'My' wonderful time with Netflix (2018-2023)

Team 0 | **Lin Yueyi** 1961779 | **Wei Yiyi** 1873027

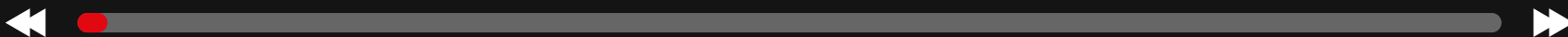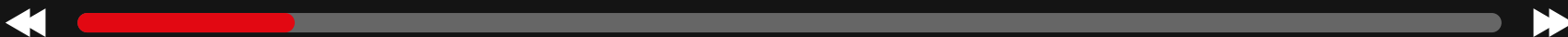▶ **Start**

+

CONTENTS

01

ABOUT 'ME'

My Time with Netflix (2018-2023)

ABOUT 'ME'

My Time with Netflix (2018-2023)
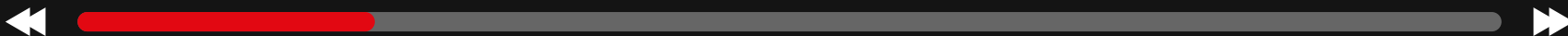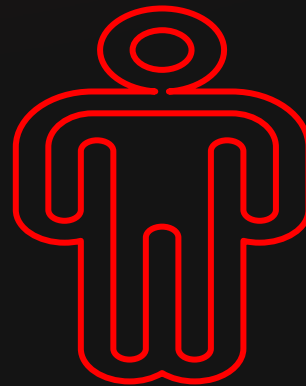
I've watched **4704** TV shows, **569** movies

I've spent **1864.11** Hours

Longest spent **02:49:04**

Maturity Rating: **16 +**

◀◀ ▶▶

Year Level

Period
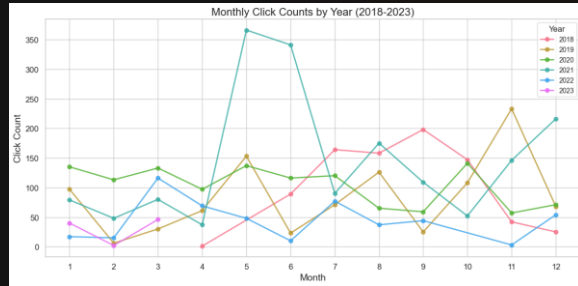
Fluctuate Monthly Trend

Always Low in Feb. Apr.

Netflix fervor increase 2018-2021

Netflix fervor declines 2022-2023
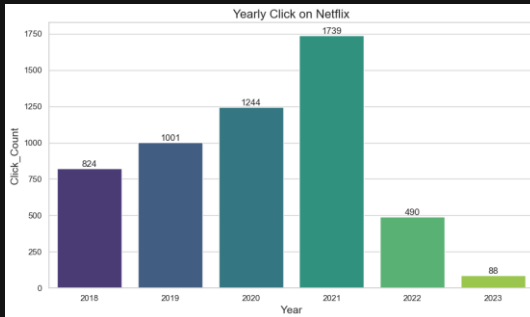
Period
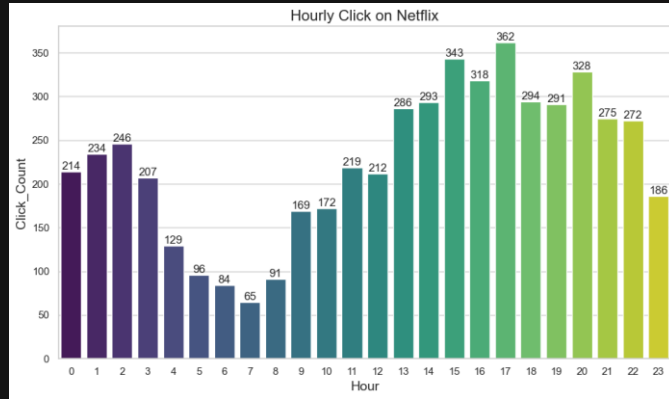
Usually Open Netflix at afternoon & night

Midnight belongs to Anime

Device

From 2018-2023,
I spent time with Netflix by **20+** devices...

