Lab 6 – Monitor factory activity

ADF provides a number of options for monitoring pipelines. The ADF UX includes a visual monitoring experience which you can use to inspect pipeline executions. You may also choose to route logs into another Azure service for longer-term storage or more sophisticated analysis. This lab uses the ADF UX monitoring experience and explores the benefits of routing log data to Azure Log Analytics.

Lab 6.1 – Create a Log Analytics workspace

To access the Azure Log Analytics service, create a Log Analytics workspace.

1. In the Azure portal, click “Create a resource” and search for “Log Analytics Workspace”. Click “Create” on the overview screen.

2. Complete the **Basics** tab like this:

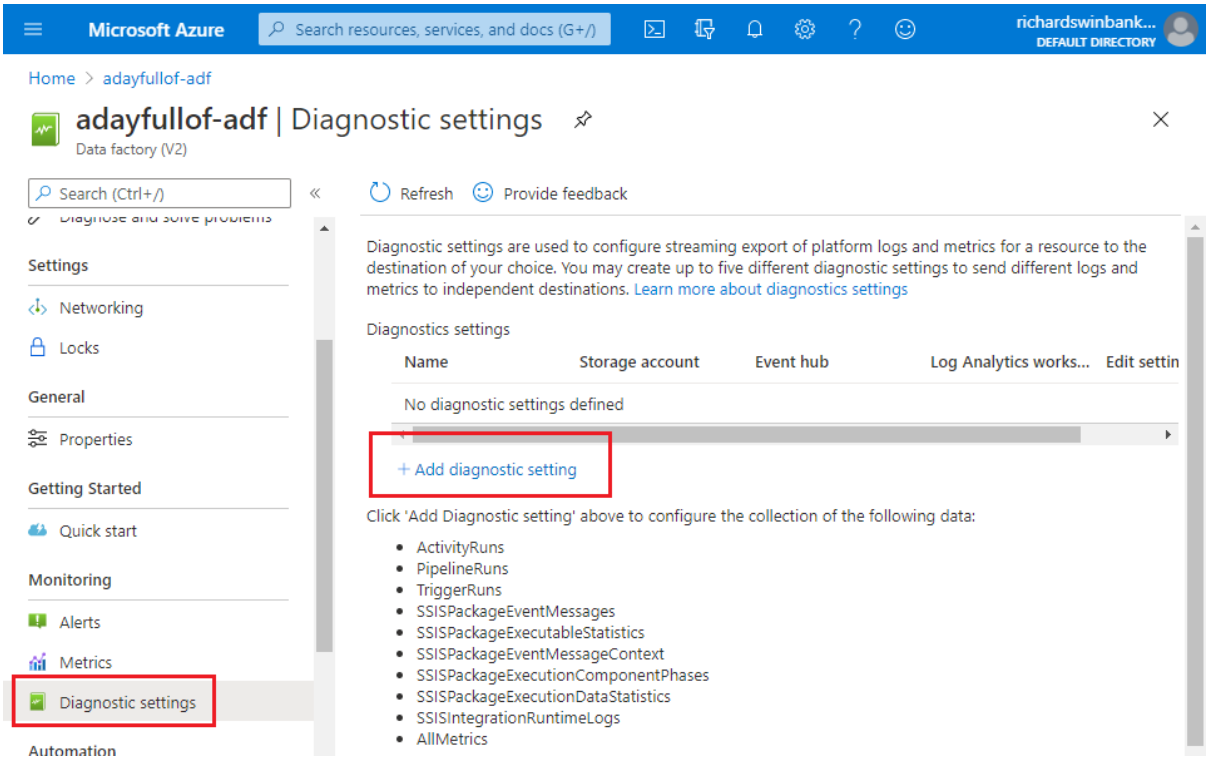
   - Choose the subscription and resource group you’ve been using throughout this lab
   - Enter a name for your workspace
   - Choose the same region you specified for your resource group, then click “OK”

   Click “Review + create”, then “Create”.

