

Project Documentation: Help Desk

A. BOGNAR

10/24/2021

Overview

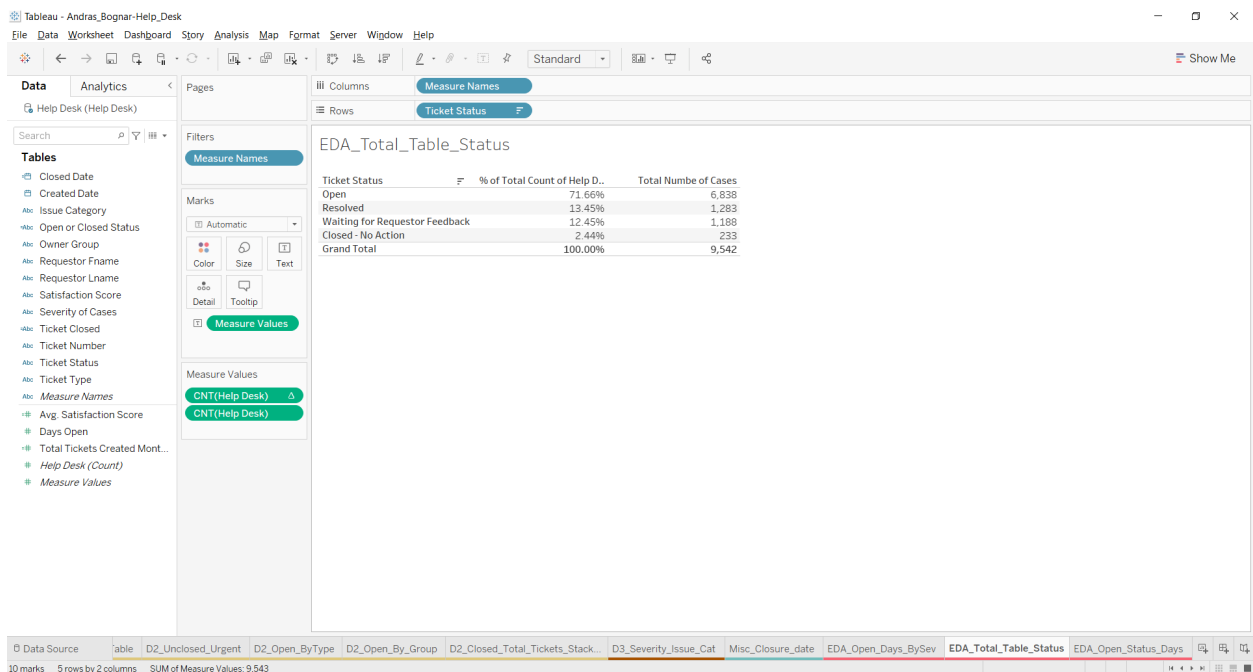
I chose the *Help Desk* dataset as I previously worked with similar data and wanted to see if I can find familiar patterns.

My original research questions were:

- How long does it take to close an IT ticket on average?
- What factors affect client satisfaction the most?
- Do longer processing times lead to lower satisfaction?

Exploratory Data Analysis

First, I investigated the data by creating simple tables and bar charts and I found several important details. To begin with, I noted that over 70% of all cases in the dataset are under Open status which is surprising since I expected most cases to be closed since the data cases were opened over 8 months. Moreover, an addition 12.5% of cases are waiting for requester input, so in total only about 16% of cases were closed in some way.



Furthermore, I investigated the variable *Days Opened* since this was crucial for my research question and I found unexpected results here as well. It turns out that cases in Open status have different number of open days, even if they were created on the same day. Since we don't have a closure day in the dataset, this means that for cases in *Open* (and similarly in *Waiting for Requester Feedback*) status, we can't confirm the processing time and we can't really tell if they are still pending or not.

Year of Created Date	Month of Created Date	Day of Created Date	Distinct count of Days Open
2020	March	1	22
2020	March	2	21
2020	March	3	21
2020	March	4	20
2020	March	5	21
2020	March	6	20
2020	March	7	19
2020	March	8	21
2020	March	9	22
2020	March	10	23
2020	March	11	22
2020	March	12	23
2020	March	13	18
2020	March	14	24
2020	March	15	23
2020	March	16	18

One possible explanation, I could think of for this discrepancy is that some type of cases might not be moved to *Closed* status after their closure. However, if that is the case, their processing times can not be tracked. Human error may also play a role in incorrect case statuses. The longest open day value in the dataset is 37 days, therefore, I could have decided to include all Open cases older than that as Closed cases. But without any contact to ask questions about the nature if these cases, I decided not to make massive changes to the dataset.