**Google Chromecast
V3 Streaming Stick**

1363 times used (most)
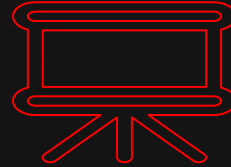
Apple iphone 11

Utilized throughout the
2020-2022

**Android
DefaultWidevineL3Tablet
Android Tablet**

Utilized throughout the
2018-2020
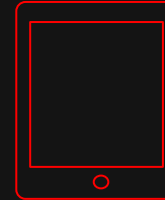
Track From 2018-2023,
I spent time with Netflix in **4** countries...

I've appeared in

Ireland 7 times during 2018.09-2018.11

Germany 1315 days from 2019 to 2023

Italy 2 times

Greece 1 times at the end of Aug. in 2019
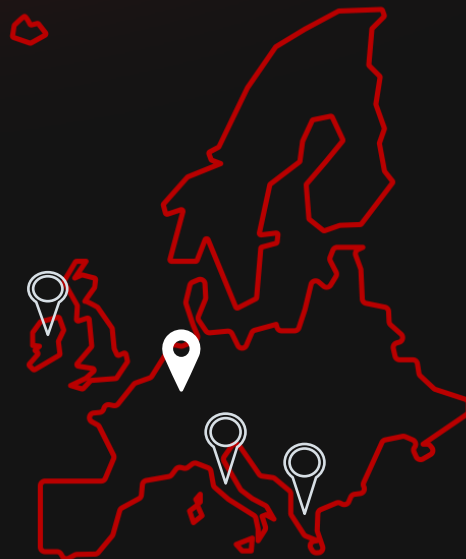
Type

Drama
30.4%
Android DefaultWidevineL3Tablet Android Tablet

Anime
27.2%
Google Chromecast V3 Streaming Stick

Comedy
13.9%
Android DefaultWidevineL3Tablet Android Tablet

Action
7.1%
Android DefaultWidevineL3Tablet Android Tablet

What is My Favorite?

Night

Android DefaultWidevineL3Tablet
Android Tablet

Rewatched 5 times
2018.08-2019.02

Favorite Movie

To All the Boys I've Loved Before

DATA WORK 02

Data Analysis Behind the Story

| **WORK PROCESS**

Data Crawling ⇒ **Data Preprocessing** ⇒ **Insights Analysis**

To merge a better own dataset

Data Cleaning

Feature Selection

Data Integration

Data Formatting

⏪ ⏩

Data Crawling

Data Crawl

Data Preview

No Netflix Account

Scrape Netflix official data from google search result

The Kaggle dataset lacks accurate title names in the user dataset

Missing keywords-title conflict

Use a regex to extract the valid title(as keywords) from the sample data's title

Case :

Staffel 1 (Trailer): Instant Dream Home

Instant Dream Home

Missing data from google search results

Netflix Dataset

Find more info on another site

Data missing from Netflix pages due to lost page

Manual supplemental data

Data Preprocessing

Step	Todo	Method	Condition
Data Cleaning	Handel missing Value	Drop	No valid info ; Duration ≤ 3s
	Detect Outliers	Keep	Representative of Real-world Scenarios; Insight Generation: Preservation of Information
Feature Selection	Feature detection	Keep valid feature	Netflix Data Dictionary
Data Integration	Merged: Netflix Dataset & User Sample Dataset	Get final valid dataset	-
Data Formatting	Date & Duration	Transfer to standard format	-

Final Valid Dataset Overview

Insights Analysis

1st

2nd

3rd

4th

Exercise

Do the basic content-based analysis

Pre-Insights

Deep dive for each feature to get the initial insights

Insights

To mine better insights in multidimensional

Result

High-level

&

Low-level

Insights Flow

INSIGHTS

1. Insights about Favorite TOP N
- Movie — 1st: To All the Boys I've Loved Before
- TV shows — 1st: The Blacklist & Naruto
- Director
- Cast
- Genre — 1st: Drama | 2nd: Anime | 3rd: Comedy
- Types — 1st: TV shows | 2nd: Movie | 3rd: HOOK

