

# VISUALIZING DATA WITH QUICKSIGHT

## INTRODUCTION

Amazon QuickSight is a cloud-based service in AWS designed to create and share data visualizations, dashboards and insights. It is useful for analyzing data stored in AWS. It can handle massive datasets. Data visualization makes complex information easy to understand, analyze and communicate and many companies and Professionals use it in their different industries. Large datasets can be overwhelming but graphs, charts and dashboards make patterns easier to see.

Netflix uses data visualization to not only enhance user experience but also improve business operations. They collect massive amounts of data on what users watch, skip or rewatch and they visualize these data to know which genres are most popular in different regions. Data visualizations help Netflix to evaluate which shows and movies are performing well.

## PREREQUISITES

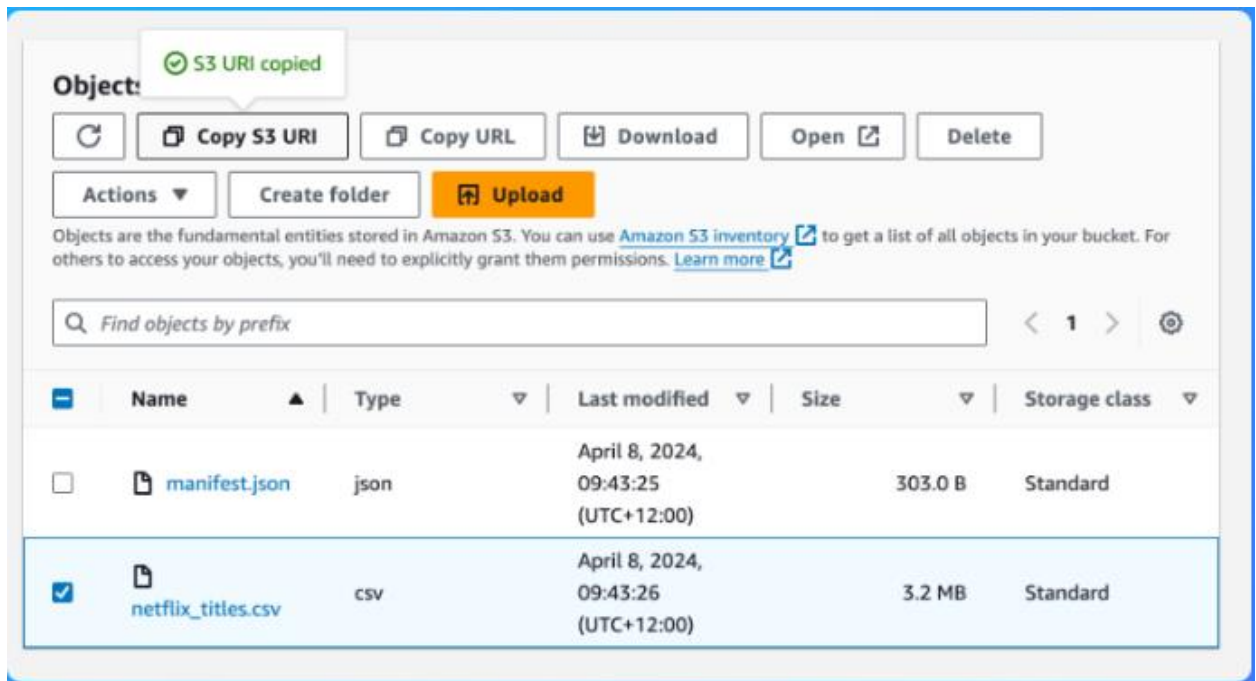
- AWS Account
- MySQL Database
- Datasets ( Netflix\_titles.csv, manifest.json)

**The following steps below can be used to visualize data in AWS using QuickSight**

### 1. STORE THE DATASETS IN AMAZON S3

In this step we are storing dataset in an Amazon S3 Bucket. QuickSight will connect to S3 to use the data in this Bucket to create visualizations.

- Log into your **AWS Account**
- Open your **S3 Console**
- Select **Create Bucket**
- Name the Bucket and select the region closest to you.
- Keep the rest of the settings as default and select **Create Bucket**.
- Upload the files ( Netflix\_titles.csv, manifest.json) into the Bucket.
- Copy the S3 URL of your file( Netflix\_titles.csv) .



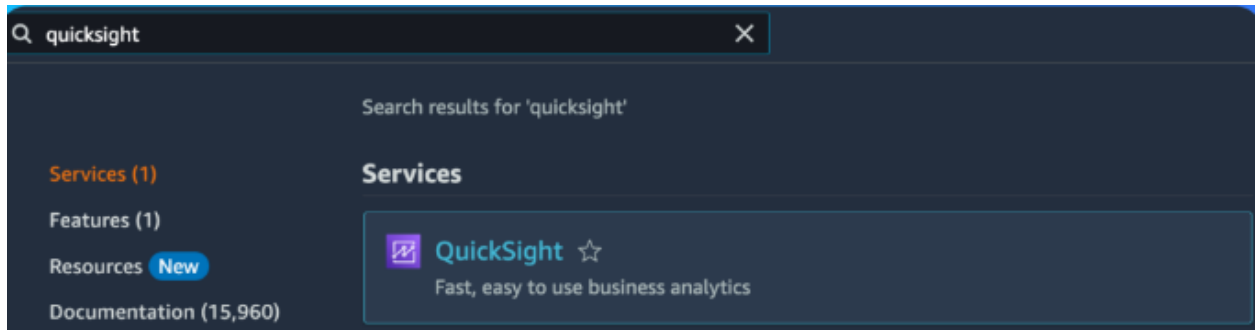
- Open your **manifest.json** file in your laptop's text editor - for example, TextEdit (Mac) or Notepad (Windows).
- Replace the URL in the **manifest.json** file with the S3 URL of your dataset.

```
manifest.json
{
  "fileLocations": [
    {
      "URIs": [
        "s3://nextwork-quicksight-project-name/netflix_titles.csv"
      ]
    }
  ],
  "globalUploadSettings": {
    "format": "CSV",
    "delimiter": ",",
    "textqualifier": "\"",
    "containsHeader": "true"
  }
}
```