Lastly, I used a simple calculated field to get the closure date of tickets. Due to the above problems, I did not use this field in my dashboards. However, when validating the field I noted that all cases were created between March and October 2020. Based on that I decided to focus my analysis on the second and third quarters of the year since this is the period where we most likely have full information already.

Filters

Marks

Rows

Ticket Number **DAY(Created Date)** **DAY(Closed Date)**

Misc_Closure_date

Ticket Number	Day of Created ..	Day of Closed Date	
949-264677-4376-5...	October 30, 2020	November 26, 2020	27
910-953563-9256-2...	October 30, 2020	November 22, 2020	23
907-320645-9010-8-K	October 30, 2020	November 14, 2020	15

Closed Date

[Created Date] + [Days Open]

The calculation is valid. 1 Dependency

ABS(number)

Returns the absolute value of the given number.

Example: ABS(-7) = 7

Figure 1: Validating Closed Date field

Accordingly, based on these new insights I reconsidered my research questions.

Revised Analytics Questions

As discussed previously, Open cases are very different from Closed ones in the dataset. Therefore, I decided to investigate them separately with different, more appropriate questions.

For Closed tickets:

- What was the average processing time?
- Were urgent cases completed faster?
- How satisfied were clients with the cases?
- Did processing time had an affect on satisfaction?

For Open tickets:

- Were urgent cases left in Open state?
- Were Open cases more common for specific services? Or specific clients?

Note that for the question of client satisfaction I only analyzed closed cases. It might be possible to change scores for tickets that are still open so that information might not be relevant.

With this separation in mind, I first created a calculated field called *Open or Closed Status* to save time not rely on ticket status when filtering.

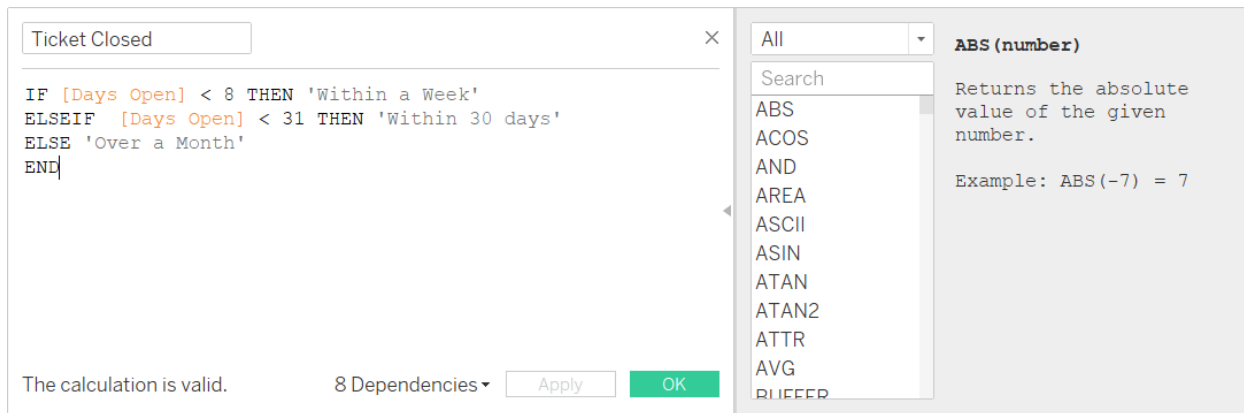


Figure 2: Calculated field for closed/open cases

A different filter that I used for every single chart in my work was quarter of creation date. As explained in the previous chapter, I analysed only cases created in Q2-3 2020. With these restrictions, I started designing my dashboard for closed cases.

Dashboard Description - 2020 Q2-Q3 Report: Closed Cases

Overview

The aim of the first dashboard is to 1. show how long it takes to close cases, 2. see how satisfied clients are with them and 3. see if there are any connections between the two.