2 Metrics of Measurement: Click Count & Viewing Time

2. Insights about Time
- Longest Viewing Time — 02:49:04
- Total Viewing Time — 1864.11 hours
- The Earliest Entry Into Netflix — 2018-04-10 13:35:28
- Viewed the Oldest Series — Das Leben des Brian(1979)

1. High Netflix usage from 2018 to 2021, sharp decline in 2022, possibly due to exploring other platforms or lack of time after 2022.
2. Viewing time fluctuates, decreases every February, increases every March. February possibly busy due to exams.
3. Prefers binge-watching after 1 PM to before 10 PM, except early mornings for Naruto anime.

3. Insights about Frequency Trend
- Time Dimensions
  - Yearly Trend
  - Monthly Trend
  - Hourly Trend

2 Metrics of Measurement: Click Count & Viewing Time

4. Insights about Devices
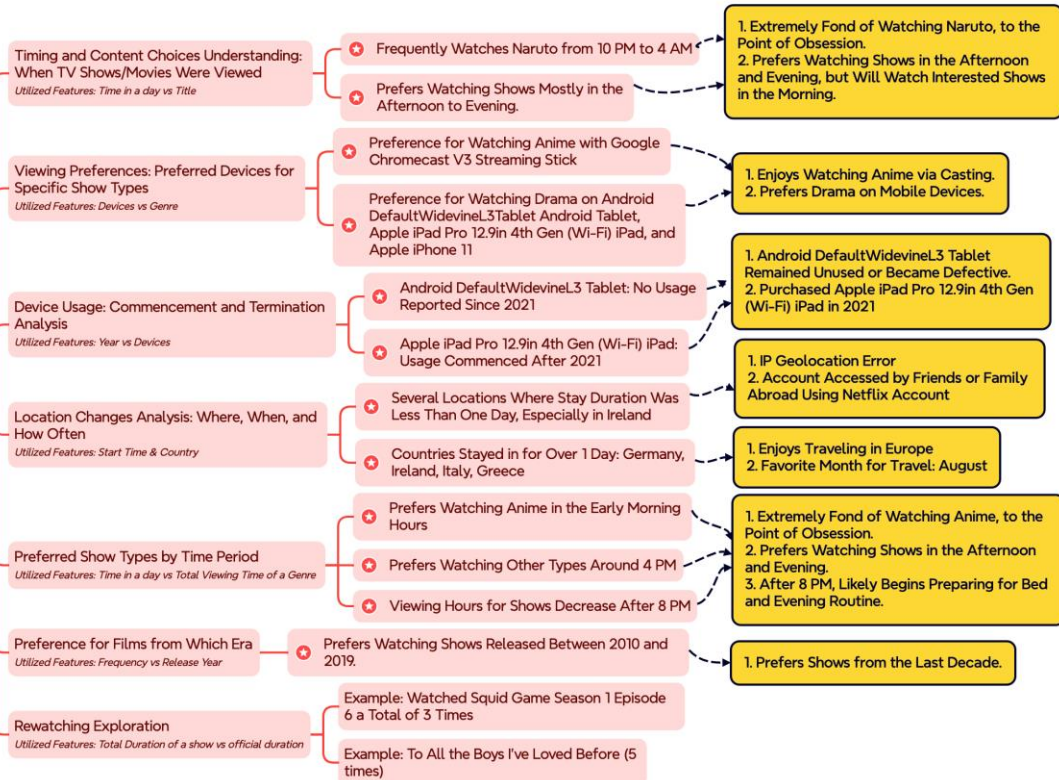- Total Count of Devices Used — 22
- Most Frequently Used Devices
  - Google Chromecast V3 Streaming Stick: Utilized 1363 Times for Streaming
  - Android DefaultWidevineL3Tablet: Utilized for 498.76 Hours on Android Tablet