Lab 6.2 – Configure diagnostic settings

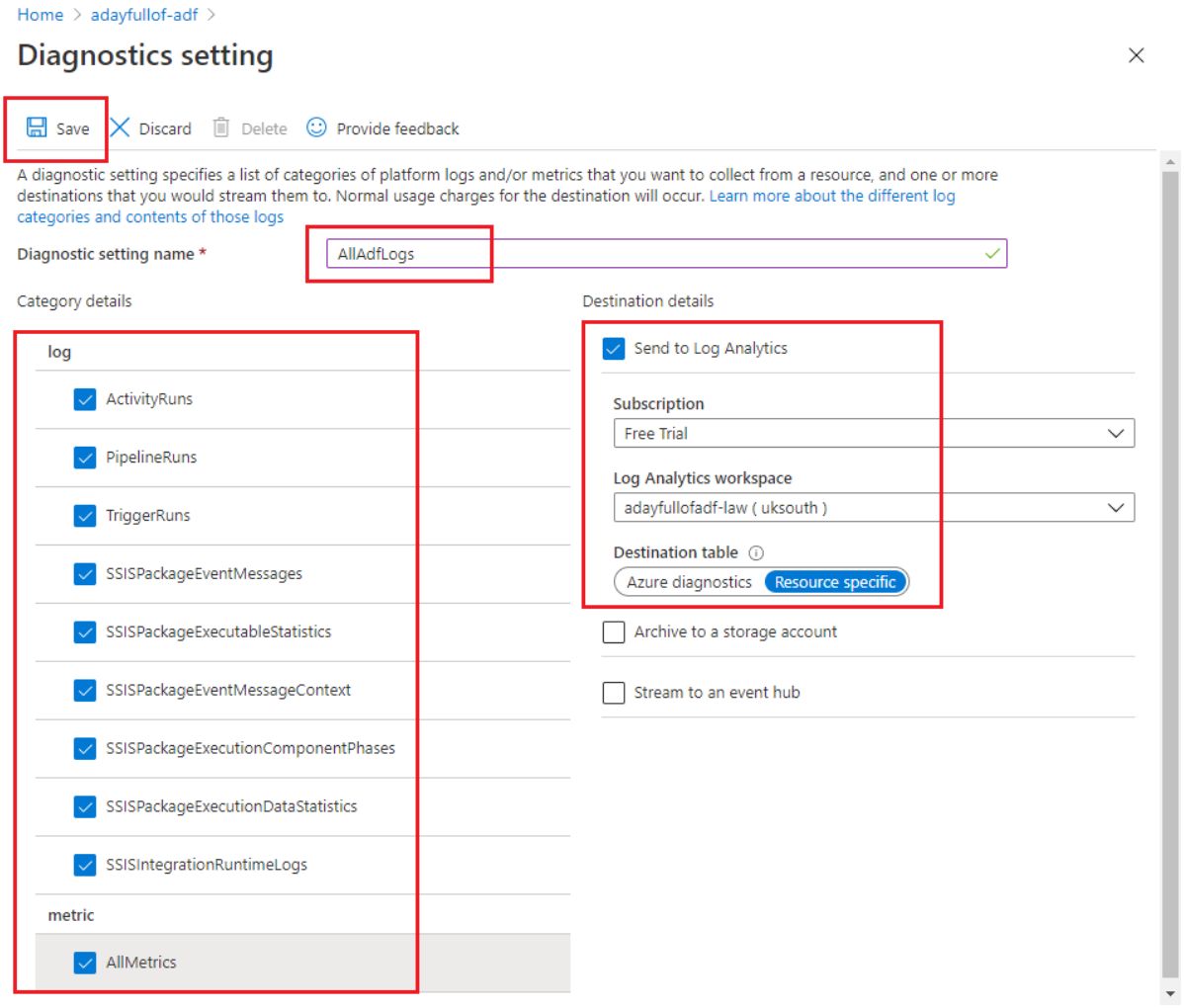
Routing log data to Log Analytics or other services is configured in your factory’s Azure portal blade.

1. On the factory’s blade in the Azure portal, select “Diagnostics settings” from the sidebar, then click “+ Add diagnostic setting”.



2. On the “Diagnostics setting” blade, give the new diagnostic setting a name. Select all categories of log data and ensure that “AllMetrics” is ticked.

3. Tick “Send to Log Analytics”. Options appear to choose the Log Analytics workspace – select the one you created in Lab 6.1. Leave “Destination table” set to “Resource specific”, then save your changes using the “Save” button in the top left.



