

CS798H : Project Milestone-2

Group Members: Ayush Yadav(210251), Pranjal Gautam(210742) & Shri Krishna Sahu (211011)

Workpiece Service: BookMyShow

Data Gathering: Web Scraping

Review Sources	Total no. of reviews	No. of Newest reviews taken	Reviews after release of latest version(i.e version 14.9.1)	1 star + 2 star rating reviews of latest version	Reviews with UI/UX based reviews	Reviews with Non trivial UI/UX reviews
Play Store	1500 k	75k	3k	2k	0.11k	0.025k
App Store	738k	20k	1k	0.75k	0.05k	0.015k
MouthShut	18.5 k	5k	0.45k	0.3k	0.02k	0.005k

Data Gathering Method:

We applied Web Scraping technique (It is the process of extracting data from websites. It involves fetching web pages using HTTP requests to the URL of the webpage we want to access, then parsing the HTML content or other structured formats using parsers like BeautifulSoup, and extracting the desired information), This can be done by manually writing the python scripts or using a scraper like Google-Play-Scraper(it is a tool meant to extract publicly available application data from Google PlayStore) for play store specific reviews.

Sampling Strategy:

To gather reviews data from our top 3 resources (**Play Store, Appstore, Mouthshut.com**) eliminating data from (**PissedConsumer and TrustPilot**) on the basis of a very low percentage of their reviews contribution in total population i.e (**approximately , < 0.1 %**).

- For example: We used Google-Play-Scraper which provides APIs to easily crawl the Google play Store for python , from the extracted app info we got **total 1500k** reviews were present then **75k Newest reviews** based on dates at which reviews were posted were selected . To cater the issues which are not yet resolved , reviews posted after the release of the latest app version (version 14.9.1) were taken into account which led to further reduction of sample size to **3k** . Then **Rating based sorting** is done to extract **1 star and 2 star reviews** to work only on negative reviews ,this less reduction to only **2k reviews** is because of ignorance of the company towards resolution of users issue in their latest version . Now to focus on our project purpose we needed only **UI/UX based issues** for which we removed reviews which were based on cancellation, refund and event management issues that led to sudden decrease in sample size to **0.11k** ,this observation also reveals the fact that the company also lacks in execution of its Business. At last as we were instructed to work only on **Non trivial issues** , we manually selected such reviews and were left with final **45 issues**.

Similar strategy of web scraping is applied to extract required reviews from other 2 sources i.e (PissedConsumer and TrustPilot) , at finally we got total of (45 Non Trivial reviews) , Reading those 45 reviews , we selected **30 reviews** based on top **7 issues** on the basis of their frequency.

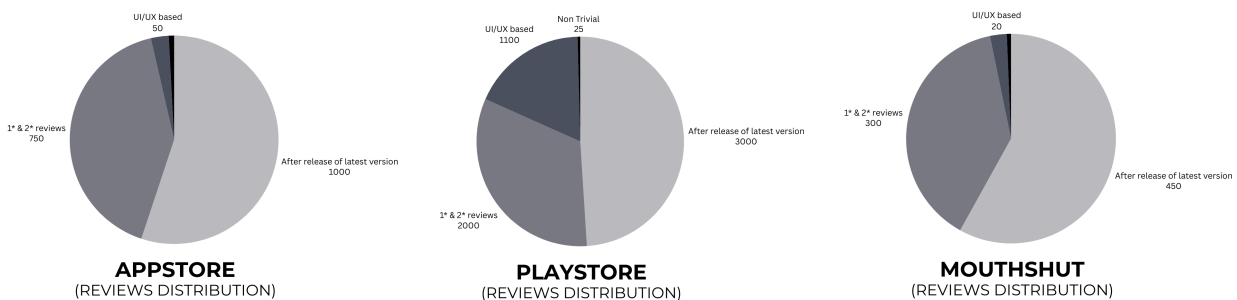
Total population size: $1500k + 738k + 18.5k = 2256.5k$

Sample Size: 30 reviews

Nature of Data Gathered:

1. Feedback and Ratings: User feedback, reviews, and ratings provided for content, recommendations, and overall user experience.
2. Demographic Information: Optional demographic data such as age, gender, location, which can help tailor recommendations and personalize the user experience.