1. Preference for Apple Devices
2. Rarely Uses TV
3. Prefers Watching Content via Casting or on Mobile Devices

2 Metrics of Measurement: Click Count & Viewing Time

5. Insights about Locations
- Visited Destinations: Exploring Travel Locations — DE (Germany), IE (Ireland), IT (Italy), GR (Greece)
- Travel Periods: Analyzing Timeframes for Vacations — Remained in Germany Since August 22, 2019, Without Further Travel

1. Resident of Germany
2. Frequent traveler within Europe
3. Ceased traveling because of COVID-19

6. Insights about Rating
- Analysis of Preferred Age-Restriction Ratings for Viewed Content — Most Viewed Films Rated for 16+

1. Age over 16
2. Possible younger siblings (watched some children's programs)

7. Insights about Tags
- Preference for Films of Which Moods — Violent, Goofy, Imaginative
- Preference for Films with Which Features — Sci-Fi & Fantasy, Family, Romantic Comedy

1. Personality: Witty, Imaginative, Optimistic, Family-oriented, Romantic, Sentimental, and Bold

Presented with xmind

Insights Flow

Insight

8. Insights about Relationship Among Multiple Attributes

Timing and Content Choices Understanding: When TV Shows/Movies Were Viewed
Utilized Features: Time in a day vs Title

- Frequently Watches Naruto from 10 PM to 4 AM
- Prefers Watching Shows Mostly in the Afternoon to Evening.

> 1. Extremely Fond of Watching Naruto, to the Point of Obsession.
> 2. Prefers Watching Shows in the Afternoon and Evening, but Will Watch Interested Shows in the Morning.

Viewing Preferences: Preferred Devices for Specific Show Types
Utilized Features: Devices vs Genre

- Preference for Watching Anime with Google Chromecast V3 Streaming Stick
- Preference for Watching Drama on Android DefaultWidevineL3Tablet Android Tablet, Apple iPad Pro 12.9in 4th Gen (Wi-Fi) iPad, and Apple iPhone 11

> 1. Enjoys Watching Anime via Casting.
> 2. Prefers Drama on Mobile Devices.

Device Usage: Commencement and Termination Analysis
Utilized Features: Year vs Devices

- Android DefaultWidevineL3 Tablet: No Usage Reported Since 2021
- Apple iPad Pro 12.9in 4th Gen (Wi-Fi) iPad: Usage Commenced After 2021

> 1. Android DefaultWidevineL3 Tablet Remained Unused or Became Defective.
> 2. Purchased Apple iPad Pro 12.9in 4th Gen (Wi-Fi) iPad in 2021

Location Changes Analysis: Where, When, and How Often
Utilized Features: Start Time & Country

- Several Locations Where Stay Duration Was Less Than One Day, Especially in Ireland
- Countries Stayed in for Over 1 Day: Germany, Ireland, Italy, Greece

> 1. IP Geolocation Error
> 2. Account Accessed by Friends or Family Abroad Using Netflix Account

> 1. Enjoys Traveling in Europe
> 2. Favorite Month for Travel: August

Preferred Show Types by Time Period
Utilized Features: Time in a day vs Total Viewing Time of a Genre

- Prefers Watching Anime in the Early Morning Hours
- Prefers Watching Other Types Around 4 PM
- Viewing Hours for Shows Decrease After 8 PM

> 1. Extremely Fond of Watching Anime, to the Point of Obsession.
> 2. Prefers Watching Shows in the Afternoon and Evening.
> 3. After 8 PM, Likely Begins Preparing for Bed and Evening Routine.

Preference for Films from Which Era
Utilized Features: Frequency vs Release Year

- Prefers Watching Shows Released Between 2010 and 2019.

> 1. Prefers Shows from the Last Decade.