Lab 6.3 – Generate log data

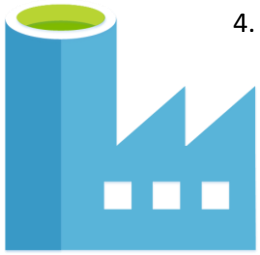
In previous labs you have mostly run pipelines using the “Debug” option in your ADF UX session. To generate log data, run some pipelines in the published environment. We haven’t covered triggers – used to start pipelines automatically – in these labs, but you can run pipelines manually in the ADF UX.

1. If your Azure Data Factory instance is Git-enabled, publish all your pipelines as you did in Lab 2.4. (If your factory is not Git-enabled, you can skip this step – you have been publishing your pipelines repeatedly to save your work).

2. Open a pipeline of your choice in the ADF UX.

3. Click “Add trigger” above the pipeline canvas and select “Trigger now”. A confirmation blade is displayed – click OK to trigger the published pipeline.

4. Repeat steps 2 and 3 a few times for different pipelines.

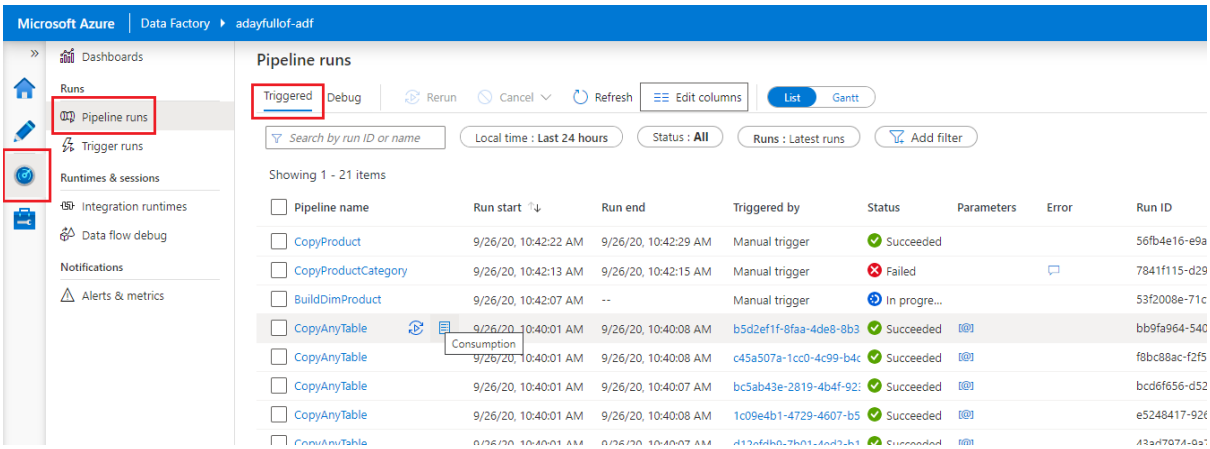


The purpose of running these pipelines is just to accumulate log data – in a production data factory instance log data is generated organically from routine pipeline runs.

Lab 6.4 – Use the ADF UX monitoring experience

The ADF UX monitoring experience allows you to inspect pipeline run history visually.

