

Data Analysis

Module 1: Data Basics + Excel Fundamentals

Week 1

- Day 1 – What is Data Analytics
- Day 2 – Types of Data
- Day 3 – Data Analyst roles
- Day 4 – Install Microsoft Excel and explore interface
- Day 5 – Rows, columns, formatting
- Day 6 – Basic formulas
- Day 7 – Practice dataset

Week 2

- Day 8 – SUM function
- Day 9 – AVERAGE function
- Day 10 – COUNT function
- Day 11 – MIN / MAX
- Day 12 – Sorting data
- Day 13 – Filtering data
- Day 14 – Practice project

Week 3

- Day 15 – Conditional formatting
- Day 16 – Data cleaning basics
- Day 17 – Remove duplicates
- Day 18 – Text functions
- Day 19 – Date functions
- Day 20 – Excel shortcuts
- Day 21 – Practice dataset

Week 4

- Day 22 – Charts
- Day 23 – Bar chart
- Day 24 – Line chart
- Day 25 – Pie chart
- Day 26 – Data visualization basics
- Day 27 – Mini dashboard
- Day 28 – Excel mini project

Module 2: Advanced Excel

Week 1

- IF function
- AND / OR functions
- Nested IF

Week 2

- VLOOKUP
- HLOOKUP
- XLOOKUP
- INDEX + MATCH

Week 3

- Pivot tables
- Pivot charts
- Data analysis using pivot tables

Week 4

- Dashboard building
- KPI analysis
- Excel automation basics

Module 3: Statistics

Week 1

- Mean, Median, Mode
- Standard deviation

Week 2

- Probability basics
- Sampling

Week 3

- Correlation
- Distribution

Week 4

- Hypothesis testing basics
- Business statistics examples

Module 4: SQL Basics

Week 1

- Database concepts
- Tables & rows
- SELECT statements

Week 2

- WHERE conditions
- ORDER BY
- LIMIT

Week 3

- GROUP BY
- HAVING
- Aggregate functions

Week 4

- JOIN basics
- INNER JOIN
- LEFT JOIN

Module 5: Advanced SQL

Week 1

- Subqueries
- Tables & rows
- SELECT statements

Week 2

- Window functions

Week 3

- ROW_NUMBER
- RANK

Week 4

- Query optimization
- SQL interview questions

Module 6 : Python Basics

Week 1

- Install Jupyter Notebook
- Variables
- Data types

Week 2

- Lists
- Dictionaries
- Loops

Week 3

- Functions
- File handling

Week 4

- Python exercises

Module 7 : Data Analysis with Python

Libraries:

- NumPy
- Pandas

Week 1

- NumPy arrays

Week 2

- Pandas DataFrame

Week 3

- Data cleaning
- Handling missing values

Week 4

- Data aggregation
- Data transformation

Module 8 : Data Visualization

Libraries:

- Matplotlib
- Seaborn

Week 1

- Line chart
- Bar chart

Week 2

- Scatter plot
- Histogram

Week 3

- Visualization storytelling

Week 4

- Power BI
- Tableau

Module 9 : Exploratory Data Analysis

Week 1

- Data exploration

Week 2

- Data cleaning pipelines

Week 3

- Feature analysis

Week 4

- Real dataset practice

Module 10 : Career + Portfolio

Week 1

- Build portfolio

Week 2

- Upload projects on GitHub

Week 3

- Resume building

Week 4

- Interview preparation

10 Real Portfolio Projects

- **Sales Data Analysis** Analyze sales trends and profit.
- Tools
- Excel
- Python
- Power BI

E-commerce Customer Analysis

- Customer purchase behavior.

Marketing Campaign Performance

- Analyze ROI of marketing campaigns.

Netflix Data Analysis

Find:

- Most popular genres
- Country distribution

COVID-19 Data Analysis

- Dataset from World Health Organization

Stock Market Analysis

- Analyze stock prices of Tesla or Apple.

Customer Churn Analysis

- Predict why customers leave a company.

HR Employee Attrition Analysis

- Find why employees resign.

YouTube Data Analysis

Dataset from YouTube.

Find:

- Trending videos
- Most viewed categories

Top 50 SQL Interview Questions

Basic SQL (1–15)

1. What is SQL?
2. Difference between SQL and NoSQL.
3. What is a database?
4. What is a table?
5. What is a primary key?
6. What is a foreign key?
7. What is the SELECT statement?
8. What is the WHERE clause?
9. Difference between WHERE and HAVING.
10. What is ORDER BY?
11. What is GROUP BY?
12. What are aggregate functions?
13. What is COUNT()?
14. What is SUM()?
15. What is AVG()?

