Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Tableau overview

Tableau is a data visualization and business intelligence software that allows users to connect to various data sources, organize, and analyze data. It enables users to create interactive and shareable dashboards that provide actionable insights through charts, graphs, and other visual representations. Tableau is widely used for making data-driven decisions, as it simplifies the process of exploring large datasets, spotting trends, and communicating data insights effectively. Key features include data connections, visual analytics, interactive dashboards, and collaborations.

Tableau public & desktop

Tableau Public is a free version of Tableau's data visualization software designed for creating, sharing, and exploring public-facing data visualizations. Tableau Public has some limitations, but it remains a powerful tool for creating interactive visualizations that anyone can access online. Tableau public is free and cloud-based.

Tableau public is **not private.** Everything you do will be published to the Tableau server and available publicly. You'll need Tableau desktop if you are working with secure data. You can download a free version of Tableau public from the Tableau website.

Tableau guide

The value of serendipity

As an undergrad, I needed a one-hour science course to fulfill a gen ed requirement. I found an independent study course and a professor who would work with me. He pulled a random book off his bookshelf and said "why don't you read this book and write a few papers about it?" Those were the only instructions. The book is Serendipity: Accidental Discoveries in Science. I loved that class. It opened my eyes to a whole new way of learning and discovering.[1]

In my experience, one of the best ways to tell data stories is to intimately learn about your data by experimenting with it and working with visualizations. This is time consuming at first. As your skills expand, however, you will become more efficient.

I will be using this guide in the workshop. This guide is not exhaustive and reflects my experience. You may already have an effective process. Or discover one that works better for you over time. You may find you don't like my file naming convention. Or come up with a better way on your own.

It is not possible to learn all of the ins and outs of Tableau in a few hours. However, you will be in a position to get started and create your visualizations and stories as soon as you return to your office.

Remember: No one knows the data better than you. So tell the story!

[1] I looked up the professor, Dr. James Carrel. Remarkable enough, he is still an active professor!

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

The survey dataset

About

This presentation will utilize fake data from a fake survey. Please don't spend time critiquing the survey questions or design. It is not perfect. It is designed to simplify the data structure you will need to complete the tasks associated with this workshop.

The dataset has three spreadsheets.

1. Results – text
2. Demographics
3. Questions

It also includes 'working spreadsheets.' You will not be using them in Tableau. They may be helpful for further analysis, however.

Dataset name: **Belonging survey dataset 24.xls**

Data clean up

Before importing data into Tableau or Power BI, you'll need to make sure to do the following. I have already done this for you.

- The number of ID's is the same in the demographic and the results tables.
- The question IDs match the question IDs in the results table.
- I like to create a pivot table tab to verify Tableaus calculations are the same as Excel.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Tableau environment & importing data

Importing data 1

- Click on Connect to a File: Microsoft Excel
  - Note how the Excel tabs are located on the left.
- Drag the **Results - text** table to the workspace.
  - Note how the Results table from Excel is replicated in the Tableau environment.

Pivoting

- Click on Qid1. CTRL+click on Qid23 so that all the questions are selected. **Do not click on SurveyID.**
- Right click anywhere in the "blue space." Then click on **Pivot**.
  - Note how you now have three columns: Survey ID, Pivot Field Names, and Pivot Field Values.
- Change the name of Pivot Field Names to **Question ID** and Pivot Field Values to **Response**.
  - Note how all the values are in one row. We will do the calculations from this row. Because each response has a question ID filter next to it, we can filter the calculations by Question ID.
- **Save the Tableau Workbook.**
  - If you are in Tableau Public, you may be prompted to save it to Tableau Public and the Tableau website will appear. That's fine.
  - If you are using Tableau Desktop, save to your computer.

