

Data Insights

Sentiment Analysis Presentation

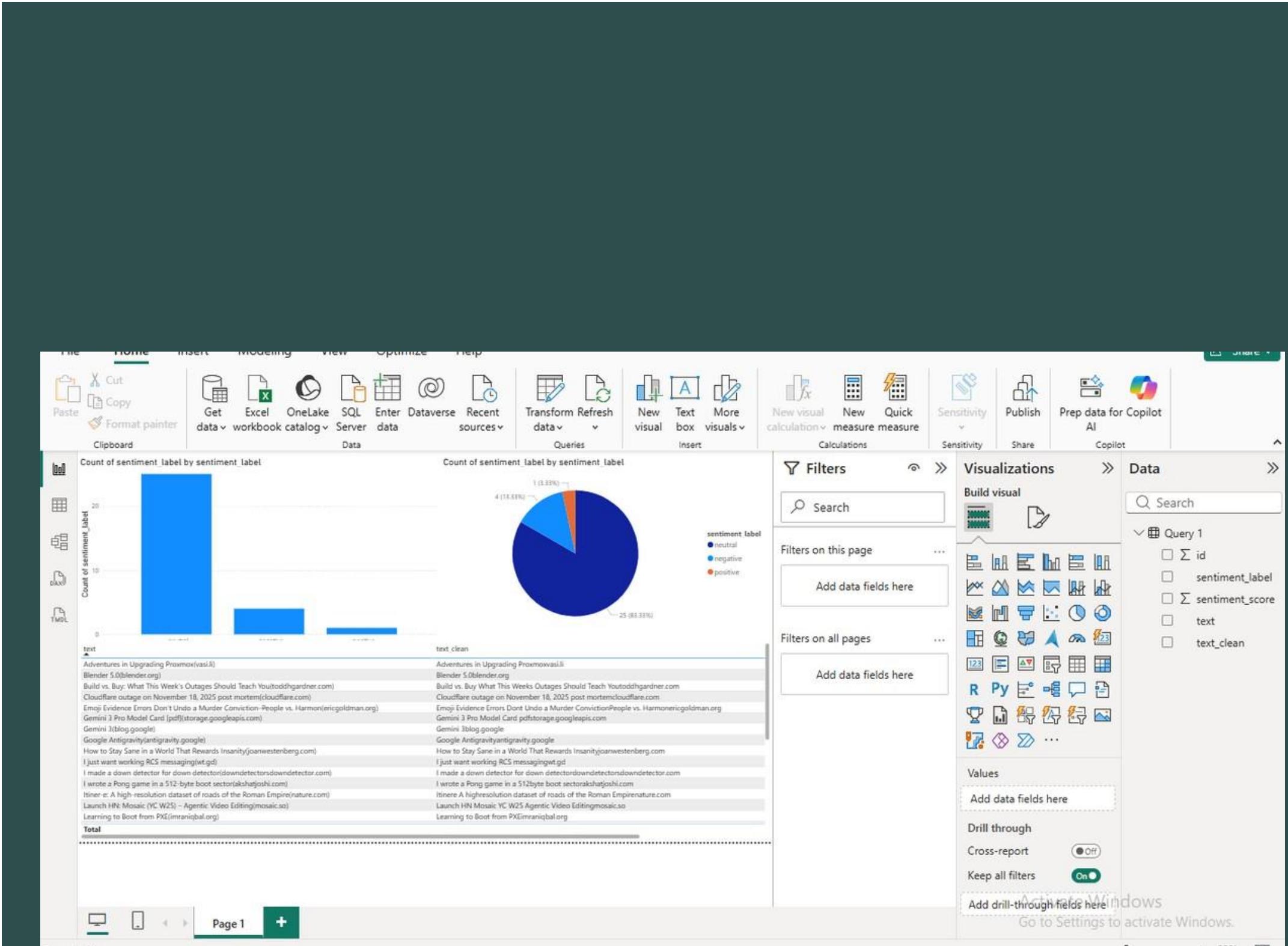
Group Members: Chandresh Patel (200595392)
Nancy Jangwal (200591978)



OVERVIEW

Project Workflow Overview

This project demonstrates an end-to-end sentiment analysis workflow using Azure services. The workflow includes: preprocessing text data, uploading CSV to Blob Storage, creating an ADF pipeline, loading data into SQL Database, and visualizing results using Power BI.



BLOB STORAGE

Uploading CSV Files

The processed final_sentiment.csv file was uploaded to Azure Blob Storage inside a dedicated container. This file serves as the input source for the Azure Data Factory pipeline.