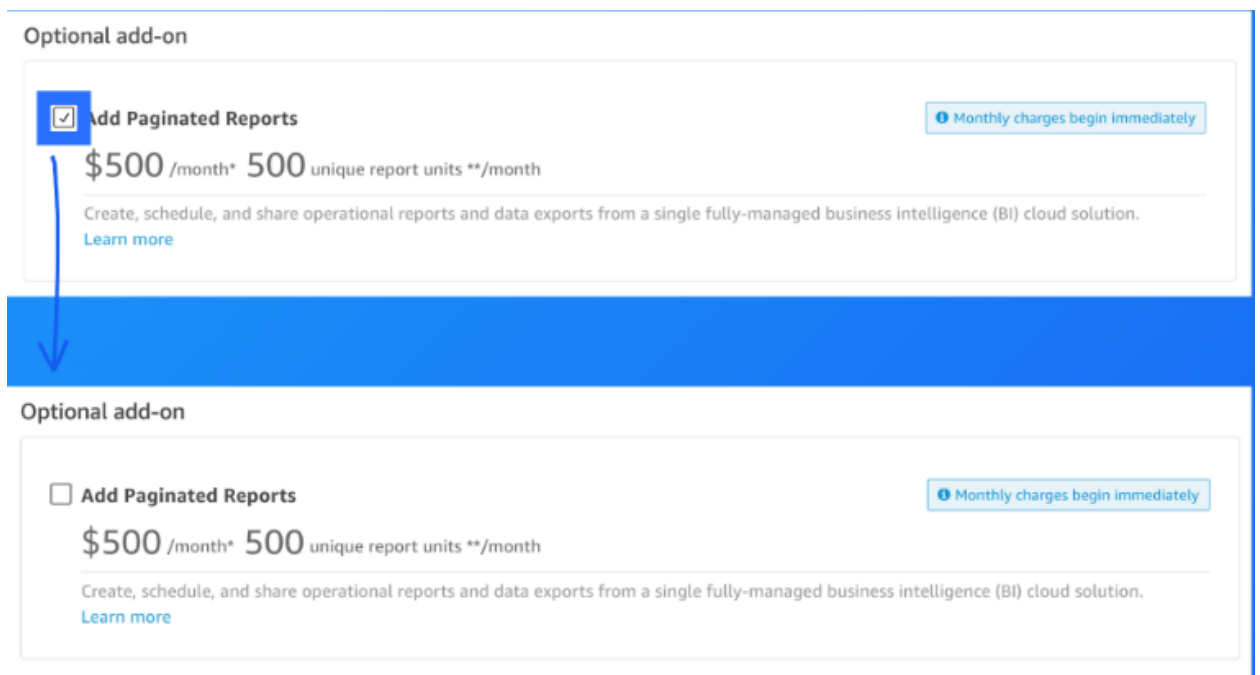
- Re-upload the edited **manifest.json** file into your bucket, which you'll notice automatically replaces the existing one.

## 2. CREATE YOUR AMAZON QUICKSIGHT ACCOUNT

- Search for Amazon QuickSight.



- Select Signup for QuickSight and signup for a free trial if you do not have an account.
- Make sure you uncheck the offer to upgrade.



- Enter your details for your QuickSight account - **make sure the email you use is the same email for your AWS account to avoid error.**

## Sign up for QuickSight

### Contact information

Email for account notifications

*Enter email address*

### Authentication method

- ☒ Use IAM federated identities & QuickSight-managed users  
Authenticate with single sign-on (SAML or OpenID Connect), AWS IAM credentials, or QuickSight credentials
- ☐ Use AWS IAM Identity Center  
Authenticate using AWS IAM Identity center  
**i** Manage access to QuickSight by assigning users and groups from IAM Identity Center. [Learn more](#)
- ☐ Use IAM federated identities only  
Authenticate with single sign-on (SAML or OpenID Connect) or AWS IAM credentials
- ☐ Use Active Directory  
Authenticate with Active Directory credentials

### QuickSight region

Select a region



US West (Oregon)



### Account info

QuickSight account name

You will need this for you and others to sign in



nextworkquicksight

- Select **Amazon S3**.
- Select **Select S3 buckets**












## QuickSight access to AWS services

Make your existing AWS data and users available in QuickSight. [Learn more](#)

### IAM Role

- ☒ Use QuickSight-managed role (default)
- ☐ Use an existing role

### Allow access and autodiscovery for these resources

- ☒  Amazon Redshift
- ☒  Amazon RDS
- ☒  IAM
- ☒  Amazon S3 (2 buckets selected)  
[Select S3 buckets](#)
- ☒  Amazon Athena  
Make sure you've chosen the right Amazon S3 buckets for QuickSight access
- ☐  Amazon S3 Storage Analytics
- ☐  AWS IoT Analytics
- ☐  Amazon OpenSearch Service
- ☐  Amazon SageMaker
- ☐  Amazon Timestream
- ☐  AWS SecretsManager  
[Select secrets](#)

- Tick the box for the S3 bucket you created.

## Select Amazon S3 buckets

×

S3 Buckets Linked To QuickSight Account

S3 Buckets You Can Access Across AWS

Select the buckets that you want QuickSight to be able to access.

Selected buckets have read only permissions by default. However, you must give write permissions for Athena Workgroup feature.

