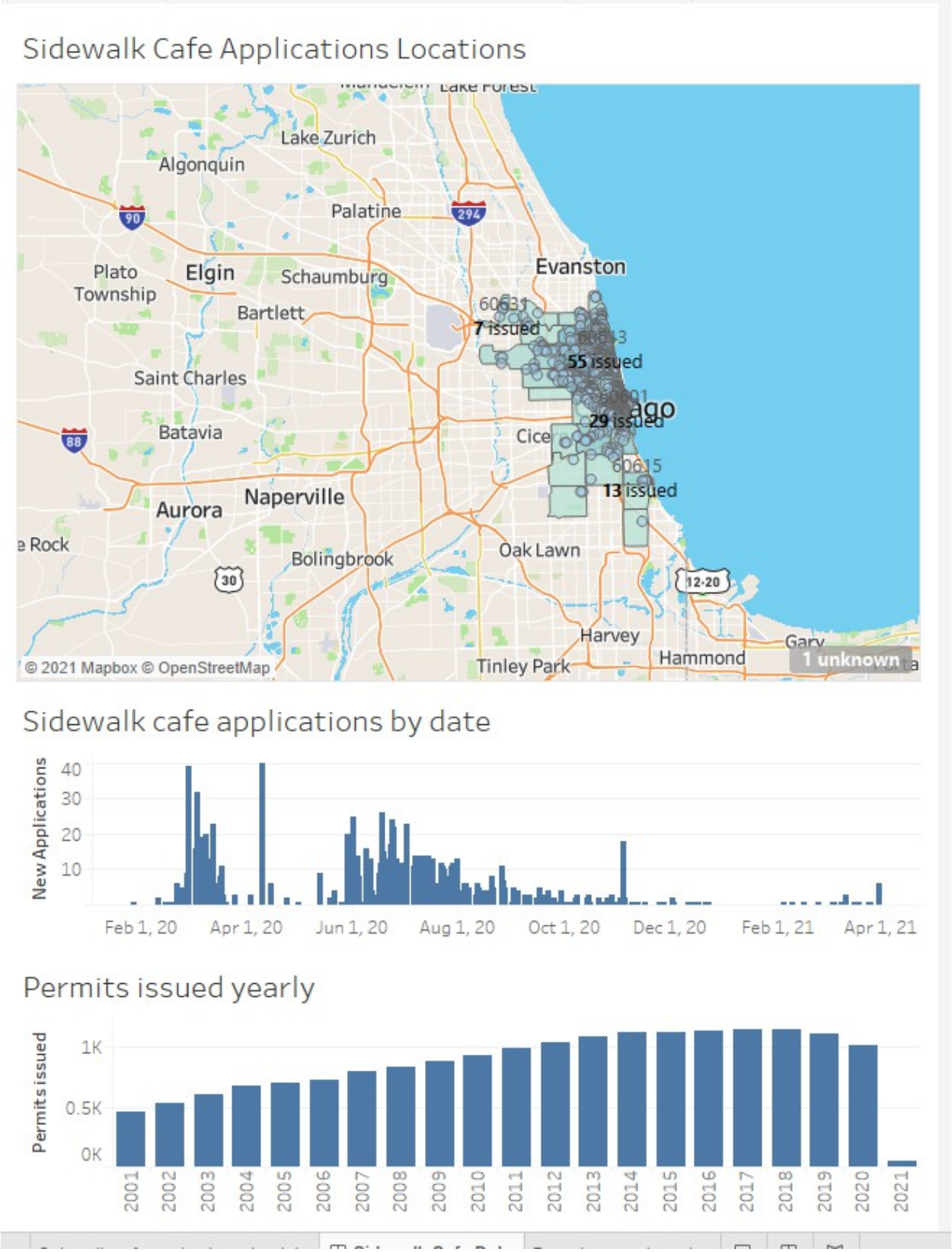
Jason Grey
LIS 883 Final Project

Comparing Power BI and Tableau for Data Visualization

Tableau and Microsoft Power BI are two of the leading data visualization products on the market. To compare the capabilities and attributes, I took two datasets and used both products to create a similar visualization from these datasets.

