Analytics: A Necessity in a Chaotic World

Jacob Broderer

Citation
Analytics: A Necessity in a Chaotic World

By

Jacob Ryan Broduer

Advisor: Dr. Carole Shook

An Honors Thesis in partial fulfillment of the requirements for the degree Bachelor of Science in Business Administration in Information Systems.

Sam M. Walton College of Business
University of Arkansas
Fayetteville, Arkansas

May 14, 2022
**Introduction and Personal Learning Objectives**

In today’s ever-changing environment, analytics and data visualization has become an essential part of any successful business. Companies are pouring money and resources into Business Analytics and Big Data, with the issues being created by the onset of the Covid-19 pandemic only serving to accelerate this trend at an astronomical rate. However, many curriculums do not do an adequate job at preparing future business professionals to utilize analytical tools and the meaningful insights that come along with them. The curators of these curriculums fail to acknowledge the fact that analytics and data visualization are no longer something unique to the data science/information systems fields, with workers in categories ranging from marketing to accounting all being in positions to take their work to the next level by implementing these skills. Many companies are also not properly leveraging the power of data analytics, meaning new employees that possess analytics knowledge and skills are in the unique position to help the firm realize it’s true potential. Power BI and Tableau play an important role in making them more well-rounded job candidates and add to the value that they will bring to a potential employer. This project was created with the purpose of exposing students to a real-world business problem and the different visualizations and techniques within the leading BI tools that from my experience as a new member of the work force can be used to address it.
Literature Review

While it might seem like the concept of big data and analytics have been hot topics for some time now, the fact is many companies have still not quite been able to recognize its full potential and take advantage. The consulting group NewVantage Partners conducted a survey in 2021 dealing with the utilization of data within various firms. First and foremost, 92% of firms surveyed said that the amount of money and resources being poured into Big Data within their organization was accelerating. To put this in perspective, this 92% is a 40% increase when compared to the results of the survey from only a year prior (Dykes). The purpose of education is to prepare students to find successful careers once their schooling comes to an end and neglecting a field growing at such a rapid pace limits the number of opportunities they can feel confident in pursuing. On top of the overall increase in spending within the industry, there are also quite a few opportunities to make a sizeable impact from an entry level position. According to the same NewVantage Partners survey mentioned above, a mere 30% of companies believe they have a “well-articulated data strategy” in place, and an even more meager 24% believe they are part of a “data-driven organization” (Dykes). Another study found similar results, noting that only 33% of companies had enough faith in their data to pull meaningful insight out of it (Vasal). These statistics highlight opportunities for two groups of students looking for their first job. Group one involves students who majored in another discipline but also took the time to learn some of the ins-and outs of data analytics. With 92% of companies saying they are accelerating their investments in Big Data/Data Analytics and the analytics field expected to increase 28% by 2026 (Shroeder) thanks to such companies, there will always be plenty of jobs available when finding work in another field proves to be more of a challenge. A student could do a lot worse than an analytics job in terms of a safety net. Group two involves students who have become
proficient in the field of analytics because of either their major or hard work done outside of normal course work. That 24% number mentioned earlier in regards to how many organizations believed they were data driven means that 76% of organizations still have yet to be able to harness the power of Big Data/Data Analytics and implement a concise data strategy. The absence of something that is this critically important to success at such a high level means there are an abundance of opportunities for young professionals to step in and catapult their careers by using data analytics to take these companies into the future. For group one, the assignments in the next section of this thesis provide an excellent way to gain some exposure to the problems presented to analysts and the tools they use to help solve them. For group two, the assignments provide a solid initial steppingstone for a career of success and innovation within the field of data analytics.
Get Started with Tableau Desktop and Power BI Desktop

Learn how to connect to data, create data visualizations, present your findings, and share your insights with others.

The tutorial you are about to complete will walk you through some of the features of Tableau Desktop and Power BI Desktop. By the time you have finished, you will have created four different visuals and combined them into a simple interactive dashboard within a Tableau and Power BI workbook. This assignment is based off of a sample of data extracted from a Microsoft SQL Server database that is owned and maintained by the Information Systems Department of the Walton College of Business. The scenario in which you will be operating focuses on the breakdown of regional sales performance as it relates to departments and brands within Dillard’s Incorporated.

DATA SET REQUIRED: Dillard’s 2004-2005 Dataset.xlsx

ESTIMATED COMPLETION TIME: 2 hours

YOUR SUBMISSION WILL BE IN THE FORM OF A MS WORD DOCUMENT, SO TAKE SCREENSHOTS WHERE PROMPTED AND PASTE THEM INTO THE DOCUMENT YOU PLAN TO SUBMIT.

