Step Aside, Tableau: The Pros and Cons of Analyzing and Reporting Ithaka S+R Survey Results Using Google Data Studio

Emily Guhde
Georgetown University, USA

Introduction

In the fall of 2017, Georgetown University Library conducted the Ithaka S+R survey on its Main, School of Continuing Studies, and Qatar campuses for undergraduate students, graduate and professional students, and faculty. Each user group received a customized survey instrument which included between 20 and 50 questions, some of which had multiple sub-questions. Almost all questions gathered categorical or ordinal data; very few questions asked for free-text responses. The categorical and ordinal data could be easily aggregated for analysis of trends across various subgroups in a spreadsheet program (like Excel), but the sheer quantity of data available would have resulted in scores of tabs and graphs. With nearly 2,000 responses to analyze and communicate to stakeholders, researchers faced the challenge of finding a presentation tool that would allow them to share the data dynamically and encourage exploration, but without overwhelming colleagues at different data visualization skill levels.

For this task, many libraries have chosen Tableau, one of the leaders in the field of analytics and business intelligence. While Tableau’s menu of data visualization options was impressive, its steep price tag and blog-like layout were less appealing, and we decided to consider other options. As a Google campus, Georgetown University relies upon the Google Apps Suite for email, collaboration tools, and synchronized cloud and desktop file storage. Our discovery of Google Data Studio’s beta version (now fully integrated as part of the Google Apps Suite) provided us with an opportunity to use a free tool that was already seamlessly integrated with our workflows, security, and file structure. However, it also presented challenges worth sharing to help others benefit from our experience.

To Avoid Drowning in Survey Data, We Identified Our Priorities

A fresh dataset is enticing to a data analyst, with all of its possibilities for data visualizations and potential findings. While it can be tempting to exhaustively explore a dataset, with a deadline for the completed report already looming in the distance, decisions about where to focus and how to present the data need to be made quickly. Our research team chose to wait until we had the raw data back from Ithaka S+R before deciding how to move forward with the analysis. Once we understood how the data were formatted, coded, and described, it was easier to give our analysis a clear direction by focusing on our previously-identified priorities. Below is a list of priorities, including our criteria for analyzing and reporting, that we considered before we began our analysis.

Priority 1: Use the dataset to actively engage with stakeholders.

With three campuses and six libraries, we needed to find ways to present slices of the data that would be relevant to each of our various stakeholders. The Ithaka S+R Survey was the largest survey conducted in over a decade by Georgetown University Library, and we wanted to ensure that the data were used vigorously before their shelf life expired. Although the survey topics were broad and the findings were relevant to the vast majority of library staff, we would have to compete for attention in a year that also saw a new dean of the library and long-awaited transition to a new ILS. To actively engage with stakeholders through the dataset, we looked for a tool that would allow us to slice the data dynamically and zoom in to study subsets of our respondents or focus on specific topics.
Priority 2: Provide data visualizations alongside analysis and context.

Early feedback from stakeholders suggested that it was meaningful to explore the dataset without analysis notes or summarized findings. In other words, our stakeholders wanted to be able to draw their own conclusions as they looked at our graphs. On the other hand, as researchers, we felt that it was important to provide analytical observations about the data, drawing attention to important trends and findings in the descriptive statistics. To achieve both views of the data, we wanted a tool that would give us a layered approach so that at first our audience would see just the data, and then they could choose to see our analysis.

Priority 3: Make it appealing to view and easy to navigate results without a start-to-finish linear requirement.

We wanted to benefit from working with the library’s communications and marketing coordinator to make our results visually appealing, with colors and logos to suggest familiarity and a cohesive package of information that looked like it was an official report of the Georgetown University Library. To get the look we were going for, we needed a tool that gave us control over the design of the digital page so that we could create a clean and simple look no matter how complicated the data. We wanted something that would feel browsable, interesting to flip through and explore, and able to stand on its own without a lengthy introduction about the technical aspects of using the tool. If someone had to be in the room with the viewer to explain how to explore the data, that was a deal-breaker. We also wanted to avoid tools that required viewers to go through every page before moving on to the next, since topics included in the survey would have varying levels of appeal to different stakeholders.

Priority 4: Use tools that are already available.

Initially, Georgetown University Library had planned to conduct the Ithaka S+R Survey in the spring of 2018. But when several administrative details aligned in support of a fall 2017 launch, we moved up our timeline by six months and committed to working with a four-week pre-launch timeline. Unsurprisingly, we were not able to spend much time considering our analysis needs prior to launch. Because we did not plan in advance, we had not included a budget proposal for any new analysis software. It might have been possible to get a few more Tableau licenses for the whole project team, but not without sacrificing other priorities for the division. By completing the analysis with tools that we already had available on all of our computers, we would save valuable time and money.

