

IS 590 DW SP20
Data Warehousing and BI

Final Project Report

**Comparative Study of Business
Intelligence Tools in
Visualizing COVID-19 US Data**

By

Gnanendra Reddy Tugu Yagama Reddy

Samuel John

Contents

1. Introduction	2
2. Tools Used	2
3. The Data	2
4. Importing the Data into the Tools.....	3
4.1 Tableau Dataset	3
4.2 Power BI Dataset.....	3
5. Insights and Analytics	4
5.1 Analytics in Tableau.....	5
5.2 Analytics in Power BI	6
6. Understanding the Dashboards	7
7. Sharing the Dashboards	8
7.1 Sharing the Dashboards using Tableau	8
7.2 Sharing the Dashboards using Power BI.....	8
8. Comparing the Tools	8
9. Conclusion	10
10. References.....	11

1. Introduction

In today's world, Business Intelligence, or BI as it is better known, is a keyword in the Industry. Business Intelligence deals with the leveraging of tools and techniques to gain insightful insights from raw data by means of mining the data, visualizing the data, creating infrastructure to manage the data and providing guidelines to businesses as to how this information can be used to make data driven decisions. ^[5]

BI tools such as Tableau provide the means to convert data into meaningful visualizations and dashboards that convey important information about the data in a way that is quick and easy to understand.

For this project, we wanted to perform a comparative study on the visualization features and ease of use of two of the most commonly used BI tools – Tableau by Tableau Software (acquired by Salesforce in 2019), and Power BI by Microsoft.

2. Tools Used

We used two BI tools for this project – Tableau Desktop and Power BI. Tableau Desktop is available as a one year licensed software to students. Power BI is available as Power BI Free, Power BI Pro, and Power BI Premium. Power BI Free is included in all Office 365 Plans, and you can sign up for Power BI Free any time you want, and we are using the Power BI free version for our analysis.

3. The Data

To properly compare the features of Tableau and Power BI, we thought it would only be fair to use data that has not been overly used before. The world is currently dealing with the spread of a pandemic named “Coronavirus Disease 2019” (abbreviated “COVID-19”), caused by the virus named “SARS-CoV-2” ^[1]. The propagation of this virus has been documented by State bodies across the world that share the numbers with the public. The Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) has been collecting this data and providing it to the public for use in analysis and research ^[2].

Initially, we looked at obtaining the data directly from the JHU COVID-19 repository on GitHub. However, both Tableau and Power BI announced the availability of datasets that have been cleaned and targeted towards analysis in their respective tools and we decided to use those datasets for this project.

The Tableau COVID-19 case data was accessible through a Web Data Connector from data.world where the data is refreshed every evening at 10 p.m. PT.

https://tableau.data.world/?dataset_name=covid-19-data-resource-hub%2F%2Fcovid-19-case-counts&query=SELECT%20*%0AFROM%20covid_19_cases&queryType=SQL

The Power BI dataset was available as online Web Data Connector from USA Facts which aggregates data from the Centers for Disease Control and Prevention (CDC), state- and local-level public health agencies and the data is updated daily at 10:30 PM Pacific Time.

https://usafactsstatic.blob.core.windows.net/public/data/covid-19/covid_confirmed_usafacts.csv

Since there are already a couple of websites that are providing data on a worldwide basis, we decided to focus our analysis on the propagation of cases within the US.

4. Importing the Data into the Tools

4.1 Tableau Dataset

Tableau provides an option to connect to an external data source using a Web Data Connector and the URL of the data source. Once the authentication is complete, Tableau imports the data from the URL and creates extracts that can be used to analyse the data.

