**Exploratory Data Analysis and Dashboard Creation for Netflix Dataset using Power BI**Sai Sai Meghna Ambati

**1. Introduction**

Netflix is a renowned streaming platform that has revolutionized the entertainment industry by providing a vast library of movies, TV shows, documentaries, and original content to millions of subscribers worldwide. One of Netflix's most significant assets is its extensive collection of user data, which it utilizes to drive its business growth and enhance the overall user experience. By analyzing user behavior, preferences, and viewing patterns, Netflix can make informed decisions and shape its content acquisition, creation, and recommendation strategies. Big data visualization plays a crucial role in Netflix by helping the company make sense of vast amounts of data and derive actionable insights. In this project, we aim to visualize complex data and make it more accessible and understandable by transforming raw data into visual representations such as charts, graphs, and infographics, businesses can quickly grasp patterns, trends, and relationships within their data. This enhanced understanding enables better decision-making and identification of growth opportunities and visualizations enable businesses to anticipate market changes, forecast demand, and align their strategies accordingly, contributing to sustainable growth.

**1.1 Motivation examples of this project**

There are several examples of how big data visualization can motivate Netflix growth:

**Audience Insights:** Big data visualization can provide valuable insights into audience behavior, preferences, and viewing patterns. By visualizing user demographics, engagement metrics, and content consumption patterns, Netflix can gain a deeper understanding of its audience. These insights can help identify untapped market segments, target specific demographics with tailored content and marketing campaigns, and drive user acquisition and retention.

**Content Performance Analysis:** Visualizing data on content performance, such as viewership, ratings, and user interactions, enables Netflix to assess the success of its shows and movies. These visualizations can reveal patterns in viewership trends, identify popular genres or themes, and highlight areas for improvement. By analyzing content performance visually, Netflix can make data-driven decisions on content acquisition, production, and promotion strategies, ultimately driving viewership and business growth.

**Competitive Analysis:** Visualizing data on competitors’ offerings, market trends, and audience preferences can help Netflix gain a competitive edge. By visually comparing market share, content libraries, user ratings, or pricing strategies, Netflix can identify opportunities for differentiation and innovation. These visualizations can guide strategic decision-making, such as acquiring specific content rights, investing in new genres or formats, or entering new markets, ultimately driving business growth and maintaining a competitive position in the industry.

**Revenue and Financial Analysis:** Visualizing data on revenue streams, pricing models, and financial performance can help Netflix identify areas for revenue optimization and cost management. By analyzing financial data visually, such as subscription plans, pricing tiers, or revenue per user, Netflix can identify opportunities to maximize revenue, optimize pricing strategies, and improve overall profitability, leading to sustainable business growth.

Big data visualization motivates Netflix's business growth by providing actionable insights into audience behavior, content performance, recommendation effectiveness, churn analysis, competitive positioning, and financial analysis. By leveraging these visualizations, Netflix can make informed decisions, drive user engagement, enhance the streaming experience, optimize content strategies, and ultimately achieve its growth objectives in the highly competitive streaming industry.

**1.2 Real Applications**

There are some real applications and components that can be included in a dashboard for business growth in the context of Netflix:

**Audience Segmentation:** Create visualizations that segment the audience based on demographics, viewing preferences, and engagement metrics. This helps in identifying target segments for personalized content and marketing strategies to drive user acquisition and retention.

**Content Performance Dashboard:** Develop visualizations that display key metrics on content viewership, ratings, and user interactions. This allows content managers and decision-makers to identify successful content, track performance over time, and make data-driven decisions on content acquisition and production.

**Competitive Analysis Dashboard:** Develop visualizations that compare Netflix's market share, content library, user ratings, and pricing strategies with competitors. This provides insights into competitive positioning and helps in identifying opportunities for differentiation and strategic decision-making.

**Revenue and Financial Analysis Dashboard:** Create visualizations that display revenue streams, subscription plans, pricing tiers, and financial performance. This helps in analyzing revenue optimization opportunities, optimizing pricing strategies, and managing costs effectively.

**Geographic Analysis Dashboard:** Design visualizations that display geographic distribution of viewership, content availability, and market trends. This helps in identifying regional preferences, tailoring content strategies, and exploring new market opportunities.

**Content Acquisition and Licensing Dashboard:** Create visualizations that track the performance of licensed content, evaluate audience demand for specific genres or themes, and analyze the return on investment for content acquisitions. This helps in making informed decisions on content licensing and optimizing content acquisition strategies.

