# Text Insights with AI/ML

Create engaging visualizations with Amazon QuickSight

September, 2022

# 1 Overview

This guide will show you how to create some simple, sample visualizations for the text inisghts generated with this prototoype using Amazon QuickSight.

# 2 Artifacts

## 2.1 Create QuickSight dashboard

In this section we will explain how you can create the QuickSight dashboard to visualize the insights obtained by this prototype.

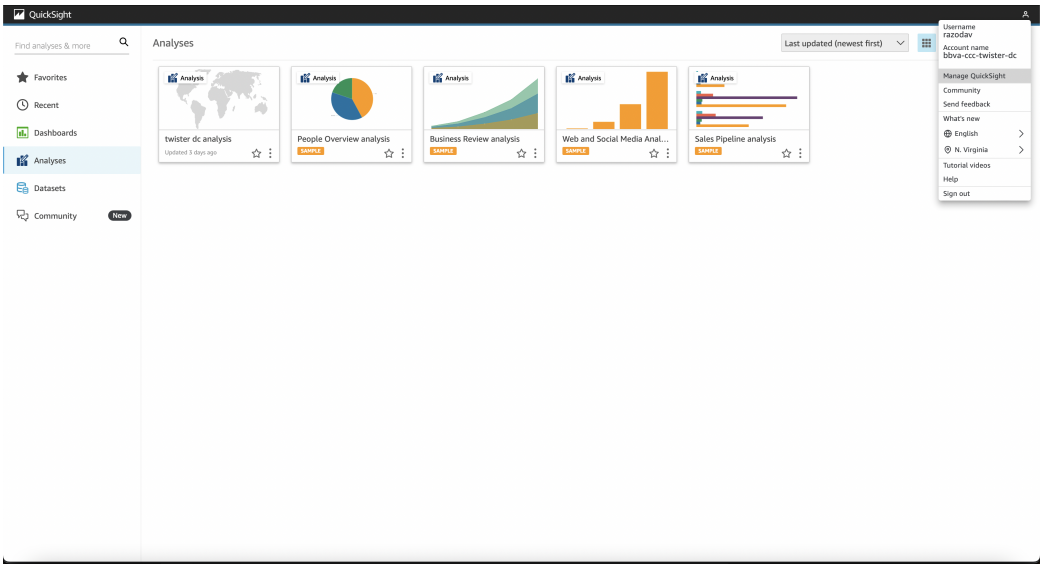
### 2.1.1 Prerequisites

In order to create the dashboard in QuickSight you will need the following:

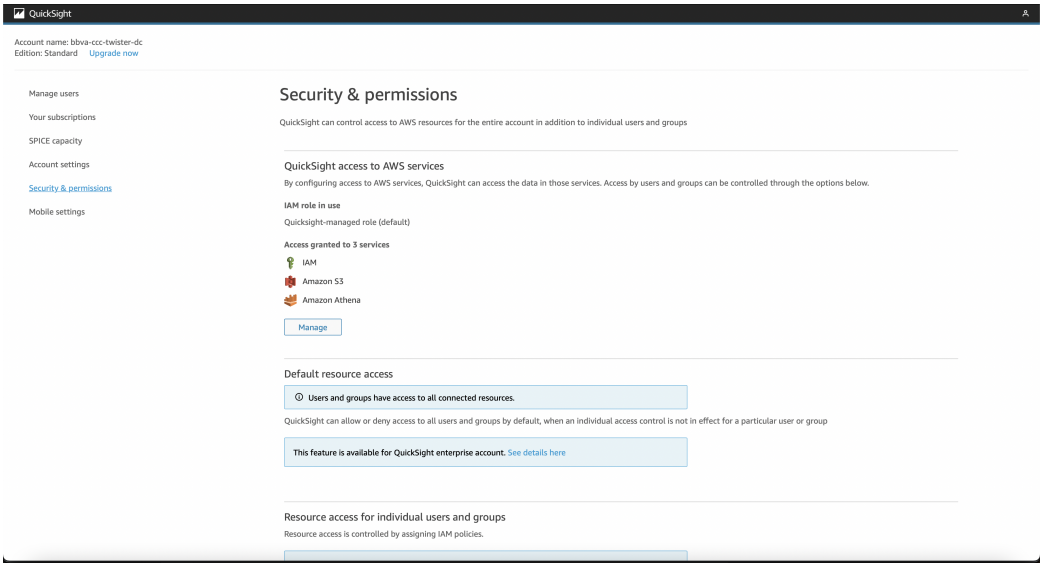
- A user with access to QuickSight
- A subscription to QuickSight

