Process Book - Boxoffice 360

Team Details

Project Repository: https://github.com/anirajk/boxoffice360
Aniraj Kesavan
aniraj@cs.utah.edu
u0996550

Ashwini Janamatti
ashwinijanamatti@gmail.com
u0996548

Process Book – Boxoffice 360
Aniraj Kesavan – u0996550
Ashwini Janamatti - u0996548

Overview and Motivation

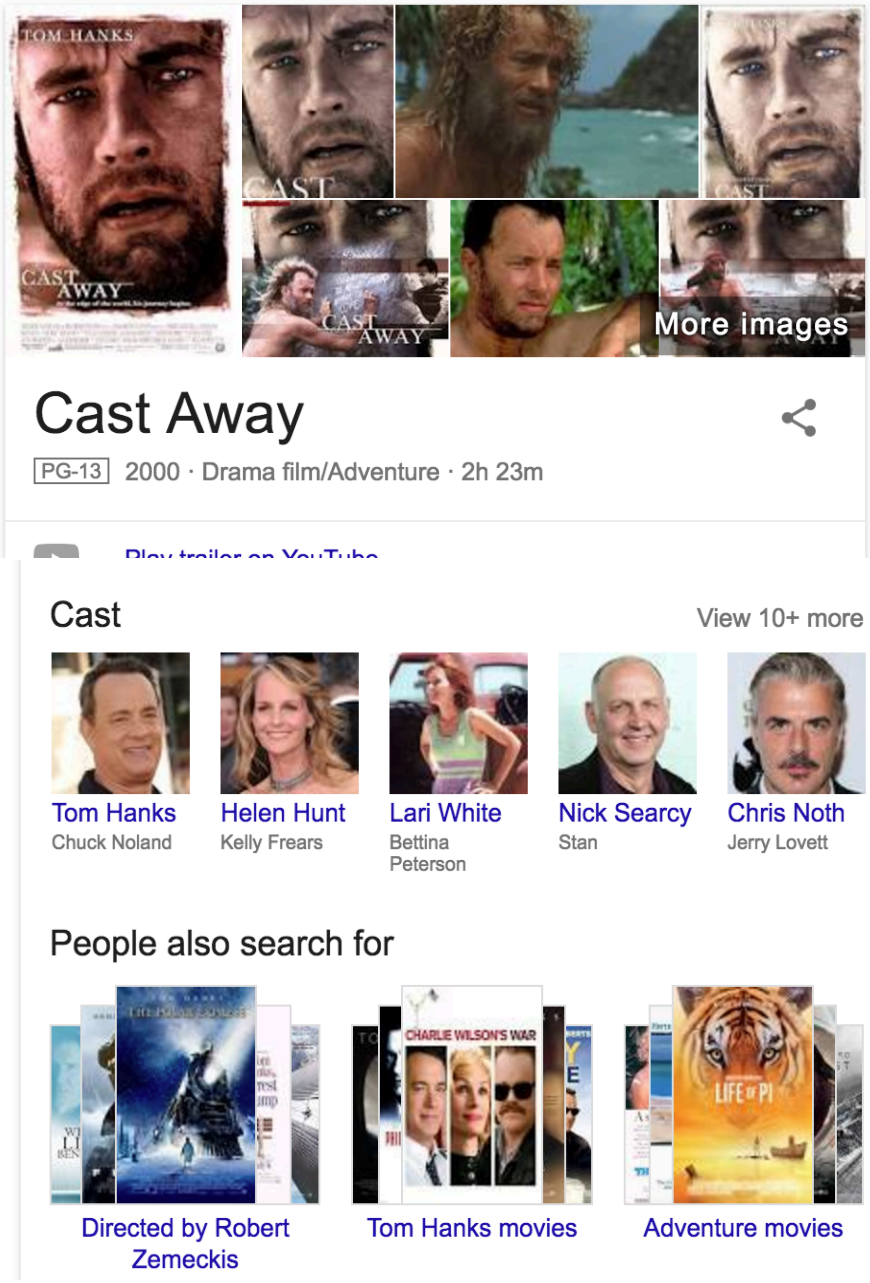
By the end of the 1880s, the introduction of lengths of celluloid photographic film and the invention of motion picture cameras, which could photograph an indefinitely long rapid sequence of images using only one lens, allowed several minutes of action to be captured and stored on a single compact reel of film. Some early films were made to be viewed by one person at a time through a "peep show" device such as the Kinetoscope. Movies have grown from paltry beginnings to complex multi year projects producing highest fidelity pictures using 3-D Stereoscopic projection etc.

