



Visualize data with QuickSight



Nikhila Reddy



Introducing Today's Project!

In this project, I will demonstrate how to use Amazon Quicksight to analyze Netflix data and generate visualizations and Insights! I'm doing this project to learn how to use cloud data services for data analysis.

Tools and concepts

Services I used were QuickSight and S3. Key concepts I learned include using manifest.json files, creating visualizations with charts and filters, and how to refresh data in QuickSight to keep dashboards up to date.

Project reflection

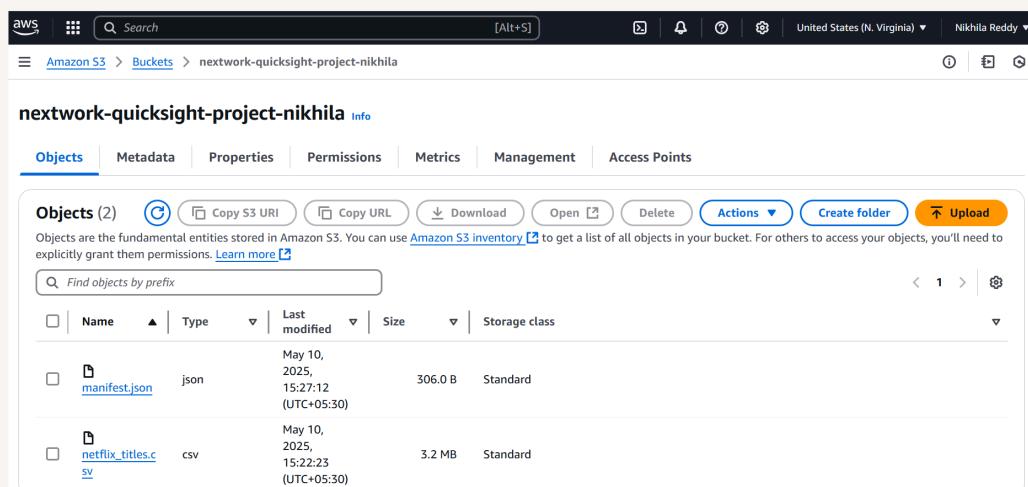
This project took me approximately 2 hours including demo time! The most challenging part was understanding how a manifest.json file works. The most rewarding part was generating a PDF of the finished visualizations to share insights.

After this project, We plan to work on day 3 of the AWS Beginners challenge. we will start this project on Monday.

Upload project files into S3

S3 is used in this project to store two files: manifest.json (which tells QuickSight about the structure and format of the data I am analyzing) and netflix_titles.csv (which is the raw data I'll be looking at today).

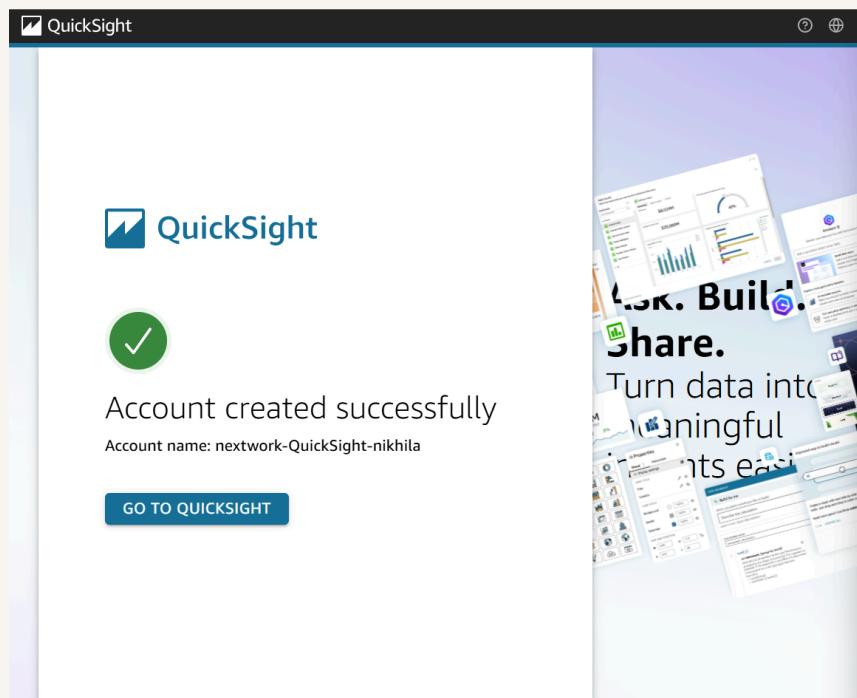
I edited the manifest.json file by updating the S3 URL to match my dataset's location. This step is important because QuickSight uses it to find the data—if it's not updated, QuickSight won't locate the dataset correctly and it will cause an error.



Create QuickSight account

Creating a QuickSight account cost \$0 as it comes with a 30 day free-trial! Make sure to UNCHECK an add-on in the sign up flow called Pixel-Perfect reports to avoid getting charged.