The dashboard has 5 graphs: figure 1. shows the important data I presented earlier: closed cases make up a small minority of all cases. Figures 2-3 focus on the average closing time of cases while figures 4-5 showcase satisfaction. As I will explain, these last 2 are connected to figure 3 to show any connection between processing time and satisfaction. I deliberately chose figures 1 and 3 to include dynamic visuals while the figures on the bottom are not meant to be clicked by the user.

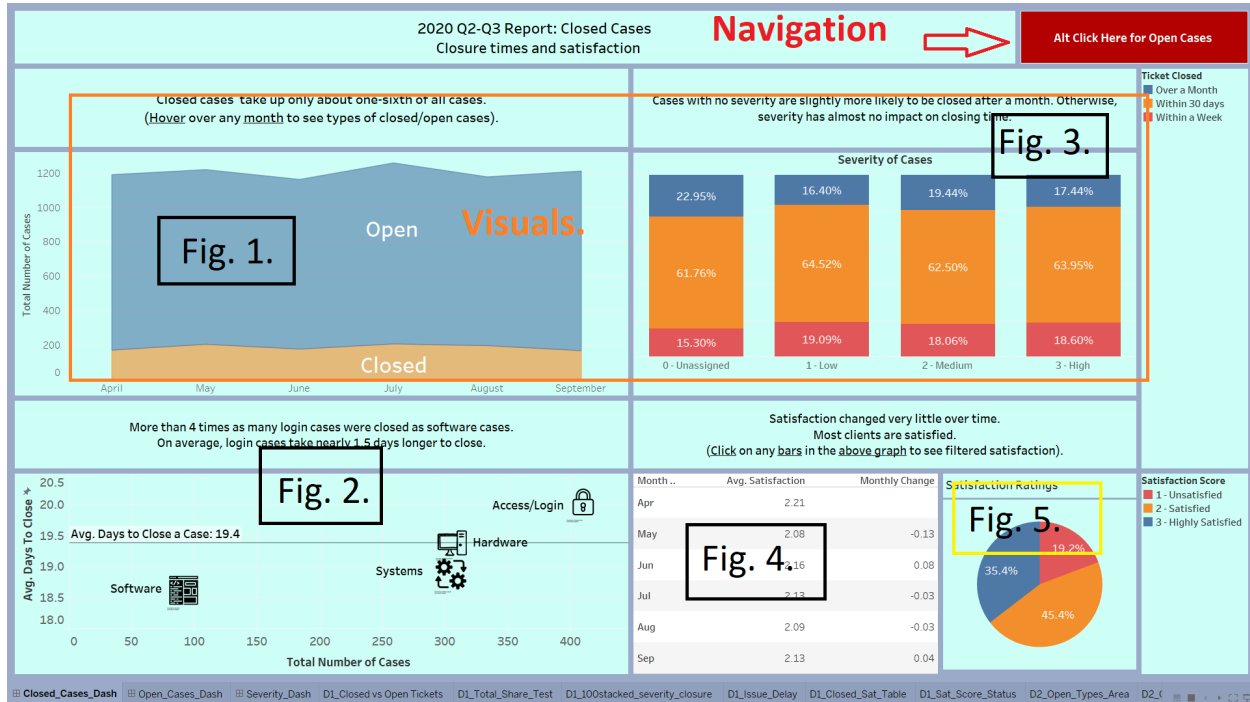


Figure 3: Figure 1 of Dashboard 1

Charts

Now let's look at the specific functionalities. While **figure 1** delivers a simple message - most cases were Open throughout Q2-3 - it also has a tooltip option. When hovering any point, a separate pie chart will show the distribution of the case states in the given month. This gives more insight, we can see that most closed cases were Resolved, most open cases are simply open. While for most visuals I tried to use the same color schemes, in this instance I made sure to use a lighter and a darker variation of the original open-closed colors to signal that these subcategories are connected.

