THE TABLEAU EBOOK

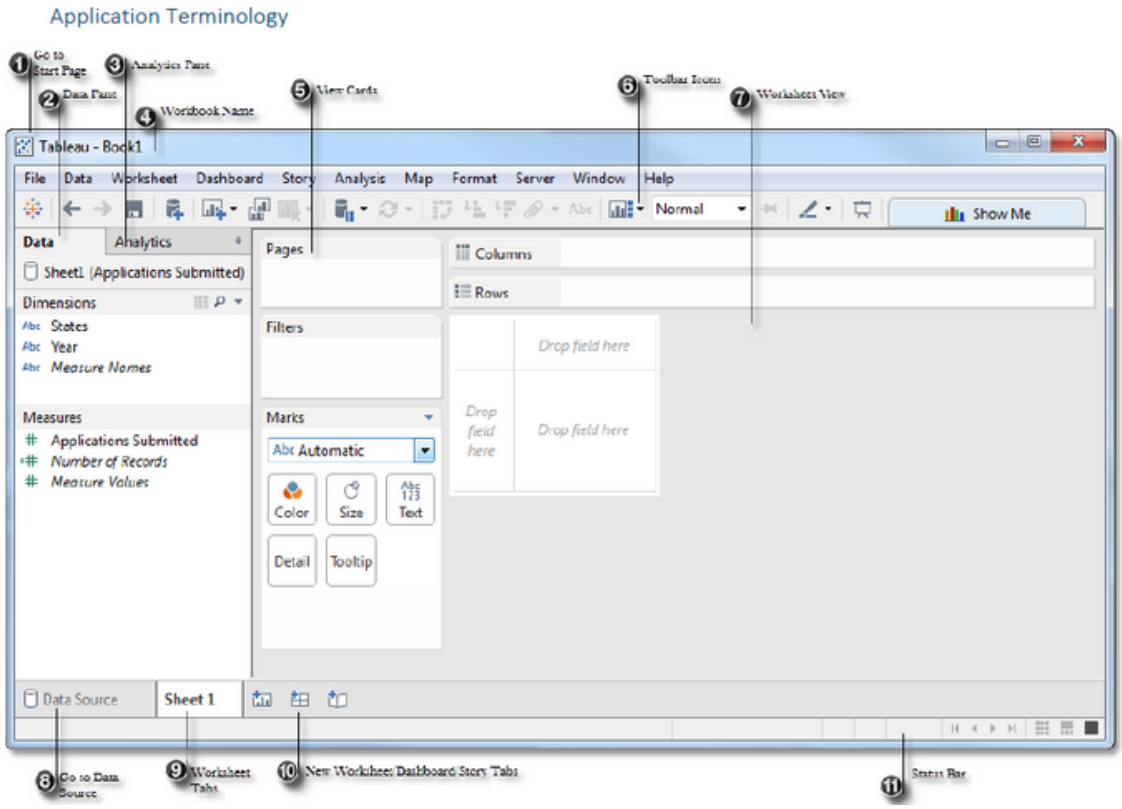
created by: Akshay Kamble

Contents:

Introduction:

Welcome to the world of Tableau, where data transforms into captivating visual stories. In today's data-driven age, the ability to analyze, interpret, and present information in a meaningful way is a skill that can unlock countless opportunities. Tableau stands as a beacon in this landscape, empowering individuals and organizations to harness the potential of their data like never before.

The Tableau Reporting Tool



1. **Go to Start Page:** Toggle between the active sheet and the Desktop Start Page.
2. **Data Pane:** Includes dimensions and measures, populated from your selected data source. May also include calculated fields, parameters, or sets.
3. **Analytics Pane:** Includes options you can use to apply reference lines, forecasts, trend lines, to add totals to crosstabs, and to build boxplots.
4. **Workbook Name:** The file name of our workbook.
5. **View Cards:** Used for modifying the worksheet.
6. **Toolbar Icons:** Icons are available for quick access to popular features.
7. **Worksheet/View:** Workspace for building your visualizations.
8. **Go to Data Source:** Returns you to the data source specification page.
9. **Worksheet Tabs:** Click to view a specific worksheet, dashboard, or story
10. **New Worksheet, Dashboard, and Story Tabs:** Click to create a new Worksheet, Dashboard, or Story.
11. **Status Bar:** Displays data about the fields and marks included in the view.

Data Sources

Tableau Desktop is a tool that allows you to connect to nearly any data source. These include Excel, Access, server sourced and cloud based data. You can also easily join multiple data sources for your visualizations. This can be useful when you need to include details from more than one source. For example, you may have performed a survey to find the effectiveness of a program. In order to determine whether the results varied by select demographic details, you might want to join the survey participant information with the personal details from a human resources database. By doing this, you could greatly enrich your reporting capabilities without overwhelming the survey takers with identifying information. This capability within Tableau allows you to move easily beyond previous reporting constraints.

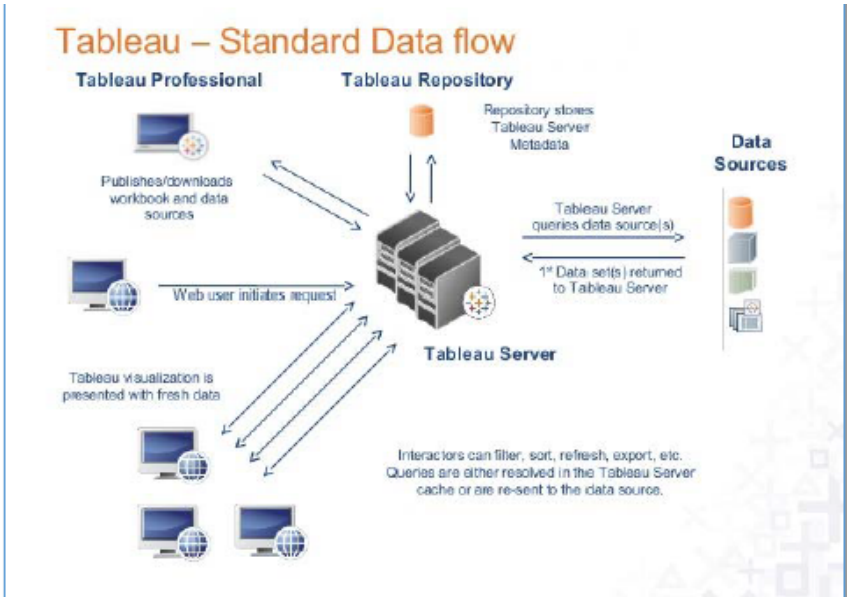
However, with this ability comes a new responsibility. It's extremely important to understand fully the data on which you are reporting. Because you will be creating new joins and providing subsequent visualizations, it's important to know that the results are accurate. As you learn, validate your data by running queries from the warehouse or in other proven methods to ensure your counts are correct. Likewise, since you have the ability to create new data connections, it is your obligation to ensure you are not sharing any confidential material. This is discussed in depth in the following section.

Tableau Server

Tableau Server is a browser based platform where you will publish your dashboards in order to share them with others around the university and beyond. Here at Princeton, the sites on Tableau server are managed in a manner very similar to the data warehouse. Each business unit will have a site and that site will be managed by a designated site administrator. The site administrator will be responsible for managing users by granting and setting access to data and dashboards on the server site.

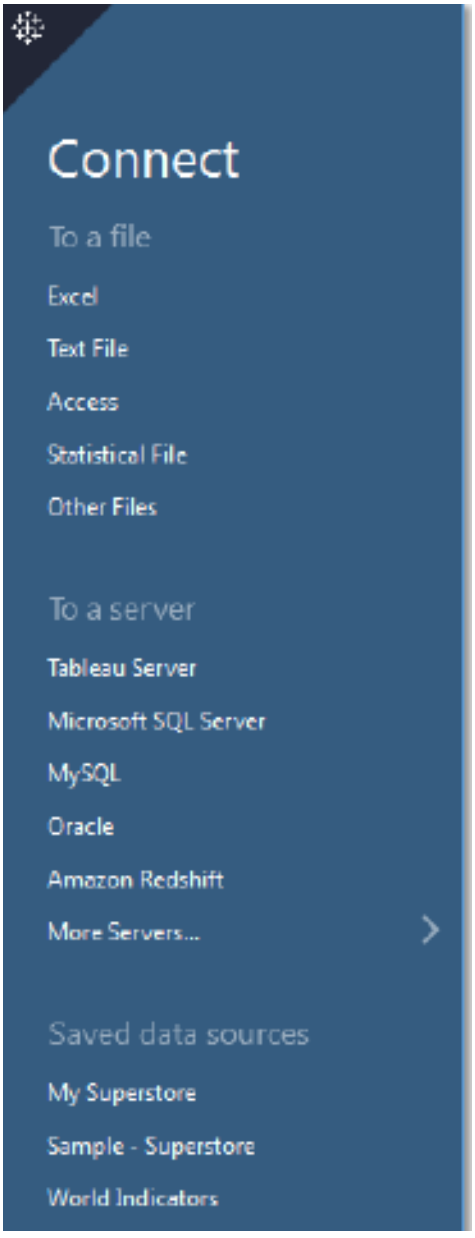
As a desktop user, when you have created dashboards you need to share, you will publish to your server site. In order to do this, you will need to have a Publisher role on the server site and access to the Project. Please note that you must not have a later version of Tableau installed on your desktop than that of the server if you would like to publish to the server (e.g. if the server is at release 9.2, desktop users with releases 9.0, 9.1, and 9.2 may publish to the server but those with release 9.3 will not). Once published, report users will be able to interact with the workbook at their designated level of access as defined by the site administrator.

As mentioned earlier, when outlining data security, there may be times when you need to embed your dashboards into external websites. If this is the case, you will need to bypass authentication. This is when you would use the Tableau Public server.



Understanding Tableau Data Sources

A data source is a reusable connection to data including connections to relational databases, cloud-hosted databases, spreadsheets, and more. When using Tableau Desktop, you may select data from a variety of locations including your local computer, server hosted or even from the cloud. In order to share a data source with others, you would first connect to it in Tableau Desktop and then publish it to the Tableau Server. The published data sources can include data or connections to live databases. The published data sources can also include layers of customizations including calculations, groups, and sets.



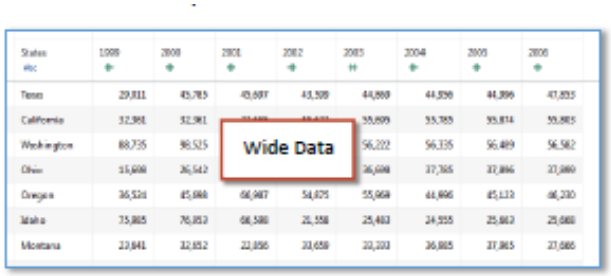
You should publish a data source when you want users to connect to the same data source from multiple workbooks. When a published data source is refreshed, workbooks using the source will reflect the changes. This facilitates consistency and accuracy in reporting.

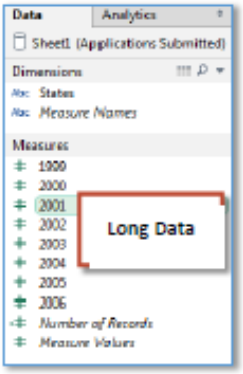
It is often useful to create a data extract in Tableau. This pulls the data from a data source into a stand-alone data set for use in Tableau Desktop. This is especially useful if performance is an issue or if you need to distribute to users without access to the data. It is important to note, however, that you cannot create an extract of a cube data source.

If you only want users to connect to a data source from a single workbook, you should embed the data source in a workbook but do not publish the source. Every published workbook has at least one embedded data source.

Tableau Data Format

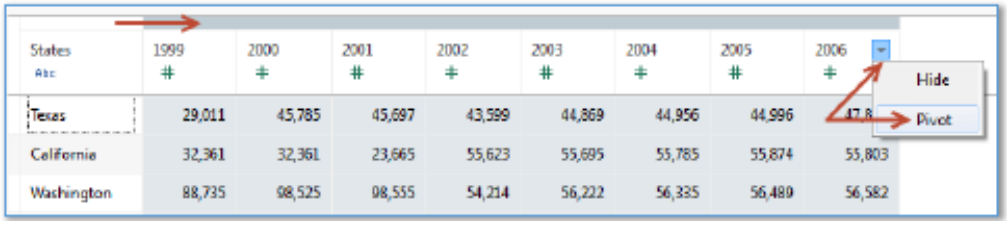
Tableau can work with data in varying formats. However to get the most out of Tableau, it is best if your data is at a detail level or "long" with a separate record for every element you are counting as opposed to compressed or "wide" with many details for each element.





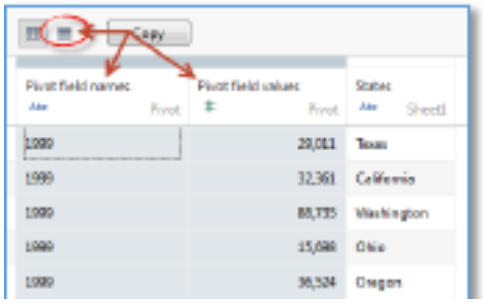
While you can work with "wide" data in Tableau, you will have reduced reporting capabilities.

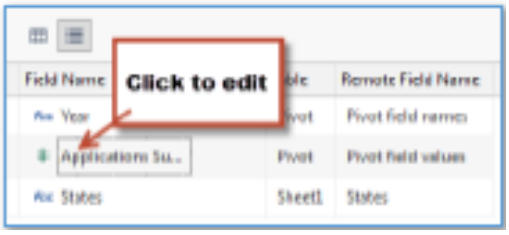
If you're starting with "wide" data, you do have the option to Pivot your data using this feature on the Data Connection page. To do this, select the columns you wish to transpose, hover over one of the selected columns, and select Pivot from the drop down arrow.



Once pivoted, your data will be in the detail or "long" format making it easier to work with.

If you have taken these steps, you will want to ensure your field names have been updated to reflect the new format appropriately, before moving to your initial worksheet.

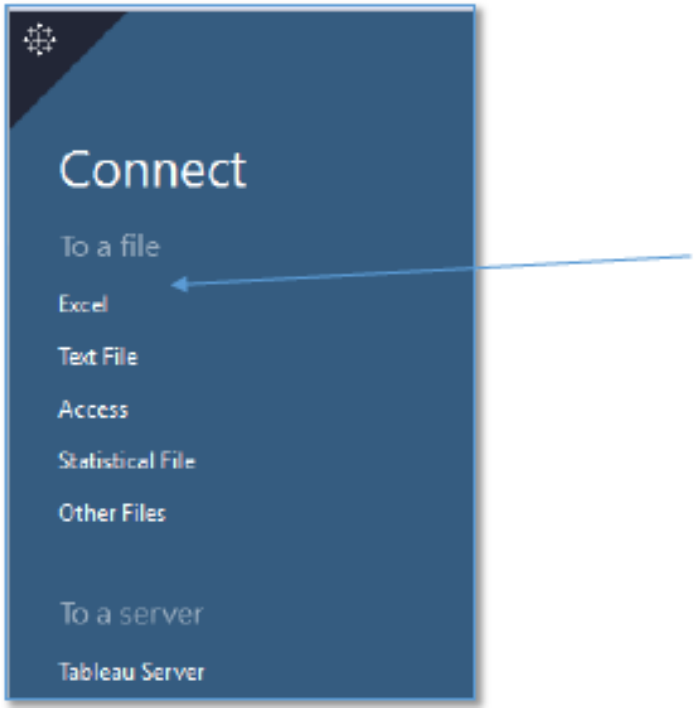




Connect to an Excel Data Source

Tableau works well with spreadsheets. In order to get familiar with the tool, you can download data from the Data Warehouse and save it locally in excel format. Once you have done this, you are ready to get started with the tool. In order to help you become familiar with the mechanics of the process you'll need, our first example will create a connection to a local excel file.

1. Start by opening Tableau. From the start page, select the database type of Excel.



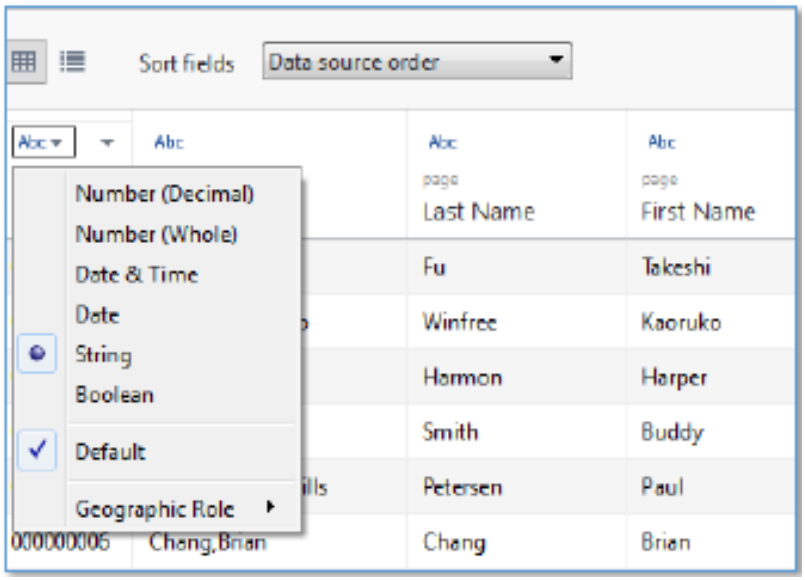
2. Choose the file you'd like to use and click Open. For this example, use Training Data Personal in the Introduction to Tableau folder on your desktop.