Importing data 2

- **Drag the Questions table to the workspace** and create a relationship with the Results table. Do not put it in the "New Base Table" box. If you do, click the left arrow in the upper left to Undo.
- To create a relationship with the new Pivoted Results table, you'll need to create a relationship with it. In the bottom left, **create a relationship** between the two tables using **Question ID** as the key.
  - You will note the **Question ID** in the **Questions** table is labeled Question ID1. That is because two columns can't exist with the same header title. So, Tableau adds the "1." It doesn't matter.
  - Note how the Question table now appears in the workspace.
- **Move the Demographics table** and match it with the Results table (blue squiggly line).
  - We will use the demographics table to filter responses by demographics.
  - Note how the Demographics table is a copy of the results table. Some users will only move over specific demographic columns. I prefer to copy the whole table because I filter results questions by other questions. This could create issues if you're working with giant datasets. But mostly, I work with surveys, so it's not a big deal.
  - Because the **SurveyId** column is in both tables, Tableau automatically creates a relationship. This only works because the number of IDs in the **Results** and **Demographics** tables are the same.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Data cleanup: Aliases and groups

- Note how our variables are string. We know this because there is an ABC above the variables. However, the first generation column has numeric data. So, it's classified as a number field.
    - First-generation is a categorical variable. So, let's change first-generation to a string variable and create aliases.
- In the Demographics table, click on the # and **change first-generation to string**. Even though it is string, the values are still numbers.
- Click on the little black arrow next to first gen. Click on **create alias.**
    - Select "continuing" for 0.
    - Select "first gen" for 1.
- **Create groups for GPA.** I created groups for <2, 2.00-2.49, 2.50-2.99, 3.00-3.49, >3.50. Select whatever group you want. Tableau will create another column for the GPA – groups.
    - It would make analysis easier to create a groups for GPA. Otherwise, your visualizations will include all possible GPAs.
    - You can go back and edit the groups or create additional groups.
    - Note how there is a new column for the GPA groups.

Save the Tableau workbook. You may be asked to create an extract or use a live connection.

- An extract is a snapshot or a copy of the data taken from the original source and stored locally in a Tableau Data Extract file (.hyper or .tde). Tableau will create a folder for you.
- A live connection means that Tableau queries the original data source in real-time, without copying or storing the data locally.

I prefer live connections for a variety of reasons. You may prefer the extract in other situations. Select cancel for the live connection.

You are now ready to visualize and tell stories!!!

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Creating Visualizations

Get organized

- Go to the first worksheet.
- Find the tiny down arrow to the right of the search box in the upper left. Change view from Group by Data Source Table to Group by Folders.
    - On the top line are dimensions. Measures are on the bottom.
    - Dimensions are fields that typically contain categorical data, which is used to slice and group your data. Dimensions usually provide context for the data and answer "who," "what," "where," and "when" questions. **Dimensions are like independent variables.**
    - Measures are fields that typically contain quantitative data and represent something that can be measured, counted, or aggregated, such as sums, averages, or counts. Measures usually answer "how much" or "how many" questions. **Measures are like dependent or outcome variables.**
- Tableau placed some dimensions in the measures section. Click on them and drag them to the dimensions section. Survey ID is a dimension.

Explore the question structure

Question organization

- Drag the **Category pill** to the row pane or the “Drop field here.” It’s the same thing.
    - Note how the categories are listed.
- Now move the **questions and subquestions** over in that order. This is your question structure.
- Move the **Category pill** to the Filters pane.
- Click on the **Category pill** in the filters pane and select Show Filter.
    - Play with the filter and see what it does the question structure table.
- Move the **Surveyid pill** to the rows pane or table viz. Tableau puts the **Surveyid** in the column pane. Move it to the row pane in front of Category.
    - Tableau lists the survey id for each question. We can see the responses for each individual.
- Move the Surveyid pill to the end in the row pane. Click on the **Surveyid pill**, select **Measure > Count Distinct**. A lot of little graphs will appear.
- In the Show Me pane, select the table.
    - See now there are 2128 responses for each question. Tableau here is counting rows, not responses. So nulls and blanks are included.
    - This is what you want to see – the number of responses matches what we have in Excel.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Visualizations