The screenshot shows the Microsoft Azure Blob Storage interface. The top navigation bar includes 'Microsoft Azure', a search bar ('Search resources, services, and docs (G+/)'), and a user account section ('Copilot', '200632295@student.ge...', 'GEORGIAN COLLEGE (GEORGIAN...)'). The main area shows a container named 'sentimentdata'. The left sidebar has options: 'Overview' (which is selected), 'Diagnose and solve problems', 'Access Control (IAM)', and 'Settings'. The right pane displays a table of blobs:

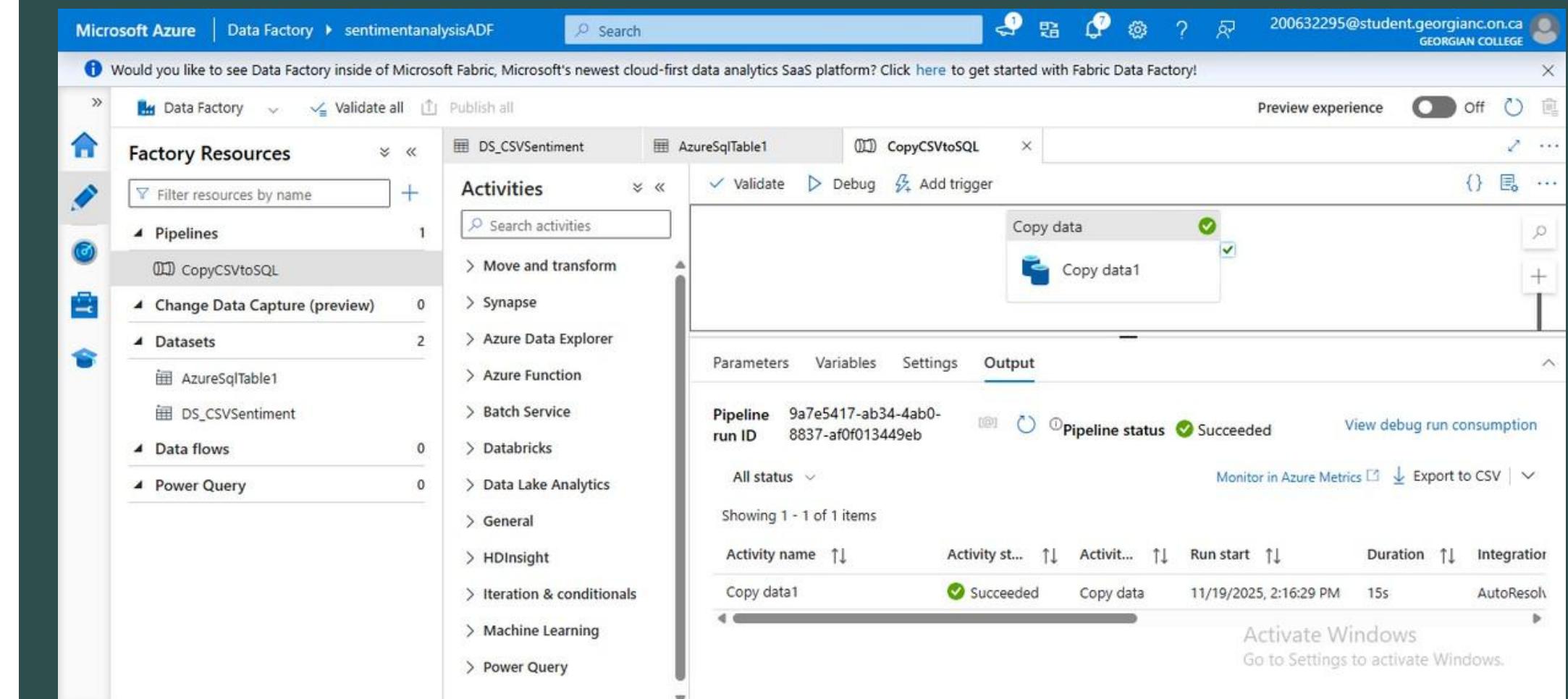
Name	Last modified	Access tier	Blob type	Size	Lease state
final_sentiment.csv	11/19/2025, 2:09:01 PM	Hot (Inferred)	Block blob	4.2 KB	Available

ADF PIPELINE

Pipeline Setup Overview

This section explains the ADF pipeline setup, detailing the Copy Data activity and how to configure linked services for seamless data movement.

- An ADF Copy Data pipeline was created to automate the transfer of the CSV file from Blob Storage into Azure SQL Database. The pipeline includes both source and sink configurations.