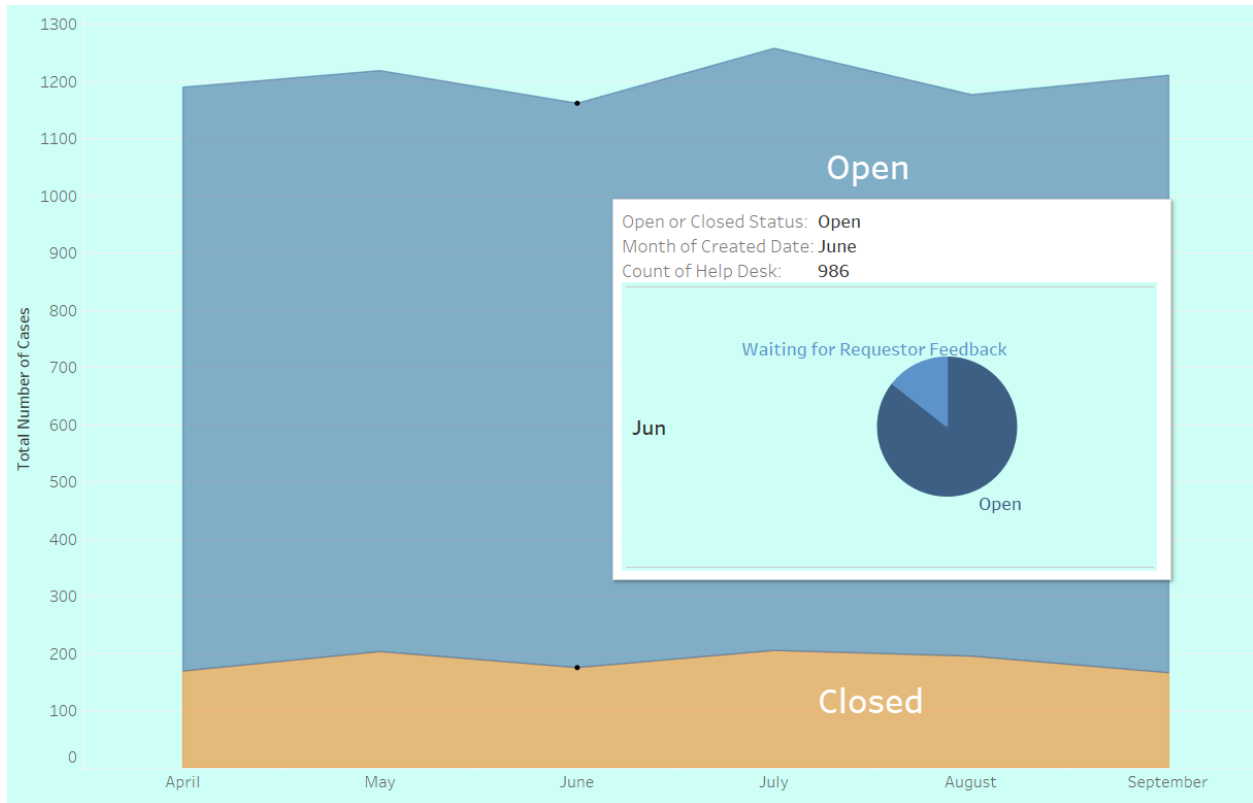


Figure 2 is the only *scatter plot* in the file and it is for a good reason, it compares two measurements: average processing time to number of cases by the service type. This is also the only figure where I **edited** the **length of** any **axis**. While it is generally considered to be good practice to leave axes starting points at 0, in this case I wanted to highlight that a seemingly small difference is rather significant. The processing time between login and software related cases differ by nearly 1.5 days. That is a big difference, even if a lot more login cases were closed. I also stressed this difference in the title of the graph to draw the viewer's attention to the changed axis. Additionally I used a **trend line** here to show the overall average processing time. Here, I had to make sure to change the base of the average from average value to total value. This is because we are not interested in the simple average of the 4 categories but their weighted mean by how many cases they each had. This will give us the overall average for all closed cases. Lastly, I should also talk about the visual representation of the categories used in the graph. I decided to use shape icons to represent the different service types. I downloaded these icons from [A Creative Commons license website](#). I used the following four images:

- Login
- Software
- Hardware
- Systems

×

Line

Band

Distribution

Box Plot

Scope

☐ Entire Table

☐ Per Pane

☒ Per Cell

Line

Value:AVG(Days Open)Total

Label:Customs to Close a Case: <Value>

Tooltip:Automatic

Line only95

Formatting

Line:

Fill Above:None

Fill Below:None

☒ Show recalculated line for highlighted or selected data points

OK

Edit Shape [Issue Category]

