BEST VIDEO CONTENT for OVER-THE-TOP (OTT) STREAMING SERVICE

MSBD5005 Data Visualization — Phase 1

Group 16: Chan Kar Chun (20729353), Lakhani Harsh Sunil (20910249),
Tsang Kai Ho (20905476), Wong Yan Ho (20605624)

TABLE OF CONTENTS

01 DATA

DATA — Overview

Streaming Platform	Number of Features	Number of Videos	TV Show Percentage	Movie Percentage
Amazon Prime Video	18	9871	14%	86%
Disney+	18	1535	27%	73%
HBO Max	18	3294	23%	77%
Netflix	18	5850	36%	64%

DATA — General Structure

Attributes	Description	Attributes	Description
id	The video ID provided on JustWatch	seasons	The number of seasons of a show
title	The name of a video	name	The name of the actor/actress of a video
type	TV show or movie	character	The character name of the actor/actress in the video
description	A brief description of a video	role	Actor/actress or director
release_year	The year of release of a video	imdb_id	The video ID on IMDb
age_certfication	Age certification for a video	imdb_score	The score of a video on IMDb
runtime	The length of a movie or an episode of a TV show	imdb_votes	The votes of a video on IMDb
genres	The genres of a video	tmdb_score	The score of a video on TMDb
production_countries	The countries produced a video	tmdb_popularity	The popularity of a video on TMDb

DATA — Source

Streaming Platform	Source URL
Amazon Prime Video	https://www.kaggle.com/datasets/victorsoeiro/hbo-max-tv-shows-and-movies
Disney+	https://www.kaggle.com/datasets/victorsoeiro/amazon-prime-tv-shows-and-movies?select=titles.csv
HBO Max	https://www.kaggle.com/datasets/victorsoeiro/amazon-prime-tv-shows-and-movies?select=titles.csv
Netflix	https://www.kaggle.com/datasets/victorsoeiro/amazon-prime-tv-shows-and-movies?select=titles.csv

PURPOSE 02

PURPOSE

BACKGROUND

Survey findings indicate that although Amazon offers more content, Netflix reigns as the more popular streaming service. This could be because of the quality of content.

GOAL

Our aim is to help Content Producers create popular content by analyzing various attributes of movies and TV shows and providing valuable insights.

TASKS

Explore the optimal combination of attributes that drive popularity in video content.

03

INITIAL PLAN

System Design

Introduction

Introduction

Overview

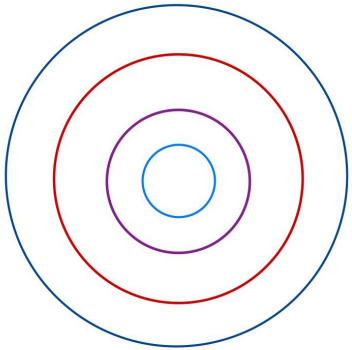
In-depth Analysis

Recommendation

Best Video Content for Over-The-Top(OTT) Streaming Service

Introduction

# of videos released by each platform



Show the data implicitly
Give a general view

# of videos released by each platform over years

	1960	1970	1980	1990	2000	2010	2020
Prime	xxx	xxx	xxx	xxx	xxx	xxx	xxx
Netflix	xxx	xxx	xxx	xxx	xxx	xxx	xxx
HBO	xxx	xxx	xxx	xxx	xxx	xxx	xxx
Disney+	xxx	xxx	xxx	xxx	xxx	xxx	xxx

Show the data explicitly
categorize by years (decades) can show the trend.