There is a staggering amount of movies made over generations and there is no good visualization tool that exists today which captures the essence of the evolution of movies, nor does a solution exist which can visualize movie data per year with the ability to sort and filter based on attributes of a movie. Solutions such as IMDB exists, but they are better suited to extract more information than general exploration

In this project, we plan to visualize a well curated dataset of over 5000 movies. We pre-processed the data using imdbs apis to enhance information and visualize both the holistic evolution of movies as well as enhanced views of movies released every year.

Related Work

The closest related work that we have seen that inspired us to develop the per year visualisations are the new and improved google movie specific search results, attaching screenshots here:


TOM HANKS
CAST AWAY
More images
Cast Away
PG-13 2000 · Drama film/Adventure · 2h 23m
Cast
View 10+ more
Tom Hanks
Chuck Noland
Helen Hunt
Kelly Frears
Lari White
Bettina Peterson
Nick Searcy
Stan
Chris Noth
Jerry Lovett
People also search for
Directed by Robert Zemeckis
CHARLIE WILSON'S WAR
Tom Hanks movies
LIFE OF PI
Adventure movies

Questions

Through visualizing the data on movies, we are trying to answer the following questions:

- ?
- ?
- ?

By answering these questions, we will be able to get a big picture of trends in terms of evolution of movies overe decades, how well movies have performed in their release year in terms of boxoffice results, current relevance (likes in social media etc), duration, ratings etc. The selective exploration of movies released every year also helps in diving more into the details and helping the user make informed opinions.

Data

- Our primary source of data was a curated dataset of over 5000 movies that we researched and found at kaggle.com
  - https://www.kaggle.com/deepmatrix/imdb-5000-movie-dataset

- Since people identify movies best with their posters, we wanted to include movie posters in our visualization, but we didn't want to store over 5000 images since that would take up serious amounts of space and hamper client loading performance. So we made use of a dynamic api provided by wemakesite creator Martin Ivanov. The api is available at
  - http://imdb.wemakesites.net/api/
- The api returned json results like the following:
  -
```
{
   "status":"success",
   "code":200,
   "message":"ok",
   "data":{
      "id":"tt0115956",
      "type":"title",
      "title":null,
      "year":null,
      "description":"The pilot of a rescue copter, Cap-
tain Karen Walden, died shortly before her helicop-
ter crew was rescued after it crashed in De-
sert Storm. It first  Written byBrian W Martz <B.Martz@Ge-
nie.com>",
      "certificate":"",
      "duration":"1h 56min",
      "released":"1996-07-12",
      "cast":[
```

```
            "Denzel Washington",
            "Meg Ryan",
            "Lou Diamond Phillips",
            "Michael Moriarty",
            "Matt Damon",
            "Bronson Pinchot",
            "Seth Gilliam",
            "Regina Taylor",
            "Zeljko Ivanek",
            "Scott Glenn",
            "Tim Guinee",
            "Tim Ransom",
            "Sean Astin",
            "Armand Darrius",
            "Mark Adair-Rios"
         ],
         "genres":[
            "action",
            "drama",
            "mystery",
            "thriller",
            "war"
         ],
         "directors":[

         ],
         "writers":[

         ],
         "image":"https:\/\/images-na.ssl-images-ama-
zon.com\/images\/M\/MV5BMTMwNDk0NjU-
zOF5BMl5BanBnXkFtZTcwNjcx-
ODIyMQ@@._V1_UY268_CR2,0,182,268_AL_.jpg",
         "review":{
            "text":"Courage un-
der fire was an okay movie, the story was good enough to ke
ep me : 6 out of 10.",
            "rating":"7\/10"
         }
      }
   }
}
```

Process Book – Boxoffice 360
Aniraj Kesavan – u0996550
Ashwini Janamatti - u0996548

