

Project Workflow: Steps to Implement ETL Process

1. Load Source Data

- Import the CSV files as source definitions into Informatica.
- Ensure the source definitions are correctly configured to reflect the structure of the input files.

2. Stage the Data

- Load the data from the source files into staging tables in the MSSQL database.
- Create staging tables in the database, mirroring the structure of the source files, including primary and foreign key constraints for data integrity.

3. Design Dimensional Target Tables

- Define and create dimensional target tables in the SQL database.
- Ensure the dimensional tables align with the required schema for analytical purposes, including necessary primary and foreign key relationships.

4. Map and Develop Workflows

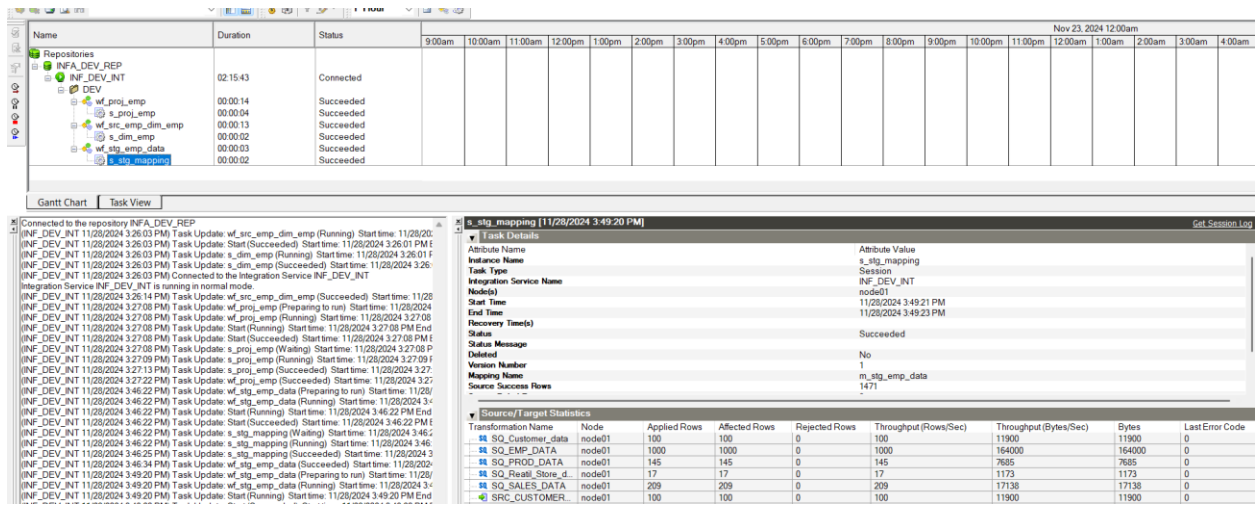
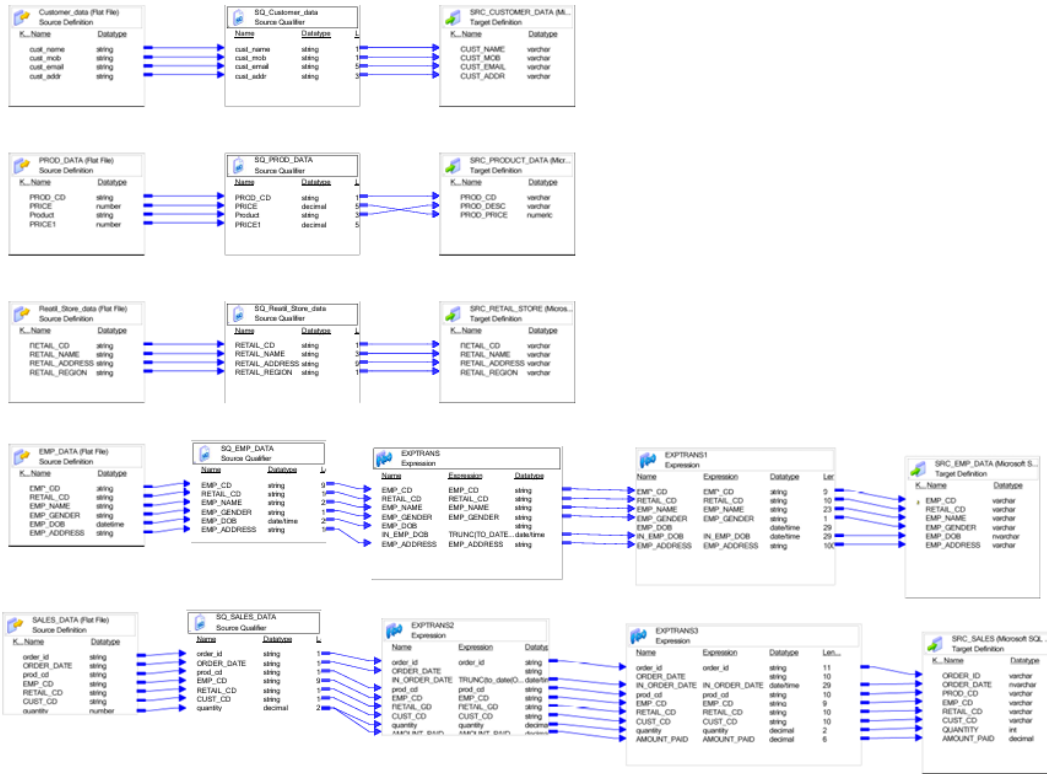
- Create workflows in Informatica for each staging source table to populate the corresponding dimensional tables.
- Implement the required transformation logic in mappings, such as:
 - **Slowly Changing Dimensions (SCD) Type 1:** Overwrite updates.
 - **SCD Type 2:** Maintain historical data.
 - **Insert-Only Logic:** For simple data addition.

