How to reveal the missing customer in tabular model / Power BI with DAX set operation

(Please note: there are many ways to compare two database tables for difference in SQL. I am focus on data in the tabular model, therefore, need to use DAX )

Missing customer is an interesting business question which business cares about. Business wants to see why or how to make them purchase again. Therefore, to identify these customer population is important.

There are quite a few posts on this topic online. Most elegant way I have seen is to use table variable to isolate different time frame and use DAX set operation to identify the target customer population. I vote for that approach is because of code readability and performance.

One thing though, most post using set operation is focusing on counting the missing customer, I tried to search if there are posts to further identify those customer. But I am out of luck to find.

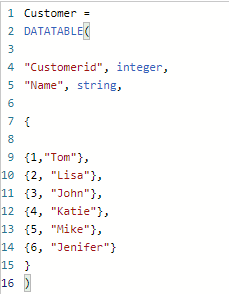
This is actually a legitimate question for business. Once you tell the business that so and so number of customers are missing. They will naturally ask who those customers are.

I tried this myself in DAX and struggled for a few weeks. This is my solutions.

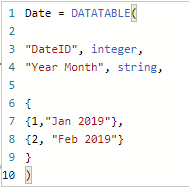
I will use a few simple tables containing some dummy data in power BI to imitate tabular model and business problem. Only 6 customers, 2 months and a few line of sale records so that it is easy to understand. The same principle applies to more complicated scenario, you just have to adjust it though.

First, let us build model in power BI desktop.

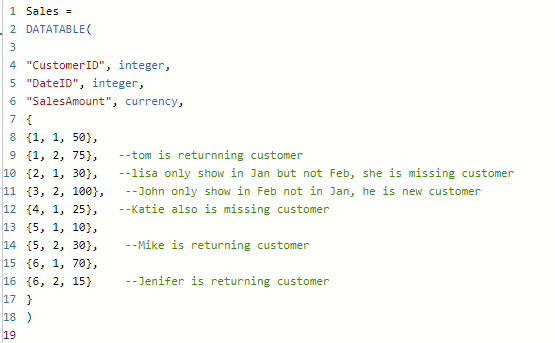
Use DAX datatable function to create customer table in power BI desktop (model->new table). This makes it easier if you want to edit it later



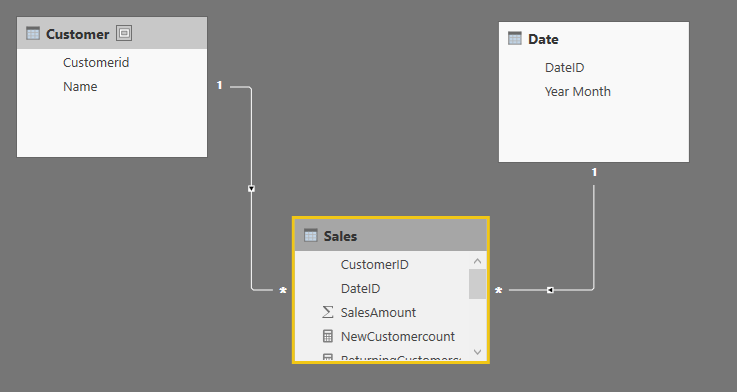
Same way, create date table



Now, let us create the fact sales table. Depending on if they purchase in Feb. I deliberately generate records so that we have 1 new customer, 2 missing customer, 3 returning customer

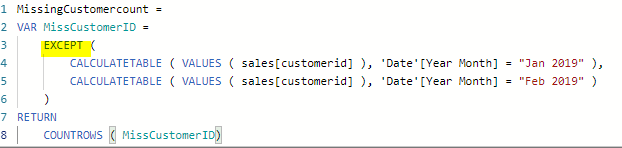


Now in diagram view, connect customer and date to sales table. You just built a smallest data warehouse in the world.



First, let us warm up by creating a missing customer count measure using table variable

Right click sales table and choose create new measure.

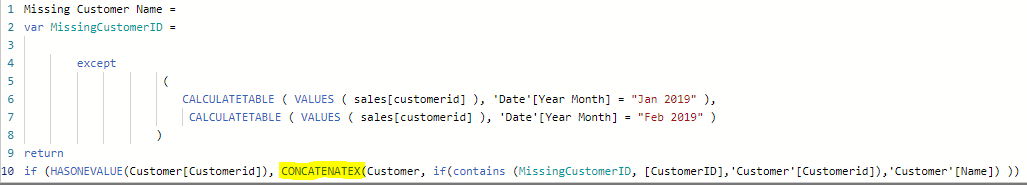


Notice the except function take two table params, records in the first table not in the second table will be kept. CustomerID in the Jan 2019 but not in Feb 2019, obviously, they belong to missing customer. This customerID pool is used to count how many inside the pool by countrows function.

Similarly, we can create returning customer and new customer.

Now how to reveal the customer who is missing? Manually, we can take that ID and search the customer name. But we need to do it programmatically since in real world, customer table can have millions of records.

It is simple to do in SQL. But for DAX that is a different ball game. The tricky part, you need to compare the customer table customerId with the population of missing id. If it is in the missing id population, give the customer Name. If not, give null value.

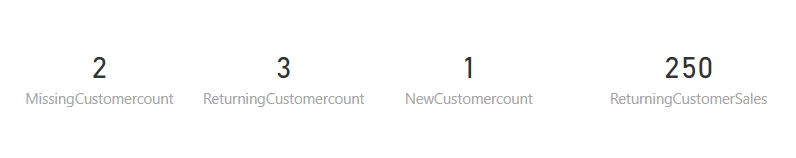


Because you only want to compare one customerID at a time. So, we use hasonevalue function. We need to iterate customer table one customerID at a time, that is why we use concatenateX, which can handle string like customer name (concatenatex is used to combine string over several row, in our case, we only have one row, we just use it as iterator and string handling ability here). I tried to use DAX maxx, but that did not take string unlike its SQL cousin. Contains function allow we compare missing customerIDs pool with the current row customer table customerID. If it is in the missing ID pool, we will give customer Name.

Similarly, we can create name measure for new, returning customer.

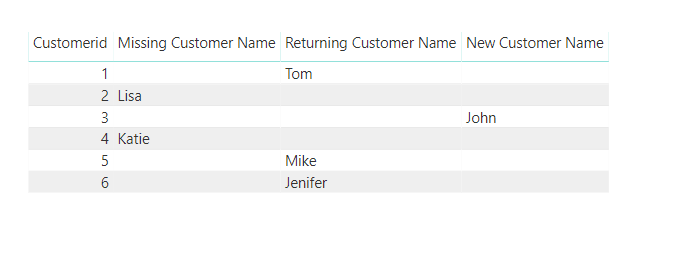
Let us put these measure in action. This is just to show idea, we don’t focus on the cosmetic side of it. So bear with me about the appearance.

For count measures, I use power BI card to show them



Seems to me, set function can identify the customer as expected.

Who are those customers?



So, the measure can correctly identify the user name.

Thank you for visiting my site. Hope you feel this useful.

Happy BI

As usual, you can download pbix file here.

Wenlei