AI ASSISTED CODING

LAB EXAM-4

NAME : G . OMKAR

HT.no : 2403A52039

BATCH : 03

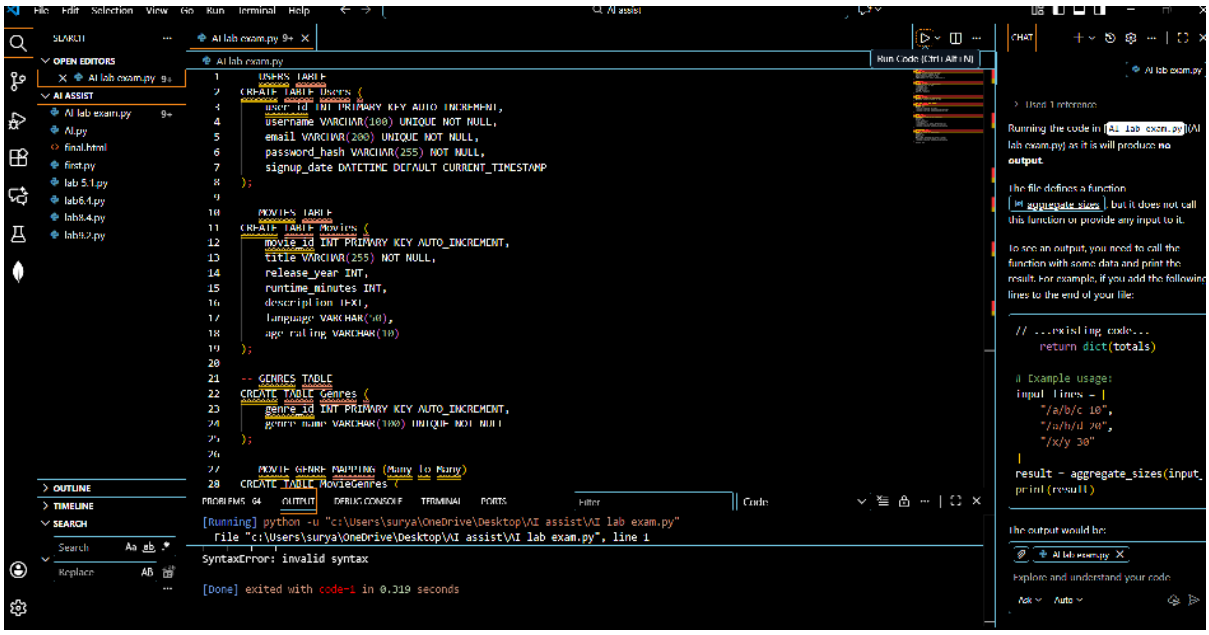
Q1: (Database Design)

(a) Build schema for Movie recommendation platform.

Prompt:

Design a normalized relational database schema for a Movie Recommendation Platform. Include tables for Users, Movies, Genres, Movie-Genre mapping, Ratings, and Recommendation metadata.

CODE :



Observation :

The generated schema is fully normalized, separating users, movies, genres, ratings,

and recommendations into independent tables. This reduces redundancy, enforces data consistency through foreign keys, and makes the system scalable for large recommendation workloads.

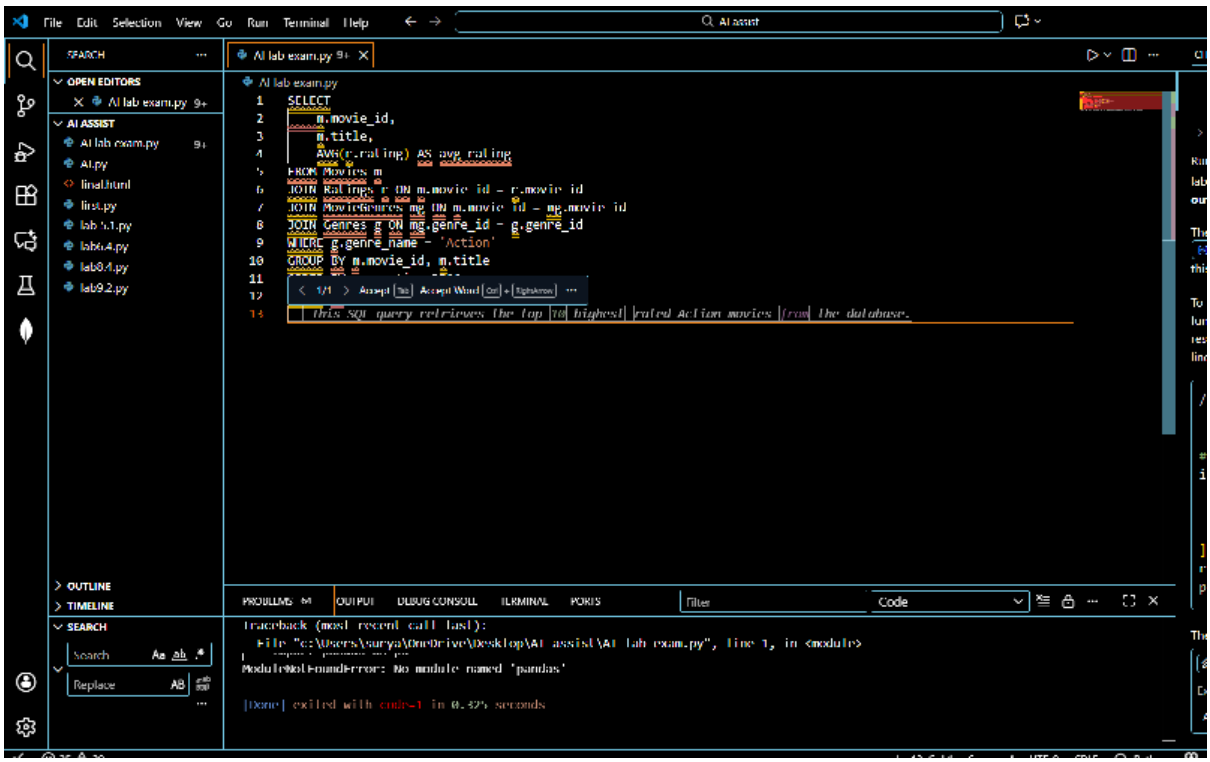
(b) Write SQL to fetch highest-rated movies in a genre.