5. Integrate with the Fact Table

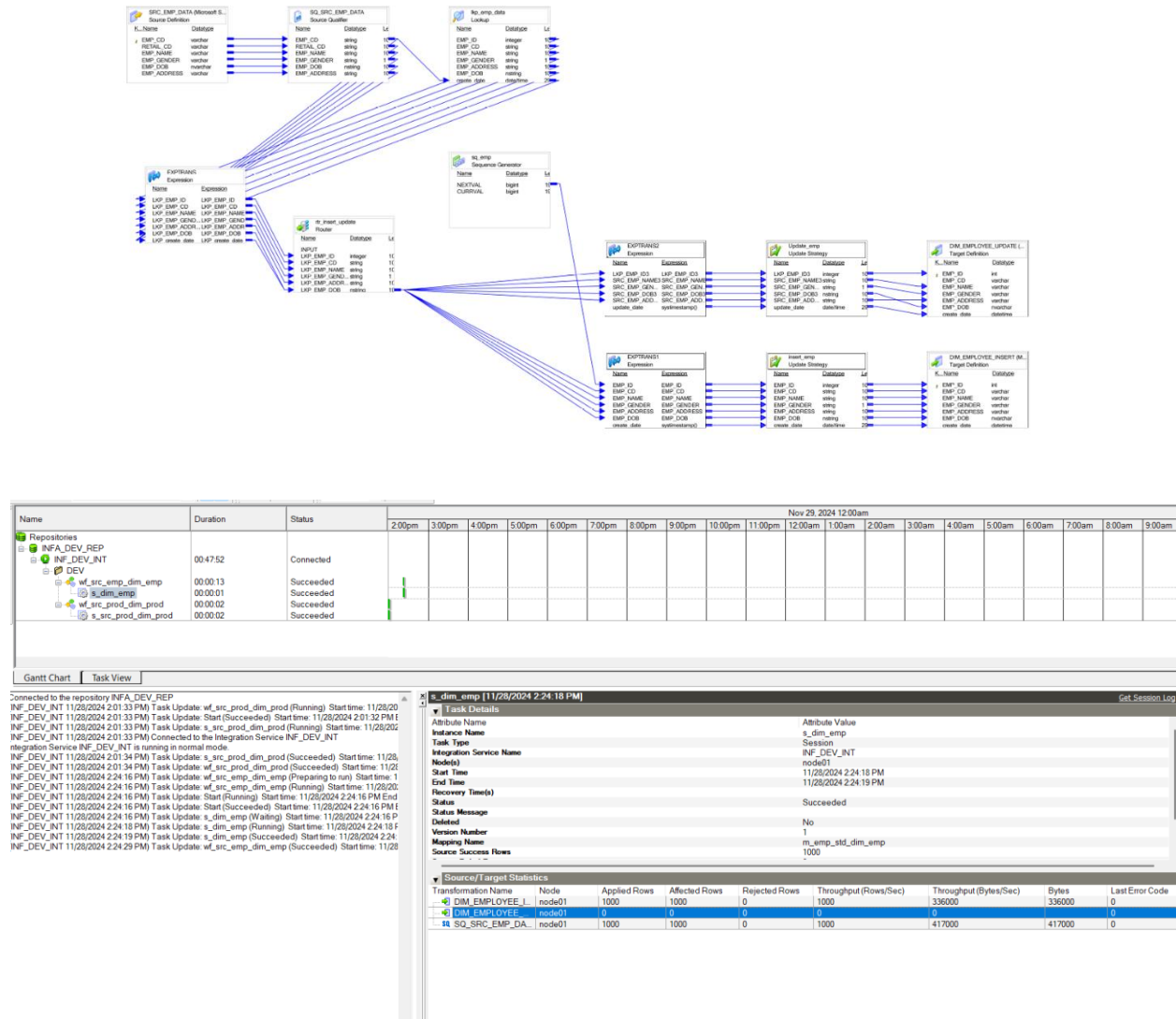
- Develop a final mapping to populate the **FACT_SALES** table.
- Use lookups on all dimensional tables to ensure accurate integration of data into the fact table.