Rewatching Exploration
Utilized Features: Total Duration of a show vs official duration

- Example: Watched Squid Game Season 1 Episode 6 a Total of 3 Times
- Example: To All the Boys I've Loved Before (5 times)

Presented with xmind

| **Weaknesses** & **Prospects**

Limited data access

our access only to a subset of Netflix user data, limiting the depth and breadth of my analysis (No account on Netflix);

Lack of context → Expand knowledge

Without a comprehensive understanding of Netflix features, our insights may be less nuanced;

Data quality issues → Improve data quality

Incomplete or inaccurate dataset could have skewed our analysis results

03

DATA VISUALIZTION

TV show & Movie Top 5

TV SHOW RANK

Title - Viewing time

① The Blacklist	135.06h	
② Naruto	135.05h	
③ Orange Is the New Black	87.74h	
④ Prison Break	75.10h	
⑤ Money Heist	64.66h	

Presented with **xmind**

MOVIE RANK

Title - Viewing time - Rewatch Times

① To All the Boys I've Loved Before	9.16h	5.49 times
② Interstellar	5.56h	1.97 times
③ The Kissing Booth	5.31h	3.01 times
④ Shooter	4.76h	1.76times
⑤ Shooter	4.75h	2.29times

Presented with **xmind**

Detection Outlier- TV show-Boxplot
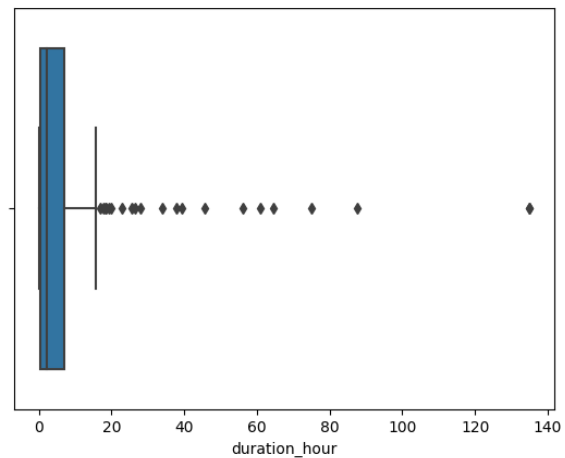
Detection Outlier- Movie-Boxplot
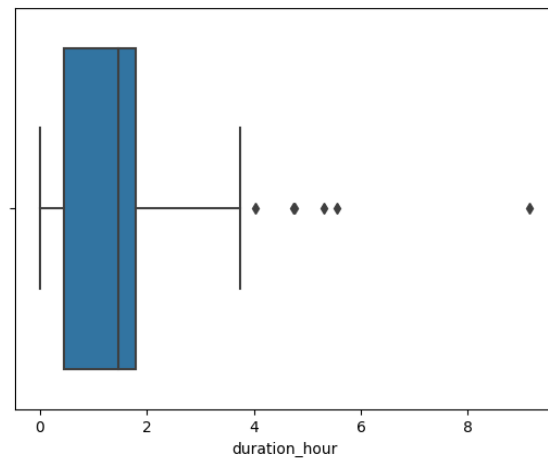
Heatmap: Year vs Device
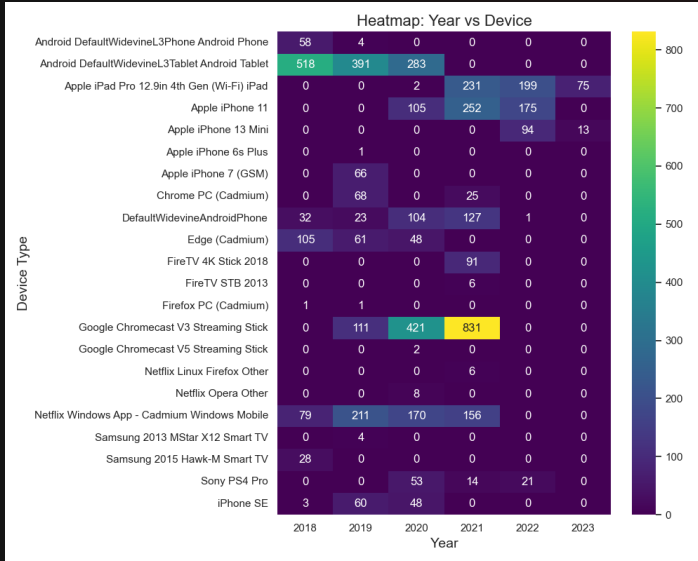
Yearly Click on Netflix

Mood show

Travel Info

Monthly Click Counts by Year (2018-2023)

