Delta Live Tables

ScholarNest

DLT – Delta Live Tables

Databricks Cloud

- Introduction to DLT
- DLT End to End Scenario
- DLT Scenario Data Setup
- Create DLT Workload using SQL
- Create and run DLT Pipeline
- Create DLT Workload using Python

ScholarNest

Databricks Delta Live Table

What is Databricks Delta Live Table (DLT)

DLT is a framework for building data processing pipelines

We use this framework in two steps.

1. Write Declarative pipeline code
2. Create a DLT pipeline using Databricks Workflow UI

Declarative ETL Tool

ETL Pipeline – How

Write code for how to do

1. How to extract data from a given source system.
2. How to transform or validate data after extracting.
3. How to load valid and invalid data after transforming.

Stitch tasks together using orchestration tools

Declarative ETL – What

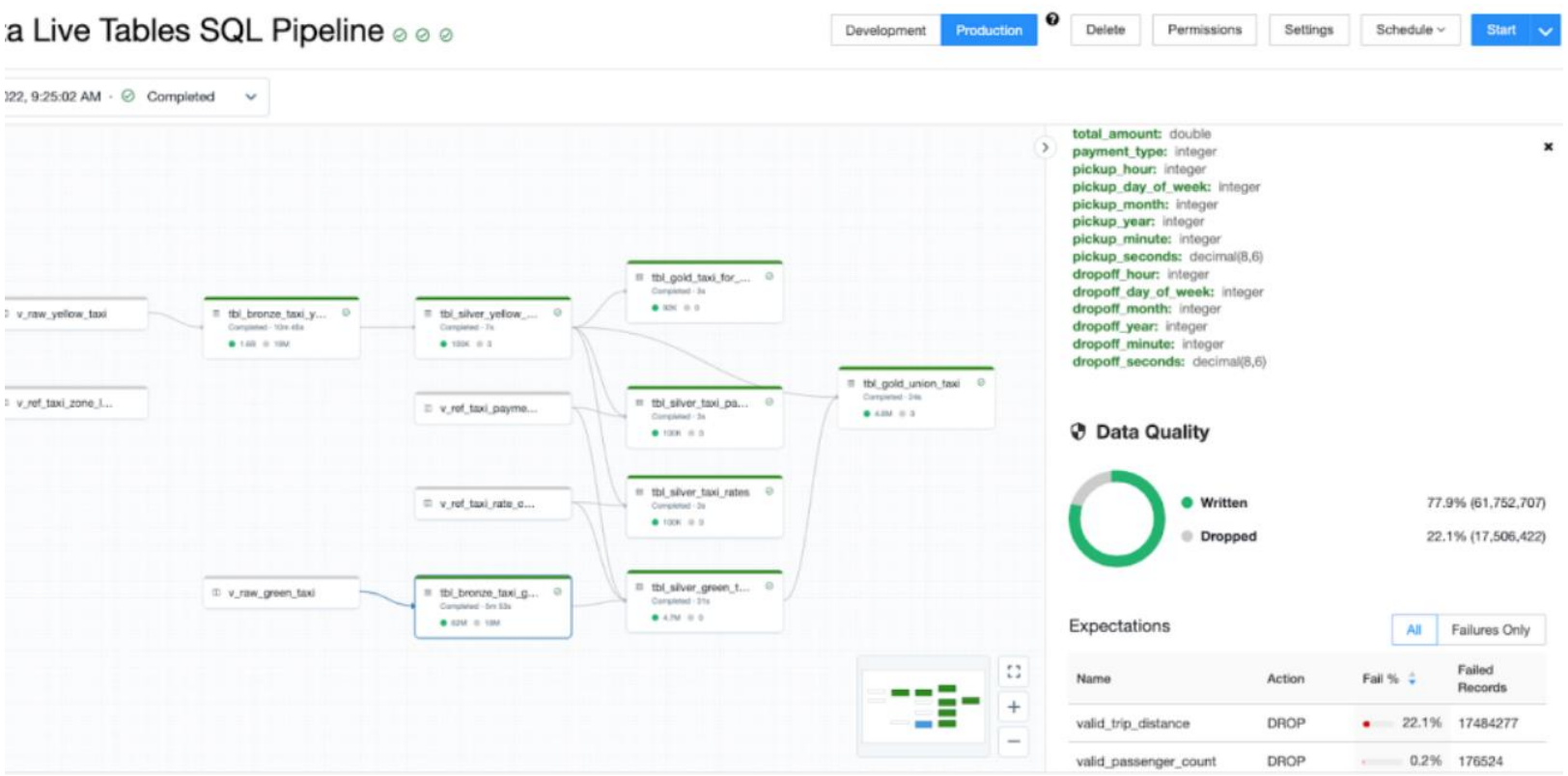
Define what to do

1. What to extract from the given source
2. What are your data quality expectations
3. What actions to take when expectations are met
4. What actions to take when expectations are not met