Output Screenshots

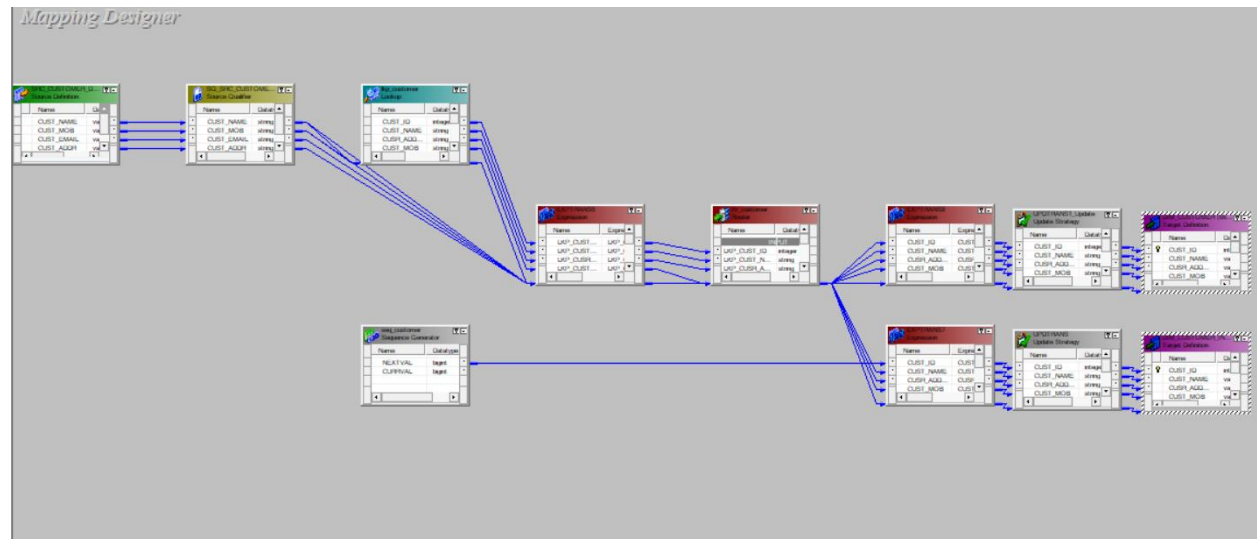
Loading the CSV files into the staging area



