**User Guide for Configuring and Running the System**

This guide will help you configure and run Docker containers, monitor tasks in Apache Airflow, and access generated reports stored in Google Drive.

## **Set Up and Run the Docker Containers**

**Prerequisites:**

* **Docker**: If you don't have Docker installed, download and install it from the official Docker website.
* **Docker Compose**: If you don't have Docker Compose installed, follow the installation instructions on the Docker Compose website.

**Steps:**

1. **Clone the repository** or navigate to the project directory where the following files are located:
   * docker-compose.yml
   * python-dockerfile (for the Python service)
   * init.sql (MySQL script)
   * packages.txt (list of Python dependencies)
2. **Build the Docker containers:**
   * Open a terminal or command prompt.
   * Navigate to the directory containing the “docker-compose.yml” file.
   * Run the following command to build the Docker images and start the containers: “**docker-compose build**”
3. **Start the Docker containers:**
   * Run: “**docker-compose up**”
   * This command will:
     1. Start **MySQL container** - Initializes the database and loads the sample data.
     2. Start **Python container** - Processes the data, generates reports, and sends notifications.
4. **Access the running containers:**
   * To view logs of a running container, use: “**docker logs <container\_name**>”

Example:

* docker logs mysql-container
* docker logs python-container

1. **Stop the containers:**
   * To stop the containers, run: “**docket-compose down**”

## **Monitor Tasks Using Airflow Web Server**

The Python container is integrated with Apache Airflow to schedule and monitor tasks.

**Prerequisites:**

* Ensure Apache Airflow is configured and running.
* Airflow scheduler and web server must be running.
* Python DAGs (Directed Acyclic Graphs) should be defined and loaded into Airflow.

**Steps:**

1. **Access Airflow Web Interface:**
   * Open a web browser and navigate to the Airflow web server. By default, Airflow’s web server runs on port 8080.
   * If you’re running Airflow inside Docker, you may need to adjust the hostname or port according to the container configuration.
2. **Log in to the Airflow web interface:**
   * Use default username (**admin**) and password (**admin**) unless you’ve configured custom credentials.
3. **Monitor Running Tasks:**
   * Once logged in, you will see your DAGs listed.
   * Click on a specific DAG to check its status (e.g., running, succeeded, failed).
   * You can click on an individual task to view detailed logs for troubleshooting.
4. **Trigger Tasks:**
   * You can manually trigger a task or DAG run from the Airflow interface by clicking on the “**Trigger Dag”** button.
   * Monitor task progress and logs as tasks execute.

## **View and Download Reports from Google Drive**

Once the Python module generates the reports, it will upload them to **Google Drive**.

**Prerequisites:**

* Ensure that the **Google Drive API** is set up and the Python script is configured to upload files to a specific folder in Google Drive.
* You should have the **Google credentials file** (credentials.json) used by the Python script for authentication.

**Steps:**

1. **Locate the Folder on Google Drive:**
   * Open your web browser and navigate to Google Drive.

* Go to the folder where the reports are uploaded. The folder ID is specified in the Python code, and you can find it in the URL after accessing the folder, such as “**https://drive.google.com/drive/folders/<folder\_id>**”

1. **Access the Report:**
   * Inside the folder, you will find the reports listed, such as “**sales\_report\_YYYYMMDD\_HHMMSS.xlsx**”.
2. **Download the Report:**
   * Right-click on the file in Google Drive and select **Download**.