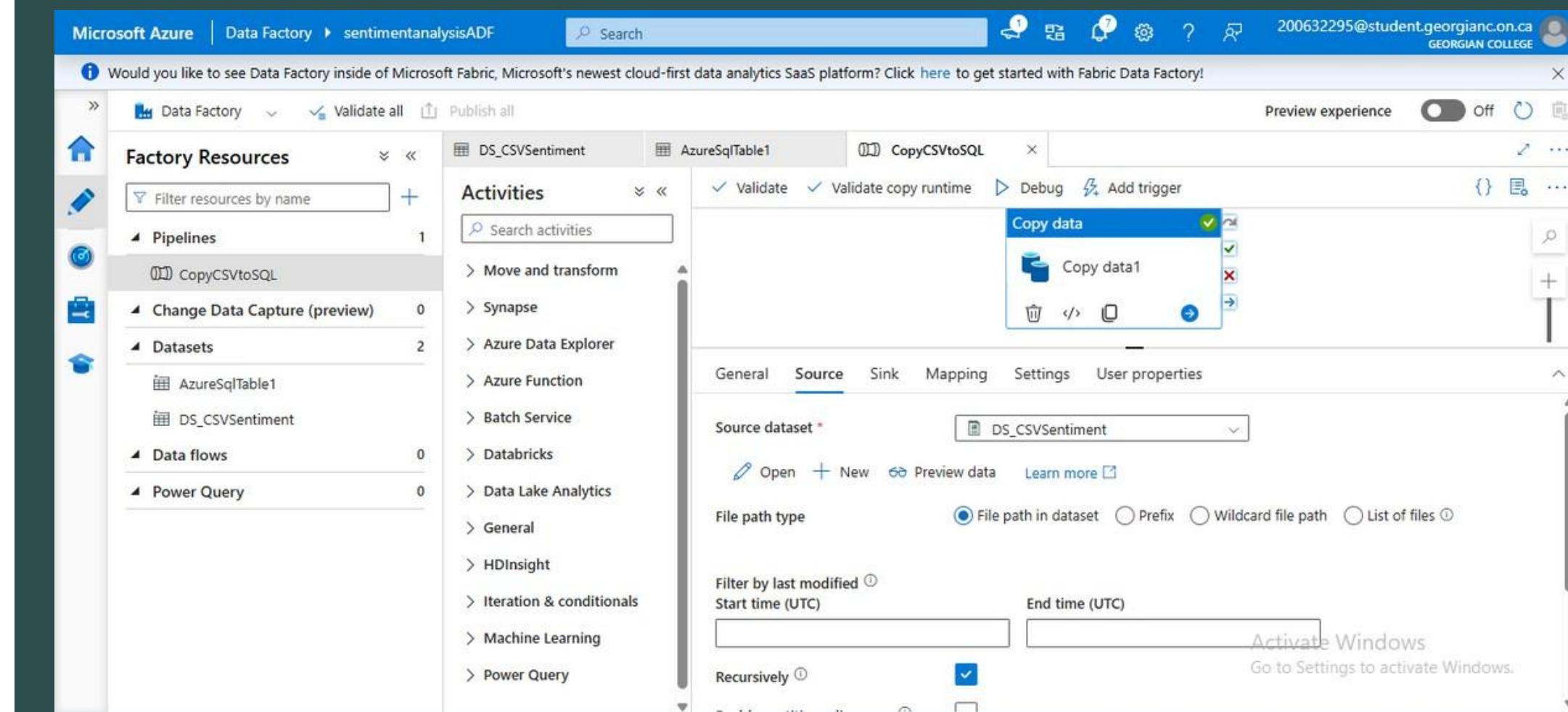


SOURCE DATASET

ADF Source Configuration

In this section, we detail the **setup of the source dataset** for our ADF pipeline, including CSV handling and schema import processes.

- The source dataset was configured to read the CSV from Blob Storage. Schema was imported automatically from the file and delimiter was set correctly.

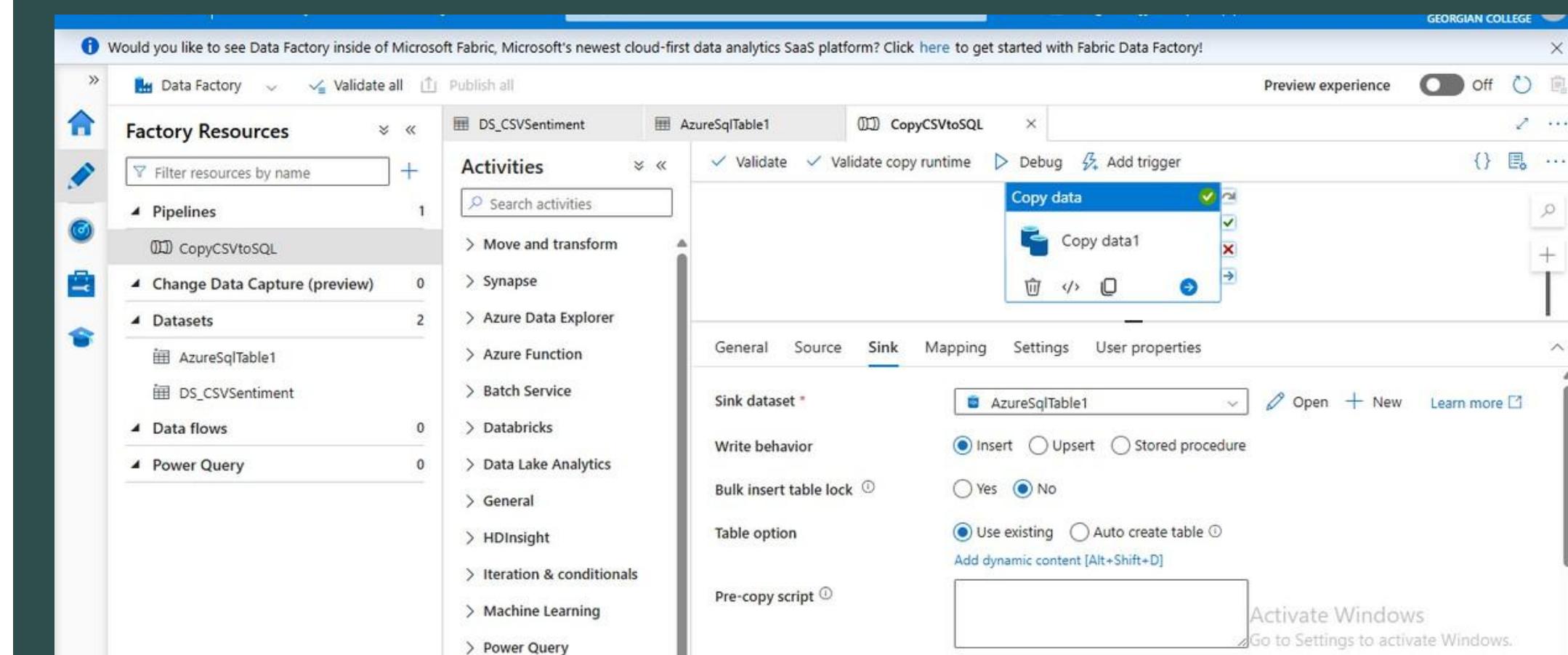


SINK DATASET

Table Mapping Overview

This section details the **mapping of tables** between the ADF pipeline and the SQL database, ensuring accurate data flow and integrity throughout the process.