3. This will bring you to the Data Connection page. This page will appear any time you select a file, server, or saved data source on the start page.

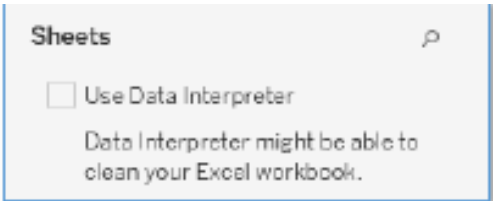
The left side of the page shows information about the data connection, including the data source type and tables within the data source.

The white box in the upper area of the page shows the tables that are being used and includes options for adjusting the join type and fields used, if any.

If you notice that data doesn't appear as expected on import, you can make adjustments to the definitions directly in the Metadata Grid. For example, an employee id might be created as a number or a date may have been imported as a string. You can change these to the appropriate definitions before going any further.



You can also turn on the Tableau Data Interpreter and it will attempt to interpret the data from your spreadsheet. When on, it will show warnings and a preview of its interpretations.



If you are connected to a Server data source, the data should already be properly defined. As a result, only the Geographical Role will be available for the field definition and the Tableau Data Interpreter will be turned off.

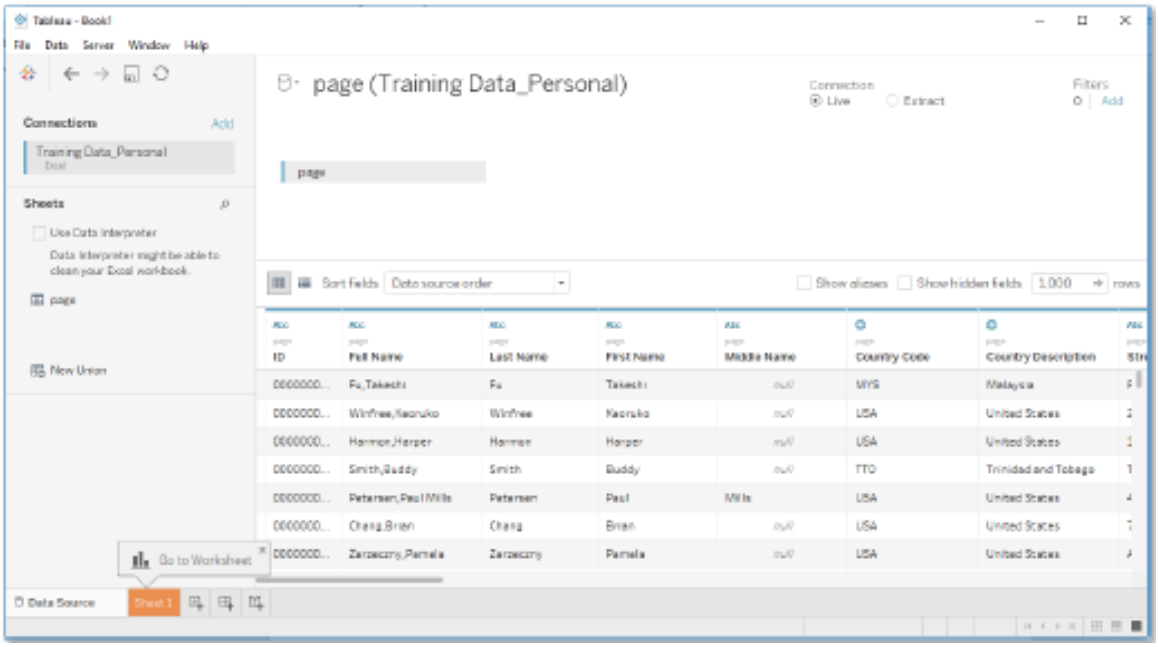
6. You can edit metadata information directly on the Data Connection page using the Metadata Grid. Alternatively, you can modify data attributes in the Data window or your worksheet.

Using the Metadata Grid, you can rename, copy, hide, manage aliases, and create calculated fields from the view. This view makes it easy to see all of the fields in a simple format and to make quick changes to the data source before going to the worksheet.

Below this, in the area toward the bottom of the page is where you can preview the data.

In the top area, under Connections, you can select whether to use a live connection or an extract.

In the top right area, under Filters, you have the option to add filters to the data source to restrict the data included.

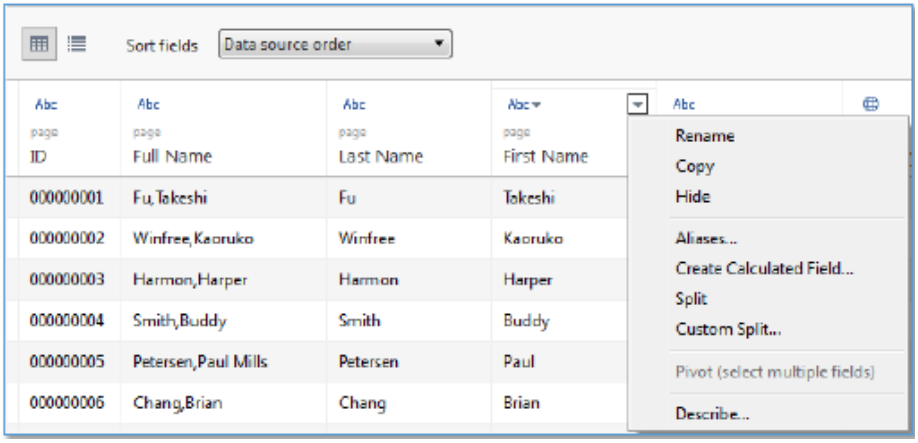


4. If you are opening a spreadsheet that has only one tab, you will see one sheet listed and the data connection will automatically be made. If there are multiple tabs, you will see a table listed for each tab. You'll need to double-click the sheet name for the connection you'd like to make.

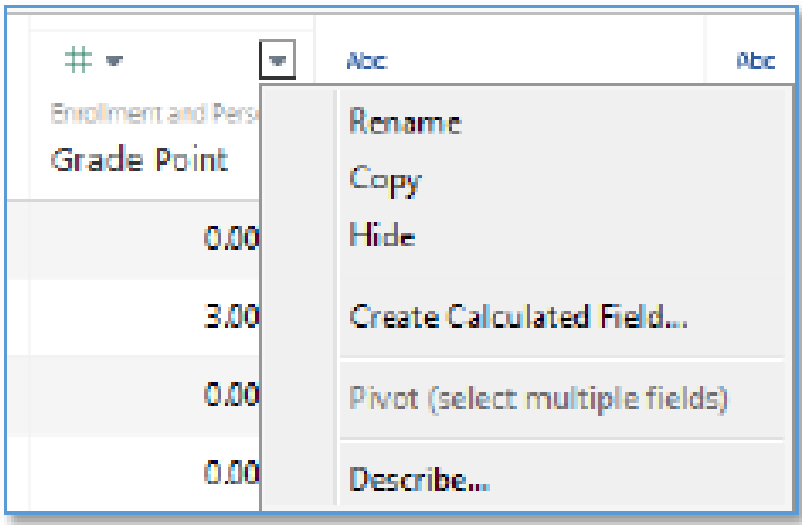
Note that the data connection is automatically set as Live. If you want to create an extract, you should change the connection type.

5. Tableau has several options to ensure your data is properly imported. By default, it will import your data with the source format.

Additional options from this view are Split, Custom Split, and Pivot. The menu options will vary depending upon the selected field type and your data source type.



The options for measures are limited to relevant actions including rename, copy, hide, and create calculated field.



Please note that if you are using a server data source, you will not be able to manage aliases and that option is removed.

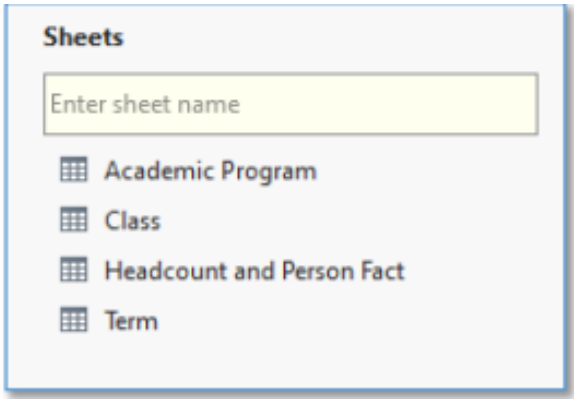
7. Once you have set your connection as desired, you are ready to start working with your data. At the bottom of the page, click the sheet under Go to Worksheet.

Connecting to an Excel File with Multiple Tabs

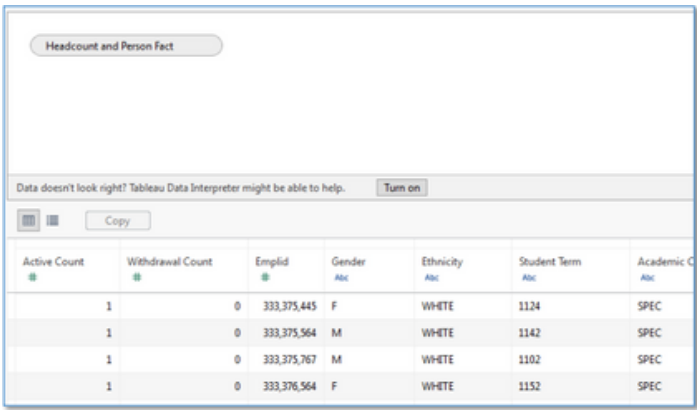
If the data you want to include in a Tableau visualization is included in a single excel file, Tableau provides the tools to set up the relationship easily. To give you some experience in our second example, we'll connect to an excel file with more than one table or tab.

1. In Tableau, on the File menu click New to open a new session. Click the Go to Start Page icon.
2. Click Excel to connect to an Excel file.
3. Select Headcnt_Fact and Dim combined.

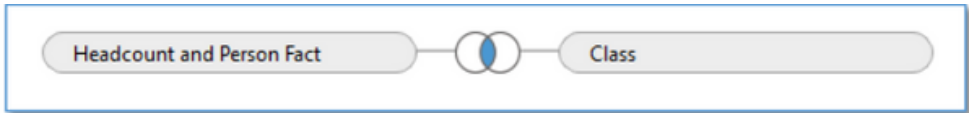
Notice there are now four tables from which to choose. Each of these represents a tab on the underlying Excel spreadsheet.



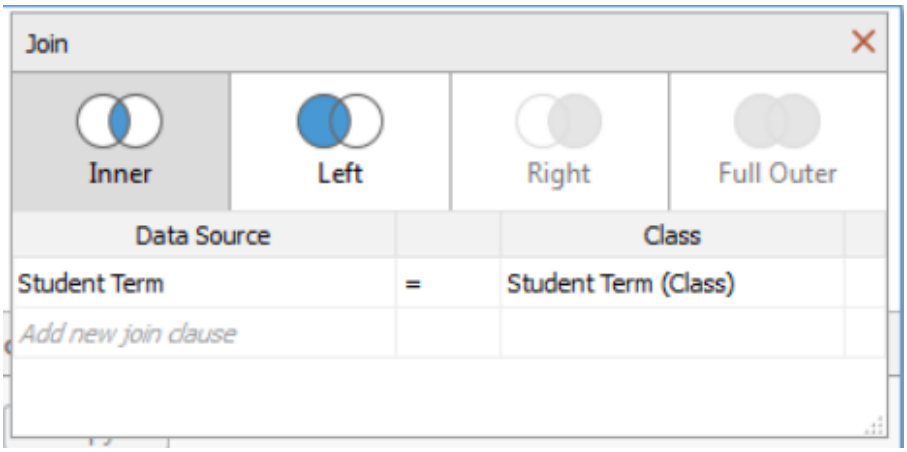
4. Double-click the table you want to use for your primary connection. We'll use Headcount and Person Fact. This will show the table in the upper area of the connection window and the details for the table into the preview pane.



5. Now select a second table to use. Double-click Class. Tableau will automatically select an inner join by default. This will work if you want only members that exist in both tables.

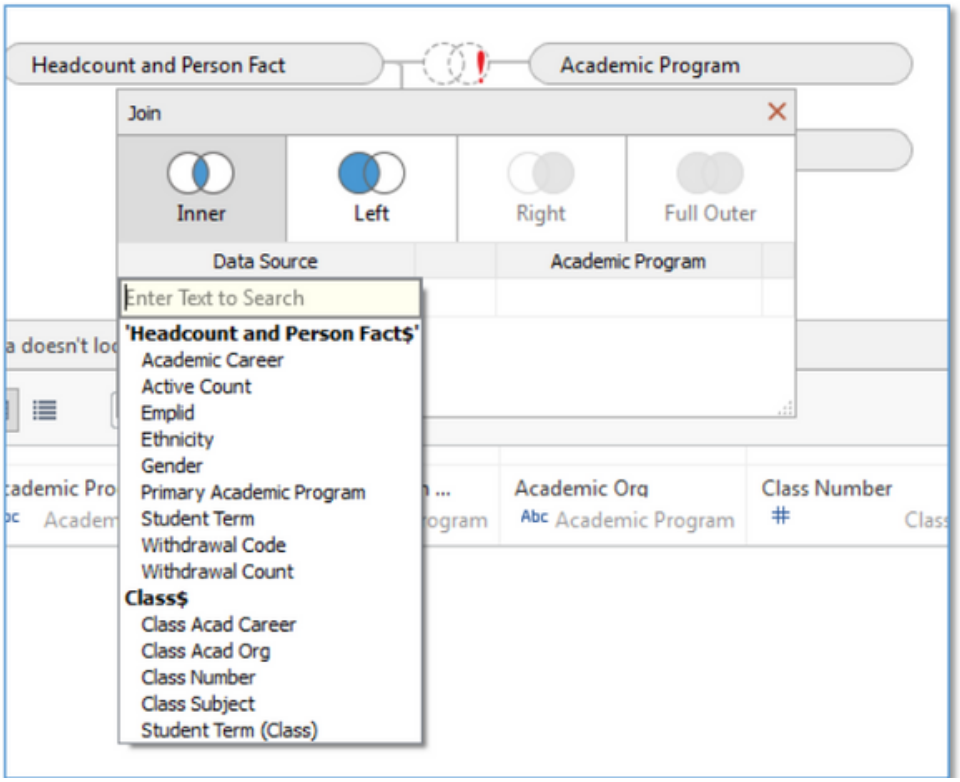


If you would prefer to drive the data in the visualization by members in your primary table, you'll want to change to join type. To do this, click the overlapping circle. This will open the Join window.
Here you have the option to select the join type you want and to edit the field(s) used in the join. Tableau will automatically try to join on matching field names. For our exercise, change the join type to a left join by clicking Left and then click the red X to close the Join window

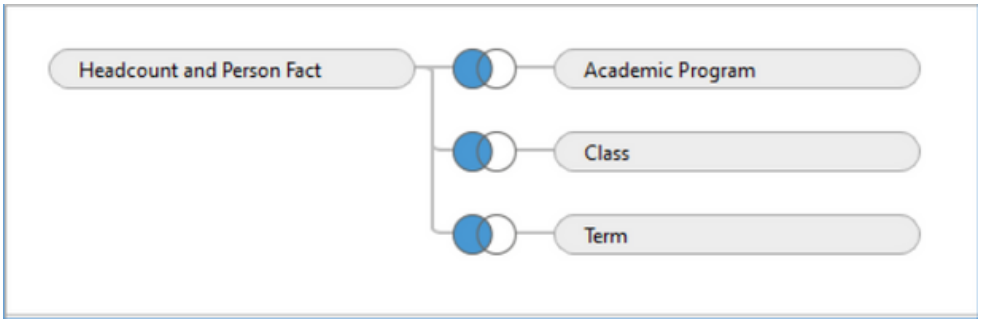


6. Next select a third table to use. Double-click the Academic Program. You'll notice several things. There is a red exclamation point indicating that the join has no join clauses.

The Join window is open by default asking you to identify the fields you want to use in the join. Click Primary Academic Program under the Headcount and Person Fact and Academic Program under Academic Program. Then click the Left join selection and the red X to close the Join window.



7. Finally, double-click the Term table and set the join type to be a left join as you did in the two previous examples.
8. When you're complete, your table connections should look like this.



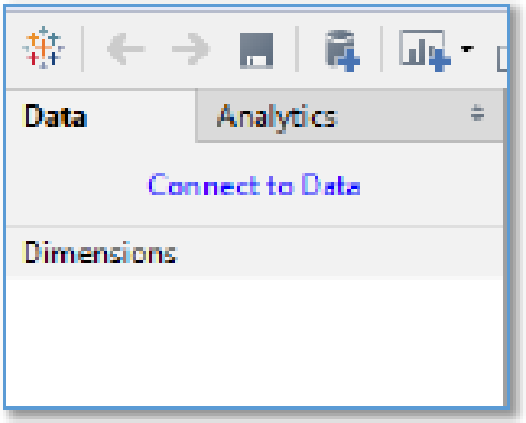
9. Once your data connections are set, click Sheet 1 to navigate to your first worksheet.
Your Data window is grouped by the Data Source table by default. Now when creating visualizations, you can include fields from all of the files.

