

Assignment 2A LLM Transcript

Muhammad Suffyan Khan

2026-02-08

LLM Transcript and Citation

Tool Used: ChatGPT (OpenAI, GPT-5.2)

Purpose: Debugging PostgreSQL connection issues and assisting with SQL-to-R workflow design

Course: DATA 607 – Assignment 2

Date Accessed: February 8 2026

Citation:

OpenAI. (2026). *ChatGPT (GPT-5)* [Large language model]. <https://chat.openai.com>.

I used ChatGPT to help debug PostgreSQL connection issues and to support the workflow of loading relational data from SQL into R for analysis. I worked iteratively by writing SQL scripts in pgAdmin, connecting to PostgreSQL from R using DBI and RPostgres, and validating outputs locally. The most useful assistance involved (1) resolving database connection-context issues, and (2) demonstrating a strategy for handling missing ratings using a complete user-movie grid.

[1] Conversation with ChatGPT (PostgreSQL connection context)

I'm able to connect to PostgreSQL from R, but when I query the users table I get "relation does not exist." What's happening?

Assistant (ChatGPT): This usually happens when R is connected to the default maintenance database (`postgres`) instead of the database containing your tables. You should connect using the database where the tables were created.

Suggested code:

```
con <- dbConnect(  
  RPostgres::Postgres(),  
  host = "localhost",  
  port = 5432,  
  dbname = "Movie _Ratings",  
  user = "postgres",  
  password = Sys.getenv("PGPASSWORD")  
)
```

[2] Conversation with ChatGPT (Handling missing ratings)

How do I demonstrate a missing-ratings strategy?

Assistant (ChatGPT): Create a full user-movie grid and then left-join ratings so that missing ratings appear as NA. This explicitly represents unseen items.

Suggested code:

```
ratings_complete <- users_df %>%  
  crossing(movies_df) %>%  
  left_join(ratings_df, by = c("user_id", "movie_id"))
```

[3] Conversation with ChatGPT (Joining tables in R)

Should I join the tables in SQL or in R?

Assistant (ChatGPT): Either approach is acceptable. Since the analysis is performed in R, joining tables using dplyr is a common data science workflow.

Suggested code:

```
ratings_joined <- ratings_df %>%  
  left_join(users_df, by = "user_id") %>%  
  left_join(movies_df, by = "movie_id")
```

ChatGPT is AI and can make mistakes.