Visualize data with QuickSight

Having previously visualized datasets using Tableau, I was eager to explore AWS QuickSight to understand its capabilities and performance. I found QuickSight to be remarkably handy and convenient, offering a more user-friendly experience for newcomers compared to Tableau. This exploration has enhanced my skills in data visualization and cloud-based analytics.

Introduction

What is Amazon QuickSight?

Amazon QuickSight is a cloud-based business intelligence platform that allows users to create interactive dashboards and visualizations. It's ideal for analyzing large datasets and quickly sharing data-driven reports, enabling organizations to make informed decisions efficiently.

How I used Amazon QuickSight in this project

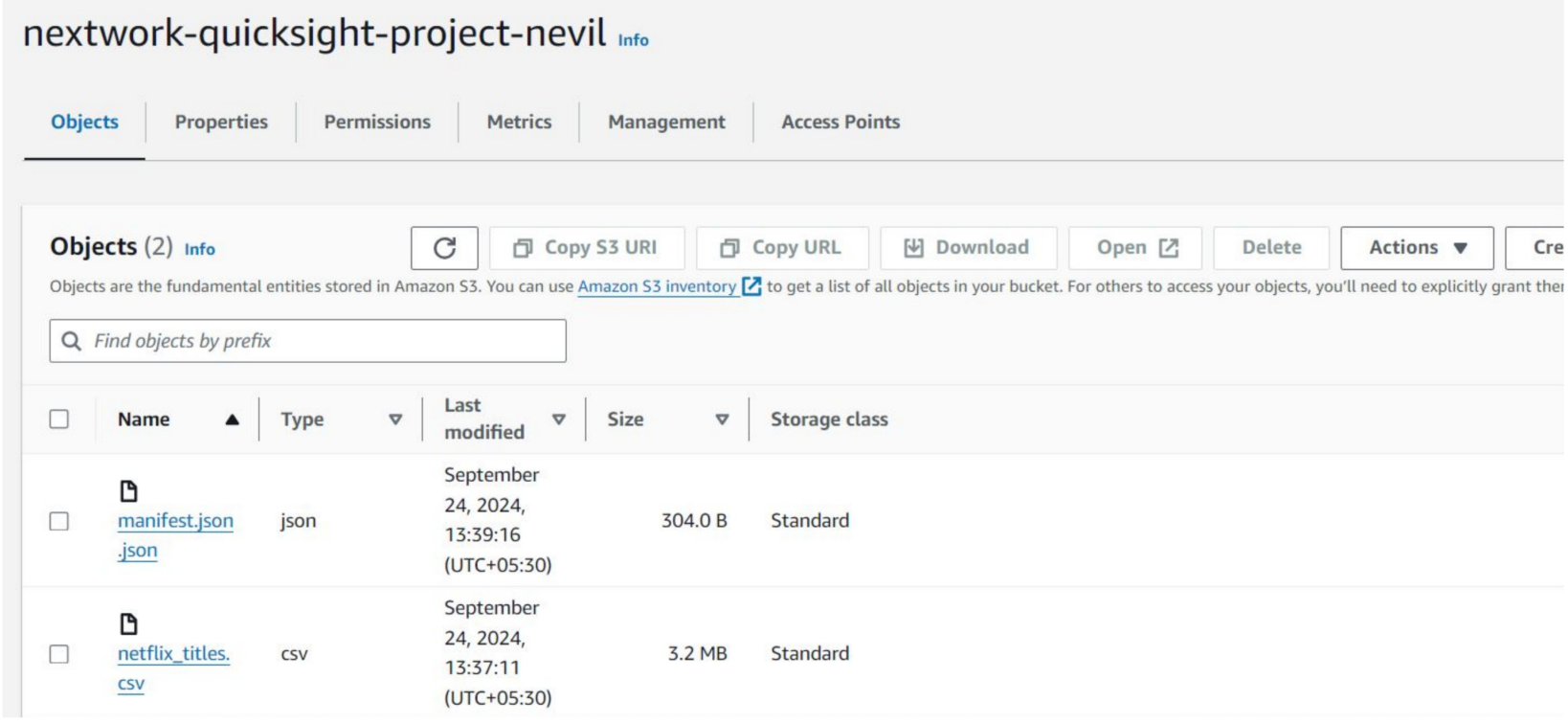
In today's project, I utilized Amazon QuickSight to connect to an S3 bucket containing Netflix data. I created various visualizations, including donut and bar charts, to analyze content by release year and type. Additionally, I applied filters for more targeted insights and successfully exported the results for further use.

One thing I didn't expect in this project was...