Connecting to More than One Separate File

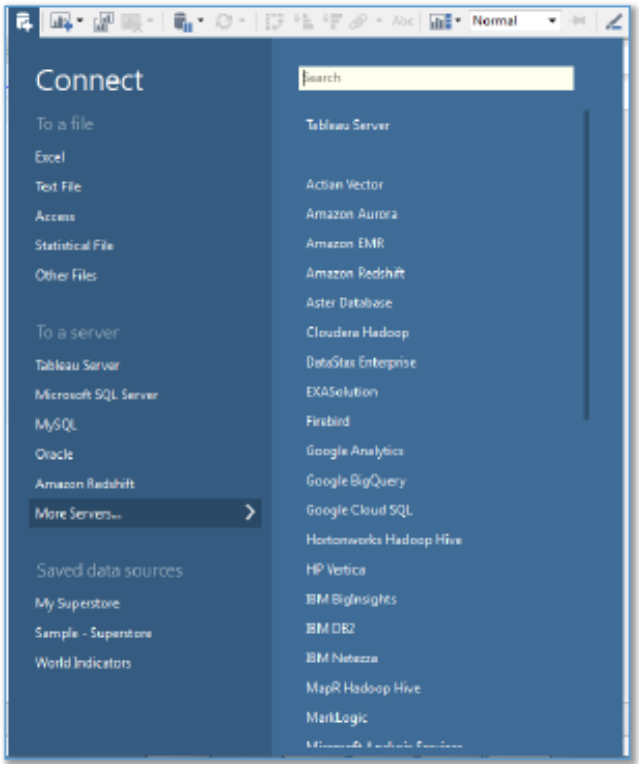
Data Blending

If you have files from more than one source you can still include these in the same visualization. You just need to take separate steps to join them in Tableau. When using a Data Blend, the secondary data is aggregated when joined to the first. You can practice this in our third example.

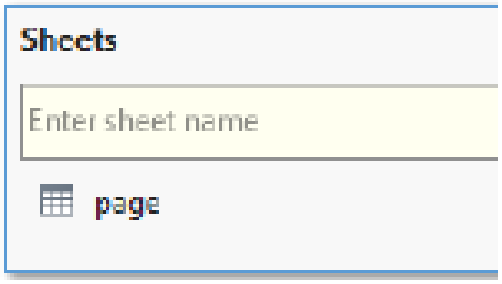
1. In Tableau, on the File menu click New to open a new session. This time we'll connect to our data sources directly in our Data window. Click Connect to Data.



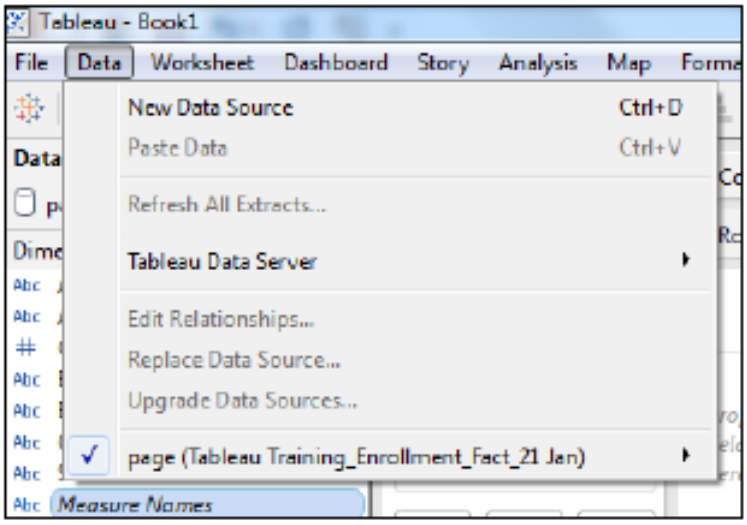
The Data Connection selection from the Start Page will open.



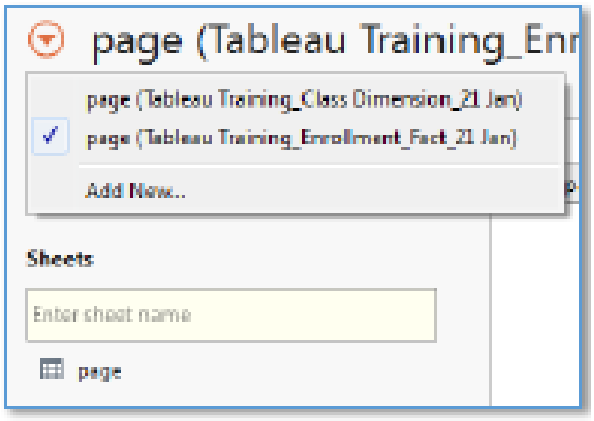
2. Click Excel and select Enrollment Fact.xlsx. This will bring you to the Data Connection page. Notice that as with our first example there is only one table available for selection.



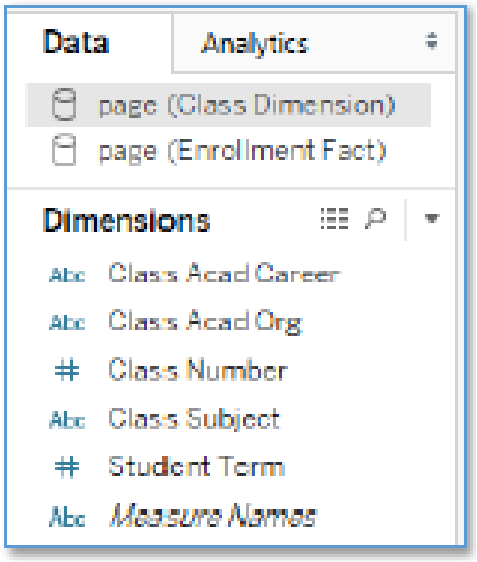
3. Click Sheet 1 to navigate to your worksheet.
4. In order to connect to a second table, on the Data menu select New Data Source.



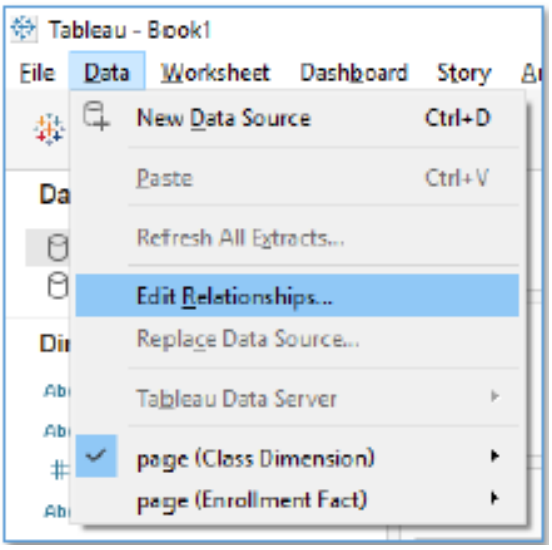
5. Select Excel and then choose Class Dimension.xlsx. This will bring you again to the Data Connection window. This time you'll notice that again there is only one table displayed but this time the preview shows the data from the Class Dimension file.
You do have the option to switch between data sources by clicking the orange arrow to the left of the data source name. However, you do not have the option to define the connections between the tables here.



6. Click Sheet 1 to return to your worksheet.
7. In the Data window notice that you now have two data sources shown, both Enrollment Fact and Class Dimension. The Dimensions and Measures will reflect the specific Data source you have selected.



8. You'll want to ensure that the two files have the correct relationships defined. To do this, click the Data menu and select Edit Relationships. Note: you can only create relationships for Dimensions. Since a relationship is required between Student Term on both files, ensure that this field is set as a Dimension.



The Relationships window will open. You can select the table you want to be primary in your join here. Tableau will automatically attempt to map the two tables based on matching column descriptions. You can choose to leave the mapping or make changes by adding, editing, or removing these selections by selecting the Custom radio button. For our example, we will accept the mapping that Tableau has suggested. Click OK to close the window.

9. Tableau defines a left outer join relationship when tables are joined in this way. When adding fields to a visualization, be mindful that you add fields from the table you want to be primary. Your results may change depending on which table is primary.

10. Drag the Dimensions Student Term and Class Number to Rows.
Enrollment Fact is now the primary file and will be indicated with a blue check mark.

Click your other file, Class Dimension, and drag the Dimension Class Subject to Rows to the right of Class Number. Notice that the field from the secondary file has an orange check mark next to it.

Additionally, the file will be indicated with an orange check mark in the Data window. The fields joined to the primary file will also have orange links next to them.

It's important to understand how the interaction between these tables can affect your results. Because of the left outer join, if you set a more restrictive table as primary your results will not include the full set of data from your other file.

11. To demonstrate this point, we'll create a second worksheet but bring the fields to our visualization in the opposite order. Click the New Worksheet icon to the right of Sheet 1.

12. First drag Class Subject from Class Dimension to Rows. Notice that the blue check mark is now next to this table indicating that it is primary. When you click Enrollment Fact, notice that the two links are shown as broken.

13. Click a link to activate it. This will set Enrollment Fact as secondary and will show an orange check mark next to the file name.

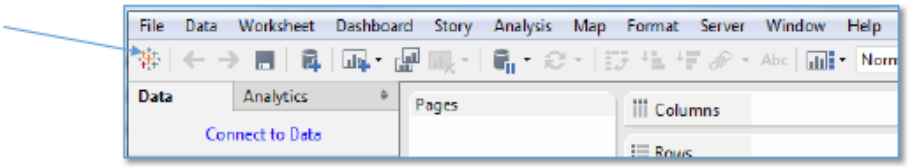
14. Drag the Dimensions Student Term and Class Number to Rows to the left of Class Subject. Notice this time the secondary file indicator is shown against these two fields.

15. Compare the two worksheets and notice the differences in the data. Even though they have the same fields in the same order, because of the join limitations they are showing different results. As you're building your visualizations, it's important to understand these mechanics in order to ensure you are delivering the expected results.

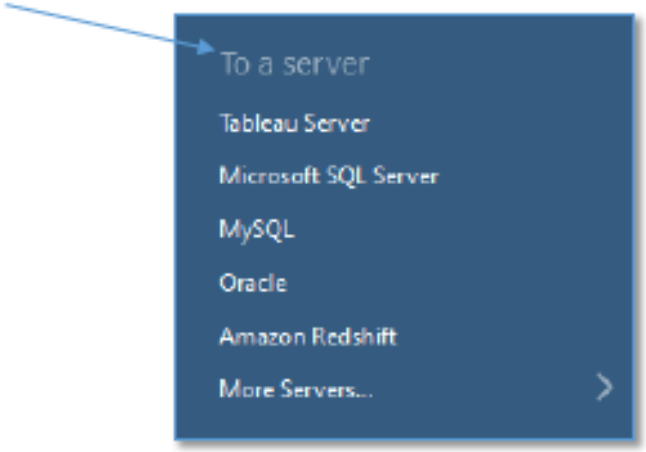
Connecting to a Server Data Source

For our next example, we'd like for you to get some practice using a data source published to the server.

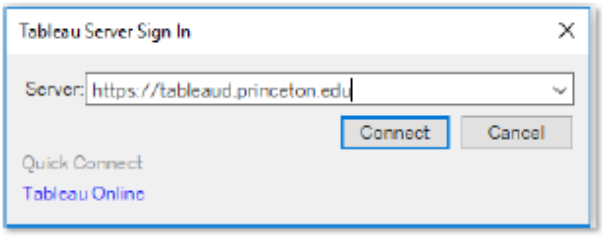
1. In Tableau, on the File menu click New to open a new session. Click the Go to Start Page icon.



2. Now select your data source. Under the To a Server section, select Tableau Server.

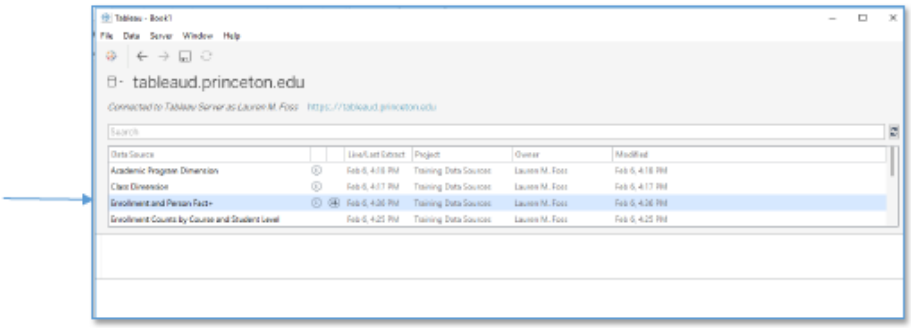


3. As this is a new session, you will need to log into the server.

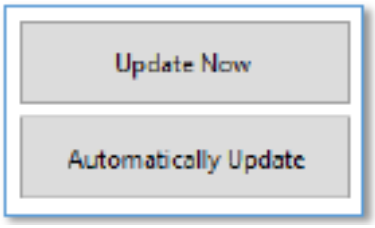


Enter your username (netid) and password. If you are authorized to use multiple sites, you'll need to select the site next. We will be using Tableau Training for this course.

4. Once logged into the site, you will see a list of available data sources. Choose the source you would like to use. We'll be using Enrollment and Person Fact + for this exercise.



5. Once your data source has been selected, the default connection type is Live or Extract. We will leave this setting as Live to allow the worksheet data update as the source on the server is updated. You will have the option of setting the refresh timing – Now or Automatically Update.



Choose Automatically Update to allow the worksheet update as the data on the server is updated.

6. When you are ready to start work, click the sheet under Go to Worksheet.

Editing and Saving a Data Source

One of the benefits of defining a data source in Tableau is that you can customize the data attributes, or metadata, for the source. For example, you might change a field's name or default properties. You might move a field from Dimensions to Measures or vice versa. You can add parameters, calculated fields, groups, hierarchies, bin, or sets. These customizations are retained while the source data table is left unchanged.

In order to review full functionality, connect to Training Data Personal in the Introduction to Tableau folder on your desktop. Once connected, click on Go to Worksheet at the bottom of the page. The following actions take place in the Data Pane on the left side of the Worksheet.

Modifying Data Attributes

There are many options for modifying your data attributes. In addition to changing the name or alias, you might want to change groupings or default options.

Organize Your Data

One thing you might choose to do is organize your Dimensions and Measures with Folders.

1. In the Data window, right-click the white space below the fields. Ensure the Group by Folder option is selected. The default is to Group by Data Source Table so you may need to change the selection.
2. In the Data window, right-click in the white space below the fields and choose Create Folder.
3. Add a name to the Create Folder dialog box, and click OK. For example, you could create an address folder to hold the various address fields.
4. Drag and drop the desired fields into the new folder. You can select multiple fields at a time using either shift + click or control + click.

Change Measure and Dimension Classification

Another thing you might change is the classification of a field from Dimension to Measure or vice versa. For example, when importing the data source into Tableau an ID field may have been interpreted as a Measure even though it is really a Dimension due to the numeric content. If this is the case, simply click the field and drag it from Measure to Dimension.

Set Default Properties for a Measure

You can customize the default properties for a measure to be different from the Tableau defaults.

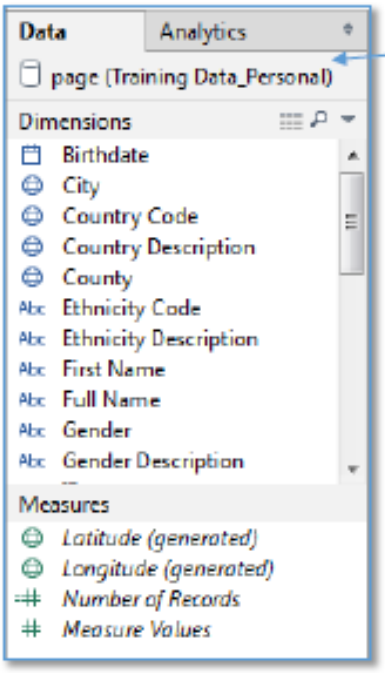
1. Hover your mouse over the Measure you want to modify, click the down arrow displayed at the right of the field name. Then choose Default Properties.
2. You can choose from Comment, Color, Number Format, Aggregation, and Total Using. Note that your options will vary depending upon the field type.
3. Change the properties as desired according to the available options.

Set Default Properties for a Dimension

1. Hover your mouse over the Dimension you want to modify, click the down arrow displayed at the right of the field name. Then choose Default Properties.
2. You can choose from Comment, Color, Shape, and Sort. Note that your options will vary depending upon the field type – date fields will also include Date Format and Fiscal Year Start.
3. Change the properties as desired according to the available options.

Saving Your Data Source

If you plan to use the data source again, you can save the data source in order to leverage these changes. They will be retained even when the underlying database or spreadsheet is update.



1. In order to save the file for personal use, on the Data pane right-click the connection you want to create as a data source and choose Add to Saved Data Sources. Changes in the data source are not inherited by workbooks created prior to the changes being made.

2. For group sharing, you will publish the data source to the Tableau server. On the Data window, right-click the connection you want to share. Choose Publish to Server. When published to the server, subsequent workbooks can inherit the data source changes.

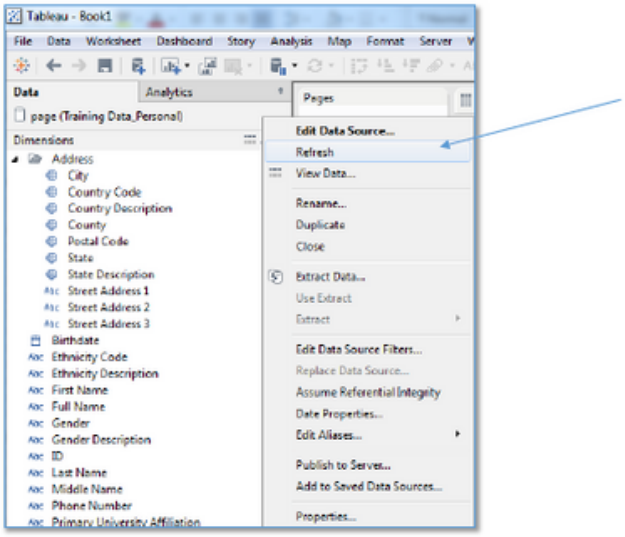
Understanding Changes to Data

When the underlying data in your visualization changes, your visualizations may be affected in different ways depending on how you connect to the data in Tableau and on the changes taking place in the underlying data.

