Lab Report: SQL Query from Table Names

Objective: The aim of this exercise was to test whether GPT-3.5-Turbo can determine the relevant tables for constructing SQL queries based on table names and their definitions. This involved feeding the model with structured prompts and evaluating its responses.

Findings:

1. **Effective Scenarios:**
   o When table names and definitions were detailed and directly aligned with the user's query, the model performed well, accurately identifying the necessary tables.
   o Clear and concise prompts, formatted with structured instructions, yielded accurate and consistent outputs.
   o In a challenging example where employees needed to be matched with their highest educational degree obtained in the same year they were hired, the model successfully identified only the relevant tables (`employees`, `studies`, and `departments`), demonstrating its ability to handle more complex and ambiguous queries.
   o For a more ambiguous query ("Who has achieved the highest educational level and is associated with the highest total salary in their department?"), the model inferred the correct set of tables (`employees`, `departments`, `studies`, `salaries`), showing that the optimized prompt improved its ability to handle unclear questions.
2. **Challenging Scenarios:**
   o **Hallucinations:** In cases where table definitions were ambiguous or overlapped conceptually, GPT occasionally inferred irrelevant tables or omitted necessary ones.
   o **Over-inclusion of Tables:** Even when the user's question was specific and did not require certain tables (e.g., `salaries` in creative variations), the model included them in the response.
   o **Omission of Tables:** For questions that involved filtering by department (e.g., "Which employees with a PhD earned more than $90,000 last year in the 'Research' department?"), the model failed to include the `departments` table. This suggests a lack of understanding of inter-table dependencies.
   o **Sensitivity to Prompt Wording:** Minor changes in the question phrasing led to inconsistencies, with the model sometimes overcompensating by adding unnecessary tables.

Key Learnings:

- **Prompt Design Matters:** A well-structured prompt, including explicit formatting, clear task delineation, and examples of relationships, is critical for obtaining accurate responses.
- **Contextual Ambiguity:** The model relies heavily on explicit definitions; vague or incomplete descriptions significantly impact accuracy.
- **Iterative Refinement:** Experimenting with variations of the same query can help identify optimal approaches and refine results.
- **Dependency Clarity:** Explicitly stating dependencies between tables in the prompt can help guide the model to better decisions.
- **Handling Ambiguity:** Optimized prompts with detailed instructions improved the model's ability to infer relationships, even when the query was unclear or incomplete.
- **Limitations of AI:** While GPT is powerful in text understanding, it lacks inherent relational database knowledge, leading to occasional logical gaps.