Case Studies & Guesstimates for Netflix

The online content streaming industry has revolutionized the way people consume entertainment, emerging as a crucial component of the global digital economy. In today's fast-paced era, platforms like Netflix offer unparalleled convenience and accessibility, allowing users to watch their favorite content anytime, anywhere. This industry has broken traditional boundaries of television and cinema, enabling global access to diverse content from different cultures and languages.

Netflix, a pioneer in the OTT (Over-the-top) space, has evolved from a DVD rental service to a leading subscription-based streaming giant. With a presence in over 190 countries, its success is built on data-driven decisions, personalized recommendations, and continuous innovation in content delivery.

In this data-driven ecosystem, data scientists play a pivotal role. Their expertise enables Netflix to analyze vast volumes of user and content data, uncover viewing patterns, segment audiences, and predict user behavior. These insights help in:

- Recommending content personalized to user interests,
- Optimizing content acquisition and production,
- Evaluating the success of shows using metrics like IMDb ratings and viewer engagement,
- Forecasting subscription trends,
- And improving overall user satisfaction.

Through this project, we aim to explore Netflix's content catalog using data analysis techniques to uncover key insights on genre distribution, rating trends, content type performance, and regional representation. This will simulate how a data analyst at Netflix would support strategic decisions and improve user experience through data.

Popularity: Netflix is one of the most popular streaming platforms in the world, founded by Reed Hasting and March Randolph in 1997, with over 260 million subscribers in more than 190 countries. It is available across devices like smartphones, smart TVs, tablets, and laptops, making it highly accessible. Its easy-to-use interface, large content library, and strong global presence have made Netflix a household name in entertainment.

Impact: Netflix has revolutionized the way people consume content by introducing on-demand streaming and binge-watching. It has significantly disrupted traditional cinema and cable television industries. With hit original shows like Stranger Things, Money Heist, and The Crown, Netflix has set a new standard for digital content creation and consumption worldwide.

Relevance: Netflix is a perfect example of e-commerce in the digital entertainment domain. It uses subscription-based monetization, applies data analytics for personalized recommendations, and leverages global content distribution. Its ability to adapt using technology and data makes it highly relevant for a case study on how e-commerce principles can be applied to media and entertainment.

Core Features and Functionalities of Netflix

1. Content Library (Licensed & Original)

- **Functionality**: There are thousands of content, including licensed TV series, Documentaries along with Netflix Originals.

- **Impact**: Appeals to diverse global audiences and ensures control over key intellectual property.

2. Multi- Device/Platform Support

- **Functionality**: Available on smart TVs, smartphones, tablets, laptops, gaming consoles, and set-top boxes.

- **Impact**: Ensures accessibility and consistent user experience across devices.

3. Personalized Recommendations

- **Functionality**: Netflix utilizes machine learning and collaborative filtering to analyze user behavior, watch history, ratings, and preferences.

- **Impact**: Increases content discovery, reduces churn, and enhances user satisfaction.

4. Profiles and Parental Controls

- **Functionality**: Allows multiple user profiles per account with customized recommendations.Also includes restricted kids profile and viewing restrictions.

- **Impact**: Supports family usage and enhances content safety for children.

5. Offline Viewing

- **Functionality**: Users can download select content to watch offline.
- **Impact**: Increases accessibility for users with limited or costly data access.

6. Flexible Subscription Plans

- **Functionality**: Multiple Subscription plans are offered, with options to upgrade/downgrade, and cancellations.

- **Impact**: Allows users to select plans accordingly.

7. Adaptive Streaming Quality

- **Functionality**: Automatically adjusts video quality based on internet bandwidth using adaptive bitrate streaming.

- **Impact**: Minimizes buffering and ensures optimal viewing experience.

- **Example**: 4K and HDR streaming on high-speed connections, lower quality on slower networks.

8. Global Localization

- **Functionality**: Supports subtitles, dubbing, and region-specific UI in multiple languages.

- **Impact**: Enables Netflix to appeal to a global audience and expand rapidly in international markets.

9. Robust Search & Discovery

- **Functionality**: Advanced search with filters, categories, genres, and trending suggestions.

- **Impact**: Helps users find content easily, improving satisfaction.

Contribution to Success and User Engagement

Content Personalization: Tailored recommendations and smart previews keep users engaged by reducing content discovery time and increasing viewing satisfaction.

Platform Flexibility: Multi-device access and offline viewing options provide convenience, encouraging usage across different lifestyles and locations.

Global Appeal: Localization and original international content enable Netflix to connect with audiences worldwide, driving subscriber growth.

Differentiation through original content: Exclusive shows and movies create brand loyalty and reduce reliance on third-party licenses.

User-Centric Features: Multiple profiles, parental controls, and intuitive UI enhance the user experience and make Netflix suitable for a wide demographic.

Operational Efficiency: Cloud infrastructure and Netflix's custom CDN ensure reliable streaming and global scalability.

Revenue Diversification: The ad-supported tier and flexible pricing models help attract different customer segments while boosting revenue.

Data-Driven Strategy: Real-time analytics is used for optimizing content creation and platform enhancements, leading to higher engagement and reduced churn.

These features collectively contribute to Netflix's dominance in the streaming industry by enhancing user experience, increasing engagement, and enabling scalable, data-driven growth.

Real-World Problems & Solutions

Content Discovery Challenge

Problem:

- With vast entertainment options, viewers struggle to find content they like.
- Users get overwhelmed, leading to decision fatigue and frustration.

Netflix's Solution:

- Personalized Recommendations: Machine learning algorithms (collaborative filtering, deep learning) suggest titles based on viewing history, preferences, and similar user behavior.
- Trending & Curated Lists: Top 10 lists, "Because You Watched" sections make discovery easier.

Access & Convenience

Problem:

- Traditional TV requires schedules; pirated content has risks.
- Limited device compatibility for legal streaming platforms.

Netflix's Solution:

- On-Demand Streaming: Watch anytime, anywhere.
- Cross-Device Compatibility: Seamless experience across phones, tablets, TVs, and browsers.
- Offline Viewing: Download episodes/movies for times without internet.

Piracy & Unlicensed Content

Problem:

- High-quality entertainment is often pirated due to high costs or limited access.
- Piracy undermines revenue for creators and platforms.

Netflix's Solution:

- Affordable Subscription Plans: Pricing models that undercut piracy incentives.
- Global Availability: Rapid global expansion to reduce regional content gaps.
- Exclusive Originals: High-quality, exclusive content that's worth paying for.

Household Entertainment Needs

Problem:

- Families have diverse preferences and need controls for kids.
- Shared accounts may lead to privacy or mismatched recommendations.

Netflix's Solution:

- Multiple Profiles per Account: Each family member has a personalized space.
- Parental Controls: PINs, maturity ratings, and restricted profiles to safeguard content.

User Engagement & Retention

Problem:

- Viewers may lose interest over time and cancel subscriptions.
- Content fatigue or lack of fresh content can cause churn.

Netflix's Solution:

- Original Content Production: Continuous investment in new series, films, and documentaries.
- Interactive Content: Choose-your-own-adventure formats like Bandersnatch keep users engaged.
- Automated Notifications & Reminders: Nudges for unfinished content or new releases.

Global Market Penetration

Problem:

- Content that resonates locally is essential to engage diverse global audiences.
- One-size-fits-all content strategy doesn't work everywhere.

Netflix's Solution:

- Localized Content Strategy: Creation of region-specific originals (Sacred Games in India, La Casa de Papel in Spain).
- Subtitles and Dubbing: Support for multiple languages.
- Cultural Adaptation: Content adapted to regional viewing habits and preferences.