Data Processing

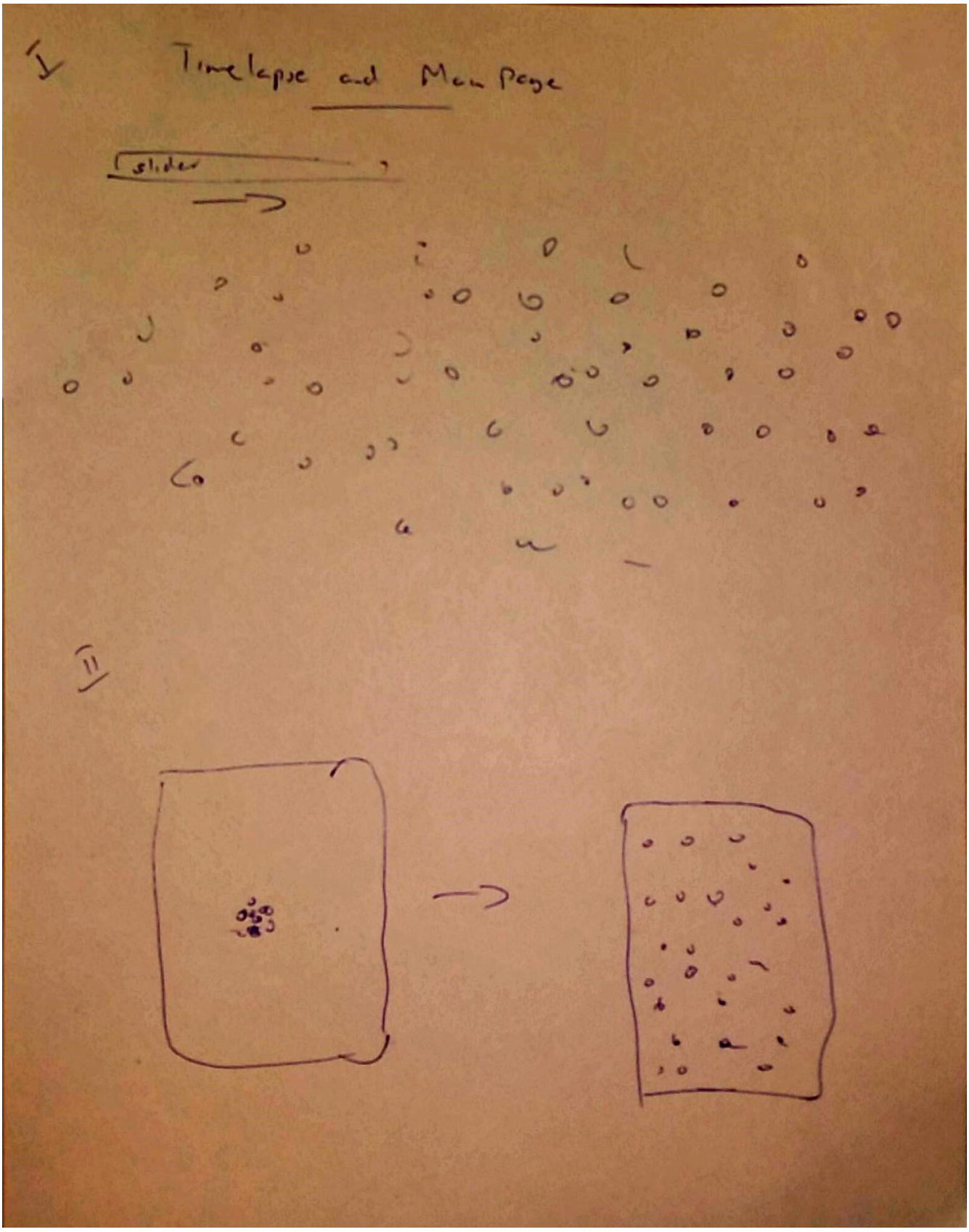
While our primary data was found from a reputable source, we still found some duplicate entries among the data source. In order to make it pluggable with any data set, we decided to implement our own logic to de-duplicate entries. This is done before any processing.

