**AWS Retail Sales Lakehouse — End-to-End Notes (Bronze → Silver → Gold)  
(No QuickSight)**

Everything up to Gold: architecture, SQL, partitioning, validation.

**Architecture & Layout**

S3: bronze/silver/gold; Glue Data Catalog for schemas; Athena for SQL.  
Format: Parquet (SNAPPY). Partition: year\_month (YYYY-MM).

**Silver Schema**

invoiceno, stockcode, description, quantity (bigint), unitprice (double), customerid (bigint), country, line\_total (double), order\_date (date), year\_month (string).

**Silver Create (schema)**

CREATE EXTERNAL TABLE retailsales\_lake\_db.silver\_retailsales (  
 invoiceno string, stockcode string, description string, quantity bigint, unitprice double,  
 customerid bigint, country string, line\_total double, order\_date date  
)  
PARTITIONED BY (year\_month string)  
STORED AS PARQUET  
LOCATION 's3://<bucket>/silver/retailsales/'  
TBLPROPERTIES ('parquet.compress'='SNAPPY');

**Populate Silver (example)**

INSERT INTO retailsales\_lake\_db.silver\_retailsales SELECT ... year\_month FROM retailsales\_lake\_db.bronze\_retailsales;

**Gold Tables (schema)**

gold\_sales\_by\_month(month\_start,date,total\_sales,double,order\_count,bigint,items\_sold,bigint) PARTITIONED BY (year\_month string)  
gold\_country\_sales\_by\_month(country,string,month\_start,date,revenue,double,items\_sold,bigint) PARTITIONED BY (year\_month)  
gold\_top\_products\_by\_month(product,string,month\_start,date,revenue,double,qty\_sold,bigint) PARTITIONED BY (year\_month)

**Gold Inserts**

Ensure year\_month is LAST in SELECT to avoid HIVE\_COLUMN\_ORDER\_MISMATCH.  
Use date\_trunc('month', order\_date) as month\_start; SUM(line\_total) as revenue/total\_sales.

**Views (optional)**

sales\_by\_month\_v, country\_sales\_by\_month\_v, top\_products\_by\_month\_v  
sales\_by\_month\_enriched\_v with LAG for MoM% in SQL.

**Validation**

SHOW PARTITIONS, COUNT(\*), LIMIT 10 selects on each Gold table.

**Troubleshooting**

Run MSCK REPAIR after new partitions. Verify S3 LOCATION. Keep partition col last in SELECT.