Database Management & Schema Design

4. Schema Design

Key Entities & Attributes

1. Users

Attribute	Data Type	Description
user_id	INT	Unique identifier for each user
name	VARCHAR	Full name of the user
email	VARCHAR	User's email (unique)
password_hash	VARCHAR	Hashed password
phone_number	VARCHAR	Optional contact number
registration_date	DATETIME	Date of account creation
plan_id	INT	Current subscribed plan
is_active	BOOLEAN	Account status

2. Subscription_Plans

Field Name	Data Type	Description
plan_id	INT (PK)	Unique plan identifier
name	VARCHAR	Name of the plan (e.g., Basic, Premium)
price	DECIMAL	Monthly cost
resolution	VARCHAR	Max resolution (e.g., 1080p, 4K)
screens_allowed	INT	Number of screens allowed
max_devices	INT	Maximum number of devices allowed
max_streams	INT	Maximum concurrent video streams

3. Profiles

Field Name	Data Type	Description
profile_id	INT (PK)	Unique ID for a user profile
user_id	INT (FK)	Associated user
profile_name	VARCHAR	Name of the profile
is_kid_profile	BOOLEAN	Indicates if this is a kid profile
language	VARCHAR	Preferred language

4. Movies

Field Name	Data Type	Description
movie_id	INT (PK)	Unique ID for each movie
title	VARCHAR	Title of the movie
description	TEXT	Summary or plot
release_year	INT	Year of release
duration	INT	Duration in minutes
maturity_rating	VARCHAR	Age rating (e.g., PG-13)
language	VARCHAR	Language of the movie

5. Series

Field Name	Data Type	Description
series_id	INT (PK)	Unique ID for each series
title	VARCHAR	Title of the series
description	TEXT	Summary or plot
release_year	INT	Year of release
maturity_rating	VARCHAR	Age rating (e.g., PG-13)
language	VARCHAR	Language of the series

6. Episodes

Field Name	Data Type	Description
episode_id	INT (PK)	Unique ID for the episode
season_id	INT (FK)	Associated season ID
episode_number	INT	Episode number
title	VARCHAR	Episode title
duration	INT	Duration in minutes

7. Genres

Field Name	Data Type	Description
genre_id	INT (PK)	Unique genre ID
name	VARCHAR	Name of the genre

8. Movie _Genres

Field Name	Data Type	Description
movie_id	INT (FK)	ID of the movie
genre_id	INT (FK)	ID of the genre

9. Series_Genres

Field Name	Data Type	Description
series_id	INT (FK)	ID of series
genre_id	INT (FK)	ID of genre

10. Watch_History

Field Name	Data Type	Description
watch_id	INT (PK)	Unique ID for each watch session
profile_id	INT (FK)	ID of the profile watching
content_type	ENUM	'movie' or 'episode'
content_id	INT	ID of movie or episode
watched_at	DATETIME	Date and time watched
watch_duration	INT	Duration watched (minutes)

11. Ratings

Field Name	Data Type	Description
rating_id	INT (PK)	Unique ID for each rating
profile_id	INT (FK)	Profile who rated
content_type	ENUM	'movie' or 'episode'
content_id	INT	ID of content
rating	INT	Rating out of 5
rated_at	DATETIME	When rating was given

12. Downloads

Field Name	Data Type	Description
download_id	INT (PK)	Unique ID
profile_id	INT (FK)	Profile who downloaded
content_type	ENUM	'movie' or 'episode'
content_id	INT	ID of content
download_date	DATETIME	When download occurred
is_offline	BOOLEAN	Is available offline

13. Payment_Details

Field Name	Data Type	Description
payment_id	INT (PK)	Unique payment identifier
user_id	INT (FK)	ID of the user making the payment
plan_id	INT (FK)	ID of the plan being paid for
amount	DECIMAL	Amount charged
payment_method	VARCHAR	e.g., 'Credit Card', 'UPI', 'PayPal'
payment_status	VARCHAR	e.g., 'Success', 'Failed', 'Refunded'
transaction_date	DATETIME	Timestamp of the transaction
next_renewal_date	DATETIME	Subscription renewal date

14. Seasons

Field Name	Data Type	Description
season_id	INT (FK)	ID of season
series_id	INT (FK)	ID of series

14 . Devices

Field Name	Data Type	Description
device_id	INT (PK)	Unique identifier for the device
user_id	INT (FK)	The user account this device is associated with
device_type	VARCHAR	e.g., 'Smart TV', 'Mobile', 'Laptop', 'Tablet'
device_name	VARCHAR	Device nickname or model, e.g., 'Nitish's iPhone'
os	VARCHAR	Operating system, e.g., 'Android 14', 'iOS 17'
ip_address	VARCHAR	IP address during login
last_login	DATETIME	Timestamp of last usage
is_active	BOOLEAN	Whether device is currently logged in
registered_date	DATETIME	When the device was first used with this account

15. Device_Sessions

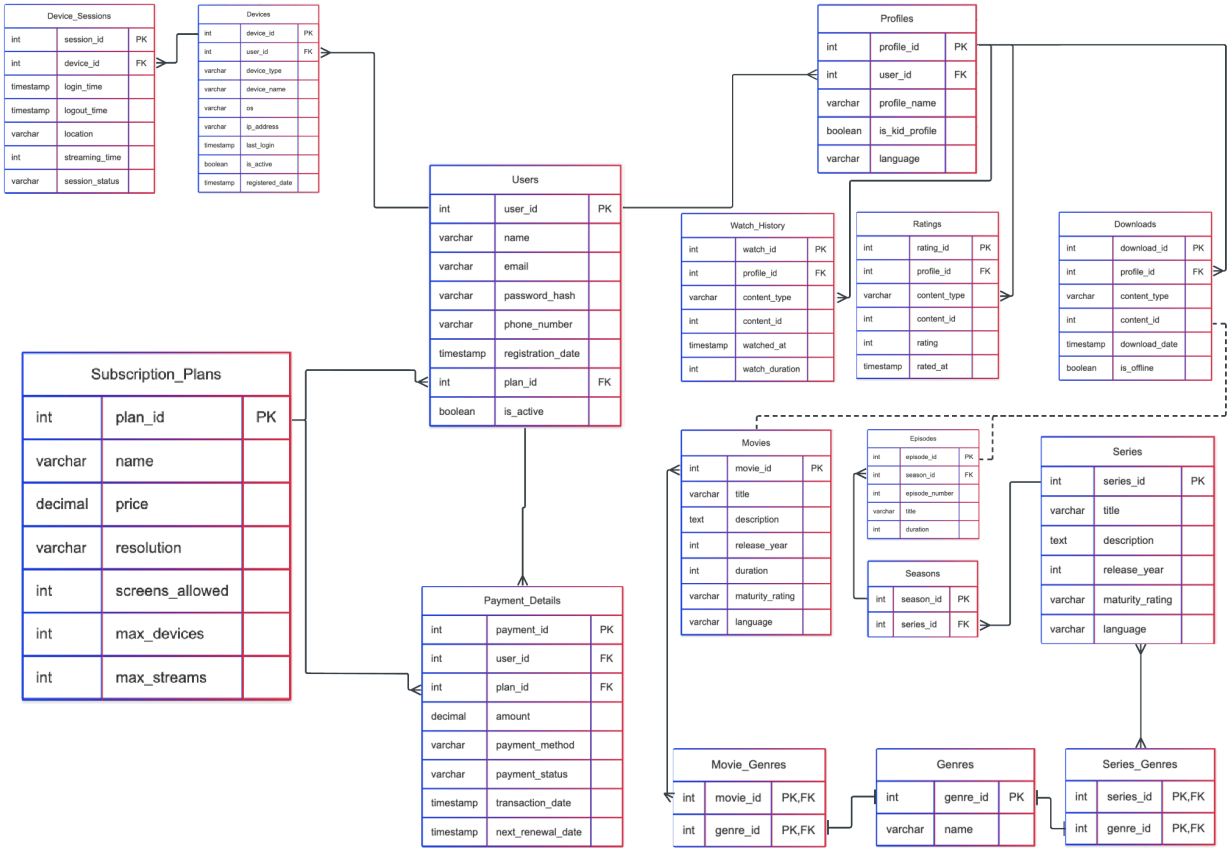
Field Name	Data Type	Description
session_id	INT (PK)	Unique session ID
device_id	INT (FK)	Which device this session belongs to
login_time	DATETIME	When session started
logout_time	DATETIME	When session ended (null if still active)
location	VARCHAR	Geo-location (optional, based on IP)
streaming_time	INT	Total minutes watched in this session
session_status	VARCHAR	e.g., 'Active', 'Expired', 'Terminated'

