# ULTIMATE DATA ANALYST ROADMAP -2025

Disclaimer - In case you are a fresher, preparing for a Data field interview, or an experienced person who wants to transition their field then you can follow this Roadmap.

## Definition of Data Analyst

A Data Analyst in 2025 is a professional who collects, processes, and interprets data to help organizations make informed decisions. With advancements in AI, automation, and cloud computing, the role has evolved beyond traditional reporting to include predictive analytics, AI-powered insights, and real-time data storytelling.

## How is a Data Analyst Different in 2025?

- **AI-Powered Analytics** – Analysts rely on AI-driven tools to generate insights faster.
- **Cloud-First Approach** – Data is processed & analyzed in cloud environments for scalability.
- **DataOps & Automation** – Manual tasks are reduced with automated pipelines.
- **Self-Service Analytics** – Business users can generate insights using AI chatbots.
- **AI & ML Awareness** – While not Data Scientists, Analysts use low-code ML tools for predictions.

## Tech-stack Required

1. **Basic Maths & Statistics**
2. **Excel**
3. **SQL & DBMS Knowledge**
4. **Visualisation Tools - Power BI/Tableau**
5. **Python & EDA**
6. **Cloud knowledge**
7. **Basic Machine Learning Algorithms**

1. ***BASIC MATHS & STATISTICS (Week-1)***

Even if you haven't touched maths in the last 5-6 years, don't worry you can begin now. It's not a Rocket-science.

**Topics to cover**

● Basic Maths :- Average, Arithmetic, Weighted average, Cumulative Sum, Percentile vs Percentage
● Statististics :- Mean, Median, Mode, Standard Deviation, Normal Distribution

**Resources**

● Introduction | Mathematics and statistics for data science and machine learning
● Complete Statistics For Data Science In 6 hours By Krish Naik
● Starter Roadmap For Learning Statistics For Data Analyst & Data Science In Hindi - For Hindi Speaking People
● https://www.khanacademy.org/math/statistics-probability/analyzing-categorical-data

**Goal**

Initially just focus on conceptual understanding of each term. You should be able to differentiate between different terms of stats.

## 2. *EXCEL (Week-2)*

### Topics to Cover

- Basic formulas: SUM, DIFF, AVERAGE, MEAN, MEDIAN, CONCATENATE
- Advance formulas: VLOOKUP, INDEX, MATCH, IF, COUNTIF, SUMIF
- REMOVE duplicates and conditional formatting
- Charts, filters, sort and slicers
- Pivot tables and pivot charts
- VBA, Macros, etc

### Resources

- **Complete Excel Tutorial for Data Analysis in 4 Hours (with FREE Files)**
- **Pivot Tables in Excel | Excel Tutorials for Beginners**

### Resources to practice

- https://www.excel-easy.com
- https://exceljet.net
- https://www.excelpracticeonline.com

### Projects in Excel - Optional

- **Full Project in Excel with Interactive Dashboard | Excel Project | Excel Project f…**
- **Excel Project | Data Analyst Portfolio Project | Finance Domain | Start to End | …**

### Goal

Learn how to manipulate data using Excel/Google Sheets, including formulas, pivot tables, and basic data visualization.

### 3. *SQL (Week-3,4,5,6)*

**Topics to Cover**

- Basic Queries:- CREATE, INSERT, UPDATE, ALTER, DELETE, DROP, TRUNCATE.
- Must Know Topics:- SELECT, WHERE, DISTINCT, LIKE, BETWEEN, ORDER BY, LIMIT, GROUP BY, HAVING CLAUSE, IMPORT, DATA TYPES.
- Advance Queries:- Date time function, Window function, Sub query, Case statement, CTE, Query Optimisation
- JOINS:- Self, Inner, Outer, Left, Right

**Resources**

- https://www.w3schools.com/sql/default.asp - For Theory
- SQL for Data Analytics - Learn SQL in 4 Hours
- SQL - Complete Course in 3 Hours | SQL One Shot using MySQL

**Resources to practice**

- https://www.hackerrank.com/domains/sql
- https://leetcode.com/studyplan/top-sql-50/
- https://datalemur.com/questions

**Goal**

- You should be able to visualize tables, joining tables and what you are extracting from tables.
- Tip:- Try making rough tables if stuck with any sql question.
- You should be able to query basic to intermediate problem statements.