I was pleasantly surprised by how effortless it was to connect data from an S3 bucket to Amazon QuickSight and promptly generate informative visualizations. The integration process proved to be more intuitive and user-friendly than I had initially expected.

Uploaded project files into S3

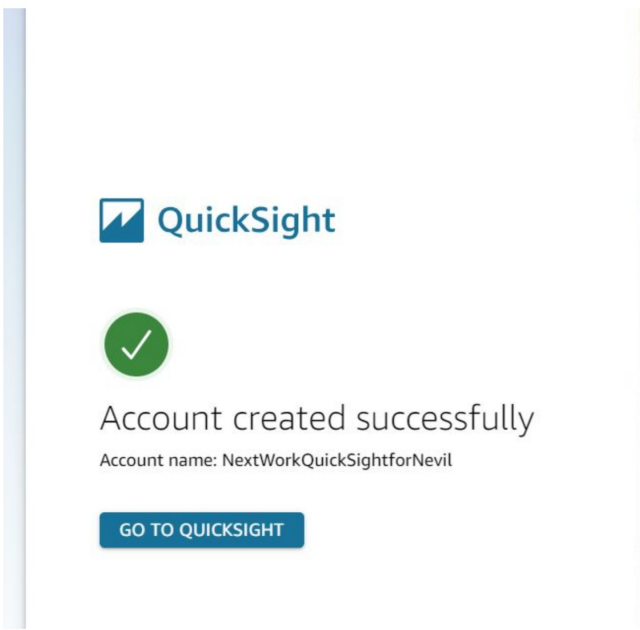
S3 stores two key files: netflix_titles.csv, a detailed Netflix dataset with title, release year, and genre information, and manifest.json, a metadata file outlining the project's configuration.
I updated the manifest.json file by modifying the CSV file's URL to accurately point to netflix_titles.csv in the S3 bucket. Editing this file is crucial, as it defines the data source; an outdated S3 URI could lead to data retrieval issues.

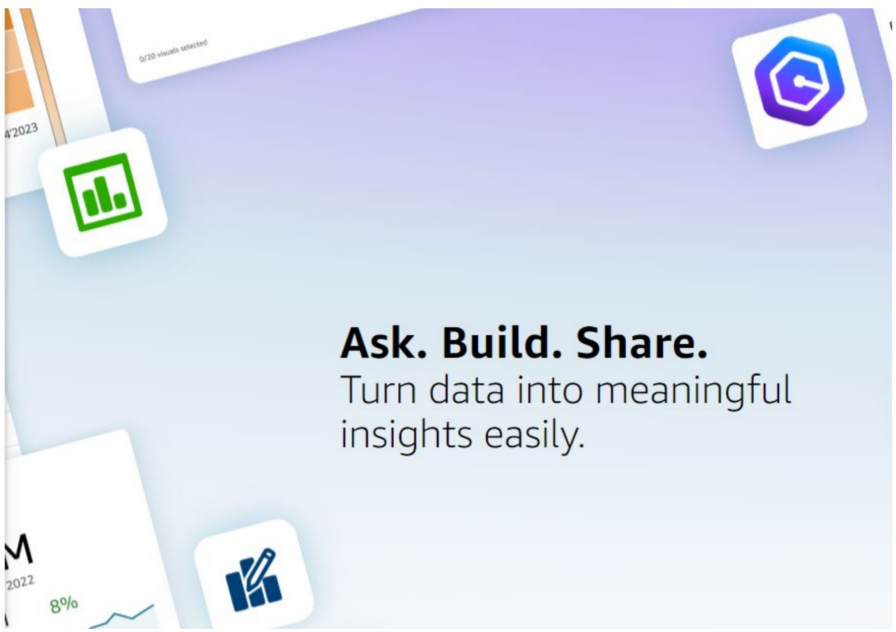


Creating QuickSight account

Creating an Amazon QuickSight account is quick and easy, taking just a few minutes. While there is a fee for the account, new users can take advantage of a free trial. After the trial, pricing options include a pay-per-session model or a subscription plan.

To set up your account, simply sign in with your AWS credentials, provide basic information, choose the desired QuickSight edition, and configure data access permissions. The process is designed for efficiency, allowing you to get started without hassle.





