Case Assignment:
Revenue Generated by a Movie Asset

Background

Your client is a Media agency. It runs a TV Network in US. The primary source of their revenue/audience comes from the Movies telecasted.

Data

The data corresponds to movies shown on a television network in the US. Two datasets have been provided

- Asset revenue data: this has the information on which asset/movie was shown at what time and the revenue generated
- ASSET_METADATA: Characteristics of the asset (Imdb ratings, tomato ratings, release date, genre etc.)

Use the two datasets and after performing the appropriate transformations (joining, cleaning up, feature engineering) to the data, create a model dataset to answer the following questions. Feel free to include some other information, like list of public holidays:

Objective:

1. Exploratory Analysis: Give good data based justifications to the questions. If you have intuitions about something try to justify that with data – that could be through a hypothesis test or even a bar graph!

2. Prepare an set of charts and tables to highlight any trends or patterns you observe in the data.

3. Create a model with the data to estimate revenue that an asset will generate if telecasted at a given time point (day-part and date).
    a. Will the revenue generated by an asset depend on external factors, like what time the movie was shown, or on factors internal to the asset or on both?
    b. Is there an interaction between when the movie was shown and asset characteristics, like genre?
    c. Is there a linear dependence of revenue on the factors you found?
    d. Is there any seasonality effect in the data? Are there any holiday effects?
    e. Can you test if a movie of a specific genre can generate more revenue if it's shown during the Prime time (fringe) of a weekend against the primetime of any weekday?