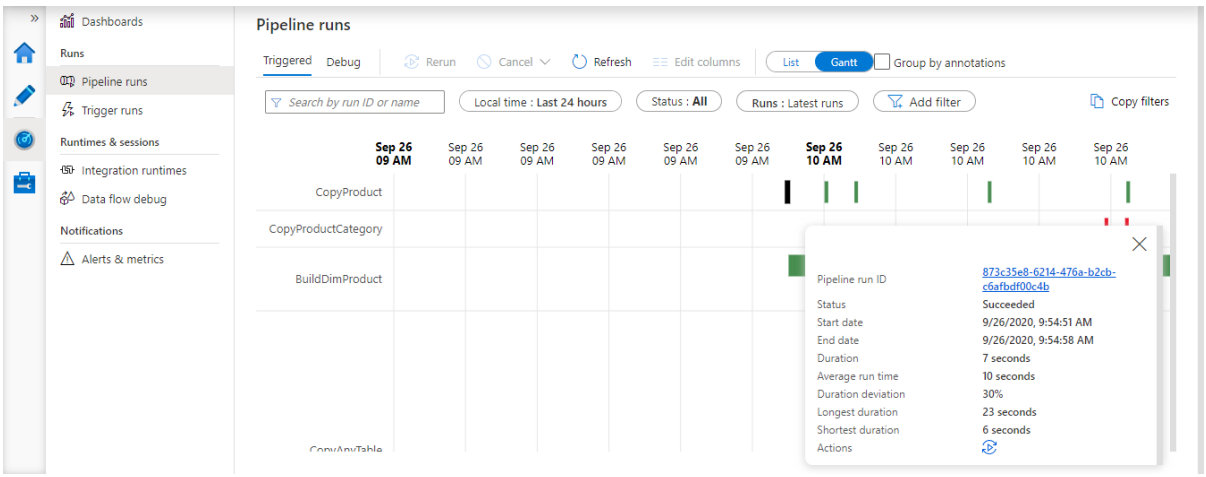
1. Open the monitoring experience in the ADF UX by clicking the Monitor button. Select "Pipeline runs" from the menu then choose the "Triggered" tab to view published pipeline runs.



2. Explore information reported in the pipeline run list. It includes:
   - Pipeline name, run ID and start & end times.
   - Pipeline run status. For failed runs, the speech bubble icon in the "Error" column provides failure information.
   - How the pipeline was triggered – in the screenshot, you can see pipelines triggered directly from the ADF UX, and the CopyAnyTable pipeline (from Lab 5.3) triggered by an Execute Pipeline activity run in the CopyAdventureWorks pipeline. Parameter values passed to each CopyAnyTable run can be inspected using the [@] link.

3. Hover over a pipeline name's name in the list – "Rerun" and "Consumption" buttons appear. Click on "Consumption" to view resources used by the pipeline's execution. The pipeline's name is a link – click on the link to view activity run information, presented in a very similar way to the debug output you see in the authoring canvas.

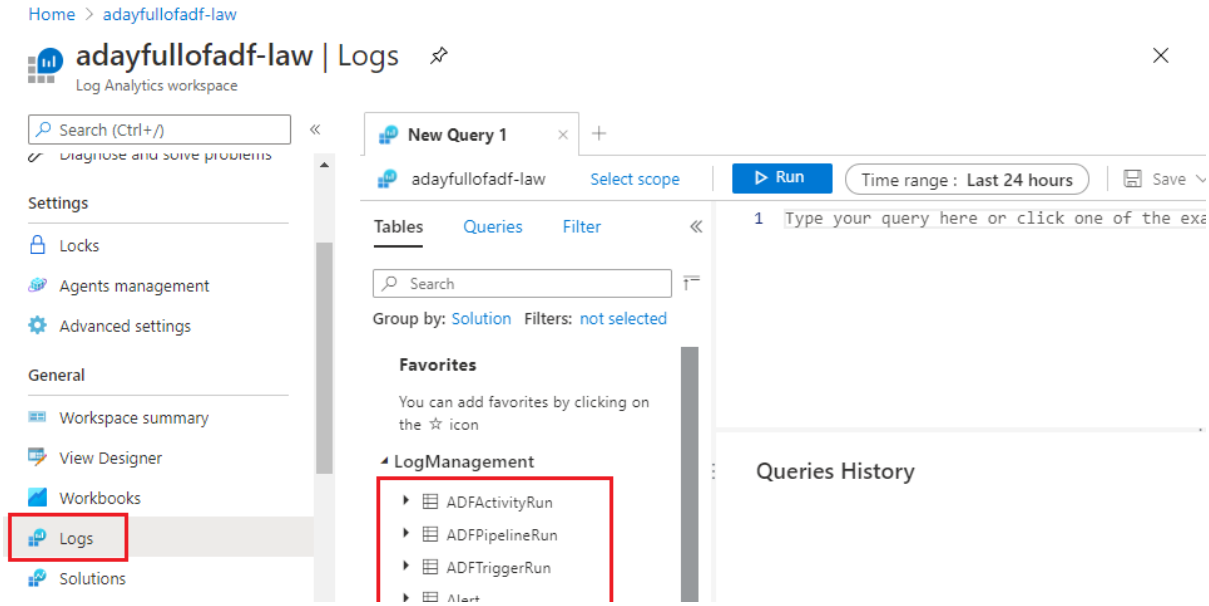
4. In the top right, move the slider from "List" to "Gantt". The Gantt chart view provides much of the same information, displaying pipeline run duration against time. Click on a bar in the Gantt chart to view detail of the pipeline's execution.