The Situation:

Your company has recently been contracted by Dillard’s to build a dashboard that will give them insights into the performance of their South-Central region (Arkansas, Texas, Oklahoma, and Louisiana). Seeing this as an opportunity to get his new employee’s feet wet, you have
been given the task of creating this dashboard. After a meeting with the client over zoom, it quickly becomes clear that they would like to view the sales performance in the context of the different Dillard’s departments and brands as they are looking at the possibility of cutting less popular brands in favor of lending more resources to their best sellers. After giving the whole thing some thought, you come up with 4 different visualizations that, when combined, could make a simple and effective dashboard.

You will create the following 4 visualizations in both Tableau and Power BI for your dashboards:

1) A tree map of the top 20 departments and their sales. This visualization will allow the user to see at a glance which departments have the most sales and, in this case, is a nice way to break away from the typical bar graph.

2) A text table of brands and their sales. This table will allow the user to drill down further from the department level to the brand level, providing them a more wholistic view of sales.

3) A geographic view of sales by state. This will give the user a high level view of how the company is performing by state.

4) A geographic view of sales filtered down to the South-Central region. Having a place where the South-Central sales performance is immediately visible is a necessity since this region is the focus of the customer request.

**Part 1: Tableau**

**STEP 1: Connect to the Dataset and Prep Data**

1) Open Tableau Desktop. This will bring you to the page below.
2) Here, you can select the type of data source you would like to connect to. Since the source we are using for this assignment is an Excel File, click the Microsoft Excel option under the "To a File" header next to the red arrow above.

3) A file explorer window should open up after doing this. Navigate to the location you have saved the Excel file *Dillard's 2004-2005 Dataset.xlsx* and open it.

4) After doing this, you will be brought to the data preparation screen below.
Here, you can see how Tableau has handled the dataset you chose to input and make modifications as you see fit.

5) As you can see from looking at the column headers, Tableau has assigned the wrong data type to some of them. The columns that will need to be changed are Trannum, Store, SKU, and Dept. Tableau is treating them as numbers, but we are using them as identifying values in our dataset, meaning they should be treated as the String data type. To change this, click the symbol next to the red arrow in the screenshot and select the String data type. Do this for each of the aforementioned columns.

6) After doing this, your Column Headers should now match the below screenshot.

If everything looks good, then you are ready to move onto the fun part: creating the visualizations.
STEP 2: The Tree Map

1) Now that our data is in a good spot, it’s time to start making our first visualization: the tree map.

2) At the bottom left of the page, click on the Sheet 1 tab. This is where our first visualization will live.

3) Sheet 1 should look something like the screenshot seen below.

On the far left of the screen next to the star, you have your dimensions (qualitative values) and measures (quantitative values) listed under the Data tab.
4) Drag `Deptdesc` from the listed of data on the left and place it in the `Rows` section at the top of the page. Do the same thing for `Sprice`.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rows</strong></td>
<td><strong>Deptdesc</strong></td>
</tr>
</tbody>
</table>

5) After doing this, direct your attention to the `Show Me` tab in the top right. This tab “shows” you the different visualizations you can use with the data you currently have on the sheet. Click the `tree map` option next to the red arrow.

6) Now, pay attention to the `Marks` tab that is directly next to the list of dimensions and measures. Each of these sections within the marks tab allow you to customize the visualization you are making. Drag `Sprice` again from the list of data on the left over to the `Label` box you see in the `Marks` section and drop it. After doing this, you should notice the sales of each department are now visible on the tree map. (Note: Do not use either of the `SUM(Sprice)` measures you see already inside the `Marks` tab)

7) However, something else also stands out. As seen in the screenshot to the right, the numbers are not being shown as dollar values. To fix this, go to `Sprice` in the list of data on the right, right click it, go down to `Default Properties`, and click `Number Format`. In the dialogue box that pops up, select `Currency (Standard)`, and then click `Ok`. Now, all the values should be formatted as a dollar value.
8) At this point, your tree map should look something like this: (on next page)

While this would work, it looks very cluttered and could benefit from some trimming down.

9) To trim things down, we need to add a filter that limits the results to only the top 20 departments based on Sprice. To do this, drag Deptdesc from the list of data on the left to the filter section, located right above the Marks tab. In the dialogue box that pops up, go to the Top tab and copy the settings found in the screenshot to the right. After doing this, click Ok.

