Cloud Data Analysis: Sentiment in the Cyprus Problem

Students: Konstantinos Evagorou, Zariza Chowdhury

Course: Data Engineering 1

Assignment: 3

1. The Problem

Objective: Compare the sentiment of Greek Cypriot versus Turkish Cypriot media regarding the recent 2025 UN negotiation deadlock.

The political stalemate hinges on two conflicting frameworks:

- **Greek Cypriot side:** Insists on a Federal framework (Crans-Montana).
- **Turkish Cypriot side:** Demands a Two-State solution and recognition of sovereignty.

Research Question: Does the sentiment of news coverage differ significantly between the two sides?

2. Data Sources

We selected two representative articles from December 2025:

- **Greek Cypriot Source:** *Politis News*
  - **Topic:** President Christodoulides' refusal to discuss anything other than the UN federal framework.
  - **Narrative:** Focus on international law and UN resolutions.
- **Turkish Cypriot Source:** *Kibris Postasi*
  - **Topic:** Ersin Tatar's refusal to negotiate without recognition of sovereign equality.
  - **Narrative:** Rejection of the federation model; emphasis on isolation and independence.

3. Methodology

The solution uses a serverless pipeline with Python (`boto3`) and three AWS services:

1. **Ingestion:** Scraped article text using `requests` and `BeautifulSoup` .
2. **Storage:** Uploaded raw text files to **AWS S3** (`/raw`).
3. **Transformation:**
   - Detected language.
   - Translated the Turkish article to English using **AWS Translate**.
   - Stored translations in S3 (`/translated`).
4. **Analysis:** Processed English text with **AWS Comprehend** to detect dominant sentiment.
5. **Visualization:** Plotted confidence scores using `matplotlib` .

4. Results

The analysis highlights a clear divergence in tone:

Source	Sentiment	Confidence	Key Drivers
Greek Cypriot	**NEUTRAL**	98%	The text relies on legalistic language ("UN resolutions," "frameworks"), which the model interprets as objective rather than emotional.
Turkish Cypriot	**MIXED**	95%	The text contains conflicting emotional markers: positive words regarding "Independence/Sovereignty" mixed with negative framing of "Isolation/Embargoes."

(See code outputs for generated charts)

5. Cost Estimation

The project stayed well within the AWS Free Tier.

- **S3:** Negligible (< 1 MB storage, ~50 requests). **$0.00**
- **AWS Translate:** ~1,000 characters processed.
  - Calculation: (1,000 / 1M) * $15.00 = **0.000015**
- **AWS Comprehend:** ~15 units (including testing).
  - Calculation: 15 * $0.0001 = **0.0015**

Total Estimated Cost: < $0.01

Appendix: