Silver Screen Analysis

Cinema Profitability

Cornelia Brinkmann - Masterschool - 2026-01-09

Project Overview

Modernizing the “ Silver Screen “ Data Pipeline

- **Current State:** Transitioned from siloed raw data (CSVs) to a centralized, automated analytics environment.
- **Core Technology:** Powered by Snowflake (Cloud Data Warehouse) and dbt Cloud (Transformation Layer).
- **Modular Architecture:** Established a best-practice 3-tier structure:
  - **Staging:** Cleansing and renaming raw source data
  - **Intermediate:** Applying complex business logic and joining disparate sources.
  - **Marts:** Delivering high-performance, business-ready tables (e.g., Cinema Profitability).

Project Objectives

Strategic Goals & Technical Milestones

- **Automated Data Integrity:** Implementation of 30+ automated tests (Unique, Not-Null, and dbt-utils) to ensure "one version of the truth."
- **Operational Efficiency:** Fully automated daily production runs, eliminating manual SQL execution and spreadsheet errors.
- **Full Data Lineage:** Created a transparent, end-to-end map of data flow, ensuring that every KPI can be traced back to its raw source.
- **Robust Modeling:** Used advanced techniques like Surrogate Keys (hashing) to handle complex relationships between movies, dates, and locations without duplicates.

dbt Cloud Structure & Model Execution

Architecture & Execution Workflow

- **Development Environment:** Utilizing dbt Cloud's Integrated Development Environment (IDE) for version-controlled SQL development.

- **Modular Foldering:**
  - models/staging: Direct mapping of Snowflake sources with light cleaning.

  - models/intermediate: The "Engine Room" where business logic is applied.

  - models/marts: Final reporting layer optimized for BI tools.

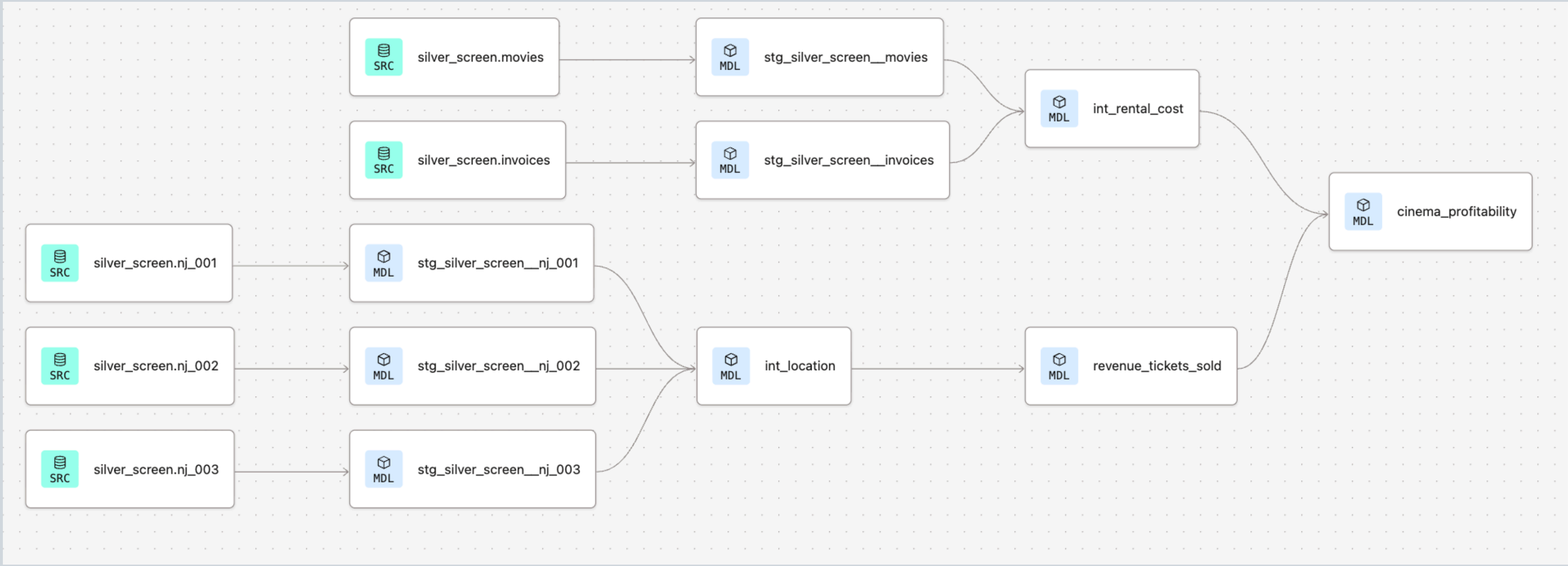
- **Execution Model:** Models are built as Views during development for speed and materialized as Tables in Production (intermediate and marts layer) for maximum query performance in Snowflake.

Lineage & Key Transformations

Visualizing the Data Journey

- **End-to-End Traceability:** The Lineage Diagram provides a live map from 5+ Raw Sources to the final Cinema Profitability mart.
- **Key Transformation Steps:** Standardization: Converting inconsistent date formats and currency fields into a unified schema.
- **Surrogate Key Generation:** Leveraging dbt_utils to create unique identifiers (MD5 hashes) for complex joins between Movies and Locations.
- **Data Enrichment:** Joining ticket sales with rental costs to calculate profitability of the three cinemas in real-time.

Lineage-Diagramm



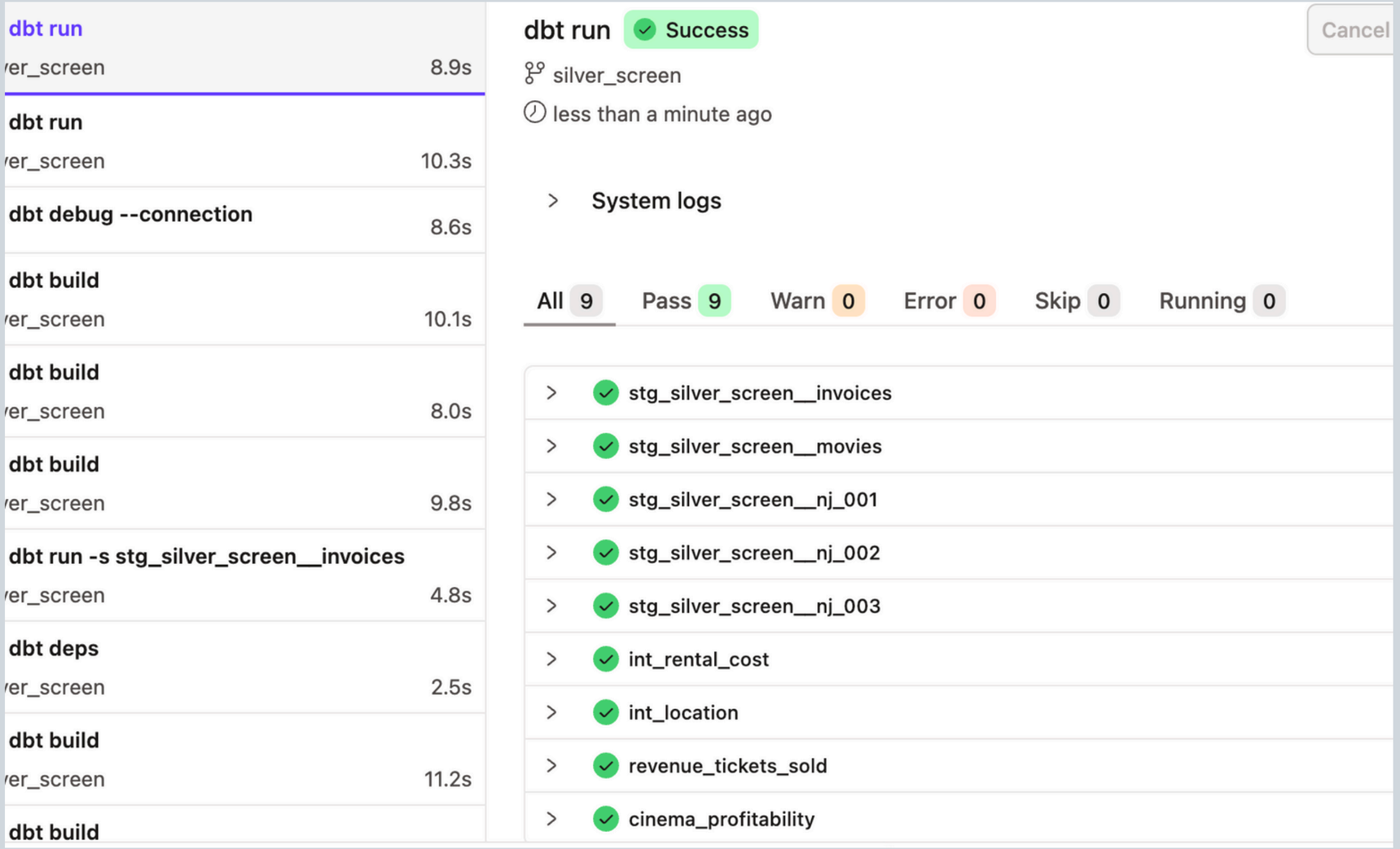
Testing & Automation

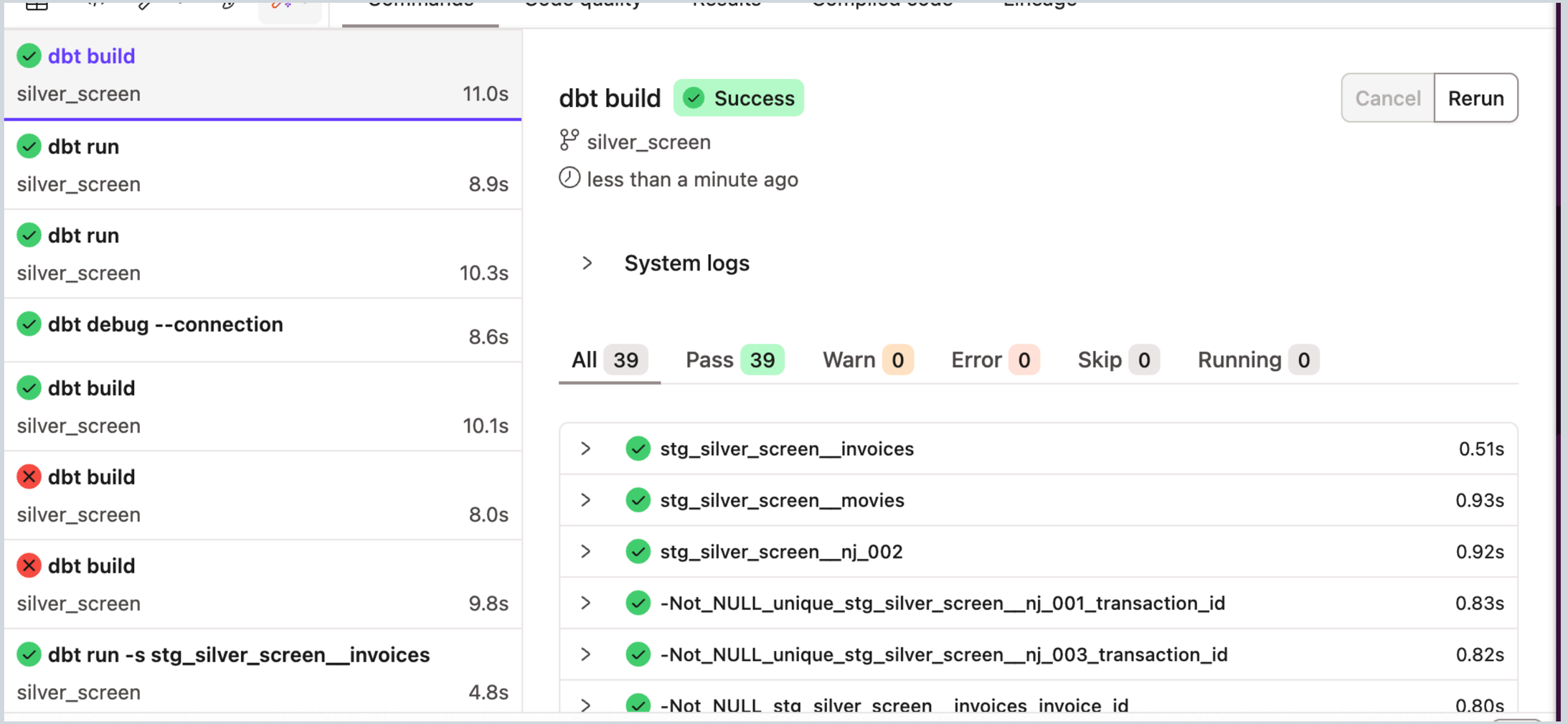
Ensuring Reliability through Automation

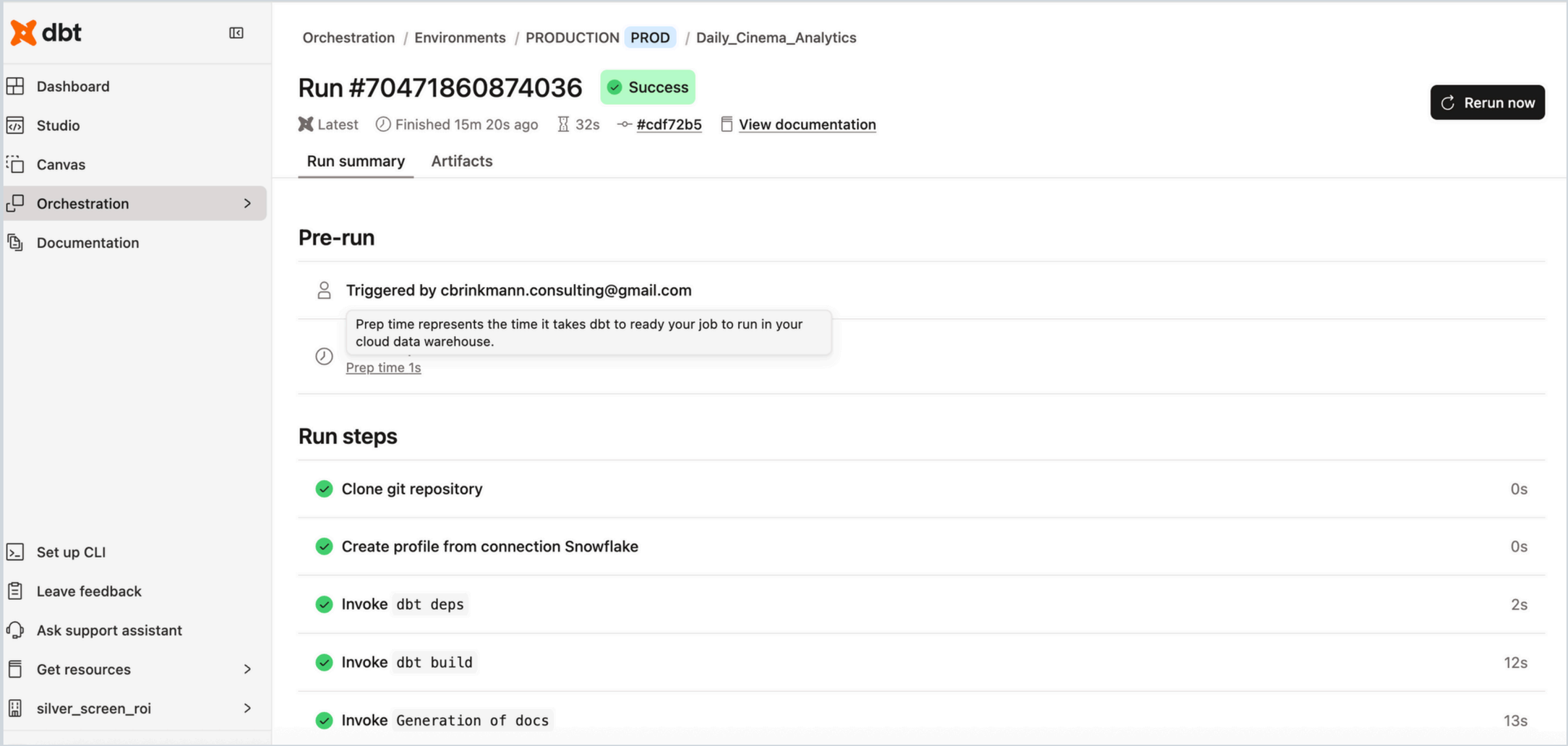
- **Proactive Testing:** Every transformation is guarded by automated tests:
  - **Generic Tests:** Checking for unique IDs and not_null values for movie_id, invoice_id and transaction_id
  - **Package-based Tests:** Using dbt-utils to validate row counts and accepted ranges to delete duplicated invoices
  (3 x intentical Invoice_id with another movie_id, location_id and release_date)

- **Job Scheduling:** * Automated production runs scheduled for **06:00 AM daily**.
  - **"Test-then-Deploy" logic:** If a test fails, the downstream tables are not updated, preventing "bad data" from reaching stakeholders.

- **Documentation:** Auto-generated documentation portal ensures that business definitions are always up to date with the code.


dbt run
ver_screen 8.9s
dbt run
ver_screen 10.3s
dbt debug --connection
8.6s
dbt build
ver_screen 10.1s
dbt build
ver_screen 8.0s
dbt build
ver_screen 9.8s
dbt run -s stg_silver_screen__invoices
ver_screen 4.8s
dbt deps
ver_screen 2.5s
dbt build
ver_screen 11.2s
dbt build
dbt run
Success
Cancel
silver_screen
less than a minute ago
System logs
All 9
Pass 9
Warn 0
Error 0
Skip 0
Running 0
stg_silver_screen__invoices
stg_silver_screen__movies
stg_silver_screen__nj_001
stg_silver_screen__nj_002
stg_silver_screen__nj_003
int_rental_cost
int_location
revenue_tickets_sold
cinema_profitability


dbt build
silver_screen
11.0s
dbt run
silver_screen
8.9s
dbt run
silver_screen
10.3s
dbt debug --connection
8.6s
dbt build
silver_screen
10.1s
dbt build
silver_screen
8.0s
dbt build
silver_screen
9.8s
dbt run -s stg_silver_screen__invoices
silver_screen
4.8s
dbt build
Success
Cancel
Rerun
silver_screen
less than a minute ago
System logs
All 39
Pass 39
Warn 0
Error 0
Skip 0
Running 0
stg_silver_screen__invoices
0.51s
stg_silver_screen__movies
0.93s
stg_silver_screen__nj_002
0.92s
-Not_NULL_unique_stg_silver_screen__nj_001_transaction_id
0.83s
-Not_NULL_unique_stg_silver_screen__nj_003_transaction_id
0.82s
0.80s



dbt build sucessfully

Final Outcome: Cinema Profitability Mart

- **Consolidated View:** Successfully merged Studio costs (RENTAL_COST) with theater performance (TOTAL_REVENUE, TOTAL_TICKETS_SOLD).
- **Granularity:** Data is now available at the Studio-Month-Location level, allowing for precise Cinema profitability analysis (as seen in the NJ_001 to NJ_003 performance breakdown).
- **Data Readiness:** The table is fully optimized for BI tools like Tableau or PowerBI, enabling instant dashboard updates.

Future Roadmap:

- **Advanced Analytics:** Integrating predictive modeling to forecast future "Rental Costs" based on historical studio trends.
- **Expanded Data Sources:** Incorporating marketing spend and concessions data (popcorn/drinks) for a 360-degree profitability view.
- **Real-time Alerts:** Setting up dbt Cloud notifications to alert the business team if specific KPIs (like Revenue per Ticket) fall below a defined threshold.

Cinema Profitability

STUDIO	MONTH	LOCATION_ID	RENTAL_COST	TOTAL_TICKETS_...	TOTAL_REVENUE...
20th Century St...	2024-05-01	NJ_001	340000	2337	82780
20th Century St...	2024-05-01	NJ_002	340000	31	61361
20th Century St...	2024-05-01	NJ_003	340000	3938	135458
20th Century St...	2024-06-01	NJ_001	340000	4608	162374
20th Century St...	2024-06-01	NJ_002	340000	61	111212
20th Century St...	2024-06-01	NJ_003	340000	7811	269242
20th Century St...	2024-08-01	NJ_001	320000	2373	83377
20th Century St...	2024-08-01	NJ_002	320000	31	64199
20th Century St...	2024-08-01	NJ_003	320000	3320	113757
20th Century St...	2024-09-01	NJ_001	320000	4731	165741
20th Century St...	2024-09-01	NJ_002	320000	61	120907
20th Century St...	2024-09-01	NJ_003	320000	7266	249775
Alcon Entertainment	2024-05-01	NJ_001	140000	1688	59880

"As you can see in the final table, we have full transparency. We can immediately identify which locations – such as NJ_003 – generate significantly higher revenues with the same rental costs. This forms the basis for our future budget planning."

THANK YOU !