**🗓️ Day-wise SQL Practical Plan for Data Engineering (10 Days)**

**✅ Goal:**

* Cover real-world SQL use cases from basic to advanced.
* Focus more on what’s used in **ETL**, **analytics**, **data warehousing**, and **interviews**.

**🔧 Tools to Use Daily:**

* **Platform**: LeetCode SQL, StrataScratch
* **RDBMS**: PostgreSQL / MySQL (locally or on DB Fiddle)
* **Dataset**: Start with [Chinook](https://github.com/lerocha/chinook-database), [AdventureWorks], or Kaggle datasets.

**📅 Day 1 – Basic SELECTs & Filtering**

**Time Investment**: 2 hrs  
**Topics**:

* SELECT, FROM, WHERE, ORDER BY
* BETWEEN, IN, LIKE, DISTINCT
* Filtering dates and strings

**Practice**:

* Top 10 most sold products
* Customers who joined in 2024
* Products containing “Pro”

**📅 Day 2 – Aggregation & Grouping**

**Time Investment**: 2 hrs  
**Topics**:

* GROUP BY, HAVING
* Aggregates: COUNT, SUM, AVG, MIN, MAX

**Practice**:

* Total revenue by month
* Average order value per customer
* Count of products sold per category

**📅 Day 3 – JOINs Deep Dive (Critical for Data Engineering)**

**Time Investment**: 3–4 hrs  
🚨 **Spend more time here**  
**Topics**:

* INNER, LEFT, RIGHT, FULL JOIN
* Self joins, multi-table joins

**Practice**:

* Orders with customer and product names
* Customers who never ordered
* Most selling category by region

**📅 Day 4 – Subqueries & Nested Queries**

**Time Investment**: 2 hrs  
**Topics**:

* Scalar, correlated subqueries
* IN, EXISTS, NOT EXISTS

**Practice**:

* Customers who placed more orders than average
* Products with price above average of category

**📅 Day 5 – Window Functions (Advanced & Very Important)**

**Time Investment**: 3 hrs  
🧠 **Critical for analytics, ADF, transformations**  
**Topics**:

* ROW\_NUMBER(), RANK(), DENSE\_RANK()
* LEAD(), LAG(), FIRST\_VALUE(), NTILE()
* Partition by and order by in window

**Practice**:

* Get 2nd highest order per customer
* Previous and next purchase date for user
* Running totals by month

**📅 Day 6 – CTEs and Temporary Views**

**Time Investment**: 2 hrs  
**Topics**:

* WITH clause (Common Table Expressions)
* Recursive CTE (optional)

**Practice**:

* Top 3 selling products per month (using CTE)
* Chain of referrals (recursive)

**📅 Day 7 – Data Cleaning & Transformation**

**Time Investment**: 3 hrs  
🧼 Real-world data cleaning for pipelines  
**Topics**:

* COALESCE, NULLIF, CASE WHEN, CAST()
* Date formatting, string manipulation

**Practice**:

* Replace nulls with “Unknown”
* Format phone numbers
* Classify users as High/Medium/Low spender

**📅 Day 8 – Data Modeling & Normalization**

**Time Investment**: 2 hrs  
**Topics**:

* Write DDLs: CREATE, ALTER, DROP
* Design tables from raw data (3NF)
* Create fact and dimension tables

**Practice**:

* Design Sales Fact table and Product Dimension
* Normalize unstructured table

**📅 Day 9 – Advanced SQL Use Cases**

**Time Investment**: 3 hrs  
**Topics**:

* Pivot/Unpivot (if DB supports)
* JSON handling, array data
* SQL in Data Warehouse (BigQuery/Snowflake syntax differences)

**Practice**:

* Create pivoted sales table (months as columns)
* Parse JSON from a column

**📅 Day 10 – Project + Mock Interview**

**Time Investment**: 4 hrs  
**Project**: Mini ETL-like SQL project

📊 Example Project:

Load sales, transform to monthly aggregates, clean nulls, join to product and customer, create final analytics view

🎯 Bonus:

* Try a SQL interview test from StrataScratch
* Mock yourself with 5 SQL questions + 1 real scenario

**⏳ Time Focus Areas (Very Important for Data Engineers)**

| **Area** | **Focus Time** | **Importance Level** |
| --- | --- | --- |
| Joins | 🔥🔥🔥 | ⭐⭐⭐⭐⭐ |
| Window Functions | 🔥🔥🔥 | ⭐⭐⭐⭐⭐ |
| CTEs & Subqueries | 🔥🔥 | ⭐⭐⭐⭐ |
| Aggregations & Filters | 🔥🔥 | ⭐⭐⭐⭐ |
| Data Cleaning | 🔥🔥 | ⭐⭐⭐ |
| Table Design (DDL) | 🔥 | ⭐⭐⭐ |

**✅ What's Next?**

Once you finish this 10-day plan, we’ll move into:

* Writing **ETL logic** using SQL
* **Integrating SQL into pipelines** (ADF, Python, Airflow)
* **AI-powered SQL queries** (Copilot, ChatGPT coding)