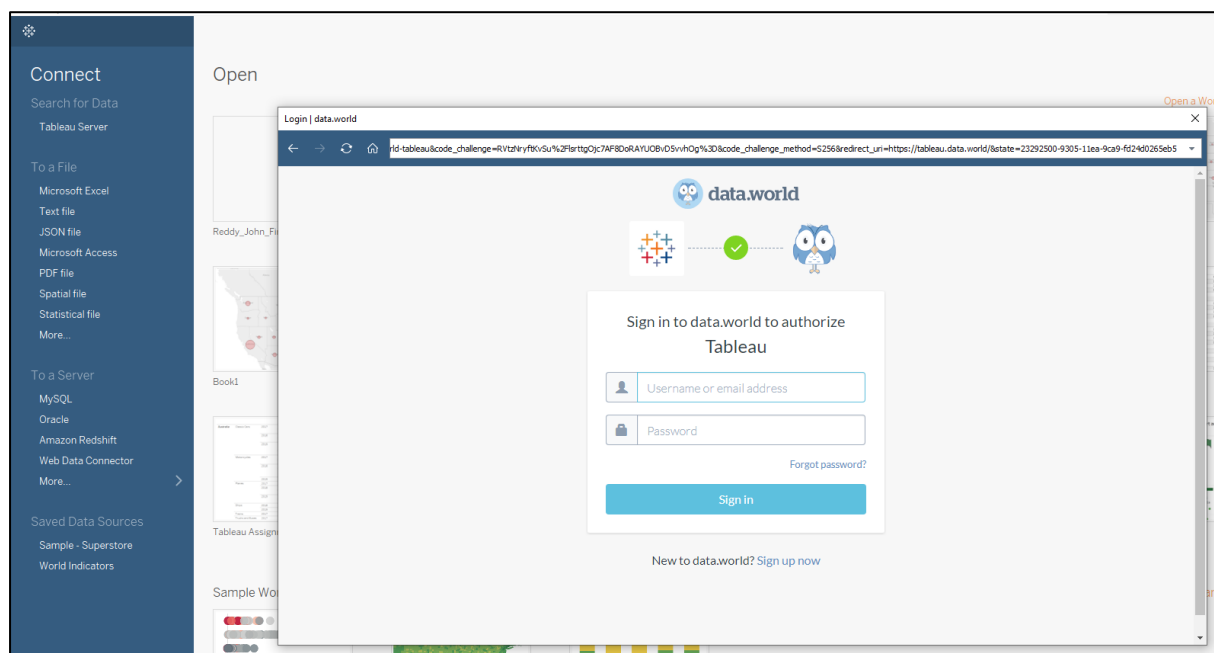


Fig 4.1.1 – Connecting to the Tableau Data Source using a Web Data Connector

4.2 Power BI Dataset

Power BI provided an option to connect to an external data source using a Web URL of the data source. Once the authentication was complete, Power BI imported the data from the URL and created extracts that could be used to analyse the data.

