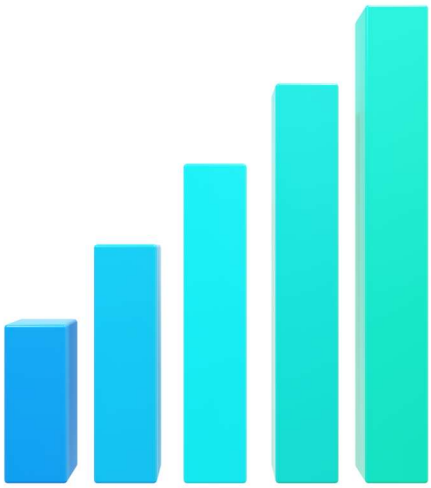

Data analytics process and Streaming Services
MENTORMINDS

ABSTRACT

Logged user data has become a highly valued asset to many Internets based services with large user bases. Being able to draw insight from this data is considered a key to gaining competitive advantages for the companies behind the services. This study aims to identify patterns in the collection of users data when interacting with, a music streaming service and a video streaming platform by data analytics tools. In the study, I studied and examined why the streaming companies need user data. I identified different types of analysis, each representing various patterns of user behavior. Also identify factors partly explaining the differences in behavior between different sessions. Finally, I demonstrate how user data helps to provide a better service.

INTRODUCTION



PROBLEM

Organizations, people, and things are generating massive amounts of data every day. In a 24-hour period, we collectively send 294 billion e-mails and 500 million tweets. We plug 3.5 billion searches into Google. Our connected cars generate a whopping four petabytes of data. Even our watches, fridges, and TVs are constantly creating and sharing data.

APPROACH

Hidden in all this data are insights that can trigger explosive business growth. The challenge is in finding them, which is where analytics comes in.

SOLUTION

Data analytics involves the manipulation and computation of large volumes of data, often from a wide variety of different sources. Manipulation and computation are performed at high velocity to identify patterns, correlations, and other useful information. Data analytics capabilities have been found to enhance the utility of the mass amounts of data that can be collected and communicated through the Company on a regular basis to enhance the Business performance and provide better service. Data analytics helps to draw insights and find opportunities in the market.

UNDERSTANDING THE CONCEPTS OF THE STUDY

DEFINITION OF DATA ANALYTICS

Data Analytics is a field of computer science that uses math, statistics, and machine learning to find meaningful patterns in data. Data analytics involves sifting through massive data sets to discover, interpret, and share new insights and knowledge.

PROCESS OF DATA ANALYSIS

Data analysis is a process of finding, collecting, cleaning, examining, and modeling data to derive useful information and insights and understand the derived information for data-driven decision-making.

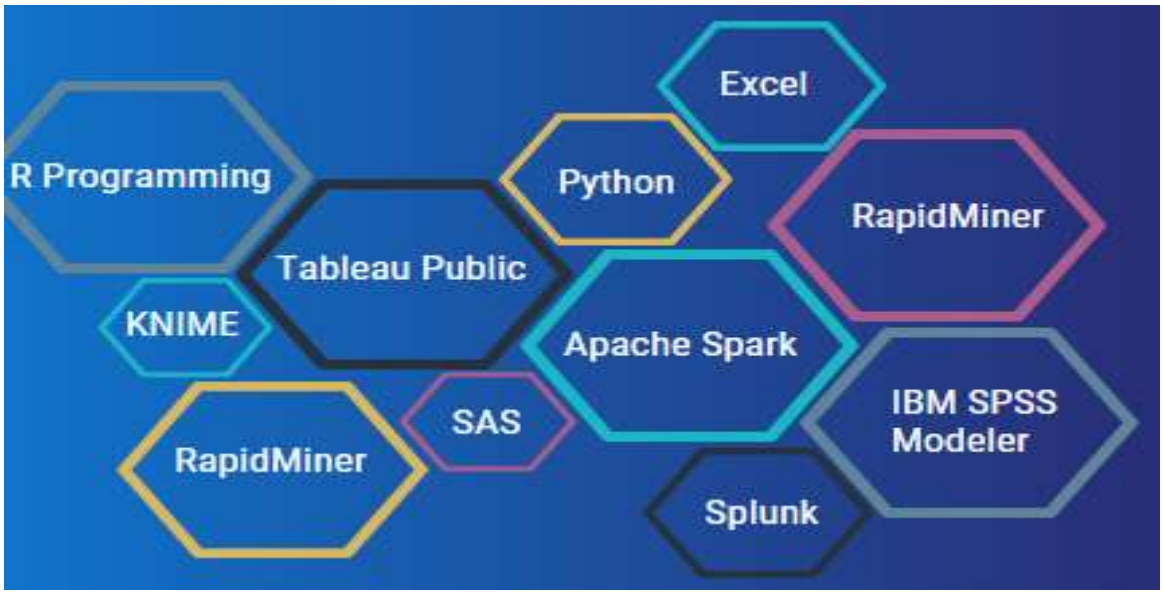
They are 5 crucial steps involved in Data Analysis process:-