10) The final change we need to make is renaming the sheet. To do this, go to Sheet 1 in
the bottom left, right click it, and go to *Rename*. Once there, change the name of the sheet to “Department Sales”. Once this step is complete, you are done with your first visualization! Your final product should look something like this: (on next page)

<table>
<thead>
<tr>
<th>Department Sales</th>
<th>KAREN K</th>
<th>COP</th>
<th>CAROLE</th>
<th>ESPRIT</th>
<th>NO FEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLOMEN</td>
<td>$105,468.16</td>
<td>$68,333.17</td>
<td>$63,776.65</td>
<td>$59,019.41</td>
<td>$59,651.89</td>
</tr>
<tr>
<td>CLINIQUE</td>
<td>$101,995.65</td>
<td>$59,624.99</td>
<td>$48,814.69</td>
<td>$48,131.24</td>
<td></td>
</tr>
<tr>
<td>BE2</td>
<td>$95,500.94</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASH</td>
<td>$85,249.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REEBOK</td>
<td>$49,649.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECHO</td>
<td>$44,479.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is Screenshot #1. Make sure to screenshot your visualization and add it to your MS Word submission.

**STEP 3: Text Table**

1) Now it is time to build a second visualization: the text table. To do this, we will need another sheet. At the bottom left of the page, click the *New Worksheet* button that is directly next to your existing sheet.
2) On this new sheet, drag the *Brand* dimension into the *Rows* section and the measure *Sprice* into the *Columns* section. After doing this, go to the *Show Me* tab’s suggested visualizations and select the *text table* option that should be in the top left corner of the tab.

3) Notice that at the top of the table, *Null* is cited as the first brand in the list. Since this category is just a catch all for records that do not possess a *Brand*, it would be best to filter it out. To do this, drag *Brand* to the *Filter* section.

   In the pop-up that appears, uncheck the box next to *Null*. After doing this, you should notice that *Null* is no longer listed as one of the categories inside of your grid.

4) Lastly, as we did with the last sheet, change the name of this sheet, this time to *Brand Sales*.

5) Now that this sheet is finished, it should look something like this:
This is Screenshot #2. Make sure to screenshot your visualization and add it to your MS Word submission.

**STEP 4: Geographic View**

1) For the Geographic View visualization, we want to see the sales for Dillard’s broken out by state so the user can get a high level look at how their company is performing across the country. Our first step will be to drag *State* to the *Rows* section and *Sprice* to the *Columns* section.
2) After doing this, go over to the *Show Me* tab and select the *map* option. (Do not select the *symbol map* option, which is located right next to *map.*).

3) Lastly, as we’ve done before, change the name of this sheet to *State Sales.*

4) Your finished map sheet should look something like this:

![State Sales Map](image)

This is Screenshot #3. Make sure to screenshot your visualization and add it to your MS Word submission.
STEP 5: South-Central Geographic View

1) Since we have already made a geographic view for sales, we can use it for this step as well. Right click on the State Sales sheet in the bottom left and select the Duplicate option.

2) The South-Central region is comprised of 4 states: Arkansas, Texas, Louisiana, and Oklahoma. Since these are the only states we need to view for our region, we need to add a filter to this duplicate sheet. Add State to the Filter tab and in the pop-up select the four states mentioned above. Lastly, rename the sheet South-Central Sales.

3) The final product should look something like this:

This is Screenshot #4. Make sure to screenshot your visualization and add it to your MS Word submission.
STEP 6: Dashboard

1) Now that we have built all the visualizations we need, it is time to use them to create the final dashboard. This is where all the visuals we have created will come together into the final product. To get started, click the New Dashboard button that is located in the bottom left right next to the New Worksheet button we were using earlier.

2) On the dashboard tab we just created, go to the Objects section in the bottom left and select the check box next to Show Dashboard Title, as we will want to display this title and need to get it on the dashboard sheet so we know how much room we have to work with.

3) Once you have done this, drag the Department Sales sheet onto the dashboard from the Sheets section on the left. After doing this, select the visualization and click the More Options arrow that appears in the top right of the visual. In this pop-up, select the floating option. This option will allow you to freely size and move the visualization in ways you could not with the default setting.

4) To maximize our free space, we can go ahead and delete the legend for the Department Sales visualization as it is not needed to understand what is going on within the visual. To do this, simply select the legend and click the x that appears in the top right corner.

5) Now, drag the remaining three visuals onto the dashboard and repeat the steps above for each of them, switching the visuals to floating and removing any legends that are automatically added to the dashboard.