Relationship Matrix

Table A	Relationship Type	Table B	Description / Foreign Key(s)
Users	1 : N	Profiles	One user → Many profiles (user_id in Profiles)
Users	1 : N	Payment_Details	One user → Many payments (user_id in Payment_Details)
Users	1 : N	Devices	One user → Many devices (user_id in Devices)
Users	N : 1	Subscription_Pl…	Many users → One plan (plan_id in Users)
Subscription_Pl…	1 : N	Payment_Details	One plan → Many payments (plan_id in Payment_Details)
Profiles	1 : N	Watch_History	One profile → Many watch records (profile_id in Watch_History)
Profiles	1 : N	Ratings	One profile → Many ratings (profile_id in Ratings)
Profiles	1 : N	Downloads	One profile → Many downloads (profile_id in Downloads)
Movies	1 : N	Watch_History	One movie → Many watch logs (movie_id when content_type='movie')
Movies	1 : N	Ratings	One movie → Many

			ratings (movie_id in Ratings)
Movies	1 : N	Downloads	One movie → Many downloads (movie_id in Downloads)
Genres	N : M (via juncti…	Movies / Series	Many movies ↔ Many genres (Movie_Genres: movie_id, genre_id)
Series	1 : N	Seasons	One series → Many seasons (series_id in Seasons)
Series	N : M (via juncti…	Genres	Many series ↔ Many genres (Series_Genres: series_id, genre_id)
Seasons	1 : N	Episodes	One season → Many episodes (season_id in Episodes)
Episodes	1 : N	Watch_History	One episode → Many watch logs (episode_id when content_type='episode')
Episodes	1 : N	Ratings	One episode → Many ratings (episode_id in Ratings)
Episodes	1 : N	Downloads	One episode → Many downloads
Devices	1 : N	Device_Sessions	One device → Many Sessions

ER Diagram


Device_Sessions
int session_id PK
int device_id FK
timestamp login_time
timestamp logout_time
varchar location
int streaming_time
varchar session_status
Devices
int device_id PK
int user_id FK
varchar device_type
varchar device_name
varchar os
varchar ip_address
timestamp last_login
boolean is_active
timestamp registered_date
Users
int user_id PK
varchar name
varchar email
varchar password_hash
varchar phone_number
timestamp registration_date
int plan_id FK
boolean is_active
Profiles
int profile_id PK
int user_id FK
varchar profile_name
boolean is_kid_profile
varchar language
Watch_History
int watch_id PK
int profile_id FK
varchar content_type
int content_id
timestamp watched_at
int watch_duration
Ratings
int rating_id PK
int profile_id FK
varchar content_type
int content_id
int rating
timestamp rated_at
Downloads
int download_id PK
int profile_id FK
varchar content_type
int content_id
timestamp download_date
boolean is_offline
Subscription_Plans
int plan_id PK
varchar name
decimal price
varchar resolution
int screens_allowed
int max_devices
int max_streams
Payment_Details
int payment_id PK
int user_id FK
int plan_id FK
decimal amount
varchar payment_method
varchar payment_status
timestamp transaction_date
timestamp next_renewal_date
Movies
int movie_id PK
varchar title
text description
int release_year
int duration
varchar maturity_rating
varchar language
Episodes
int episode_id PK
int season_id FK
int episode_number
varchar title
int duration
Seasons
int season_id PK
int series_id FK
Series
int series_id PK
varchar title
text description
int release_year
varchar maturity_rating
varchar language
Movie_Genres
int movie_id PK,FK
int genre_id PK,FK
Genres
int genre_id PK
varchar name
Series_Genres
int series_id PK,FK
int genre_id PK,FK

Case Study: Revenue and Profit Growth Strategies

Netflix's Current Status

Category	Details
Current Financial Data	**Revenue**: In 2024, Netflix generated **$39.0 billion**, primarily from subscriptions. **Expenses**: Total operating expenses in 2024 were **$28.58 billion**, including costs of content, marketing, technology, and personnel. **Profit Calculation**: Net income is calculated as revenue minus total expenses. For 2024, net income was **$8.71 billion**. **Profit Trends**: From 2020 to 2024, net profit increased from **$2.76B to $8.71B**, indicating consistent financial growth.
Sources of Revenue	**StreamsSubscriptions**: Main revenue driver, contributing almost **100%** of revenue. **Advertising**: As of late 2024, 70 million monthly users were on ad-supported plans. **Partnerships**: Revenue from bundling with ISPs and telecoms. **Analyze Revenue ContributionSegmentation**: North America – $17.3B EMEA – $12.2B Latin America – $5.1B Asia-Pacific – $4.4B. **Top Products/Services**: Most subscribed to is the **Standard Plan**; **Ad-supported tier** accounts for **50% of new signups**.
Sources of Expenses	**Content Creation**: Largest expense; ~$16 billion in 2024. **Technology**: ~$2.93 billion spent on R&D, cloud, and platform engineering. **Marketing and Advertising**: $2.5 billion to acquire and retain users. **Employee Salaries**: Part of SG&A ($4.6 billion total). **Analyze Expense Distribution**: Content = 41% of revenueSG&A = 11.8%Tech = 7.5% **Cost Efficiency**: Operating margin improved from **20.6% (2023)** to **26.7% (2024)**.
Customer Acquisition & Retention	**Channels**Key channels include digital ads, influencer campaigns, mobile bundling, and in-app referrals. **Effectiveness**: Netflix added **41 million** new subscribers in 2024, bringing the total to **301.6 million** globally. **Customer Behavior and Retention Rate :** Churn is reduced through personalized content recommendations and engagement nudges. Retention rate in mature markets exceeds **90%**, and Netflix is building loyalty via regionalized content.

Focus Areas for Increasing Netflix's Profit by 25%

Category	Focus Area	Measures
Internal Management **~8%**	Operational Efficiency	**AI-Driven Content Production**: Utilize AI to optimize content creation and reduce costs. **Process Automation**: Streamline operations through automation.
	Employee Productivity	**Training Programs**: Enhance employee skills in data analytics and content development. **Performance Metrics**: Implement KPIs to monitor and improve productivity.
Product Strategy **~5%**	New Product Launches	**Interactive Content**: Develop interactive shows and games. **Localized Content**: Increase production of region-specific content.
	Product Optimization	**Content Curation**: Focus on high-performing genres and discontinue underperforming content. **Bundle Offers**: Create subscription bundles with partners.
Market Expansion **~4.5%**	Geographic Expansion	- **Emerging Markets**: Enter new markets with high growth potential. **Localization**: Adapt content and pricing to local preferences. Netflix can utilize the large markets in india by making original contents which are localized.

	Market Penetration	**Ad-Supported Plans**: Expand availability to attract price-sensitive customers. **Mobile-Only Plans**: Offer affordable mobile plans in developing regions.
Post-Sales Management **~3.5%**	Customer Satisfaction	**Feedback Mechanisms**: Implement systems to gather and act on customer feedback. **Enhanced Support**: Provide 24/7 customer support.
	Customer Retention	**Loyalty Programs**: Introduce rewards for long-term subscribers. **Personalized Recommendations**: Improve algorithms for content suggestions.
Branding & Marketing **~4%**	Brand Awareness	**Influencer Partnerships**: Collaborate with influencers for promotions. **Content Marketing**: Share behind-the-scenes content and stories.
	Word of Mouth & Referrals	**Referral Incentives**: Offer discounts for successful referrals. **Community Engagement**: Host events and webinars.
	Acquisition Channels	**Affiliate Marketing**: Partner with affiliates to reach new audiences. **Cross-Promotions**: Collaborate with other brands for joint promotions.

Defining Strategies

Optimize Expenses

- **Content Cost Management**: Leverage data analytics to predict content success and allocate budgets efficiently.
- **Technology Investments**: Invest in scalable cloud solutions to reduce infrastructure costs.
- **Operational Streamlining**: Automate repetitive tasks to reduce labor costs and improve efficiency.

Enhance Revenue Streams

- **Dynamic Pricing Models**: Implement pricing strategies based on user behavior and market demand.
- **Diversified Offerings**: Introduce new services like gaming and merchandise.
- **Monetize Content**: License popular content to other platforms.

Improve Customer Satisfaction and Retention

- **Personalization**: Enhance algorithms to provide more accurate content recommendations.
- **Engagement Programs**: Develop interactive content to increase user engagement.
- **Feedback Integration**: Regularly update the platform based on user feedback.