Step 1: Define why you need data analysis
Step 2: Collect data
Step 3: Clean unnecessary data
Step 4: Perform data analysis
Step 5: Data visualization
Step 6: Interpret the results

TOOLS AND TECHNIUES USED IN ANALYSE DATA

Data Analysis Tools

There are several data analysis tools available in the market, each with its own set of functions. The selection of tools should always be based on the type of analysis performed and the type of data worked. Here is a list of a few compelling tools for Data Analysis.



1. **Excel:** It has various compelling features, and with additional plugins installed, it can handle a massive amount of data. So, if you have data that does not come near the significant data margin, Excel can be a versatile tool for data analysis.
2. **Tableau:** It falls under the BI Tool category, made for the sole purpose of data analysis. The essence of Tableau is the Pivot Table and Pivot Chart and works towards representing data in the most user-friendly way. It additionally has a data cleaning feature along with brilliant analytical functions.
3. **Power BI:** It initially started as a plugin for Excel, but later on, detached from it to develop in one of the most data analytics tools. It comes in three versions: Free, Pro, and Premium. Its PowerPivot and DAX language can implement sophisticated advanced analytics similar to writing Excel formulas.
4. **R & Python:** These are programming languages that are very powerful and flexible. R is best at statistical analysis, such as normal distribution, cluster classification algorithms, and regression analysis. It also performs individual predictive analyses like customer behaviour, spending, items preferred by him based on his browsing history, and more. It also involves concepts of machine learning and artificial intelligence.
5. **SAS:** It is a programming language for data analytics and data manipulation, which can easily access data from any source. SAS has introduced a broad set of customer

profiling products for web, social media, and marketing analytics. It can predict their behaviours, manage, and optimize communications.

6. **RapidMiner**: RapidMiner is a platform for data processing, building Machine Learning models, and deployment.

7. **Apache Spark:** Spark Is an integrated analytics engine for Big Data processing designed for developers, researchers, and data scientists.

- **Data Analysis Techniques**

There are different techniques for data analysis, depending on the question, the type of data, and the amount of data collected. Therefore, different methods of data analysis can be categorized as follows:

Techniques based on Mathematics and Statistics

1. Descriptive Analysis
2. Dispersion Analysis
3. Regression Analysis
4. Factor Analysis
5. Discriminant Analysis
6. Time Series Analysis
7. Sentiment Analysis