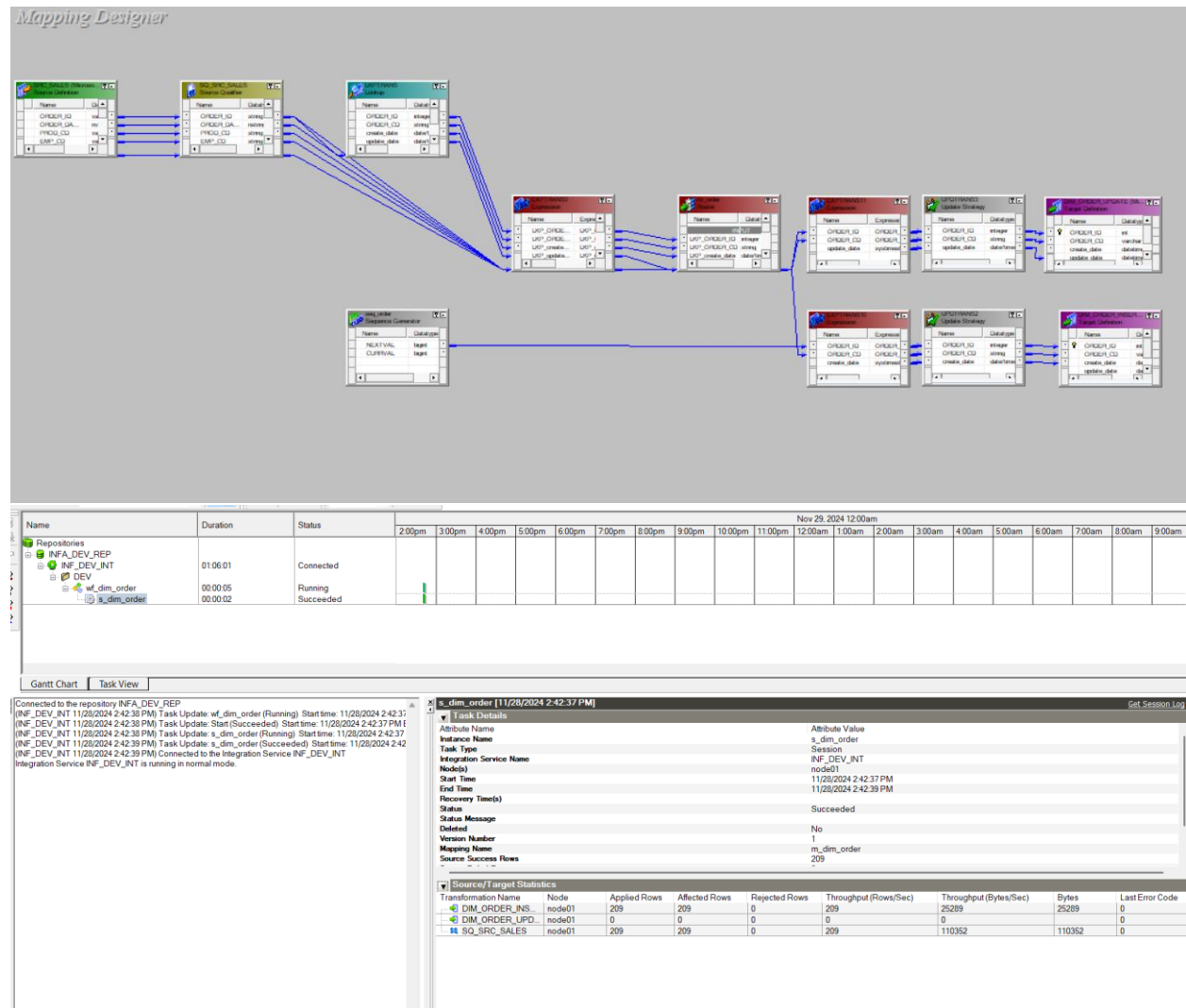
Emp Dimension:



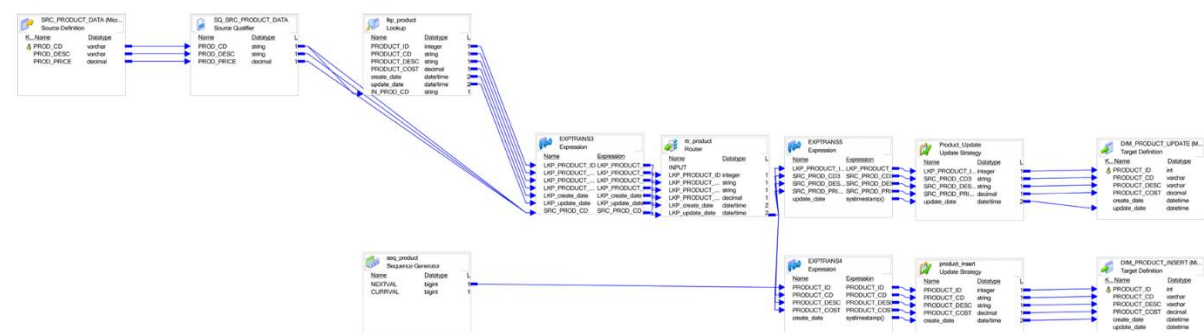
Customer Dimension:



