Overall Steps I've Followed

1. **Read Task Instructions**
    - Understood that the goal was to write SQL queries to analyze an e-commerce dataset, demonstrate joins, subqueries, indexes, views, and NULL handling, then submit via GitHub.
2. **Setup**
    - Chose **SQLite** with **DB Browser** as the SQL tool.
    - Created a **sample dataset** (`customers` & `orders` tables) with realistic values.
3. **Dataset Preparation**
    - Exported sample dataset to `sample_dataset.sql` for easy import into SQLite.
    - Imported this SQL file into a new SQLite database in DB Browser.
4. **Writing Queries**
    - Wrote **SQLite-friendly queries** covering:
        - WHERE vs HAVING
        - INNER JOIN, LEFT JOIN, RIGHT JOIN alternative
        - Average Revenue Per User (ARPU)
        - Subqueries
        - Index optimization with multiple examples (`idx_order_date`, `idx_customer_date`)
        - Creating views (`customer_spending`)
        - Handling NULL values with `COALESCE`
5. **Optimization Section**
    - Avoided re-creating an existing index (`idx_customer_id`) to prevent errors.

- Added new indexes for performance tuning.

6. **Testing & Capturing Outputs**
   - Ran queries in DB Browser's "Execute SQL" tab.
   - Took screenshots of each query output for submission.
7. **Documentation**
   - Created a clear `README.md` explaining:
     - Objectives
     - Queries implemented
     - Dataset used
     - How to run the project
   - Used a structured naming scheme for screenshots (e.g., `q1-where-vs-having.png`).
8. **Packaging for GitHub**
   - Prepared folder structure:
     - `task4_queries_sqlite.sql`
     - `sample_dataset.sql`
     - `README.md`
     - `screenshots-task4/` (screenshots inside)
   - Chose repo name: **`task4-sql-data-analysis`**
   - Wrote short GitHub description.
9. **Next Step**
   - Pushed the prepared folder to GitHub.
   - Submitted GitHub repo link in Google Form before deadline.

Summary

You have successfully prepared a complete **SQL Data Analysis project** in SQLite for Task 4, including a custom dataset, fully working queries, multiple join types, subqueries, indexing for optimization, views, and NULL handling. You've documented everything in a `README.md`, organized outputs into a screenshot folder, and are now ready to upload the complete package to a GitHub repo for submission.