Techniques based on Artificial Intelligence and Machine Learning

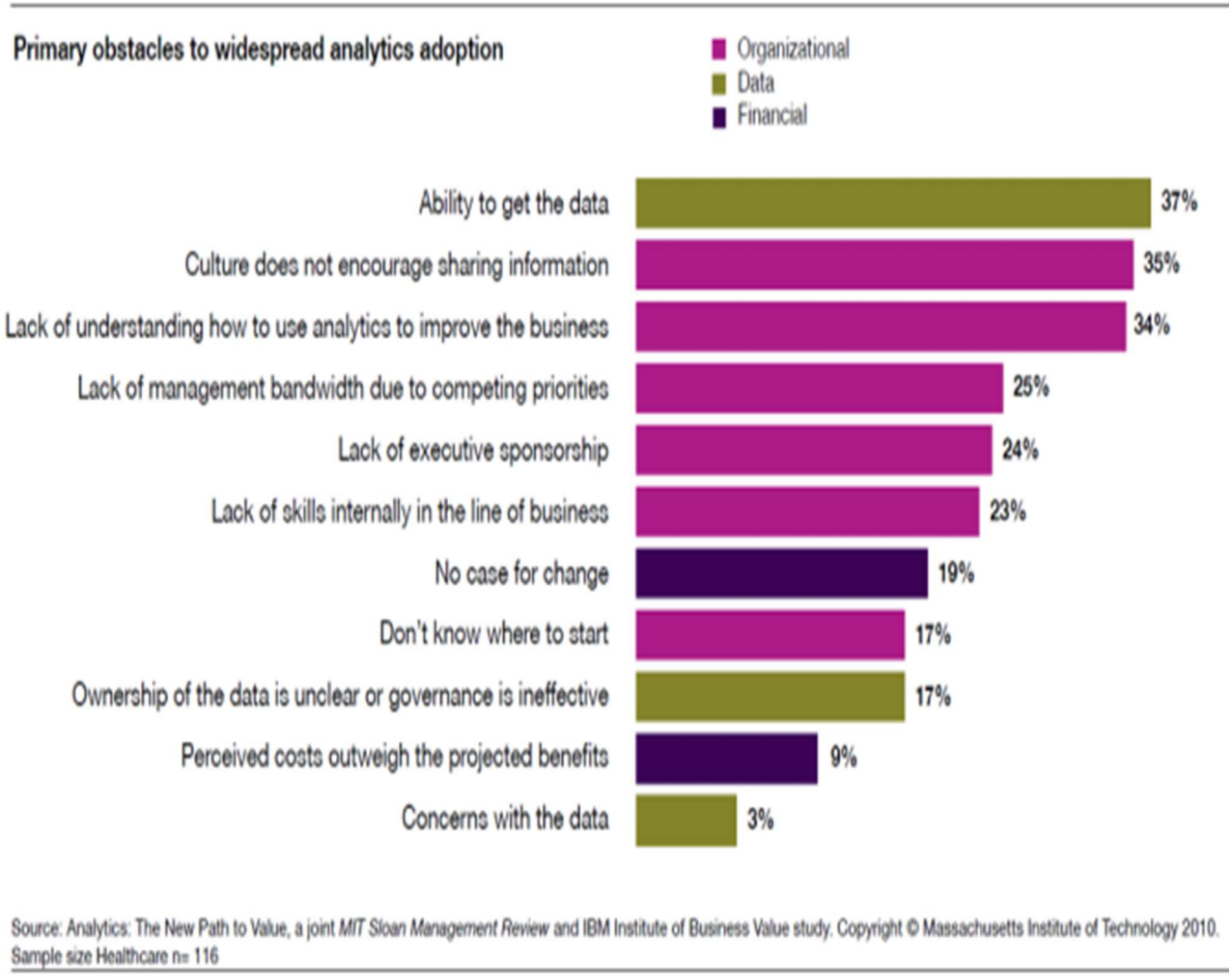
1. Artificial Neural Networks
2. Decision Trees
3. Evolutionary Programming
4. Fuzzy Logic

Techniques based on Visualization and Graphs

1. Column Charts, Bar Charts
2. Line graph
3. Area Graph
4. Pie chart
5. Funnel chart
6. Word Cloud Chart
7. Gantt Chart
8. Radar chart
9. Scatter plot
10. Bubble Chart
11. Gauge
12. Frame Diagram
13. Rectangular Tree Diagram

BARRIERS TO EFFECTIVE ANALYSIS

1. Data Discovery Challenges.
2. Getting Voluminous Data into The Data Platform
3. Uncertainty Of Data Management Landscape
4. Need For Synchronization Across Disparate Data Sources
5. Integrating Disparate Data
6. Getting Meaningful Insights Through the Use of Data Analytics
7. Data Storage and Quality
8. Security And Privacy of Data
9. Hiring and Retaining Workers with data skills