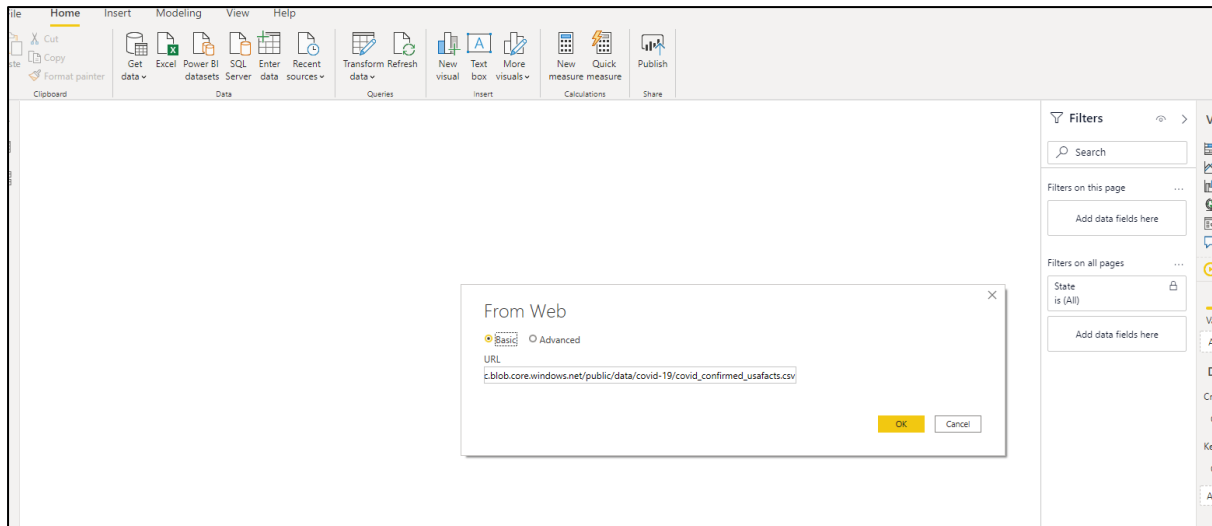


Fig 4.2.1 – Connecting to the Power BI data source using a web URL

We could also edit and transform the data before loading the data according to our needs.

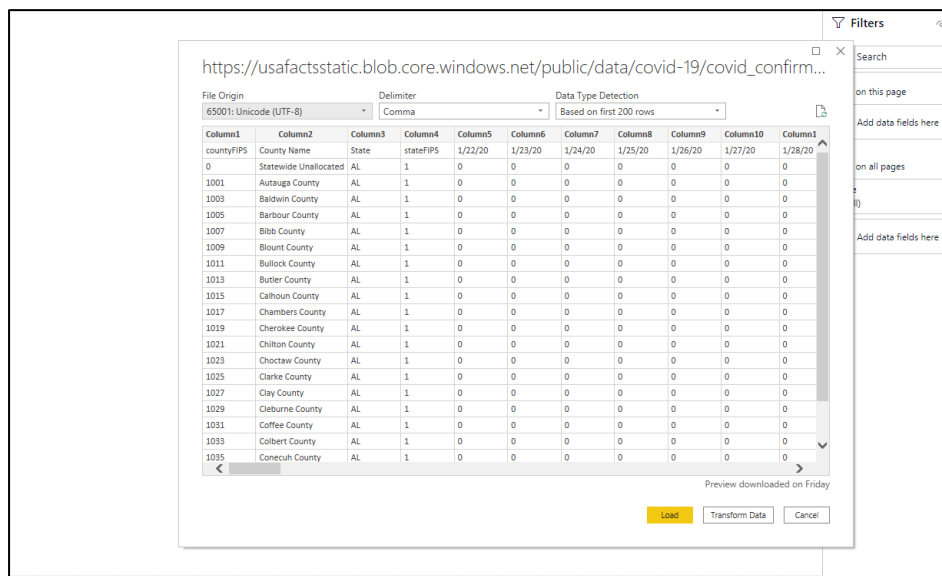


Fig 4.2.2 – Power BI displaying the data to be loaded with options to edit and transform the data

5. Insights and Analytics

We decided to create dashboards in both the BI Tools that were as similar to each other as possible. This way, we could see how creating the same visualizations in each tool differed in terms of ease of conversion of the data into each visualization and the grain that was available for analysis in each tool. Further, it let us compare the visual impact the visualizations had in terms of how professional they looked and how easy it was to embed the graphs and charts created into dashboards which were the final analytical objects we aimed to produce from each tool.

5.1 Analytics in Tableau

Tableau did a good job when it came to converting the raw data into individual visualization objects. It automatically converted fields such as dates into multiple level of grain such as Year, Month and Date. This made it possible to filter out data by date which was important in this case since we wanted to convey the latest information on the number of cases. It also provided grain based on Country and County data which was further represented by automatically generated latitude and longitudes which could be used to map the data onto world maps. Using these features, we were able to build a dashboard as seen in Fig 5.1.1 and Fig 5.1.2