System Design Overview

Introduction

Overview

In-depth Analysis

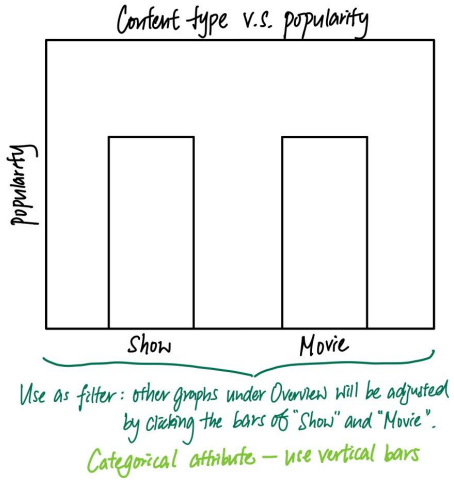
Recommendation

Overview — • Describe our dataset
• compare candidate attributes with "popularity"

Description of data
- # of videos, attributes • details of attributes
- popularity measure : Popularity Index per Content (PIC)

PIC = avg(normalized IMDB votes, normalized TMDB popularity)

Heat map
to see the correlation of each attribute



Use as filter: other graphs under Overview will be adjusted by clicking the bars of "Show" and "Movie".

Categorical attributes — use vertical bars

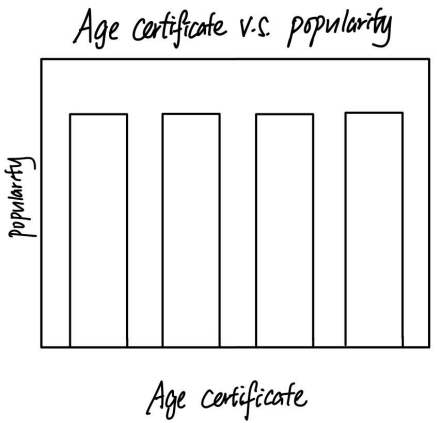
System Design Overview

Introduction

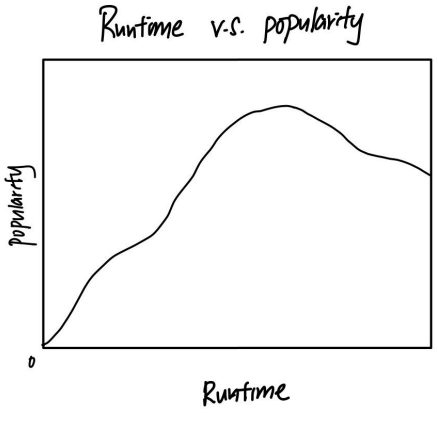
Overview

In-depth Analysis

Recommendation

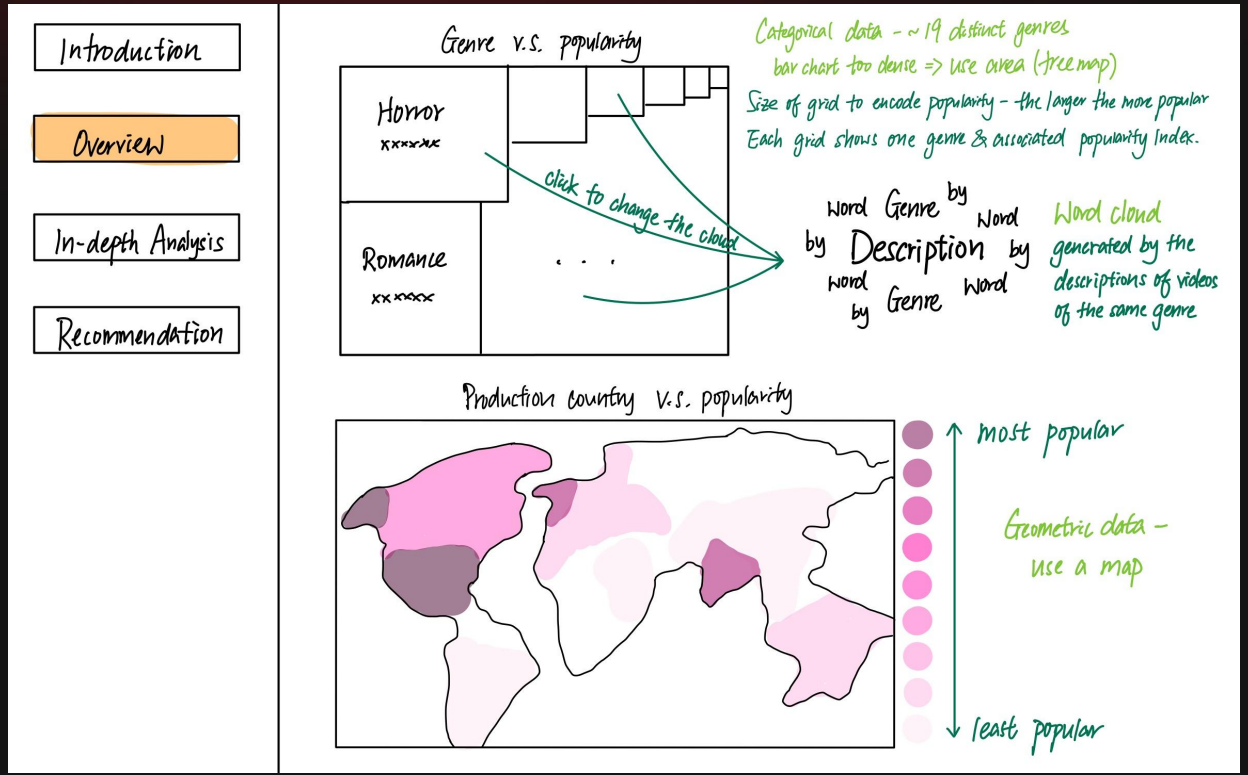


Categorical data - use vertical bars



