

# 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 Quick-Sight account.

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



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

QuickSight							
ount name: bbva-ccc-twister-dc tion: Standard Upgrade now							
Manage users	Security & permissions						
Your subscriptions	QuickSight can control access to AWS resources for the entire account in addition to individual users and groups						
SPICE capacity							
Account settings	QuickSight access to AWS services						
Security & permissions	By configuring access to AWS services, QuickSight can access the data in those services. Access by users and groups can be controlled through the options below.						
Mobile settings	IAM role in use Quicksight-managed role (default)						
	Access granted to 3 services						
	<b>腎</b> IAM						
	Amazon S3						
	💐 Amazon Athena						
	Manage						
	Default resource access						
	O Users and groups have access to all connected resources.						
	QuickSight can allow or deny access to all users and groups by default, when an individual access control is not in effect for a particular user or group						
	This feature is available for QuickSight enterprise account. See details here						
	Resource access for individual users and groups						
	Resource access is controlled by assigning IAM policies.						

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.

ke your existing AWS data and users available in QuickSight.	aker your misting AWS data and users available in QuickSight.   Select the buckets that you want QuickSight to be able to access.   Concol   Amazon Rothint   Amazon Rothint   Amazon Staff Backets solectal)   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the buckets that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Select the bucket that you want QuickSight to be able to access.   Selec	Beter ensisting AWS data and users available in QuickSight.     Select the buckets that you want QuickSight to be able to access.       Select the buckets that you want QuickSight.     Select the buckets that you want QuickSight to be able to access.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that you want QuickSight.       Image: Comparison RDS     Select the buckets that the top particle top	feature.
waters and autodiscovery for these resources       Selected buckets have read only permissions by default. However, you must give write permissions for Athena Workgroup features.         I Amazon Rechift       Is stead:         I Amazon Rock here       Is stead:         I Amazon S3 I Suckets selected)       Is stead:         I Amazon XHena       Write permission for Athena Workgroup         I Amazon Athena       Is stead:         I Amazon XHena       Is stead:         I Amazon S3 Stoage Analytics       Is stead:         I Amazon S3 Stoage Analytics       Is tead:         I Amazon S3 Stoage Analytics       Is tead:         I Amazon SageMaker       Is tead:         I Amazon SageMaker       Is tead:         I tead:       Is tead:         I tead:       Is tead:         I Amazon SageMaker       Is tead:         I tead:       Is tead:         I tead:       Is tead:         I Amazon SageMaker       Is tead:         I Amazon SageMaker       Is tead:         I Ama	Betcade Mackets have read only permissions by default. However, you must give write permissions for Athena Workgroup Fastures.         Image: Annance Redshift:	Biowaccess and autodicovery for these resources     Selected buckets have read only permissions by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Select all       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Anazon Redshift:     Status by default. However, you must give write permissions for Athena Workgroup I       Image: Athena Factor Backet and the anazon Status for Quite     Image: Athena Factor Backet and the anazon Status for Quite       Image: Athena Backet Anazon Status for Quite     Image: Athena Factor Backet and the anazon Status for Quite       Image: Athena Backet Anazon Status for Quite     Image: Athena Factor Backet and the anazon Status for Quite	feature.
	Cancel     Cancel     Sett all     Sett	Compares Redolfrit     Compares Redolfri	2
Imacon RDS     S3 Bucket     Write permission for Athena Workgroup       Imacon RDS     S3 Bucket     Write permission for Athena Workgroup       Imacon RDS     Imacon RDS     Imacon RDS       Imacon Athena     Imacon RDS Athena Yorkgroup     Imacon RDS       Imacon Athena     Imacon RDS     Imacon RDS       Imacon RDS     Imacon RDS	Statuscie     Statuscie     Write permission for Athena Workgroup       P AMA     Identified athena results 1730874955       Amazon Athena     Statuscie       Amazon S3 Storage Analytics     Sagemaker-studio-41730874955       Amazon S3 Storage Analytics     Sagemaker-studio-41730874955       Amazon S3 Storage Analytics     Itest-classification-backend-athenaresultsbucket-1cvollinkjhf0       Amazon Timestram     Itest-classification-backend-athenaresultsbucket-430mmq46jpip       Store     Cancel	Amazon RDS     Amazon RDS     S3 Bucket     Write permission For Athena Workgroup     Amazon RDS     Amazo	>