These real applications demonstrate how data visualization can be leveraged in building a dashboard for Netflix's business growth. In our project, we aim to build a dashboard which has the trends and viewership of Netflix all through the year and with that data businesses can gain insights and take actions to improve the quality of the platform in terms of all aspects.

**2. Project Description**

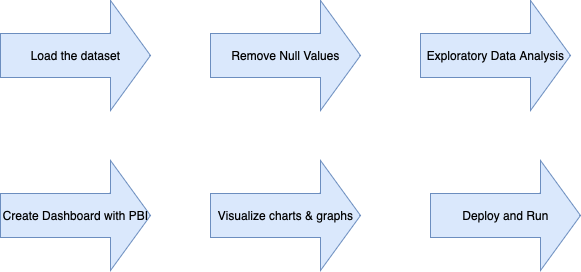
**2.1 Brief description of the project**

The goal of this project is to perform exploratory data analysis (EDA) on a dataset from Netflix and create an interactive dashboard using Power BI. The dashboard will provide valuable insights into last year's viewing trends, trending content, rating and other relevant metrics that can drive business growth for Netflix. The dataset contains information on Netflix shows and movies, including titles, genres, release dates, cast, directors, ratings, and viewership statistics.

During the EDA phase, trends in viewership over the past year will be analyzed, identifying popular genres and understanding viewing preferences during specific periods. The popularity of actors and directors will be assessed, exploring their influence on viewership and highlighting successful collaborations. Genre analysis, ratings, reviews, geographic analysis, and seasonal trends will also be examined.

Using Power BI, a visually appealing and interactive dashboard will be created to present the insights gained from the EDA. The dashboard will include visualizations such as line charts, bar charts, donut charts and more to effectively communicate the findings. It will showcase viewership trends, genre distribution, user ratings and reviews, and interactive filters for exploration.

The insights derived from the dashboard will drive actionable decisions for business growth. Recommendations may include optimizing content acquisition, personalizing recommendations, targeting marketing efforts, refining content strategies, and improving user retention.



*Fig 2.1.1 Flow of the project*

**2.2 Challenges and Technical Contributions**

The project faces challenges related to data quality, scalability, privacy, and analyzing complex relationships. The technical contributions include advanced data visualization techniques, analytics for viewership trends, integration of streaming data, and scalable data processing solutions. These contributions address new problems and offer innovative solutions to derive valuable insights and drive business growth in the context of Netflix. With the help of powerful tools like Power BI, we can visualize big data and source fruitful information out of it. We believe that a dashboard should be self-explanatory, and we aim to build a dashboard like the same.

**3. Background**

**3.1 Related Papers**

**1. Making Smarter Business Decisions with Big Data: A Netflix Case Study**

The author has articulately stated how Netflix gained valuable insights into viewer preferences by collecting and analyzing vast amounts of data, including user ratings, viewing habits, and customer engagement. Statistical techniques like Collaborative filtering and data analysis were employed to identify patterns and correlations within the dataset, while visualization techniques likely played a role in presenting insights in a visually appealing and easily understandable manner.

<http://oceansofdata.edc.org/making-smarter-business-decisions-big-data-netflix-case-study>

**2. Big Data Analytics for Business Coursework- Big Data Analytics in the OTT industry**

The author emphasizes on enhancing business growth through effective visualization and analysis of big data in the context of Netflix's user behavior and content strategies. The report highlights the use of statistical techniques, such as regression and classification, to achieve personalization and recommendation systems . It also emphasizes the importance of data visualization in understanding and presenting business insights. The visualization techniques highlighted in the coursework dashboards, graphs, infographics, maps, and charts.

<https://www.researchgate.net/publication/359510861_Big_Data_Analytics_for_Business_Coursework-_Big_Data_Analytics_in_the_OTT_industry_NETFLIX_case_study>

**3. Data Visualization, The Key to Netflix's success: An Analysis**

Data visualization plays a crucial role in decision-making processes. Therefore, the author highlights on how Netflix aims to understand customer preferences, personalize content recommendations, and improve the overall user experience by leveraging vast amount of user data. As a part of the visualization process complex data sets can be transformed into visually appealing and easily understandable charts, graphs, and interactive dashboards. These visual representations would help to gain actionable insights and make data-driven decisions.