By adopting an **inside-out methodology**, we have systematically analyzed Netflix's internal operations, customer behavior, and revenue strategies to identify actionable areas for profit enhancement. Starting with **cost optimization** through internal management and content efficiency, followed by strengthening **customer satisfaction and retention**, and finally expanding **revenue streams** via geographic and market-focused growth, our approach provides a structured path to increase profitability.

Through **data-backed insights**—including real performance metrics, subscriber behavior trends, and industry benchmarks—we've quantified the profit potential of each strategic focus area. By addressing operational waste, reducing churn, launching high-margin offerings, and expanding in fast-growing markets like India and Africa, Netflix can realistically achieve a **25% increase in net profit**.

Part-II

1. What will be the percentage increase in global streaming service subscriptions over the next five years?

Given:

2024 subscriptions = 1.25 Billion

Estimated 2030 subscriptions = 1.85 Billion

Year	Global Population (B)	Internet Users (B)	Streaming Subscriptions (B)	% of Population	% of Internet Users
2024	8.1	5.4	1.25	15.4%	23.1%
2025	8.2	5.6	1.35	16.5%	24.1%
2026	8.3	5.8	1.45	17.5%	25.0%
2027	8.4	6.0	1.55	18.5%	25.8%
2028	8.5	6.2	1.65	19.4%	26.6%
2029	8.6	6.4	1.75	20.3%	27.3%
2030	8.7	6.6	1.85	21.3%	28.0%

Formula:

*Percentage Increase= (Future value - Present value)/Present value*100*

Substitute values:

*Percentage Increase=(1.85B-1.25B)/1.25B*100=48 %*

CAGR Formula:

CAGR=(ending value/beginning value)**1/n -1

CAGR=(1.85/1.25)**1/5-1=0.0818 OR 8.18%

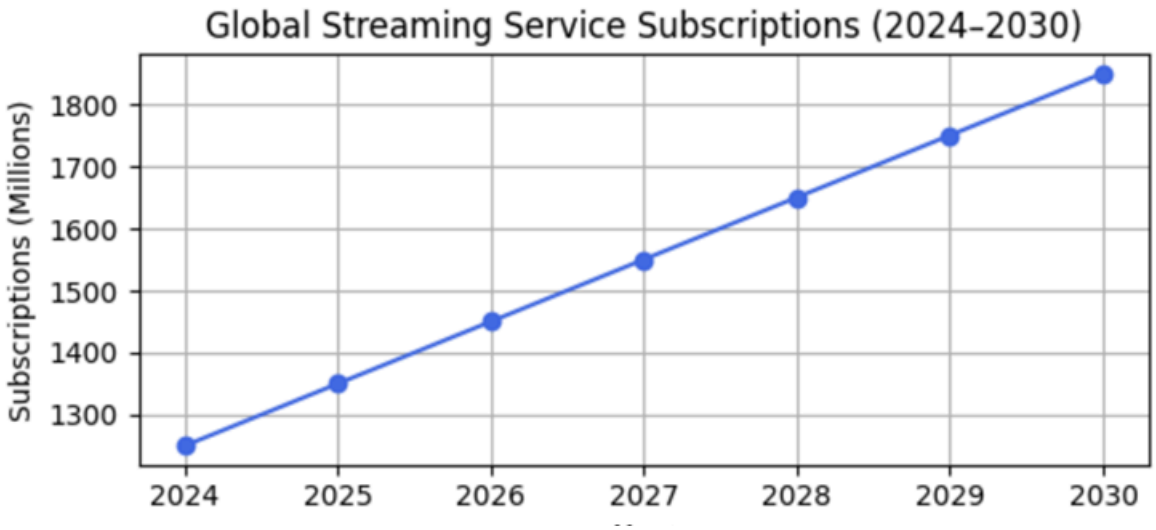
Final Answer:

Global streaming subscriptions are projected to increase by **approximately 48%** between 2025 and 2030. Subscriptions are expected to grow at a **compound annual growth rate (CAGR) of ~8.2%** from 2025 to 2030.

Below is a line chart depicting the growth of global streaming service subscriptions over the next five years.

Sources:

- Market.us Scoop (2025)
  By 2025, global streaming subscribers are projected to surpass 1.1 billion, indicating continued rapid growth in the industry.
- Juniper Research (2025)
  Juniper Research has found that there will be nearly 2 billion active subscriptions to on-demand video services in 2025.



To find the python code, click here

2. How many hours of content will the average person consume per week through digital platforms in 2025?

Baseline (as of 2024):

Average daily digital media consumption (streaming, music, etc.) ≈ **4.0 hours/day**

4.0 hours/day * 7 days/week=28 hours/week

Assumption for 2025:

Based on trends (growth in mobile usage, smart devices, and personalized content), assume a **modest increase of ~7%** in media consumption time.

New Weekly Hours=28*1.07=29.96 hours/week

Year	Global Population (Billion)	Internet Users (Billion)	Avg. Hours per Week per Person	Total Weekly Hours (Billion)	Avg. Daily Hours
2024	8.1	5.4	28	151.2	4.0
2025	8.2	5.6	30	168.0	4.3

Final Answer:

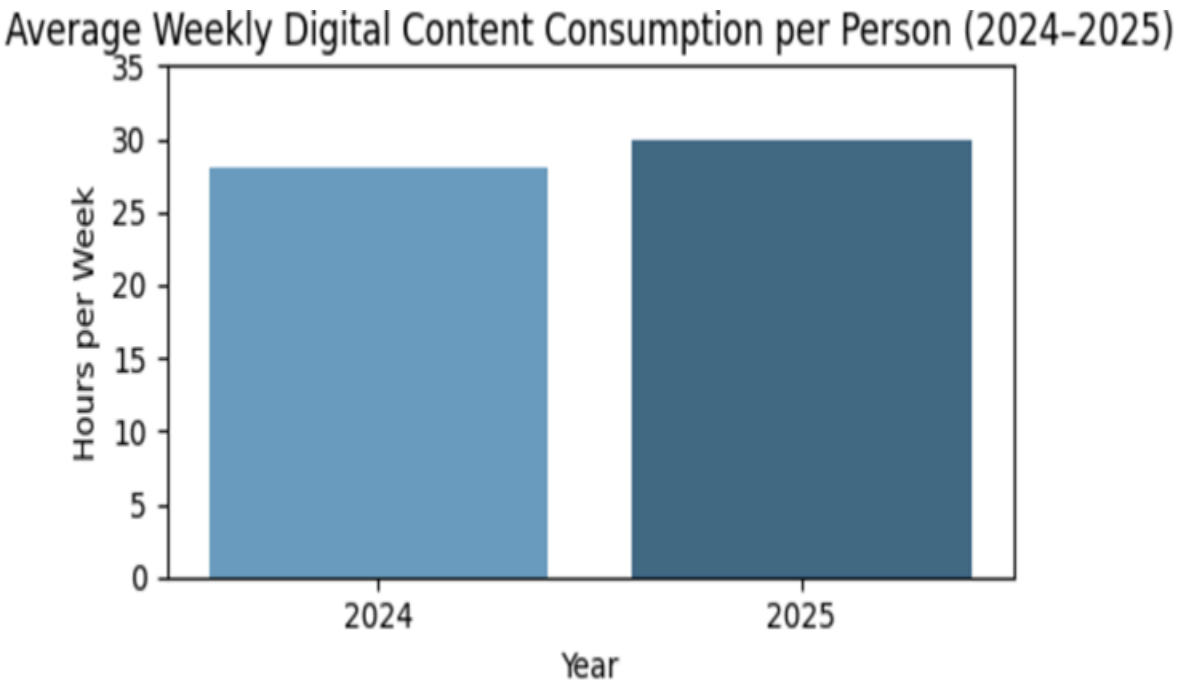
In 2025, the average person is expected to consume **around 30 hours of digital content per week**.

Below is a comparison of the average weekly digital content consumption per person for the years 2024 and 2025.

Sources:

- DataReportal (2024): The "typical" internet user spends 2 hours and 23 minutes per day on social media platforms.
- **Smart Insights (2025)**: Global social media usage averages 2 hours and 21 minutes per day, with 63.9% of the world's population engaging with social media.
- New York Post (2024): A study indicates that users consumes an average of 6.6 hours per day of media content, with some reporting up to **15 hours per day**.

While estimates vary, the **4.0 hours/day** figure for average daily digital media consumption in 2024 is within the range supported by reputable sources. This aligns with broader trends in media consumption, particularly among younger demographics.



