The initial database structure suffered from several critical flaws that hindered effective data analysis. First, product information was incorrectly stored in the Status field of the Orders table, making it impossible to distinguish between actual order statuses (e.g., "Processing") and product names (e.g., "Enterprise Switch 10"). This improper schema design led to unreliable queries, as filtering for real order statuses required manual exclusion of product names. Additionally, the absence of a dedicated Products table meant there was no centralized way to track inventory, pricing, or product categories, forcing analysts to rely on inconsistent text fields.

Second, the RMA (returns) table lacked proper relational integrity. A significant issue was that 38,161 RMA records had NULL CustomerID values, making it impossible to trace which customers were returning products. Without foreign key constraints, the database allowed orphaned records, leading to incomplete return analyses. Furthermore, the RMA table did not store ShippingAddress or regional data, making geographic return trends difficult to analyze unless joined with the Customers table—which itself had inconsistent state entries (e.g., "MA" vs. "Massachusetts").

Third, the database lacked normalization. The Orders table did not have an OrderItems junction table, meaning each order could not contain multiple products. This forced product names into a single Status field, creating duplicate entries and preventing granular sales/return analysis. Queries attempting to calculate return rates had to use inefficient subqueries or joins, slowing down performance. Additionally, there were no constraints to validate dates, resulting in nonsensical entries like 0000-00-00 and 2069-12-08.

Finally, reporting was severely limited. Without proper indexing, queries on large tables (e.g., RMA with thousands of records) were slow. The absence of views or stored procedures meant analysts had to write repetitive, complex joins for basic metrics. Data validation was minimal, leading to dirty data—such as Massachusetts orders appearing as zero due to mismatched state abbreviations ("MA" vs. "Massachusetts"). These issues made it nearly impossible to generate accurate sales or return reports without manual cleanup.