Select Data Item:

- Access/Login
- Hardware
- Software
- Systems

Select Shape Palette:

Issues

[Computer Monitor Icon] [Padlock Icon] [Document with Magnifying Glass Icon] [Gear Icon]

Assign Palette Reload Shapes

Reset OK Cancel Apply

Moving onto **Figure 3**, here I used a *100% stacked bar chart* to compare the time period spent on each case by Severity. Instead of average time spent, here I introduced a new **calculated field** where I used an IF function to group the cases in three categories: Closed within a week, within 30 days and cases closed after more than a month. These categories seemed sensible for this purpose as the longest values in days open was 37 as mentioned before. Also, I expected that cases taking longer would have lower satisfaction levels and that more severe cases would be dealt with in less time. What I found however, was that other than a small difference in unassigned cases, cases with 1-2-3 severity were generally closed in the same time frame (most within a month).

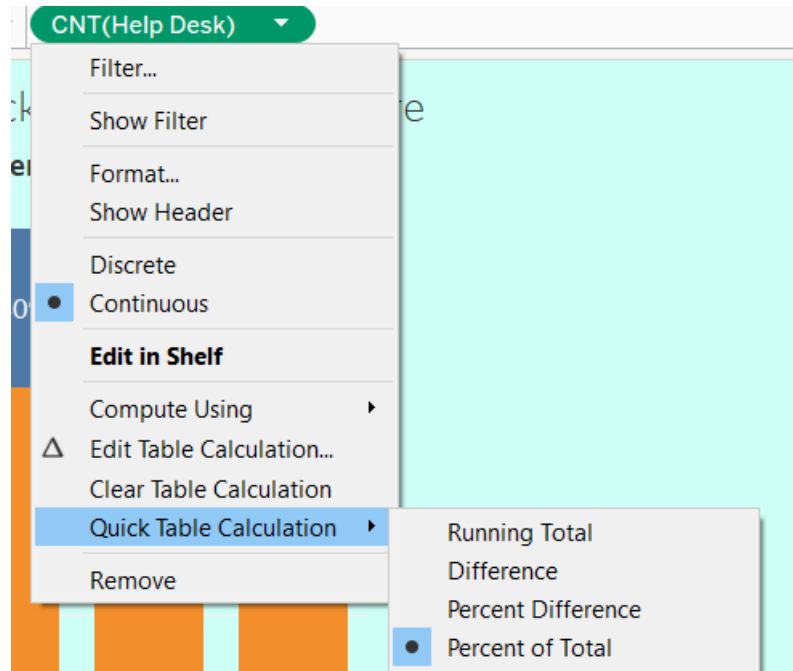


Figure 6: Quick Table Calc used for 100% bar graph

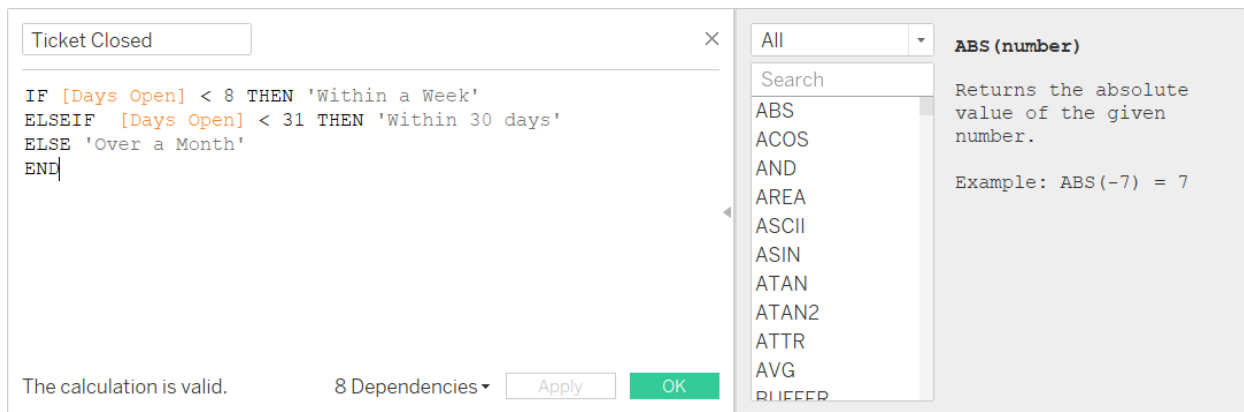


Figure 7: Calculated field about processing times

Still, I made good use of Figure 3 by connecting it to **Figures 4 and 5**. A simple table and a bar graph to investigate satisfaction levels, they are filtered through an **action** by the severity and closing time period selected in Figure 3. This allows the viewer to see whether either of these variables have an affect on satisfaction. For my satisfaction measurement, I created another calculated field called **Avg. Satisfaction Score**. This is another IF function, where I made sure to not include any scores of 0 - *Unknown* since these

are not actual values and adding the zeroes would distort the final average. The table of Figure 4 proved to be a simple solution to show that there is no significant change in satisfaction over time, I used another QTC of difference between each month.

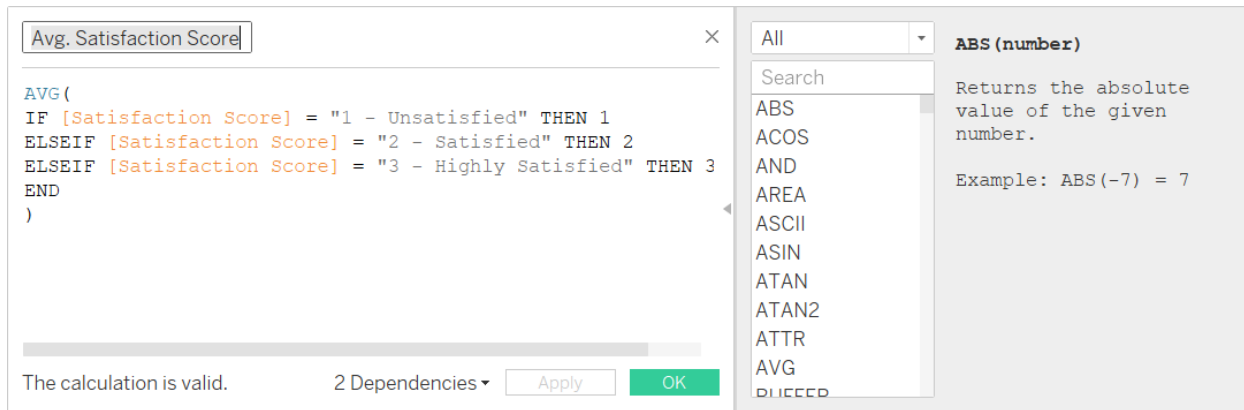


