databricks Academy

DAWD 03-4 - Demo - Notifying Stakeholders

The two fields below are used to customize queries used in this course. Enter your schema (database) name and username, and press "Enter" to populate necessary information in the queries on this page.

Schema Name:	hive_metastore.class_013_odg7_da_dawd
Username:	class+013@databricks.com

Lesson Objectives

At the end of this lesson, you will be able to:

- Describe how to configure alerts
- Describe how to share queries and dashboards with stakeholders
- Describe how to refresh dashboards

Query Refresh Schedule

You can use scheduled query executions to keep your dashboards updated or to enable routine alerts. Let's make a query and put it on a refresh schedule.

1. Run the query below.
2. Name the query by clicking "New Query" and typing “Gym Logs”

3. Click "Save"

   The query needs to be saved with a descriptive name, so we can reference it later in this lesson. To refresh this query automatically:

4. Click "Schedule" (upper-right corner of query window)
5. Click the dropdown that says "Never" and note that the query can be scheduled for many different times
6. Click "Cancel"

   WARNING: If the refresh rate is less than the SQL Warehouse 'Auto Stop' parameter, the Warehouse will run indefinitely.

```
USE hive_metastore.class_013_odg7_da_dawd;
SELECT gym, count(*) number_of_visits


FROM gym_logs
GROUP BY gym
ORDER BY gym;
```

Copy

Alerts

Alerts allow you to configure notifications when a field returned by a scheduled query meets a specific threshold. Although we just configured a refresh schedule for our query, the Alert runs on its own schedule.

To create an Alert:

7. Click "Alerts" in the sidebar menu
8. Click "Create Alert"
9. From the Query dropdown, select our query: "Gym Logs"
10. Use the dropdown to change the "Value" column to `number_of_visits` and change "Threshold" to 1
11. Change "Refresh" to Every 1 minute

    The default destination is the user's email address. The alert is triggered when the count of the top row in the query's results is greater than 1.

Let's add some data to trigger the alert.

12. Run the code below.

This code will ingest the remaining gym log data from the object store. This will increase the number of gym visits past our threshold and trigger the alert.

Something to note with regard to configuring Alerts and Refresh Schedules: Every time they run, the SQL Warehouse will start (if it's stopped), run the query, and go into an idle state. Once the Auto Stop time has expired, the SQL Warehouse will stop. If the refresh schedule is set to a lower time limit than the SQL Warehouse's Auto Stop time, the Warehouse will never stop. This may increase costs.

13. Delete the alert and change the refresh schedule back to "Never".

```
USE hive_metastore.class_013_odg7_da_dawd;
COPY INTO gym_logs
    FROM 'wasb://courseware@dbacademy.blob.core.windows.net/data-analysis-with-
databricks/v02/gym_logs/gym_logs_json'
    FILEFORMAT = JSON;
```

Copy

Sharing Queries

We can share queries with other members of the team:

13. Back in the Query Editor, click "Share"

The "Manage Permissions" dialogue appears. If you do not have permission to change settings, all options will be greyed out. Note that, as the owner of the query, you have "Can manage" permissions. You can share the query with users and groups who are configured in your workspace. These users and groups can have either "Can run" or "Can edit" permissions. Those with "Can edit" permissions can also run the query. In order to allow "Can edit" permissions, the Credentials drop down must be changed to "Run as Viewer". Click inside the input box, and a dropdown will show all users and groups with whom the query can be shared.

14. Select a user or group

15. Select either "Can run" or "Can edit" permissions
16. Close the dialogue

Note that any "Can edit" permissions that were granted must be revoked before the credential type for the query can be changed back to "Run as owner".

Sharing Dashboards

Sharing dashboards is exactly the same as sharing queries. Click "Share" from any dashboard to update sharing permissions.

Refreshing Dashboards and Sharing Results

We can set a refresh schedule for a dashboard and, optionally, share the results with others.

17. From any dashboard, click "Schedule"
18. Drop down "Refresh" and select a refresh interval
19. Optionally, select a SQL Warehouse to use to refresh the dashboard
20. Set any Subscribers to be notified of dashboard results
21. Ensure that "Enabled" is set to on
22. Click "Save"

When you are finished with the dashboard refresh schedule, go ahead and disable it.

WARNING: If the Dashboard refresh interval is less than the SQL Warehouse 'Auto Stop' parameter, the Warehouse will run indefinitely.

1.2.13