Visualize Data with QuickSight

Karen Wong

linkedin.com/in/wongkayankaren

Powered by NextWork.org

Introducing Today's Project!

What is Amazon QuickSight?

Amazon QuickSight is a cloud-based business intelligence tool that helps you visualize and analyze data easily. It turns complex data into interactive dashboards and reports, making it simple to gain insights and share them with others. It's fast, scalable, and user-friendly.

How I used Amazon QuickSight in this project?

I utilized Amazon QuickSight to retrieve the dataset I uploaded to S3. I then crafted an interactive dashboard, which allowed me to visualize and analyze the data for deeper insights.

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

I received unexpected bills from QuickSight, even though I followed the project instructions to uncheck the optional upgrade. I recall encountering a disconnection error when trying to retrieve files from S3, despite having configured the related settings. I suspect that the settings I applied when creating my QuickSight account were not retained.

This project took me...

I spent a total of 4 hours on this project: 1.5 hours following project instructions, 1.5 hours exploring additional features including Calculated Fields, Insights and Generative BI, and 1 hour delving into official documentation and refining my work here.

in Karen Wong
linkedin.com/in/wongkayankaren

Architecture Overview

Visualize Data with QuickSight
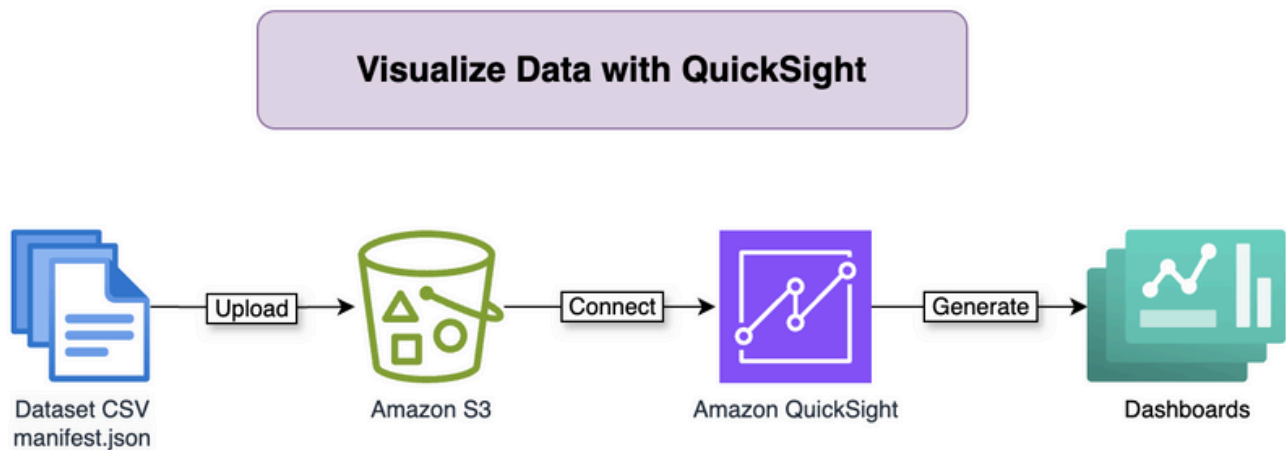
A dataset in CSV format, along with a "manifest.json" file, are uploaded to Amazon S3. The files are then connected to Amazon QuickSight, which ultimately generates interactive dashboards, presenting the data as visualizations like charts and graphs.

Karen Wong
linkedin.com/in/wongkayankaren

Upload Project Files into S3

S3 is used in this project to store two files: "netflix_titles.csv" and "manifest.json". The CSV file contains Netflix show data, including show_id, type, title, and director, while the JSON file contains metadata such as file format, delimiter, and S3 URIs.