☒ Select all

S3 Bucket	Write permission for Athena Workgroup
<input checked="" type="checkbox"/> nextwork-quicksight-project-name	<input type="checkbox"/>

- Select **Create**.



Account created successfully

Account name: nextworkquicksightekom

GO TO QUICKSIGHT

---

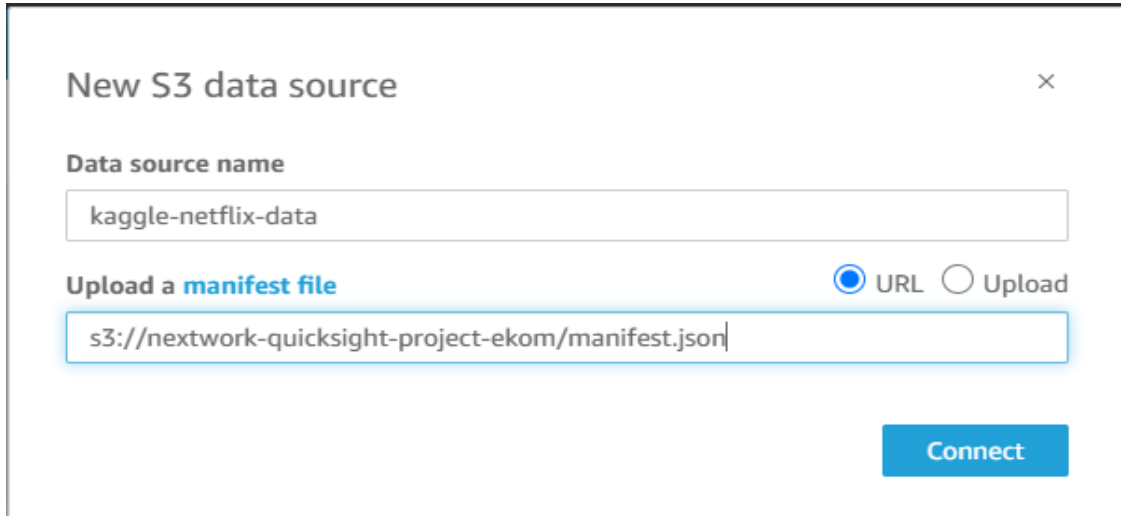
You've created an Amazon QuickSight account!

- Select **Go to Amazon QuickSight** to keep going.

### 3. CONNECT YOUR S3 BUCKET TO AMAZON QUICKSIGHT

- From the left hand navigation bar, select **Datasets**, then **New dataset**.
- Select **S3**.
- For the first field (source name), enter “**kaggle-netflix-data**”
- Open a **new tab** to open your AWS Management Console again. Head back to your S3 bucket.
- Select the checkbox next to **manifest.json**, then select **Copy S3 URL**

- Enter the S3 URL to your **manifest.json** file.



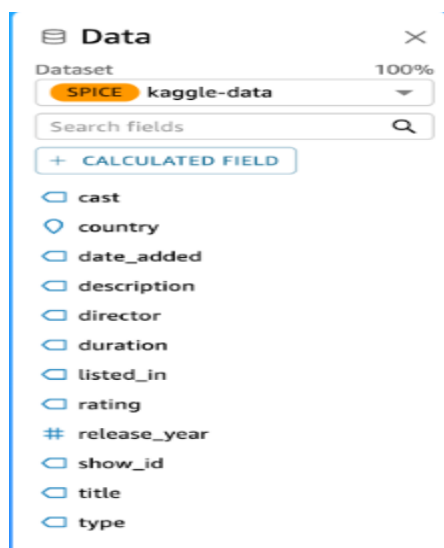
- Select **Connect**.

**SUCCESS!**

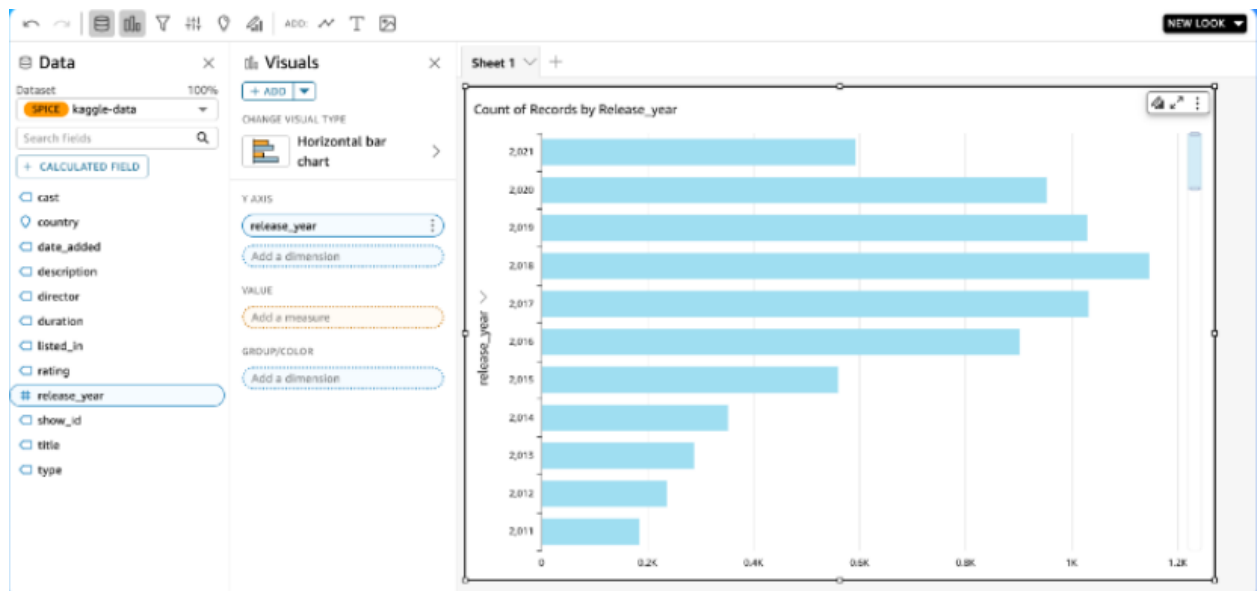
- Select **Visualize**.
- Select **Create**.
- Select **Interactive sheet** to start creating visualisations.

