**Introduction**

In this project, I researched the Netflix Dataset and how it grew over time. My dataset includes information about the entire movie and TV library from 2008 to 2021 on the streaming service.

The biggest trend I found was that growth was relatively slow until about 2015--I can confidently say that after, Netflix exploded its potential growth--2019 was the highest year of growth with nearly 2000 new titles. In fact, over 75% of the Netflix library was added within the last five years (2017-2021), which makes sense because this is when Netflix was looking to solidify its international presence and access for diversified global audiences while trying to surpass potential competitors in the streaming market.

I also learned that Netflix has more movies than TV shows in its library and that a lot of its content is relatively new. I could visualize some of these findings easily through the various charts I made such as a pie chart that showed the annual dispersion of content.

Overall, this was an insightful experience to see how such a platform grows over time with changes in opportunity for access.

**Libraries Used:**

* Numpy
* Pandas
* Matplotlib
* SeaBorn
* Custom

**Data Cleaning:**

* We have imported data through CSV file of NETFLIX
* Data Cleaning Process

1. Duplicates were identified and removed
2. NULL values were handled using appropriate methods

**Conclusion**

* Problem Solving

This project to enhance decision-making processes by extracting, cleaning and Analysing Netflix Dataset

* Key Findings

1)US Leads in Content Volume  
2) International Content is Growing Rapidly  
3) Dramas and Comedies are Most Popular Genres  
4) Recent Titles Added Faster than Older Ones

* How These Findings Inform Business Decisions

1) Strategic **Content Investment**These insights help Netflix allocate budgets effectively by focusing on popular genres and expanding international content libraries.

2) Targeted **Regional Growth**  
Identifying regions with growing demand guides decisions on producing local content and strengthening presence in emerging markets.

3) Enhanced **User Retention**  
Understanding viewer preferences supports better content recommendations, increasing user engagement and satisfaction.

4) Balanced **Content Acquisition**  
Analysing release patterns helps balance classic title acquisitions with timely additions of new content to remain competitive.

* Limitations and Challenges

1) Data **Limitations**  
The available dataset may not cover all countries, titles, or updated information, which can affect the depth of insights.

2) Dynamic **Content Library**  
Netflix’s catalogue changes frequently, so static data snapshots may quickly become outdated.

3) Genre **Overlaps**Titles often belong to multiple genres, which can complicate clear categorization and analysis.

4) Processing **Large Data**  
Handling and analysing large volumes of data requires strong computational resources and time.

* Acknowledgments

I would like to express my sincere gratitude to my mentors and faculty for their valuable guidance and support throughout this project.  
I also appreciate the use of publicly available Netflix datasets and the open-source tools that made this analysis possible.  
Finally, I thank my peers and reviewers for their constructive feedback and encouragement.