Figure 8: Calculated field about client satisfaction

This concludes the elements of Dashboard 1, by using the *navigation button* in the top right corner, one can move Dashboard 2.

Dashboard Description - 2020 Q2-Q3 Report: Open Cases

Overview

Dashboard 2 is all about the open status cases and aims to answer 2 questions: Are urgent cases left in Open status and what kind of services are kept open?

This dashboard also contains 5 graphs and have 2 interactions.

Charts

Figures 1 and 2 give general information about open cases over time and in total. The table is used for **cross filtering** across the dashboard to see how big the difference is between the two types.

Figure 3 is the first standard line chart, but it combines 2 graphs through **synchronized dual axes**. In addition, it contains a **Fixed calculated field** as the purpose of the graph is to compare urgent open cases to all open cases over time. To do that, the Fixed functions will prevent the first line to be filtered by severity, only by creation date. The difference between the two lines show that the share of urgent open cases is stable over time. I included the max/min values which turn out to be the first and last day values to highlight this relationship.

Figure 4 is a simple stacked bar chart showing the distribution of open cases by service type. In addition, it contains another **action** which takes the viewer to Dashboard 3 where they can see the severity break down of the selected service type. I included this graph in a separate dashboard as this is a very specific connection with not as much impact as some of the other figures. The dashboard 3 also allows me to add a navigation button.

Lastly, **Figure 5** shows an interesting data, the amount cases that are waiting for response by which group they are sitting with. In this graph particularly you can see the effect of filtering by case type, the ranking of the 5 groups completely changes by the two categories.

This concludes the second dashboard and the visuals in the workbook.