<https://www.toucantoco.com/en/blog/data-visualization-the-key-to-netflixs-success-an-analysis>

**4. Data Visualization for Developing Effective Performance Dashboard with Power BI**

This research paper aims at exploring the use of Microsoft Power BI for data visualization. The paper highlights the importance of data visualization in understanding trends and patterns and its impact on decision making. Power BI is identified as a popular tool for analyzing data and provides different visualization methods such as stacked column charts, tables, and maps. The research focuses on creating a Power BI dashboard to connect different visual graphs and gain valuable insights.

<https://ieeexplore.ieee.org/document/10100169>

**5. What is the Importance of Power BI in Analytics, And How It Can Transform Your Business?**

The author articulates on how Power BI is a powerful tool that helps businesses grow by providing data visualization, real-time analytics, self-service analytics, data integration, collaboration and sharing capabilities, mobile accessibility, scalability, and flexibility. With Power BI, businesses can easily visualize and understand their data, monitor key performance indicators, empower business users to analyze data independently, integrate data from multiple sources, collaborate effectively, access data on the go, and scale the solution as their needs evolve. By leveraging these features, businesses can make informed decisions, gain actionable insights, and drive growth and profitability.

<https://medium.com/geekculture/what-is-the-importance-of-power-bi-in-analytics-and-how-it-can-transform-your-business-3c874d752db0>

**3.2 Software Tools**

**Jupyter Notebook:** For the current project we are using Jupyter as an IDE which is one of the most widely used IDE in the field of Data Science.

**Libraries Used:** Since the coding and most of the project are built using Python, here’s a brief list of the libraries that have been used during this entire project.

* NumPy
* Pandas
* Seaborn
* Matplotlib

**Power BI:** It is a business intelligence (BI) tool developed by Microsoft that allows users to visualize and analyze data from various sources. It provides interactive dashboards, reports, and data exploration capabilities, enabling organizations to gain insights and make data-driven decisions. It’s helpful in Data Analysis, Visualization and Integration.

**Kaggle:** The dataset hasn’t been sourced from an existing DBMS, whereas it has been pooled from one of the largest coding platforms which have an abundant amount of data available for data enthusiasts to learn and explore a wider horizon in the field of data science and data visualizations.

**3.3 Required Hardware**

There are four core requirements to meet the hardware requirements,

* A fast and reliable processor,
  + For Windows – Intel Core i5 and Above
  + For MacOS – Apple M1 or M2
* 8GB RAM (Random Access Memory) Preferred.
* 32GB of available Hard Disk Space.
* A strong internet connection.
* This project was built on a Mac OS where the processor is Apple 21 with 8 GB of RAM and 232 GB of Available Hard Disk Space.

**3.4 Related Programming Skills**

* Python has been formally used in the entire analysis.
* Libraries such as NumPy, Pandas, Seaborn, Matplotlib and Scikit Learn have been used to address the study.
* NumPy and Pandas are the base packages to build arrays, lists, dictionaries, data frames, data transformations, clean data etc.
* Seaborn and Matplotlib are the visualization packages which would be used in the analysis to develop a visual understanding of the attributes.
* Bar Charts, Box Plots, Histograms, Pie Charts, Donut Charts, Line Charts are a few of the visualization strategies that would be used in this project.

**4. Problem Definition**

**4.1 Mathematical Concepts**

While the project statement does not involve specific formal mathematical problems or definitions, we can provide some relevant mathematical concepts and definitions related to data analysis and visualization in the context of the project. Here are a few examples:

**Exploratory Data Analysis (EDA):** Exploratory Data Analysis is a statistical approach that involves summarizing and visualizing data to gain insights, identify patterns, and detect anomalies. It utilizes techniques such as descriptive statistics, data visualization, and graphical representations to understand the underlying structure and relationships within the dataset.

**Data Visualization:** Data visualization is the graphical representation of data to communicate information effectively. It uses visual elements, such as charts, graphs, and maps, to represent data patterns, relationships, and trends. Visualizations aid in understanding complex data, identifying patterns, and conveying insights in a visually appealing and intuitive manner.

While these mathematical concepts are relevant to the project, it's important to note that the project statement primarily involves data analysis and visualization techniques rather than formal mathematical problem statements.