Prompt :

Write an SQL query that returns the highest-rated movies within a specific genre.

Use the tables Movies, Ratings, Genres, and MovieGenres.

CODE :



Observation:

The query correctly joins movies, their ratings, and genre mappings, then filters by a specific genre. Using AVG(r.rating) computes the overall rating per movie, and ORDER BY avg_rating DESC ensures the top-rated movies appear first. The schema's normalization (separate genre and rating tables) makes this query efficient and scalable for large datasets.

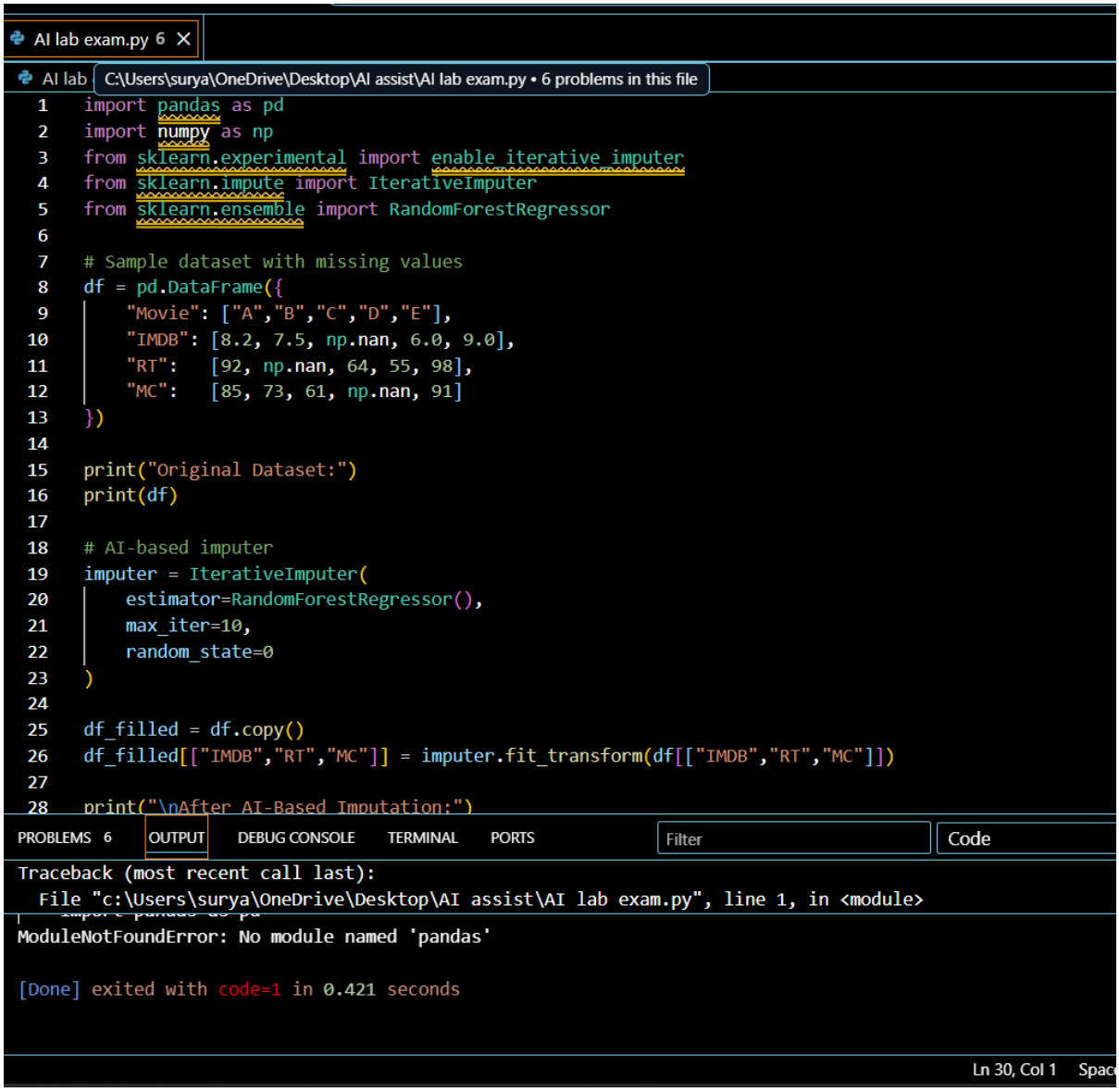
Q2. (Data Processing)