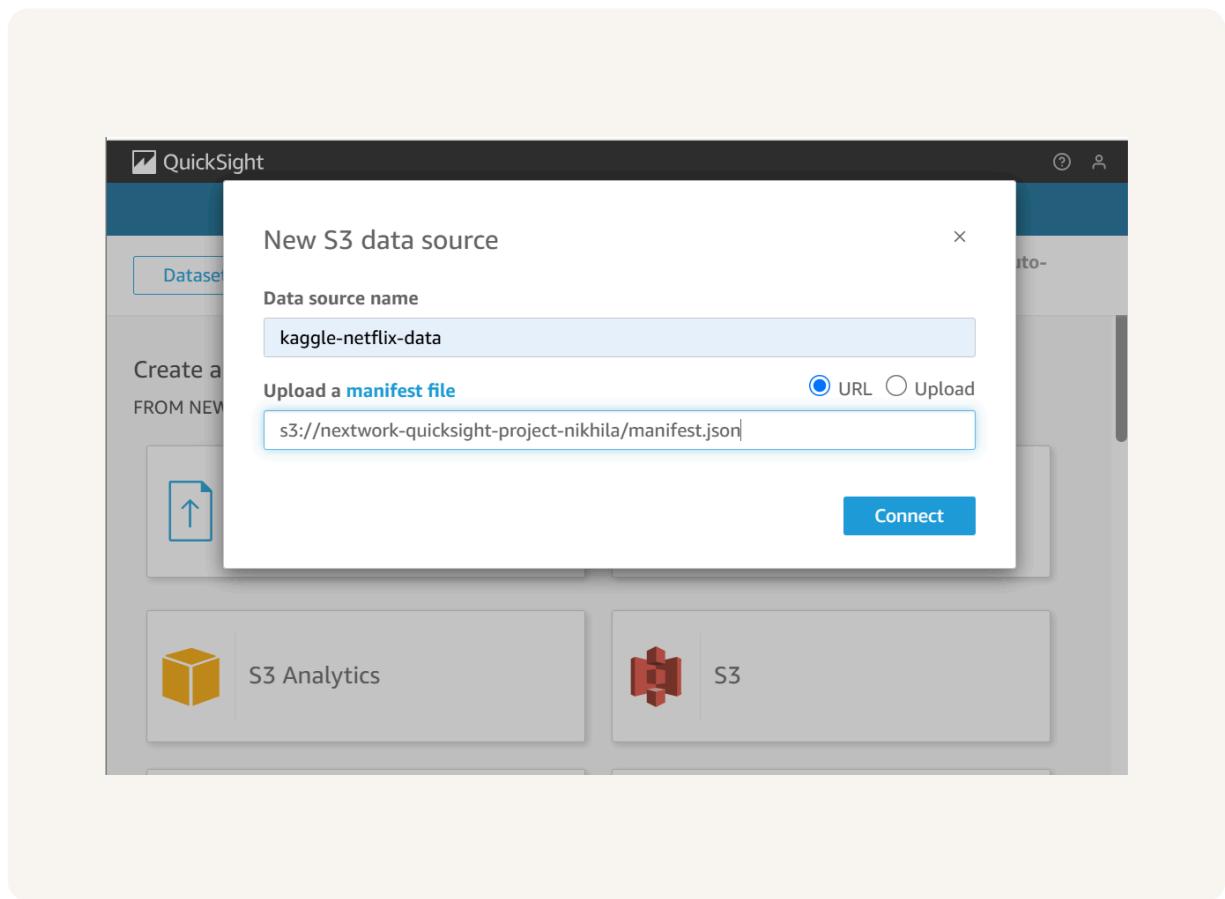
Creating an account took me about 5 mins including setting up the S3 bucket permissions.



Download the Dataset

I connected the S3 bucket to QuickSight by visiting the datasets page. There were so many options for data sources we could connect to, like databases and external tools (even platforms like Salesforce), and we selected S3!

The manifest.json file was important in this step because it tells QuickSight how to read the data. It explains that we're uploading a .CSV file and using commas as separators—without it, QuickSight might not understand how to process the data.



My first visualization

To create visualizations on QuickSight, I simply had to click on data fields like release_year and it auto-generated a chart that fits the data. You can also drag fields to "Group by" or "Y-axis" to control how the chart shows the data.

The chart/graph shown here is a breakdown of release years of Netflix content—basically, how many shows or movies came out each year. There are over 8800 titles in total, and 2019 stands out as the year with the highest number of releases.

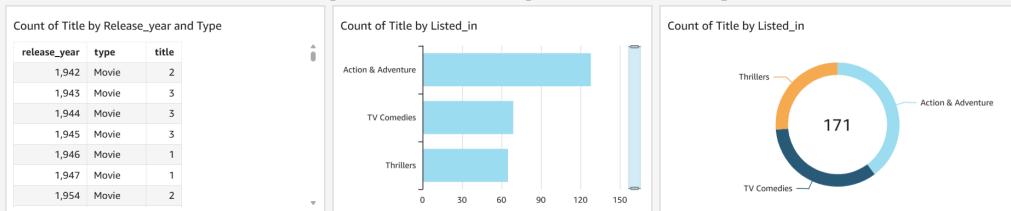
I created this graph by simply clicking on the "release_year" data label and switching the default bar chart to a donut chart. I also added a bar chart version to compare how the same data looks in different visual



Using filters

Filters are useful for narrowing down our data to the part we want to focus on, and in this case, we used them to zoom in on specific categories. We also added a filter to only show content released from 2015 and beyond.

This visualization is a breakdown of TV shows and movies in three categories—Action & Adventure, TV Comedies, and Thrillers. I added a filter using the "listed_on" label so only these categories appear in the chart for focused analysis.



Setting up a dashboard

As a finishing touch, we updated the titles of the charts in our dashboard to make them easy to read. The default names just showed data labels, but the new titles clearly explain what each chart is about and what insights it's showing.

Did you know you could export your dashboard as PDFs too? I did this by selecting publish and then "Generate PDF" on the top right hand corner of the QuickSight analysis.





nextwork.org

The place to learn & showcase your skills

Check out nextwork.org for more projects

