

Analysis of Streaming Services Data

Exploring data, distributions, and correlations in the global streaming market





Dataset Background

Data Description

The data focuses on streaming services and includes platform-level data such as:

- Countries Available
- Monthly/Annual Price
- Number of Subscribers
- Additional metrics supporting business and behavioral analysis

Analysis Goal

To perform Exploratory Data Analysis (EDA), examine key distributions, and investigate statistical relationships (correlations) between business variables—in order to derive operational and business insights.

Granularity

High Granularity Level: Each row in the table details one streaming service



Service Comparison

Enables comparison between services
(Benchmarking)



Distribution Calculations

Enables distribution calculations,
rankings, and correlation calculations
between different service
characteristics



Country Analysis

In the countries_available column,
there might be multiple countries for a
single service—therefore, for country
distribution analysis, the field can be
broken down to a "country-service"
level.

Initial Exploration, Cleaning and Quality Checks (EDA + Data Validation)

Quality Checks Performed (as executed in code)

- Checking for missing values:
`df_streaming.isnull().values.any() +
df_streaming.isnull().sum()`
- Checking structure and data types: `df_streaming.info()`
- Checking for duplicates: `df_streaming.duplicated().any()`

Feature Engineering / Auxiliary Columns

Adding a reference column for average monthly price:

```
df_streaming["month price avg"] =  
df_streaming["monthly_price_usd"].mean()
```

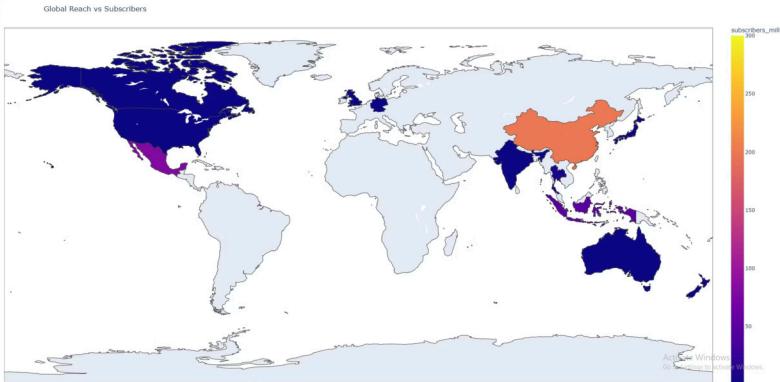
This column served as a comparison line/Benchmark in analyses and visualizations.

- Data Quality Summary:** After the initial checks and review, the data was found to be **suitable for analysis** (without missing values/duplicates according to our checks), and therefore we proceeded to the research questions and correlations.

Streaming Service Distribution by Country

Research Question

What is the distribution of countries where streaming services are available, and which country leads in the number of services?



33%
India
**Leading country in number
of streaming services**

Key Finding

The country with the most streaming services in the data is **India**, with almost **a third** of all services in the table.

Conclusion

India represents a large and dominant streaming market, with a multitude of platforms and high availability.

Most Popular Streaming Service by Number of Subscribers

Research Question

Which is the most popular streaming service in terms of subscriber count?

300M

MX Player

Registered Subscribers

Key Finding

The most popular service is **MX Player** with approximately **300 million** registered subscribers.

Conclusion

MX Player conspicuously leads in user base, representing a service with high adoption and broad market penetration.

Does monthly price affect the number of subscribers?

Research Question

Is there a relationship/correlation between the monthly price and the number of subscribers?

Hypothesis

We expected a relationship (for example: higher price → fewer subscribers), and therefore expected a significant correlation.

Conclusion

There is no significant correlative relationship between monthly price and number of subscribers in our data.

0.05

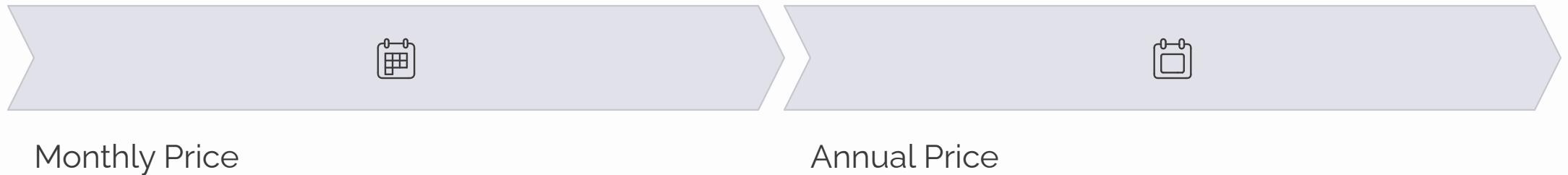
Correlation

Key Finding

Positive Correlation: Monthly Price vs. Annual Price

Research Question

Does the monthly price "affect" the annual price (i.e., is there a correlation between them)?



Key Finding

A very strong positive correlation was found.

Explanation

The annual price is derived from the monthly price, so the correlation is expected and even "almost perfect".

Conclusion

The data is logically consistent: annual pricing directly aligns with monthly pricing.

New Content vs. Monthly Price and Subscribers vs. Churn Rate

Research Question

Is "content innovation" (new movies/series) related to the monthly price? Are services with more new content more expensive?

Hypothesis

We expected a high positive correlation (more new → more expensive).

Key Finding

A **positive but low correlation** was found.

Conclusion

Despite expectations, no strong link was found between the new content index and the monthly price in the data.

Research Question

Is there a connection between the number of subscribers and the churn rate/likelihood?

Hypothesis

We assumed that the more subscribers there are, the more churn there might be (in absolute terms), and therefore expected some connection.

Key Finding

A **negative correlation** was found between the number of subscribers and the churn rate.

Conclusion

As the number of subscribers increases, the churn rate tends to be lower—which may indicate higher stability in larger platforms.

Summary of Key Insights

India Leads

India leads in the number of streaming services (almost a third of the data).

MX Player

MX Player is the leading service by number of subscribers (~300M).

Price and Subscribers

There is no significant correlation between monthly price and number of subscribers ($r \approx 0.05$).

Monthly and Annual Price

There is a strong positive correlation between monthly price and annual price (direct derivative).

New Content

New content was not found to have a strong correlative effect on the monthly price.

Platform Stability

A negative correlation was found between the number of subscribers and churn rate (larger platforms are more stable).