#### 4. CREATE YOUR FIRST QUICKSIGHT VISUALIZATION

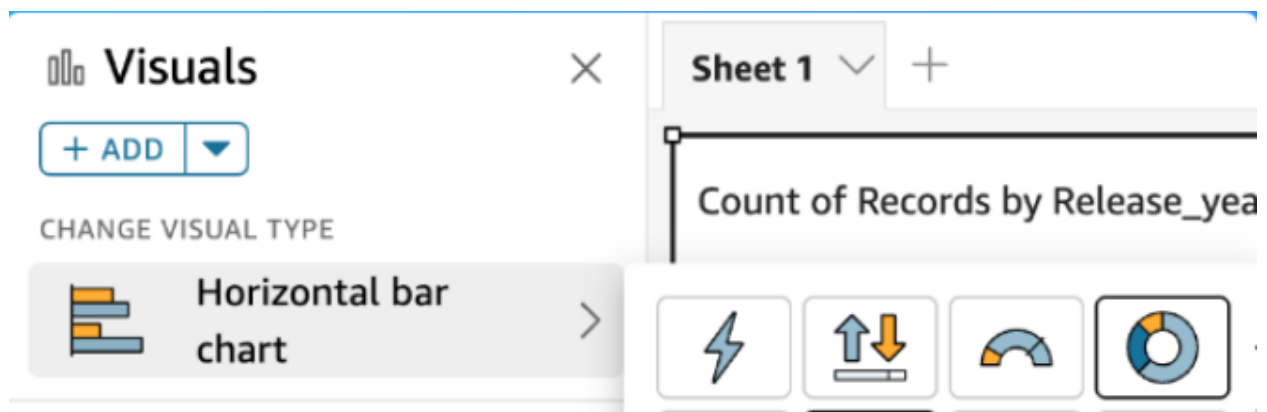
With QuickSight, you can sort, filter, and customise your data to create visualisations. You can also experiment with different types of graphs like bar charts, pie charts, line graphs, etc. You can see on the left hand panel that the dataset's fields are already imported.



- Drag fields into the graph to create visualisations.
- Drag **release\_year** into the Y-Axis heading. Now you can see a breakdown on the year that these Netflix-featured TV shows and movies were released.

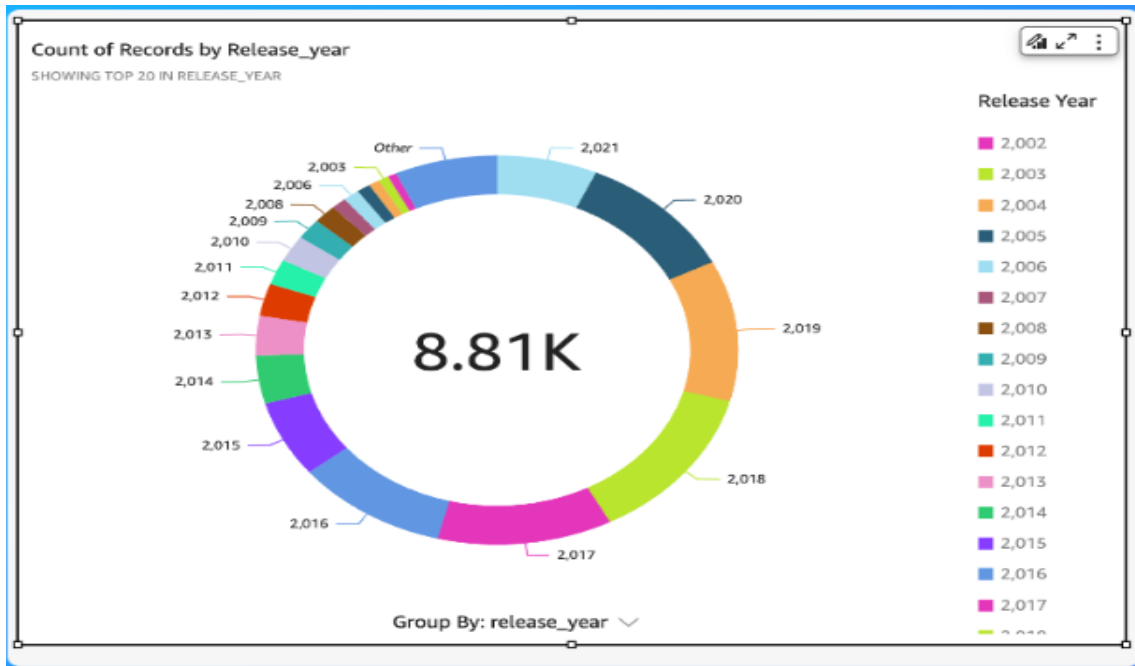


- You can create different types of charts too. See what happens when you change your chart to a donut chart!