Connection Type

If you are using a live data connection, changes to the underlying data will be reflected when you open the visualization or refresh the data connection in a visualization that is already open. If, however, you are using an extract changes made to the data will not be reflected in the visualization until you refresh the extract.

If you are using an extract and want to refresh the data, right-click the data source connection in the Data pane and choose Refresh.



Data Change Type

If the changes made to the data are only in the values of the data, these changes are reflected in the values displayed in your visualizations but will not break the visualizations. However if the changes are made to the structure of your data, such as removing or renaming fields, visualizations using these changed fields will break.

In the case where a field has been changed and results in a broken visualization, the field will appear in red where it is used with an exclamation point in the Data window. The visualization will be greyed out and cannot be interacted with until the field in error is fixed or removed. New field descriptions or new fields will be added to the Data window.

Filtering

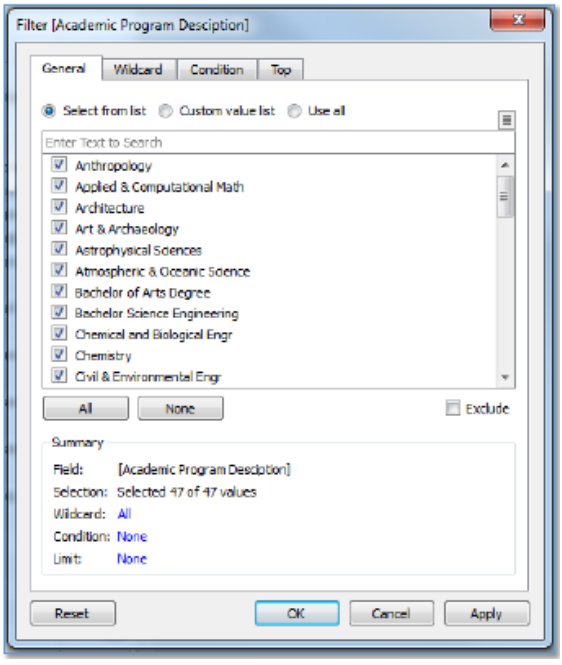
Filtering allows you to narrow the data shown in a view and increases focus on relevant information. There are several ways in which you may add filtering within Tableau. These include using the Filter Shelf, Interactive (Quick Filter), and filtering within a View. The filtering options on both the Filter Shelf and Quick Filters will vary depending upon the field type (dimension, measure or date dimension) chosen. Filtering directly within a view allows one to narrow the view by keeping or excluding the selections. Whichever method chosen, filtering is helpful when focusing a view on specified data.

Filtering on a Dimension

A dimension filter is used to view the data values of a smaller set of the dimension's members.

We'll explore the creation of a dimension filer here using the server data source Headcount and Person Fact +.

1. Drag a dimension, Academic Program Description, to Filters.
2. When the dimension has been placed on the Filters shelf, the Filter dialog box will automatically open.

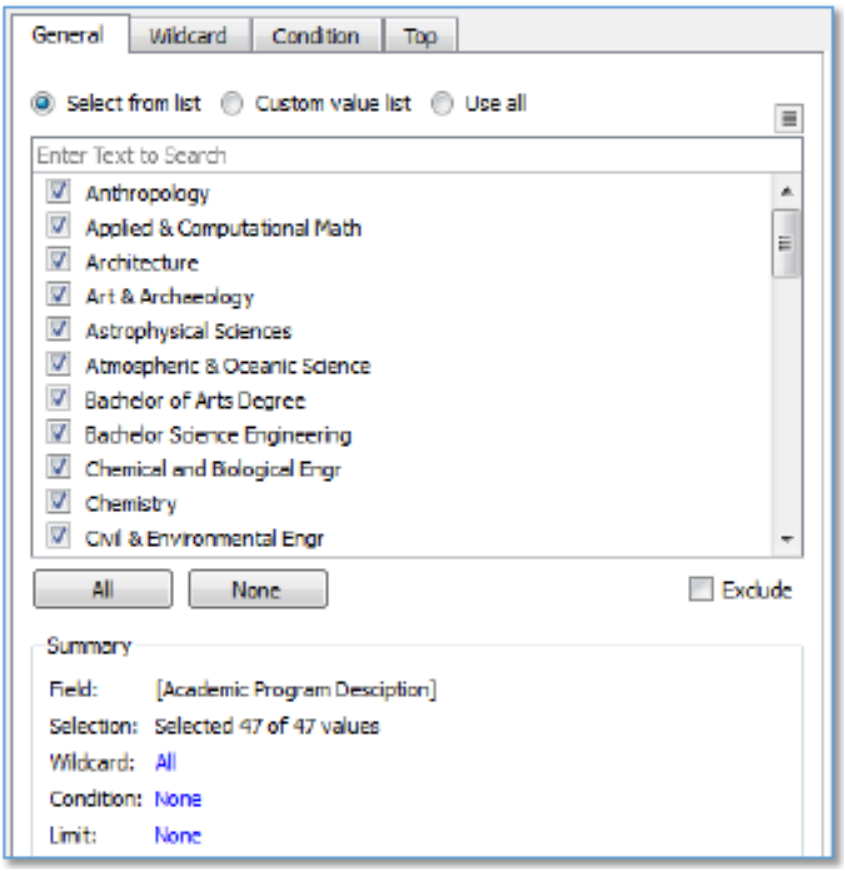


There are several filter option tabs that can be completed. It should be noted that these are cumulative. Settings from each tab are considered with "AND" logic so will further refine your results. Clicking on Reset on any tab within the Filter window will reset all selections (not just the values on the current tab) to their original values.

The four tabs include the following options.

General

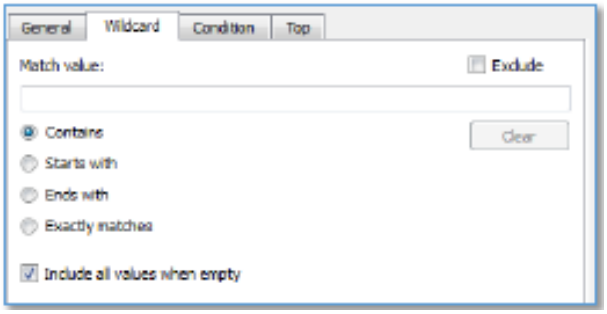
This option shows the members of the selected dimension you can select for inclusion or exclusion. You can also create a custom value list.



At the bottom of the General tab is a summary of all conditions included in the filter. Each has a hot link that will take you to the relevant tab (Wildcard, Condition, or Limit).

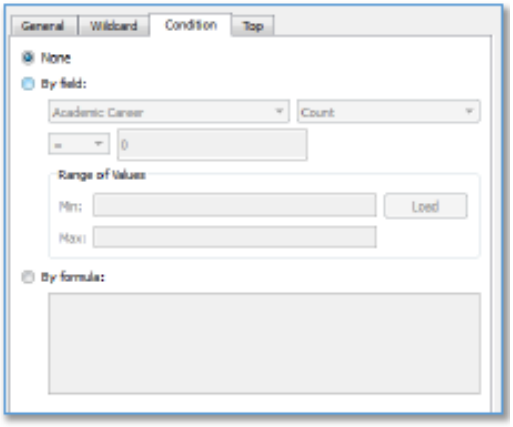
Wildcard

This option sets up a wildcard inclusion or exclusion of dimension members matching the value entered.



Condition

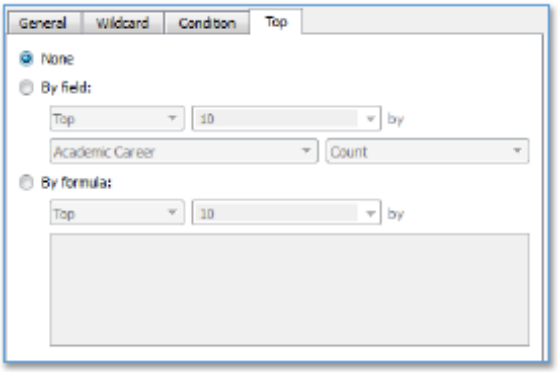
This option filters values based upon specified conditions including a field, range of values or specified formula.



Top

This option filters by the top or bottom “N” where N is determined by the value of specified fields or formula. For example, you could say the top 10 Academic Programs by Grade Point Average Median.

3. Once you have made your selections, click OK.

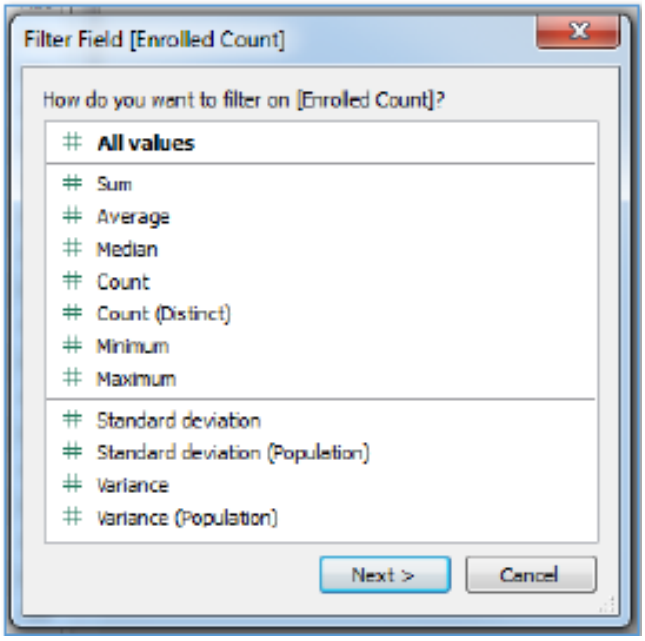


Filtering on a Measure

A measure filter is used to show only the values that meet your filter criteria.

Let's explore the creation of a measure filer here using the server data source Enrollment and Person Fact +.

1. Drag a measure, Enrolled Count, to Filters.
2. When the measure has been placed on the Filters shelf, you will be asked to select the method of aggregation for the measure. After making your selection, click Next.



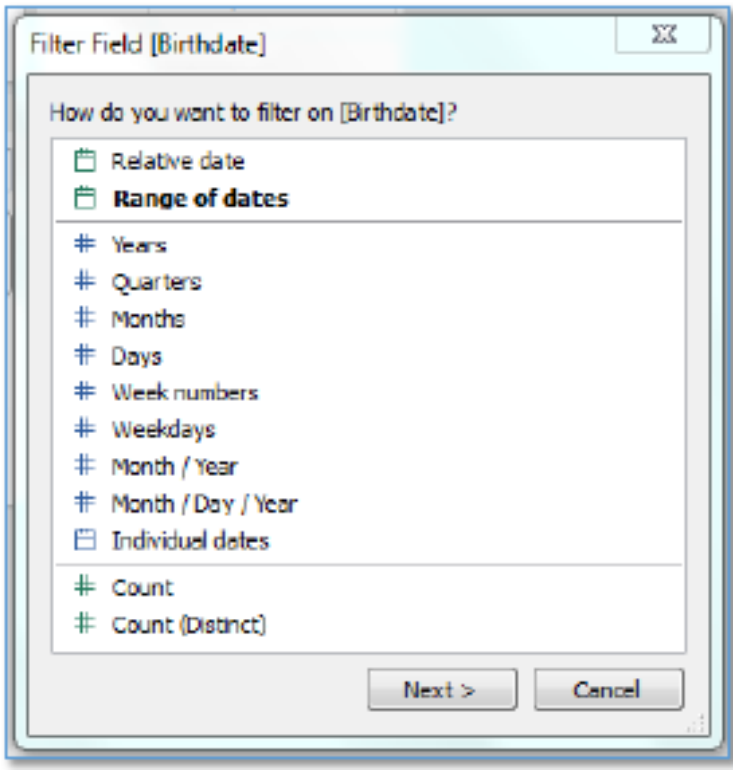
3. Whichever selection you make, the Filter dialog box again offers several different options for filtering.

Creating Date Filters

A date filter is used to filter a data subset for specific date or time criteria. You can choose to filter for a specific range of dates or for a discrete date/time.

Let's explore the creation of a date filer here using the Training Data Personal spreadsheet.

1. Open a new Workbook and select the Training Data Personal spreadsheet from the Introduction to Tableau folder on your desktop.
2. Drag a date, Birthdate, to Filters.



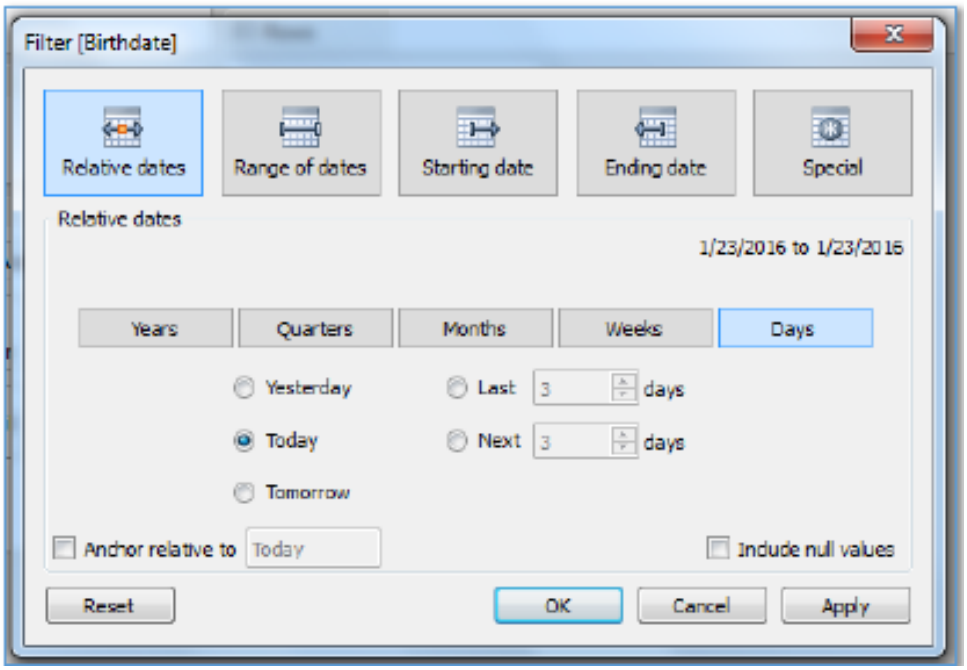
3. In the Filter Field dialog box, select the type of date filter you want to use – Relative date, Range of dates or one of the discrete or continuous options available.

The different date filter options are described below.

Relative Date

Use a relative date filter to update data dynamically with time. For example, set your filter to show data for 6 months before or after the current date. The data shown in your view will be updated daily.

In the Filter Field dialog box, choose Relative date and then click Next. In Filter dialog box, choose the time unit you want to use for the relative date range (Years, Quarters, Months, Weeks or Days)



Next, set the specific time to target (Previous, This or Next) based upon the time unit you've selected. You can also specify Last N or Next N.

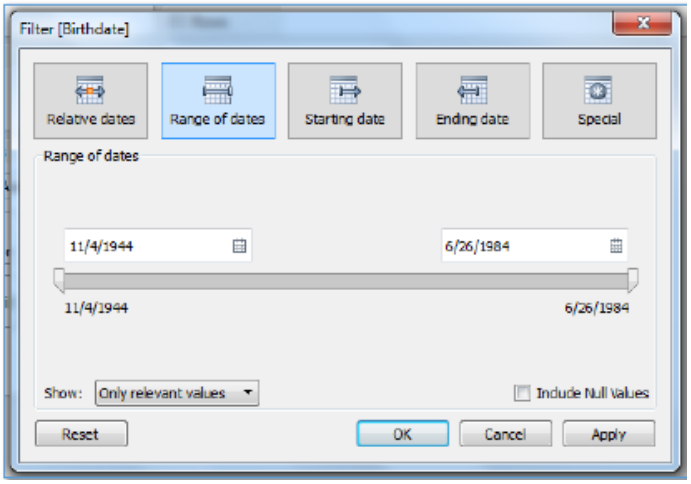
You can also anchor your filter to a specific anchor date by selecting Anchor relative to and setting a specific date.

Once you've completed your selections, click OK.

Range of Dates

Use the range of dates filter to specify a span of time. You can also specify a starting or ending date.

In the Filter Field dialog box choose Range of dates and then click Next.



Next, set the specific start and end dates to target. To do this you can either use the slider or select the dates from the drop down menus. Once these have been set, click OK.

