Project-2 Data Modeler

Schema Type:

A Star Schema data model was implemented with Sales_Fact as the central fact table. All dimension tables (Customer_Dim, Product_Dim, Region_Dim, and Date_Dim) are directly connected to the Sales_Fact table, following best practices for analytical models. The Returns_Fact table was modeled as a secondary fact table to represent return transactions, creating a controlled snowflake-like extension while maintaining model clarity.

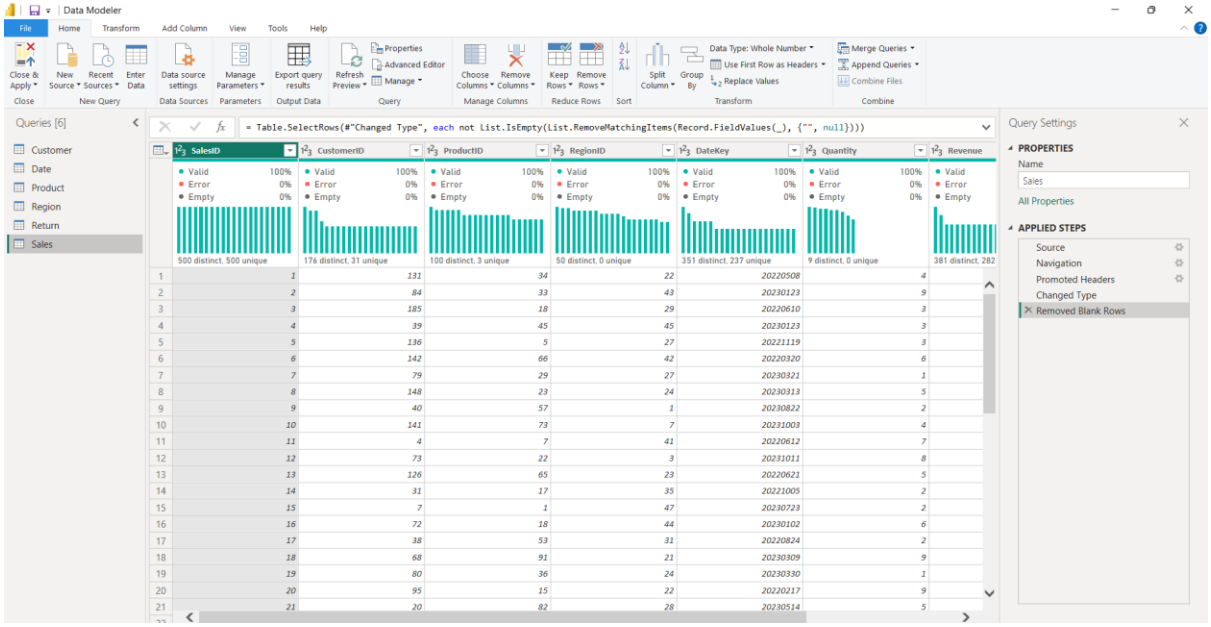
Relationship Logic:

All relationships were manually defined using primary key and foreign key mappings. Each dimension table has a one-to-many (1:*) relationship with the Sales_Fact table. Single-direction cross-filtering was used across the model to ensure predictable filter flow and optimal performance. An inactive relationship was intentionally created between Returns_Fact and Date_Dim using the ReturnDateKey to demonstrate handling of multiple date relationships within the same model.

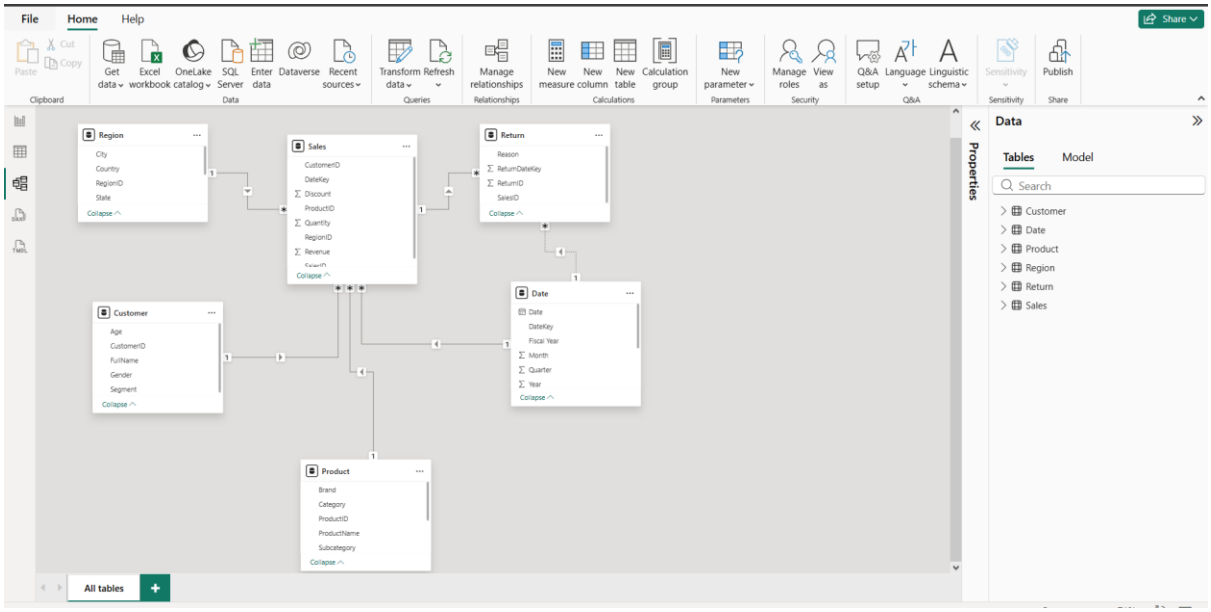
Issues Faced & Resolution:

An ambiguous filter path arose due to multiple relationships between fact tables and the Date_Dim table. This was resolved by keeping the Sales_Fact–Date_Dim relationship active and setting the Returns_Fact–Date_Dim relationship as inactive. This approach eliminated ambiguity while preserving analytical flexibility and aligns with industry-standard dimensional modeling practices.

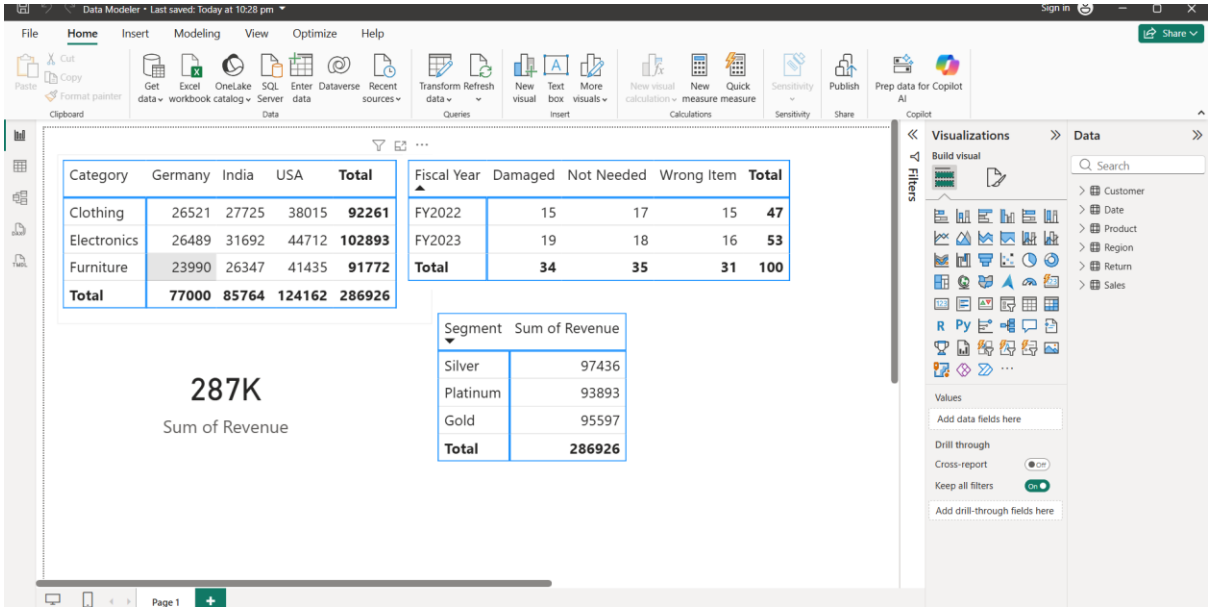
Transformation applied:



Snowflake Structure:



Matrix:



Category	Germany	India	USA	Total
Clothing	26521	27725	38015	**92261**
Electronics	26489	31692	44712	**102893**
Furniture	23990	26347	41435	**91772**
Total	**77000**	**85764**	**124162**	**286926**

Fiscal Year	Damaged	Not Needed	Wrong Item	Total
FY2022	15	17	15	**47**
FY2023	19	18	16	**53**
Total	**34**	**35**	**31**	**100**

Segment	Sum of Revenue
Silver	97436
Platinum	93893
Gold	95597
Total	**286926**