**4.2 Challenges and tackling of problems**

**Data Quality and Cleaning:** The Netflix dataset may contain missing values, inconsistencies, or errors. Ensuring data quality and cleaning the dataset is crucial to avoid biases and inaccuracies in the analysis. It requires careful handling of missing data, addressing outliers, and ensuring data consistency. Implementing data cleaning techniques such as imputation methods for missing data, outlier detection and treatment, and standardization of data formats. Conduct thorough data validation and verification processes to ensure data quality.

**Visualization Design and Interpretation:** Designing effective visualizations that convey insights accurately and intuitively can be challenging. Presenting the data in a visually appealing and understandable manner requires thoughtful visualization design choices and proper interpretation of the visualizations. Utilizing principles of data visualization and visual design to create visually compelling and informative dashboards. Considering appropriate chart types, color schemes, and interactive elements to enhance understanding. Clearly annotating and providing context to aid interpretation and avoid misinterpretation of visualizations.

**Actionable Insights and Recommendations:** Deriving actionable insights and translating them into effective business recommendations can be challenging. The insights need to be meaningful, relevant, and actionable for driving business growth. Clearly defining the business objectives and align the analysis with the desired outcomes. Interpret the insights within the context of the business and identify specific actions or strategies that can be implemented based on the insights. Collaborate with relevant stakeholders to validate the recommendations and ensure their practicality and feasibility.

By recognizing and addressing these challenges, the project can ensure accurate and meaningful analysis, effective visualization, and actionable insights that can contribute to driving business growth for Netflix.

**4.3 Brief summary of general solutions**

Implementing data cleaning techniques to handle missing values, inconsistencies, and errors in the dataset. Designing visually appealing and interactive dashboards using tools like Power BI to present insights effectively. Utilize appropriate chart types, color schemes, and interactive elements to enhance understanding and engagement. Derive meaningful and actionable insights from the analysis, aligned with the business objectives. Translate insights into concrete recommendations that can drive business growth, considering factors such as content acquisition, personalization, marketing strategies, and user retention.

**5. The Proposed Techniques**

**5.1 Description of Data**

***show\_id*** – Unique ID which refers to each row in the dataset.

***type*** – Content type in Netflix.

***title*** – Title of the content.

***Director*** – Director of the content.

***Cast*** – Cast in the content.

***Country*** – Where the content is released.

***date\_added*** – Date when it uploaded in Netflix.

***release\_year*** - Year when the content is released.

***Rating*** – Rating for the content.

***Duration*** – Duration of the content.

***listed\_in*** – Genre of the content.

***Description*** – Brief description about the content.

**5.2 Details of Major Techniques**

This section defines the techniques, such as graphs, and methods used to provide a comprehensive understanding of the data and the work done.

* **Count Plot:** Plots a bar chart based on the count of observations in each attribute.
* **Histogram:** Provides the distribution of numerical values in a feature, enabling us to determine if that feature follows a normal distribution or not.
* **Box Plot:** Useful for identifying outliers in the feature, it also provides information such as mean, maximum, and minimum values in that column.
  1. **Exploratory Data Analysis**

In the first step of Exploratory Data Analysis (EDA), we will import the important libraries used in the process.

**A computer screen with text and images

Description automatically generated**

* Load the Netflix dataset into the Python environment and have a glimpse of how the data looks.

**A black background with white text

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

* Next, have a look at the number of rows and columns, column names and the datatype of it.

**A screenshot of a computer

Description automatically generated**

* Check the summary statistics of the columns and for duplicates.

****

**A black screen with green text

Description automatically generated**

* In summary, the dataset has 8807 rows, 12 columns. *id* and *description* columns are useless so we can drop them. We have null values, so we need to clean the dataset. There are no full duplicates in the dataset.
  1. **Cleaning Dataset**
     1. **Cleaning Country Column**
* The *Country* column in the dataset has data inconsistency as the countries are not separated with a special character.
* We solve this inconsistency by using ‘,’ to separate the countries and have unique names for all countries.

**A screenshot of a computer program

Description automatically generated**

* + 1. **Cleaning Rating Column**
* The values in the *Rating* column have inconsistent values such as ’66 min’, ’84 min’ and more, so we replace that with NaN values.

A screenshot of a computer program

Description automatically generated

* + 1. **Cleaning Director and Cast columns**
* *Director* column has null values in case of director name is not known. So, we introduced a new category ‘Unknown’ to define where a director name is not known we use this category.
* Similarly in *Cast* column, we face the same issue, so we used the same ‘Unknown’ column.
* *Date Added* column has both month and year together. We split that into *month* and *year* columns so that it will be able to use the data without any dependency.