Starting Date or Ending Date

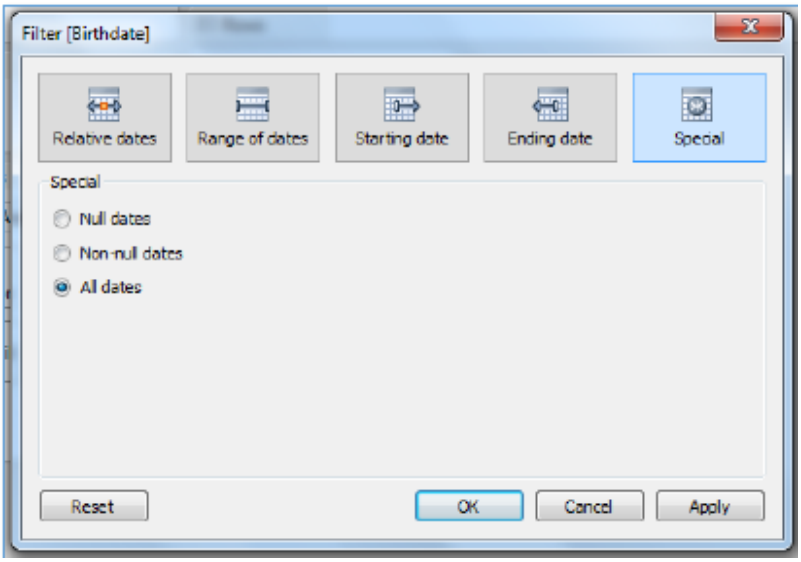
Both Starting Date and Ending Date filters originate with the selection of Range of dates in the Filter Field dialog box. Once selected click Next.

In the Filter dialog box, choose the Starting date or Ending date.

Once set, click OK.

Special Date Filter

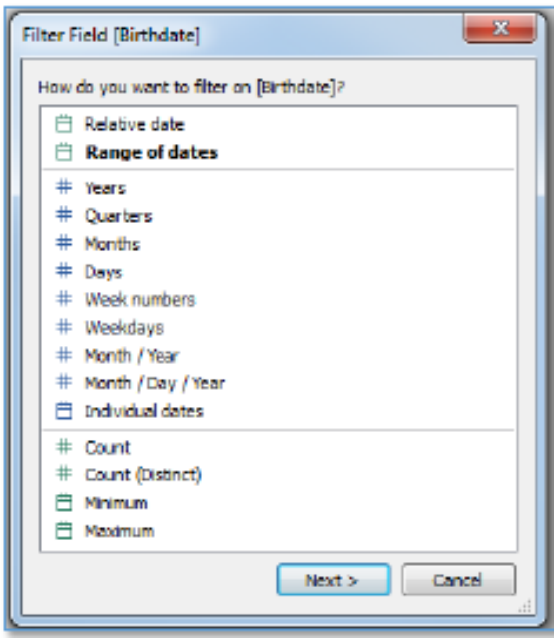
You can also set a filter to see null dates or non-null dates.



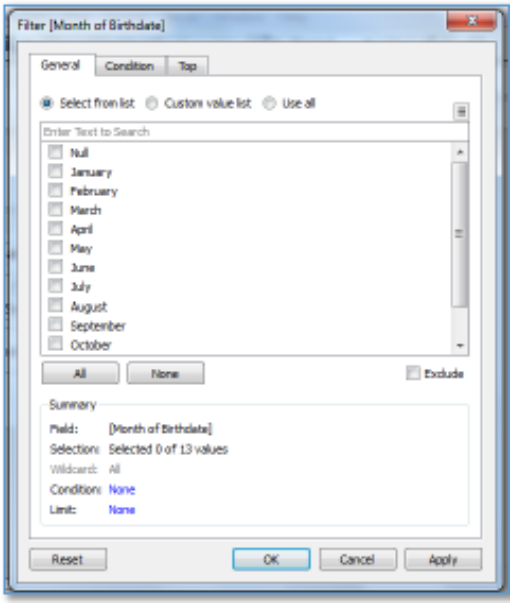
Other Date Filters

Following these same steps you can also set filters for the discrete portions of the selected date (those options shown in blue).

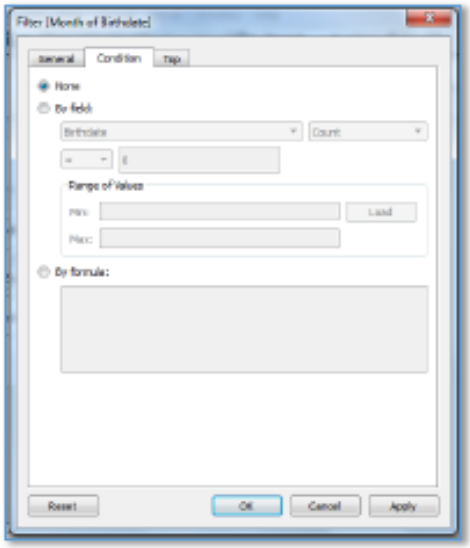
In the Filter Field dialog box, select the date part you want to use as a filter and then click Next.

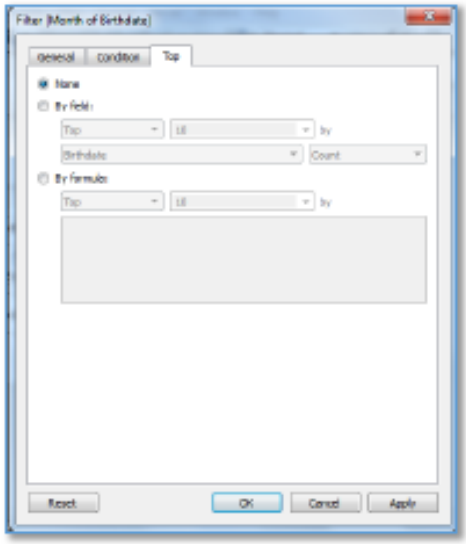


In the Filter dialog box, select the members you want to use as a filter. This is a list based upon your selection in the first dialog box. For example, if you choose Months, the months of the year are displayed.



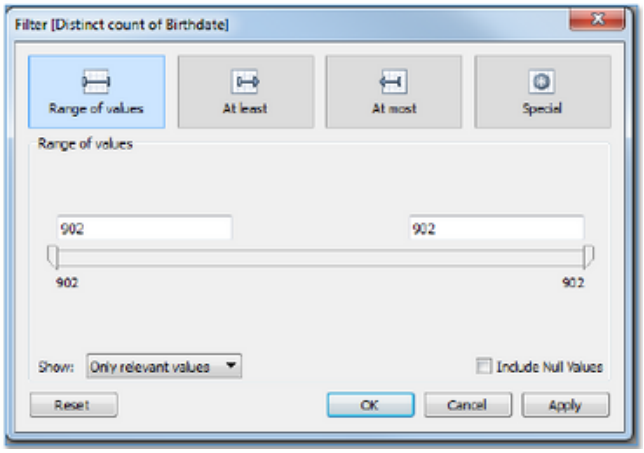
Similar to the dimensional filtering, you can set the filter on General, Condition, or Top (or Bottom) selections. These are shown below.





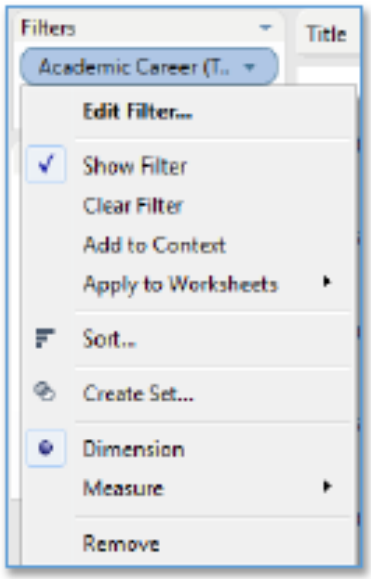
Once you have made your selections, click OK.

Alternatively, you can also filter based upon continuous date values (those in green). When making this selection, the filter options are the same as when you choose to filter based on a measure including Range of values, At least, At most, and Special.

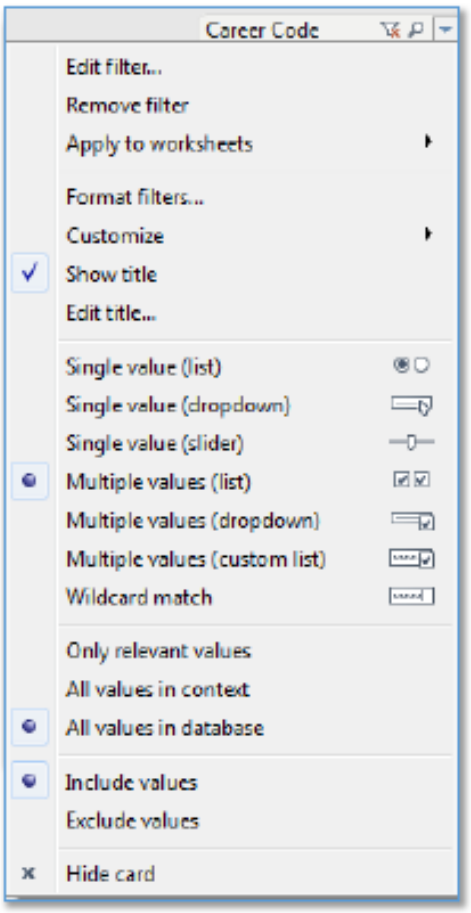


Quick Filters

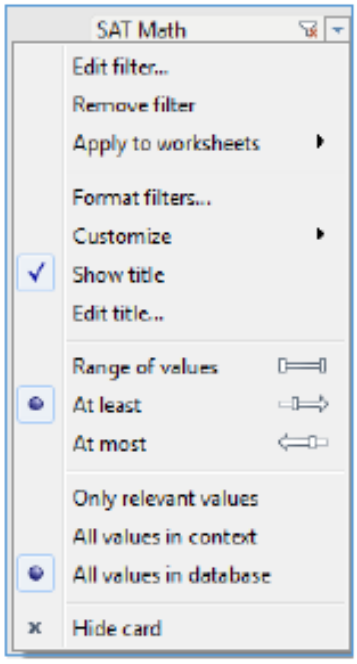
Quick filters allow your report users to manipulate the data in the view. Once a filter has been added you can easily add a quick filter by right-clicking on the field on the Filter shelf and choosing Show Filter.



Once a Quick Filter has been added, you can customize the appearance of the filter by clicking on the drop down arrow of the filter and making the desired modifications. As with filter options on the Filter Shelf, the options for customization the Quick Filter vary depending upon the data type.

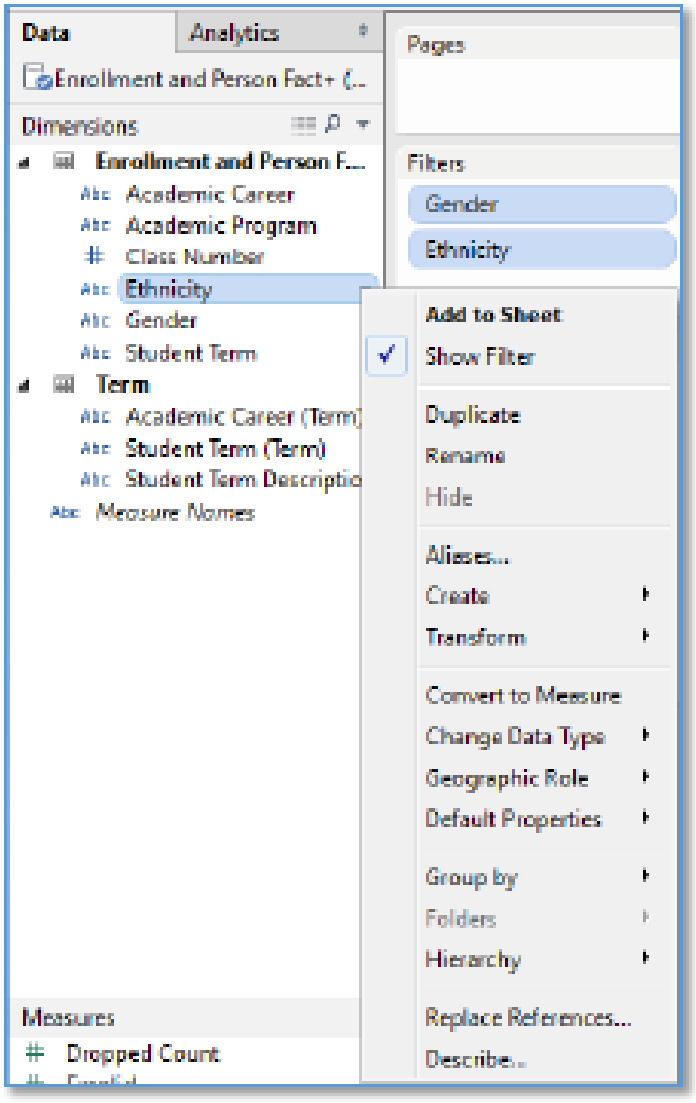


Dimensions include selection options for customization including Single or Multiple Value (List), Single or Multiple Value (Dropdown), Single Value (Slider), and Wildcard Match. Measures, on the other hand, include various range selections including Range of values, At least, and At most. Both options allow for selections to include Only relevant values, All values in context, or All values in the database.



Quick filters for dates are very similar to both Dimension and Measure filters depending upon whether discrete or continuous date values have been selected.

An alternate to quickly adding a filter and showing the quick filter is to right-click directly on the field in the Data pane. Choose the option for Show Filter. This will add both the Quick Filter and the field on the Filter shelf. Taking this option, however, automatically adds the field with all values selected. Any customization would need to be performed by editing the filter either through the Filter shelf or the Quick Filter.



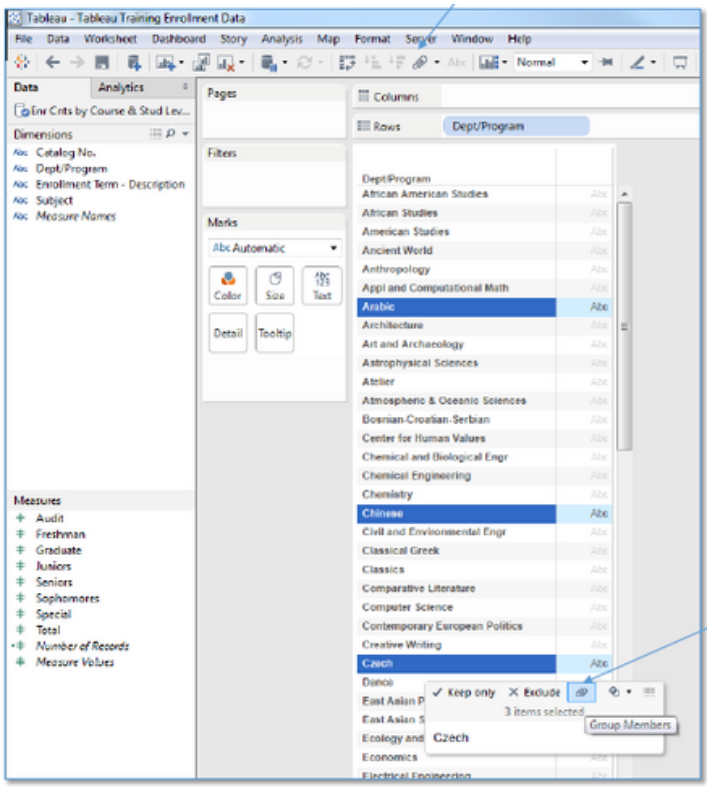
Groups

Groups are useful in displaying data when there are large numbers of dimension members or when you find the need to combine dimension members into a single set. For example, when creating a visualization of enrollments, it might be useful to combine the departments into groups such as Languages, Natural Sciences, Physical Sciences, etc. Similarly, if your data set has elements that might mean the same thing such as “NJ” and “New Jersey” you might choose to create a group including members of both dimensions that would present a more logical display. Whichever the reason, groups are easily created within Tableau and once created are available as a Dimension for selection just like any other Dimension derived from a data source.

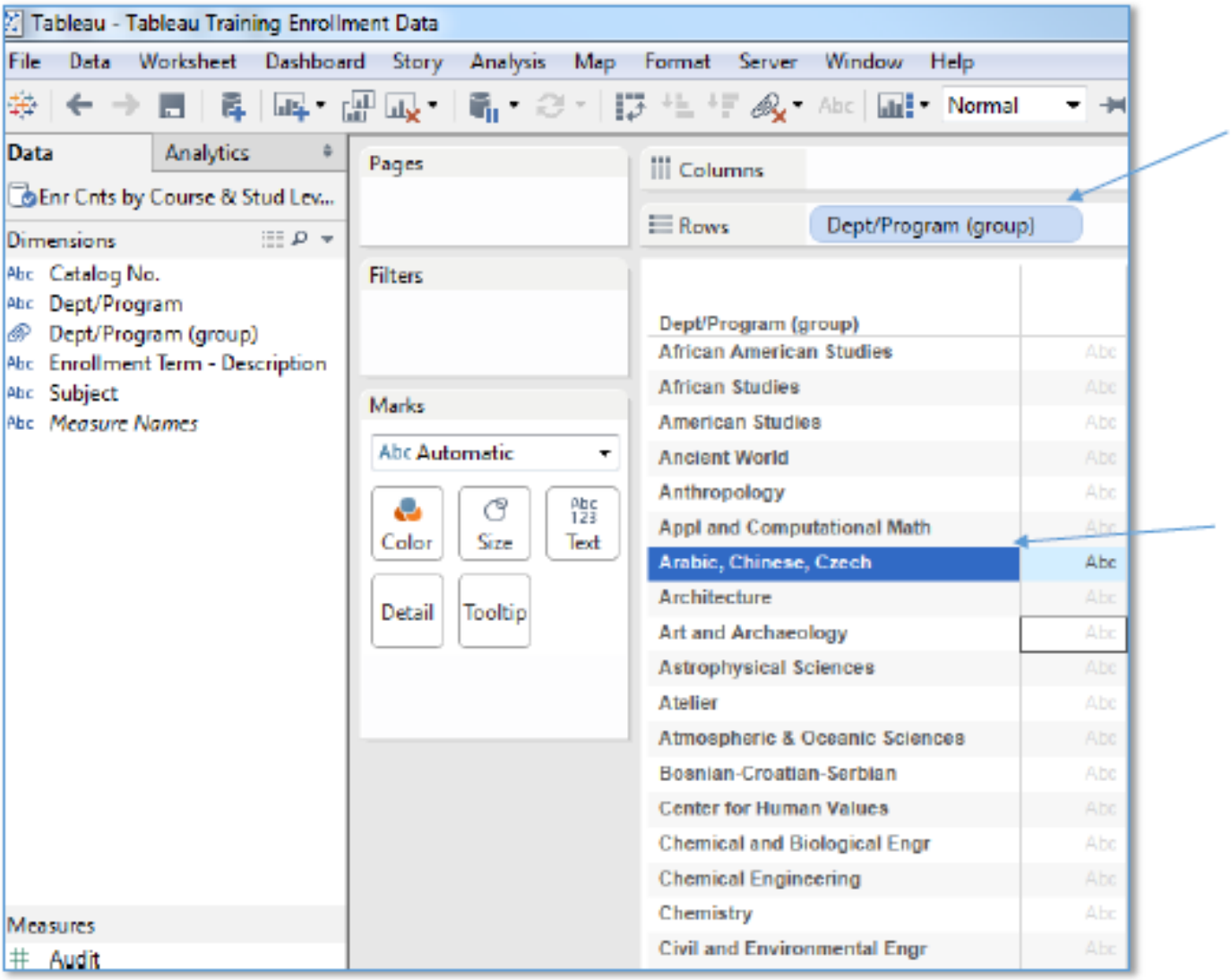
There are three simple ways to create a group: creating a group directly in a View, from the Data Pane, or by using a Visual Grouping.