W M       Idm:sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample.sample-states/sample-statestates-sample-statestates-sample-states-states-s	Image: Mark Signed and Signed Andrea Signed and Signed Andrea Signed and Signed Andrea Signed and Signed Andrea Signed and Signed Andrea Signe		
Image: State Stat	Anazon 53 (E buckets sloetschi)     Salect 53 buckets     asgemaker-studio-417308874955 bim?regaugg     asgemaker-studio-417308874955     asgemaker-studio-417308874955     asgemaker-studio-417308874955     text-classification-backend-athmaneut/tbucket-1cvedisribph0     def Anazon 7 SageMaker     def Anazon SageMaker     text-classification-backend-athmaneut/tbucket-foredisplipitie     text-classification-backend-athmaneut/tbucket-foredisplipitie     def Anazon SageMaker     text-classification-backend-athmaneut/tbucket-foredisplipitie     text-classification-backend-athmaneut/tbucke	Constant S3 (E Nockets selected)     aligemaker-studio-417308374955-biom7eqacougq     defamation Athena     Nates uny poly of home the right Ansator S3 Subjects for Quick     text-classification-backend-athenaresultsbucket-1 cooldinbjthD	
Annacon Athena Nake unsy yooker doors the right Amazon 53 Suckets for Quick     Amazon 53 Suckets for Quick     Amazon 53 Suckets Analystics     Amazon 52 Suckets Analystics     Amazon OpenSearch Service     Amazon SageMaker     Amazon Timestream      Size     Cancet     Cancet     Cancet     Cancet	Arnazon Athena Pala aus you'd charan the right Annazon S3 backets for Quick     Annazon Athena Pala aus you'd charan the right Annazon S3 backets for Quick     Annazon S30 Marker     Annazon S30 Marker     Annazon Timestram      Sove     Cancel     Innob	Amazon Athena	
Image: Strange Analytics     text-classification-backed-athenersultsbucket-1oodlarklph0       Image: Strange Analytics     text-classification-backed-athenersultsbucket-1oodlarklph0       Image: Strange Analytics     text-classification-backed-athenersultsbucket-45mmp46jpb       Image: Strange Analytics     text-classification-backed-athenersultsbucket-45mmp46jpb       Image: Strange Analytics     text-classification-backed-athenersultsbucket-qsc5gl01a82       Image: Strange Analytics     text-classification-backed-tweetsbucket-11hs1ig06kad4       Image: Strange Analytics     Ext-classification-backed-tweetsbucket-11hs1ig06kad4       Strate     Entect	Amazon S3 Storage Analytics     Amazon S3 Storage Analytics     Amazon S4 Storage Analyti	text-classification-backend-athenaresultsbucket-1cvodlsnbjxh0     text-classification-backend-athenaresultsbucket-1cvodlsnbjxh0	
WYS IoT Analytiss     I test-classification-backend-athenaresultsbucket-45nmmq46gipio       Amazon OpenSearch Service     I test-classification-backend-loggingbucket-ggr5gi01882       Imazon SageMaker     I test-classification-backend-loggingbucket-ggr5gi01882       Imazon Timestream     I test-classification-backend-tweetsbucket-11w3lq05lade4       Size     Cancet	2 AVX5 lof Xnalytics     rext-classification-backend-athenareult/sbucket-45mmp44jgio       2 AVX5 lof Xnalytics     rext-classification-backend-athenareult/sbucket-gor:gip(lat)2       3 AVX5 norman     rext-classification-backend-ethenareult/sbucket-11w3lcp6kade4       3 By Annazon Timestram     Cancel		
Image: Same Service     Inter-classification-backend-loggingbucket-ggr5ig01682       Image: Same Service     Inter-classification-backend-tweetsbucket-11w3iq05iad64       Image: Same Service     Inter-classification-backend-tweetsbucket-11w3iq05iad64       Same Cancel     Finish	Image: Amazen OpenSearch Service     Image: text-classification-backend-loggingbacket-ggc5j0li682       Image: Amazen Timestream     Image: text-classification-backend-tweetsbacket-11w3lq06kade4       Sore     Cancel	🖉 AWS IoT Analytics 🛛 🗸 text-classification-backend-athenaresultsbucket-4Snmmq46jgio	
Amazon SagetAaker     z test-classification-backend-tweetsbucket-11w3lq06kade4     Gancet     Gancet     Gancet	Same     Cancel	Mrazon OpenSearch Service text-classification-backend-loggingbucket-ggc5jg0ils82	
Sive Cancel Finish	Save Cancel Findsh	(iii) Amazon SageMaker     (iv) Amazon SageMaker     (iv) text-classification-backend-tweetsbucket-11w3lq06kade4	
Save Cancel Finida	Save Cancel Finish	📴 Amazon Timestream	
		Save Cancel Cancel	Finish