Create the viz

- Create a new worksheet.
- Move **Category pill** to filter, then show filter. When asked to select from the list, only select categories related to the survey responses, not demographics. They are: Belonging, Interactions, Learning impact, and Satisfaction.
- Move the **Category**, **Question** and the and **Subquestion pills** to the row pane.
    - You can move the order of the categories by clicking and dragging.
- Let's only look at the Belonging questions. So select the belonging category in the filter, or deselect the other categories. The three belonging questions will appear. (You can delete the category and subquestion pills).
- Move the **Surveyid pill** to the column pane. Add all members. Change surveyid to discrete if necessary. Even though they are numbers, IDs are categorical and are discrete..
- Click on the **Surveyid pill** and select Measure > Count (Distinct). A bar chart with blue bars will appear.
    - This only shows the number of people (2128) for each question. We want the number of individuals per response.
- So, move the **Response pill** to the row pane. Now you can see the number of individuals who responded to each question.
- Reorder the responses in a logical way. Click on each one and move it in the order you prefer.
    - You'll want to see each response by color.
- Move the **Response pill** to the Color box in the Marks Section. Note how the bars all have different colors.
- We don't care about the null, so right click on Null and exclude it from the chart.

Assign a color palette

- Tableau has placed a color legend on the right. Click on the down arrow and select **Edit Colors.**
    - Select a palette. I went with **Summer**. The colors may not match what you want. Tableau assigns colors based on the order in which they appear in the chart.
    - So, click and drag to reorder to the responses. Strongly agree on top, followed by agree, strongly disagree, and degree on the bottom.
    - Note how the extreme values have a darker, bolder hue.
- Assign palettes for the Interactions, Learning Impact, and Satisfaction questions.
    - Unselect Belonging in the Category filter and select Interactions. Note how the **Subquestion** pill is missing. We didn't need it for belonging, but we do for interactions. So move the **Subquestion** pill back to the row pane. Since this is a 7-point scale, we are going to use a gradient hue. Drag the questions in reverse order, with Excellent on the bottom. I went with purple. Then, place the bars back in order.
    - Assign palette for Learning Impact.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

- Assign palette for **Satisfaction**. Notice how the **Satisfaction** questions are present as well as the overall satisfaction question. We want the overall satisfaction question to have a different color scheme.
- So, move the **Questions** pill to the filter pane. Click show filter. Note how two questions are selected. **Deselect How satisfied are you with the following college services?**
- Remove the **Questions** pill from the **Filters**. Note how both questions now appear.

Create a stacked bar chart

- Click on **Belonging** in the Filter window.
- Drag the **Response pill** out of the pane or delete. Stacked bars are automatically created.
- Drag to make the bars thicker and chart larger.
- You can also move the questions around in an order that makes sense.

Create percent of totals

- Click on the arrow in the **CNTD(Surveyid) pill**. Select Quick Table calculation. Select **Percent of Total.**
- Select **Compute** using from the **Response pill**. Select Table across. Now, each bar adds up to 100%.
- Right click on the **X axis** and change to something more intuitive.
- In the **Edit Colors** pane, move the responses around in an order that's more intuitive for you.
- Right click on the **subquestions** in the chart. Select **Format**. In the header, select 12 point font and bold.
- Explore the other categories and make adjustments.

Save your file!

Demographics

- Right click on the stacked bar tab. Click duplicate. An identical worksheet is created. Name it **Demo Gender.**
- Move the **gender id pill** to the rows column.
- Slide the **gender id pill** in front of the questions pill in the row pane.
- To fit the chart into the window, select **Entire View** from the drop down menu (top).
- In the **Marks** window, select label. Note labels now appear on the bars.
- Right click on a data label and select **Format**.
- In the format window, select the little arrow by **Fields**. Select the **% of Total CNTD** category.
- In the pane tab, select numbers > percent. Select 0 decimals.
- Remove the **gender pill** and move over the first gen pill. Place the **first gen pill** after question. Note any changes.
- Play around with **GPA groups**, **Race**, and other demographic categories and note any trends.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Time graphs

- Duplicate the **demo gender** tab.
- Move **survey year** next to **first gen**.
    - Note now little slope graphs are created.
    - Because year is classified as continuous, 2022 and 2023 are included.
- Move **Question**, **First gen status**, and Year to the Column pane. Move **CNTD(Surveyid)** to the Rows pane.
- Remove **first gen** from the column pane and marks panes.
- Change **Compute** using to table (down) from the **CNTD(Surveyid)** pill.
- Right click **Question** on top of chart and select hide.

