



10Alytics

Driving Efficiency in Energy Sector: A Case Study of ETL Pipeline Implementation through API Integration Case Study For Nunu Energies

FACILITATOR: Dr. Ughele Ehibhahiemen Nelson





NUNU ENERGIES

Nunu Energies, a frontrunner in the energy sector, is dedicated to pioneering efficiency and innovation within its operations. Recognizing the pivotal role of data management in achieving these objectives, Nunu Energies has embarked on a transformative initiative to elevate its data infrastructure. This endeavor revolves around the implementation of an ETL (Extract, Transform, Load) pipeline, augmented by seamless API integration. This case study presents a comprehensive exploration of this endeavor, shedding light on the challenges encountered, the strategic solutions devised, and the substantial benefits reaped by Nunu Energies.

BUSINESS PROBLEM STATEMENT

In the contemporary energy landscape, Nunu Energies confronts the formidable challenge of wrangling and harnessing the vast troves of data generated across its diverse operational facets. Traditional data handling methodologies, reliant on manual extraction, transformation, and loading processes, have proven inherently inefficient and error-prone. This operational bottleneck not only impedes the agility of decision-making but also compromises the overall efficacy of Nunu Energies' endeavors.

Furthermore, the imperative to incorporate external data streams, particularly from sources like the Bloomberg API, further complicates the data integration landscape. Nunu Energies seeks a scalable and sustainable solution to automate its data workflows, ensuring unfettered access to timely insights crucial for informed decision-making and operational agility.

OBJECTIVES

- Develop and implement an ETL pipeline to automate data workflows across Nunu Energies' energy operations.
- Integrate data from Bloomberg API into the ETL pipeline to enrich internal datasets with real-time market insights.
- Ensure data quality and consistency through robust data transformation processes.
- Optimize data storage and retrieval mechanisms to facilitate fast and efficient access to insights.
- Enhance scalability and flexibility to accommodate future data growth and evolving business needs.

BENEFITS

- Increased operational efficiency through automation of data workflows, reducing manual effort and minimizing errors.
- Timely access to accurate insights derived from integrated Bloomberg API data, enabling proactive decision-making.
- Enhanced data quality and consistency, improving trust and reliability in analytical outputs.
- Scalable infrastructure design facilitates seamless adaptation to changing business requirements and data volumes.
- Improved resource utilization, as data engineers can focus on strategic initiatives rather than repetitive data management tasks.

The Dataset

- You can access the dataset here from the api url ⇒ [Link](#)

Database Schema Design:

Agriculture Table:

name
Units
Price
Change
%Change
Contract
Time (EST)

Energy Table:

name
Units
Price
Change
%Change
Contract
Time (EST)

Commodities Table:

name
Value
Change
%Change
High
Low
Time (EST)

Metals Table:

name
Value
Change
%Change
High
Low
Time (EST)

TECH STACK

For this project, the following tech stack will be utilized:

- Python: for scripting and automation of ETL processes.
- SQL: for data querying, manipulation, and transformation.
- Bloomberg API: for accessing real-time market data. [⇒ ⇒ LINK HERE](#)
- AWS S3: for scalable and durable storage of raw and processed data.
- AWS Redshift: for high-performance data warehousing and analytics.

SUBMISSION CRITERIA

The following documents should be submitted upon completion of the task

- Entity Relationship Diagrams (ERDs)
- Data Architecture
- Python codes (notebooks/scripts)
- Presentation slides
- Any Supporting documents

Business Questions

If you wanted to optimize your existing architecture?

- **What would you do different?**
- **and why?**

GOODLUCK