I updated the "manifest.json" file by changing "URIs" to the S3 URI of my uploaded CSV file. Editing this file is important because it links the data in the S3 bucket to QuickSight and specifies how to read data in the file.
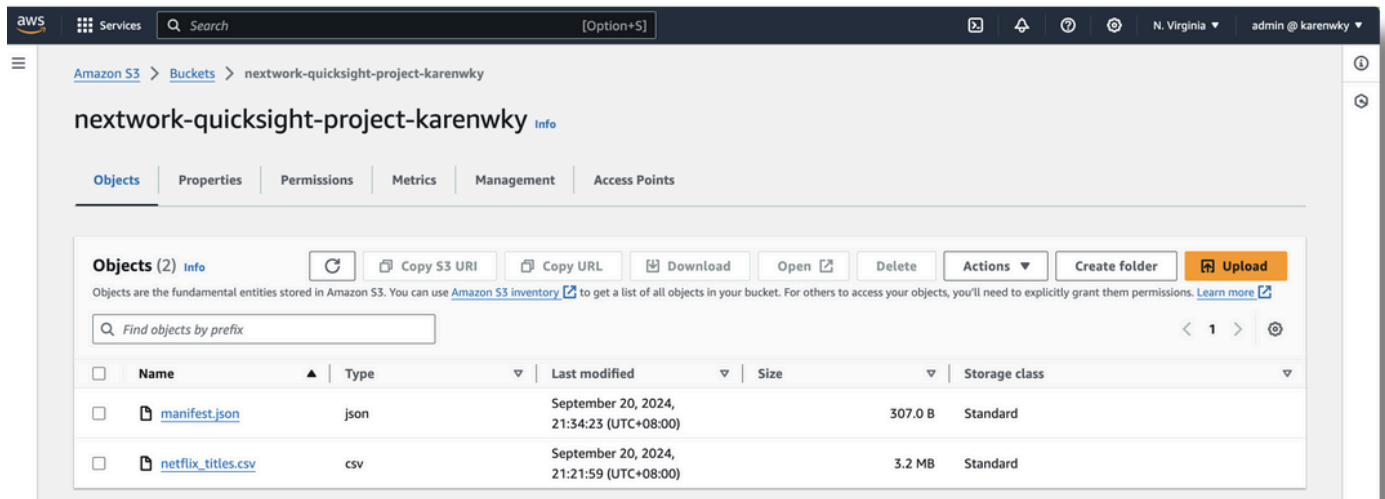
Create QuickSight Account

Amazon offers a 30-day free trial of the Amazon QuickSight