Source: Analytics: The New Path to Value, a joint *MIT Sloan Management Review* and IBM Institute of Business Value study. Copyright © Massachusetts Institute of Technology 2010.
Sample size Healthcare n= 116

RESEARCH WEBSITES

1. **SPOTIFY (music streaming service)**

Spotify- a music streaming app loved and used by millions of users daily. The app is quite successful in the music industry and is the largest on-demand music service globally. The latest technology like Big Data plays a tremendous role in making Spotify the leading music streaming app. Presently, Spotify's net worth is more than **$25 billion** and has an enormous impact on the music industry by using data in the most dynamic ways. Spotify creates a playlist automatically according to the user choice; provide a personal touch- it all can be possible using Big Data Analytics.

- **What user data do Spotify collect?**

1. Personal data collected when signing up for the Spotify Service or when user update account
2. Personal data collected through use of the Spotify Service
3. Personal data that user choose to give Spotify
4. Personal data Spotify collect from third party sources

1. Personal data collected when signing up for the Spotify Service or when user update account	
User Data	profile name email address phone number date of birth gender street address country
2. Personal data collected through use of the Spotify Service	
Usage Data	• Information about your Spotify Service Option. • Your actions with the Spotify Service (including date and time), such as: ✓ search queries ✓ streaming history ✓ playlists user creates ✓ User library ✓ Browsing history ✓ account settings ✓ interactions with other Spotify users ✓ User use of third-party services, devices and applications in connection with the Spotify Service Use technical data Examples include: • URL information • Online identifiers such as cookie data and IP addresses • Information about the devices you use such as: ✓ device IDs ✓ network connection type (e.g., wife, 4G, LTE, Bluetooth) ✓ provider

	✓ network and device performance ✓ browser type ✓ language ✓ information enabling digital rights management ✓ operating system ✓ Spotify application version Information (such as the device name, device identifiers, brand and version) which enables us to discover and connect with third party devices and applications such as those listed below: ✓ devices on your Wi-Fi___33 network (such as speakers) which can connect to the Spotify Service ✓ devices made available by your operating system when connecting via Bluetooth, plugin, and installation ✓ Spotify partner applications to determine whether the application is installed on your device
3. Personal data that user choose to give Spotify	
Voice Data	If voice features are available in user market and only where user gives permission, Spotify collect voice data (audio clips of user voice).
Payment and Purchase Data	The exact personal data collected will vary depending on the payment method, but will include information such as: • name • date of birth • credit or debit card type, expiration date, and certain digits of your card number Note: For security, Spotify never store your full card number • ZIP/postal code • mobile phone number • entails of your purchase and payment history
Contests, Surveys, and Sweepstakes Data	When user complete any forms, respond to a survey or questionnaire, or take part in a contest or sweepstake, Spotify collect the personal data you provide.

4. Personal data Spotify collect from third party sources	

Third party applications, services and devices user connect to Spotify account	These third-party apps, services or devices may include:
	• social media
	• devices including:
	  ○ audio (e.g., speakers and headphones)
	  ○ smart watches
	  ○ televisions
	  ○ mobile phones and tablets
	  ○ automotive (e.g., cars)
	  ○ games consoles
	• services or platforms such as voice assistants

Payment partners	If user choose to pay Spotify by invoice, Spotify may get data from payment partners. This allows Spotify to: • send user invoices • process user payment • give user what they purchased
Advertising and marketing partners	• cookie id • mobile device id • email address

➢ **Why do they need this user data?**

- To provide the personalized Spotify services
- To understand, diagnose, trouble shoot and fix issues with Spotify services
- To evaluate and develop new features, technologies and improvement to the Spotify services
- To comply with legal obligations and law enforcement request
- For marketing, promotions and advertising purpose
- To conduct business planning, reporting and forecasting
- To process your payments
- To conduct research contests, surveys and sweepstakes
- To detect and prevent fraud, including fraudulent payments and fraudulent use of Spotify services.
- To conduct research, contests, surveys, and sweepstakes.

1. **Netflix (Video Streaming Platform)**

Founded in 1997 as a subscription mail-order DVD company, Netflix has grown to be the top digital streaming platform, with over 160 million subscribers worldwide. The streaming giant has steadily grown over the past two decades using insights from its treasure trove of user data to personalize content recommendations and inform content curation.