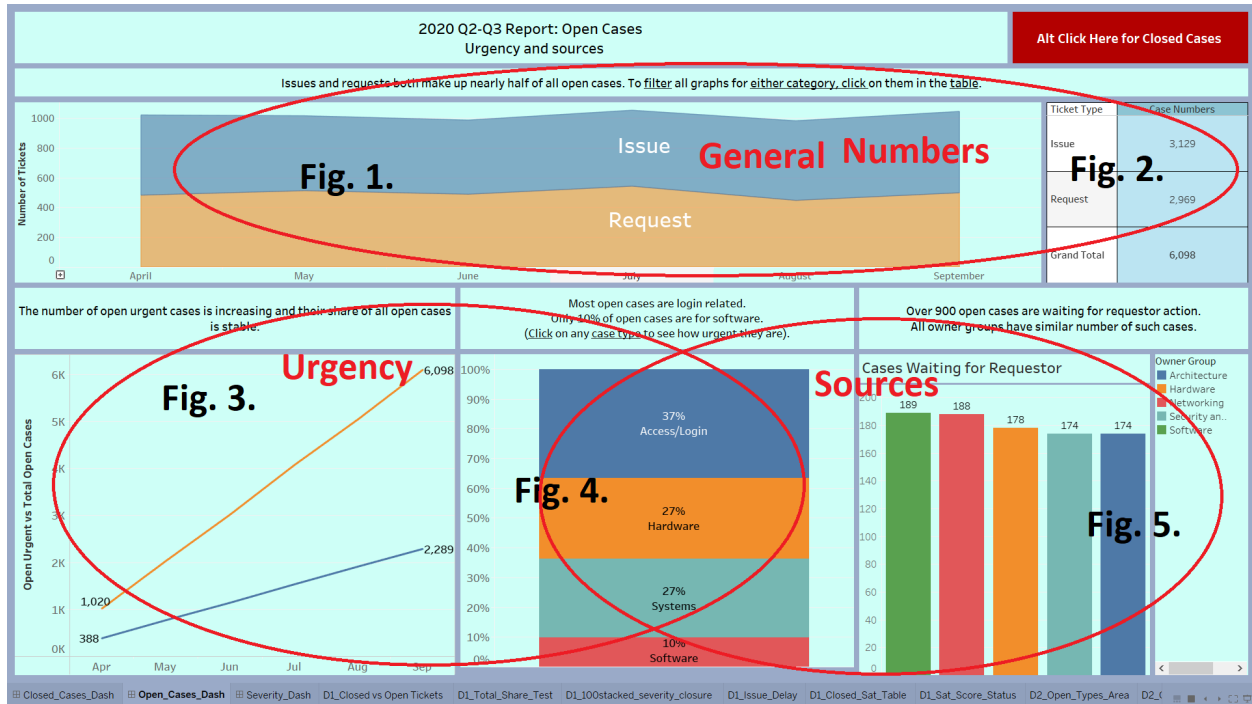


Figure 9: Dashboard 2 Overview

Tickets Created Monthly

{ FIXED [Open or Closed Status] , DATEPART('month', [Created Date]) : COUNT([Help Desk]) }

The calculation is valid.

3 Dependencies ▾

Apply OK

Figure 10: Fixed Calculated Field to avoid filtering

Edit Filter Action

Name: Severity_of_Issues

Source Sheets

Help Desk (Help Desk)

- ☐ D2_Closed_Total_Tickets_Stacked
- ☐ D2_Open_By_Group
- ☒ D2_Open_ByType
- ☐ D2_Open_Type_Table
- ☐ D2_Open_Types_Area

Run action on:

☒ Single-select only

Target Sheets

Severity_Dash

- ☒ D3_Severity_Issue_Cat

Clearing the selection will:

☒ Keep filtered values

☐ Show all values

☐ Exclude all values

Target Filters

☐ Selected Fields ☒ All Fields

Source Field	Target Field	Target Data Source

Add Filter... Edit... Remove

OK Cancel

Figure 11: Action setup

Conclusion

In conclusion, the answers to my rearch questions are:

- Cases were closed on average in 19 days with significant difference in process types.
- More severe cases were generally not treated more urgently.
- Most customers were satisfied and processing time did not heavily impact their ratings.
- Severe cases are often left in open sate which may require further investigation.
- Most open cases are login related across all working groups.