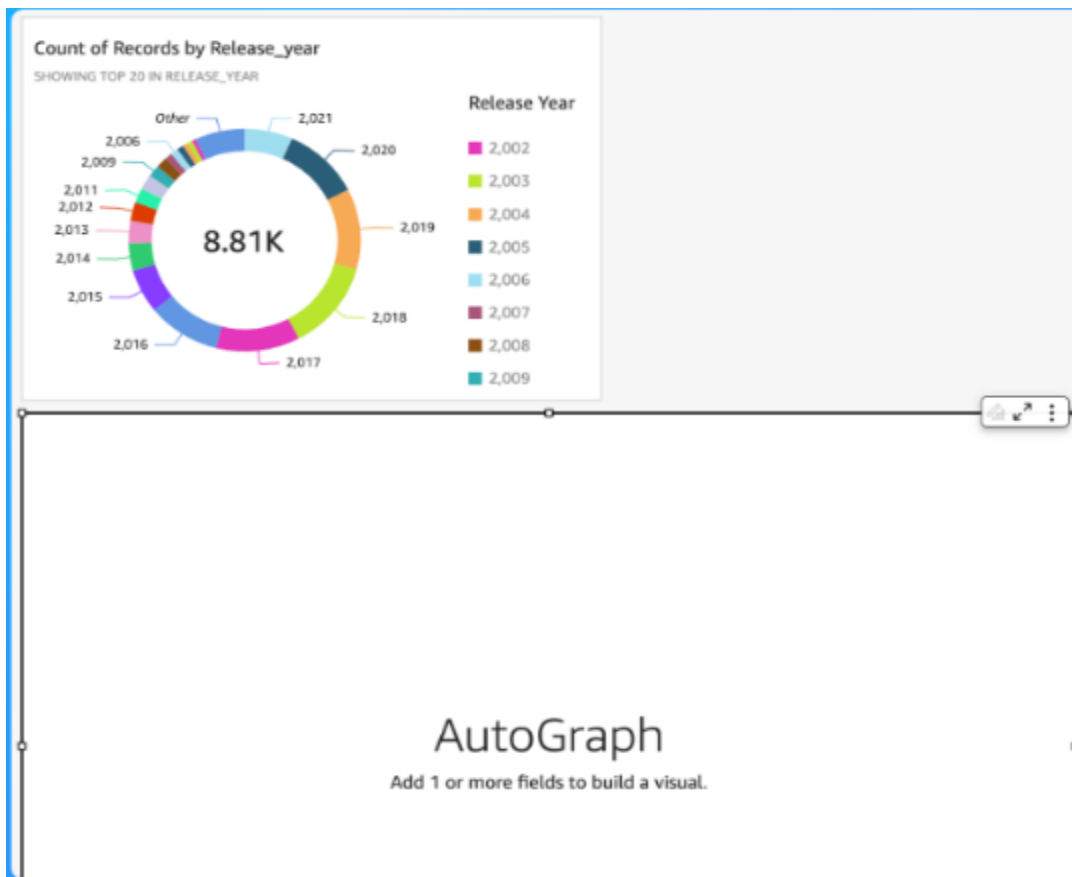


- Let's save this in a dashboard. Click on the frame surrounding your lovely donut chart, and click on the white boxes at the edges to resize it.



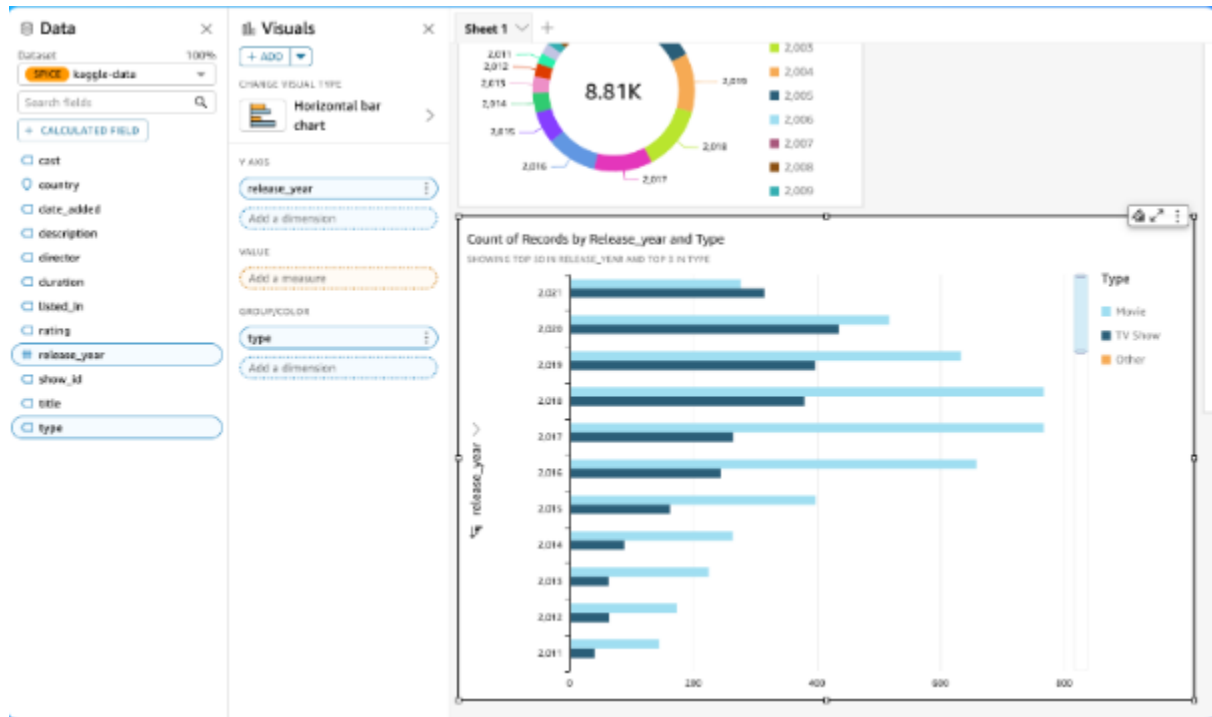


- Now let's create a new visual, select + **ADD** under the **Visuals** heading on your middle navigation bar, and you'll see another blank frame pop out.