6) Now, instead of placing stand-alone filters on the dashboard for the user to select the Departments and Brands they want to see, we are going to turn the Department Sales and Brand Sales visualizations into filters. To do this, click on the Department Sales visual and select the Use as Filter funnel button in the top right.
7) Lastly, rename the Dashboard *Sales Summary* and then right click on the title at the top of the dashboard and click *Edit Title*. In the pop-up, select the option to center the title that can be found on the pop-up’s ribbon.

8) You can choose to position your visuals on the dashboard however you see fit and your final project should look like some variation of the following:

![Dashboard Image]

9) For your last step, click the *Polomen* square on your *Department Sales* tree map to activate the filter. After doing so, your results should look something like this: (on next page)
This is Screenshot #5. Make sure to screenshot your visualization and add it to your MS Word submission.

*Bonus:*

1) After showing the dashboard to the customer, he decides he likes the layout but
is not crazy about the color scheme and asks you to show him something else. When asked what exactly he wants, he says, “Surprise me.” On your own, customize the colors of each visualization and provide the client with your revamped dashboard. *(If you choose to do this, screenshot your final product and add it at the end of your submission with the label “Bonus”).*

**Part 2: Power BI**

**STEP 1: Connect to the Dataset and Prep Data**

1) Open up Power BI Desktop. You will see the screen shown below.

![Power BI Desktop](image)

2) Click on the *Import data from Excel* option. In the File Explorer window that opens, navigate to wherever you stored the Dillard’s dataset. Once you find it, open the dataset. You should then see the following window: (on next page)
3) In this window, select the checkmark next to Sheet 1 and then click the Load button in the bottom left of the window.

4) Now that we have the data loaded, we need to set all the Dimensions and Measures within our dataset to the proper data types. To do this, navigate to the Data button located near the top left corner of the screen.
5) Power BI has given the columns *Trannum, Store, ZIP, SKU,* and *DEPT* the wrong data type. We are using these columns as ID’s, which means they should all be utilizing the *Text* format. To change a column’s Data Type, go to the *Column Tools* tab on the ribbon, click the header of the column you are changing, and select *Text* in the *Data Type* drop down in the *Structure section*. Do this for each of the columns listed above.

6) For the *Sprice* column, the data type is correct. However, the formatting needs to be changed to *Currency* as the column is referencing sales. To do this, go to the *Formatting* section on the same *Column Tools* tab from above and use the *Format* dropdown to select *Currency*. Directly underneath this dropdown, use the decimal places box to change the number of places to 2.

**STEP 2: Tree Map**

1) On the right side of the screen in the *Visualizations* section, click on the *Treemap* option. After doing this, select the checkbox next to *Deptdesc* and *Sprice* in the *Fields* section that should be right next to the *Visualizations* section.
2) Now, you will notice that each department within the treemap has been given a unique color. We need to change this to a gradient scale that uses different shades of the same color to illustrate which departments have the most sales. To do this, click the *Format* tab illustrated by the paint roller in the screenshot above.

3) In this *Format* section, find the *Data Colors* option and expand it. At the top of the list within the dropdown there should be an *Advanced controls* option. Click on it.

4) In the window that opens, copy the settings in the screenshot below. This will accomplish the goal stated in step 2.
Notice that when you look at the graph, each square in the treemap only has the department name listed within it. We want to also see the sales for each department. Within the same Format tab, find the Data Labels option and switch it to On. At this point, your treemap should look similar to this: (on next page)
6) As you can see, it is not practical to have this many departments on the graph as things get far too cluttered. To fix this, we need to filter the number of departments down to the top 20 based on Sprice. Open the filter tab that’s on the right next to Visualizations. Within this tab, there should already be a filter for DeptDesc. Set all the controls within this filter to match those seen in the screenshot to the right. After doing this, the treemap should only be displaying 20 departments.

7) Lastly, within the Format section, go to the Title option and type "Department Sales" within the Title text box. After doing this, your final visualization should look something like this: (on next page)
This is Screenshot #1. Make sure to screenshot your visualization and add it to your MS Word submission.

**STEP 3: Table**
1) For our next visualization, we will create a table that shows each brand and its sales. Make sure you do not still have the treemap visualization selected, go to the Visualizations tab, and select the Table visualization.

2) In the Fields section, select the check boxes next to Brand and Sprice. Notice that there is a row at the top of the grid that is missing a Brand value but has sales attributed to it. This represents entries with a null brand.