Integrated orchestration

Define what to do

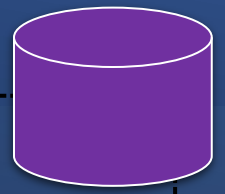
1. What to extract from the given source
2. What are your data quality expectations
3. What actions to take when expectations are met
4. What actions to take when expectations are not met

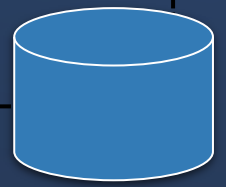


Declarative ETL Tool

ScholarNest

Databricks Delta Live Table





Lakehouse

Data Engineering Workspace

Spark

Spark

Spark

SQL WS

Unity Catalog - Data Governance

Bronze

Silver

Gold

/

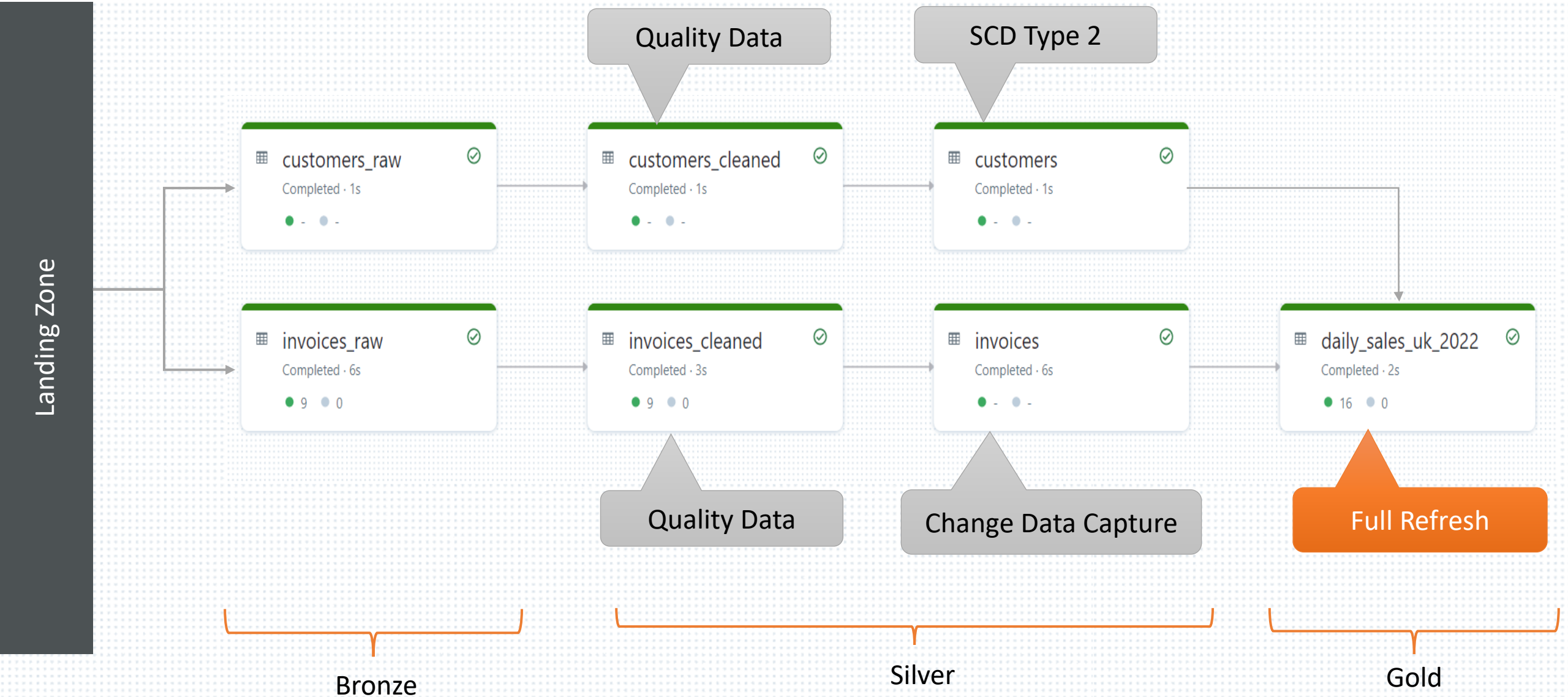
Raw

Processed

Analysis

Azure Storage Account

Azure Cloud Platform


Landing Zone
customers_raw
Completed · 1s
customers_cleaned
Completed · 1s
customers
Completed · 1s
invoices_raw
Completed · 6s
9 0
invoices_cleaned
Completed · 3s
9 0
invoices
Completed · 6s
daily_sales_uk_2022
Completed · 2s
16 0
Quality Data
SCD Type 2
Quality Data
Change Data Capture
Full Refresh
Bronze
Silver
Gold