You may follow the instructions in Setting up QuickSight to set up you QuickSight account.

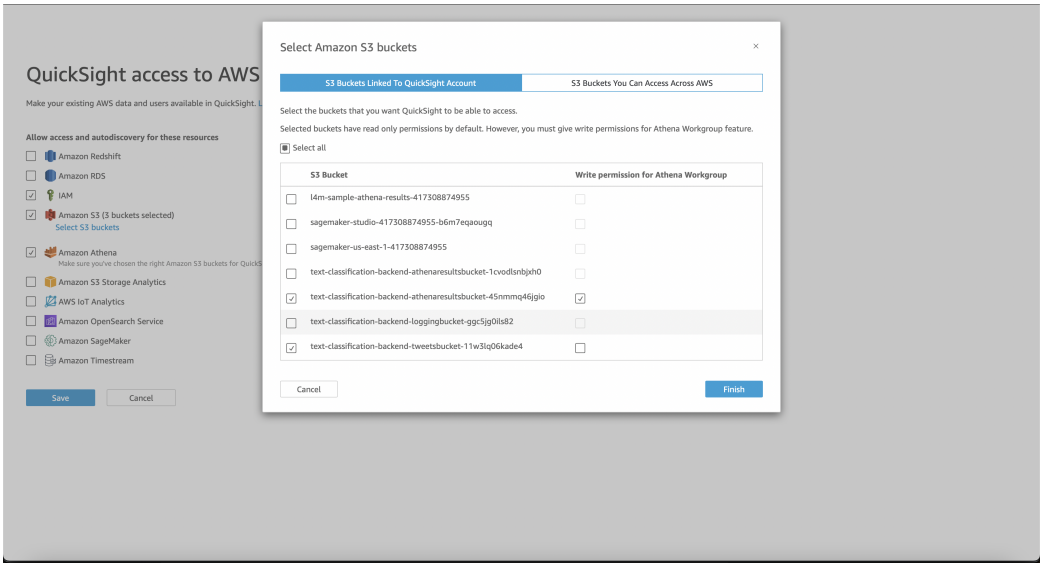
Configure access to your data sources. Head to the top-right corner of QuickSight and click on the user icon then click on "Manage QuickSight".




Go to security and permissions and then click on the "Manage button".