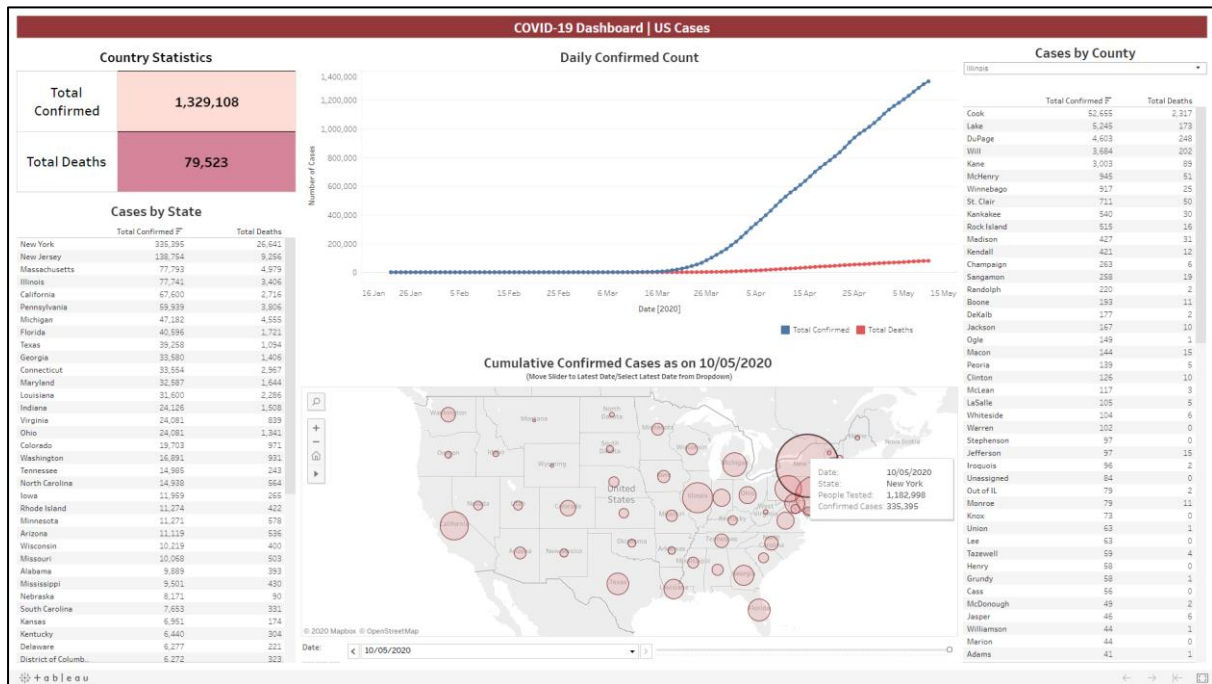


Fig 5.1.1 – Tableau Dashboard for COVID-19 Cases in the US (PC view)