The first dataset I chose to work with is “Sidewalk Cafe Permits - Current”, which lists all of the businesses who have requested sidewalk cafe permits from the city, starting in January of 2020. My assumption was that sidewalk cafe applications would have increased dramatically as a result of the pandemic restrictions, but it turns out this assumption was incorrect.. I started with Tableau, and built this dashboard with it:

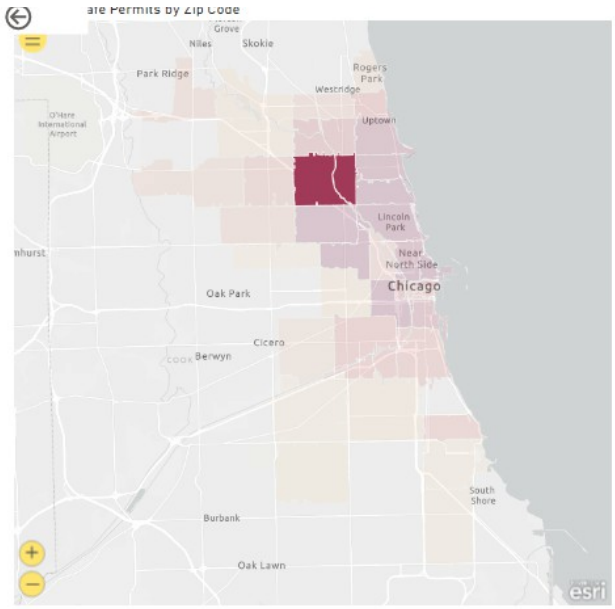


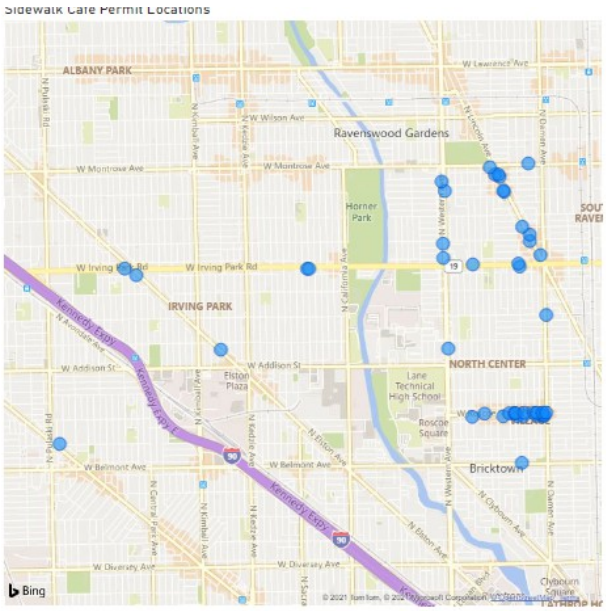
The live version of the dashboard is here:

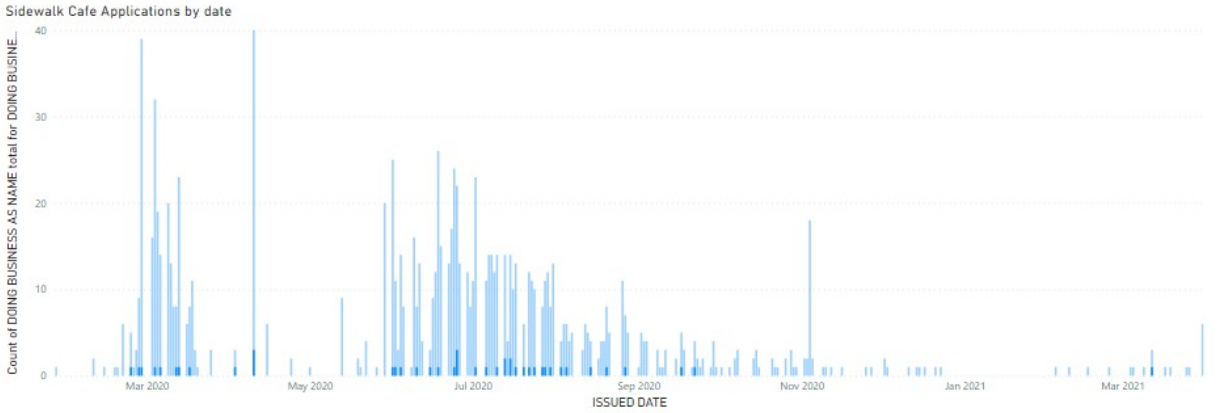
https://public.tableau.com/profile/jason.grey#!/vizhome/ChicagoSidewalkCafeApplications/SidewalkCafeData