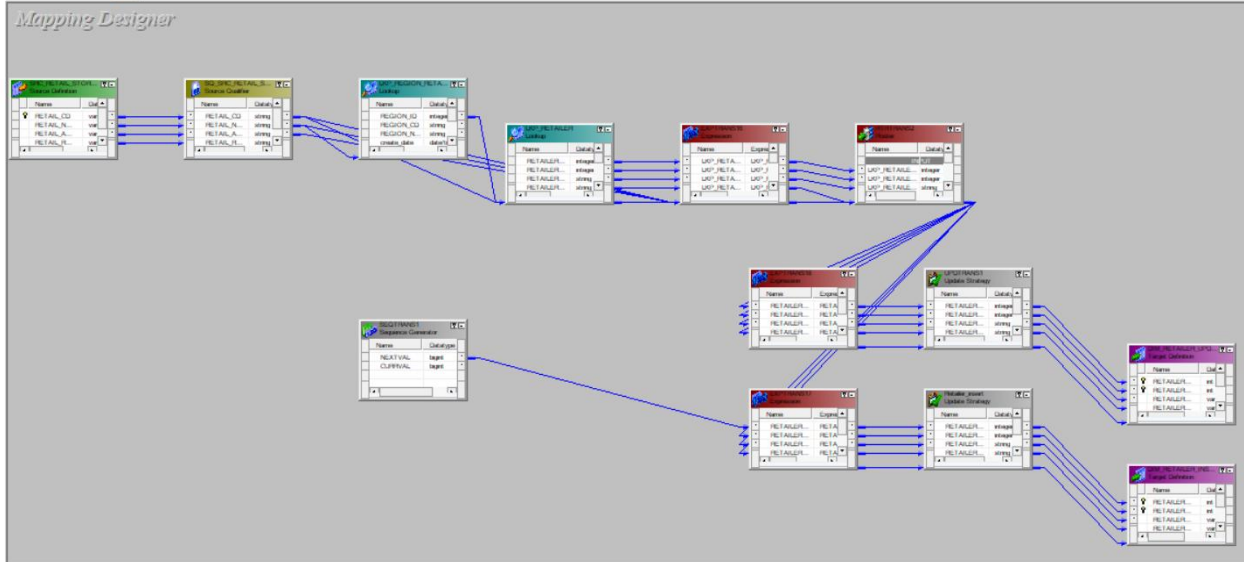
Order Dimension



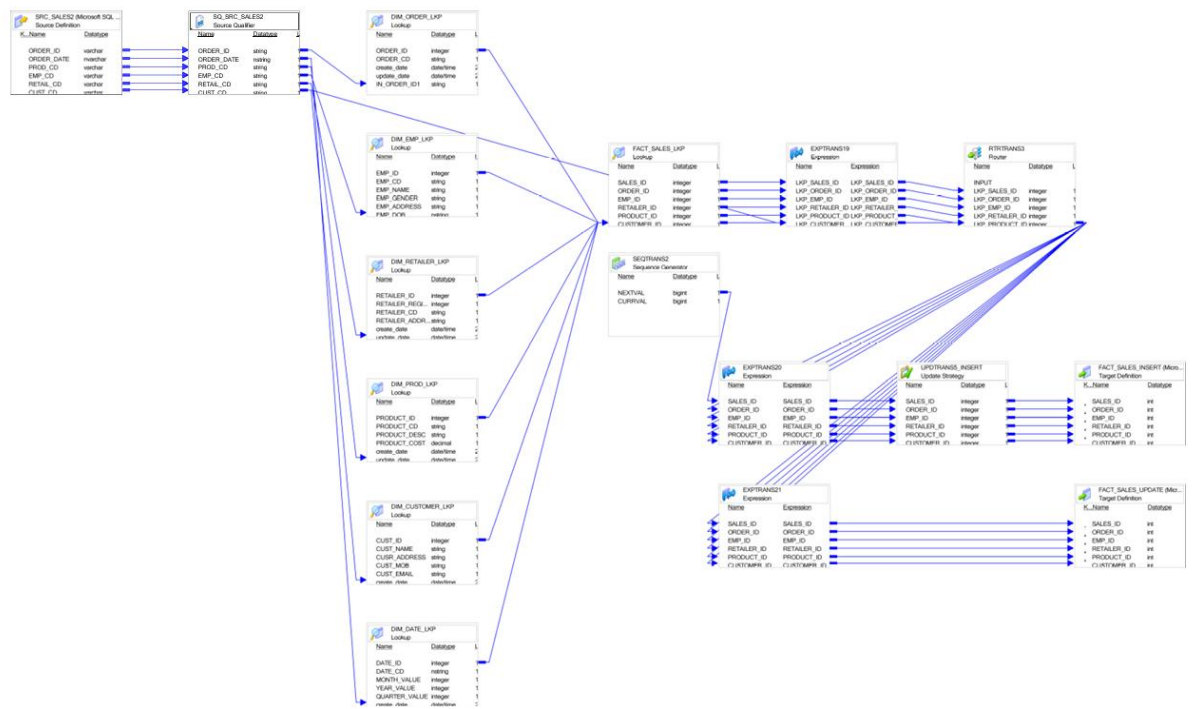
Product Dimension

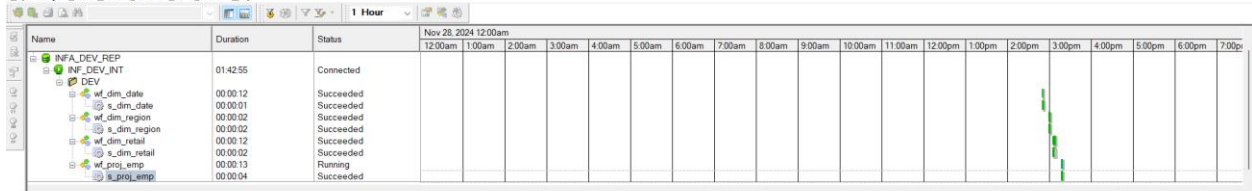






Fact Sale Dimension





Gantt Chart Task View

INF_DEV_INT 11/28/2024 2:52:24 PM Task Update s_dim_date (Running) Starttime: 11/28/2024 2:52:23 F
Integration Service INF_DEV_INT is running in normal mode.
INF_DEV_INT 11/28/2024 2:52:24 PM Task Update s_dim_date (Succeeded) Starttime: 11/28/2024 2:52:
INF_DEV_INT 11/28/2024 2:52:35 PM Task Update wt_dim_date (Succeeded) Starttime: 11/28/2024 2:52:
INF_DEV_INT 11/28/2024 3:01:12 PM Task Update wt_dim_region (Preparing to run) Starttime: 11/28/2024
INF_DEV_INT 11/28/2024 3:01:12 PM Task Update wt_dim_region (Running) Starttime: 11/28/2024 3:01:1
INF_DEV_INT 11/28/2024 3:01:12 PM Task Update s_dim_region (Waiting) Starttime: 11/28/2024 3:01:12
INF_DEV_INT 11/28/2024 3:01:12 PM Task Update s_dim_region (Running) Starttime: 11/28/2024 3:01:12
INF_DEV_INT 11/28/2024 3:01:14 PM Task Update s_dim_region (Succeeded) Starttime: 11/28/2024 3:0
INF_DEV_INT 11/28/2024 3:01:14 PM Task Update wt_dim_region (Succeeded) Starttime: 11/28/2024 3:1
INF_DEV_INT 11/28/2024 3:07:18 PM Task Update wt_dim_retail (Running) Starttime: 11/28/2024 3:07:18
INF_DEV_INT 11/28/2024 3:07:18 PM Task Update s_dim_retail (Preparing to run) Starttime: 11/28/2024
INF_DEV_INT 11/28/2024 3:07:18 PM Task Update s_dim_retail (Running) Starttime: 11/28/2024 3:07:18 PM E
INF_DEV_INT 11/28/2024 3:07:18 PM Task Update wt_dim_retail (Waiting) Starttime: 11/28/2024 3:07:18 F
INF_DEV_INT 11/28/2024 3:07:18 PM Task Update s_dim_retail (Running) Starttime: 11/28/2024 3:07:18 F
INF_DEV_INT 11/28/2024 3:07:21 PM Task Update s_dim_retail (Succeeded) Starttime: 11/28/2024 3:07:
INF_DEV_INT 11/28/2024 3:07:30 PM Task Update wt_dim_retail (Succeeded) Starttime: 11/28/2024 3:0
INF_DEV_INT 11/28/2024 3:09:05 PM Restart Task Request acknowledged
INF_DEV_INT 11/28/2024 3:09:05 PM Task Update wt_dim_retail (Preparing to run) Starttime: 11/28/2024
