

6-Month Data Scientist Roadmap (Beginner → Job-Ready)

Month 1: Core Python + Math Foundation

Goals:

- Get strong in Python basics
- Build solid math & stats foundation

Learn:

- Python (loops, lists, functions)
- Libraries: numpy, pandas
- Math: mean, median, variance
- Statistics: probability, hypothesis testing

Projects:

- Student marks analyzer
- COVID data exploration

Month 2: Data Visualization + EDA

Goals:

- Learn data exploration & visualization

Learn:

- matplotlib, seaborn, plotly
- Data cleaning: NaN, duplicates, outliers

Projects:

- IPL dataset analysis
- Netflix dataset visualization

Month 3: Machine Learning Basics

Goals:

- Understand ML & build models

Learn:

- scikit-learn
- Algorithms: Linear/Logistic Regression, Decision Trees
- Model evaluation: accuracy, confusion matrix

Projects:

- House Price Prediction
- Loan Approval Prediction

Month 4: Advanced ML + Feature Engineering

Goals:

- Learn advanced algorithms & preprocessing

Learn:

- KNN, SVM, XGBoost
- Feature scaling, encoding

Projects:

- Credit card fraud detection
- Customer churn prediction

Month 5: Specialization + Real-World Projects

Choose one specialization:

- NLP: text classification, sentiment analysis
- Computer Vision: CNNs
- Analytics: Power BI dashboards

Projects:

- Twitter Sentiment Analysis
- Image Classifier
- Sales Dashboard

Month 6: Portfolio + Interview Prep

Goals:

- Build portfolio & prepare for jobs

Tasks:

- Create 3–5 GitHub projects
- Build portfolio website
- Polish LinkedIn & resume
- Practice interview Qs

Bonus:

- Kaggle competitions
- Apply for internships

Tech Stack Summary:

Area	Tools
Programming	Python, R (optional)
Data	Pandas, NumPy, SQL
Visualization	Matplotlib, Seaborn, Power BI, Tableau
ML	Scikit-learn, XGBoost
Deep Learning	TensorFlow, Keras
Deployment	Streamlit, Flask
Version Control	Git, GitHub