Enterprise Edition for up to 4 users per account. It is priced by user role, providing flexibility and simplicity to choose the pricing model that best suits your business needs.

Creating a QuickSight account took me around one minute, and the process was straightforward and user-friendly.
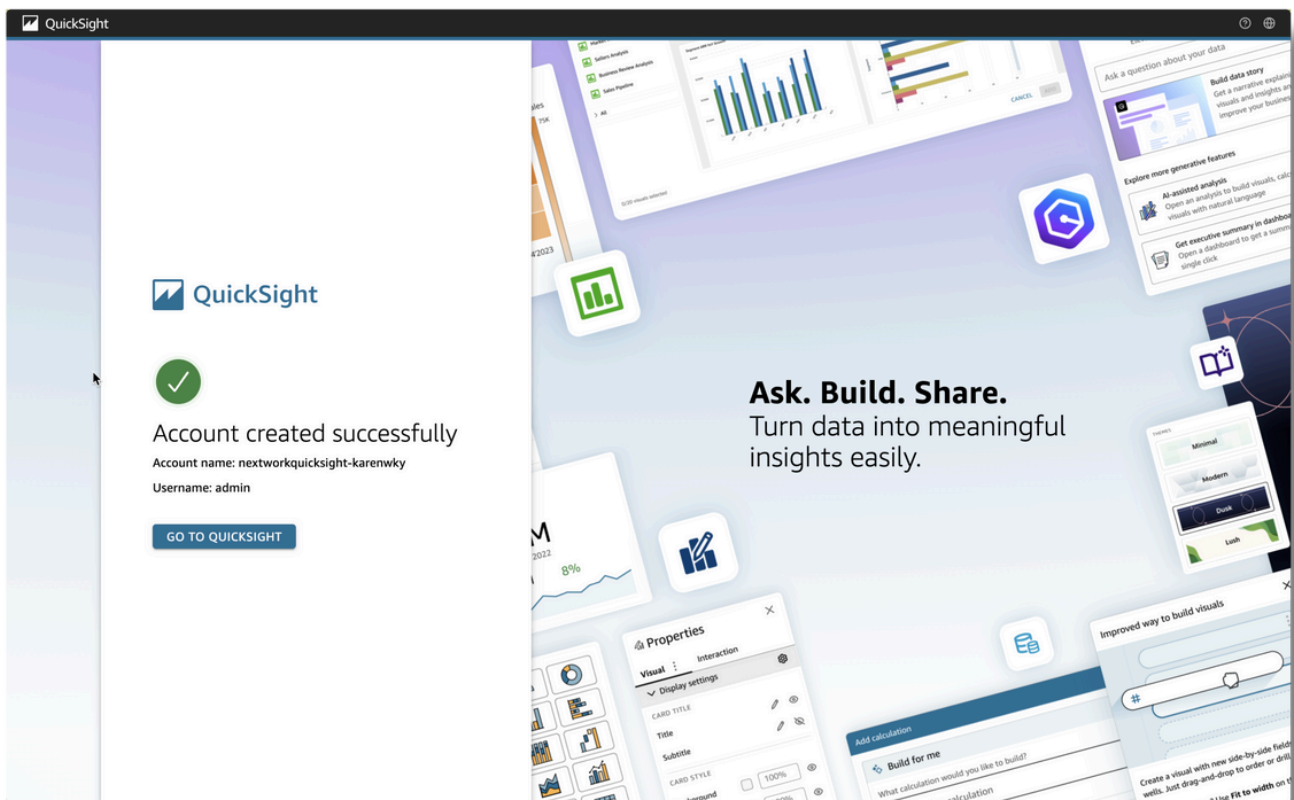
in **Karen Wong**