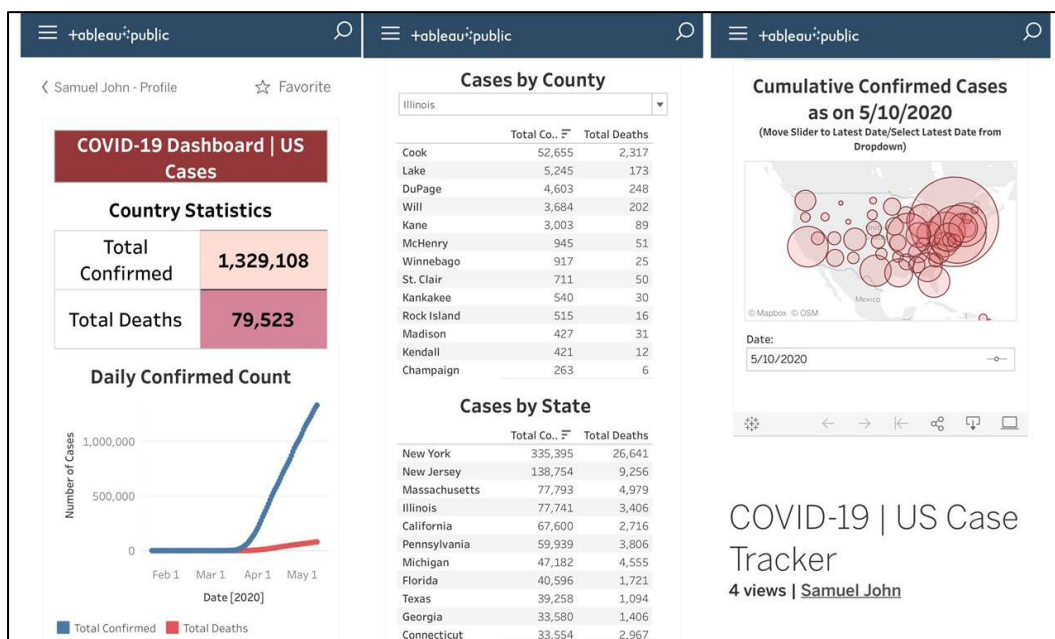


Fig 5.1.2 – Tableau Dashboard for COVID-19 Cases in the US (Mobile view)

5.2 Analytics in Power BI

Power BI also did a good job when it came to converting the raw data into individual visualization objects apart from a few hiccups. It automatically converted fields such as dates into multiple level of grain such as Year, Month and Date and this made it possible to filter out data by date which was important in this case since we wanted to convey the latest information on the number of cases. But when it came to some fields like latitude and longitude it considers these fields as whole decimal values which made it difficult for using these values in Maps to get their locations. Power BI is excellent in manipulating and handling large datasets as many functions which are used in excel can be used here.

Using these features, we were able to build a dashboard as seen in Fig 5.2.1 and Fig 5.2.2