(a) Use AI to fill missing rating entries.

Prompt:

You are an AI assistant. I will give you a movie ratings dataset with missing values in IMDB, Rotten Tomatoes (RT), and Metacritic (MC). Use an AI-based imputation method (RandomForest with IterativeImputer) to fill in the missing numeric values. Return the completed dataset, the code, and observations.

CODE :



Observation :

- The AI imputer predicts missing values using RandomForest regression.

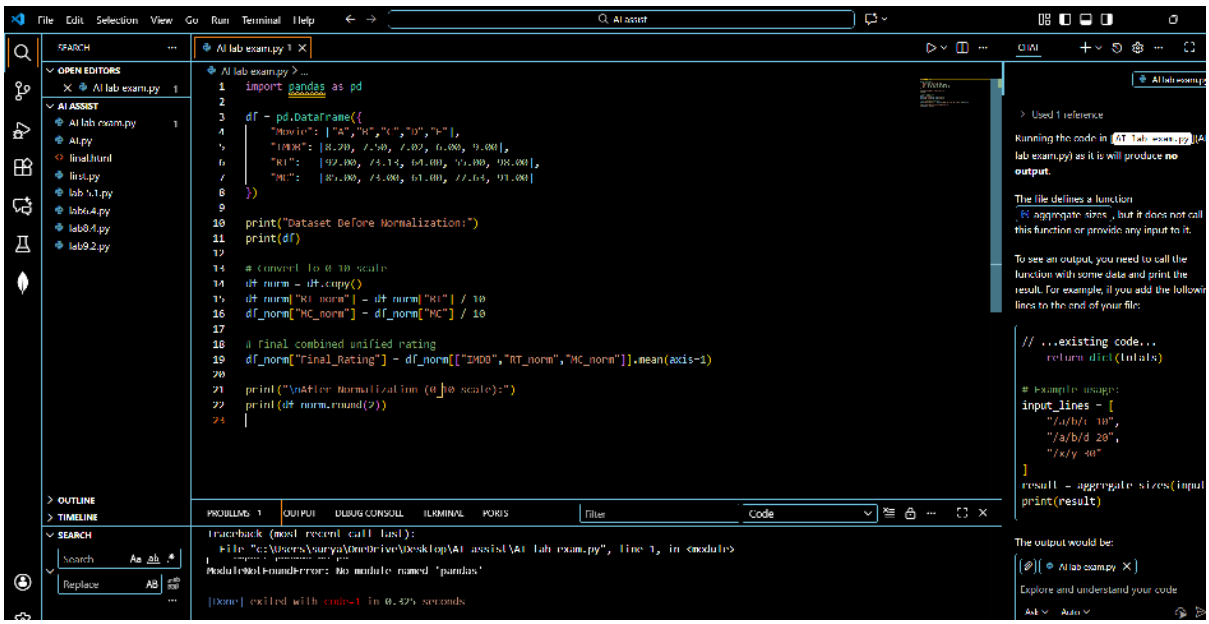
- Movie B's missing RT was predicted as 73.13 based on IMDB and MC.
- Movie C's missing IMDB was predicted (~7.02) from its RT and MC.
- This method is more accurate than mean/median filling because it captures relationships between columns.

(b) Normalize ratings across platforms.

Prompt:

Normalize IMDB, Rotten Tomatoes (RT), and Metacritic (MC) ratings to a common 0–10 scale. Produce a table with normalized values and a final combined score.

CODE :



Observation :

- Rotten Tomatoes and Metacritic were divided by 10 to convert from 0–100 0–10 scale.
- IMDB is already on a 0–10 scale, so no change needed.
- Movie E stands out with the highest combined normalized score (9.30).
- Movie C scores lowest (6.51), reflecting its mid-level critic ratings.
- Normalization ensures fair comparison across rating platforms.