3) We do not want to see null values in the graph. Go to the Filters section and expand the filter for Brand. Set the filter so that every checkbox is selected other than the one that is next to the null brand.

4) Following the same steps as the previous visualization, change the title of this table to Brand Sales. Your now-finished table should look something like this:
This is Screenshot #2. Make sure to screenshot your visualization and add it to your MS Word submission.

**STEP 4: Geographic View**

1) Make sure none of your existing visualizations are selected and click on the *Filled Map* visualization in the *Visualizations* section on the right.

2) Select the check boxes next to *State* and *Sprice* in the *Fields* section.

3) Go to the *Format* tab indicated by the paint roller and expand the *Data colors* option. Click on the *fx* button that will open a window that will allow you to color the states based on sales.

4) In this window, use the same settings shown in the screenshot below. There is no need to change the color used to indicate the highest *Sprice*, but you need to change the color that is used to indicate the lowest *Sprice* to the *Theme color 1, 60% lighter* color.
5) After doing this, stay inside the *Format* tab and expand the *Map controls* option. Within this option, switch the button next to *Zoom buttons* to on. Now, on the visualization, use the zoom button to get a better view of states. Change the title to *State Sales* and the finished visual should look like this: (on next page)
STEP 5: South-Central Geographic View

1) This next map will just be a copy of the one above with a little more filtering. Copy and paste the previous visualization onto the page.

2) In the new visualization, go to the Filters section and change the State filter so that it only has four states selected: Texas, Arkansas, Louisiana, and Oklahoma.

3) Change the title of this visualization to South-Central Sales. Your final product should look something like this: (on next page)
STEP 6: Dashboard

1) Since we have all of the visualizations already on the same page, this step will just involve adding a title and arranging the visualizations. To add a title, go to the Insert tab at the top and insert a textbox onto the sheet. In this textbox, type Sales Summary as this will be the title of our dashboard. I used size 24 Ariel font that
was bolded, but feel free to use whatever you think will look best.

2) Now, move everything around on the page until you are happy with the layout. Collapsing the *Filter, Visualization, and Fields* tab can be helpful when trying to get an accurate view of the way things will look. Once you have things looking good, click on the *Polomen* department within the tree map. As you can see, the visualization acts as a filter for the rest of the dashboard. This same feature can be found on the other visuals as well. While it might not be exact, your dashboard should look something like this (KEEP FILTER SELECTED): (on next page)

This is Screenshot #5. Make sure to screenshot your visualization and add it to your MS Word submission.
*Bonus:

1) After showing the dashboard to the customer, he decides he likes the layout but is not crazy about the color scheme and asks you to show him something else. When asked what exactly he wants, he says, “Surprise me.” On your own, customize the colors of each visualization and provide the client with your revamped dashboard. (If you choose to do this, screenshot your final product and add it at the end of your submission with the label “Bonus”.)
Analysis/Implications

Having finished the project and being able to look at final results, I believe everything that was tried along the way worked the way it was intended to. I was able to create a step-by-step assignment that introduces students to Tableau and allows them to get hands on experience with the software. I was also able to replicate these same steps inside of Power BI, exposing students to yet another tool and allowing them to compare the strengths and weaknesses of each platform. Furthermore, the project utilized visualizations and software capabilities like using tree maps as filters on a dashboard that are not typically part of entry level assignments. This fact will allow students to set themselves apart when demonstrating their skills to their professors and potential employers alike. In the long run, I am confident that this project will allow students to feel more capable and confident when asked to display their skills and give them a creativity edge over others who have partaken in the standard “tried and true” data visualization assignments.

Conclusion and Personal Evaluation

The most stressful part of the job searching process is the uncertainty that comes along with it. For me, it was very easy to feel underqualified for jobs that, in reality, were right up my alley. This feeling of inadequacy is what makes the interviewing process so stressful as you have no idea what to expect and are afraid the shortcomings that are present in your head will be exposed in front of a potential employer. Having already been through this interview process and been hired in a full-time role, I can honestly say that the two assignments contained in this thesis provide a glimpse into the actions one might be asked to perform during the hiring process. For
the role I currently hold, I was sent an Excel spreadsheet and asked to prepare a dashboard that would be critiqued during my formal interview. The workflow to these projects is very similar, with the only major difference being there is a set of directions to guide the student along the way as they were made as introductions to the BI tools. I believe giving students confidence in their own abilities and preparedness for what comes after they graduate should be the ultimate goal of any successful curriculum, and this thesis would serve as a meaningful contribution to any such curriculum.
References