INF_DEV_INT 11/28/2024 3:09:05 PM Restart Task INFO: Starting execution of workflow [wt_dim_retail] in
INF_DEV_INT 11/28/2024 3:09:05 PM Task Update wt_dim_retail (Running) Starttime: 11/28/2024 3:09:05
INF_DEV_INT 11/28/2024 3:09:05 PM Task Update s_dim_retail (Waiting) Starttime: 11/28/2024 3:09:05 F
INF_DEV_INT 11/28/2024 3:09:05 PM Task Update s_dim_retail (Running) Starttime: 11/28/2024 3:09:05 F
INF_DEV_INT 11/28/2024 3:09:07 PM Task Update s_dim_retail (Succeeded) Starttime: 11/28/2024 3:09:
INF_DEV_INT 11/28/2024 3:09:17 PM Task Update wt_dim_retail (Succeeded) Starttime: 11/28/2024 3:09:
INF_DEV_INT 11/28/2024 3:09:17 PM Restart Task INFO: Workflow [wt_dim_retail] Execution succeeded
INF_DEV_INT 11/28/2024 3:19:23 PM Task Update wt_proj_emp (Preparing to run) Starttime: 11/28/2024
INF_DEV_INT 11/28/2024 3:19:23 PM Task Update wt_proj_emp (Running) Starttime: 11/28/2024 3:19:23
INF_DEV_INT 11/28/2024 3:19:23 PM Task Update Start (Running) Starttime: 11/28/2024 3:19:23 PM End
INF_DEV_INT 11/28/2024 3:19:23 PM Task Update Start (Succeeded) Starttime: 11/28/2024 3:19:23 PM E
INF_DEV_INT 11/28/2024 3:19:23 PM Task Update s_proj_emp (Waiting) Starttime: 11/28/2024 3:19:23 P
INF_DEV_INT 11/28/2024 3:19:24 PM Task Update s_proj_emp (Running) Starttime: 11/28/2024 3:19:24 F

s_proj_emp [11/28/2024 3:19:24 PM]

Get Session Log

Task Details

Attribute Name	Attribute Value
Instance Name	s_proj_emp
Task Type	Session
Integration Service Name	INF_DEV_INT
Node(s)	node01
Start Time	11/28/2024 3:19:24 PM
End Time	11/28/2024 3:19:28 PM
Recovery Time(s)	
Status	Succeeded
Status Message	
Deleted	No
Version Number	1
Mapping Name	m_project
Source Success Rows	209

Source/Target Statistics

Transformation Name	Node	Applied Rows	Affected Rows	Rejected Rows	Throughput (Rows/Sec)	Throughput (Bytes/Sec)	Bytes	Last Error Code
FACT_SALES_IP	node01	209	209	0	209	11704	11704	0
SO_SRC_SALES2	node01	209	209	0	209	110352	110352	0