Priority 5: Protect data and share selectively.

The library shares reports and data internally via its Staff Wiki (intranet), but we do not typically share our data and reports publicly through our website. While we planned to share our Ithaka S+R Survey results more broadly than most of our previous assessment projects, we knew that we did not want to make the data publicly available. To maintain transparency and openness about the results among staff while maintaining them securely, we would need a tool that was password-protected and that would not reveal the entirety of the underlying dataset to viewers.

Google Data Studio, Explained

Google Data Studio (GDS) is a data visualization and dashboard-building tool that uses data connectors as an active pipeline between a dataset and the interactive front-end interface. The dataset can live on the web, in a Google Sheet, or as a .csv file. As the data in the background update, GDS reflects the changes. Data can be cleaned and formatted using functions within the GDS editor’s interface, or data can be prepared ahead of time in Excel or Google Sheets. Permissions, both for collaborative editing and sharing with viewers, work similarly to other tools in the Google Apps Suite. When we were considering GDS for the Ithaka S+R Survey project, it was still in beta, but in September 2018, Google moved GDS out of its beta phase, and it continues to roll out updates to the editing and viewing interfaces.
The Benefits of Using Google Data Studio to Meet our Priorities

While we were somewhat concerned that attempting to use a product in beta for analysis and reporting would backfire, it seemed to be the best fit given our priorities, and we were curious enough about Google Data Studio to give it a try. We worked with the data in GDS for about three weeks, then requested some feedback from our Assessment Steering Committee. Their positive feedback gave us the green light to move forward with analyzing the whole dataset in GDS. Overall, our experience was mostly positive and we were able to meet our priorities for analysis and reporting. Below is a summary of the benefits of GDS that the research team found most appealing.

GDS offers customizable features for presenting basic data visualizations, with multiple filters to isolate data within subgroups. For the Ithaka S+R Survey results, we created filters for “school,” “primary library” (i.e., the library location that respondents indicated that they visited most often), and “gender.” This tailored view of the data allows users to actively engage with the dataset, deciding whether to zoom in tightly on a specific subgroup or zoom out to get a broad view of the data. Instead of making separate static reports for each potential group of stakeholders, we created filter buttons to allow users to choose their own view of the data. It is worth noting that there is a learning curve for structuring the dataset appropriately for the filters to work, but within GDS there are examples to follow. We also used scorecards to display dynamic n counts, which update automatically as filters are applied or removed. Image 1 shows how these features are combined on a typical report page.

Image 1: Sample GDS graph from the Ithaka S+R Survey

Features include customized formatting, multiple filters, automatically updating n count, tooltips, data download option, and links to other pages.
The layout of Google Data Studio is like a slide deck with a unique URL for each slide. For some of our users, flipping through the slide deck to make their own observations was their preference. But for users who wanted a prepared analysis, we included a link on most slides that would pop-out an appendix slide with an additional view of the data, as well as a few digital “sticky notes” highlighting our observations about the results to the question. By layering the analysis behind the results, GDS gave us a way to present the analysis and its context alongside robust views of the data. Image 2 shows a close-up of a sticky note from one of our report’s appendices.

Image 2: Detail of GDS sticky note highlighting an observation from the data

While sticky notes are not a standard “shape” for GDS, a simple image with a CC0 license can be easily imported and then layered behind a text box.

While extensive training was not necessary for users to explore the dataset on their own, the library assessment team held a one-hour, hands-on training session for the library’s Leadership Council, demonstrating how to explore the data and use filters to make observations within subgroups. We also hosted several trainings open to all library staff, though these were only lightly attended. Thanks to feedback from staff members, we added more navigational details to the report, such as question numbers and overlap indicators within the tables of contents, to facilitate comparisons within topics and across surveys.
Many of the questions about library spaces were the same (or similar) on all three surveys. In this view of the Graduate and Professional Student Survey table of contents, we are directing users to also explore the Undergraduate Student [u] and Faculty [f] Surveys. The original survey question numbers are included to help viewers who may be interested in looking up the exact phrasing of a question.

Google Data Studio allowed us to use the visual identity of Georgetown University to customize the look and feel of the report. Banners, logos, and even the colors of the bars on our bar graph were easily customized with Hex codes. To help with navigation, we used headers to indicate the name of the survey instrument and the section and topic of the survey. Viewers can browse, flip, and explore as if they were looking at a digital magazine. Viewers can also use the table of contents to select which topic they want to jump to. At the bottom of every page, we included a link to the table of contents, as well as a link to advance them forward or backward within the slide deck. We connected all three survey instruments to one digital “cover page” (see Image 4) so that all three sets of results were available via one link.
The cover page was set up as a separate report in GDS with just one page of content. This page linked out to the three sets of survey results, each contained within its separate GDS report.