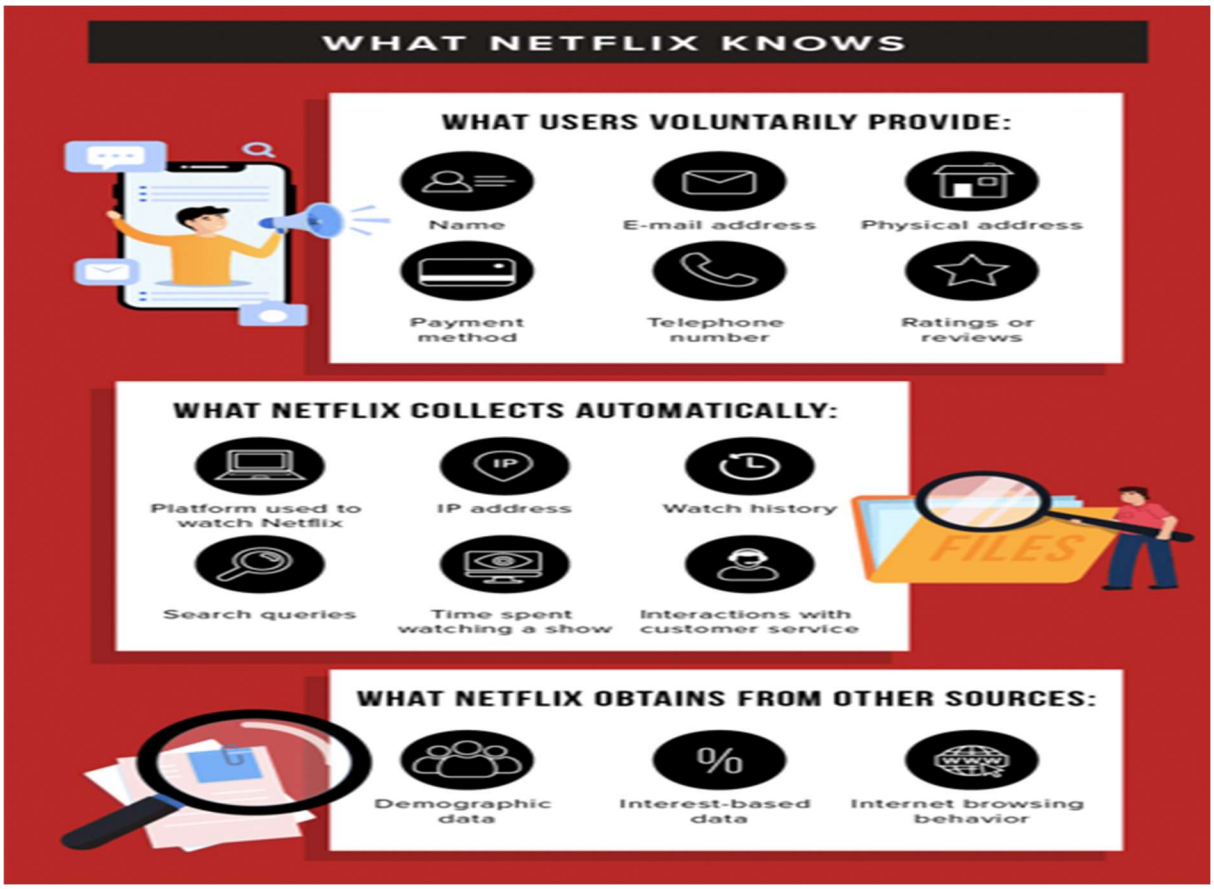
Netflix is the world's leading premium media streaming platform, operating in nearly every country in the world. It was one of the first players in the streaming industry when it transitioned in 2007, and the bet has paid off with hundreds of millions of subscribers worldwide.

➢ **What user data do they collect?**

- **Information user provide to Netflix:** Netflix collect information user provide to them, which includes:
  - User name, email address, payment method(s), telephone number, and other identifiers user might use (such as an in-game name).
  - information when user choose to provide ratings, taste preferences, account settings (including preferences set in the "Account" section of our website).