Data Analysis



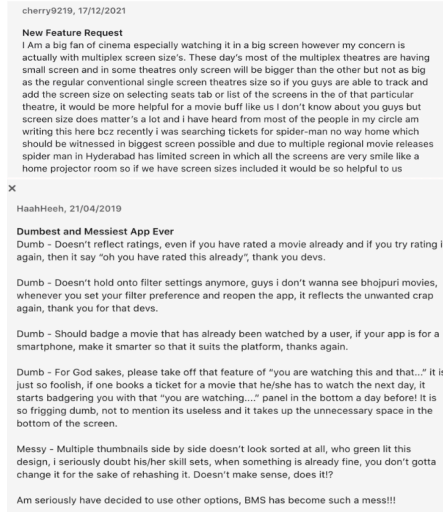
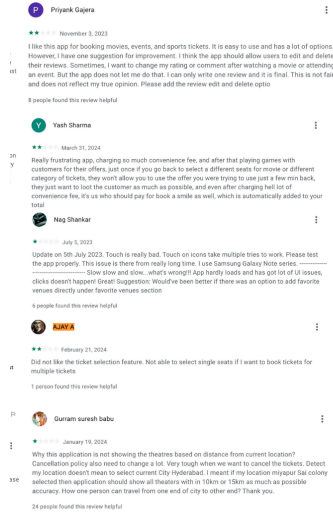
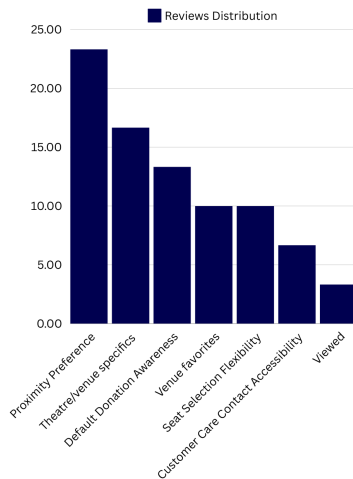
After collecting and categorizing 30 reviews, we applied thematic analysis to identify recurring issues and concerns. Each problem was tagged accordingly, and we calculated the percentage of occurrences for each tagged problem. Using a bar graph to visualize the frequency of these problems, we identified 5 reviews for in-depth analysis to explore potential solutions.

Benefits of Thematic Analysis:

- Identifies key user concerns and frustrations.
- Provides direction for further investigation through targeted surveys or user interviews.

Limitations:

- Thematic analysis can be subjective, depending on the analyst's interpretation.
- Deeper insights might require more advanced qualitative coding techniques.



Results

- **Proximity Preference:**
BMS only gives a filter of the city and not distance from your current location. So, at times users end up booking an event/movie very far.
- **Default Donation Awareness:**
During billing, BMS adds a charity donation of Rs.2 by default(with an option to remove) and this goes behind the eyes of many users.
- **Customer Care Contact Accessibility:**
Customer care number is not available on the website and requires a lot of searching to get it.
- **Theatre/venue specifics:**
At many times the venue stands under par from the set expectations leading to an overall bad experience for users like small screen size. BMS should provide details about the venue like screen size, reviews of the venue which will facilitate the users to make a better choice.
- **Viewed(Badge a movie):**
Movies or events that have once been watched by the user to be tagged as watched.
- **Venue Favorites:**
There should be an option to add favorite venues and the updates of venues should be given to the user.
- **Seat Selection Flexibility:**
Whenever you book multiple seats together, once you select one seat, other seats get automatically chosen consecutively and you cannot choose other seats according to your choice. And many times different people in the group have different priorities.
- **Change in seat location:**
Users want to be given an option to change the seat location after the seat is once booked.

Prioritization

We have selected the following problems to work on, in order of priority:

1. **Proximity Preference:**
This problem exists mostly for ultra big cities like Mumbai, Bangalore, Delhi. Also most consumers usually know the location of the venue before booking the ticket. Hence, not the top-top priority but a big issue to cater to.
2. **Theater/venue specifics:**

This is a very beneficial feature to add. It is right of the customer to know exactly what they are paying for.

3. **Seat Selection Flexibility:**

This is an issue that we even faced while using the platform. As this was a bad personal experience our human nature tilts us to solve this problem first. Also this issue is the biggest because we feel if this is not fixed, few users have to book tickets in 2-3 iterations which could be done in a single time or they compromise on their choices.

4. **Default Donation Awareness:**

We also feel this is in the gray area of being unethical and it should be given as an option not included by default. As its use is for good purpose, it can wait xD.

5. **Venue favorites :**

This is least prioritized because we feel this is not some huge issue that is faced by consumers or big aid to the consumers.

This is an issue that we even faced while using the platform. As this was a bad personal experience our human nature tilts us to solve this problem first. Also this issue is the biggest because we feel if this is not fixed, few users have to book tickets in 2-3 iterations which could be done in a single time or they compromise on their choices.

Next are the reasons why we didn't go for other issues. Staying limited in this project work is also one of the good reasons for excluding other problems.

1. **Change in seat location**

We didn't choose this because this is not an issue that many people face. People usually decide the best seat available while booking only. There was only 1 review regarding it.

2. **Customer Care Contact Accessibility**

Providing customer care numbers was a trivial change that's why we didn't use it.

3. **Viewed(Badge a movie)**

This is irrelevant for a platform like BMS. Usually a movie stays in the theater for 3-4 weeks and an event happens only once in a while. It is a very important feature for platforms like Netflix but irrelevant for BMS

1 liner description of our approach to solve the problems prioritized:

1. **Distance of Venue:**

We can add a slider filter as seen in the app like(Tinder) of distance and then accordingly the site will show events/movies in a set radius.

2. **Theater venue specifics:**

We can add specifics like screen size, photos,videos of theater and event venue etc while booking.

3. **Individual seat selection while booking multiple tickets together:**

We will give them an option to select different seats during seat map selection.

4. **Charity fee:**

We will give a plane face emoji as default and will give a choice to the user during time of booking to click it to make it a smiling face emoji and then only 2 rs donation will be added in total.

5. **Venue favorite:**

We will make a feature where you can add a venue to your favorites and whenever there is a new event at the venue, notification would be given to the user and even during further bookings these venues will be kept at priority.

Appendix

 **CS798_Milestone2_ThematicAnalysis**

Here's the code:  **CS798_Milestone2_WebScraping_Code**