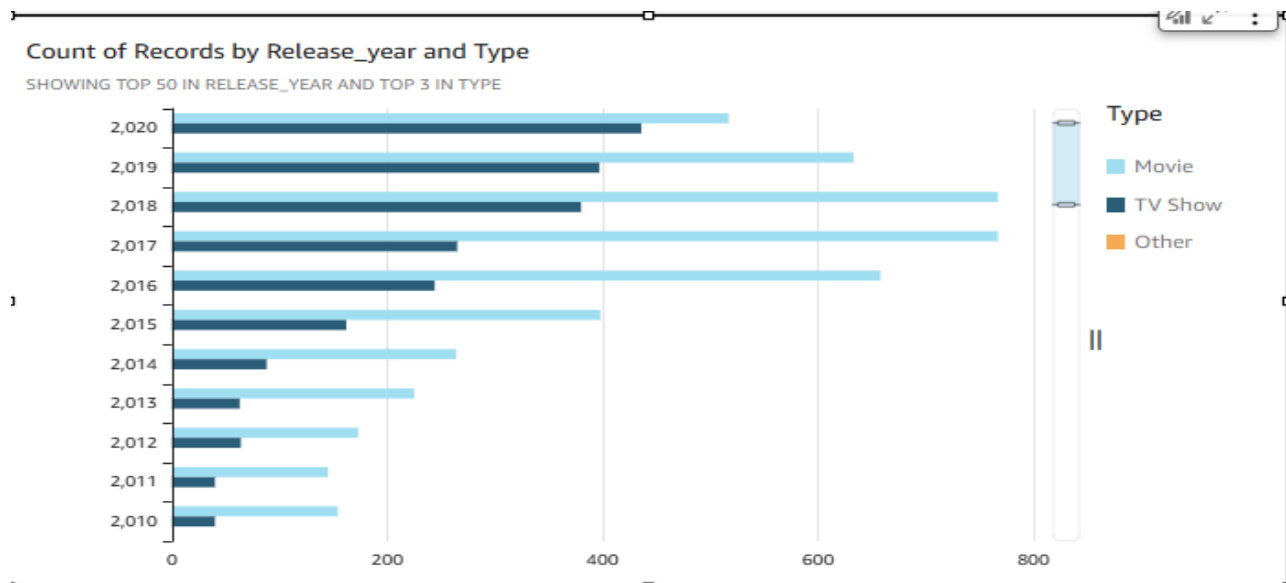


Let's see what we could do to see a breakdown of TV shows vs movies for every year

- Drag the **release\_year** label into the **Y Axis** heading.
- Next, drag the **type** label into the **Group/Color** heading.

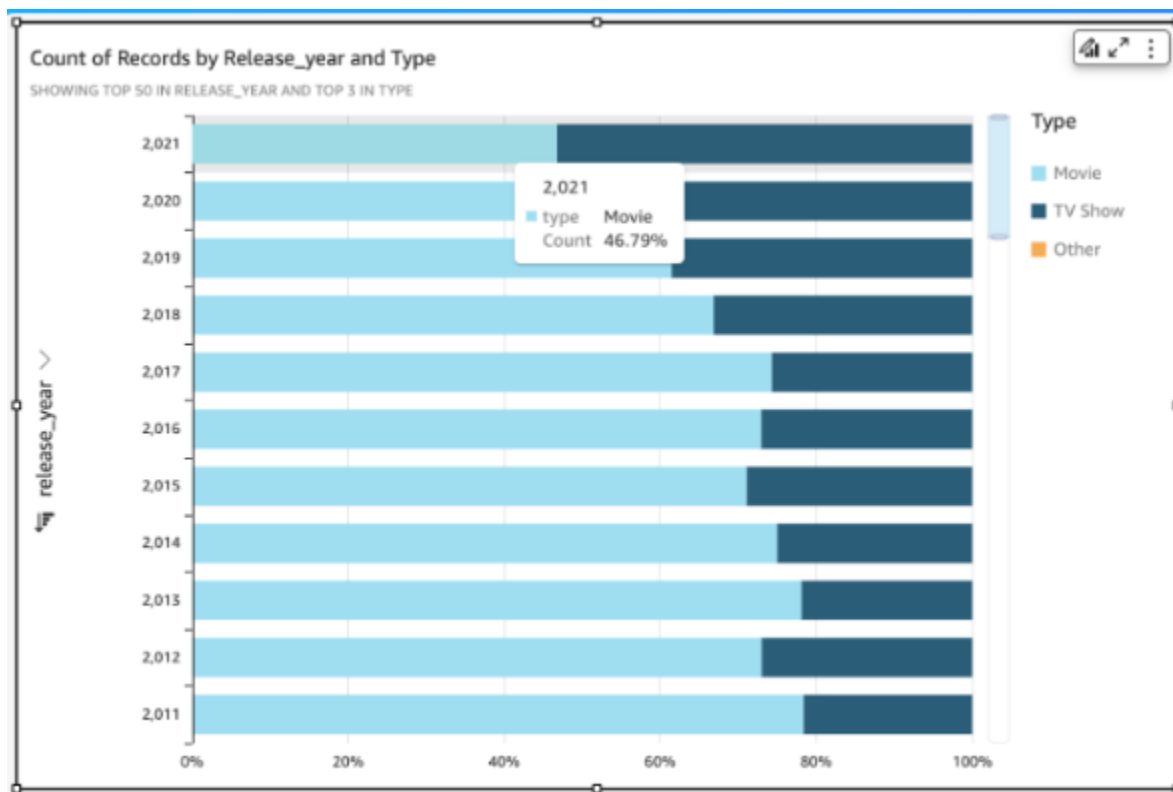


- Resize your bar chart so it's the same width as the pie chart. If you click on the frame and hover over the black edges, you can also start to move your chart around the screen.