### 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.

ZuickSight			ዳ				
Datasets Create a Dataset FROM NEW DATA SOURCES Upload a file (cox, so, stf, stf, sta, peri) Salesforce Connect to Salest	New Athena data source Data source name Inister-dc Athana workgroup (primary ] Validated SSL is enabled	X Crote data source	SPICE capacity for this region: STMB of 11GB				
RDS Redshift Auto-discovered	Redshift Manual connect	MySQL	PostgreSQL				
ORACLE SQL Server	Aurora	A MariaDB	Presto				
Spork <sup>7</sup> Spark Teradata Provided by Terad	snowflake	AWS IoT Analytics	Amazon OpenSearch Ser Successor to Amazon Elasticsearch Ser				
Timestream Exasol	GitHub	Twitter	Jira				
NOW ServiceNow Adobe Analy	tics						
FROM EXISTING DATA SOURCES							

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

VuickSight			٨
Datasets Create a Dataset FROM NEW DATA SOURCES	Choose your table twister-dc Catalog: contain sets of databases.	×	SPICE capacity for this region: 51MB of 11GB
Upload a file (csv. tsv. dt, dt, dsv, pon)	andoaceasay Tababase: contain sets of tables. tweets Tables: contain the data you can visualize.	~	Athena
RDS Redshift Auto-discovered	<ul> <li>phrases</li> <li>tweets</li> </ul>	QL	PostgreSQL
ORACLE SQL Server	Edit/Preview data Use custom SQL	Select	Presto
Spark Spark Teradata	Snowflake	AWS IoT Analytics	Amazon OpenSearch Ser Successor to Amazon Elasticsearch Ser
Timestream X Exasol	GitHub	Twitter	Jira
NOW ServiceNow Adobe Analy	ics		
FROM EXISTING DATA SOURCES			

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.

QuickSight				٨
Datasets Create a Dataset FROM NEW DATA SOURCES Upload a file (cre. tre. off. eff. stor, port)	Finish dat Table: Data source: Schema: Salesforce Convect to Salesfore Convect to Salesfore	aset creation tweets tweets tweets SPRCE for quicker analytics	x	SPICE capacity for this region: \$1MB of 11GB
RDS	Redshift Auto-discovered	w data Manual connect	Visustize y-QL	PostgreSQL
ORACLE	SQL Server	Aurora	AmariaDB	Presto
spaik Spark	Provided by Teradata	Snowflake	AWS IoT Analytics	Amazon OpenSearch Ser Successor to Amazon Elasticsearch Ser
Timestream	Exasol	GitHub	<b>Y</b> Twitter	Jira
NOW ServiceNow	Adobe Analytics			
FROM EXISTING DATA SOURCES				

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.

QuickSight		Å
New dataset	weets ×	SPICE capacity for this region: 51MB of 11GB
Your Datasets	sicc Dataset 10.2MB	
phrases weets weets	100% success 11290 rows were imported to SPICE 0 rows were skipped	s Pipeline Business Review
	Last refreshed: 9 days ago View history	
Web and Social Media A	Refresh now Schedule a refresh	
	<ul> <li>Incremental refresh configuration</li> <li>Date column: partition_timestamp</li> <li>Window size: 1 hour</li> </ul>	
	Data source name: TwisterDC Database name: ATHENA	
	Delete dataset Share	
	Edit dataset Duplicate dataset Create analysis Use in a new dataset	

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".

🛛 🖬 Quid	kSight   介 tweets analysis						٨
+ ADD ~							
00a Visualize	Dataset 🖉	Field wells	Choose dataset to add		× Q		
<b>∀</b> Filter	Fields list Search fields Q	Sheet 1 V +	• phrases	🖊 Athena	SHCE		
141 Parameters	<ul> <li>category_type</li> <li>category_type_model_result</li> <li>category_type_score</li> <li>count</li> </ul>		People Overview     Sales Pipeline	🏚 S3	SPICE		
R Actions	Contraction Contraction Contraction		Business Review	🏚 53	SPICE		
or Themes	Iongitude     model     motification	o Choose 1 or more fields	Cancel	<b>10</b> 23	Select		
Settings	partition_timestamp     platform     platform     semiment     source     text     text_clean     text_clean     text_sem	c					
- R	· use Valual types · · · · · · · · · · · · · · · · · · ·						
Community							_

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.

