■ Data Analytics SQL Roadmap (Basic → Advanced)

1. SQL Foundations (Basics)

- Introduction to databases & SQL
- SQL types: DDL, DML, DCL, TCL
- Creating and managing databases
- Creating tables (constraints: PRIMARY KEY, FOREIGN KEY, UNIQUE, NOT NULL, DEFAULT, CHECK)
- Data types: numeric, text, date/time, boolean, JSON (vendor-specific)
- Basic queries: SELECT, WHERE, ORDER BY, LIMIT/TOP
- Operators: =, !=, >, <, >=, <=, BETWEEN, IN, LIKE, IS NULL
- INSERT, UPDATE, DELETE

2. Data Retrieval & Functions

- String functions: UPPER, LOWER, TRIM, CONCAT, SUBSTRING, REPLACE
- Date/Time functions: NOW, GETDATE, DATEADD, DATEDIFF, EXTRACT, YEAR, MONTH, DAY
- Numeric functions: ROUND, CEIL, FLOOR, ABS, POWER, MOD
- Handling NULLs: COALESCE, NULLIF, IFNULL

3. Aggregation & Grouping

- Aggregate functions: COUNT, SUM, AVG, MIN, MAX
- GROUP BY basics
- Multiple column grouping
- Filtering groups with HAVING
- Order of execution: WHERE → GROUP BY → HAVING → SELECT → ORDER BY

4. Joins (Combining Data)

- INNER JOIN, LEFT JOIN, RIGHT JOIN, FULL OUTER JOIN, CROSS JOIN
- Self Join (hierarchical/parent-child data)
- Join with multiple tables

5. Subqueries & Set Operations

- Subqueries in WHERE, FROM, SELECT
- Correlated subqueries
- EXISTS vs IN
- Set operations: UNION vs UNION ALL, INTERSECT, EXCEPT/MINUS

6. Conditional Logic

- CASE WHEN (simple & searched case expressions)
- Nested CASE statements
- IF, IIF (vendor-specific)
- NULL handling: COALESCE, ISNULL, NULLIF

7. Intermediate SQL

- Common Table Expressions (CTEs)
- Recursive CTEs (basic hierarchy)
- Temporary tables, Views, Materialized Views
- Pivot and Unpivot queries
- Ranking functions: ROW_NUMBER, RANK, DENSE_RANK, NTILE

8. Window Functions (Analytics Core)

- OVER() clause basics
- Aggregate window functions: SUM, AVG, COUNT, MIN, MAX
- Ranking functions: ROW_NUMBER, RANK, DENSE_RANK
- Time-series: LAG, LEAD
- Running totals & moving averages
- Frame clauses (ROWS/RANGE for sliding windows)

9. Advanced Joins & Analytics Queries

- Multi-table joins with complex conditions
- Self joins with hierarchical queries
- Nested joins with subqueries + CTEs
- Multi-level aggregations (rollups, cubes, grouping sets)
- A/B testing queries (control vs treatment groups)

10. Performance Optimization

- Indexing (clustered vs non-clustered, composite indexes)
- Query execution plans (EXPLAIN, EXPLAIN ANALYZE)
- Optimizing joins vs subqueries
- Partitioning & sharding large datasets
- Materialized views for optimization

11. Data Analytics SQL Use Cases

- Cohort analysis (signup date grouping)
- Retention & churn metrics
- Customer segmentation (RFM analysis)
- Funnel analysis (conversion stages)
- Trend analysis (monthly/quarterly/yearly growth)
- Event analysis (logs, clicks, transactions)

12. Advanced Topics

- Recursive CTEs (org charts, tree structures)
- Advanced Frame clauses
- Pivot/Unpivot for dynamic reporting
- JSON/XML data handling
- Stored procedures & user-defined functions
- Triggers for automation
- Data warehouse concepts: OLTP vs OLAP, Star schema, Snowflake schema

13. Big Data & Cloud SQL

- BigQuery (Google), Redshift (AWS), Snowflake, Azure SQL
- Partitioned & clustered tables
- Distributed SQL engines: Hive, Presto, SparkSQL
- Querying billions of rows efficiently
- Integration with BI tools (Power BI, Tableau, Looker)