySpark basics
- **Wed:** Databricks/Colab Pro setup
- **Thu:** Scaling & optimization
- **Fri:** End-to-end ML + Spark + API project

🚃Month 3 — Generative AI, LLMs & Interview Prep

Goal: Work with LLMs, LangChain, RAG, and prepare for interviews.

Week 9: NLP Foundations

- **Mon:** Text preprocessing, embeddings
- **Tue:** Sentiment analysis
- **Wed:** Transformers (HuggingFace)
- **Thu:** Fine-tuning DistilBERT
- **Fri:** Project: Text classification app

Week 10: Generative AI & LangChain

- **Mon:** LLM basics (GPT, instruction tuning)
- **Tue:** Prompt engineering
- **Wed:** LangChain basics
- **Thu:** Vector DBs (FAISS, Pinecone, Chroma)
- **Fri:** RAG pipeline project

Week 11: Capstone Project

- **Mon:** Project scoping & dataset selection
- **Tue:** Data preprocessing & EDA
- **Wed:** Model training & evaluation
- **Thu:** Deployment (API + Docker)
- **Fri:** Documentation & storytelling

Week 12: Interview & Profile Prep

- **Mon:** Core DS/ML interview questions
- **Tue:** Case study practice
- **Wed:** System design for ML pipelines
- **Thu:** Mock interview session
- **Fri:** Final GitHub + LinkedIn polish

Deliverables by End of 12 Weeks

- 5–6 polished GitHub repos (Python, ML, GenAI)
- At least 1 capstone project deployed with API + Docker
- Documented learning roadmap & notes in GitHub
- Recruiter-ready LinkedIn & GitHub profile
- Interview prep (Q&A, case studies, mock interviews)

✨Tip: Treat each week as a **mini-sprint**. By Friday, aim to push a clean notebook/repo to GitHub so your portfolio grows steadily.