linkedin.com/in/wongkayankaren

Connect the Dataset

I connected the S3 bucket to QuickSight by visiting the "Datasets" page, clicking "New dataset", selecting "S3", and then entering the "Data source name" along with the S3 URI of the uploaded "manifest.json" file.

The "manifest.json" file was important in this step because it serves as the connection between the uploaded CSV file in the S3 bucket and QuickSight, allowing data retrieval for analysis.
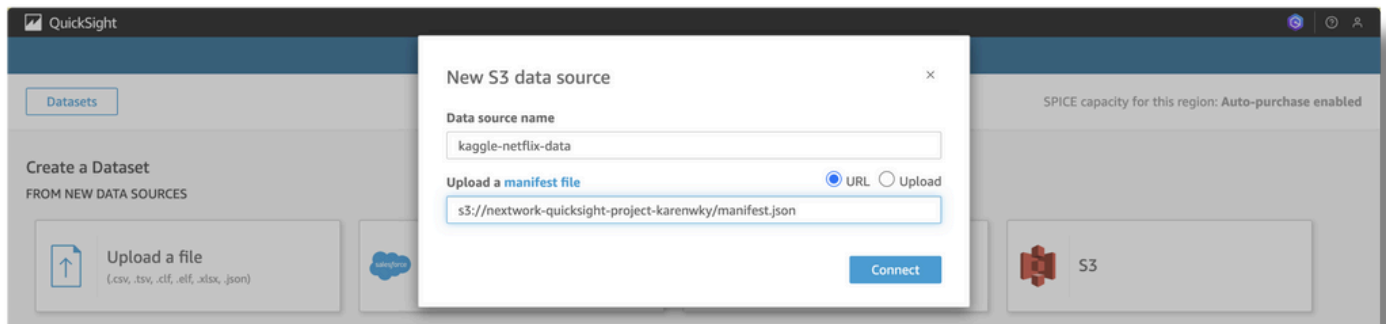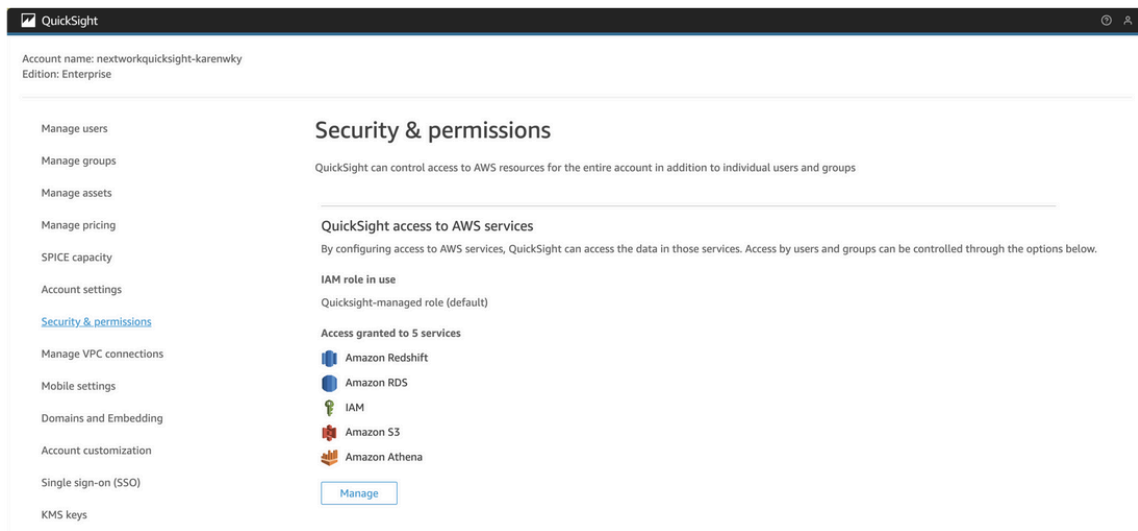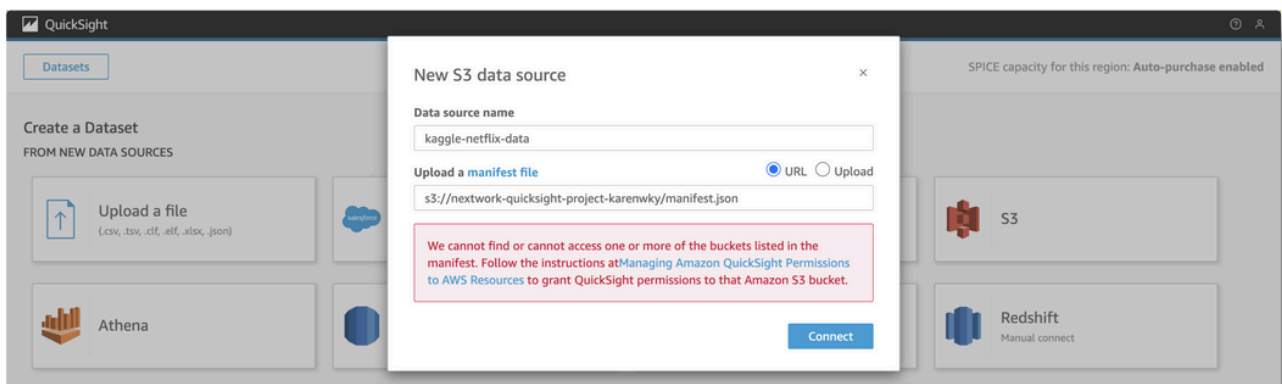

Karen Wong
linkedin.com/in/wongkayankaren

Error and Solution

An error popped up saying the buckets listed in the manifest are not found and can't be accessed. This is because the connection between S3 and QuickSight is not established.