- The sink dataset was connected to the Azure SQL Database and mapped to the SentimentOutput table. This allowed the CSV data to be inserted into SQL Server.



MAPPING

Column Mapping Overview

This section outlines the process of mapping columns between the source CSV data and the destination SQL database to ensure accurate data transfer.

- The mapping step ensures that every CSV column is correctly matched to the SQL table: `text`, `text_clean`, `sentiment_score`, `sentiment_label`.

The screenshot shows the Microsoft Azure Data Factory Pipeline Editor. On the left, the 'Factory Resources' sidebar lists 'Pipelines', 'Datasets', and 'Data flows'. The main area displays a pipeline named 'CopyCSVtoSQL' with two stages: 'DS_CSVSentiment' (source) and 'AzureSqlTable1' (sink). A 'Copy data' activity is selected. The 'Mapping' tab is active, showing four columns being mapped from the source to the sink:

Source	Type	Destination	Type
<code>text</code>	abc String	<code>text</code>	abc String
<code>text_clean</code>	abc String	<code>text_clean</code>	abc String
<code>sentiment_score</code>	abc String	<code>sentiment_score</code>	abc String
<code>sentiment_label</code>	abc String	<code>sentiment_label</code>	abc String

EXECUTION SUCCESS

Pipeline Run Results

The pipeline successfully executed, resulting in the insertion of multiple rows into the SQL database. This indicates correct data flow throughout the ETL process.

This screenshot shows the Microsoft Fabric Pipeline runs page. The left sidebar includes options like Dashboards, Runs, Pipeline runs (which is selected), Trigger runs, Change Data Capture, Runtimes & sessions, Integration runtimes, Data flow debug, Notifications, and Alerts & metrics. The main area is titled "Pipeline runs" and shows a table of recent runs. The table columns are: Pipeline name, Run start, Run end, Duration, Status, Triggered by, Run ID, and Parameters. All listed runs are marked as "Succeeded" and triggered manually. The status bar at the bottom right indicates "Last refreshed 0 minutes ago".

Pipeline name	Run start	Run end	Duration	Status	Triggered by	Run ID	Parameters
CopyCSVtoSQL	11/19/2025, 2:16:2	11/19/2025, 2:16:4	23s	Succeeded	Manual trigger	9a7e5417-ab34-4ab...	
CopyCSVtoSQL	11/19/2025, 2:08:4	11/19/2025, 2:09:0	22s	Succeeded	Manual trigger	63acce56-ce90-4968...	
CopyCSVtoSQL	11/19/2025, 2:07:4	11/19/2025, 2:08:0	20s	Succeeded	Manual trigger	8269a73b-22e6-430...	
CopyCSVtoSQL	11/19/2025, 1:36:4	11/19/2025, 1:37:0	21s	Succeeded	Manual trigger	186d78b2-371e-4f7...	
CopyCSVtoSQL	11/19/2025, 1:32:2	11/19/2025, 1:32:4	20s	Succeeded	Manual trigger	dd56e31-e58b-47cf...	

Activate Windows
Go to Settings to activate Windows.

This screenshot shows the Microsoft Fabric Activity runs page for a specific pipeline run. The left sidebar is identical to the previous screenshot. The main area shows a summary for a recent debug run of the "CopyCSVtoSQL" pipeline. It highlights a single activity named "Copy data1" which succeeded. Below this, the "Activity runs" section lists one item with details: Pipeline run ID, Activity name, Status, Start time, Duration, and Integration runtime. The status bar at the bottom right indicates "Last refreshed 0 minutes ago".

Activity name	Activity st...	Activit...	Run start	Duration	Integration runtime
Copy data1	Succeeded	Copy data	11/19/2025, 2:16:29 PM	15s	AutoResolveIntegrationRuntime (Canada Central)