- **Information Netflix collect automatically:**
  - Information about user and use of service
  - User interactions with app and advertising

- Information regarding user network, network devices, and computer or other Netflix capable devices user might use to access Netflix service (such as gaming systems, smart TVs, mobile devices, set top boxes, and other streaming media devices).



- **Information from partners:**
  - search queries and commands applicable to Netflix that you make through Partner devices or voice assistant platforms;
  - service activation information such as your email address or other contact information;
  - IP addresses, device IDs or other unique identifiers, as well as associated pre-paid promotion, billing and user interface information, that support user authentication, the Netflix service registration experience, Partner payment processing, and the presentation of Netflix content to you through portions of the Partner user interface.

- **Information from other sources:**
    - Service providers that help Netflix determine a location based on User IP address in order to customize service and for other uses consistent.
    - Security service providers who provide Netflix with information to secure systems, prevent fraud and help us protect the security of Netflix accounts.
    - Payment service providers who provide Netflix with payment or balance information, or updates to that information, based on their relationship with you;
    - Online and offline data providers, from which Netflix obtain aggregated demographic, interest based and online advertising related data;

➢ Why do they need this user data?

Netflix use information to provide, analyse, administer, enhance and personalize services and marketing efforts, to manage member referrals, to process user registration, orders and payments, and to communicate with on these and other topics.

- PDetermine user general geographic location, provide localized content, provide user with customized and personalized viewing recommendations for movies and TV shows
- Coordinate with Partners on making the Netflix service available to members and providing information to non-members about the availability of the Netflix service.
- Secure Netflix systems, prevent fraud and help us protect the security of Netflix accounts;
- Prevent, detect and investigate potentially prohibited or illegal activities, including fraud, and to enforce Netflix terms.
- Analyse and understand audience, improve service (including user interface experiences and service performance) and optimize content selection, recommendation algorithms and delivery;
- Communicate with users concerning of service so that user can receive news about Netflix, details about new features and content available on Netflix, special offers, promotional announcements, consumer surveys, and to assist users with operational requests such as password reset requests.

FINDINGS AND ANALYSIS

- **How do end using this data to create a better service?**

1. **Developing Personalized Content**

A crucial approach Spotify applies to adopt the data generated by their users is to use it for developing content that every user will regard to be exclusive to their unique tastes.

2. **Digitizing the taste of the user**

The daily taste profile of the listener is also incorporated in Spotify's playlists named **"Daily Mixes"**. These playlists are different from the music genres which the user normally incline towards and are generally composed of songs which the user has added to their playlists or saved, or which have been created by the artists which the user has included in their present playlists or any fresh artists or albums which the user is unfamiliar with.

3. **For Enhanced Marketing through targeted ads**

While enhancing the experience of customers, Spotify has also been able to adopt a humongous section of data generated through its users for the purpose of updating their ad campaigns and targeting their customers in a more compelling manner.

4. Continuously updating its system

The streaming platform had stated that their free users would no longer be required to solely shuffle through music on their application. Rather, their users were now allowed the liberty of exploring 15 of the platform's well-known playlists which included the platform's popular **"Discover Weekly"** as well as **"Rap Caviar"**.