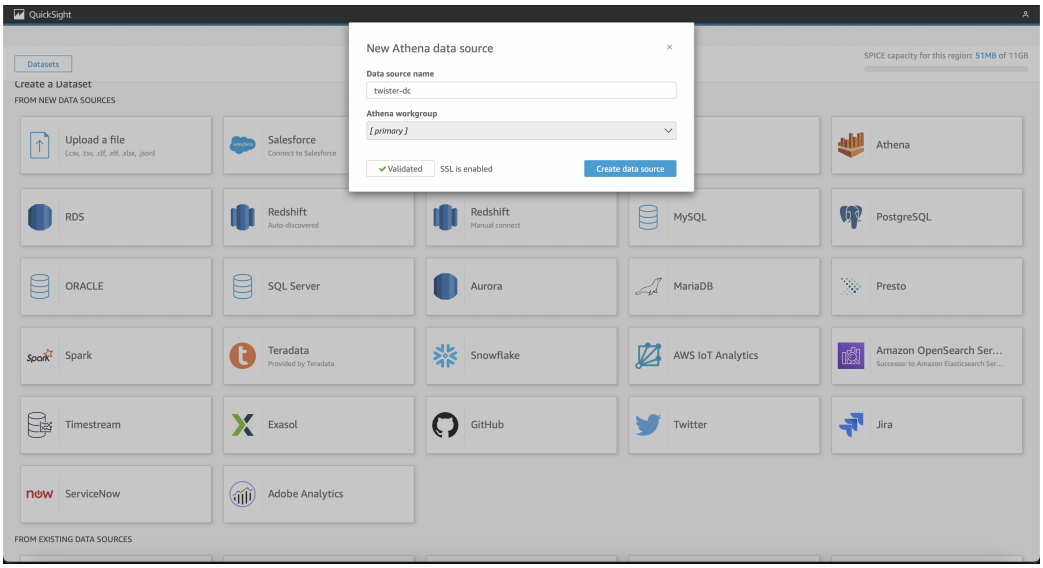

On the screen check the boxes of IAM, Athena and S3. Click on the "Select S3 buckets" button and check the boxes for the Athena results bucket and Tweets bucket, also check the "Write permission for Athena workgroup" bucket for the Athena bucket. The name of the buckets can be found in the backend stack outputs as explained in the README file.



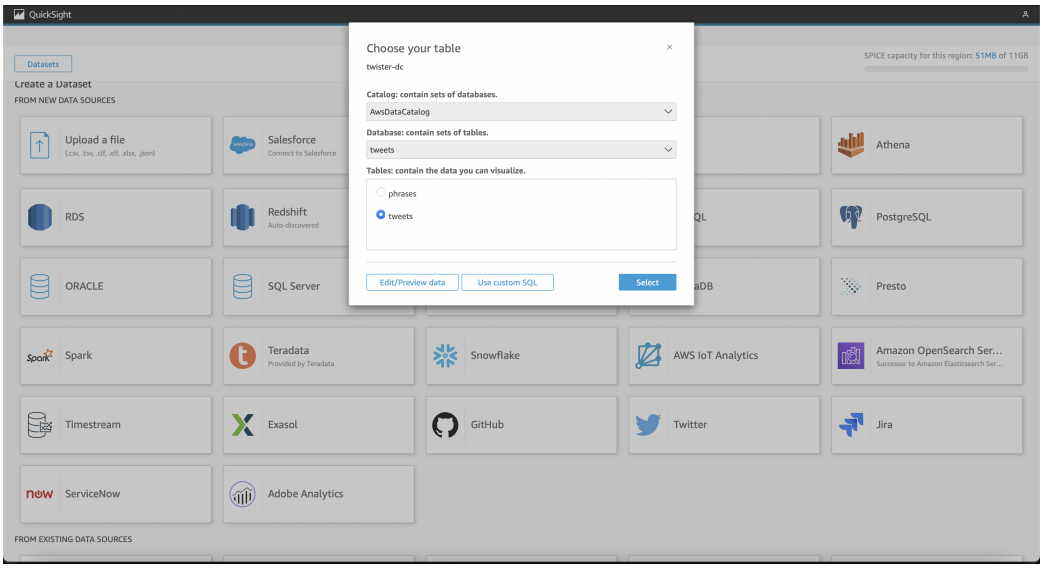

## 2.1.2 Create datasets

We now need to create the datasets. Head back to QuickSight main screen and navigate to Datasets and click on "Create new dataset", on the new screen