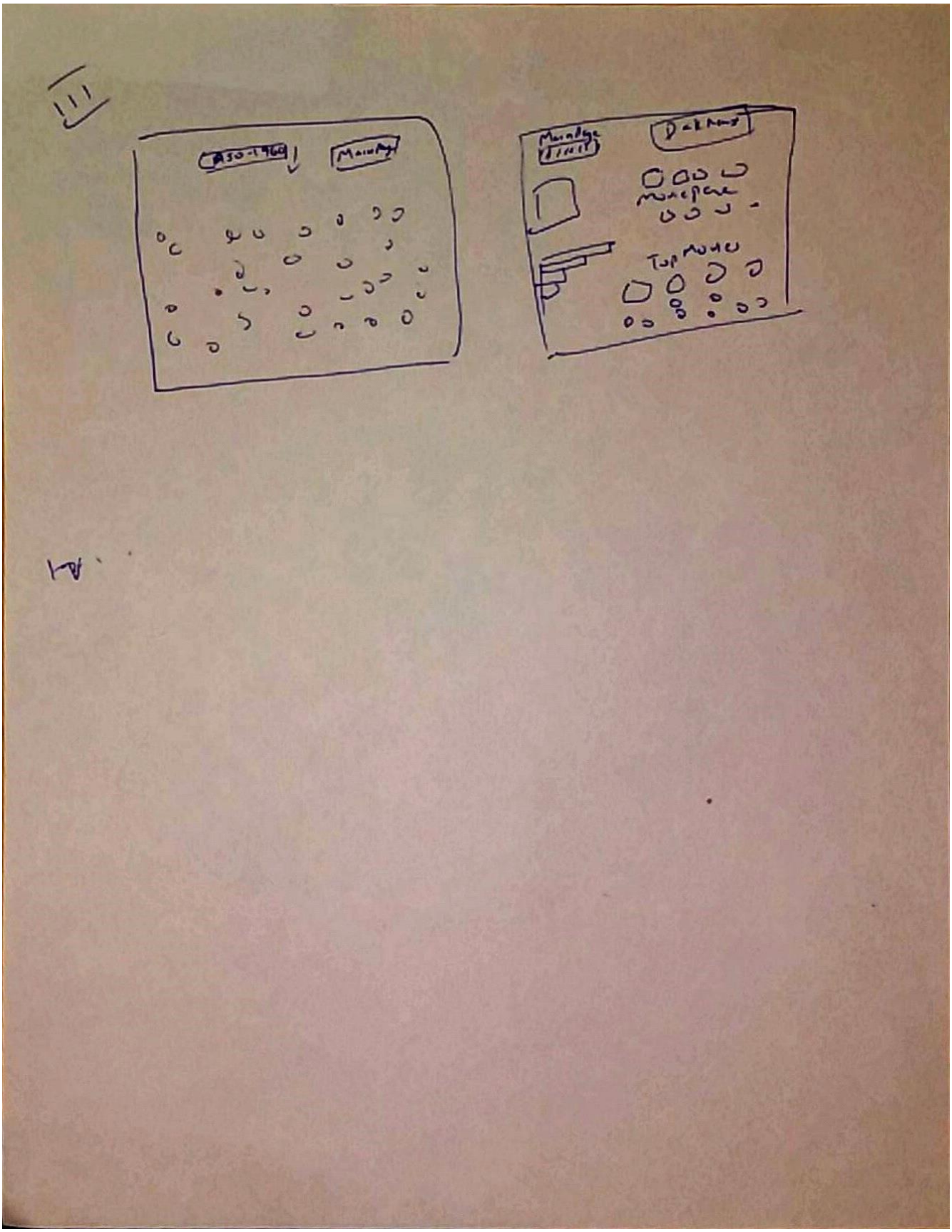
We were initially hitting the imdb api asynchronously whenever we wanted to load poster images, this involved a lot of network transfer. We decided we could pre-fetch the image url and tack that on to the data, so we did a run of data enhancement to enrich the data.