Lab 6.5 – Query Log Analytics data

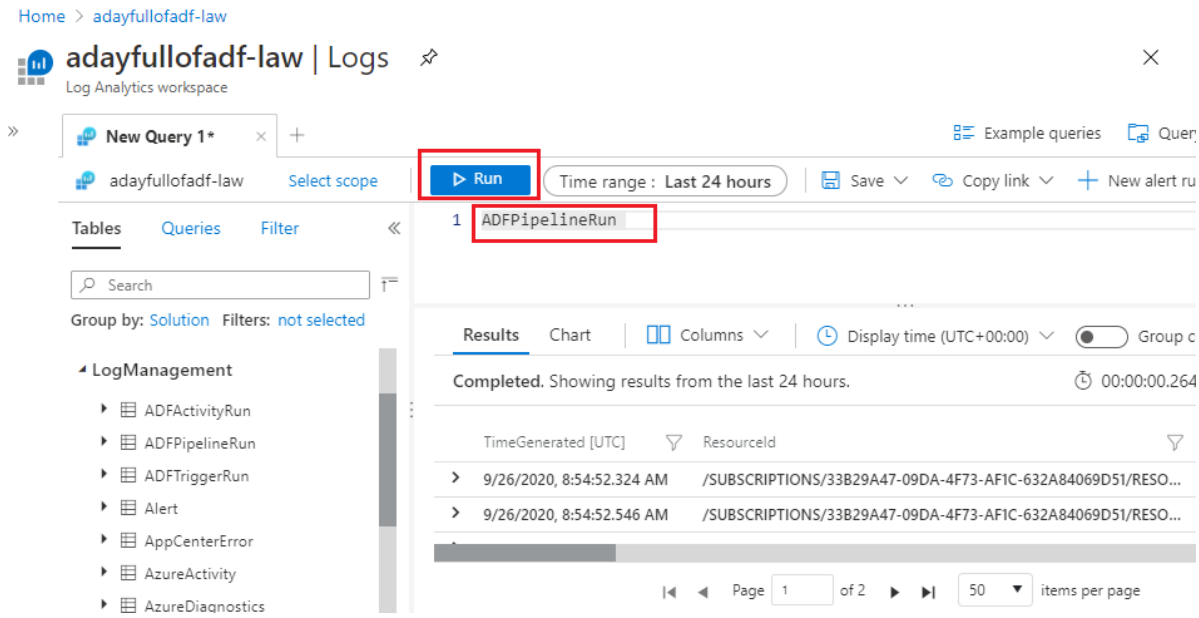
Pipeline execution data from Lab 6.3 has also been pushed to Log Analytics because of the diagnostic setting you created in Lab 6.2.

1. Open your Log Analytics workspace in the Azure portal and select the “Logs” item from the sidebar. Close any example query dialog that appears. A query interface is displayed. Collapse the portal sidebar if you need more space.



2. The query interface includes a list of tables and a query pane. Double-click on the “ADFPipelineRun” table name to add it to the query pane.

3. Log Analytics queries are written in Kusto. The table name on its own is a valid Kusto query – it means “select everything from table AdfPipelineRun”. Click “Run” to run the query.