Continuous data - use line

System Design Overview



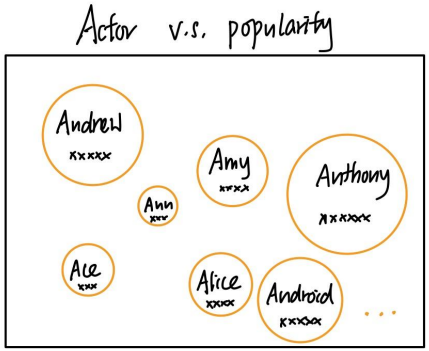
System Design Overview

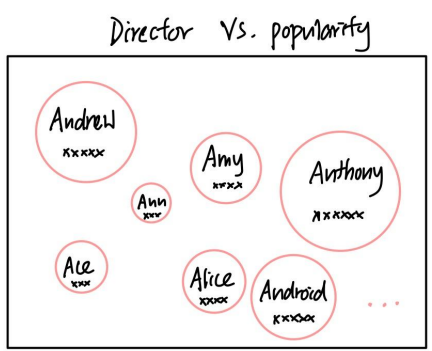
Introduction

Overview

In-depth Analysis

Recommendation





Actors & directors can form a network — use nodes

Size of node to encode popularity: the larger the more popular

Each node indicates an actor/a director

displays their name & average popularity

System Design

In-depth Analysis

Introduction

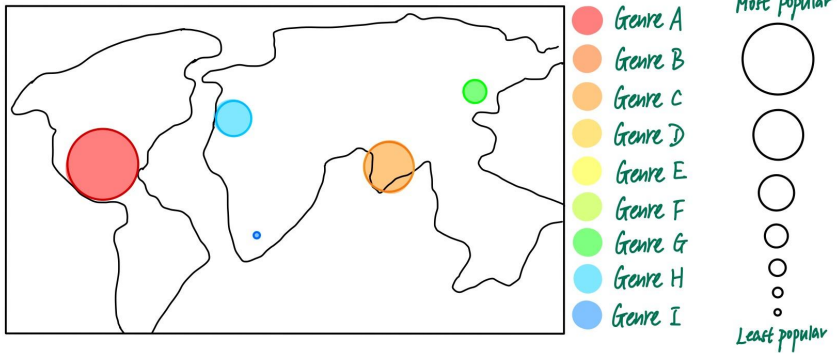
Overview

In-depth Analysis

Recommendation

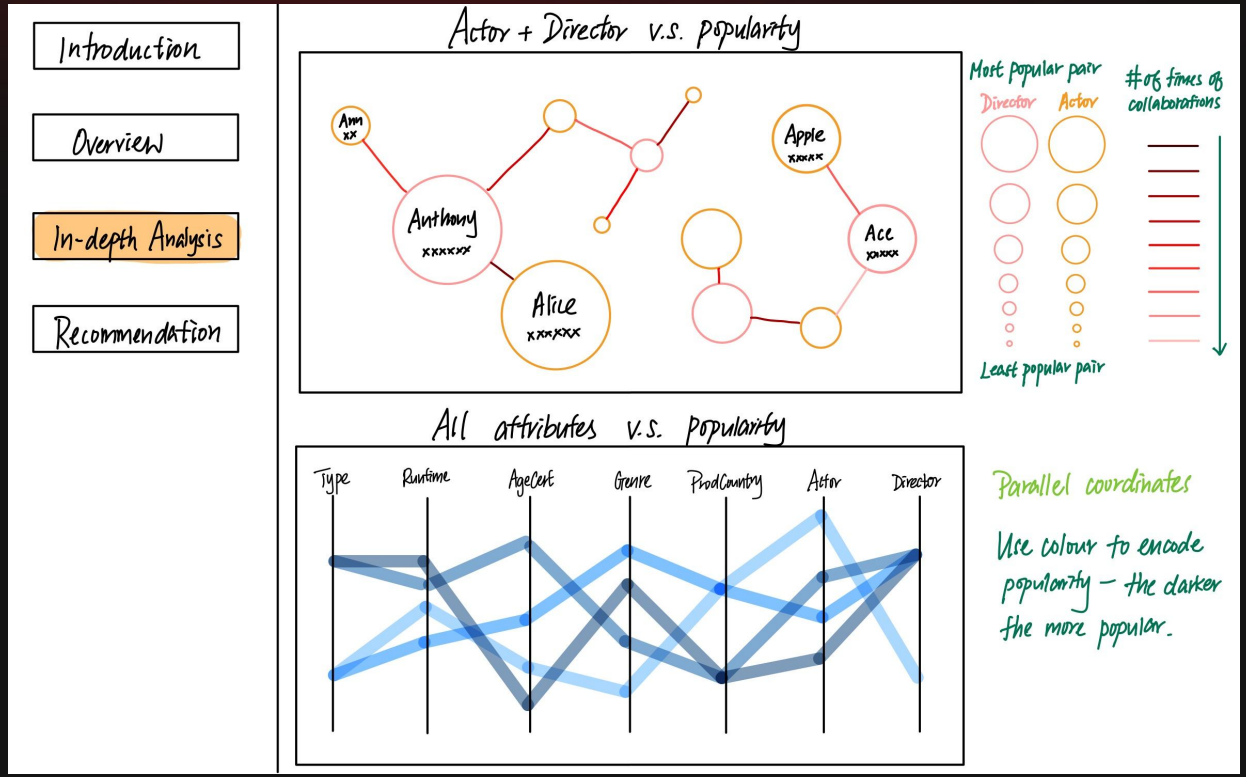
In-depth Analysis – Instead of viewing each attribute independently, we believe attributes can be interdependent and may bring synergy effect to "popularity"

Production country + Genre V.S. popularity



System Design

In-depth Analysis



System Design

Recommendation

Introduction

Overview

In-depth Analysis

Recommendation

Recommendation — • Potential popular video combinations
• popular video choice simulator

Potential combinations a popular video has :

①

②

③

Popular Video Choice Simulator

Filter 1 (Input/select)

Filter 2 (Input/select)

Results

Filter : One of the attributes discussed above

Yet to decide the number of filters

04

IMPLEMENTATION PLAN

Tool – Python

- Perform data preprocessing tasks including data cleaning, data transformation, and data integration

Tool – Streamlit

- Python-based interface with real-time data processing capabilities
- Ability to incorporate python based visualization libraries such as Matplotlib Plotly and Seaborn

Tool – Tableau

- Perform initial Exploratory Data Analysis (EDA) to get an understanding of the data