Let's try answering some questions on QuickSight. For instance

- I quite like the breakdown of TV shows/movies for each release year. Would it be possible to stack movies and TV shows in the same row, so you can visualise the % of each?
- **SOLUTION:** Picking up where you left off in the previous task, change the graph type to a Horizontal stacked 100% bar chart!

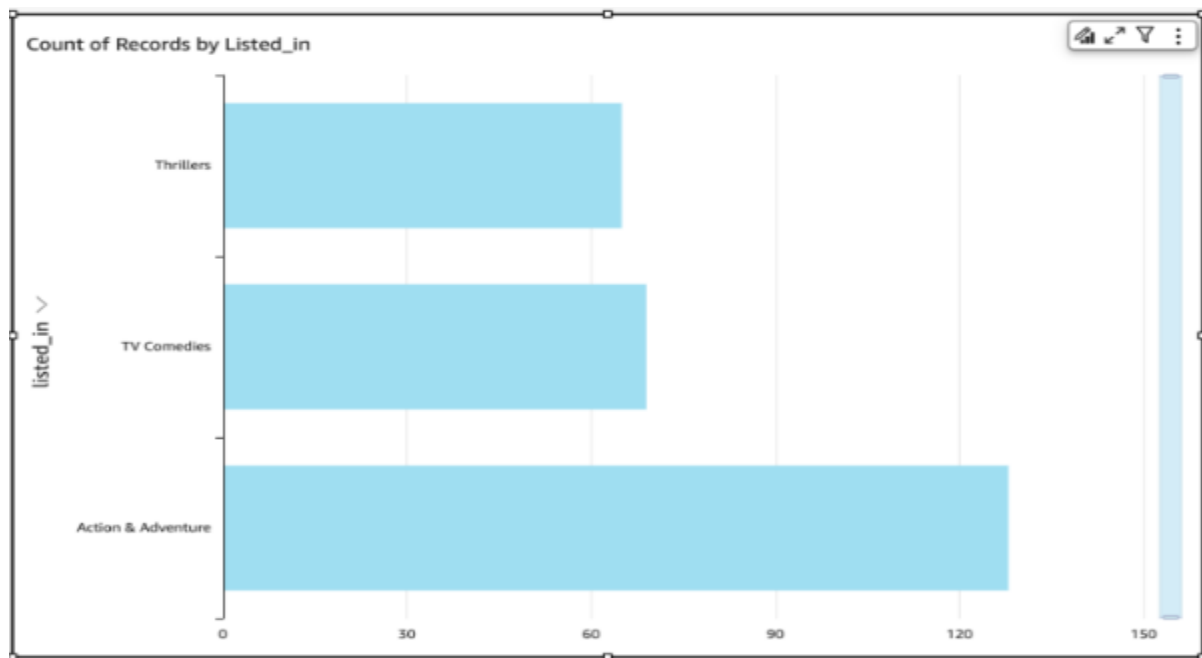


- [Add a new chart instead of replacing the current one]  
Now can you show me the same thing in a table? i.e. please show me the number of movies vs TV shows per release year in a table.
- **SOLUTION:** Change your visual type to **Table**, then add **release\_year** as your **Group By** label. Add **title** as your **Value** metric. Add the **type** label as a dimension!
- [Add a new chart instead of replacing the current one]  
On what day did Netflix add the largest number of movies/TV shows to their catalogue?

- **SOLUTION:** To do this, move the **date\_added** label to both the **Y Axis** and **Value** headings. Then, next to the **Value** heading, click on the three dots and select **Sort order: Descending**. Use your mouse to scroll to the top bar - it's January 1, 2020!

**Answer:** January 1, 2020. Netflix added 109 TV shows/movies to their catalogue.

- [Add a new chart instead of replacing the current one]  
Of the TV shows and movies featured, how many were listed as 'Action & Adventure', 'TV Comedies', or 'Thrillers'? For simplicity, ignore the TV shows and movies that have multiple listings/categories.



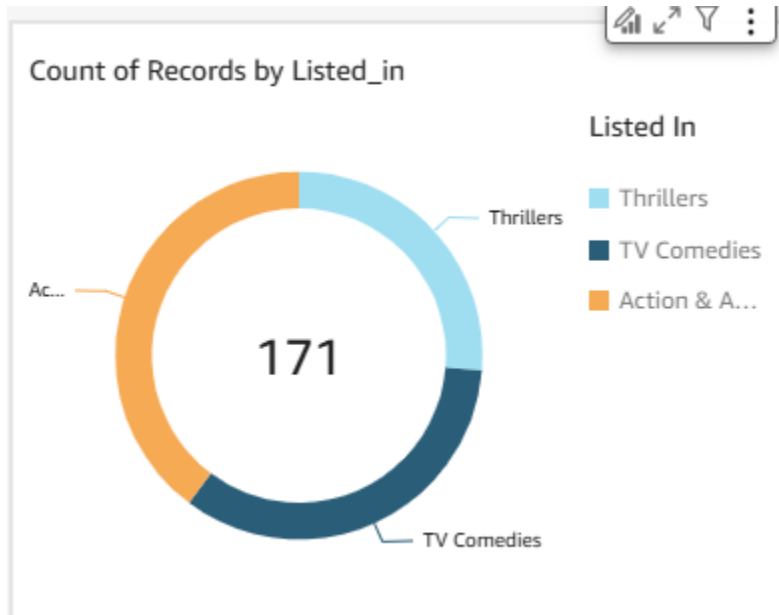
- **SOLUTION:** To do this, add a filter for the **listed\_in** field. Click on the created filter to edit it: Deselect the **Select all** checkbox, then select these three tags (use the Search values search box to help you):
  - Action & Adventure
  - Thrillers
  - TV Comedies
- Select **Apply** when you're done.

