

/*

Mkhanyisi Gamedze
Insight Squared - Product Support Associate interview
Task: Resolving Client Issue

2076169348

*/

Problem Description:

Here's the low-down: we are often working with data that's unfamiliar, sometimes even while we're on the phone with the customer. As we've done with other candidates, I'd like to give you a chance to show how you work with that sort of situation by giving you this hypothetical scenario.

You've received this email from a customer today:

"Thomas E and Antoine each had calls in July that are missing from the reports. Please explain??? Tim"

You have the data linked below, which is all of their activities from a 30 day date range. This is the data that sits directly behind the InsightSquared reports, so it's a result of both the raw data from the source system, and of the configuration we have in place that determines how that data is interpreted before being written into these tables.

Please write up a response that you would send to Tim, explaining as much about the problem as you can figure out from the data. (If there's something you see in the data that you don't understand, just make your best guess at what it means. If there's a guess where you feel like you are taking an especially random shot in the dark, feel free to flag that for me.) The data: <https://cl.ly/499b40a05410>

Non Technical Response

Hi Tim,

You called us yesterday about missing data in our reports for Thomas Edison and Antoine Lavoisier on calls they made in July s reported by our Insight-Squared Product. I forwarded this request to the technical team here at InsightSquared and they looked through our system to figure out the call history of those users in the past month(30 days) and to see whether any calls were made by them.

After reviewing the request, the team is happy to say that no calls were missed on the recordings on our internal system and reports as well. Our records showed that both Thomas Edison and Antoine Lavoisier were on the employees list and for the recorded Updated Activity list, their **owner_id** were used to record their activity. Please check the product records again by searching the activity records by Owner ID's as that is how our system works. The activity sheet is for the past 30 days including both July and June dates, of which our source system focused more on for searching.

Please try searching the Updated activity by the suggested parameters and see if this problem persists. If any further issues arise, please feel free to call our Customer Support team as they would be very happy to help you.

Thank you and have a great day.

Best,
Product Support at Insight Squared

Technical Overview of Solution

Given the caller's enquiry about service and misunderstanding they had. My initial intuition was to automate this validation process and check whether the suggested ID's exist within the records. Using IPython notebooks and Python library Pandas, I load the CSV data provided into dataframes and then do my processing. This file is linked with the other documents. My solutions start off by Reviewing the data within the both datatables for Employees and Update Activity. Both data tables hold relational data and so I looked into column headers in both tables to see what relates the data. The **id** in Column 1 for employees is the same ID used to record **owner_id** in the activity table, and thus this would act as a good tool to match foreign keys in tools like SQL.

Since I worked with python, I selected all names matching the given search names in order to validate that these users were active in the first period. The employees table gave all the related information for each user and from these I copied the id for selecting the **call activity details**. Searching for the call activity, I specifically set the month of **July = 7** as an additional parameter in the activity sheet. The output results of those calls resulted in 194 recorded calls for Antoine and 119 recorded calls for Thomas Edison.

The user did not specify more so which specific dates are missing and so the conclusion is that these records exist given the proof.

I made assumptions about the nature of the product and reporting methods used to conjecture my findings on this case.