

ⓘ This prerelease documentation is confidential and is provided under the terms of your nondisclosure agreement with Amazon Web Services (AWS) or other agreement governing your receipt of AWS confidential information.

Walkthrough: Using Amazon QuickSight Q to create a Topic and Make Natural-Language Friendly Datasets

[PDF \(amazon-quicksight-user.pdf#quicksight-q-walkthrough\)](#)

[Kindle \(https://www.amazon.com/dp/B076453HGY\)](https://www.amazon.com/dp/B076453HGY)

[RSS \(amazon-quicksight-doc-release-notes.rss\)](#)

Amazon QuickSight Q is a natural language query (NLQ) tool that uses metadata and machine learning algorithms to answer business questions instantly. Amazon QuickSight Q uses Topics to link datasets and metadata so you can ask questions naturally.

This walkthrough will cover how to conceptualize, develop, and test a Amazon QuickSight Q topic for a music streaming service. We'll use a provided data model, dataset, and data dictionary to begin using Amazon QuickSight Q topics as quickly as possible.

To demonstrate how to use Amazon QuickSight Q and Topics, we'll walk through how to develop Topics and make datasets natural language friendly. Then, we'll analyze and test Topics to produce visualizations instantly.

By the end of the walkthrough, you'll have an Amazon QuickSight Q Topic similar to [tk - Example section], which visualizes subscription data.

This walkthrough should take approximately 45 minutes and will cover the following steps:

1. Step 1: Anticipate Questions and Identify Data

We'll use questions to identify necessary data fields and define key terms. We'll also walk through leveraging your data model for answering questions.

2. Step 2: Create a Topic

We'll determine potential topics from our list of questions and define the scope of our topic.

3. Step 3: Update Your Topic to be Natural Language Friendly

We'll add friendly names, synonyms, and metadata to ensure Amazon QuickSight Q can make connections between your fields.

4. Step 4: Test Your Topic

We'll go over potential questions to test with and improve Amazon QuickSight Q's answers.

Prerequisites

Before you begin this tutorial, you should have completed Getting Started with QuickSight Q and downloaded the files we will be using in this walkthrough:

- Users 2018-2019.csv

- Subscribers 2018-2019.csv
- Artists 2018-2019.csv
- Songs 2018-2019.csv
- Streams 2018-2019.csv
- Stakeholder Activity Topic - Data Dictionary 2018-2019.csv
- Stakeholder Activity Topic - List of Questions

Walkthrough Scenario

Imagine you are an analyst at a music streaming company hoping to improve subscriptions and analyze traffic across platforms (mobile, desktop, app). With data on user traffic and campaigns, you want to allow your coworkers to efficiently get answers to their questions about the data.

Music Streaming Service Example

The data model is simplified for the purpose of this walkthrough. The data model has 5 tables that provide crucial information to Amazon QuickSight Q:

- Users Table (Users 2018-2019.csv) – This table records the information about stakeholders with accounts for the music streaming service. The following information about each individual stakeholder is provided: User ID, Country, Join Date, and Age.
- Subscribers Table (Subscribers 2018-2019.csv) – This table records the information about stakeholders with paid annual memberships for the music streaming service. The following information about each individual subscriber is provided: Subscriber ID, Cost, Start Date, User ID.
- Artists Table (Artists 2018-2019.csv) – This table records the information about music content creators for the music streaming service. The following information about each individual artist is provided: Artist ID, Name, Monthly Listeners, Monthly Payment.
- Songs Table (Songs 2018-2019.csv) – This table records the information about songs for the music streaming service. The following information about each individual song is provided: Song ID, Release Date, Genre, Total Streams, Royalties, and Artist ID.
- Streams Table (Streams 2018-2019.csv) – This table records the information about streams for the music streaming service. The following information about each individual stream is provided: Date and Time, Platform, Song ID, User ID.
- Stakeholder Activity Topic - Data Dictionary 2018-2019.csv – This CSV file records information about each table and field with descriptions about each field's metadata.
- Stakeholder Activity Topic - List of Questions – This CSV file records the questions we will develop in this walkthrough and potential testing.

Using the data model above, we will create a topic about Stakeholder Activity from 2018-2019. For the purpose of this walkthrough, we will define stakeholder activity as an interaction with the content being streamed or the stakeholder's account.

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.