Connecting QuickSight to S3 Bucket

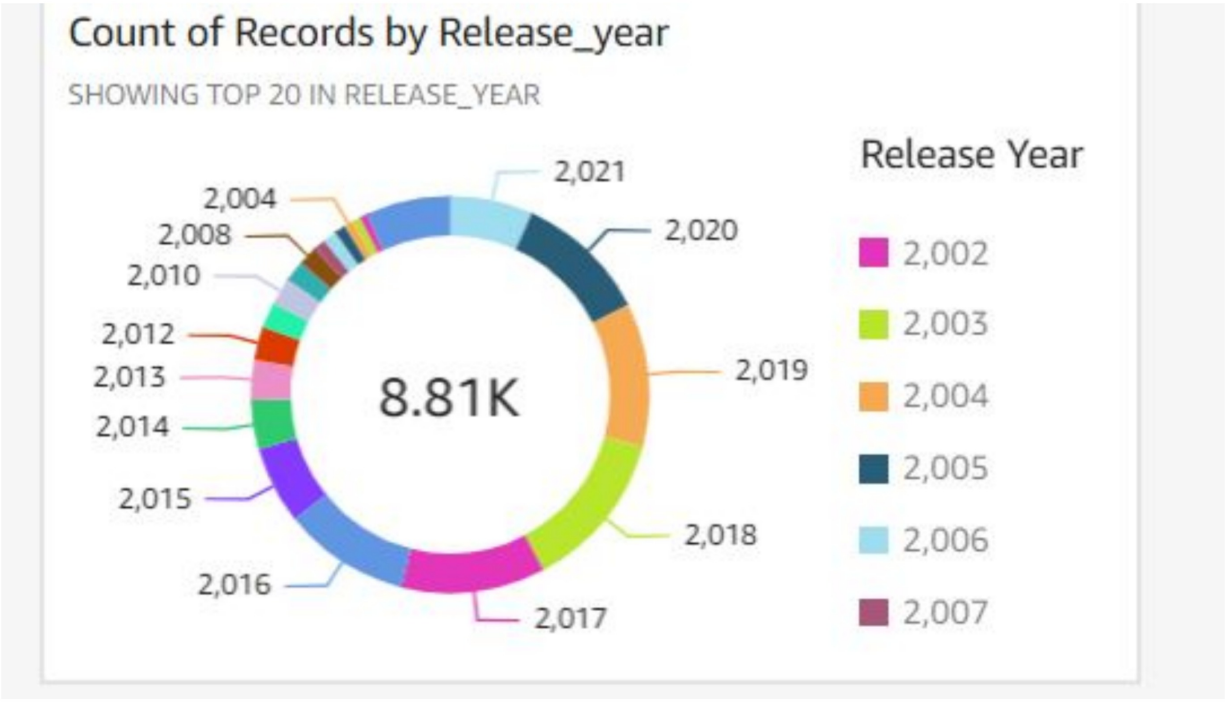
I connected the S3 bucket to QuickSight by visiting the Datasets page. From there, I clicked on "New Dataset" selected S3 as the data source, and entered the S3 bucket details.

The manifest.json file acts like a map for Amazon QuickSight, guiding it to the location of your data files and detailing how they are organized. It also describes the structure of each piece of data, enabling QuickSight to interpret and visualize the information accurately in charts and graphs. Without this map, QuickSight could become confused and fail to display your data correctly.
This is why updating the URI is so critical; an outdated path would lead QuickSight to the wrong file.