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

* In conclusion, we have separated the countries in country column. There are wrong values in rating column, so we replaced it with right values. There are a lot of null values in director we impute the null values with unknown. I used mode() to impute null values with most frequent values in '*country*', *'date\_added*', '*rating*', '*duratio*n'. I added month and year columns to the data set. Now the dataset has no null values.
  1. **Data Visualization**

We have cleaned the dataset; with that we did some interpretation with the data having a few datasets related questions. We have used Python to visualize these questions to have an overview on the dataset.

* + 1. **Which content type has the highest ratio?**
* Netflix dataset has two types of content namely Movies and TV Shows. To have a clear picture on which contest is more in the platform we did a bar chart to visualize it.

**A screenshot of a computer screen

Description automatically generated**

* It’s clearly seen that the number of movies is approximately twice the number of TV shows. In this it is evident that making more content in movies will be helpful for the business growth than investing more in TV shows.
  + 1. **Which months have least releases?**
* To see which months, have least releases, we have used the heatmap to have an overview.

**A screen shot of a computer

Description automatically generated**

**A graph of different colored squares

Description automatically generated**

* Based on the last complete year 2020 the best months to add new content are February, March, and May.
  + 1. **Which genre of Movies or TV shows has more content?**
* Next, we checked on what genre has more content in Movies and TV shows.

**A screenshot of a computer

Description automatically generated**

**A screen shot of a computer

Description automatically generated**

**A screen shot of a computer program

Description automatically generated**

**A graph of a number of squares

Description automatically generated**

* In movies, International movie genre has more content.

**A screenshot of a computer program

Description automatically generated**

**A graph showing a bar graph

Description automatically generated**

* In TV shows, International TV shows has more content.
  + 1. **In which year number of movies or TV shows has been released?**
* We have used lineplot and countplot to see in which year a greater number of movies and TV shows are released.

**A screen shot of a computer code

Description automatically generated**

**A graph of a number of years

Description automatically generated**

**A black background with blue text

Description automatically generated**

**A graph showing the time of a year

Description automatically generated**

**A black background with colorful text

Description automatically generated**

**A graph of growth in different colors

Description automatically generated**

* We can see a clear spike in movies in 2018 and TV shows in 2020. During Covid, people has started watching more of TV shows than movies.
  + 1. **Which country has the most release?**
* To have an overview of the demographics, we picked up this question to check which country has a greater number of releases.

**A computer screen with colorful text

Description automatically generated**

**A blue pie chart with white text

Description automatically generated**

* United States holding most of the release with more than 60% with India in the second place with almost 20%.
  + 1. **What is the most popular rating in Movies or TV shows?**
* With respect to ratings, we checked on which rating is the most popular among the two content types.

**A computer screen with colorful text

Description automatically generated**

**A graph of a bar graph

Description automatically generated**

* Movies and TV shows TV-MA or in simple terms Adults has the most popular rating.
* With covering almost every column and interpreting with it, we can see a clear picture of how the data is and how it behaves.

**6. Visual Applications**

**6.1 GUI Design**

Netflix Dataset Dashboard has been designed using Microsoft Power BI. The dashboard aims to provide insightful visualizations and analytics based on the Netflix dataset while maintaining the aesthetic appeal consistent with the Netflix brand.

**A screenshot of a computer screen

Description automatically generated**

**Data Integration:** The dashboard integrates the Netflix dataset, which includes information about movies, TV shows, genres, release dates, ratings, and more.

**Visualizations:** The dashboard features a variety of visually appealing charts and graphs to showcase key insights. Some of the visualizations include bar charts, stacked bar charts, stacked area charts, and donut charts.

**Filters and Interactivity:** Users can interact with the dashboard using intuitive filters and slicers to analyze data based on specific criteria, such as content type and year.

**Netflix Brand Palette:** The color palette used in the dashboard is carefully chosen to align with the colors associated with the Netflix brand. This consistency ensures a cohesive user experience with a familiar look and feel.

**Performance Metrics:** The dashboard includes key performance metrics, such as the top genre that has performed, trend by title type, content percentage in the platform and ratings.

**User-friendly Interface:** The user interface is designed to be intuitive and user-friendly, making it easy for both analysts and casual users to explore the data effortlessly.