Counting All Shows by Release Year

Genre Duration on Netflix
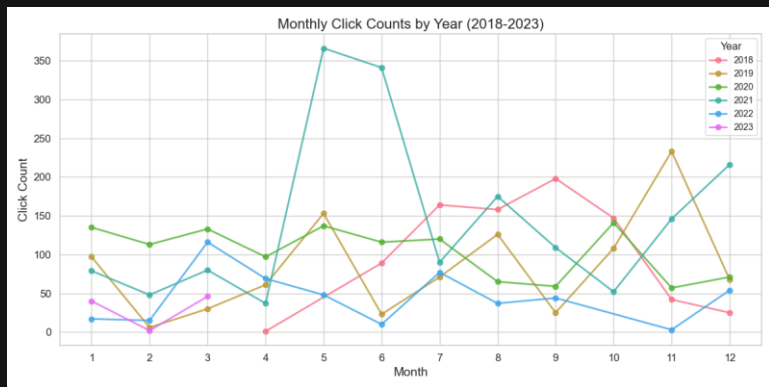
Counting All Shows by Release Year
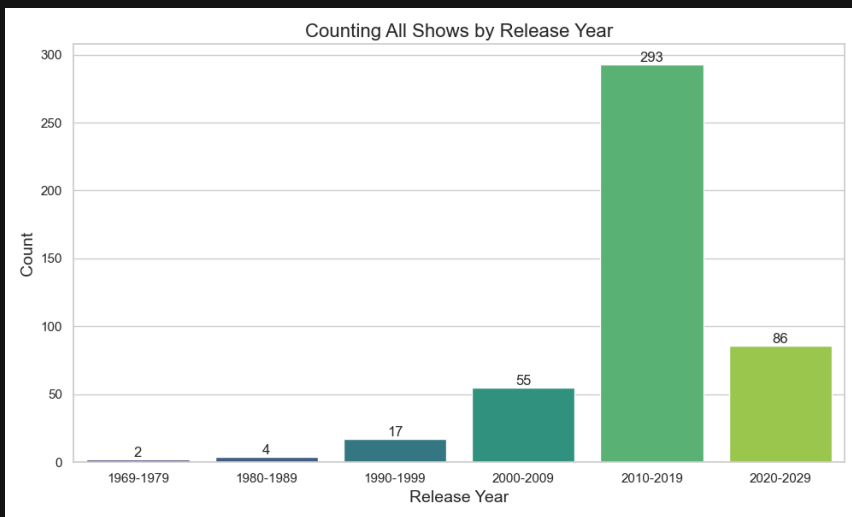
Genre Duration on Netflix
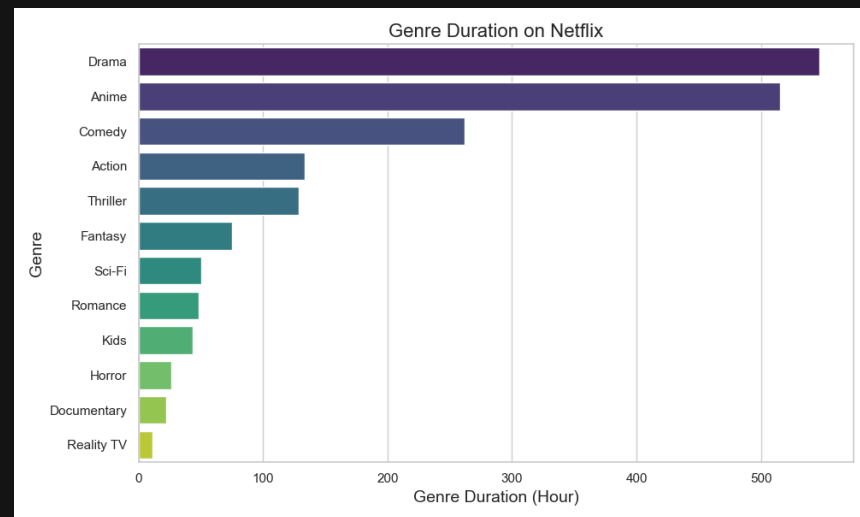
THANKS!

Happy to Questions ☺