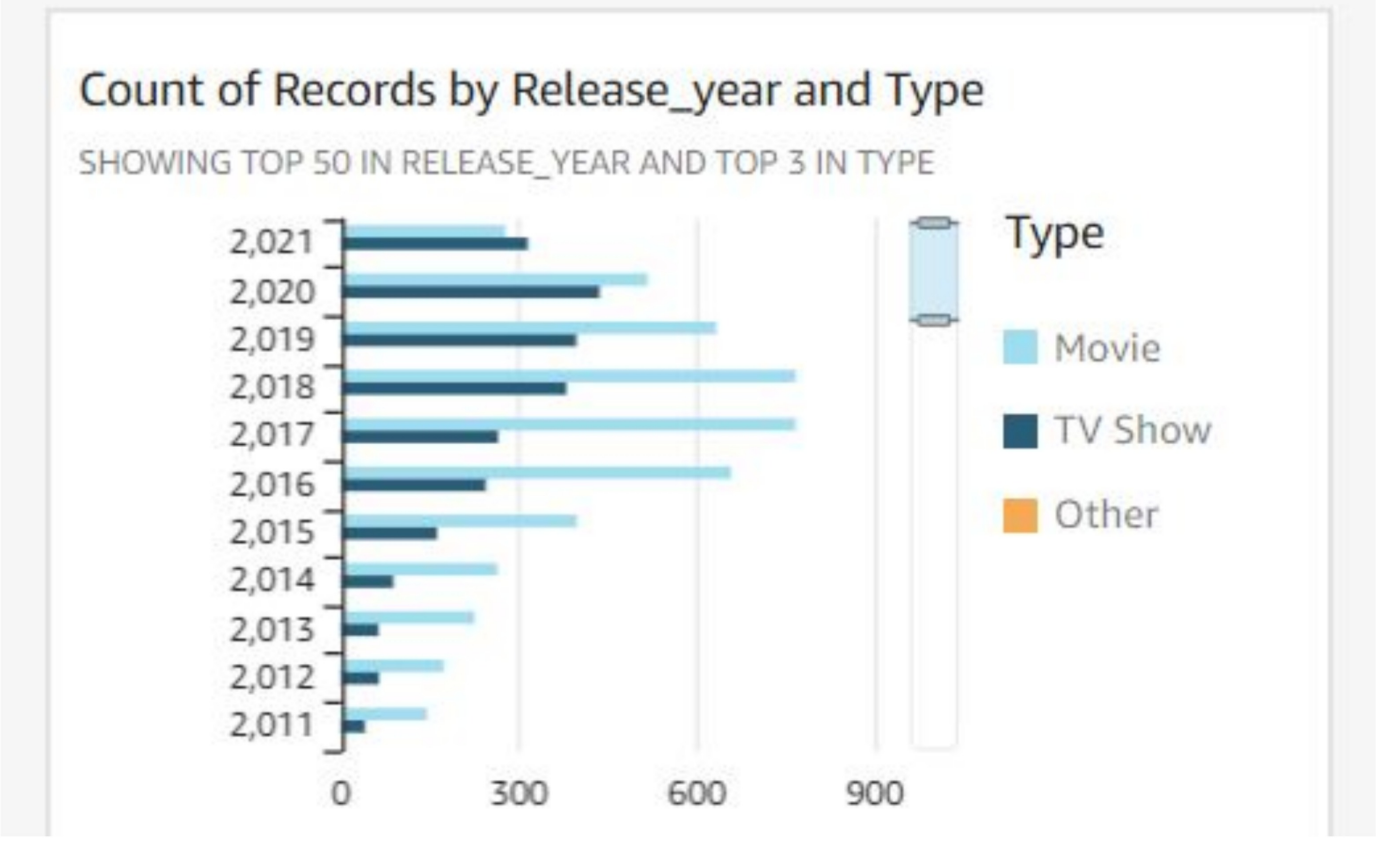
First visualization

To create visualizations on QuickSight, I first imported the dataset, then selected the fields I wanted to analyze. I chose the visualization type (e.g., bar chart, pie chart), customized it with filters, and adjusted settings to explore and present.

The chart shown here is a breakdown of Netflix content by release year and type. The release year donut chart displays the proportion of movies and TV shows released each year. The horizontal bar chart shows the count of movies versus TV shows.



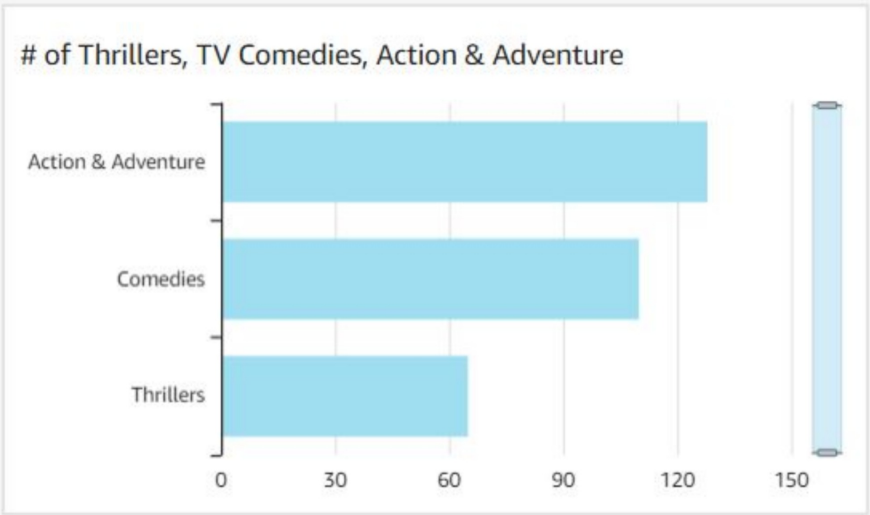
I created this graph by dragging and dropping release year onto the Y-axis and type into the Group/Color field. This setup allowed me to visualize the distribution of movies and TV shows by release year, with different colors representing the content.