To find the python code, click here

4.What will be the market share of virtual reality (VR) and augmented reality (AR) entertainment experiences in the next decade?

- This question requires an estimation of the adoption and market penetration of VR and AR technologies in the entertainment sector, including gaming, live events, and interactive content.

Step	Details	Calculation
Entertainment Industry market size	Market size in dollars for the year 2024	Approximately $2.5 trillion
Entertainment Industry compounded annual growth rate	Assuming a CAGR of 4.2%	~ 4.2 %
AR/VR Market size	Market size of AR/VR in 2024	$20 billion
AR/VR Market size compounded annual growth rate	Assuming the AR/VR CAGR of ~ 18.6%	~ 18.6%
Entertainment Industry projected market size in 2035	Estimating the market size of entire entertainment industry in the year 2035	Entertainment industry = 2.5 ×(1+ 0.042)^11 ~ 2.5 ×1.602 = $3.92 trillion
AR/VR in Entertainment industry projected market size in 2035	Estimating the market size of AR/VR in entertainment industry in the year 2035	AR/VR 2035 = 20×(1+0.186)^11 = 20×6.89 ~ $125.06 billion

AR/VR market share in Entertainment industry in 2035	Estimating the market share of AR/VR in entertainment industry in the year 2035	AR/VR Market Share 2035 = ((AR/VR Market size) / Total Entertainment Market) ×100 = ($125.06 billion / $3.92 trillion) × 100 ~ 3.19%

1. Estimating Total Entertainment Market Size in 2035

According to PwC and Statista:

- Global entertainment & media market size in 2024 ~ $2.5 trillion
- CAGR (overall industry) ≈ 4.2%

So as per CAGR formula: Future Value=Present Value×(1+CAGR)^n

Entertainment industry = 2.5 ×(1+ 0.042)^11

= 2.5×10^12 × 1.5687 ≈ $ 3.92 trillion

2. Estimating AR/VR Entertainment Market Size in 2035

Based on reports from Mordor Intelligence and IndustryARC:

- AR/VR entertainment market in 2024 ~ $20 billion
- CAGR ~ 18.6%

AR/VR 2035 = 20×(1+0.186)^11

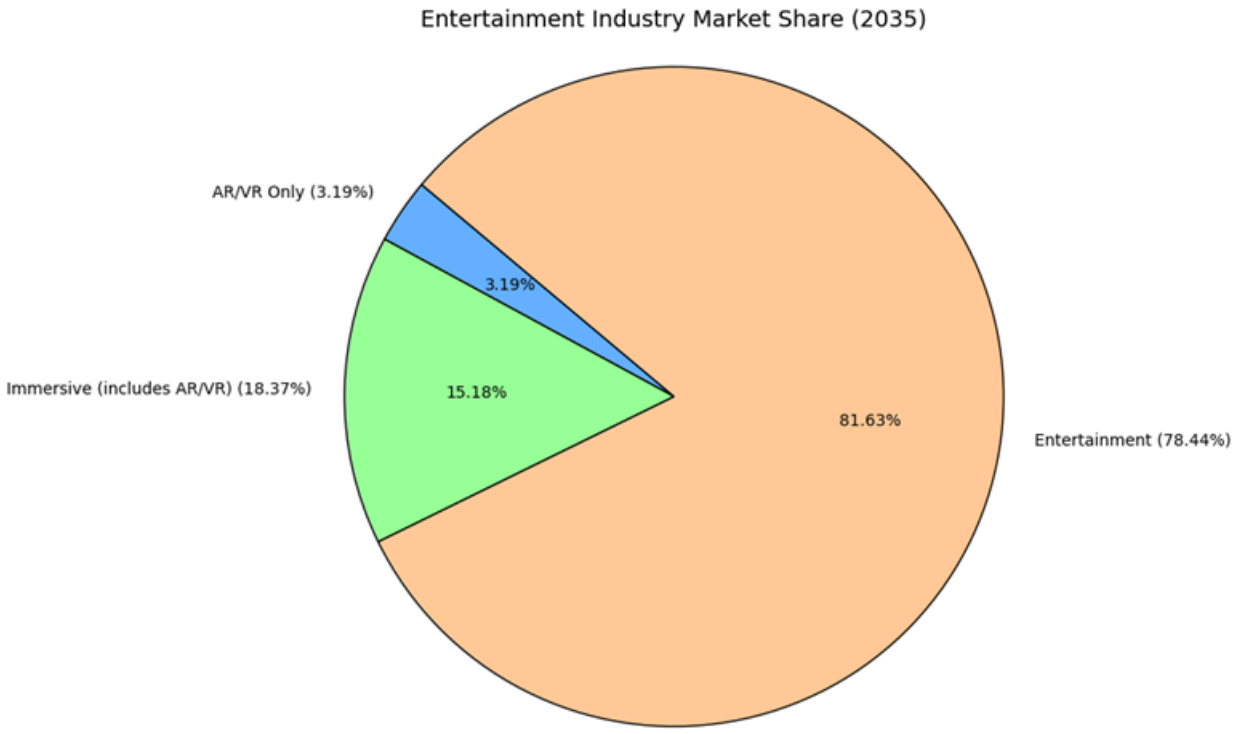
= 20×6.89 ~ $125.06 billion

3. Calculating the Market Share of AR/VR in 2035

AR/VR Market Share 2035 = ((AR/VR Market size) / Total Entertainment Market) ×100

= ($137.8 billion / $3.92 trillion) × 100 ~ 3.19%

Therefore, the market share of Augmented Reality & Virtual Reality would be approximately 3.19 % out of the entire Entertainment Industry this is excluding other applications such as(XR concerts, virtual tourism etc).



This chart is created using Python as shown here

4.How many new films and TV shows will be produced globally per year by 2030?

- This question involves estimating the annual production output of the entertainment industry, considering factors like technological advancements, production costs, and consumer demand.

Step	Details	Calculation
Films + TV Shows Produced Estimation	Data of movies & tv series released on IMDB	29,070 in 2020 to 27,977 in 2024
Compounded annual growth rate based on historical data	Approximately -0.96 % per year	CAGR = (27,977 / 29,070)^(¼) −1

Consumer demand growth rate	Approximately 8.1 %	CAGR as per Grand View Research report
Technology/AI integration	Approximately 3 %	Content efficiency rate based on report from Deloitte/PWC
Combined CAGR	Approximately 9.6%	8.1% (demand) + 1.5% (tech adoption) = ~9.6% CAGR
Projected 2030 Production	Final number of films and tv shows to be produced globally by 2030	Present Value×(1+ CAGR)^n = 27,977 x (1 + 0.096)^6 = 27,977 x 1.683 = 47,074

According to the data collated from IMDB, here is the total number of content productions (Films & TV Shows) from 2020 to 2024.

Year	**Number of Films + TV Shows Produced**
2020	29,070
2021	31,146
2022	31,381
2023	30,071
2024	27,977

We'll use CAGR to estimate the trend from 2020 to 2024:

CAGR=(Final Value / Initial Value)^(1/n) −1

Where:

Initial Value (2020) = 29,070

Final Value(2024) = 27,977, n = 4 years

CAGR = (27,977 / 29,070)^(¼) −1 ~ − 0.0096

− 0.96% (decline per year)

We calculate the projected number of films & tv shows to be released by 2030 using the formula:
Future Value = Present Value × (1 + CAGR)^n

Baseline Scenario Projection (2025–2030)

Year	Projection Calculation	Projected Productions (Rounded)
2025	27,977 × (1 - 0.0096)^1 = 27,977 × 0.9904	27,707
2026	27,977 × (1 - 0.0096)^2 = 27,977 × 0.9808	27,441
2027	27,977 × (1 - 0.0096)^3 = 27,977 × 0.9714	27,176
2028	27,977 × (1 - 0.0096)^4 = 27,977 × 0.9620	26,917
2029	27,977 × (1 - 0.0096)^5 = 27,977 × 0.9527	26,662
2030	27,977 × (1 - 0.0096)^6 = 27,977 × 0.9435	26,411

Based on rising demand, we assume a consumer demand growth (~ 8.1% CAGR till 2030), based on (Grand View Research)