Dot plot

- Create new worksheet.
- Move the **Category** pill to **Filters**. Select Show **Filter** and select **Belonging**.
- Move **Calc Belong pill** to the columns pane.
- Change the calculation in the **Calc Belong pill** from Sum to Avg.
- In the **Show Me** tab, select the dot plot.
- Move Avg**(Calc belonging)** to the Columns tab.
- Move the **Questions** pill to **Rows**.
- Right click Edit Axis.
- Select 2.3 as the fixed start and 3.3 as the fixed end.
- Change Axis Titles to Average Belonging Score.
- Pull down on the chart with your mouse to make it larger.
- In the **Marks** pane, change the shape (drop down menu) to circle. The dots are now filled in with colors.
- Let's make the question larger. Right click on the header (where the questions are) and select Format. In the Header tab, select Font and change font size to 12 and bold.
- Move the **Race pill** to the Row pane, after questions.
- Right click on Null and select Hide.
- Remove **Question** from the **Marks** window (if it is there).
- Drag the **Race pill** to the colors pane. Note how each race category is a distinct color.
- Move race to before **Question** and note the different alignment.

If you want to add data labels:

- Click on labels and select **Show Mark Labels**. Under horizontal alignment, select center.
- Right click on one of the data labels and select format. In the **Format** Font window, select "Avg Calc Belonging" from the Fields drop down (there's a little black upside triangle next to it).
- In the Pane tab, go to Default. There, select Numbers > custom > one decimal.

Save your file!!!

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Dashboards & Stories

Dashboard

- In the bottom left **Objects** pane, select **Show Dashboard Title** and type **Belonging Dashboard** (or whatever you want).
- Under the title, add an object. Select height as 10. In the Layout tab, select black for background. Now you have a line between the title and dashboard.
- Go back to **DB – belonging**. Add First-gen status, race, gender ID, Discipline, GPA term groups, and **Overall Satisfaction** to the filters. Do not add year. Click on show filter for each one.
- Move the **DB – Belonging** sheet to the dashboard. Notice how the filters are also included.
    - Delete the category filter.
    - The filters only apply to the **Belonging** chart. So we need to change the filters to apply to all the charts.
- For each filter, change to "apply to all using this data source." It's in the little black arrow.
- Move the **DB - Interactions** sheet to the right of Sense of Belonging.
- Move DB – satisfaction with services below sense of belonging and **DB – Learning below Interactions** with others.
- Play around with the size in the layout until you get a dashboard you like. See mine as an example.

Stories

- Create worksheets with stories you want to tell. Once you have them, add a new story.
- To create a story, click on **Blank** on the New story point pane.
- Add the workbook and write the story.

Empowering Insights: Data Storytelling with Tableau

Workshop Guide

Ryan Smith, Illinois State University, 10/27/24, Assessment Institute, Indianapolis IN

Calculated fields

```
IF [Response] = "A lot" THEN 4
ELSEIF [Response] = "Some" THEN 3
ELSEIF [Response] = "A little" THEN 2
ELSEIF [Response] = "None" THEN 1
ELSE 0
END
```

```
IF [Response] = "Excellent" THEN 4
ELSEIF [Response] = "Good" THEN 3
ELSEIF [Response] = "Okay" THEN 2
ELSEIF [Response] = "Poor" THEN 1
ELSE 0
END
```

```
IF [Response] = "Excellent" THEN 7
ELSEIF [Response] = "Very good" THEN 6
ELSEIF [Response] = "Good" THEN 5
ELSEIF [Response] = "Fair" THEN 4
ELSEIF [Response] = "Poor" THEN 3
ELSEIF [Response] = "Very poor" THEN 2
ELSEIF [Response] = "Terrible" THEN 1
ELSE 0
END
```

```
IF [Response] = "Very satisfied" THEN 4
ELSEIF [Response] = "Contently satisfied" THEN 3
ELSEIF [Response] = "Somewhat satisfied" THEN 2
ELSEIF [Response] = "Not satisfied at all" THEN 1
ELSE 0
END
```