I started with the map - working with maps in Tableau isn't super-obvious; Tableau makes a lot of decisions for you based on the data you give it, and to generate a map you set your rows Latitude data field and your columns to a Longitude data field, which defines the size and shape of your map based on the minimum and maximum values. Not especially intuitive, but the documentation for Tableau is easy to navigate and understand. Other geographic information like zip codes can then be dropped in place as data points ("marks" in Tableau's jargon). Editing the appearance of marks and tooltips is pretty easy and flexible in Tableau. Creating the date range chart was also very simple and intuitive. The thing that took me the longest to figure out was how to put multiple datasets on the same map; eventually I discovered how to make a "dual-axis chart", so I could add the color-coded Zip codes and the dots representing locations on the same map. (I think you are limited to two, though, which is unfortunate.) My main frustration with this report is not being able to control or disable the click-filtering; for instance when you focus on a certain zip code, the location dots disappear, which I didn't want but can't seem to avoid. I also couldn't make filtering work between the two charts, perhaps because they don't use exactly the same set of data.

I then tried to do the same chart in Power BI :







https://app.powerbi.com/groups/me/reports/749c264e-94d3-4b4a-8a76-7379006ab30b?ctid=e363050e-fa18-48f7-aefc-7db1230b452a

The main disappointment here was not being able to combine both datasets on the same map, at least not easily. I used the Arcgis Maps for Power BI visualization for the Zip Code map, because it had the ability to fill in Zip Code areas, and the Maps visualization for the second one with the locations. To combine both into one map, I think I would have had to find or create a custom file with the shapes of the Zip code areas to use as data in the visualization - so not impossible, but much less easy then with Tableau. I couldn't figure out how to add labels to the color-coded zip code areas in ArcGis like I did in Tableau. The ArcGis Map visualization also required a second level of authentication, which makes sharing the map more complicated - with just the basic non-organizational account that I signed up for at ArcGis, I was unable to make a map that could be publicly shared with anyone as an embedded iframe. I also couldn't figure out how to change the color palette in the ArcGis map.

One thing that did work better with Power BI was the filtering interaction between the visualizations, that happened without any effort on my part and made up for not being able to combine the two datasets into the same map.

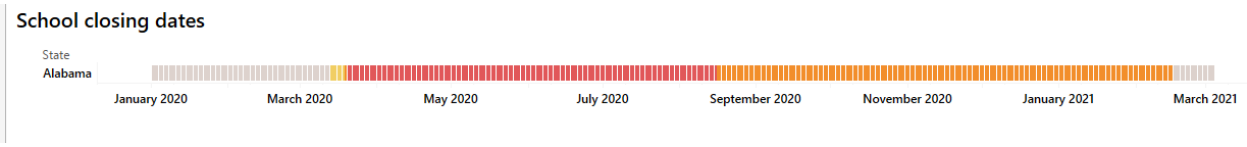
Creating and styling the tooltips was more complicated in Power BI, but also more flexible; to create anything more than a very simple tooltip you can create a set of visualizations (a 'report') specifically to use as a tooltip in your main report. (In my sidewalk cafe report, the ArcGis map on the left shows the default tooltip, the location map on the right shows a custom tooltip.)

My overall impression based on this first comparison was that Power BI has depth and can do a lot of things, but has less out-of-the-box user friendliness compared to Tableau, and is both more structured and more rigid.