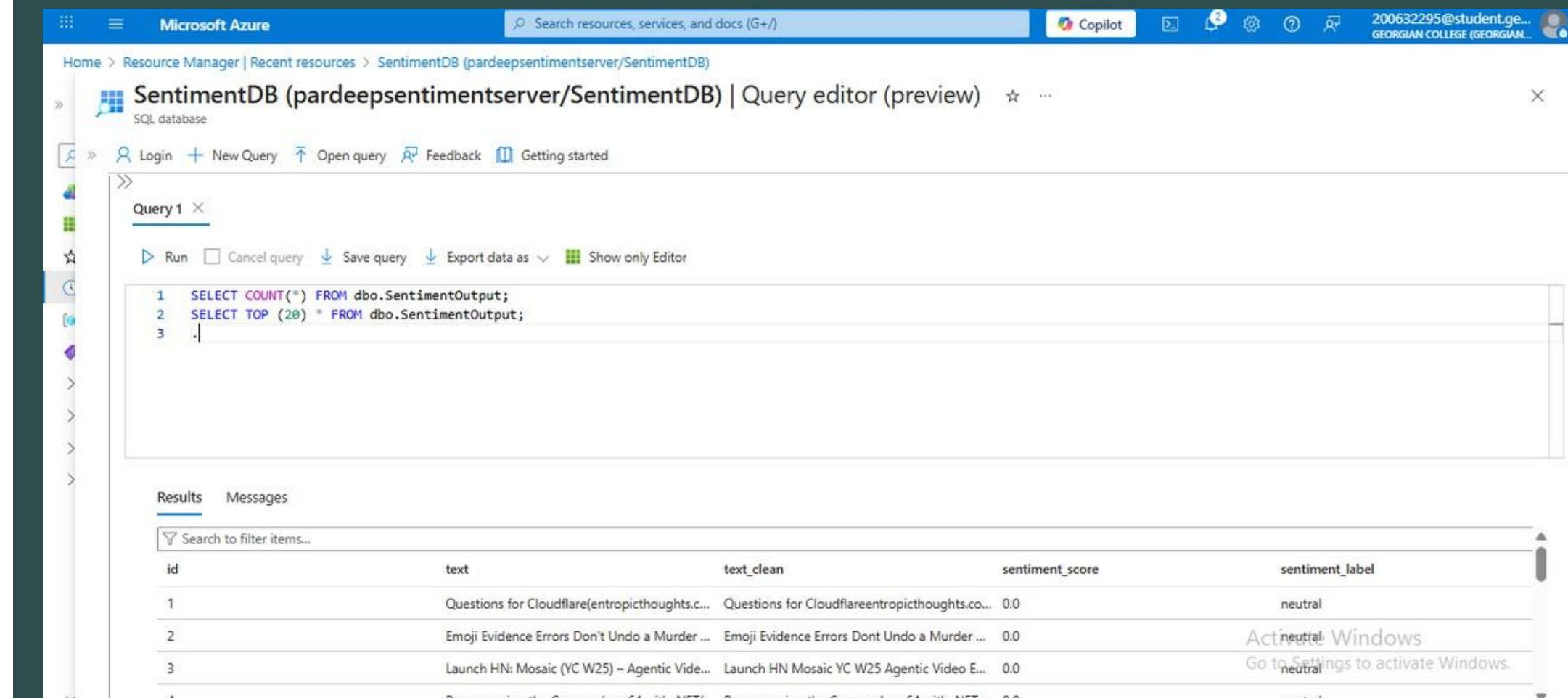
Activate Windows
Go to Settings to activate Windows.

SQL VALIDATION

SQL Query Results

This section highlights the importance of validating SQL data after extraction. Ensure accuracy by reviewing query results to confirm successful data loads into the database.

- Data was validated in Azure SQL by running queries such as `SELECT COUNT(*)` and `SELECT TOP(20)`. This confirms successful data insertion.



The screenshot shows the Microsoft Azure Query editor interface. The title bar reads "SentimentDB (pardeepsentimentserver/SentimentDB) | Query editor (preview)". The query pane contains the following SQL code:

```
1 SELECT COUNT(*) FROM dbo.SentimentOutput;
2 SELECT TOP (20) * FROM dbo.SentimentOutput;
3 .
```

The results pane displays the output of the second query, showing four rows of sentiment analysis data:

