Twitter Sentiment Analysis

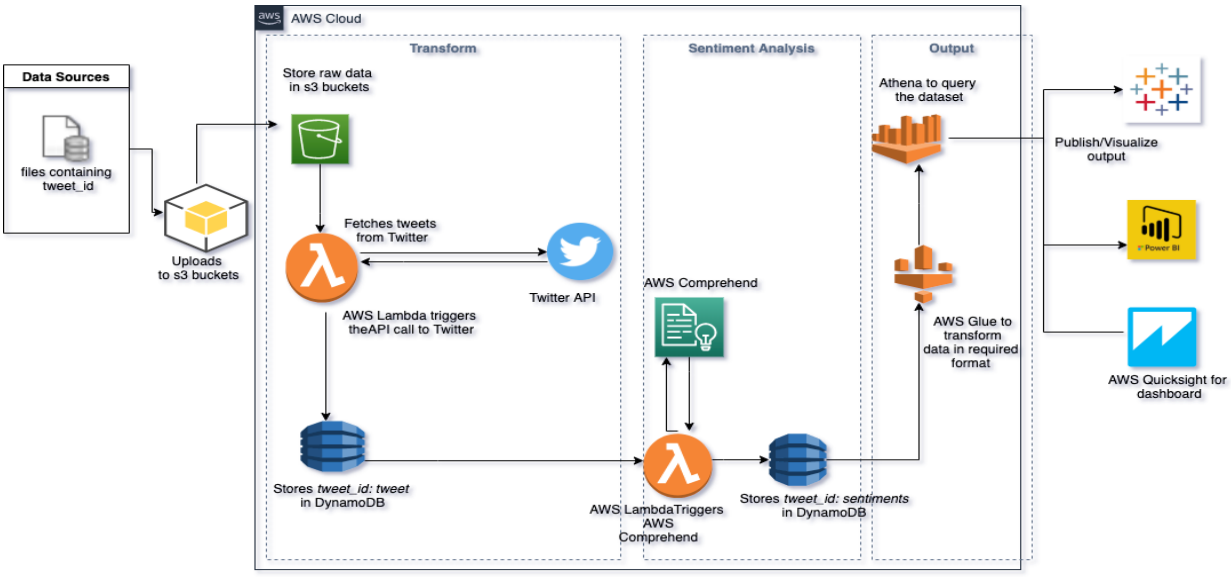
Goal

We want to do sentiment analysis on the various tweets for certain global events of importance in the past

Flow

1. The input compressed files are decompressed and uploaded to AWS S3 to be processed by downstream systems.
2. *tweet_ids* are read from the S3 bucket and corresponding tweets are fetched from Twitter by API calls .
3. AWS Lambda triggers API call to twitter once there is a dataset in S3.
4. The *tweet_ids* along with the tweets are stored in DynamoDB.
5. AWS Lambda triggers comprehend to fetch the sentiment associated with the tweets.
6. The tweets are read from DynamoDb and AWS Comprehend is used to derive the sentiments for the tweets.
7. *tweet_ids* along with the sentiment are stored in DynamoDB (Key-value pair).
8. AWS Glue helps to transform the data in a particular format.
9. AWS Athena will be used to query the dataset for various analysis.
10. AWS Quicksight will be used for visualizations.

Architecture Diagram



Note: The architecture diagram can change in the next iteration