Deloitte, PWC reports suggest AI-driven production efficiency could boost annual content output (by approximately 3% CAGR),

Since growth demand may already involve tech investment), let's use a **conservative combined CAGR**:

Combined CAGR = 8.1% (demand) + 1.5% (tech adoption) = ~9.6% CAGR

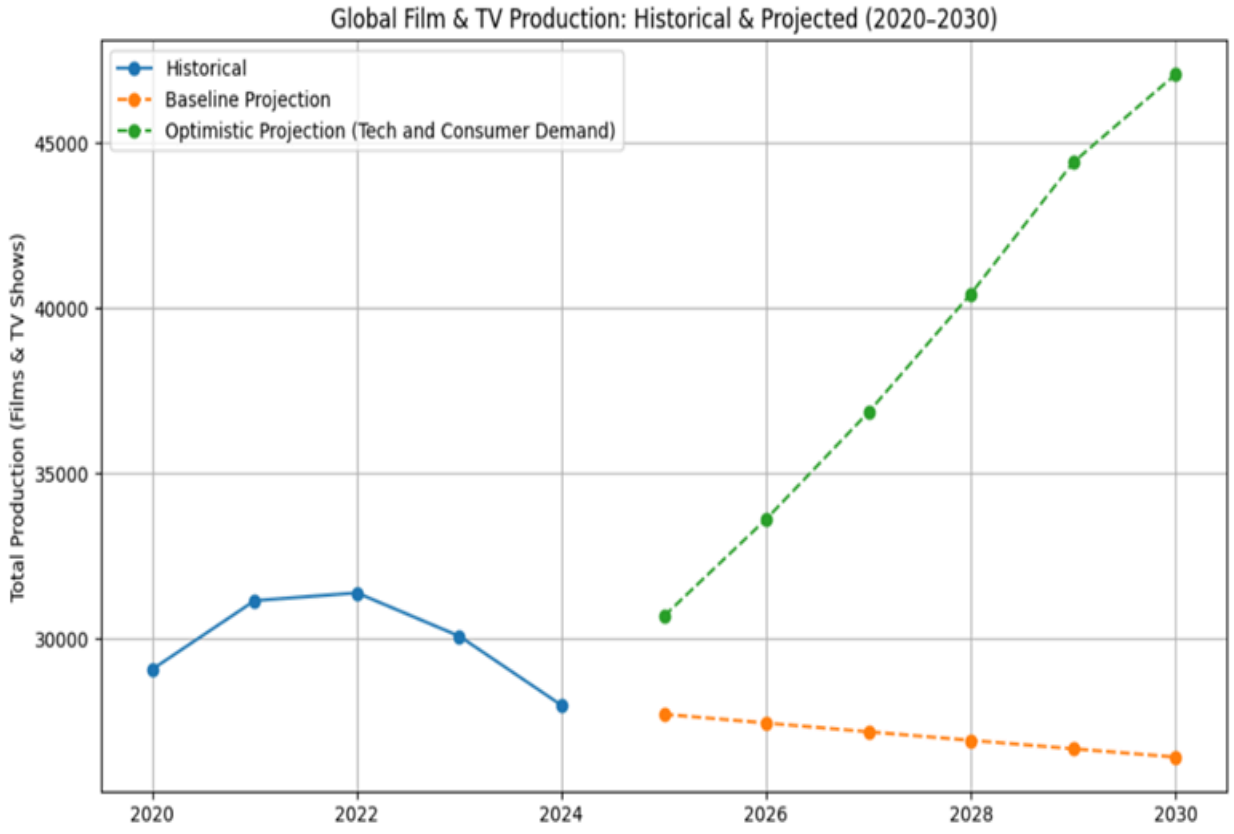
We'll now use the **compound growth formula**:

Future Value= Present Value×(1+ cagr)^n where

Present Value = 27,977(No.of films& shows in 2024), r = 0.096 (9.6%), t = 6 (from 2024 to 2030)

Projected 2030 Production = 27,977 x (1 + 0.096)^6
= 27,977 x 1.683
= 47,074

By 2030, approximately 47,074 new films and TV shows could be produced globally per year, assuming a 9.6% annual growth rate driven by rising consumer demand and AI-enabled production efficiencies.



This chart is created using Python as shown here

5. Percentage of global box office revenue generated by international markets in the next five years

- This question requires predicting the contribution of non-domestic markets to the overall box office revenue, taking into account trends in international film distribution and audience preferences.

Answer :

Source = https://en.wikipedia.org/wiki/Box_Office_Mojo ,

Historical Data Overview Box office revenue (In Billion USD)			
Year	**Revenue**	**Domestic**	**Non-Domestic**
2020	14.07 Billion	1.95 billion (13%)	12.12 billion (77%)
2021	21.9 Billion	4.48 billion (20%)	17.42 billion (80%)
2022	26.0 Billion	7.37 billion (28%)	18.63 billion (72%)
2023	33.9 Billion	8.9 billion (26%)	25.00 billion (74%)
2024	32.3 Billion (estimated)	8.57 billion (26%)	23.73 billion (74%)

CAGR Calculation (2020 to 2024) :

Formula = (Final Value / Initial Value) ^ (1 / n) - 1

Where : Initial Value (2020) = 14.07 billion , Final Value = (2024) = 32.3 billion

Calculation :

CAGR = (32.3 / 14.07) ^ (1/4) - 1 = 0.233 * 100

= 23.3%

Future Global Revenue Forecast(2025 - 2029):

Formula

Revenue n = Base Revenue x (1+CAGR)^n

Where n = numbers of years after 2024

Non Domestic Revenue Estimation Based On Previous Study :

Calculating avg of non domestic revenue percentage =

(86.1 + 79.6 + 71.6 + 73.7 + 73.5) / 5 = 76.9 %

We can take 70 % as a guess estimation for this because of following

It's slightly lower than recent trend, offering margin of safety
Accounts for possible decline in international dominance (due to regional streaming growth, localization, etc.)
Keeps forecasting conservative and realistic.

Forecasted Revenue Of Next 5 Years(2025 - 2029)

Years	Formula Implementation	Forecasted Revenue	Non-Domestic (70%)
2025	32.3 × (1 + 0.233)^1	39.81 billion	27.87 billion
2026	32.3 × (1 + 0.233)^2	49.08 billion	34.36 billon
2027	32.3 × (1 + 0.233)^3	60.48 billion	42.34 billion
2028	32.3 × (1 + 0.233)^4	74.48 billion	52.13 billion
2029	32.3 × (1 + 0.233)^5	91.68 billion	64.18 billion

Based on recent trends and a conservative estimate of 70% international contribution, this analysis projects the non-domestic box office revenue for the next five years using CAGR. The results offer a realistic outlook for global market performance and can guide strategic planning in the film industry.

Part - III

Scenario 1 : An entertainment company wants to analyse the behaviour of users who signed up for a **premium streaming subscription** in the past year. They want to track how many of these users renew their subscription in the months following their initial sign-up.

Question 1:
How would you calculate the **monthly retention rate** for each cohort of users who signed up in different months?

- **Hint**: Group users by the month they signed up (cohort), and calculate the percentage of users who renew their subscription in subsequent months.

Question 2:
If you notice that users tend to drop off after the third month, what strategies would you propose to improve **long-term retention**?

- **Hint**: Consider offering special discounts, exclusive content, or personalised recommendations to keep users engaged after the initial months.

Solution 1 :

To calculate the monthly retention rate for each user cohort:

1. Define Cohorts: Group users by their `signup_month` (the month they initially subscribed).
2. Calculate Initial Cohort Size: For each `signup_month` cohort, count the total number of unique users.
3. Track Active Users Per Month: For each `signup_month` cohort, and for each subsequent month (Month 1, Month 2, etc., relative to signup):
    - Identify the actual calendar month (e.g., if signup was Jan 2024, Month 3 is Apr 2024).
    - Count how many users from that *original cohort* were still actively subscribed (i.e., their subscription covered that specific calendar month).
4. Calculate Retention Rate: Divide the number of active users in each subsequent month by the initial cohort size, then multiply by 100 to get a percentage.