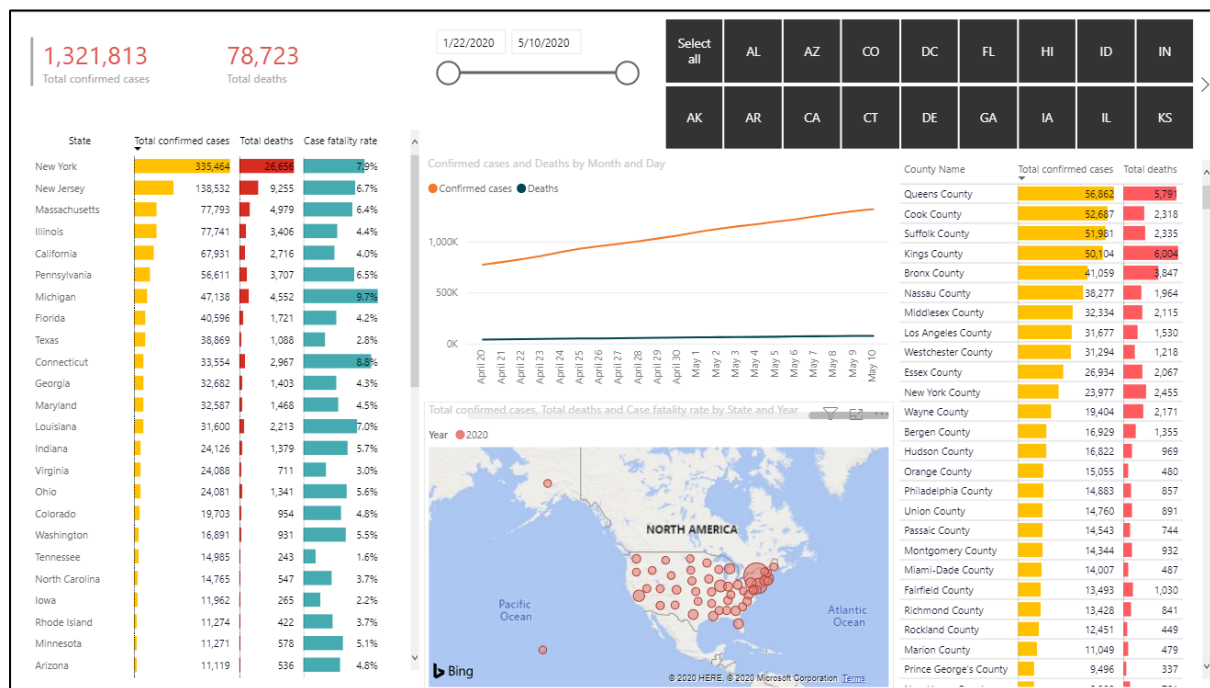


Fig 5.2.1 – Power BI Dashboard for COVID-19 Cases in the US

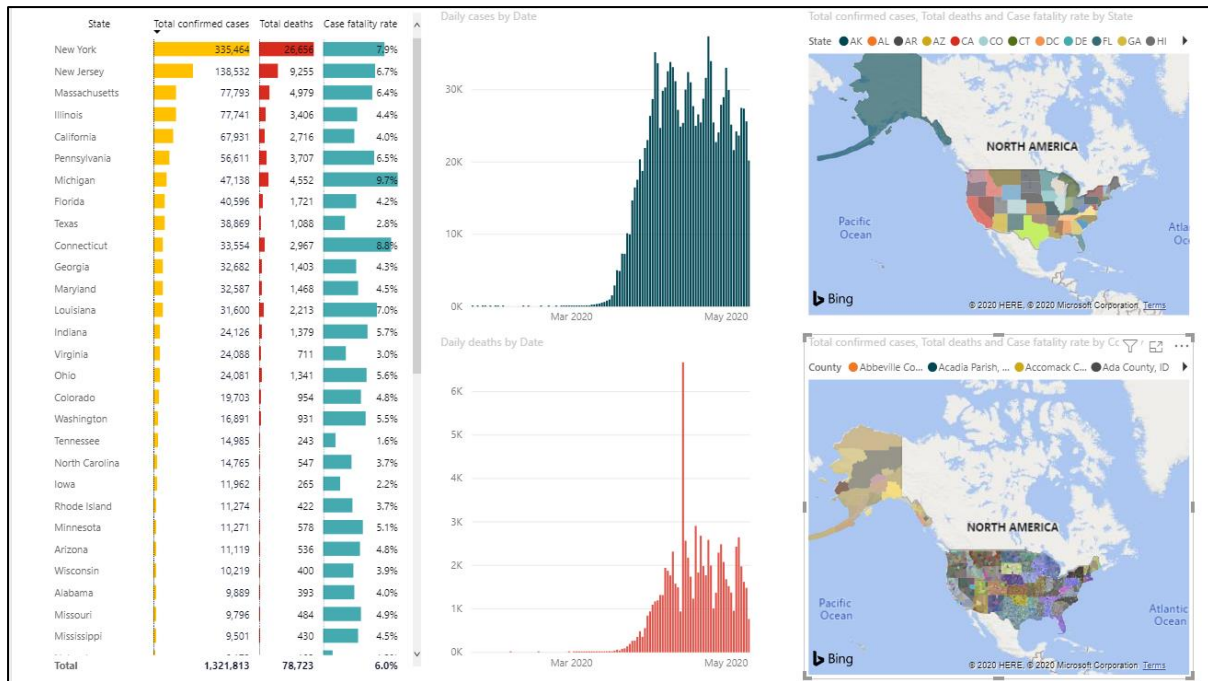


Fig 5.2.1 – Power BI Dashboard for COVID-19 Cases in the US

6. Understanding the Dashboards

The dashboards that we created using the BI tools provided information regarding the total number of confirmed cases in the US, both for the Total Confirmed Positive Cases of COVID-19 and the Total Confirmed Number of Deaths in the country.