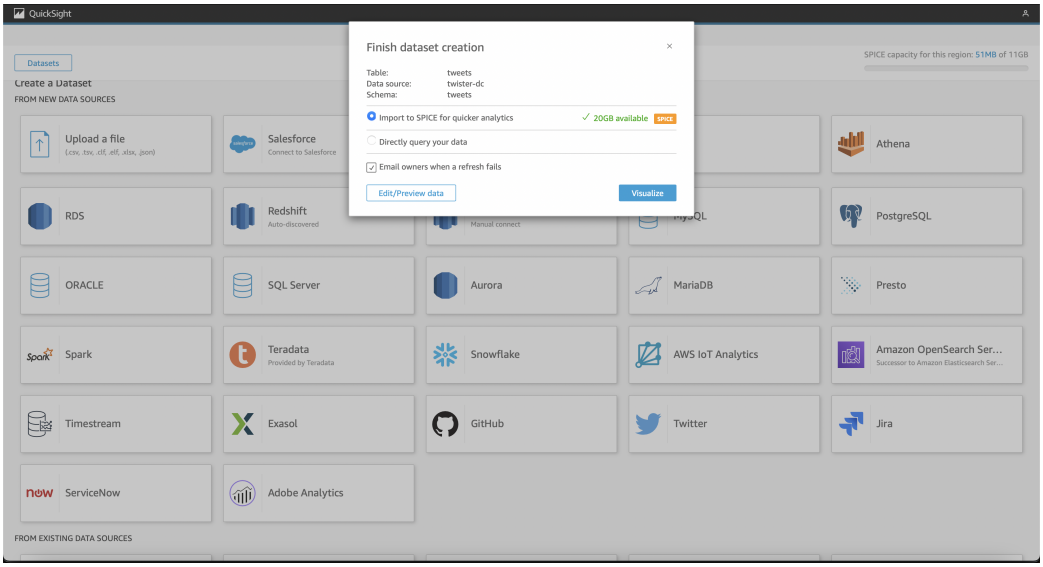
select the "Athena" option. Name your data source and select the "primary" workgroup.




In the next step you will configure your table. Choose the AWSDataCatalog and select the "tweets" table and choose the tweets table.




On the next screen select the "Import to SPICE for quicker analytics" options and click on the Visualize button. This will create an Analysis but we will disregard it for now.

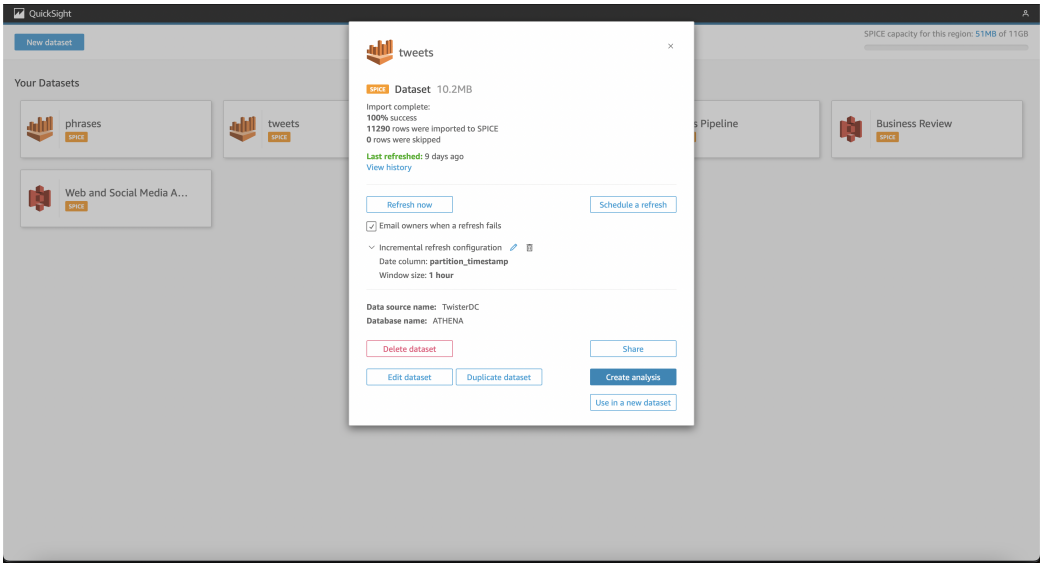



Repeat the same process to add tha "phrases" table.