Cohort Analysis Metrics Table – Netflix Subscription Retention

Step	Details	Metric	Details
Define Cohorts	Group users based on the month they started their **first Netflix subscription**.	Subscription Start Month	E.g., Jan 2024 cohort includes all users who subscribed in January 2024.
Track Metrics	Monitor key metrics like renewal rate, watch activity, and churn over time.	Monthly Renewal Rate	Percentage of users in each cohort who renew their subscription each month.
		Average Watch Time	Average number of hours watched per user in each cohort per month.
Analyze Data	Compare how different cohorts behave across months to spot trends.	Retention Rate	% of original cohort still subscribed after 1, 3, 6 months.
		Engagement Rate	% of users who streamed at least 1 title per month.

Step 1: Data Collection

User ID	Subscription Start	Subscription End	Activity Month	Watch Hours
1	2024-01-05	2024-04-05	2024-01	30
1	2024-01-05	2024-04-05	2024-02	40
1	2024-01-05	2024-04-05	2024-03	50
2	2024-01-18	2024-03-18	2024-01	20
2	2024-01-18	2024-03-18	2024-02	60
3	2024-02-10	2024-04-10	2024-02	35
3	2024-02-10	2024-04-10	2024-03	25
4	2024-02-15	2024-03-15	2024-02	15
5	2024-03-05	2024-05-05	2024-03	45
5	2024-03-05	2024-05-05	2024-04	50

Step 2: Define Cohorts

Cohort	User IDs	Cohort Start Date	Cohort End Date
Jan 2024	1, 2	2024-01-01	2024-01-31
Feb 2024	3, 4	2024-02-01	2024-02-28
Mar 2024	5	2024-03-01	2024-03-31

Step 3: Track Metrics

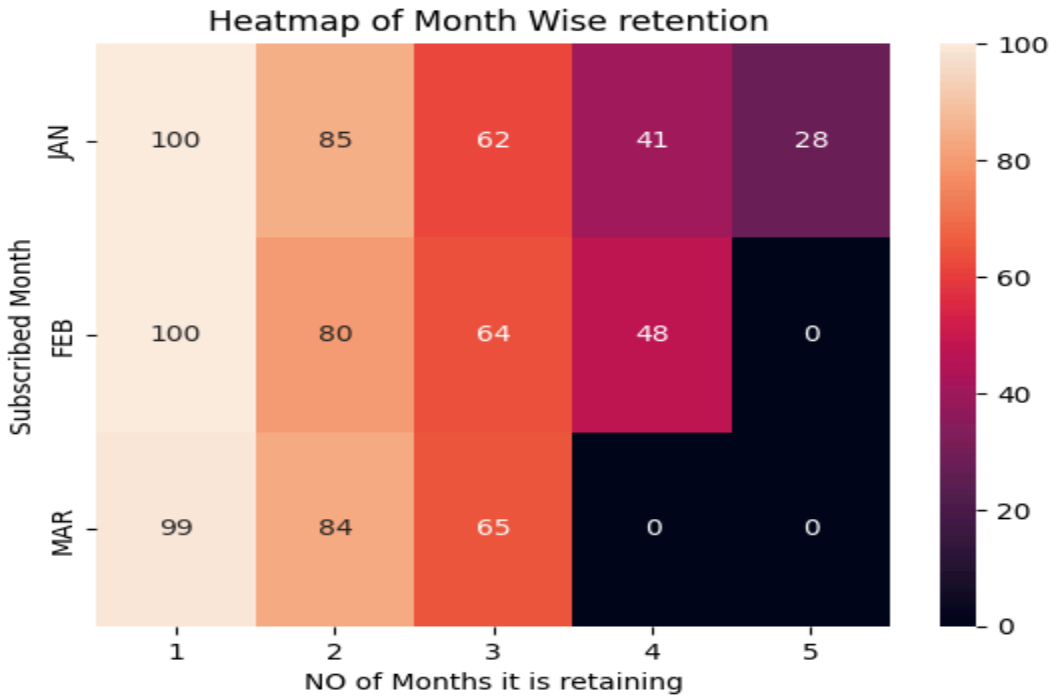
Metric	Jan 2024 Cohort	Feb 2024 Cohort	Mar 2024 Cohort
Number of Users	2	2	1
Initial Watch Hours	30, 20	35, 15	45
Repeat Activity Months	2	1	1
Total Watch Hours	120	75	95
Retention After 1 Month	100% (2/2)	50% (1/2)	100% (1/1)

Step 4: Analyse Patterns

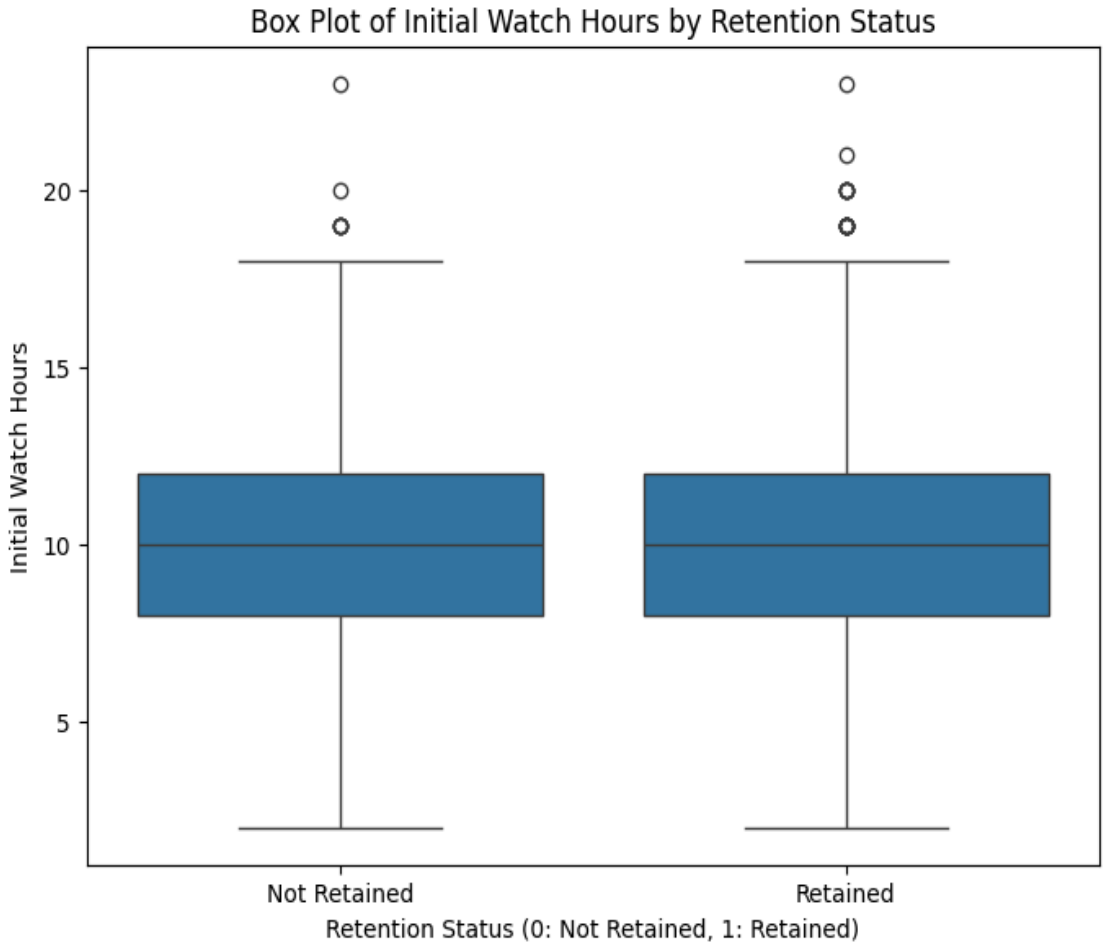
Pattern	Jan 2024 Cohort	Feb 2024 Cohort	Mar 2024 Cohort
Avg Initial Watch Hours	25	25	45
Avg Total Watch Hours/User	60	37.5	95
Retention After 2nd Month	50%	0%	—
Engagement Trend	Increasing	Declining	Consistent

[Colab link for The Entire Analysis of The Data]