4. ***Power BI/ Tableau (Week-7,8,9)***

**Topics to Cover**

● For Power BI - https://medium.com/@AnweshaB/18-important-topics-to-cover-in-power-bi-e225c97c1ba1
● For Tableau - https://medium.com/@shravan1998/important-concepts-in-tableau-you-should-know-114075a2f4ee

**Resources for Power BI**

● Power BI Full Course for FREE with Practical Projects [3 Hours] | Power BI Tutorial …
● Complete Power BI in 10 Hours | PowerBI For Data Analysis (Hindi) #powerbi #data…
● Power BI Tutorial For Beginners 2025 | Power BI Dashboard Project | Power BI Tuto…

**Resources for Tableau**

● Tableau Full Course - in 3 Hours | Become a Data Visualization Rockstar | Beginner …
● Tableau Full Course with Project – Master Data Visualization in 3 Hours (Beginner L…

**Goal**

Create advanced visualizations and interactive dashboards.

### 5. *Python & EDA (Week-10,11,12,13)*

**Topics to Cover**

● Introduction to Python:
- Variables, data types, and basic operations.
- Control structures (if statements, loops).
- Functions and modules.
- Working with Data Structures (List, Tuples, Dictionaries, Sets)
● NumPy:
- Array creation and manipulation.
- Mathematical operations on arrays.
- Indexing and slicing.
● Pandas:
- Series and DataFrame basics.
- Data cleaning and manipulation.
- Grouping and aggregation.
● Matplotlib:
- Creating Basic Plots
- Working with Figures & Axes
- Subplots & Multiple Plots
● Seaborn:
- Basic Plots in Seaborn
- Statistical Data Visualization
- Working with Categorical Data
● Data Cleaning and Preprocessing:
- Handling missing data.
- Removing duplicates.
- Data normalization and scaling.
● EDA (Exploratory Data Analysis)

**Resources**

● Python Tutorial For Beginners in Hindi | Complete Python Course 🔥
● Python Tutorial for Beginners - Full Course (with Notes & Practice Questions)
● Python Pandas Tutorial 2: Dataframe Basics
● numpy tutorial - introduction | numpy array vs python list
● Matplotlib Tutorial 1 - Introduction and Installation
● Python SEABORN Tutorial [HINDI] | Learn Seaborn in 3 Hours - Complete Course
● https://courses.analyticsvidhya.com/courses/pandas-for-data-analysis-in-python
● https://www.w3schools.com/python/default.asp - For Theory
● Learn Exploratory Data Analysis (EDA) from Scratch | EDA in 5 hours | Satyajit Patt…

**Resources to Practice**
● https://pynative.com/python-exercises-with-solutions/
● https://www.hackerrank.com/domains/python

- Leetcode Weekly Contest

**Complete at least 3-4 case study from below playlists**

- https://youtube.com/playlist?list=PL_1pt6K-CLoDMEbYy2PcZuITWEjqMfyoA

**Python with EDA Project (Optional)**

- Python Project for Data Analysis- Exploratory Data Analysis | Data Analyst Project
- New York Airbnb EDA Project with Python | Data Analytics Python Resume Project | …
- Python Project For Data Analysis- Exploratory Data Analysis (EDA) End-to-End Proj…

### 6. *Cloud Knowledge* ***(Week-14,15)***

**Topics to Cover**

● Cloud Storage & Data Warehousing
● Data Processing & ETL (Extract, Transform, Load)
● Business Intelligence & Data Visualization
- Power BI (Azure), Tableau (AWS/GCP), Google Looker Studio
- Connecting BI tools to cloud databases
● Cloud Security & Access Management

