### Data Structure for Movie Recommendation System

#### CSV and JSON Formats

Each file iteration increases in complexity and messiness. Below is the explanation for the data structure:

1. \*\*Iteration 1 (Simple and Clean):\*\*

- Columns/Keys: Title, Genre, Year, Rating, Actors

- All fields are complete, clean, and consistent.

2. \*\*Iteration 2 (More Complex):\*\*

- Adds a new column/key: Runtime (movie runtime in minutes).

- Includes additional genres and actors with varied counts.

3. \*\*Iteration 3 (Messy Data):\*\*

- Includes missing values (e.g., None in Python).

- Year and Rating columns contain mixed types (e.g., strings and numbers).

- Inconsistent formats (e.g., "90 minutes" for runtime).

4. \*\*Iteration 4 (Extremely Messy):\*\*

- Contains typos (e.g., "SciFi" instead of "Sci-Fi").

- Duplicate entries (e.g., "Inception" twice).

- Irrelevant data (e.g., BoxOffice column).

- Severe inconsistencies in format and missing fields.

#### Purpose:

These files simulate real-world data challenges, allowing users to experience data cleaning, handling inconsistencies, and improving their prompts iteratively.