They also provided the these daily counts in terms of a graph which made it easy to see the trend in the case counts, which in this case clearly show how the number of confirmed cases were quite low till mid-March 2020, before seeing a gradual spike in numbers that has kept rising each day since. It also outlines the difference in the number of Confirmed Cases and the Number of Confirmed Deaths in the country per day.

The Map provides a visual representation of the density of cases in each state and the tooltip allows user to view information about the number of cases that were tested in each state vs the number of confirmed cases in that state. Further, the dropdown menu allows users to view the visualization for a particular date and the slider allows users to view the change in the density as the dates progress.

The table on the left hand side dashboard provides viewers with quick information regarding the numbers of confirmed cases vs confirmed deaths in each state, sorted in descending order of confirmed cases.

The table on the right hand side of the dashboard allows users to select a particular state and view the statistics of confirmed cases vs confirmed deaths per county in that state, also arranged in descending order of the number of confirmed cases.

Each of these visualizations provide the latest numbers and the numbers are updated as the data source is refreshed.

Overall, the data presented in these dashboards gives users a good idea as to how the pandemic is affecting the country as a whole as well as each state and each county individually while providing an insight into the trend in the change of the count of these cases as whole.

7. Sharing the Dashboards

Since these dashboards were created with an aim to help provide quick insights into the propagation of the COVID-19 pandemic to the public, we thought it would be ideal to share these dashboards on a public forum where people could view the dashboards or use them for their own analysis. Each of the BI tools provided us with different options to do this.

7.1 Sharing the Dashboards using Tableau

Tableau provided us with two easy options to share the dashboards from within the tool itself. The first option was to publish the dashboard using Tableau Server. Tableau Server allows for the creation of a website that hosts the dashboard with the option to share the dashboard with others who have a Tableau Server account. Unfortunately, this is a paid option and only a 14-day trial was available to us. The second option was to share the dashboard to Tableau Public, which allows anyone access to the dashboard, with the option to download the workbook itself for their own analysis. ^[4] We used Tableau Public to share our dashboard, which can be accessed using this link – <https://tabsoft.co/2YV5Dpu>.

7.2 Sharing the Dashboards using Power BI

Power BI desktop provided us with an option to publish to Power BI online service and to share content inside or outside our organization, we need a Power BI Pro license. We can only share the dashboard by entering the full email addresses for individuals, distribution groups, or security groups. We can't share with dynamic distribution lists. When we share with people outside our organization, they receive an email with a link to the shared dashboard or report. They must sign into Power BI to see what you shared. If they don't have a Power BI Pro license, they can sign up for a license when they click the link. So sharing is not so fun in Power BI unless you have a pro license. ^[3]

8. Comparing the Tools

Feature	Tableau	Power BI
Loading the Dataset	Tableau connected quickly to the dataset online, however it took a while to extract the data into the tool. Post the extraction, the tool hung each time we clicked on the Data Source tab because it automatically initialized a fetch query to the data server. This caused the tool to crash multiple times.	Loading the dataset into Power BI was easy and quick. There were no issues while refreshing the data from the web URL and while editing the data.
Data Wrangling	Tableau was relatively easy to use since it automatically converted fields such as country/state into longitude and	Power BI has the capability to handle very large datasets and provides all the excel editing functionalities. For our case, the