To resolve the error, select "Manage QuickSight" after clicking the user icon in the top right corner. Then, select "Security & permissions", and click "Manage" to grant access to S3.
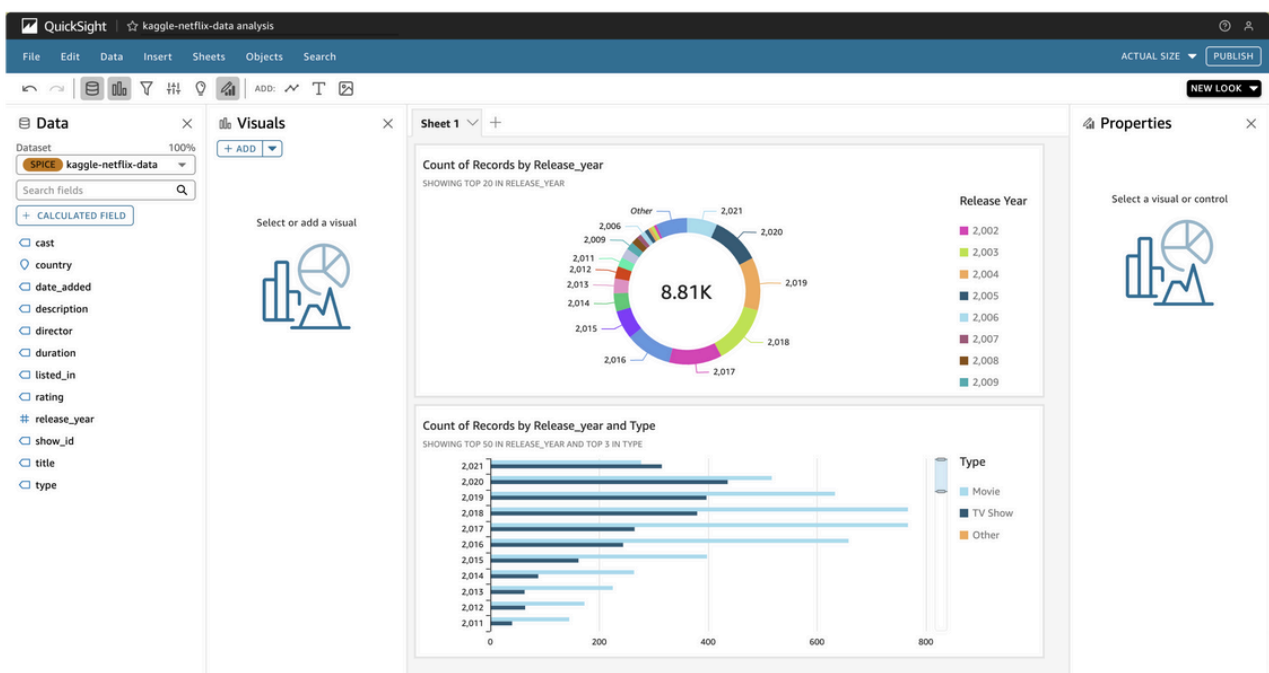

Karen Wong

linkedin.com/in/wongkayankaren

The First Visualization

To create visualizations in QuickSight, I clicked "ADD" under "Visuals" section to add a blank AutoGraph. Then, I dragged and dropped data fields from "Data" section to create the visualization, which automatically defines the visual type. I can also change the visual type and add multiple data fields into the visualization.

The graph shown here is a breakdown of Netflix show records by release year, displayed as a donut chart. Additionally, a horizontal bar chart presents the same data, complemented by show type. A legend is used to indicate colors for different data categories.

I created the donut chart using "release_year" as dimension. For the horizontal bar chart, I used "release_year" for y-axis and "type" for grouping.
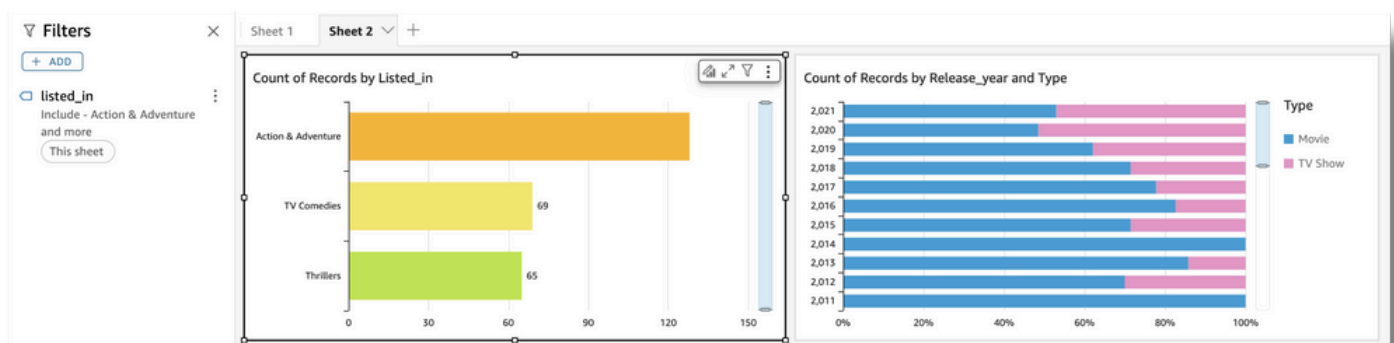
Karen Wong
linkedin.com/in/wongkayankaren

Using Filters

Filters are useful for displaying only a portion of the desired data. We can determine the filter scope to be applied to selected visualizations or dashboard sheets. By selecting filter type, such as filter list or custom filter, we can choose which data values to include or exclude.

