Dear AIQ Backend Engineer,

I have found the reason for the data discrepancy between Johnny’s Jeans and our platform. It seems that our system either failed to retrieve a file or failed to update our data with a file.

Please see the attached summary.py as a reference. Using delta0.csv as a starting point, I subsequently updated the data with delta files 1-4. From the resulting data, the sum of the total spend was inline with the Johnny’s Jeans finance team at $33,904,525. However, the same calculation of total spend from our AIQ-user-summary.csv was $29,162,044. If we just look at the first five rows of the resulting dataframes, we can already see that our AIQ summary failed to update certain rows. As a result, our AIQ summary was incomplete as it did not contain all the updated data.

Continuing, please see the attached issue.py. I computed four copies of the datasets, omitting one update file in each. From there, I calculated the sum of the total spend, and we can see that after omitting the delta2 update file, the result of the total sum is identical to the AIQ summary ($29,162,044). Furthermore, looking at the first five rows of the data frame with the delta2 update file omitted, we can see that it is identical to AIQ-user-summary. Thus we can come to the conclusion that the data discrepancy is caused by a failure to incorporate delta2.csv into our data for Johnny’s Jeans.

I hope this has clarified the issue.

Kind regards,

Chloe