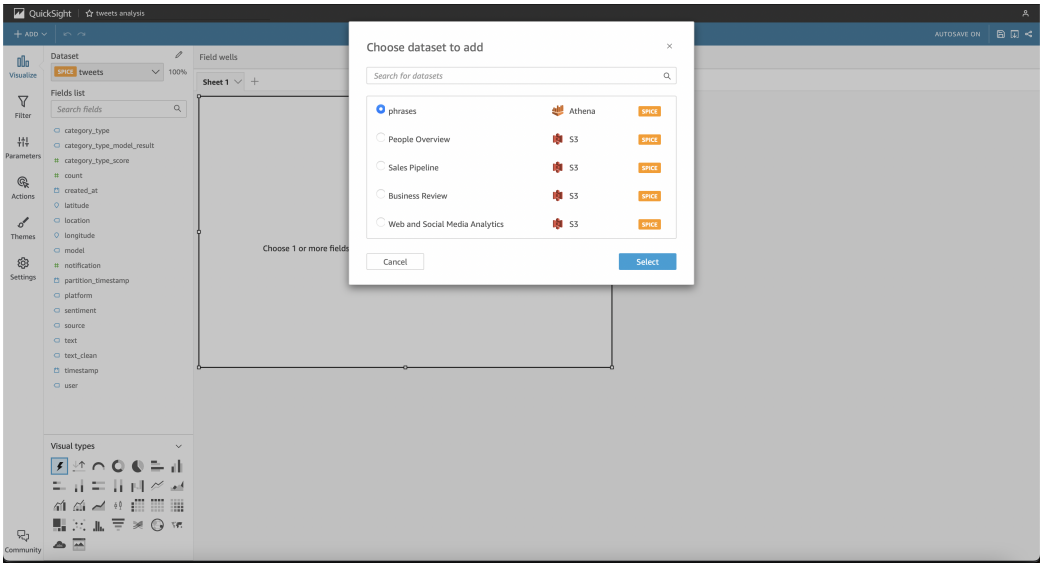
Additionally, you can configure dataset periodic refreshes to keep your dashboard up to date with the latest data.

### 2.1.3 Create QuickSight analysis

To create the visuals we first need to create a new QuickSight analysis. From the main QuickSight screen click on the "Analyses" section and then click on the "New Analysis" button on the top-right corner. In the next screen select the "tweets" dataset and click on the "Create analysis" button. This will create a new QuickSight analysis where you can create visuals to obtain insights from the processed data.



