FINAL PROJECT PROPOSAL

MBAI 5100: Business Analytics

Group Members:

Hemaswi Gorrela - 101037597

Qian Yang - 101033621

Ngan Nguyen Le Thu - 101037583

Wan-Ting Chang - 101037722

1. Business Question:

How can we predict the future ticket sales and revenue of different movies in a cinema based on historical data of sales, showtime schedules, and pricing trends?

2. Datasets to be Used:

We will use a dataset from Kaggle.com to answer our business question.

Here is a link to the dataset: Sarder, H. (2023). *Cinema hall ticket sales and customer behavior* [Data set]. Kaggle. Retrieved from https://www.kaggle.com/datasets/himelsarder/cinema-hall-ticket-sales-and-customer-behavior/data

3. High-Level Description of the dataset

- An overview of this dataset:
  - Total observations (rows): 1440
  - Total variables (columns): 7
- Columns Overview:
  - Ticket_ID - unique ticket ID for the customers.
  - Ticket_Price - the price paid per ticket.
  - Age - age of the customer, useful for understanding audience demographics.
  - Movie_Genre - genre of the movie watched, which may influence demand and pricing.
  - Seat_Type - indicates whether the seat was regular, premium, or VIP, affecting ticket price.
  - Number_of_Person - number of people included in the transaction, linked to group sales behavior.
  - Purchase_Again - whether the customer made repeat purchases, relevant for customer loyalty analysis.

Ticket_ID	Age	Ticket_Price	Movie_Genre	Seat_Type	Number_of_Person	Purchase_Again
N4369	55	12.27	Comedy	Standard	7	No
B8091	35	19.02	Drama	Standard	Alone	Yes
V6341	55	22.52	Horror	VIP	3	No
B3243	53	23.01	Drama	Standard	6	Yes
I3814	30	21.81	Comedy	VIP	4	Yes
E5655	28	11.58	Horror	VIP	Alone	Yes
P1526	50	22.91	Action	Standard	Alone	Yes
V4726	44	23.09	Sci-Fi	Premium	7	Yes
A2029	46	12.12	Sci-Fi	Standard	Alone	Yes

The most important variables that we believe can help us with our goal are Ticket_Price, Movie_Genre, and Number_of_Persons. Ticket_Price directly affects overall sales and reflects how pricing influences customer demand. Movie_Genre helps identify which types of films attract larger audiences and generate higher revenue. Number_of_Person indicates the group size per transaction, which plays a significant role in determining total ticket volume.

4. Similar Projects:

A similar project by Apala et al. (2013), Prediction of Movies Box Office Performance Using Social Media, used data from Twitter, YouTube, and IMDb to classify movies as hits, neutrals, or flops based on social media engagement and sentiment. However, our project uses a different dataset focuseing on cinema ticket sales and customer behavior, which provides transaction-level insights. Instead of relying on social media buzz, we aim to predict ticket sales or revenue trends using customer demographics, showtimes, and movie genres. This method offers realistic, data-backed insights that help improve cinema operations and marketing strategies, emphasizing efficiency and decision-making over predicting movie popularity.

Here is the link to this project:

Apala, K. R., Roy, S., & Choudhury, T. (2013). *Prediction of movies box office performance using social media.* Retrieved from: https://www.researchgate.net/publication/262249589_Prediction_of_movies_box_office_performance_using_social_media