Create a Group Directly in a View

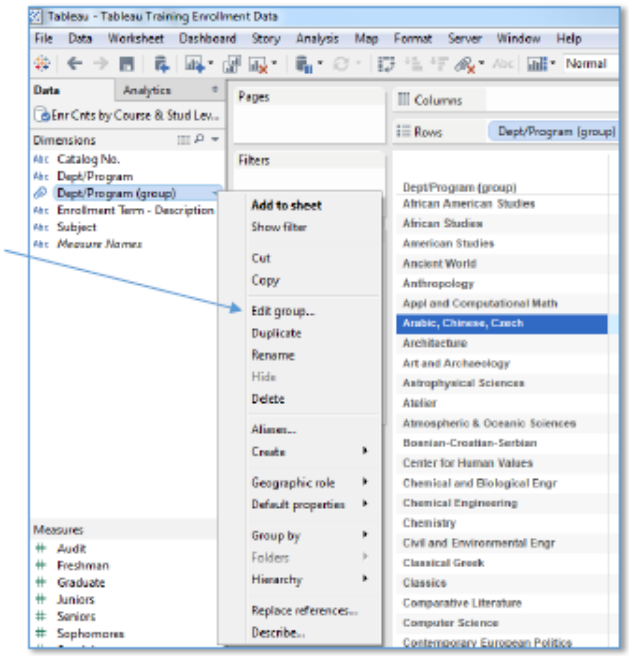
Select Dimension labels you would like to combine using either SHIFT+click or CTRL+click to select the dimension labels to be grouped. Use the Group Icon on the Toolbar or by right-clicking.

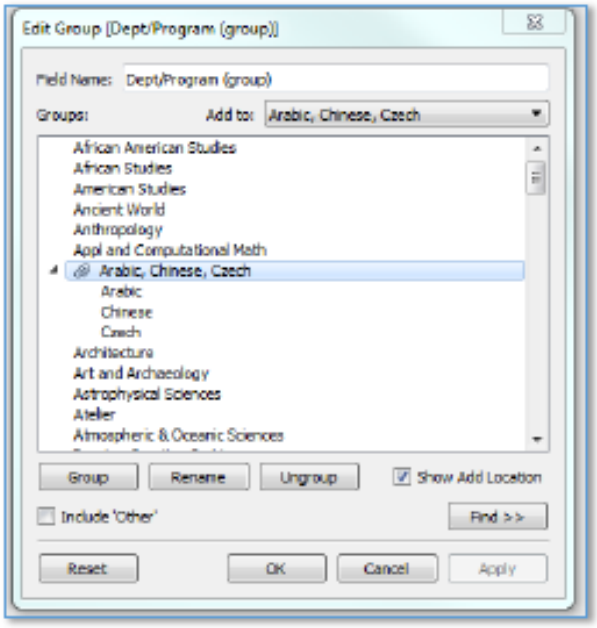


The new group will appear in the Data pane and the selected Dimension members will now be grouped.



In order to add or remove group members, right-click on the group in the Data pane and choose Edit Group...





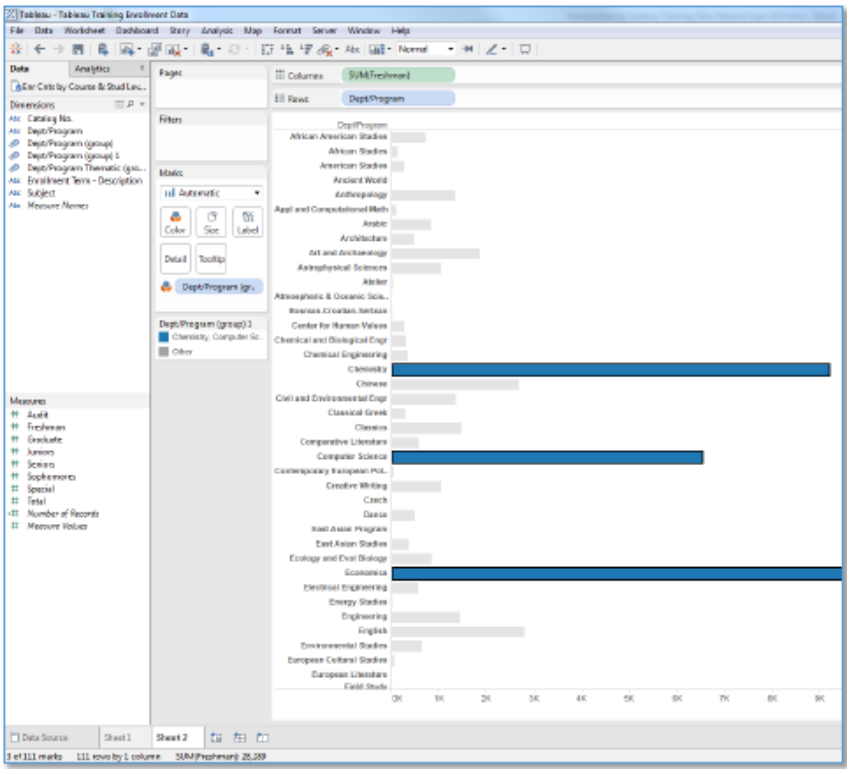
Create a Group from the Data Pane

Alternatively, you may create a group by starting from the Dimension in the Data pane directly. Start by hovering your mouse over the Dimension with which you'd like to work. An arrow will be shown to the right of the Dimension name. Click on this down arrow and select Create and then Group…

The Create Group Dialog box will be shown. Select the dimension members you'd like to group and select Group. You may customize the appearance by renaming both your groups and the name of the grouping as it will appear in the Data pane.

Create a Group Using a Visual Grouping

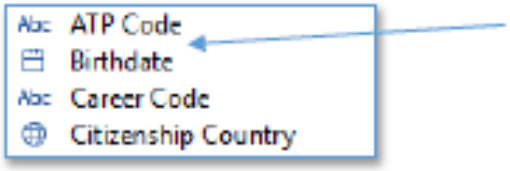
Select a group of Marks in the view and click the Group icon on either the toolbar or the tooltip. Selecting group members in this way allows you to group on a single dimension and show items as members of a group without losing their labels or individuality.



When creating the group in this manner, the new group is placed on Color and all other Dimensions in the view are designated as "Other."

Dates in Tableau

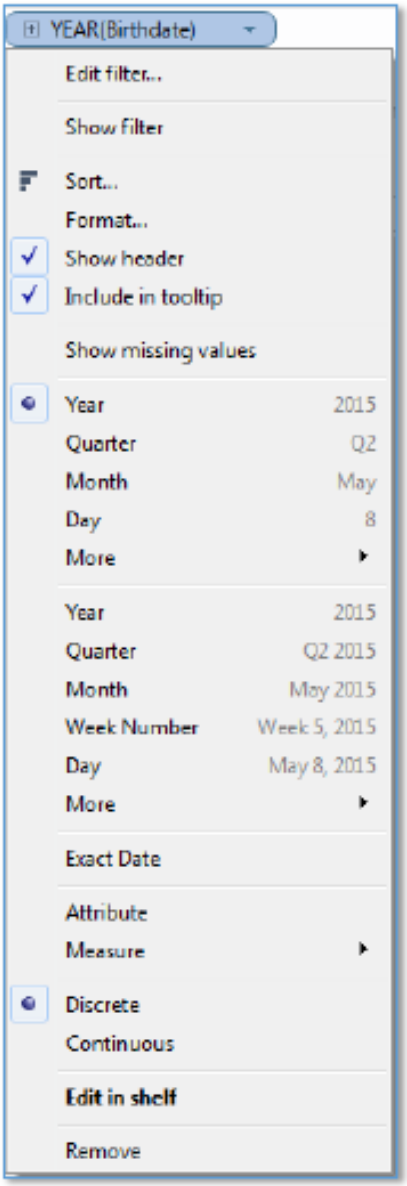
Tableau automatically defines dates and times as Dimensions for relational data sources. They are identified in the Data Pane by a calendar icon.



When brought to a shelf in a view, the field name automatically reflects the default date level reflecting the level of multiple date instances. What this means is that if the dates in your data source span several years, the default date will be year. However, if all dates are within the same year or month, the default level will be months and days respectively. The level may be changed by hovering over the date on the shelf, clicking on the down arrow, and selecting the desired level (year, quarter, month, day, etc.).

Discrete Dates

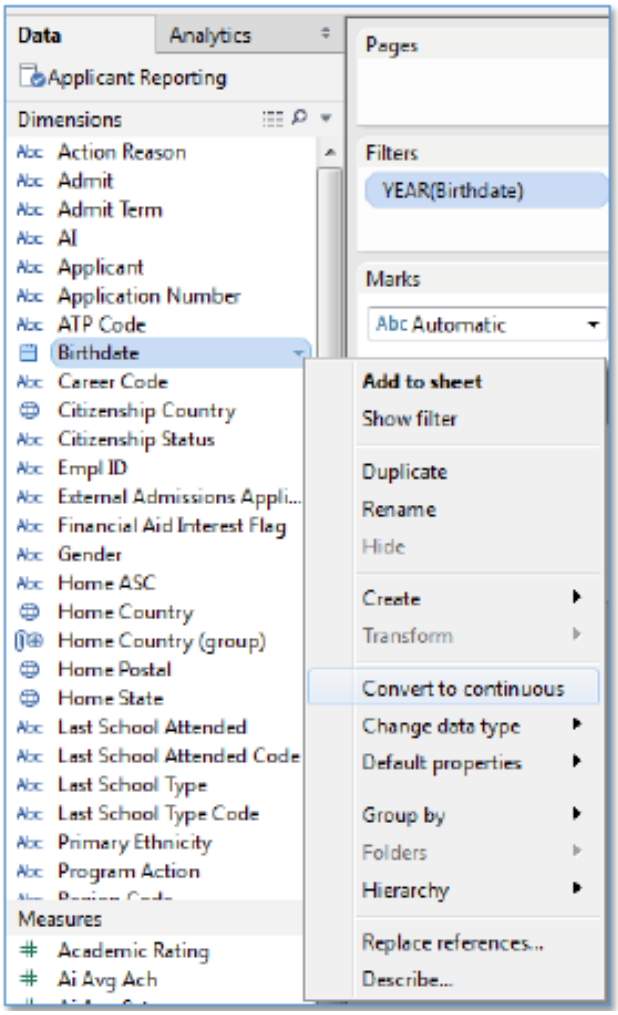
Continuous Dates



Discrete vs. Continuous Dates

Dates are characterized in one of two ways within views. They may be apportioned as discrete units within a defined category (e.g. October) independent of time. Alternatively, they may be representative of a continuous progression of time. The categorization selected will affect the way in which the date behaves and the options available when added to a view.

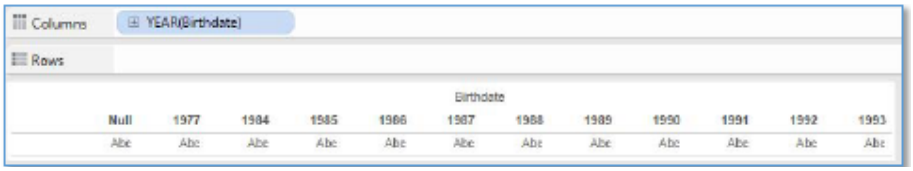
Discrete dates are unique in that the individual parts (year, quarter, month, and day) are automatically set up in a hierarchy. These parts are independent of linear time. As such, these parts may be drilled into and out of and arranged on a view in the order desired on a single shelf or even on different shelves.



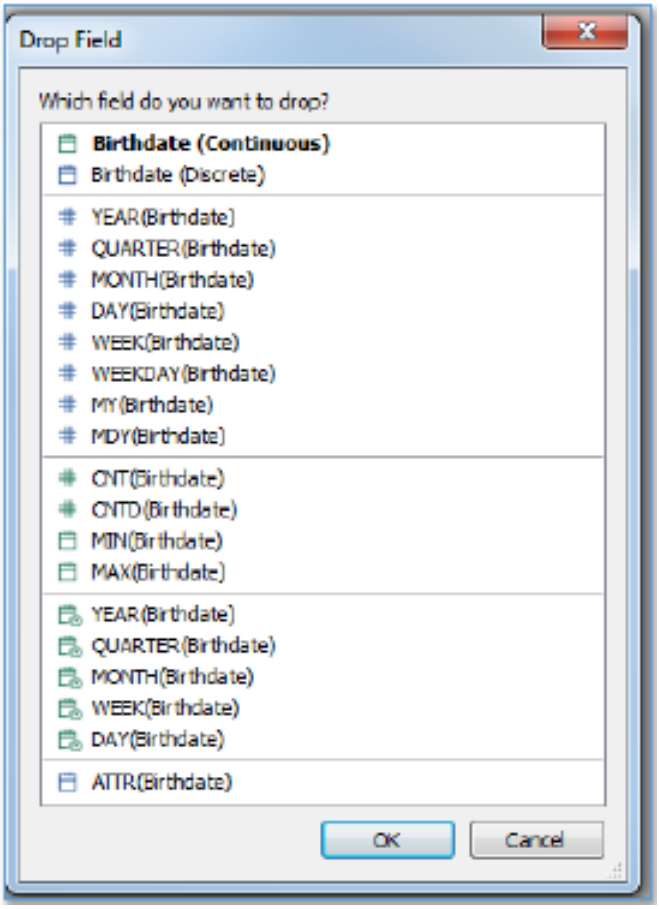
Continuous dates always reflect the progression of time. Though the level may be narrowed (e.g. from years to quarters to months, etc.), these are not distinct entities as they are with discrete dates. Once a level has been selected, the view will reflect the corresponding level of detail. One may drill in but not out of the continuous date detail.

By default, Tableau will define date and date/time fields as discrete. This setting can be changed by clicking on the arrow next to the field name and selecting Convert to continuous.

When dragging the date field onto a view, the resulting Pill type will reflect the default type (continuous or discrete) that has been set. As with all Tableau indicators, green indicates continuous and blue discrete.



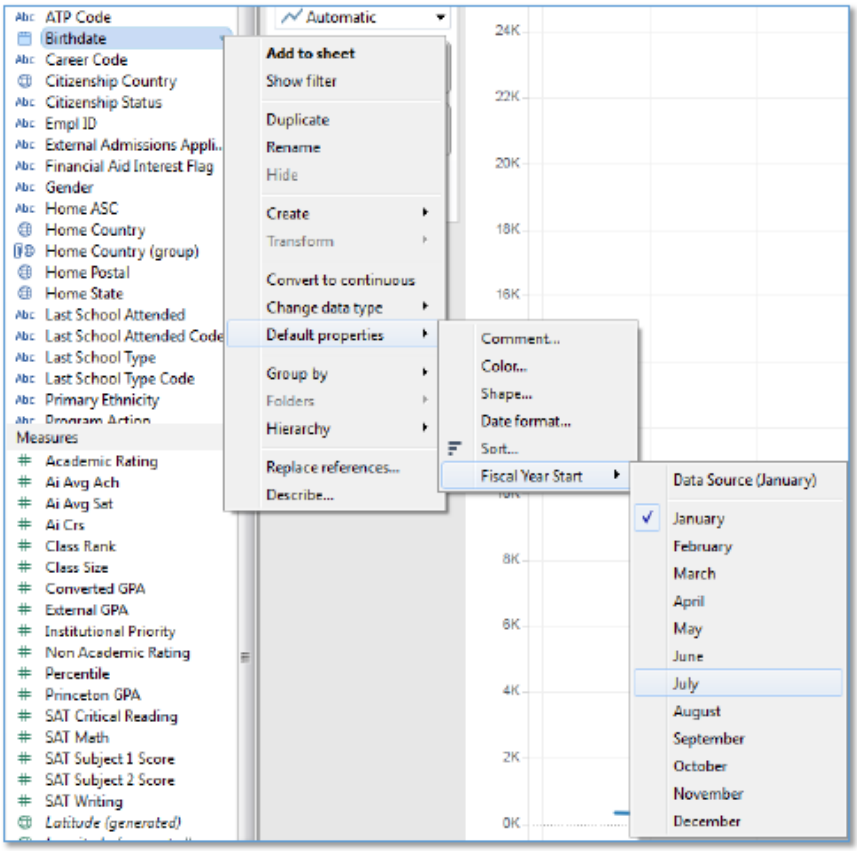
Alternatively using the mouse right click to drag the date onto a shelf, the specific date type and level of detail may be selected at the time the Date Dimension is added to the view.