**Resources**

● Azure Full Course - Learn Microsoft Azure in 8 Hours | Azure Tutorial For Beginners …
● https://www.youtube.com/live/m6ozQnqit50?si=WrEmpcaNUz1QR7F8
● AWS Tutorial For Beginners | AWS Full Course - Learn AWS In 10 Hours | AWS Trai…

🔗 **Free Courses & Guides:**

● https://cloud.google.com/learn/certification/cloud-digital-leader
● https://aws.amazon.com/training/classroom/aws-cloud-practitioner-essentials/
● Microsoft Learn - Azure Fundamentals

7. ***Machine Learning Basics (Week-16,17)***

**Topics to Cover**

- Supervised Learning (Regression & Classification)
- Unsupervised Learning (Clustering & Dimensionality Reduction)
- Feature Engineering & Data Preprocessing
- Model Evaluation & Performance Metrics
- Time Series Analysis (For Forecasting)

**Resources**

- Complete Machine Learning In 6 Hours| Krish Naik
- Python Machine Learning Tutorial (Data Science)
- Complete ML Machine Learning in One Shot (5 Hours) | Semester Exam | In Hindi

**Goal**

As a Data Analyst, focus on practical ML techniques like Regression, Classification, Clustering, and Time Series Analysis, along with strong feature engineering and model evaluation skills.

**Must-Do 20 Data Analyst Interview Questions (2025 Edition)**

If you're preparing for a Data Analyst interview, these are the most important questions across SQL, Excel, Python, Statistics, and Business Intelligence.

**SQL Interview Questions**

1. What is the difference between INNER JOIN, LEFT JOIN, RIGHT JOIN, and FULL JOIN?
2. How do you find duplicate records in a table? (Write an SQL query)
3. How do you rank rows without using the ROW_NUMBER() function?
4. What is the difference between WHERE and HAVING clauses in SQL?
5. Write an SQL query to calculate the running total of a sales column.

**Python for Data Analysis**

6. How do you handle missing values in a dataset using Python?
7. What is the difference between apply(), map(), and lambda functions in Pandas?
8. How do you merge two datasets in Pandas? (Explain different types of joins)
9. Explain the difference between a list, tuple, and dictionary in Python.
10. How do you group data in Pandas and perform aggregate functions?

**Statistics & Probability**

11. What is the difference between correlation and causation?
12. What is p-value, and how is it used in hypothesis testing?
13. What is standard deviation, and why is it important in data analysis?
14. Explain the Central Limit Theorem (CLT) and its significance.
15. What is A/B testing, and how do you determine if a test is successful?

**Excel & Data Visualization**

16. What is a Pivot Table, and how is it used in data analysis?
17. How do you create a dynamic dashboard in Excel using Power Query and Pivot Tables?
18. What are different types of charts in data visualization, and when should you use them?
19. How do you remove duplicates and clean data in Excel?

**Business Intelligence & Problem-Solving**

20. You are given a dataset with missing values, duplicates, and inconsistent formats. How would you clean and prepare it for analysis?

**Next Steps**

- Practice SQL Queries on real datasets (LeetCode SQL)
- Work on Data Projects (e.g., customer churn, sales forecasting)
- Revise Statistics Concepts (Khan Academy, StatQuest YouTube)
- Practice with Real-World Datasets (Kaggle Datasets)

**Important Note**

**Once you are done with 60-70% of your syllabus, prepare your resume & start applying along-side.**

***ALL THE VERY BEST FOR YOUR PREPARATION!!!***

Follow on Youtube - Noodle Brain

Follow on Instagram - Noodle Brain - Instagram

Follow on Linkedin - https://www.linkedin.com/in/mansi-r-6a4115169/

Topmate - Book 1:1 session with me