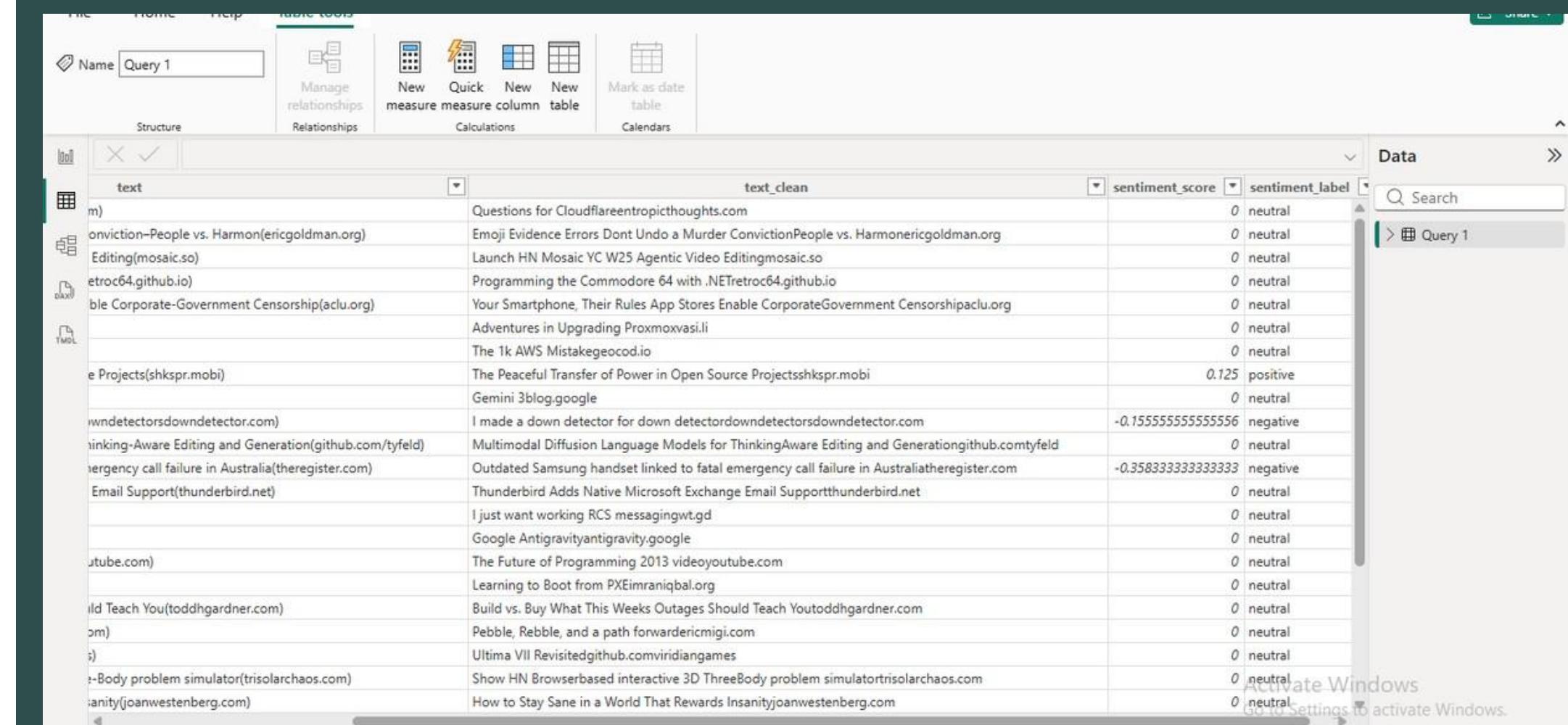
id	text	text_clean	sentiment_score	sentiment_label
1	Questions for Cloudflare(entropicthoughts.c...	Questions for Cloudflareentropicthoughts.co...	0.0	neutral
2	Emoji Evidence Errors Don't Undo a Murder ...	Emoji Evidence Errors Dont Undo a Murder ...	0.0	neutral
3	Launch HN: Mosaic (YC W25) – Agentic Vide...	Launch HN Mosaic YC W25 Agentic Video E...	0.0	neutral
4	Programming the Commodore 64 with .NET/...	Programming the Commodore 64 with NETc...	0.0	neutral

POWER BI

Importing Data from SQL

In this section, we will discuss how to seamlessly import data from the SQL database into Power BI for effective data visualization and reporting.

- Power BI was connected to the SQL Database table. The SentimentOutput table was imported successfully for further visualization.