The availability of Google Data Studio as part of the Google Apps Suite, even though it was in beta throughout our analysis and reporting phase, was of enormous benefit to us. Because it is a Google product and we are a Google campus, it fit in well with the systems we were already using to analyze and store our data. Viewers had seamless access to the reports without downloading any new software or setting up an account with a third party. It was available on all of our machines via web browser at no any additional cost. It also meant that collaboration was easy to coordinate because, like many Google Apps, GDS allows for multiple report editors. Like other Google Apps tools that we use, GDS report access is controlled using individual email addresses or Google Groups. We could also grant temporary access to people outside of the organization for demonstration purposes. GDS does allow report viewers to download data from its interface, but the data are downloaded at the aggregated level and are not linked back to individual respondents in the underlying dataset. See Image 5 for a view of the permissions interface.
The options to share with collaborative editors and viewers are very similar to other Google Apps, but there is not currently a way to give permission for anyone to add comments within GDS.

To ease our minds about using a product still in beta, we discussed an exit strategy that we would use if Google decided to shut down GDS while we were in the middle of developing the report. Fortunately, because the underlying dataset lives in a Google Sheet and feeds to GDS, we knew that most of our analysis work would be preserved if we needed to leave GDS behind. We also learned that GDS exports nicely to a PDF (see Image 6) without additional formatting, and while we did not design the report with a print version in mind, it was nice to be able to provide a printed PDF copy as part of the welcome packet for our new dean of the library.
The export-to-PDF feature is relatively new to GDS. While a big time-saver (previously, print-to-PDF had to be executed on each page separately), we have noticed occasional glitches with the display of data visualizations on the exported PDF.

Room for Improvement
As a tool for actively engaging with stakeholders, we expected that viewers would explore the dataset on their own. Unfortunately, in a couple of cases, we were alerted by colleagues that a graph (or a series of graphs) was unexpectedly broken. Because GDS was in beta at the time, frequent and unannounced updates would occasionally cause our graphs to break because of updated data formatting restrictions. Now that GDS has moved into full production mode, we have not experienced any breakage in months. See Image 7 for a view of a broken GDS graph.
Seeing an error message is never a good sign, but we quickly learned that fixing our graphs was much easier when we were consistent with how we formatted our data fields behind the scenes in GDS. Once we identified the solution for fixing one graph, we could easily apply that approach to all broken graphs.

The data visualization filters have behaved predictably and without glitches, but one disadvantage of the page-level filter is that, if you are interested in seeing all results for one subgroup throughout the entire report, you have to select that filter for every data visualization separately. If the filters could be placed at a higher level and then applied across the report, it would be a big time-saver for report viewers.

Having the appendices pop-out to provide additional context and analysis was a major benefit of GDS. But with multiple GDS tabs open at one time, the processing speed often slows down, and a viewer might wait 10–30 seconds for the graph to appear in a separate tab. This delay was frustrating when editing the graphs and page layout. Completing a series of quick updates across all pages on the report seemed to take more time than it was worth for minor adjustments.

In terms of formatting options, font selection has been improved recently, but the options for formatting within the data visualizations are still rather limited. For instance, the graph legend cannot be stretched across the entire length of the graph canvas; instead, the legend awkwardly begins wrapping at the Y-axis. It would also be beneficial to have more shapes available, although inserting an image of any shape from an external source and then sizing it to fit your needs works well in most instances.

For collaboration, Google Data Studio only allows the report owner to connect and reconnect a data source. Any user with editing privileges can access the underlying dataset (if it has been shared with them), and update the content in the data source, but if the structure of the data source changes, only the owner of the report can refresh the fields within GDS. This was inconvenient for our team, because it meant that we had to wait for the graph’s “owner” to refresh the dataset. However, we established a workflow to manage the inconvenience.
Likely Future Use of Google Data Studio at Georgetown University Library

A year after we began working with Google Data Studio, we have continued to find additional reasons to keep using it.

We are using GDS for smaller assessment project reports on a regular basis. By using filters, we can quickly answer questions about how a subset of our participants responded to a particular question. We can add the analysis right alongside the descriptive statistics and include recommendations for future actions in the final page. After completing hundreds of graphs and tables for the Ithaka S+R Survey results, every other report feels like a snap. In summer 2019, Library Assessment plans to support the library’s culture of assessment by offering its first Google Data Studio training for library staff.

As for the Ithaka S+R Survey Results, our Assessment Steering Committee is working through the data with a deadline of April 2019 to identify actionable findings and prepare a list of recommendations. We have been able to complete our work efficiently in Google Data Studio, and in general, we think that the risk of exploring this new tool for the Ithaka S+R Survey analysis and reporting has paid off.

—Copyright 2019 Emily Guhde