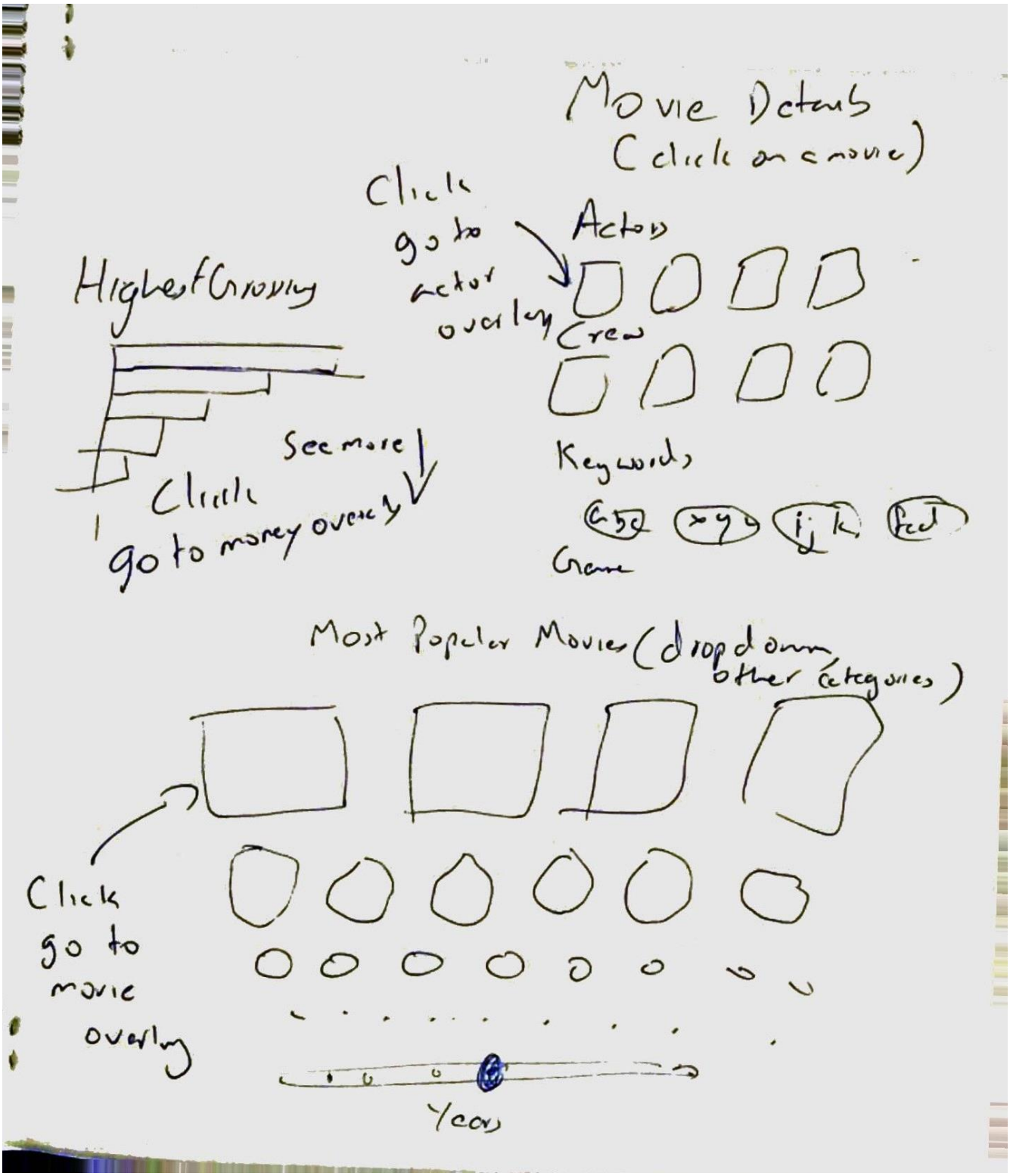
While enriching the data, we also found that there are other useful information from the response that was otherwise missing from our primary data source, so we merged our primary data with the newly found data.

