Team,

After combing through the data in the products, customers, and sales tables I would like to share some insights.

Each table contained null, or *empty*, values which is concerning for two main reasons:

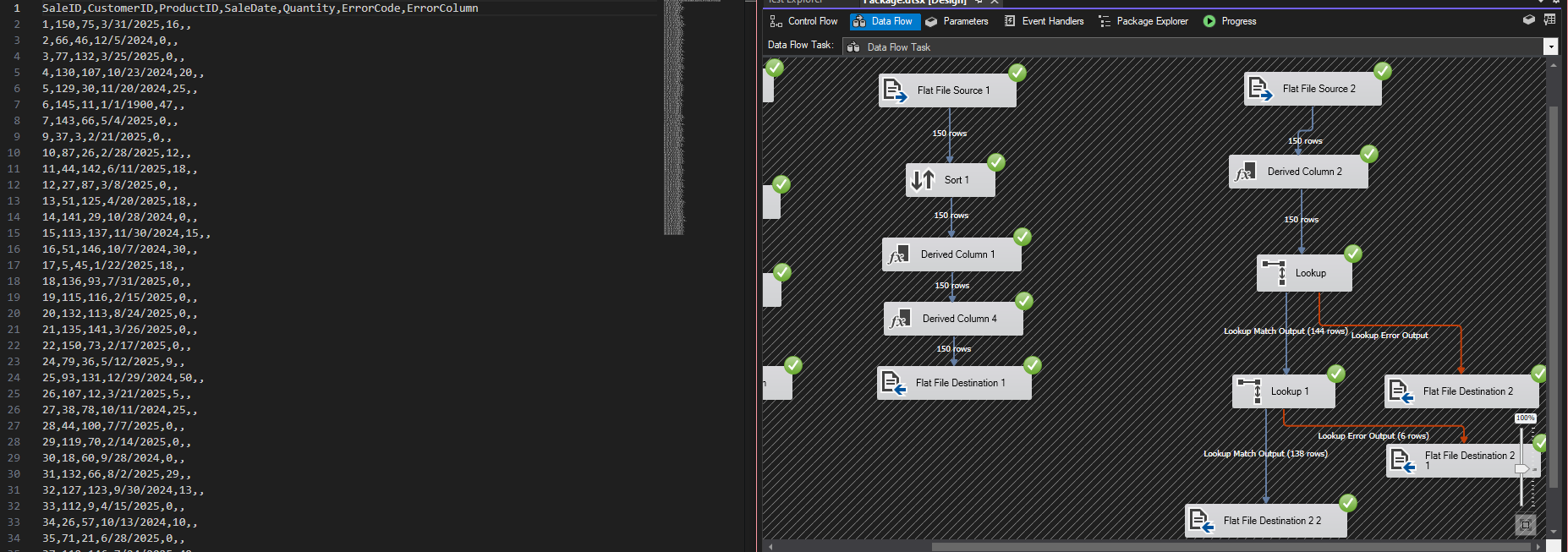
* 1) Where is this data being stored if at all? The potential loss of this data can be detrimental to operations and accurate business intelligence.
* 2) The holes left in this data can be hazardous to company systems if not handled properly.

While I am unable to fix the first problem (auditing team, be ready), I have come up with a solution for the second. I have settled on several placeholder values that are easily identifiable and can be easily integrated into the current system without disrupting reports.

While auditing the keys in the sales table, I came across a few product IDs that do not match anything we have in our system. This could have happened for many reasons, but I will leave it to the auditing team to figure out if these products still exist in the company and we will proceed from there. Once we establish if these products are still in use, I will have more to share on this.

Other than the major issues discussed above, I am continuing to fix spelling and capitalization inconsistencies as I come across them. As I mentioned before, these errors compromise the integrity of the database and make reports less reliable. As we move forward, I expect to update the applications we use to utilize more automation and validate input before it reaches the database. This way these errors will be much less likely to occur again.

Below is the Sales data after cleaning:



Let me know if you have any questions.

Best,

Connor Gaskill