**6.2 Design Modules**

Let’s break down the dashboard and explore what’s in the design.

**Title Type**

This is a dropdown created using slicers where the user can choose the content type either movies or TV shows in general.

**A white background with black lines

Description automatically generated**

**Year**

This is created using slicers with the filter condition ‘between’ to slide through the years and see what happened during the years.

**A red line on a white background

Description automatically generated**

**Top 10 Genre**

Top 10 Genre describes which genre performed the more and an aggregate function is used to sum it up and the data values is shown. This will help in promoting the business to produce more interesting content in least liked genre.

**A red and white text on a white background

Description automatically generated**

**Content Percentage**

This described how much ratio of contents we have in movies and TV shows. This can improve in making critical business decisions in producing new contents.

**A red circle with numbers and a white background

Description automatically generated**

**Trend by Title Type**

This shows when did the contents viewership peaked. As you can see during 2020 TV shows are at the highest as people were more interested in watching TV shows than movies.

**A graph showing a number of different types of data

Description automatically generated**

**Rating**

This says which rated movies are liked by people and helps in taking business decisions accordingly.

**A graph of red squares

Description automatically generated**

**7. Experimental Evaluation**

**7.1 Description of Dataset**

The Netflix dataset obtained from Kaggle, containing information about Netflix shows and movies, including titles, genres, release dates, cast, directors and ratings.

**7.2 Competitors**

**Baseline Method:** A basic approach for viewership trend analysis and content popularity assessment could be comparing overall viewership trends and top-rated content based on simple metrics like total views and average ratings.

**Existing Techniques:** Existing algorithms or machine learning models for recommendation systems or predictive analytics can be considered as competitors.

**7.3 Parameter Settings**

**Data Cleaning Parameters:** Deciding on specific techniques for handling missing values, outlier treatment, and data transformation. For instance, mean imputation for missing values, Z-score normalization, or Min-Max scaling for standardization.

**Visualization Parameters:** Determining the types of visualizations to use for different analyses, such as line charts for viewership trends, bar charts for genre distribution, and bubble charts for actor/director popularity.

**7.4 Evaluation Measures**

**Recommendation Metrics:** For personalized recommendation systems, evaluate their effectiveness using metrics such as Click-Through Rate (CTR), Conversion Rate, or Precision-Recall metrics.

**Business Impact Metrics:** Assessing the impact of the dashboard insights on business growth metrics, such as user acquisition, retention rates, or revenue growth, to measure the effectiveness of the strategies implemented based on the dashboard's recommendations.

**8. Future Work**

Future work for this project could involve further enhancing the dashboard and analysis, exploring new dimensions of data, and expanding the scope of insights and recommendations. Here are some potential areas of future work:

**Real-time Data Integration:** Currently, the dashboard may be based on historical data. Future work could involve integrating real-time data streams to provide up-to-date insights on viewership trends and user interactions.

**Machine Learning for Content Production:** Explore the use of machine learning for content production. Analyze successful shows and movies to identify patterns and attributes that contribute to their popularity, assisting in the creation of new content that aligns with viewers' interests.

**Content Curation Optimization:** Use data-driven approaches to optimize content curation. Analyze the success of curated collections and make data-informed decisions on the content to be included in these collections.

**Geo-location Analysis:** Perform a more in-depth geographic analysis to understand regional viewing preferences and adapt content offerings for different markets.

**User Engagement Analysis:** Dive deeper into user engagement metrics to identify factors that influence high user engagement, such as binge-watching behavior, and tailor strategies to encourage continuous content consumption.

These future work ideas aim to further enrich the analysis, improve personalized recommendations, optimize content strategies, and enhance the dashboard's capabilities to drive business growth and provide a better user experience for Netflix subscribers.

**9. References**

1. <http://oceansofdata.edc.org/making-smarter-business-decisions-big-data-netflix-case-study>
2. <https://www.researchgate.net/publication/359510861_Big_Data_Analytics_for_Business_Coursework-_Big_Data_Analytics_in_the_OTT_industry_NETFLIX_case_study>
3. <https://www.toucantoco.com/en/blog/data-visualization-the-key-to-netflixs-success-an-analysis>
4. <https://ieeexplore.ieee.org/document/10100169>
5. <https://medium.com/geekculture/what-is-the-importance-of-power-bi-in-analytics-and-how-it-can-transform-your-business-3c874d752db0>