

Netflix Data Analysis — Interview Notes

1. Business Problem

The goal of this project was to understand what kind of content Netflix focuses on and how it has evolved over time. Questions I wanted to answer: - Does Netflix have more movies or TV shows? - Which countries produce the most content? - What type of ratings are most common? - How has content production changed over the years? These insights help Netflix make business decisions about which countries, genres, or content types to invest in.

2. Full Project Explanation (Storytelling – 2–3 Minutes)

I worked on a project called Netflix Data Analysis, which I did completely using SQL. I started with the Netflix dataset from Kaggle, which includes details of movies and TV shows — like title, director, country, cast, release year, and rating. I chose this dataset because it's real-world data and perfect to practice data cleaning, SQL queries, and storytelling. First, I loaded the dataset into MySQL using the import wizard and created a table called 'netflix'. Then I focused on data cleaning — I found missing values in columns like director, cast, and country. I replaced those with 'Unknown' or 'Not Specified'. I also checked for duplicate titles and removed them, and I used TRIM() to clean unwanted spaces. After cleaning, the data was ready for analysis. Next, I performed data analysis using SQL. I divided it into two levels — basic and intermediate. In the basic analysis, I found: - The total number of titles on Netflix - The count of movies vs TV shows - Number of titles released per year - Top countries producing the most content - Most common content ratings In the intermediate analysis, I explored: - Which directors had the most titles - Which countries focus more on TV shows - What were the major content trends over the years All this was done using simple SQL commands like GROUP BY, COUNT(), ORDER BY, and HAVING. I didn't use Python or window functions — everything was handled through SQL queries. From my analysis, I found that: - Netflix has more movies than TV shows, but TV shows are growing faster. - The United States and India are top content-producing countries. - Most content is rated TV-MA, meaning it's mainly for adults. - The years 2018–2020 had the highest number of releases, showing Netflix's fastest growth period. These findings give Netflix a clear idea of what's trending and where their content strength lies. From a business perspective, such analysis helps Netflix understand what type of content people prefer, which countries perform best, and where to invest more. Personally, I learned how to: - Clean and prepare raw data in SQL - Handle missing and duplicate values - Write efficient queries - Convert raw numbers into meaningful insights It also helped me improve how I communicate technical results in a simple story. In the future, I'd like to add IMDb ratings, create Power BI dashboards, and apply advanced SQL features like window functions to get deeper insights. So overall, this project taught me how to take real-world data, clean it, analyze it, and tell a clear story using SQL — just like how it's done in real data engineering or analytics roles.

3. Short Summary (1-Minute Version)

I worked on a project called Netflix Data Analysis using SQL. The main business goal was to find what type of content Netflix focuses on — like whether it has more movies or TV shows, which countries produce the most content, and how those trends changed over time. I used the Netflix dataset from Kaggle, cleaned it by handling missing values, removing duplicates, and ensuring consistency. Then I performed analysis using SQL queries with GROUP BY, COUNT(), and ORDER BY. I found that Netflix has more movies than TV shows, but TV shows are growing quickly. The US and India are top producers, and most content is adult-rated (TV-MA). Between 2018 and 2020, Netflix added the most titles. This project helped me strengthen my SQL, data cleaning, and storytelling skills — and understand how to generate business insights from real data.