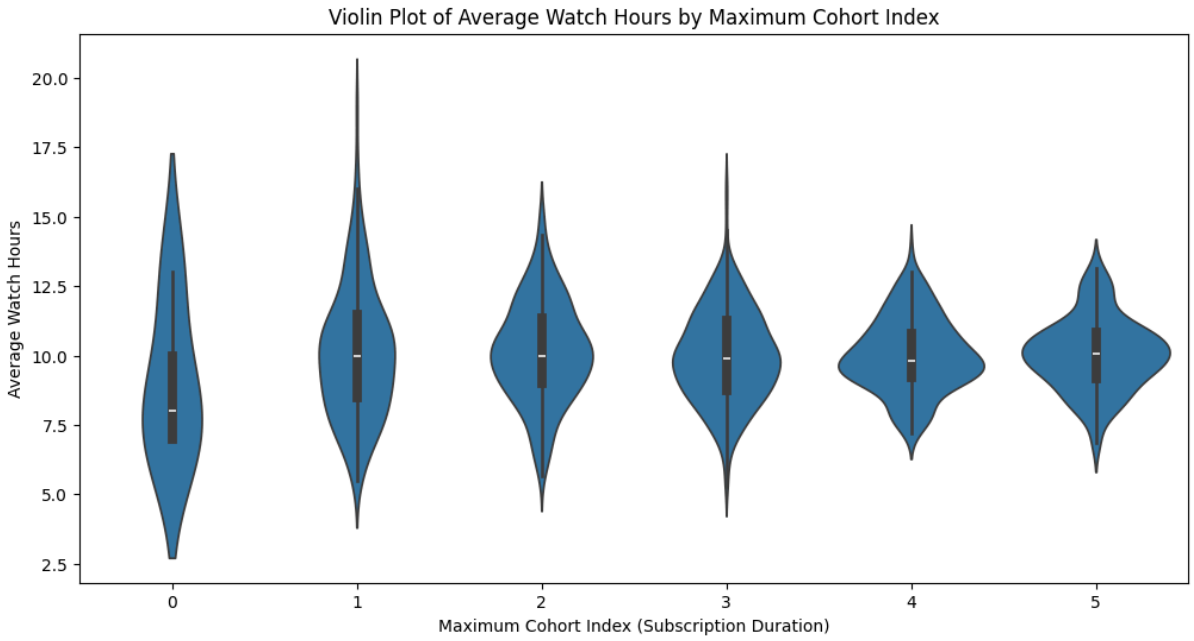
HeatMap Showing Monthly Retention



Boxplot Showing How Watch Hours Can affect the Retention Status



Violin Plot of Average Watch Hours by no of months a user is Subscribed



Step 5 : Actionable Insights

Insights	Actions
Retention decreases over time: The retention heatmap shows a decline in retention rate as the cohort index increases.	Implement strategies to re-engage users in later months, such as personalized content recommendations or targeted promotions.
Watch hours vary by subscription duration: The box and violin plots suggest a relationship between how long a user subscribes and their average watch hours.	Analyze the watch patterns of long-term subscribers to identify common characteristics and tailor content or features to encourage similar behavior in newer users.
Initial watch hours may influence retention: The box plot comparing initial watch hours for retained and not-retained users shows a potential difference.	Investigate if higher initial watch hours correlate with increased retention and consider onboarding strategies that encourage early engagement.

Solution 2 :

After the analysis if we look at the heatmap of the retention rates it's clearly saying that the retention rates drops after the third month . so Here are some proposed strategies to improve long term retention.

Category	Strategy	Description
Content Engagement	**Personalized Recommendations**	Leverage AI to suggest highly relevant content based on individual viewing history and preferences.
	Exclusive/Early Access Content	Offer unique shows, behind-the-scenes content, or early access to new releases for long-term subscribers.
	Consistent Content Pipeline	Ensure a steady stream of new and engaging original content to prevent boredom.
User Experience & Value	**Targeted Re-engagement Offers**	Provide personalized discounts or promotions around the 2-3 month mark to at-risk users.
	Highlight Value Reinforcement	Periodically communicate new features, added content, or viewing milestones to remind users of subscription benefits.
	Seamless Multi-Device Experience	Ensure consistent, high-quality streaming and user experience across all devices.

Community & Interaction	**Gamification & Rewards**	Implement badges or loyalty points for continuous subscription or engagement (e.g., "6-Month Streamer").
	Community Features	Foster user interaction through forums, watch parties, or polls related to content.
Proactive Support	**Automated Nudge Campaigns**	Send personalized in-app notifications or emails prompting users to explore new content or features.
	Solicit & Act on Feedback	Proactively gather feedback from users, especially around the 2-3 month mark, to address pain points.

Scenario 2:

A gaming company is testing two different **tutorial designs** for new users in its mobile game. **Version A** is a brief, text-based tutorial, while **Version B** is an interactive, step-by-step guide. They want to see which tutorial leads to better **user retention** and higher **conversion rates** (i.e., users making in-game purchases).

Question 1:
Design an **A/B test** to evaluate which tutorial version leads to better retention and conversion rates. What metrics would you use to measure success?

- **Hint**: Track retention (e.g., percentage of users who return after one week) and conversion rates (e.g., percentage of users who make their first purchase within 7 days).

Question 2:
If **Version B** (interactive guide) shows higher conversion rates but slightly lower retention, how would you balance these results when making a recommendation to the business?

- **Hint**: Consider the long-term value of higher conversions versus the potential impact of reduced retention. You might suggest refining Version B to improve retention without sacrificing conversions.

Solution 1 :

Hypothesis:

> ***"Users who go through an interactive, step-by-step tutorial (Version B) will have higher retention and conversion rates than those who go through a text-based tutorial (Version A)."***

A/B Test Design :

Variants :

Variant A (Control): Text-based tutorial

Variant B (Treatment): Interactive tutorial

Selection :

To ensure that their is no selection bias we have decided to assign random assignment with 50% to variant A and 50% to variant B .

Duration

Run the experiment for **14 days**, tracking user behavior for 7 days after onboarding.

Isolation

Ensure users in each group are isolated from other influences like promotions, ads, or referral bonuses.

Success Metrics

Primary Metrics

- **Day 7 Retention Rate** = (Users who return on Day 7) / (Total users in group)
- **7-Day Conversion Rate** = (Users who make at least one in-game purchase within 7 days) / (Total users in group)

Secondary Metrics

- Average time spent per session
- Average number of levels completed

Result Interpretation

Summary Table

Variant	Users	Retention Rate	Conversion Rate
A	10,000	43.94%	10.2%
B	10,000	42.27%	15.4%

Colab link to for the testing analysis

Statistical Significance Analysis

- **Retention p-value**: 0.0178
  → Statistically significant at **5% level** → *The difference in retention is likely not due to random chance.*

- **Conversion p-value**: 0.0000
  → **Highly statistically significant** (typically this means $p < 0.0001$)
  → *The difference in conversion is extremely unlikely to be due to random chance.*

Recommendation

Based on the analysis:

- **Adopt Version B** (interactive tutorial):
    - It significantly increases conversion (by ~5.2 percentage points).
    - Though retention drops slightly (~2.8%), the difference is statistically significant but relatively small.
- **Mitigation plan for retention**:
    - Add onboarding checklists or reward nudges post-tutorial
    - Send reminder notifications to bring users back on Day 3 and Day 7

Final Recommendation: Implement Version B to drive early revenue growth, while continuing to optimize it to maintain or improve retention.

Solution 2 :

If Version B (the interactive tutorial) results in a significantly higher conversion rate but slightly lower retention, the recommendation should focus on balancing short-term revenue gains with long-term user engagement. Since the increase in conversion is quite substantial and statistically significant (15.4% vs. 10.2%, p = 0.0000), it indicates that Version B is effective in encouraging users to make early in-game purchases. This is beneficial for the business because higher conversions lead to improved early monetization, better return on ad spend, and faster revenue growth. On the other hand, the drop in Day 7 retention (from 45.1% to 42.3%, p = 0.0178) is also statistically significant, although the difference is smaller in magnitude. Lower retention could affect the long-term lifetime value of users and might lead to increased marketing costs to retain or reacquire users. Therefore, the best course of action is to implement Version B due to its strong performance in conversions, but simultaneously work on optimizing it to improve retention. Some ways to do this include adding post-tutorial engagement prompts, gamifying early progress with rewards, or introducing reminder notifications to encourage users to return. These additions can help maintain the conversion benefit of Version B while addressing the retention gap. This balanced approach ensures the business benefits from higher early revenue without compromising user engagement in the long run.