

Data 226 – Data Warehouse and Pipeline

Prof. Keeyong Han

Assignment 6

Date: 28th October 2025

Question 1:

ETL PIPELINE

The screenshot shows the Airflow web interface for the ETL pipeline. The top navigation bar includes links for Airflow, DAGs, Cluster Activity, Datasets, Security, Browse, Admin, and Docs. The time is 23:03 UTC. The main page displays the DAG 'BuildELT_CAS' with a run ID of '2025-10-28T23:15:00Z-05:00'. A summary table provides details about the DAG runs, showing 2 total runs, 2 successful runs, and a mean run duration of 00:00:40.

Total Runs Deployed	2
Total Success	2
First Run Start	2025-10-28, 23:00:29 UTC
Last Run Start	2025-10-28, 23:00:29 UTC
Max Run Duration	00:00:40
Mean Run Duration	00:00:40
Min Run Duration	00:00:40

Question 2:

ELT PIPELINE:

The screenshot shows the Airflow web interface for the ELT pipeline. The top navigation bar includes links for Airflow, DAGs, Cluster Activity, Datasets, Security, Browse, Admin, and Docs. The time is 23:04 UTC. The main page displays the DAG 'BuildELT_CAS' with a run ID of '2025-10-28T23:03:00Z-05:00'. A summary table provides details about the DAG runs, showing 3 total runs, 1 successful run, and a mean run duration of 00:00:35.

Total Runs Deployed	3
Total Success	1
Total Failed	2
First Run Start	2025-10-28, 23:00:14 UTC
Last Run Start	2025-10-28, 23:03:17 UTC
Max Run Duration	00:00:45
Mean Run Duration	00:00:42
Min Run Duration	00:00:35

QUESTION 3:

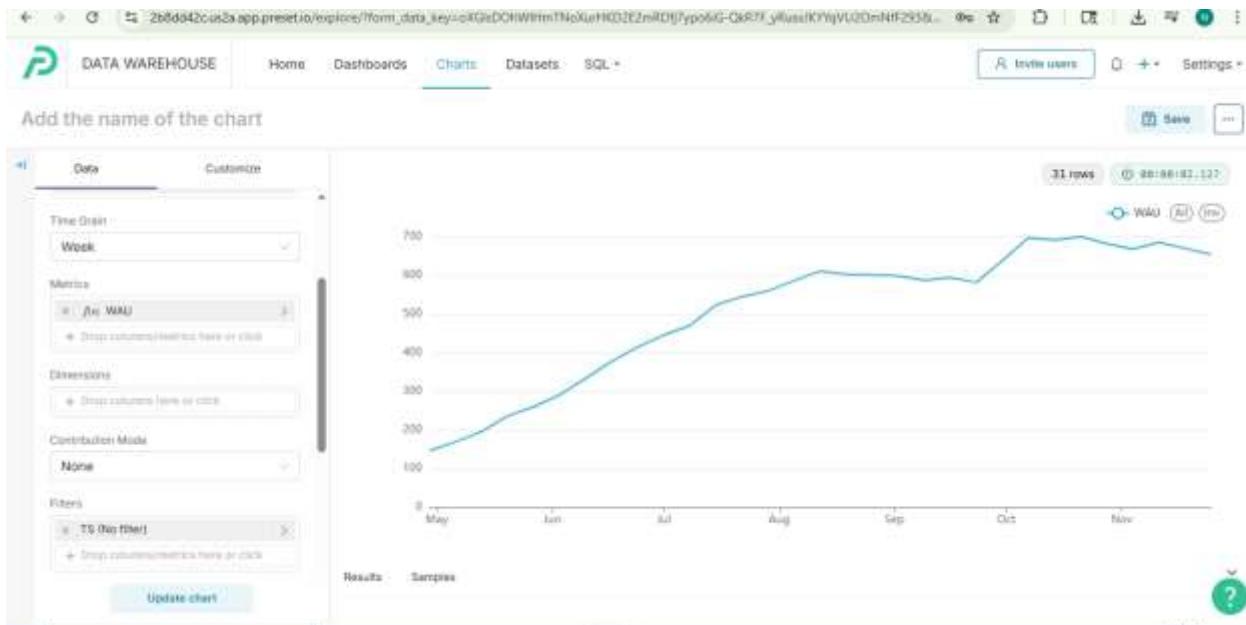
DATASET SCREENSHOT

The screenshot shows a web-based data management interface. At the top, there are tabs for Home, Dashboards, Charts, Datasets (which is selected), and SQL. Below the tabs, there are search and filter fields for Name, Type, Database, Schema, and Owner. A "Build select" button and a "Dataset" button are also present. The main area displays a table of datasets with columns: Name, Type, Database, Schema, Owners, Last modified, and Actions. The datasets listed are:

Name	Type	Database	Schema	Owners	Last modified	Actions
session_summary	Physical	Snowflake	analytics	NC	2 minutes ago	
COVID Vaccines	Physical	examples	public		3 hours ago	
FCC 2018 Survey	Physical	examples	public		3 hours ago	
Flights	Physical	examples	public		3 hours ago	
Sample Geodata	Physical	examples	public		3 hours ago	
San Francisco Population Polygons	Physical	examples	public		3 hours ago	
San Francisco BART Lines	Physical	examples	public		3 hours ago	
Slack Channel Combinations and Users	Virtual	examples	public		3 hours ago	

QUESTION 4:

WAU CHART



GITHUB LINK: <https://github.com/NAMAN-CHHEDA/DATA-226-ASSIGNEMLT-6>