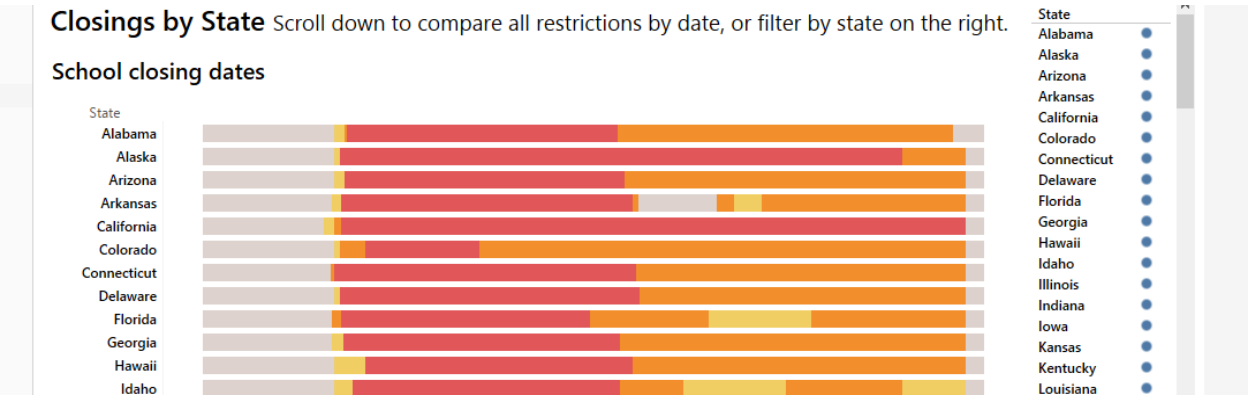
The second dataset that I decided to work with was from Google Bigquery, although the data itself was compiled by the Blavatnik School of Government at Oxford University. The dataset is described here: :https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker

The data tracks the policy responses of various governments to the Covid-19 pandemic, in particular the dates that certain policies such as restrictions on public activities were implemented. Quantifying this data requires making certain simplifications; each category of public restriction is divided into 3-4 levels of strictness (recommended vs. required actions, for instance) and coded to a scale. For my project, I used the data from Bigquery rather than the more up-to-date information from the Github repo, for convenience. I used a subset of the available data set that included information for the US only, categorized by state, and six policy actions that seemed most meaningful.

I wasn't completely sure what the best way to visualize this data was, so I began by importing the data into Tableau and playing around with it to see what kind of charts it would build by default. Using date as the column axis was an obvious choice, but what to use for dimensions and how to filter and organize the data was less so. By randomly trying out different options, I ended up with something like this:



Using the region_name field (the individual states) as both a "row" value and a filter, and representing the ordinal school_closing value as a color, created the event-driven timeline graphic I was looking for. Putting detailed information in the tooltip made the visualization fairly self-explanatory, and the compact graphic made it easy to show comparison:

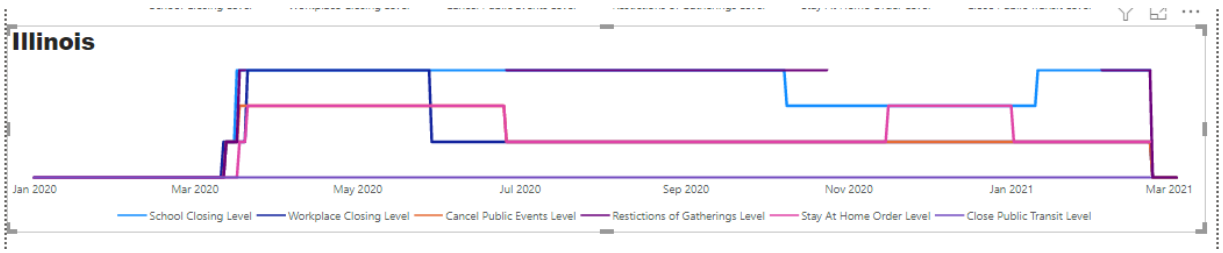


I arranged this basic style of presentation in a couple of different ways, to visually compare different dimensions:
https://public.tableau.com/profile/jason.grey#!/vizhome/CovidClosingsbyState/CovidClosingsStory