Defining a Fiscal Year

By default, the fiscal year in Tableau is defined as January through December. Since the fiscal year at Princeton University runs from July through June, it might be beneficial to adjust this setting. As long as the start is the first date of a month, Tableau allows for the fiscal year setting to be adjusted.

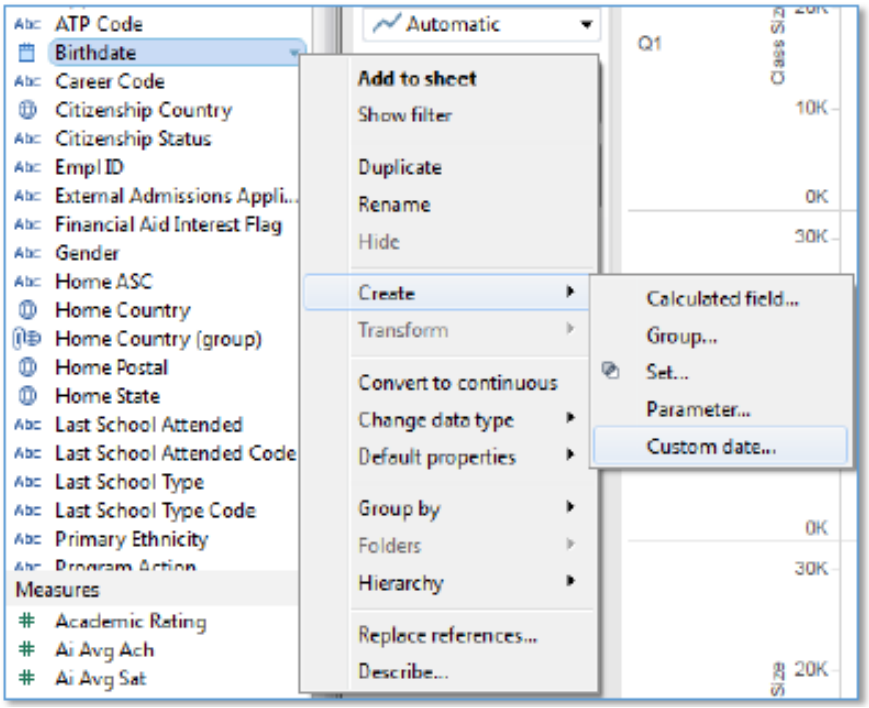
To adjust the fiscal year setting, select the Date Dimension to be changed. Click on the down arrow, select Default Properties, Fiscal Year Start, and then the desired month.



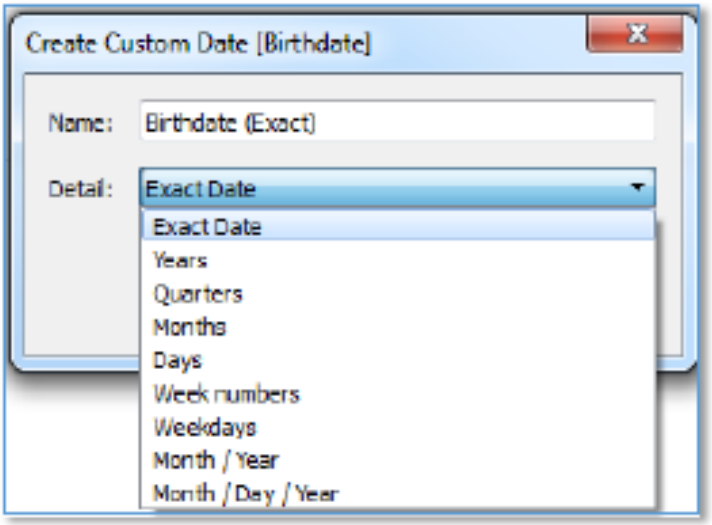
Custom Dates

Tableau has one additional feature available for use when working with dates. If there is a date part frequently used, it may be saved as a Custom Date. The saved part may be discrete (date part) or continuous (date value) and can be used without having to adjust the default setting or specific use. If discrete, the custom date will also not have the default hierarchical nature available for a standard date.

A Custom Date is created by selecting the Date Dimension to be changed. Click on the down arrow, select Create and Custom date...



The Create Custom Date dialog box is used to set a meaningful Name, level of Detail, and Date Part or Date Value as desired.



Bar Charts

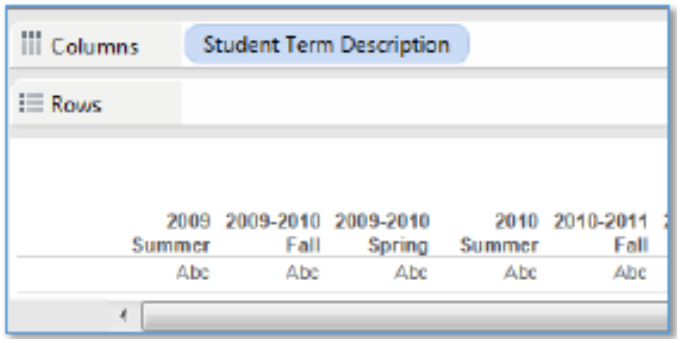
A bar chart is a good choice of visualization when you are comparing different groups or to track changes over time. It's good to note that when tracking changes over time, the bar chart is best when the change is large. If not another type of visualization might be better.

There are several simple steps that can be followed to create different types of bar charts in Tableau. We'll step through a few in our exercises.

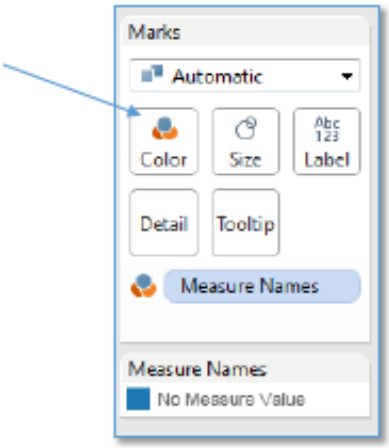
Exercise: Create a Stacked Bar Chart with a Separate Bar for Each Dimension

1. First connect to the server data source Enrollment and Person Fact +. Create a blank worksheet. To do this you can either create a new worksheet using the New Worksheet icon or by right-clicking on your current sheet and selecting New Worksheet.

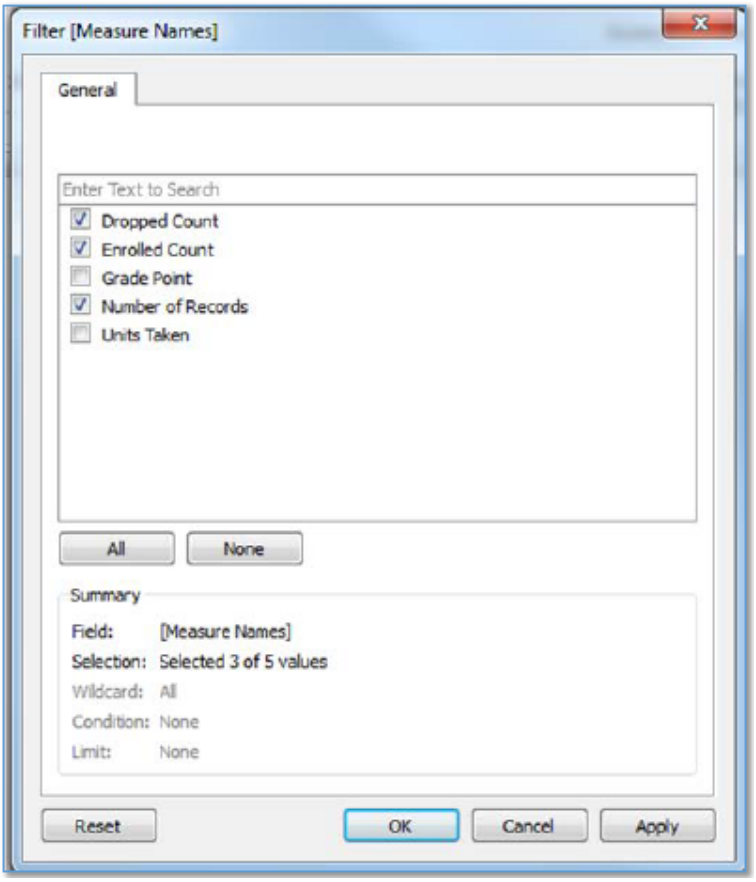
2. Drag the dimension Student Term Description to Columns.



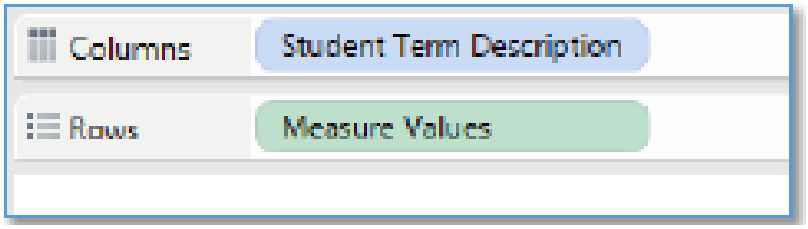
3. Drag Measure Names to Color on the Marks Card.



4. On the Marks Card, right-click Measure Names and select Filter. Select the check boxes for the measures you would like to display. Then click OK.



5. From the Measures pane, drag Measure Values to Rows.



You will now have a stacked bar chart with each Dimension with its own bar.

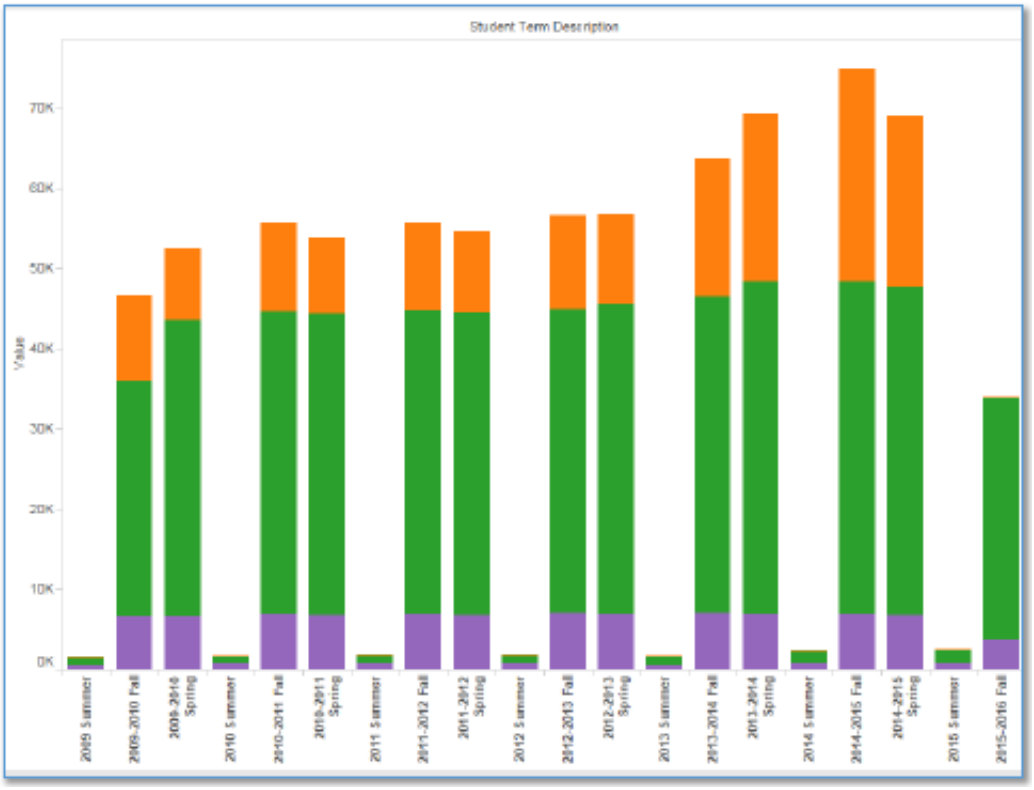


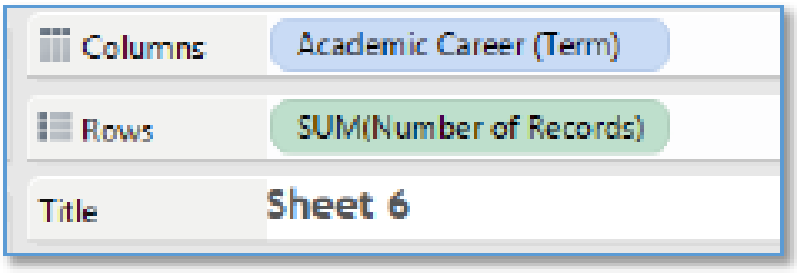
Chart Customization

Once you've created a chart there are many ways you can add customization. These include changing the title, customizing the axes, editing aliases, editing the colors, changing the sort order, adding filters, and many more actions.

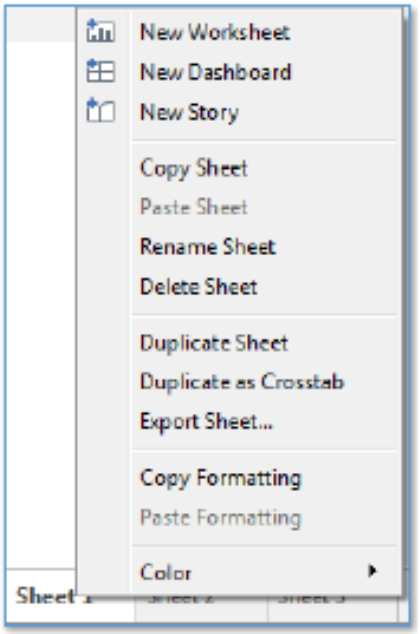
In order to follow along if you don't already have it open, please open the workbook Training – Bar Charts xx you saved earlier.

Changing the Chart Title

By default, the title of your chart will be the sheet name and is displayed in Tableau just below Columns and Rows.

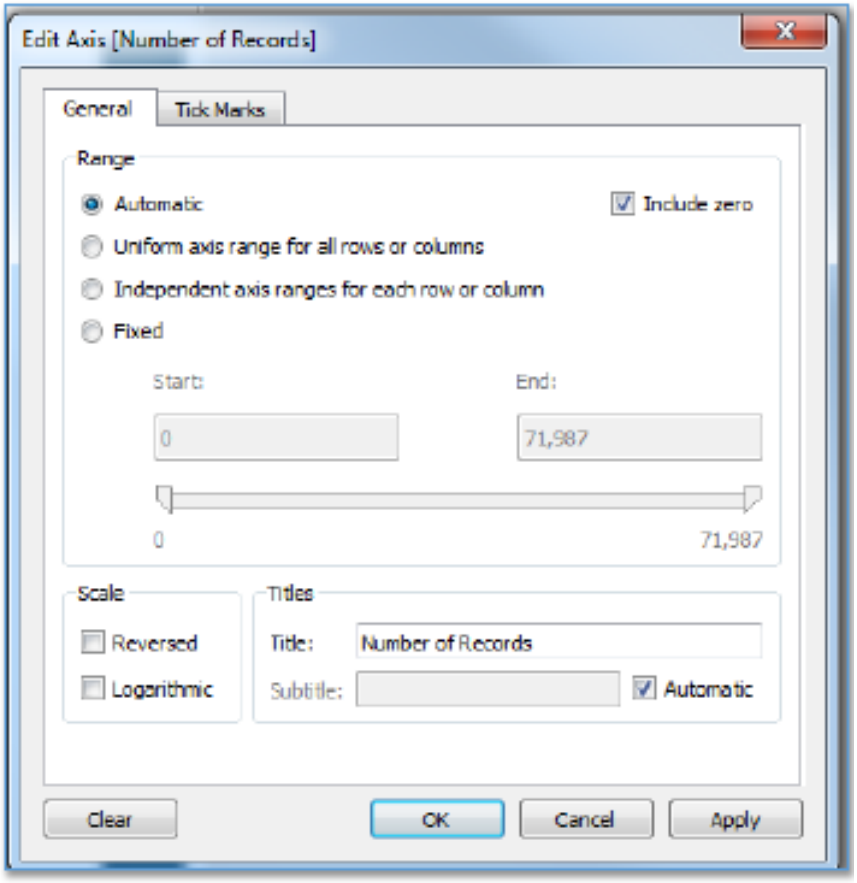


You can edit the sheet name by right clicking on the sheet tab and selecting Rename Sheet. This will allow you to edit the sheet name directly on the tab.



Customize Axis

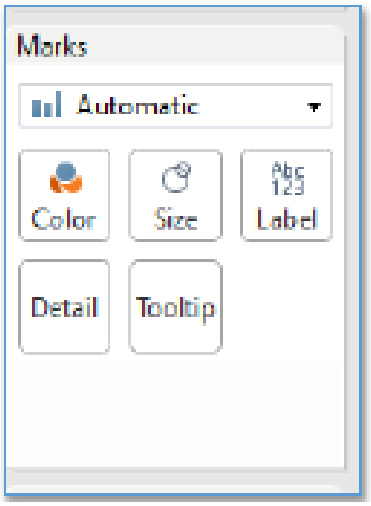
If you have a continuous measure, you will be able to edit your axis. If the values are not continuous, they are treated as headers and can be adjusted by editing the alias as described below. In order to edit an axis, right-click the axis name and select Edit Axis.