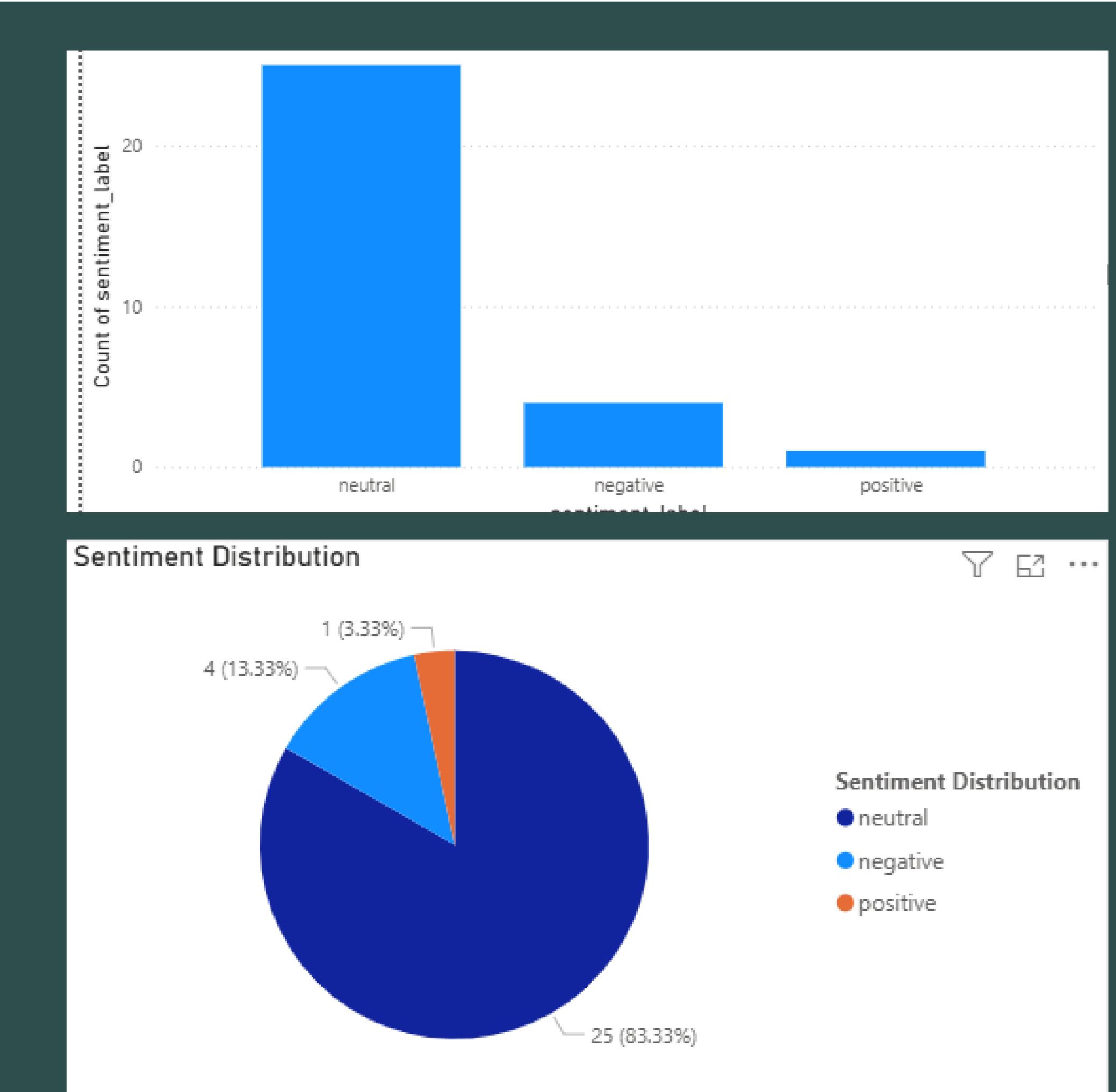


text	text_clean	sentiment_score	sentiment_label
m)	Questions for Cloudflareentropicthoughts.com	0	neutral
onviction-People vs. Harmon(ericgoldman.org)	Emoji Evidence Errors Dont Undo a Murder ConvictionPeople vs. Harmonericgoldman.org	0	neutral
Editing(mosaic.so)	Launch HN Mosaic YC W25 Agentic Video Editingmosaic.so	0	neutral
etroc64.github.io)	Programming the Commodore 64 with .NETretroc64.github.io	0	neutral
ble Corporate-Government Censorship(aclu.org)	Your Smartphone, Their Rules App Stores Enable CorporateGovernment Censorshipaclu.org	0	neutral
	Adventures in Upgrading Proxmoxvasili	0	neutral
The 1k AWS Mistakegeocod.io	The 1k AWS Mistakegeocod.io	0	neutral
e Projects(shkspr.mobi)	The Peaceful Transfer of Power in Open Source Projectsshkspr.mobi	0.125	positive
wndetectorsdowndetector.com)	Gemini 3blog.google	0	neutral
hinking-Aware Editing and Generation(github.com/tyfeld)	I made a down detector for down detectordowndetector.com	-0.15555555555556	negative
ergency call failure in Australia(theregister.com)	Multimodal Diffusion Language Models for ThinkingAware Editing and Generationgithub.comtyfeld	0	neutral
Email Support(thunderbird.net)	Outdated Samsung handset linked to fatal emergency call failure in Australiatheregister.com	-0.35833333333333	negative
utube.com)	Thunderbird Adds Native Microsoft Exchange Email Supportthunderbird.net	0	neutral
ld Teach You(toddhgardner.com)	I just want working RCS messagingwt.gd	0	neutral
om)	Google Antigravityantigravity.google	0	neutral
s)	The Future of Programming 2013 videoyoutube.com	0	neutral
3-Body problem simulator(trisolarchaos.com)	Learning to Boot from PXEimraniqbal.org	0	neutral
anity(joanwestenberg.com)	Build vs. Buy What This Weeks Outages Should Teach Youtoddhgardner.com	0	neutral
	Pebble, Pebble, and a path forwardericmig.com	0	neutral
	Ultima VII Revisitedgithub.comviridiangames	0	neutral
	Show HN Browserbased interactive 3D ThreeBody problem simulatortrisolarchaos.com	0	neutral
	How to Stay Sane in a World That Rewards Insanityjoanwestenberg.com	0	neutral

VISUALS

Power BI Charts

This section showcases the **dynamic bar and pie charts** generated in Power BI, illustrating key sentiment analysis metrics and insights from the data.



FINAL TABLE

Power BI Visuals – Table

This section illustrates the **finalized data table** in Power BI, showcasing key metrics and insights derived from the ETL process for effective analysis.

text_clean	Sum of sentiment_score	sentiment_label
ie Corporate Government Censorship(actu.org)	0.00	neutral
Your Smartphone, Their Rules App Stores Enable Corporate Government Censorshipactu.org	0.00	neutral
Ultima VII Revisitedgithub.comvldangames	0.00	neutral
mail Support(thunderbird.net)	0.00	neutral
Thunderbird Adds Native Microsoft Exchange Email Supportthunderbird.net	0.00	neutral
Projects(shkspr.mobi)	0.13	positive
ube.com)	0.00	neutral
The Future of Programming 2013 videoyoutube.com	0.00	neutral
uce a science fiction anthology(compellingsciencefiction.com)	0.00	neutral
The code and open source tools I used to produce a science fiction anthologycompellingsciencefiction.com	0.00	neutral
The Tk AWS Mistakegeocod.io	0.00	neutral
Body problem simulator(trisolarchaos.com)	0.00	neutral
Show HN Browserbased interactive 3D ThreeBody problem simulatortrisolarchaos.com	0.00	neutral
Questions for Cloudflareentropichthoughts.com	0.00	neutral
roc64.github.io)	0.00	neutral
Programming the Commodore 64 with .NETretroc64.github.io	0.00	neutral
Pimped Amiga 500pimyretro.org	0.00	neutral
n)	0.00	neutral
Pebble, Pebble, and a path forwardericmigj.com	0.00	neutral
ergency call failure in Australia(theregister.com)	0.36	negative
Outdated Samsung handset linked to fatal emergency call failure in Australiatheregister.com	0.00	neutral
king Aware Editing and Generation/github.com/tfloyd)	0.00	neutral
Multimodal Diffusion Language Models for Thinking Aware Editing and Generationgithub.comtfloyd	0.00	neutral
com/toddmaustin)	0.40	negative
	-1.19	