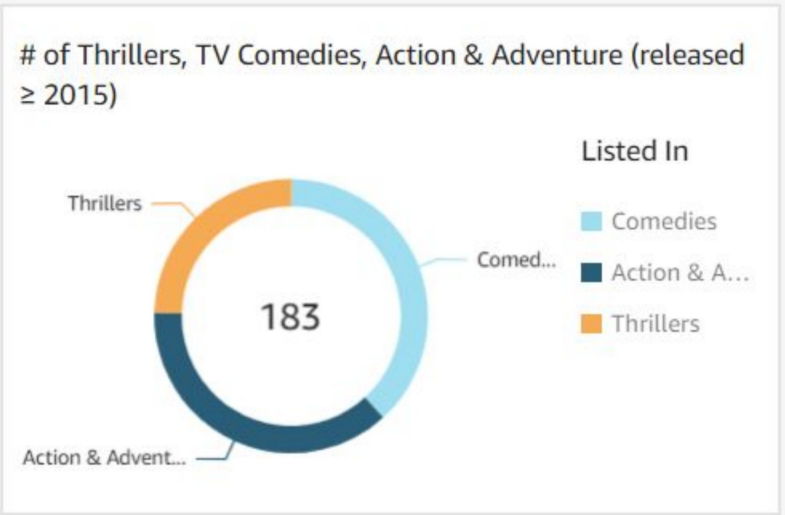


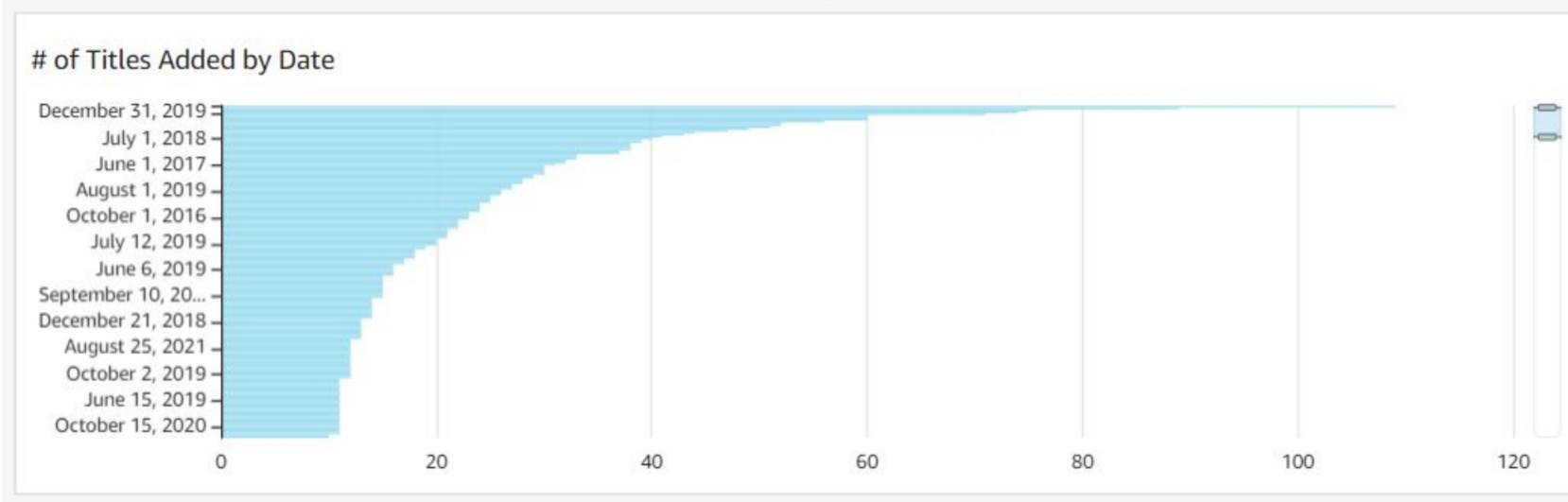
Using filters

Filters are useful for refining large datasets by focusing on specific criteria, such as dates, categories, or conditions. They help eliminate irrelevant data, highlight key trends, and provide more accurate insights, making analysis more precise.

This visualization breaks down Netflix content by release year and type. I added a filter by release year to focus on specific time periods, enabling us to analyze how the number of movies and TV shows has changed over time.







Setting up a dashboard

As a finishing touch, I adjusted the formatting of the visualizations, ensuring consistency in colors, labels, and titles. I also applied relevant filters, reviewed the layout for clarity, and added descriptive titles.

I did this by navigating to the dashboard view in QuickSight, clicking on the Share dropdown menu, and selecting Export to PDF. This option allowed me to download a PDF version of the dashboard for easy sharing and presentation.

Final Dashboard