Intermediate SQL (16–35)

16. What is a JOIN?
17. Types of joins.
18. Difference between INNER JOIN and LEFT JOIN.
19. What is RIGHT JOIN?
20. What is FULL JOIN?
21. What is a subquery?
22. What is a correlated subquery?
23. What is UNION?
24. UNION vs UNION ALL.
25. What is DISTINCT?
26. What is LIMIT?
27. What is OFFSET?
28. What are constraints?
29. What is NOT NULL?
30. What is UNIQUE?
31. What is DEFAULT?
32. What is CASE statement?
33. What is COALESCE?
34. What is NULL?
35. Difference between DELETE and TRUNCATE.

Advanced SQL (36–50)

36. What is a window function?
37. What is ROW_NUMBER()?
38. What is RANK()?
39. Difference between RANK and DENSE_RANK.
40. What is a Common Table Expression (CTE)?
41. What is a stored procedure?
42. What is a trigger?
43. What is indexing?
44. Clustered vs Non-clustered index.
45. What is query optimization?
46. What is normalization?
47. What are normal forms?
48. What is denormalization?
49. What is transaction?
50. What is ACID property?

Top 50 Python Interview Questions

Basic Python (1–15)

1. What is Python?
2. Python features.
3. What are variables?
4. Data types in Python.
5. What is a list?
6. What is a tuple?
7. What is a dictionary?
8. What is a set?
9. Difference between list and tuple.
10. What is a loop?
11. For loop vs while loop.
12. What is indentation in Python?
13. What is a function?
14. What are arguments?
15. What is return statement?

Intermediate Python (16–35)

16. What is list comprehension?
17. What is lambda function?
18. What is map()?
19. What is filter()?
20. What is reduce()?
21. What is a module?
22. What is a package?
23. What is exception?

24. What is try-except?
25. What is file handling?
26. What is JSON in Python?
27. What is API?
28. What is virtual environment?
29. What is pip?
30. What is NumPy?
31. What is Pandas?
32. What is DataFrame?
33. What is Series in Pandas?
34. How to handle missing values?
35. How to read CSV files?

Common libraries:

- NumPy
- Pandas

Advanced Python (36–50)

36. What is object-oriented programming?
37. What is a class?
38. What is an object?
39. What is inheritance?
40. What is encapsulation?
41. What is polymorphism?
42. What is multithreading?
43. What is multiprocessing?
44. What is decorator?
45. What is generator?
46. What is iterator?
47. Difference between shallow and deep copy.
48. What is memory management in Python?
49. What is GIL (Global Interpreter Lock)?
50. What are Python frameworks?

50 Datasets for Data Analytics Practice

Business / Sales

1. Superstore sales dataset
2. E-commerce sales
3. Retail transactions
4. Customer purchase dataset
5. Online store sales
6. Walmart sales data
7. Amazon product data
8. Marketing campaign dataset
9. Customer lifetime value dataset
10. Retail inventory dataset

Entertainment

- 11. Movies dataset
- 12. IMDB ratings dataset
- 13. Netflix titles dataset
- 14. Spotify songs dataset
- 15. YouTube trending videos dataset

Gaming

- 46. Video game sales dataset
- 47. Steam games dataset
- 48. Esports dataset
- 49. Game ratings dataset
- 50. Gaming revenue dataset

Finance

- 16. Stock market dataset
- 17. Cryptocurrency dataset
- 18. Credit card transactions
- 19. Bank marketing dataset
- 20. Loan prediction dataset

HR Analytics

- 21. Employee attrition dataset
- 22. HR analytics dataset
- 23. Employee salary dataset
- 24. Job satisfaction dataset
- 25. Workforce diversity dataset

Healthcare

- 26. COVID-19 dataset
- 27. Diabetes dataset
- 28. Hospital patient dataset
- 29. Heart disease dataset
- 30. Medical insurance dataset

Transportation

- 31. Uber trip dataset
- 32. Taxi trips dataset
- 33. Airline delays dataset
- 34. Flight dataset
- 35. Traffic accidents dataset

Miscellaneous

- 36. Global temperature dataset
- 37. Population dataset
- 38. Education dataset
- 39. Crime dataset
- 40. Housing prices dataset

Technology

- 41. Mobile phone usage dataset
- 42. App store dataset
- 43. Social media dataset
- 44. Website traffic dataset
- 45. Online reviews dataset