4. This Kusto query identifies the ten most recent failed pipeline runs:

```
ADFPipelineRun
  | where Status == 'Failed'
  | project PipelineName, End
  | order by End desc
  | limit 1
```

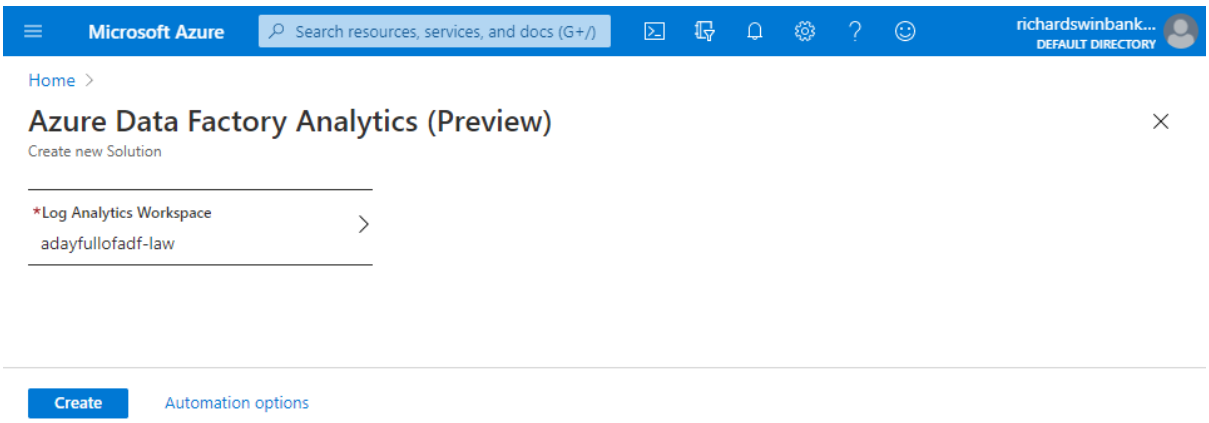
   Run this or other queries to explore the logged data. The query pane provides intellisense options to help you get started quickly.

Lab 6.6 – Use Log Analytics workbooks

Using Log Analytics for ADF reporting has a number of advantages, for example log retention beyond ADF's 45-day maximum, the ability to integrate ADF log data with logs from other Azure services or access to enhanced query support using Kusto. Another advantage is Log Analytics' in-built visualisation support using workbooks. A quick way to get started is with the "Azure Data Factory Analytics" solution available from the Azure Marketplace.

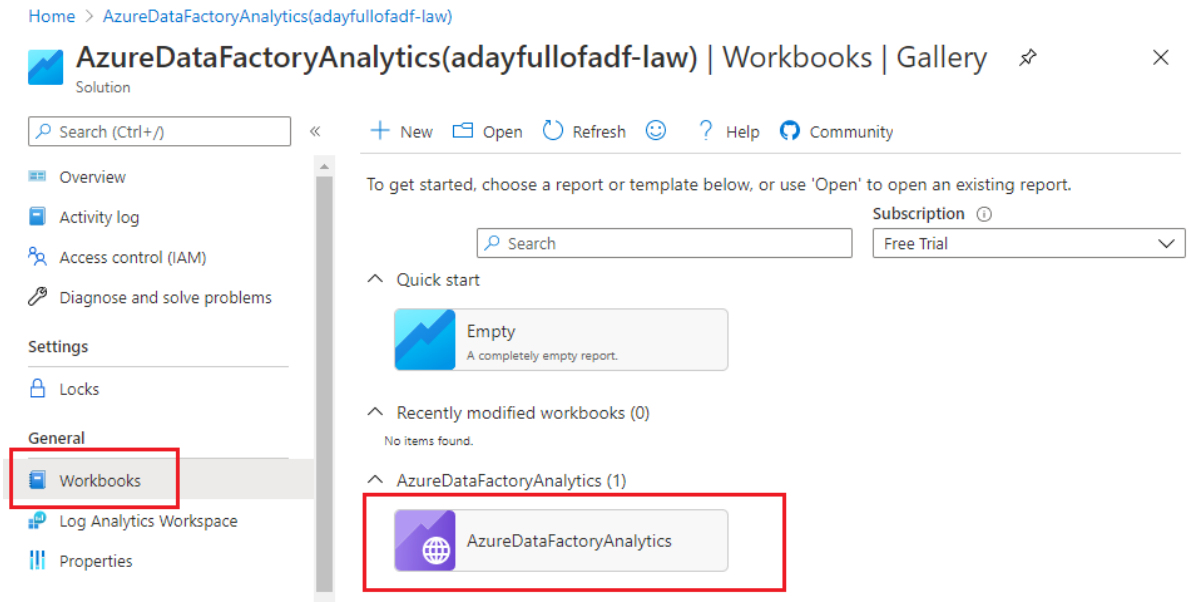
1. Browse to the Azure Data Factory Management Solution Service Pack in the Azure Marketplace. Click "GET IT NOW", then click "Continue" to accept T&Cs.