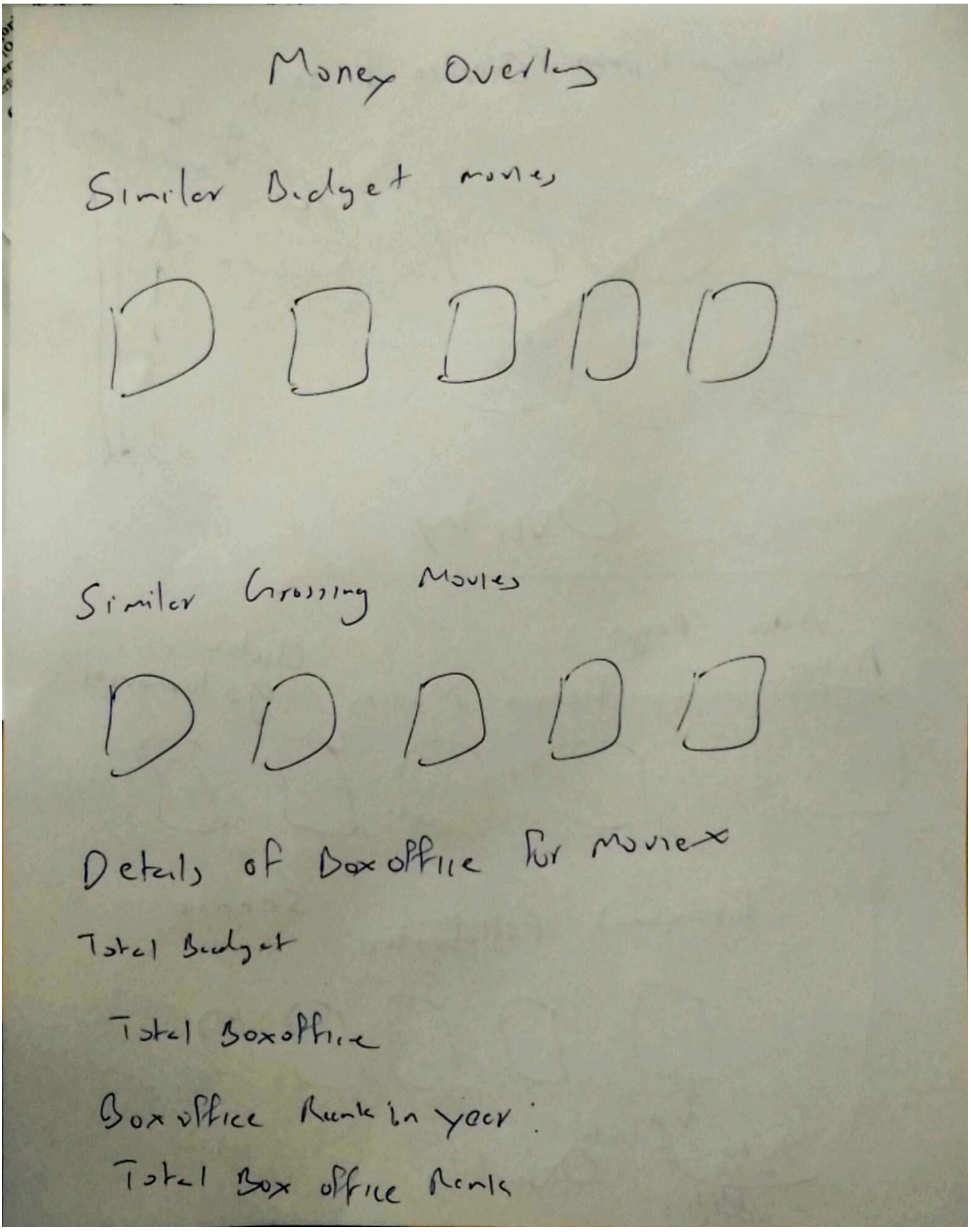
Exploratory Data Analysis

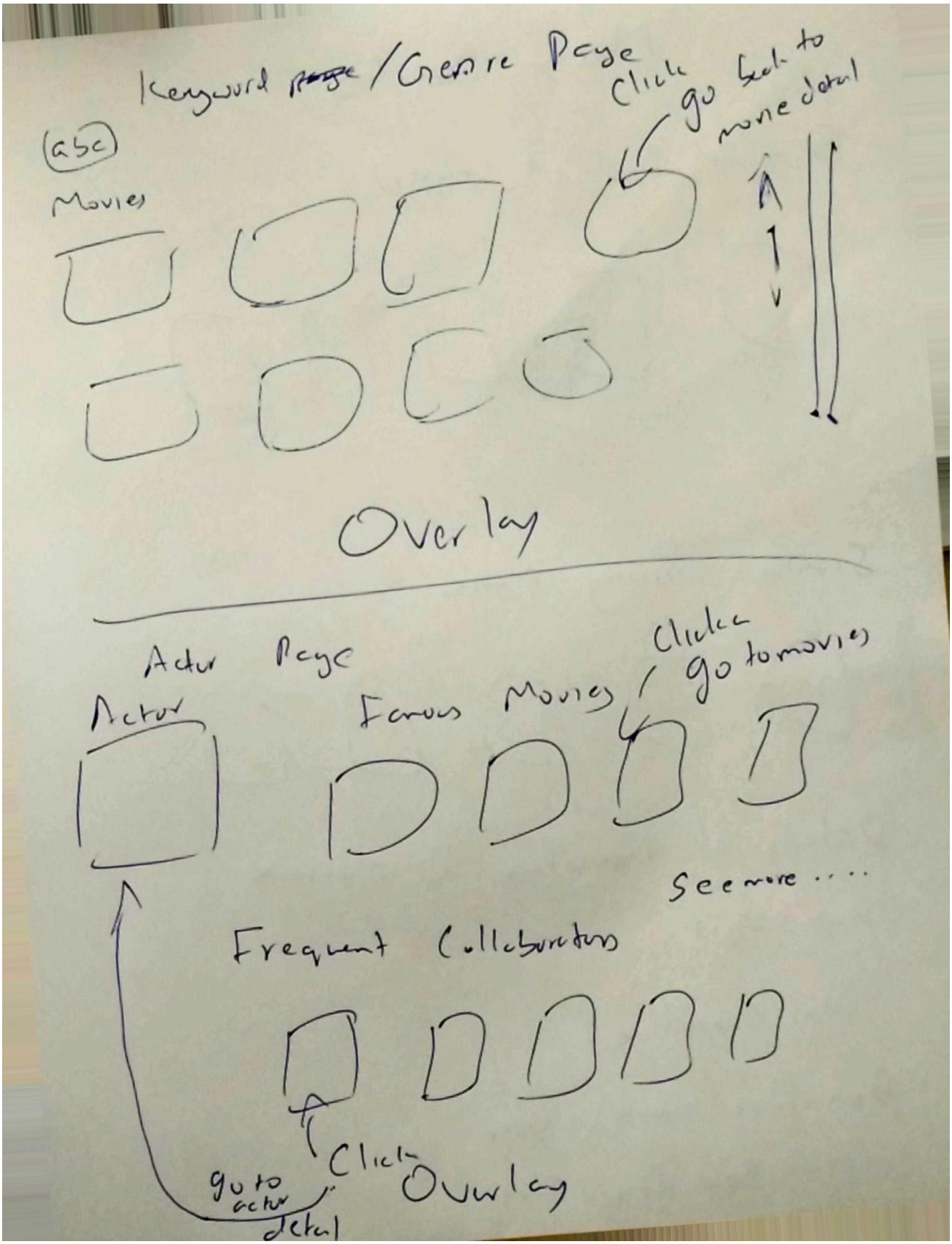
Design Evolution


I
Timelapse and Map Page
slider
II




Movie Details
(click on a movie)
Click
go to
actor
overlay
Actors
Crew
Highest Grossing
See more
Click
go to money overlay
Keywords
Genre
Most Popular Movies (dropdown
other categories)
Click
go to
movie
overlay
Years


Money Overlay
Similar Budget movies
Similar Grossing Movies
Details of Boxoffice for movie
Total Budget
Total Boxoffice
Boxoffice Rank in year:
Total Box office Rank


Keyword / Genre Page
Click go back to movie detail
(asc)
Movies
Overlay
Actor Page
Actor
Famous Movies
Click go to movies
See more ....
Frequent Collaborators
Click Overlay
go to actor detail

Process Book – Boxoffice 360
Aniraj Kesavan – u0996550
Ashwini Janamatti - u0996548

Must-Have features

Optional features

Implementation

1. **Timelapse**
2. Per Year.

Evaluation