Data Science Learning Pathway: Beginners

Math Foundation

1 Linear Algebra
basic operations, matrices, vectors

2 Calculas
derivatives, integrals, basic optimization

3 Statistics
descriptive statistics, basic distributions, hypothesis testing

Programming

4 Python Basics
syntax, data structures, control flow, functions

5 Numpy
arrays, basic operations

6 Pandas
Series, DataFrames, basic data manipulation

7 Matplotlib
basic visualizations

Data Manipulation & Analysis

8 SQL Basics
SELECT, FROM, WHERE, GROUP BY, JOIN

9 Data Cleaning
handling missing values, basic outlier detection

10 Basic EDA
summary statistics, simple visualizations

Machine Learning

11 Supervised Learning
linear regression, logistic regression, decision trees

12 Model Evaluation Basics
accuracy, precision, recall, overfitting/underfitting

- ● Math
- ● Programming
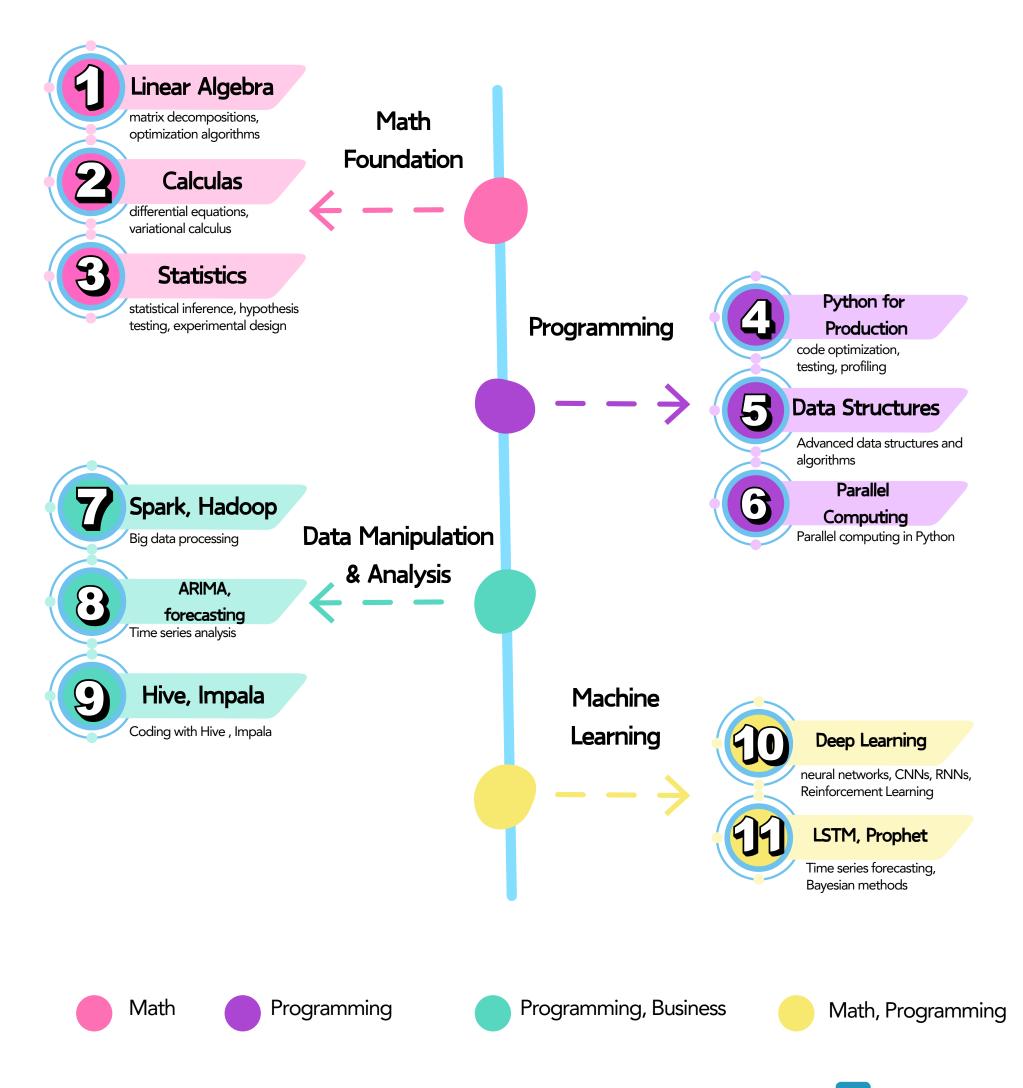- ● Programming, Business
- ● Math, Programming

DineshLal

Data Science Learning Pathway: Intermediate

Math Foundation

1 Linear Algebra
linear transformations, eigenvalues/eigenvectors

2 Calculas
multivariable calculus, advanced optimization

3 Statistics & Probability
advanced distributions, confidence intervals, Bayesian inference

Programming

4 OOP Python
Object-oriented programming in Python

5 Advanced NumPy/Pandas
indexing, grouping, merging

6 Data visualization
advanced Matplotlib, Seaborn

Data Manipulation & Analysis

7 Advanced SQL
subqueries, window functions, stored procedures

8 Feature engineering
creating new features, handling categorical variables

9 Advanced EDA
distribution analysis, correlation analysis

Machine Learning

10 Supervised Learning
ensemble methods, support vector machine

11 Unsupervised Learning
k-means clustering, PCA

● Math   ● Programming   ● Programming, Business   ● Math, Programming

DineshLal

Data Science Learning Pathway: Advanced

Math Foundation

1 Linear Algebra
matrix decompositions, optimization algorithms

2 Calculas
differential equations, variational calculus

3 Statistics
statistical inference, hypothesis testing, experimental design

Programming

4 Python for Production
code optimization, testing, profiling

5 Data Structures
Advanced data structures and algorithms

6 Parallel Computing
Parallel computing in Python

Data Manipulation & Analysis

7 Spark, Hadoop
Big data processing

8 ARIMA, forecasting
Time series analysis

9 Hive, Impala
Coding with Hive , Impala

Machine Learning

10 Deep Learning
neural networks, CNNs, RNNs, Reinforcement Learning

11 LSTM, Prophet
Time series forecasting, Bayesian methods

- Math
- Programming
- Programming, Business
- Math, Programming

DineshLal

Data Science Learning Pathway: Other Important Areas

Business

- Domain knowledge specific to your industry
- Metrics and KPIs relevant to your business goals
- Communicating results to stakeholders

Soft Skills

- Communication (explaining technical concepts to non-technical audiences)
- Storytelling (creating narratives with data)
- Critical thinking and problem-solving
- Teamwork and collaboration