

› Earnings Call Trending Themes In Your Snowflake

Save hours catching up with earnings calls by getting trending topics in near real-time.

This recipe demonstrates how to store earnings call themes extracted via FactSet's Natural Language Processing (NLP) API to your Snowflake. It also helps you visualize data using word cloud as an example. Automating the process of extracting themes from transcripts allows you to receive updates more efficiently and rapidly.



To start, please visit the [FactSet Snowflake marketplace](#) or contact a sales representative to start a trial. Once you get access, create a Snowflake Stream out of FactSet's CallStreet transcripts table and a local table to store data. Then, write a Python script to feed earnings call transcripts into the FactSet NLP API and retrieve earnings call themes.

Snowflake is not the only option to extract CallStreet transcripts. Please ask a FactSet sales representative to learn more about other options if you prefer a different delivery mechanism to consume the transcripts.

Time Required: 2+ hours **Difficulty:** Moderate

STEPS

1 SUBSCRIBE TO FACTSET SNOWFLAKE

Request a trial from the [FactSet Snowflake marketplace](#).

2 CREATE A SNOWFLAKE TABLE

Create a table to store the transcript themes.

3 FETCH TRANSCRIPT IN SNOWFLAKE

Create a Snowflake Stream to connect to FactSet's CallStreet transcripts table.

4 EXTRACT TRANSCRIPT THEMES

Write a Python script to extract transcripts from the stream and run FactSet NLP API to return the themes.

5 STORE AND VISUALIZE THE THEMES

Save the themes into the local table created in Step 2 and use a word cloud to visualize the themes.

Ingredients

[FactSet Natural Language Processing API](#)
[Snowflake](#)
[Python WordCloud Library](#)

1 SUBSCRIBE TO FACTSET SNOWFLAKE

Read [this document](#) to learn how to access FactSet Snowflake and visit the [FactSet Snowflake marketplace](#) to request a trial to start the subscription process.

2 CREATE A SNOWFLAKE TABLE

If you have not yet created a table that will store transcript themes, create a table by running the **create table** Snowflake SQL query. Consider any columns that should be included in the table before creating a table. If you need an idea, check out the image in Step 3. Confirm that the table is created.

3 FETCH TRANSCRIPT IN SNOWFLAKE

We will use [Snowpark Python API](#) to interact with Snowflake from our Python script. If you haven't previously installed the Snowpark API package, please follow [Snowflake's official documentation](#).

Once FactSet grants permission, run the **create stream** Snowflake SQL query to retrieve transcripts data from the FactSet Snowflake table, `FDS_SEMI_STRUCTURED.TRANSCRIPTS.TRANSCRIPTS`. To learn about Snowflake Stream service, please check out their [Introduction to Streams page](#).

	ID	DATE	RAW_XML	TRANSCRIPT_TYPE	VERSION_ID	METADATA\$ACTION	METADATA\$ISUPDATE...	METADATA\$ROW_ID
1	2824035	2023-04-24 00:00:00.000	<transcript id="2824035" product="Corre...	CorrectedTranscript	6559305	INSERT	FALSE	51f7444eba3fa78f754e2234c776a45e9325fb20
2	2818734	2023-04-20 00:00:00.000	<transcript id="2818734" product="Corre...	CorrectedTranscript	6559216	INSERT	FALSE	6ab66488363d8c65b9ec5b49f1fd96a774702fd1

Then, build a Python script to

- Create a [Snowflake Session](#) with Snowpark API with your credential and database information.
- Fetch the transcripts using the session and extract *ID* and *RAW_XML* columns.
- To filter the transcripts based on the company's entity ID, we must join the FactSet table `FDS.EVT_V1.CE_REPORTS`, and the Stream we created.
- Extract the **event title** and **transcript** from the management discussion section.

4 EXTRACT TRANSCRIPT THEMES

Run FactSet NLP API with the transcript from Step 3 and get the themes. First, you should set up an authentication to use the FactSet API. Check out the [Getting Started page](#) for more information. After successfully configuring the credentials, use the *themes* endpoint to generate the transcript themes.

```
{
  "data": [
    {
      "themeText": "different business areas",
      "themeScore": 0.92
    },
    ...
  ],
  "meta": {}
}
```

5 STORE AND VISUALIZE THE THEMES

Create a Snowflake data frame with the themes and merge the data frame to the table created in Step 2. The themes returned in Step 4 is the list of JSON objects, which is difficult to read the trends. Use the *WordCloud* Python library or any other visualization package you prefer to consume the data easily. Then, it can be distributed either via email or an internal tool.

	ID	DATE	FACTSET_ENTITY_ID	TITLE	THEMES	...
1	2818734	2023-04-20 00:00:00.000	0017G5-E	Q4 2023 Earnings Call	["incremental year-over-year revenue", "company's actual future results", "Seemap", "significant new business", "one...	
2	2824035	2023-04-24 00:00:00.000	05HLL9-E	Q1 2023 Earnings Call	["significant land sales", "different business areas", "second element", "double-digit increase", "positive contribution"	