The bar chart breaks down the show count across different categories, while the stacked bar chart displays the proportion of show types for each release year. I added a filter to apply to the entire sheet, including only "Action & Adventure", "TV Comedies", and "Thrillers" in the "listed_in" data field.

Karen Wong
linkedin.com/in/wongkayankaren

Calculated Fields

Calculated fields allow users to create new data points based on existing data within a dataset. These fields enable more complex analyses by applying mathematical operations, logical expressions, or string manipulations. Users can add calculated fields to a dataset during data preparation or from the analysis page.

I created new data points, "TV Show" and "Movie", from the "type" data field, each filtered to display only the specific type of data. Then, I used these new data points to create a data table showing the number of shows for each type by release year.

Edit calculated field		Sum of Movie and Sum of Tv Show by Release_year		

TV Show ✏

```
1  contains(type, "TV Show", CASE_SENSITIVE)
```

Movie ✏

```
1  contains(type, "Movie", CASE_SENSITIVE)
```

release_y...	Movie	TV Show
2,021	9	8
2,020	16	17
2,019	18	11
2,018	20	8
2,017	21	6
2,016	19	4
2,015	10	4
2,014	6	0
2,013	6	1

in Karen Wong
linkedin.com/in/wongkayankaren

Suggested Insights

Apart from charts and graphs, users can also select "Insights" as visualization type. This includes features such as aggregation, top and bottom ranked data, and the number of unique values displayed as text based on the data fields in each visualization. Users can also define customized insights by selecting their desired calculations.
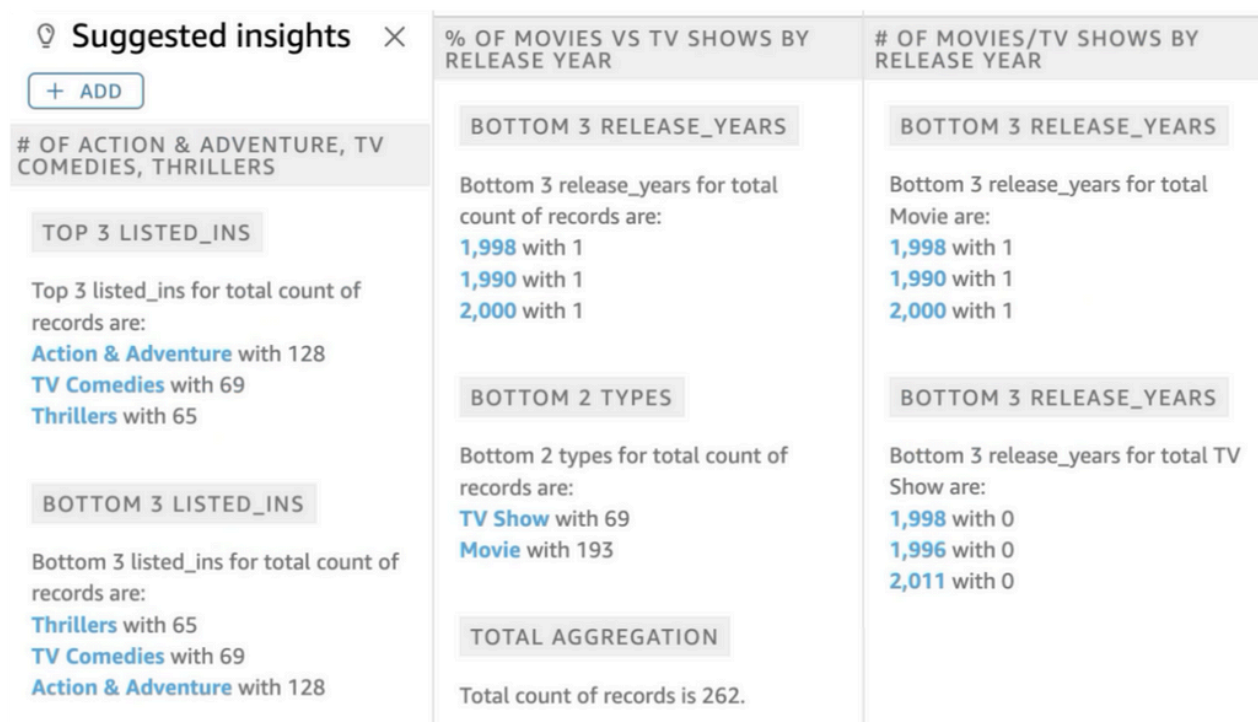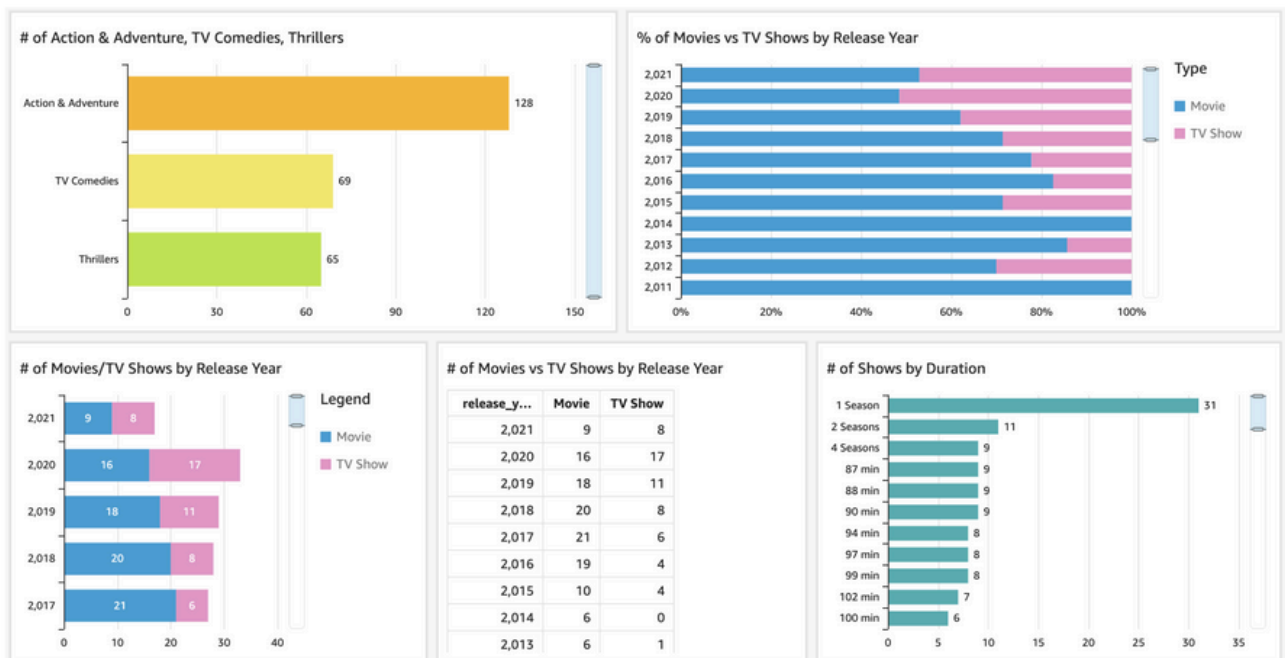
in **Karen Wong**
linkedin.com/in/wongkayankaren

Setting up a Dashboard

As a finishing touch, I edited the visualization titles by double-clicking on them. The font size and style can be selected, with a character limit of 120.