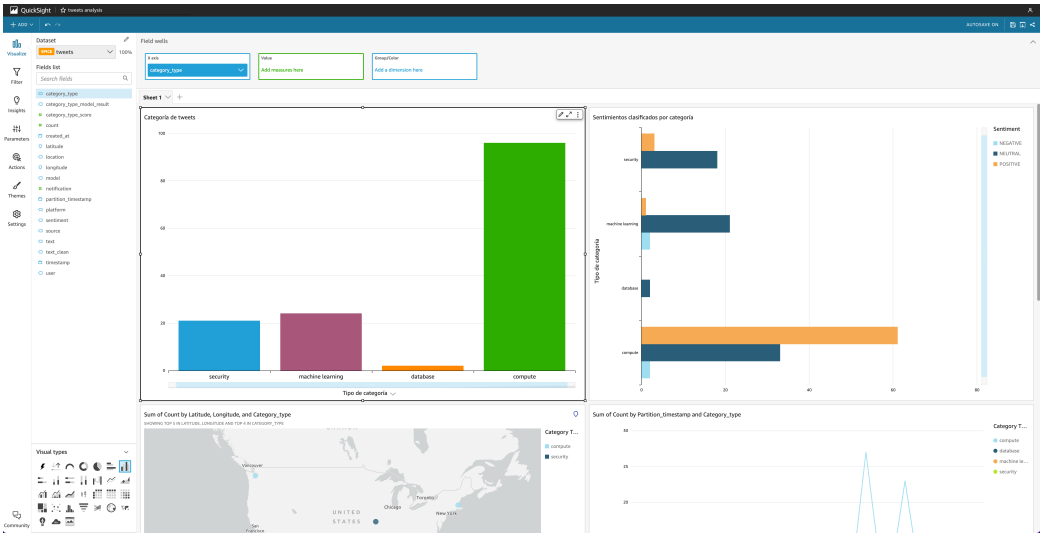

Let's now add the "phrases" dataset to the analysis. Click on the little pencil icon next to the Dataset section and click on "Add dataset". From the screen select the "phrases" dataset and click "Select".




We are now ready to create some data visualizations.

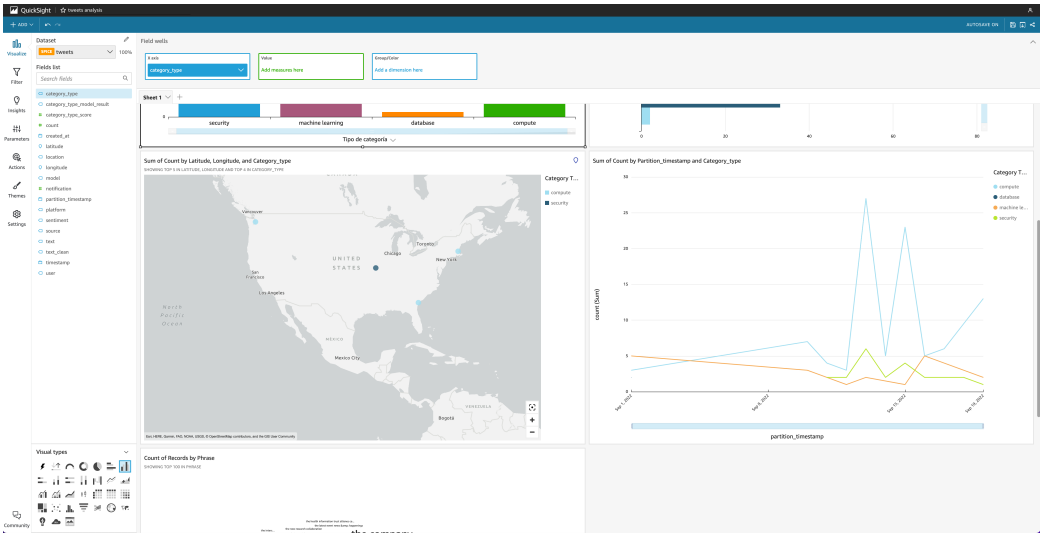
We first create a bar plot to visualize the total number of times a category has appeared. From the "Fields list" select the "category_type" field, this will automatically create a bar plot like the following.

For the second visual lets create the same bar plot but splitting the data by sentiment. Select the fields "category_type" for the Y-axis and "sentiment" for the GroupColor.




The third visual will be a 2D map to pinpoint where are these mentions being made. Select the fields "longitude" and "latitude" for the Geospatial, "count" for the Size and "category_type" for the Color.

We will now create a word cloud to know what are the customers talking about the most. For the word cloud we will use the "phrases" table so first select the "phrases" table from the Dataset dropdown menu. Now select the "phrase" field from the fields list and choose "Word Cloud" from the Visual Types.




Finally, let's create a line chart to visualize the progression over time of the volume of mentions made per category, it will also help us identify anomalous spots. Select the "tweets" dataset again and choose the fields "created_at" as X axis with an aggregate by HOUR, "count" as Value with a SUM aggregate, and "category_type" as Color.