Project 2: Multidimensional Data Visualization Design

In this assignment, you will design a visualization for a small data set and provide a rigorous rationale for your design choices. You are free to use any graphics or charting tool

Data Set: Movies Data Mash-Up

- Movies sales information from Box Office Mojo website.
- Movies rating from IMDB.
- Any other online movie data you can find.

Assignment

Your task is to design visualization that you believe effectively analysis and communicates the data and provide a short write-up (no more than 4 paragraphs) describing your design. While you must use the data set given, note that you are free to transform the data as you see fit. Some sample Data cleanup and star schema-building tutorial with PowerBI are provided in the links below, but you are free to use any ETL or Visualization tool to clean and build the data model. Those tutorials are just examples and you are not required to follow.

https://radacad.com/build-your-first-star-schema-model-in-action-power-bi-modeling-basics

https://radacad.com/get-started-with-power-query-movies-data-mash-up

Develop some interesting questions about the selected data set – put yourself in the shoes of a data analyst, and think about all the different kinds of analysis tasks that a person might want to perform. Do NOT make all of your questions be about correlations or min or max values.

4. Write a report –

Part 1 – List your minimal three questions and answers, along with a screen shot showing the visualization you used to answer each question. One page per question – screen shot and narrative. Each question should be answered with a different visualization.

Part 2 – Critique the visualization tool . What are the tool's strengths and weaknesses? For what kinds of user tasks is the tool particularly well suited? Focus more here on the visualization techniques as opposed to the particular user interface quirks, though you should feel free to comment on UI aspects when they are particularly good or bad. Describe characteristics of the UI using the concepts and terminology you have learned in class.

In your write-up, you should provide a rigorous rationale for your design decisions. Document the visual encodings you used and why they are appropriate for the data. These decisions include the choice of visualization type, size, color, scale, and other visual elements, as well as the use of sorting or other data transformations. How do these decisions facilitate effective communication?

You are also free to incorporate external data as you see fit. Your chart image should be interpretable without recourse to your short write-up. Do not forget to include title, axis labels or legends as needed!

Submit a text file describing the visualization and your design rationale (≤4 paragraphs), a copy of your visualization in PDF format, and the source code for the visualization.

Extra credit: Build ETL process to transform the source data into star schema, upload your ETL code, and final schema and data.