After publishing the dashboard, click "Export" then "Generate PDF" on the upper right corner. QuickSight will generate your dashboards into pdf file and you can download the file after few seconds wait.
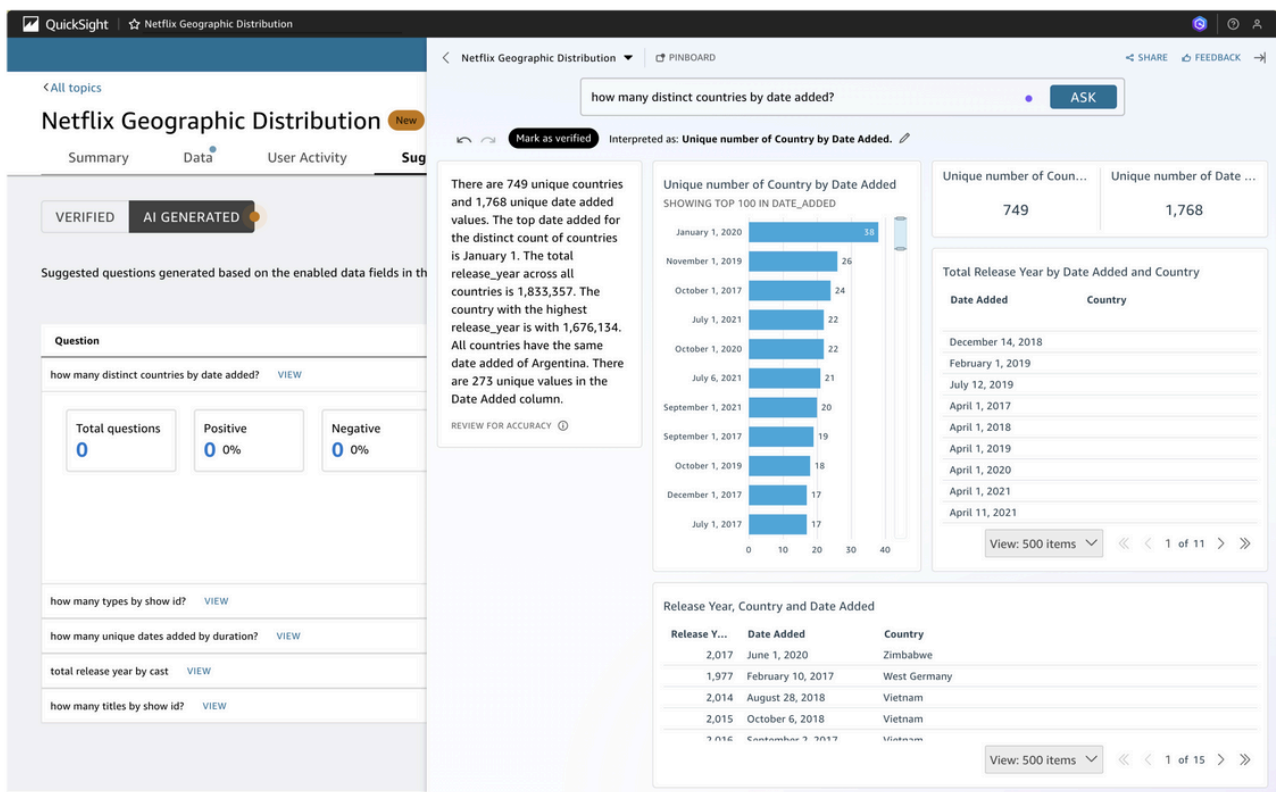
Karen Wong
linkedin.com/in/wongkayankaren

Generative BI with Amazon Q

Amazon Q offers powerful generative BI features that enhance data exploration and visualization. It enables users to ask questions in natural language, generating insightful reports and dashboards effortlessly. The ability to automatically suggest relevant visualizations based on queries is impressive. I'm particularly amazed by the Q Topics feature, which intelligently identifies key themes in data, making analysis intuitive and efficient.

 Karen Wong

linkedin.com/in/wongkayankaren

References

- [Amazon QuickSight - Business Intelligence Tools - AWS](#)

- [Adding calculations - Amazon QuickSight](#)

- [Working with insights in Amazon QuickSight - Amazon QuickSight](#)

- [Unlock the power of Generative BI with Amazon Q in QuickSight | Amazon Web Services](#)

- [Using Generative BI with Amazon Q in QuickSight - Amazon QuickSight](#)

- [What is S3 URI in Simple Storage Service? | AWS re:Post](#)

- [Quicksight error when chosing url for manifest file - Question & Answer - Amazon QuickSight Community](#)

- [NextWork - Visualize data with QuickSight](#)

in Karen Wong
linkedin.com/in/wongkayankaren