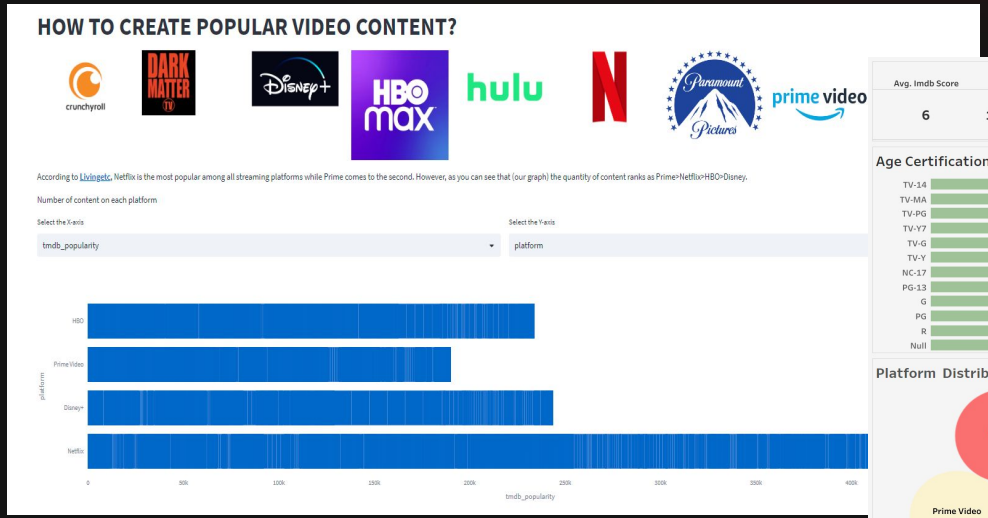
Overall Structure

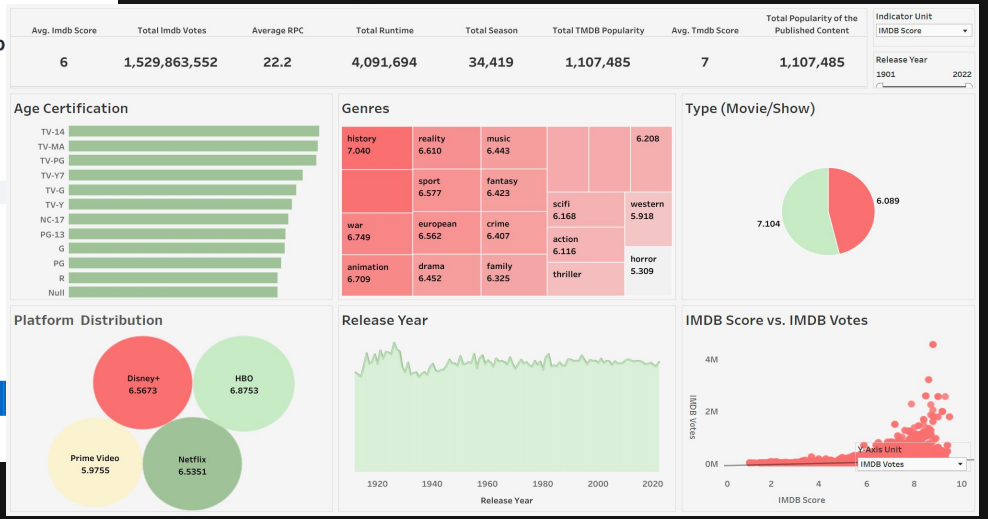
- Develop a dynamic web app with Streamlit, integrating Python libraries and Tableau visualizations for a seamless and interactive user experience