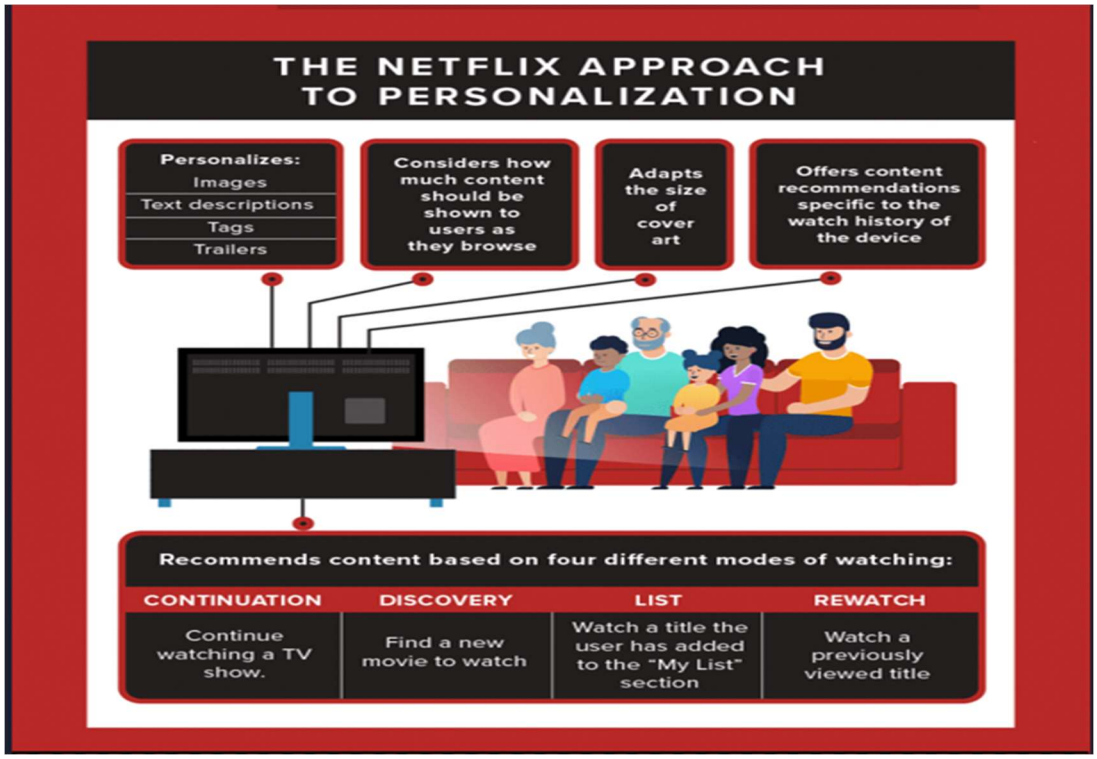
5. Spotify Wrapped

This is Spotify's year-end report, a tradition that supplies each user of the platform with an excuse to share their music taste on social media without any hesitation.

6. Forecasting and Predictions

Spotify uses big data not just for personalized services and recommendations but also to engage users apart from music streaming. It uses big data for interesting activities on the app.
The platform was able to do this by taking into account data from various sources, such as user's listening preference, album and song streaming. This helped Spotify to gauge the popularity of the music. If you are wondering whether the predictions were accurate, 4 out of 6 predictions were perfectly accurate.

- **How Netflix end using this data to create a better service?**



1. **Personalized Recommendations.**

Netflix uses the data in different ways to better understand each viewer. They do this in different ways. Of course, they track the movies and shows that the viewer sees, but that`s just a small part of the data they collect. Basically, Netflix used the data to verify that it had "repeat customers.

2. **Trends**

Netflix uses personal data to personalize each customer's suggestions, but uses the data at the macro level to understand which shows are trending.

Netflix uses the trendline at the top of the home page to inform customers of new programs to watch. This data also helps you determine which programs need to be retained, updated, and removed from the service. Netflix can make business decisions faster with confidence in real-time data that can be integrated into any business plan, regardless of industry.

3. **Original Content**

Netflix has gone from a hub that streams the content of other companies, to an original content powerhouse, with shows such as Stranger Things, Fuller House, House of Cards, and Orange Is the New Black driving a huge percentage of their overall views. Their use of data helps drive what original projects to pursue and produce. The data they collect gives an extremely clear look into their audience, what they are looking for, and how much of it they want.

4. **Marketing**

Netflix uses their data in their marketing as well. You may have noticed that the image of your favorite show changes from time to time. This is not by accident. Netflix uses data to choose the best image to use, and when to change it

5. **Optimized streaming quality**

Netflix makes use of past viewed data for predicting bandwidth usage in order to help the service decide when it should cache regional servers to ensure prompt load times during high demand. Thus, the service predicts which show is to be streamed in a certain location and caches the content in the nearby server when the internet traffic is minimal. This is done to ensure that the content is streamed without any buffering to maximize customer satisfaction.