	latitude and further broke down these fields in terms of grain.	longitude and latitude fields were considered as decimal values and not recognized by Bing maps which caused a little hiccup.
Ease of Use	The drag and drop feature in Tableau made it easy to add fields and filters to the visualizations. The “Show Me” feature provided options as to the best visualizations to use for the data selected, which made it easier to create the visualizations.	The most talked-about advantage of Power BI over any other data visualization tool is its attractive, intuitive, and interactive visualizations. We could use Power BI’s easy drag-and-drop functionality to add different visualizations in a report. Also, we could interact with visualization by applying filters, making selections in it, etc. It is very easy to create and understand data through visualizations in Power BI
Dashboard Creation	Tableau made it easy to create a dashboard that automatically resized itself to the dimensions of the device being used. It also allowed us the option to edit fields on the dashboard itself, which then reflected across the tabs.	In Power BI desktop though, there is no separate dashboard option, all visualizations need to be copy pasted on to a sheet to create a dashboard. It is easier to adjust, reshape and resize visualizations on the Power BI dashboard than the Tableau dashboard. It is much more flexible
Sharing the Dashboard	Tableau provided multiple options to share the dashboard with others. These options were integrated into the software itself and required only valid credentials to the servers to create shareable dashboards. However, the servers themselves had limited feature support which took away features such as the play and speed options for the slider.	We can’t share the dashboard with others unless you have a Power BI pro license and you can share by entering the full email addresses for individuals, distribution groups, or security groups. You can't share with dynamic distribution lists.
Troubleshooting	As mentioned above, we faced several crashes due to queries to the data servers. Tableau was able to recover the latest version of our dashboard each time the tool crashed.	We didn’t face any other issues apart from the Bing maps taking a huge amount of time to load and its quite understandable as we are dealing with a large dataset.
Support	There are multiple sources of support online including official Tableau pages that provide in-	Power BI gets upgrades from Microsoft every month. Microsoft has made a user

	detail explanations on how to use the tool. This made it easy to work around roadblocks that we faced while creating visualizations.	community where users can upload their suggestions and tweaks about Power BI. The suggestions that get the most rating by other users have a good chance of getting included in the next update of Power BI.
--	--	--

9. Conclusion

By using the BI tools Tableau and Power BI to create dashboards for the COVID-19 cases in the US, we were able to gage the features and ease-of-use of both these tools as well as compare the functioning of these tools to each other.

It was apparent that both tools were able to internally modify the data for efficient visualization, however, Tableau had an edge of Power BI in more than one way. Tableau was more user friendly and most of the tasks were handled automatically by the tools. Visualizations could be selected and dropped into the dashboard and were automatically re-sized and snapped to the dashboard in Tableau while Power BI required us to manually re-size the dock windows. Further, Tableau provided options such as Fit-to-window which allowed for the easy alignment of the data within each block while Power BI did not provide such options.

The visualizations in both Tableau and Power BI provided the right information as required, however, the ones in Tableau did seem to have a touch more of a professional look to them.

Tableau also provided us with the means to share our work with the public through a one-click, single sign on option within the tool while Power BI only provided us the option to share our work with an internal team with the added requirement that anyone who needed to view the work had to have an account of their own.

Overall, we enjoyed working with both the BI tools Tableau and Power BI, but if we had to choose a winner, it would be Tableau due to the advanced features that were provided, the fact that it was easier to use and because it doesn't have dependencies unlike Power BI that was dependent on Bing for maps, had a clunkier front end and required a lot of user actions to create dashboards that were similar to the ones in Tableau.

10. References

- [1] CDC. (2020, February 11). *Coronavirus Disease 2019 (COVID-19)*. Centers for Disease Control and Prevention. <https://www.cdc.gov/coronavirus/2019-ncov/index.html>
- [2] *Home*. (n.d.). Johns Hopkins Coronavirus Resource Center. Retrieved May 11, 2020, from <https://coronavirus.jhu.edu/>
- [3] maggiesMSFT. (n.d.). *Ways to share your work in Power BI - Power BI*. Retrieved May 11, 2020, from <https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-how-to-collaborate-distribute-dashboards-reports>
- [4] *Sharing Workbooks with People Who Don't Have Tableau Desktop* | Tableau Software. (n.d.). Retrieved May 11, 2020, from <https://kb.tableau.com/articles/howto/sharing-workbooks-without-tableau-desktop>
- [5] *What is business intelligence? Your guide to BI and why it matters*. (n.d.). Tableau Software. Retrieved May 11, 2020, from <https://www.tableau.com/learn/articles/business-intelligence>