05

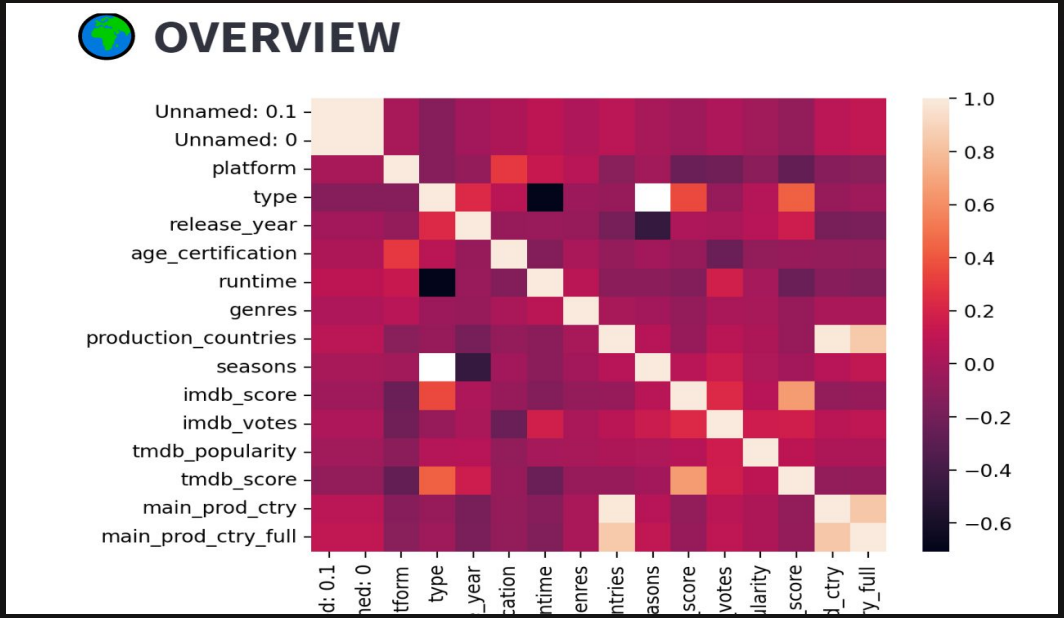
INITIAL RESULTS

Introduction





Overview

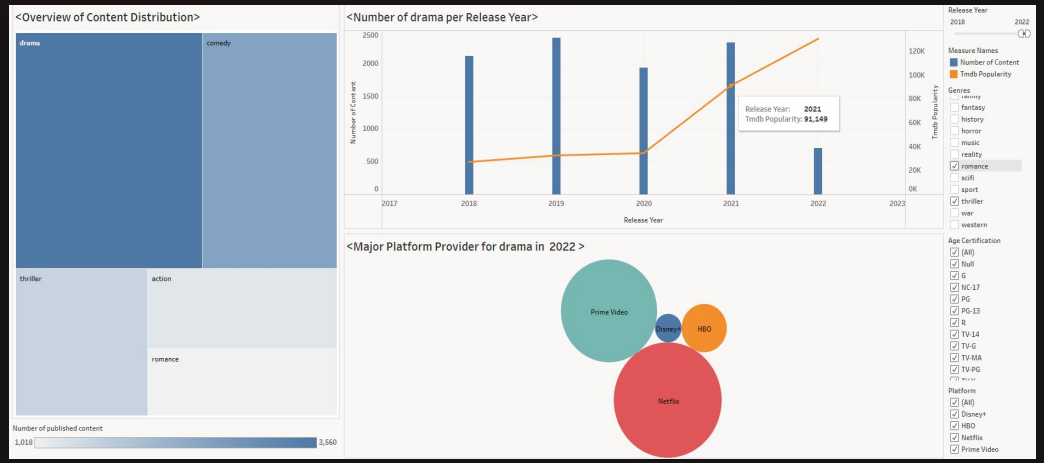

OVERVIEW
Unnamed: 0.1
Unnamed: 0
platform
type
release_year
age_certification
runtime
genres
production_countries
seasons
imdb_score
imdb_votes
tmdb_popularity
tmdb_score
main_prod_ctry
main_prod_ctry_full
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

In-depth Analysis

Key Insights:

1. Different Content Category has its own PRC
2. OTT Platform Strategy based on PRC

	2018-2022	
	Popularity Ratio per Content (PRC)	**Best Platform Seller**
Drama	34	Netflix
Comedy	41.6	Netflix
Thriller	75	Prime Video
Action	100	Prime Video
Romance	40	Netflix



Conclusion:

If you are a Thriller producer, Prime Video will be your partner if you want your content get popular

THANK YOU!

Please feel free to raise your questions.