SQL for Data Analytics: Comprehensive Guide on Data Types, Keys, and Terms

# 1. Introduction to SQL for Data Analytics

Structured Query Language (SQL) is the backbone of data manipulation and retrieval in relational database systems. For any data analyst, mastering SQL is fundamental to extracting insights from data. This guide explores the foundational components of SQL: data types, keys, and common terminologies.

# 2. SQL Data Types

# 2.1 Numeric Data Types

INT: Integer numbers, no decimals. E.g., Age, Count  
SMALLINT: Smaller integer range. E.g., Status codes  
BIGINT: Larger range. E.g., Population  
DECIMAL(p,s)/NUMERIC(p,s): Fixed-point. E.g., Currency  
FLOAT: Approximate float. E.g., Scientific values  
REAL/DOUBLE PRECISION: Double-precision float. E.g., Analytics

# 2.2 Character/String Data Types

CHAR(n): Fixed-length. E.g., Gender  
VARCHAR(n): Variable-length. E.g., Name, Email  
TEXT: Unlimited text. E.g., Description

# 2.3 Date and Time Data Types

DATE: Date only. E.g., Birthdate  
TIME: Time only. E.g., Clock-in  
DATETIME/TIMESTAMP: Date and time. E.g., Log  
YEAR: Year only. E.g., Graduation year

# 2.4 Binary Data Types

BINARY: Fixed-length binary. E.g., Encryption key  
VARBINARY: Variable-length. E.g., Multimedia

# 2.5 Boolean Data Type

BOOLEAN/BOOL: TRUE or FALSE. E.g., IsActive

# 3. SQL Keys

# 3.1 Primary Key

Uniquely identifies each record. Must be NOT NULL and unique. Used for IDs.

# 3.2 Foreign Key

Links tables. Refers to a primary key in another table. Used in relationships.

# 3.3 Unique Key

Ensures values are unique. Allows NULL once. E.g., Email.

# 3.4 Composite Key

Multiple columns form a key. E.g., For many-to-many join tables.

# 3.5 Candidate Key

Any column(s) that can be a primary key.

# 3.6 Alternate Key

Candidate key not chosen as primary key.

# 3.7 Surrogate Key

Auto-incremented key used instead of natural key.

# 4. Essential SQL Terms and Concepts

# 4.1 Table

Collection of related data.

# 4.2 Column

Vertical field defining an attribute.

# 4.3 Row (Record)

Single data entry in a table.

# 4.4 Schema

Structure of tables, views, etc.

# 4.5 Index

Speeds up searches.

# 4.6 Constraint

Rules for data integrity: PK, FK, NOT NULL, UNIQUE, CHECK.

# 4.7 View

Virtual table from a query.

# 4.8 Stored Procedure

Precompiled task or logic.

# 4.9 Trigger

Executes on table/view events.

# 4.10 Join

INNER JOIN: Matched rows only  
LEFT JOIN: All left, matched right  
RIGHT JOIN: All right, matched left  
FULL OUTER JOIN: All matched/unmatched

# 4.11 Normalization

Reduces redundancy.

# 4.12 Denormalization

Improves speed by redundancy.

# 4.13 NULL

Represents missing data.

# 4.14 Alias

Temporary name for tables/columns.

# 4.15 Subquery

Query inside another query.

# 5. Best Practices

- Use precise data types  
- Apply constraints  
- Index wisely  
- Use JOINs for performance  
- Regular backups

# 6. Conclusion

Understanding SQL data types, keys, and terms forms the foundation of data analytics. Mastery enables efficient data structuring, querying, and insight generation.