## Job Scheduling Documentation for Current Weather ETL Process

**1. Overview**

The job scheduling documentation outlines the setup for the current weather data ETL process, which retrieves weather data from the OpenWeather API hourly, transforms it, and loads it into a specified storage solution.

**2. Job Scheduler**

The process is scheduled using Apache Airflow, which allows for easy management and monitoring of ETL jobs.

**3. Schedule Frequency**

The current weather data extraction job is scheduled to run hourly, ensuring that the latest weather information is retrieved regularly.

**4. DAG Configuration**

The Directed Acyclic Graph (DAG) is configured with the following settings:

* **Start Date**: September 30, 2024
* **Default Arguments**:
  + Owner: 3Signet
  + Depends on Past: False
  + Email Notifications: Alerts will be sent to olawumisalaam@gmail.com
  + Retries: 2
  + Retry Delay: 2 minutes

**5. Task Dependencies**

The job consists of three main tasks:

1. **HttpSensor**: Checks if the weather API is ready to accept requests.
2. **SimpleHttpOperator**: Extracts the weather data from the OpenWeather API.
3. **PythonOperator**: Transforms the extracted data and loads it into a CSV file.

The sequence of task execution is as follows:

* The is\_weather\_api\_ready task is executed first to ensure the API is available.
* Upon successful completion, the extract\_weather\_data task retrieves the current weather data.
* Finally, the transform\_load\_weather\_data task processes and saves the data.

**6. Monitoring and Alerts**

Airflow provides a user interface for monitoring job execution. Users can view the status of each task and its logs. In case of failures, email alerts are sent to the designated email address, allowing for timely intervention.

**7. Conclusion**

This documentation serves as a guide for understanding the job scheduling for the current weather ETL process.