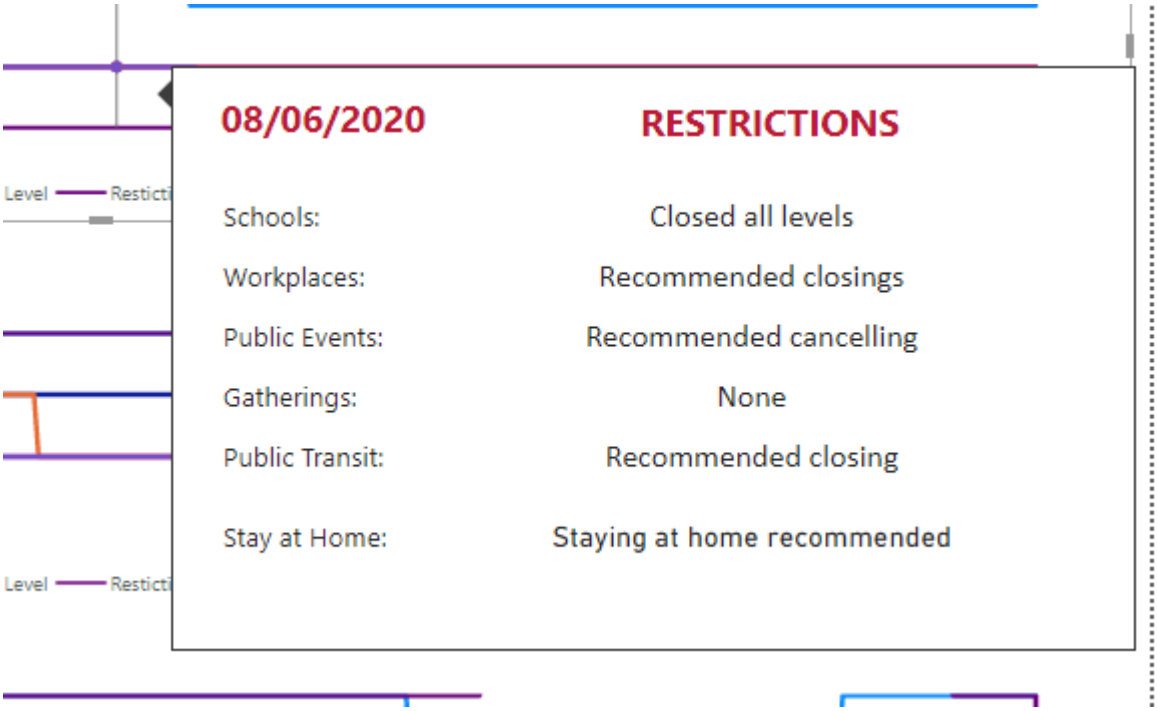
Figuring out how to set up the filtering was probably the part that took me the most effort - understanding some of Tableau's core concepts, like the difference between "Measures" and "Dimensions", reduces the amount of frustrated clicking that you end up doing considerably. While the basic logic of of Tableau is fairly straightforward, the UI packs a lot of features into a compact workspace and it can be hard to figure out how to do things without looking them up. Looking them up requires knowing what the correct term is for the feature you are trying to work with, which also rewards a careful reading of the documentation.

The hierarchy of Worksheet->Dashboard->Story that Tableau uses to organize content is immediately familiar to anyone who has worked with content management on the web, and lets you move back and forth between detail and big-picture thinking about the presentation. The layout features are a bit awkward; I spent a considerable amount of time on tasks like setting heights and widths that seemed like they could have had a better set of default behaviors. There's no option to set content frames to percentages rather than fixed sizes, something that was particularly annoying when building a "small multiple" style of presentation like in the second tab of my story.

My Power BI version of this data ended up looking very different; I was unable to find a visualization widget in Power BI that would display the kind of color-coded timeline that the Tableau chart uses. Of the standard set of widgets that can be added to Power BI, a Gantt chart is the most similar, but after downloading the three Gantt chart vizes available in Power Bi, none of them were able to properly interpret the dataset; they were all set up to display date ranges rather than information about specific dates. I ended up going with the basic Line Chart viz:

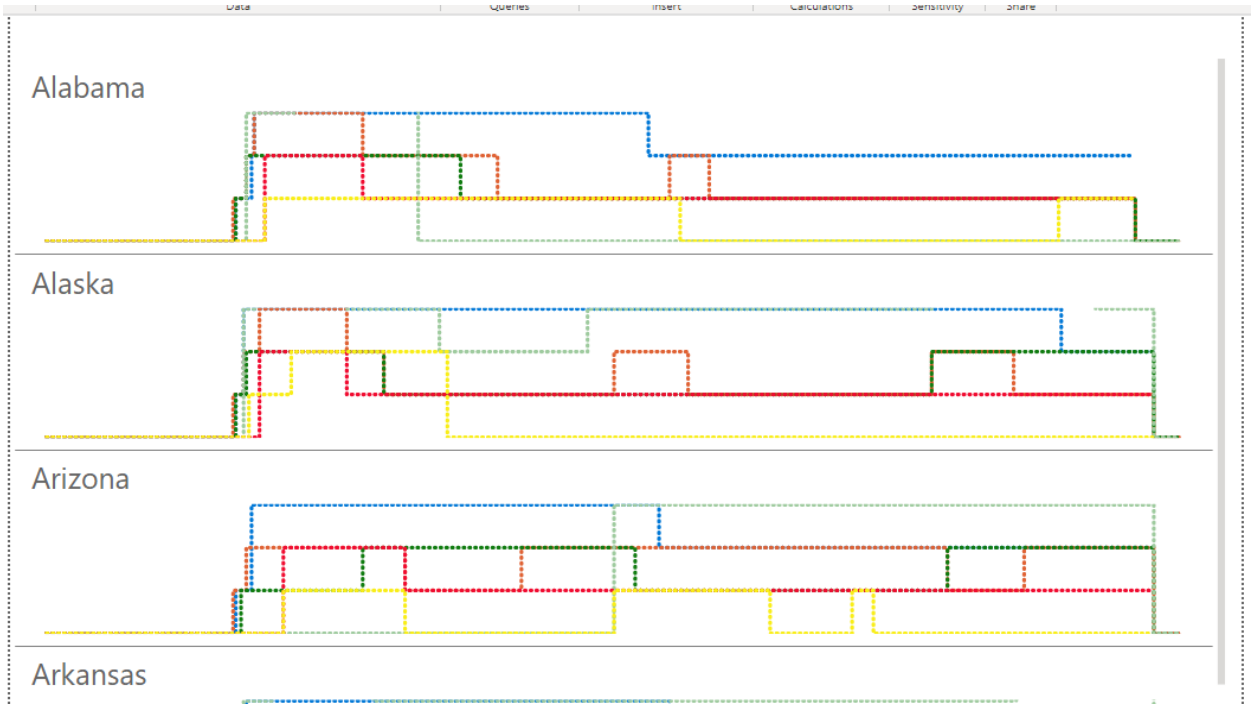


which is an even more compact display of the six timelines for each state, but packs so much in that a lot of detail is lost. Making the lines dotted and setting the colors scheme to highly contrasting colors helped with this problem a little bit, but it's still not ideal. Once again, the contextual information explaining the meaning of the levels is placed in a tooltip:



To get the descriptions in the tooltips, I had to add what Bi calls "calculated columns" to my datasource, which basically involves writing a formula to say "If this row is X, put Y in the new row". Creating an alias like this seems to me like something that logically belongs in the presentation layer, not the data, so I found this a questionable product design decision. I was also unable to create aliases for things like data labels, like changing "date_region" to "State".

Aside from not being able to get a visual style that was appealing, the main problem I encountered was again that the filtering was difficult, but in BI I encountered some problems that I was not able to solve. Dynamic filtering that is embedded in the visualization itself, so that an end user who is not a "report editor" can compare individual elements dimensions, exists but is pretty simplistic and poorly integrated with other elements. The Line Chart gives the option to split data into "small multiples" by a field, in the case of my data the obvious one being state/region_name:



But neither the dynamic filtering nor the custom Tooltip integrated with the small multiple display, so it became useless. The tooltip aggregates all values that are displayed rather than the specific values for the multiple it is hovering over. The alternative to using small multiples in the viz was to copy the viz 50 times and apply a filter to each one - tedious for sure, and also not quite adequate, because filtering individual visualizations out of the report doesn't collapse the space between them in the display, so the intended purpose of providing a useful visual comparison between two different states doesn't work.

So here is the final product (it should be viewable by ygao@dom.edu):
https://app.powerbi.com/groups/me/reports/2fa2158e-baf3-4ab1-a403-0b3b69cdf50c/ReportSection
This is a much less visually intuitive and effective visualization, and lacks the comparative power of the one built in Tableau.

CONCLUSIONS

tableau public	Microsoft | Power BI
Greater design flexibility, easy to create with	Better source data manipulation capabilities
Better sharing features, especially for the open web.	Integration with other Microsoft products, particularly Excel
Better looking, more modern design	Less flexibility with design, mostly relies on prepackaged visualizations made for a specific purpose.
Less advanced data analysis features	A bit of a resource hog; runs kind of slow on the average computer
Some UI elements are unintuitive, and some features are “buried” and hard to find.	Complicated to share outside of the user’s organization.

I found Tableau to be a lot more fun and creative to work with, but if I had a really complex data set that I needed to work with, the Power Query features of Power BI would be essential. In general, Tableau is more of a Data Visualization Builder, and Power BI is more of a Data Manipulation and Display Tool.