2. A "Create new solution" page opens in the Azure portal. If your Log Analytics workspace is not pre-selected, click "Log Analytics Workspace" and choose it from the list. Click "Create".

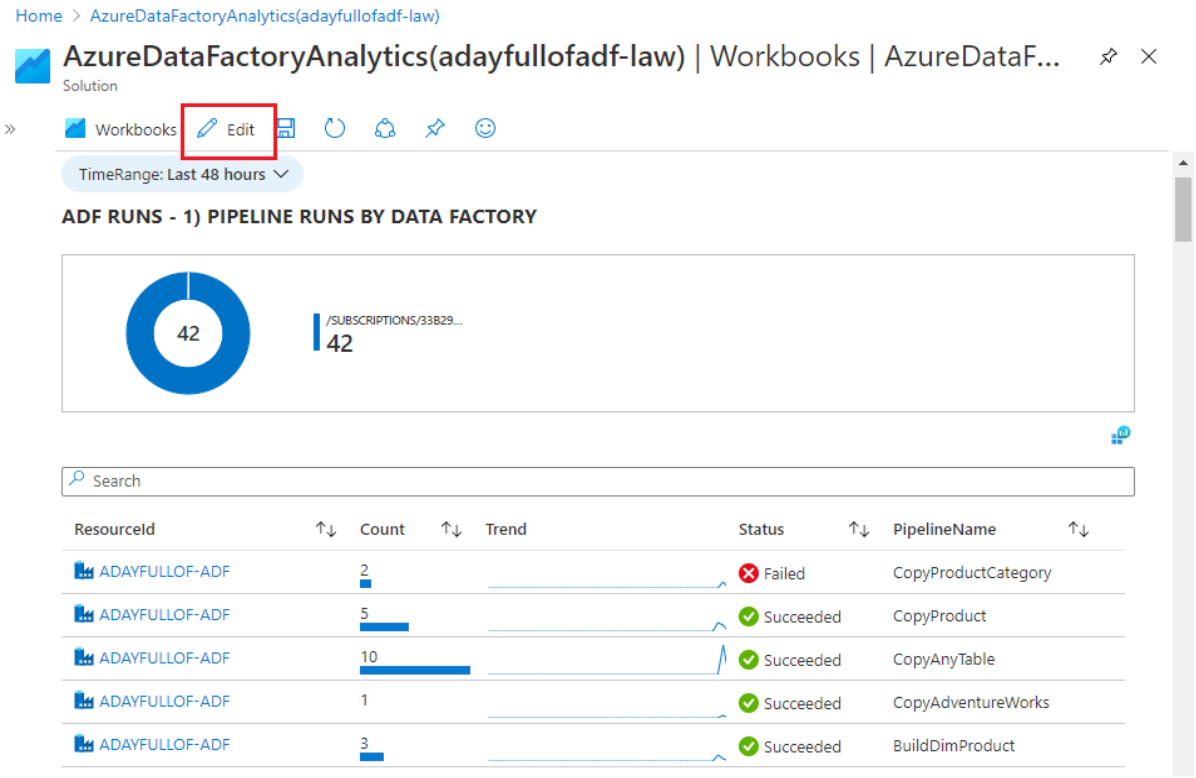


3. When confirmation of deployment is displayed, click "Go to resource". If you don't see the notification, you can find the solution using the "Solutions" item in the Log Analytics workspace blade.

4. On the "AzureDataFactoryAnalytics" blade, select "Workbooks" from the sidebar then click the "AzureDataFactoryAnalytics" tile.



5. Once installed, you can edit the Workbook as you wish. Here I have reduced the default size of the pie chart and filtered out "Queued" entries from the list of statuses reported from Log Analytics.



Recap

In Lab 6 you:

- created a Log Analytics workspace
- configured the workspace to receive log data from ADF
- generated log data for a series of published pipelines
- interacted with pipeline logs using the ADF UX, Log Analytics queries in Kusto and Log Analytics workbooks.