**Answer:**  $65 + 69 + 128 = 262$

- [Add a new chart instead of replacing the current one!]  
Of the TV shows and movies with the listing 'Action & Adventure', 'TV Comedies', or 'Thrillers', how many were released on 2015 or after?
- **SOLUTION:** To pick up where you left off, you can simply click on the three dots and select duplicate visual. This pops out an identical chart that gets you started right away.

Next, let's make sure this is filtered to items that were released from 2015 or after. Select the **Filter** icon and add a **release\_year** filter. Make sure only the years 2015 and over are selected, then select **Apply**.

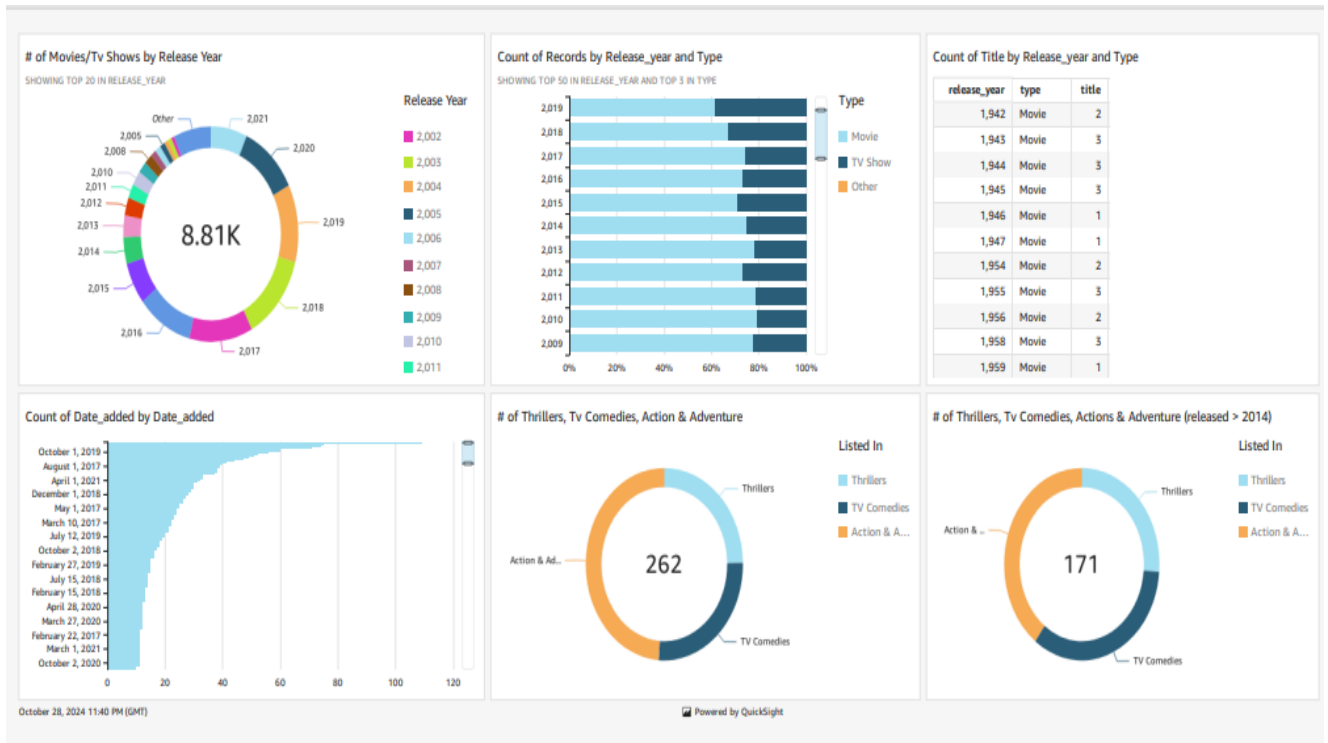
**Answer:**  $68 + 58 + 45 = 171$



- On the top right hand corner, select **Publish**. This will make your dashboard public so you can share it with your data team lead.
- Give your dashboard a name and click on that **Publish dashboard** button.

To have a copy of your dashboard, on the top right hand corner, select the **Export** icon.

- Select **Generate PDFs**.
- Wait for the PDF to be ready, then select **Download** when you see that green banner pop up!
- Niceeeee, now you see a wonderful PDF version of the dashboard you've just created.



## 5. CLEAN UP YOUR RESOURCES!

### Don't forget to...

- Terminate your QuickSight account
- Delete your S3 bucket

### To terminate your QuickSight account

- After returning to the home page, select the user icon on the top right corner, and then Manage QuickSight:
- Select Account Settings from the left hand navigation panel, and then Manage at the bottom of the page.
- **Toggle off** account termination, and then type confirm before clicking Delete account.

## Account termination

QuickSight account name

ambernexwork

### Account termination protection ⓘ

Account termination protection is an extra safe-guard to help prevent accidental deletion of accounts.



Account termination protection is off.

### Delete account

Deleting this account can't be undone and will permanently delete all users, dashboards, analyses, along with other related data.

Type "confirm" to delete this account

Cancel

Delete account

**Success!**

## Unsubscribe successful

All content has been deleted. You won't be charged for any further QuickSight usage.

Sign up for QuickSight

### To delete your S3 bucket

- Select your bucket, then choose Delete.
- You'll notice that you can't delete the bucket! There are still objects inside. Select Empty bucket.
- Type permanently delete in the text input field.
- Nice! Once that's deleted, you should be able to delete the bucket itself.