Conclusion

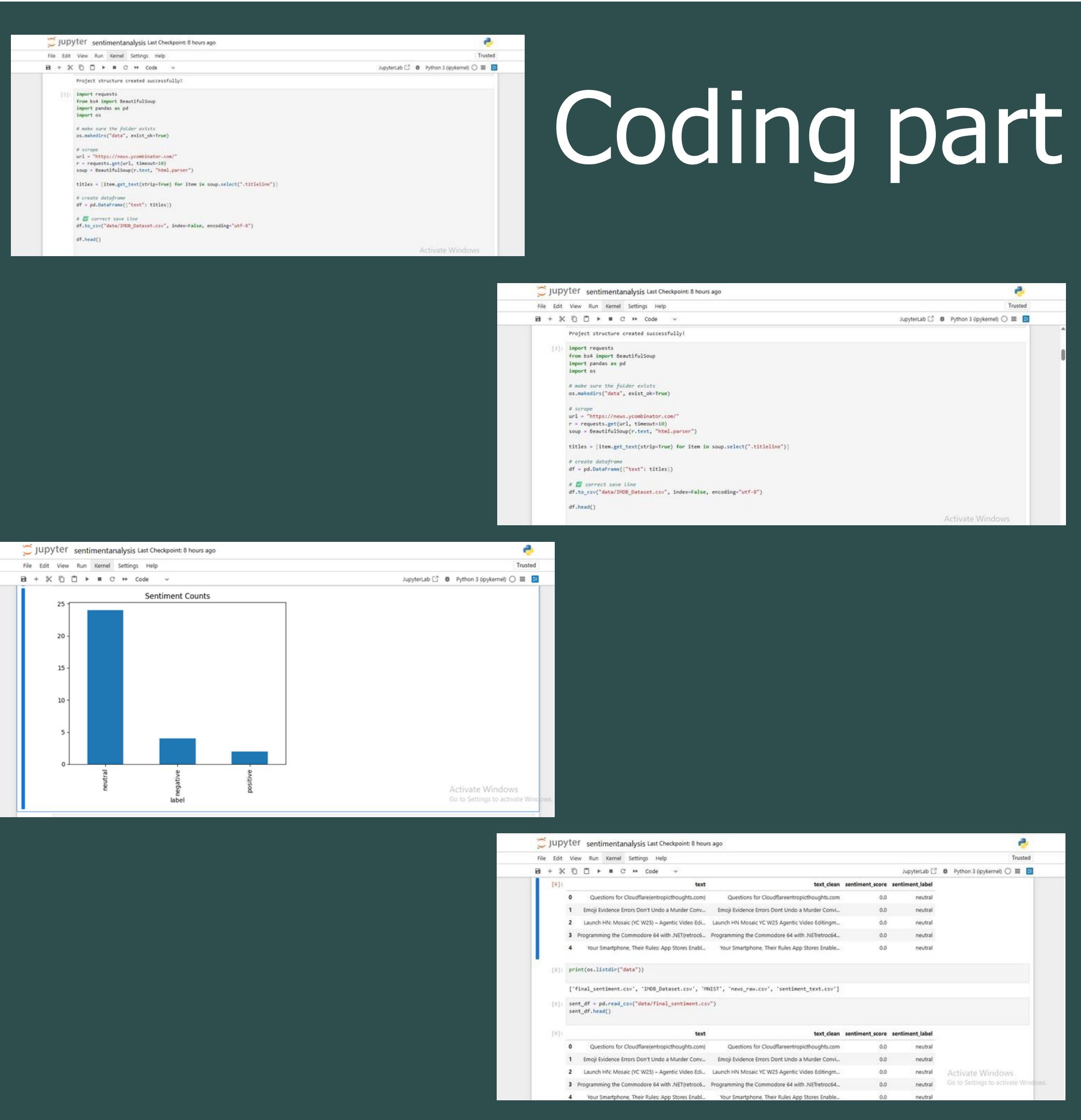
This end-to-end pipeline demonstrates data ingestion, processing, storage, and visualization using Azure services. The workflow is scalable and suitable for real-world sentiment analysis scenarios.

Reference:

- Loria, S. (2020). TextBlob Documentation.
- Microsoft Azure (2025). Blob Storage, Data Factory & SQL DB Docs.
- Hacker News (2025). News Headlines Source.
- McKinney, W. (2018). Python for Data Analysis.
- Microsoft (2025). Power BI Documentation.

Github Link:

https://github.com/chandresh2508/Sentiment_Analysis



Thank you!