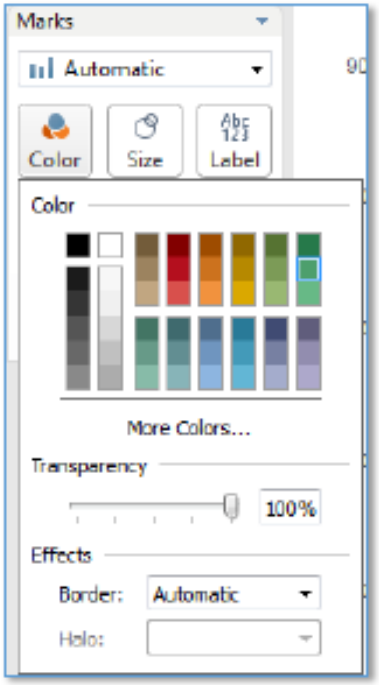
You have the ability to update the title of the axis, change the range, and update the tick marks.

Edit colors

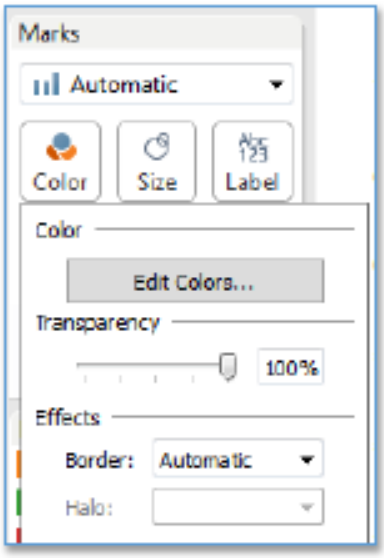
You may choose to change the palate to be different from the default Tableau provides. To do this click the Color shelf on the Marks card.



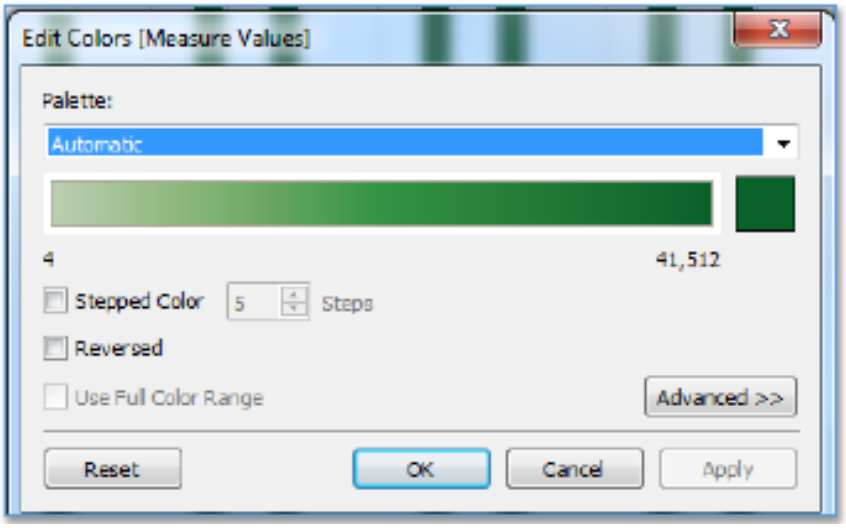
If you only have a single color on your chart, you will be able to choose a new color immediately.



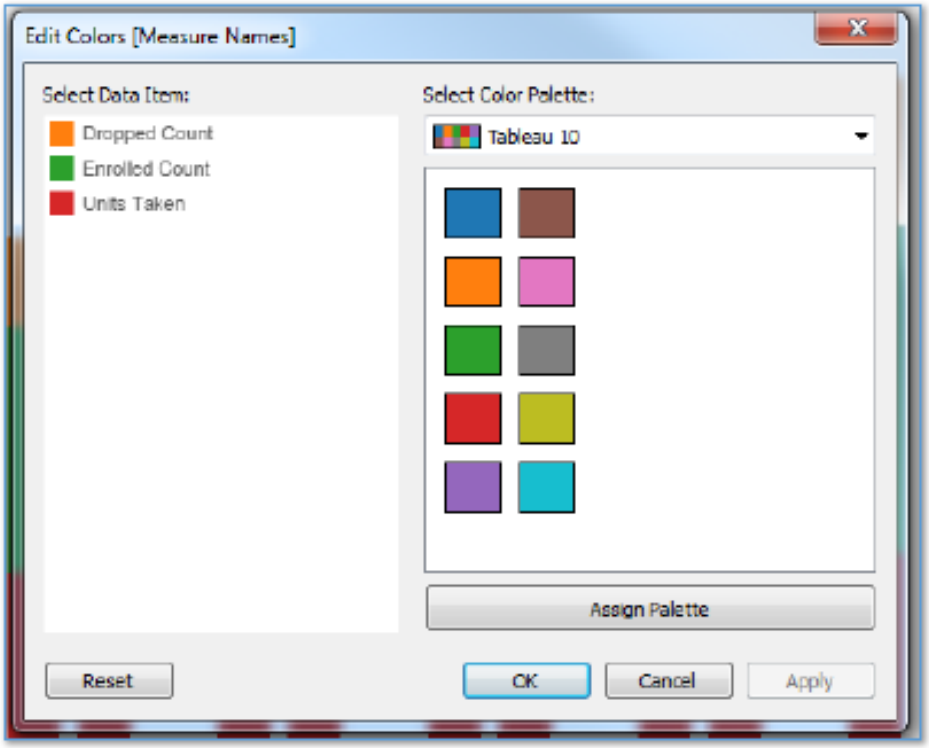
Alternatively, if you have a continuum or multiple colors, click Edit Colors.



If you have a continuous range represented by your colors, you will have the option to choose from a variety of continuous palettes.



On the other hand, if your colors represent discrete values, you can choose specific colors for each value as shown below.

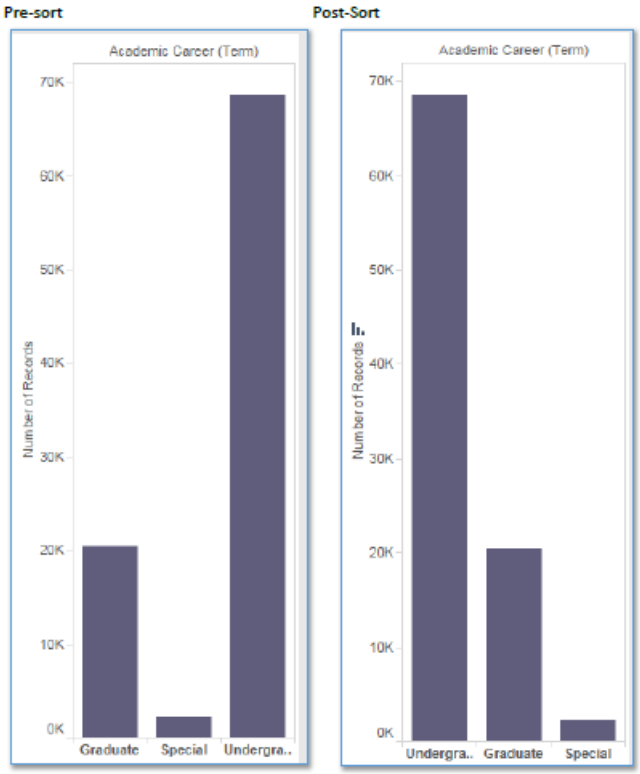


Change sort order

Tableau offers several options for changing sort order. Using the saved workbook, Tableau Training – Bar Charts xx, explore these different methods.

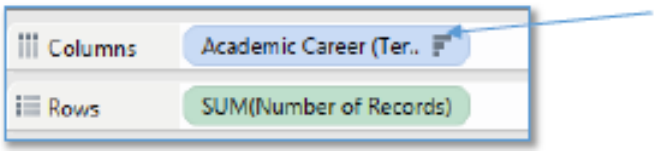
First, you can quickly change your chart's sort by clicking on either the Sort Measure Names ascending by Measure Values or the Sort Measure Names descending by Measure Values icons. Both icons are shown to the right here.

Alternatively, you will notice that if you place your mouse over an axis that is sortable, the sort icon will appear indicating that you can change the sort order. If you click the sort icon, the chart sort will change.

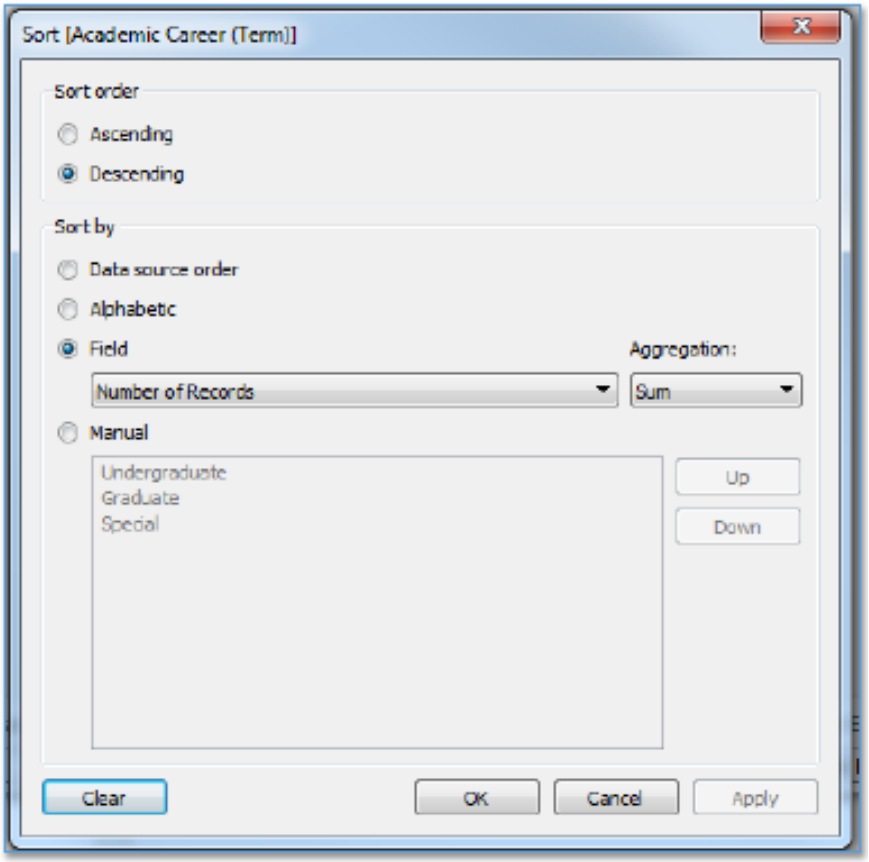


When sorted notice the image that appears on the axis indicating that the data has been sorted. Clicking on this image will cycle through different sorts – descending, ascending and clear the sort.

Also notice that there is an image next to the dimension name indicating that the values have been sorted.



The third and final option for sorting is performed by clicking on the down arrow of a discrete field. Click Sort from the menu. Using this method allows you to review further details of a sort. Using the Sort dialog box you may select the sort order and sort detail including whether or not you wish to have a manual sort.

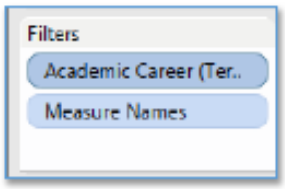


Note: there are some cases where Tableau uses a Manual sort when you have chosen to sort your chart automatically. As a result, it's a good idea to review the results when you have added a sort.

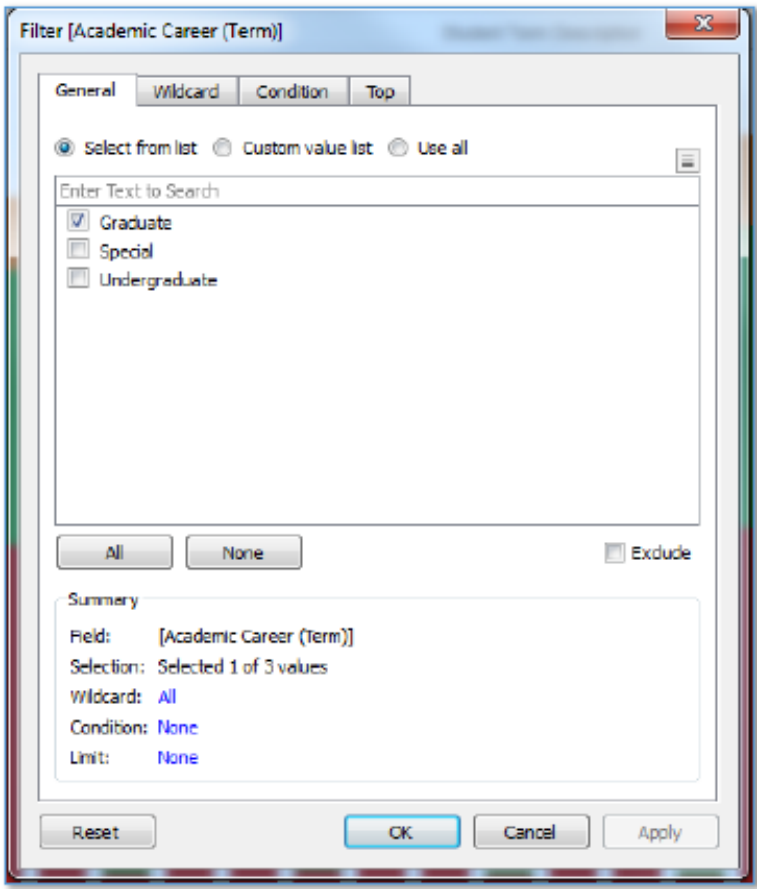
Add filters

Filters have been previously discussed but their use within charts is worth noting again.

As a reminder, adding a dimension or measure allows you to narrow the amount of data included in your visualization. For example, dragging Academic Career onto the Filters Shelf, you can select from the available values.



Depending upon the data available within the dimension selected there are multiple filter options available.



Heat Maps

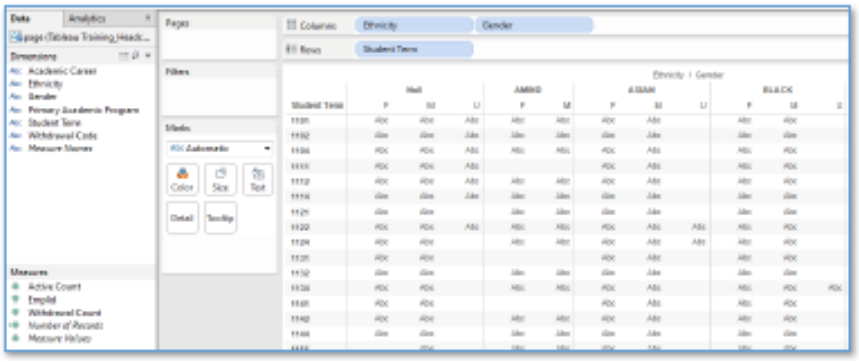
Heat maps are a great way to compare categorical data using color. They are typically constructed as a table using colored squares to represent the data, and a continuous range of colors. Heat maps allow you to see variations in the data via variations in color.

In Tableau, you create a heat map by placing one or more dimensions on the Columns shelf and one or more dimensions on the Rows shelf. You then select the mark type as Square and place a measure of interest on the Color shelf.

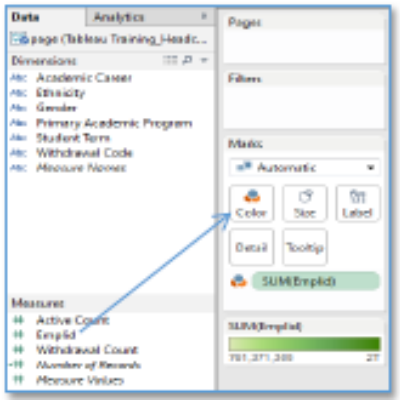
Create a Heat Map

We'll explore the creation of a Heat Map here using the Headcount and Person Fact spreadsheet from the Introduction to Tableau folder on your desktop.

1. Drag the Ethnicity dimension to Columns.
2. Drag the Gender dimension to the right of the Ethnicity dimension on Columns.
3. Drag the Student Term dimensions to Rows.

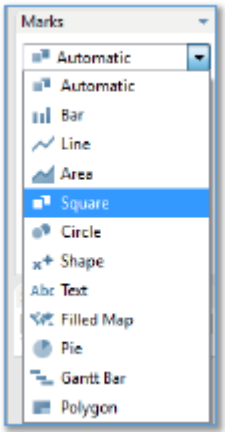


4. Drag the Emplid measure to the Color shelf.



The measure is automatically aggregated as a summation. The color legend reflects the continuous data range.

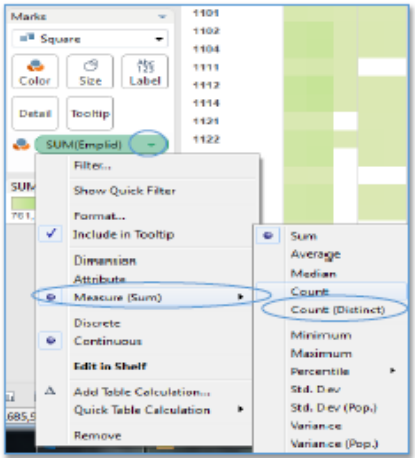
4. Change the Mark type to be Square by clicking on the down arrow to the right of the text Automatic and choosing Square.



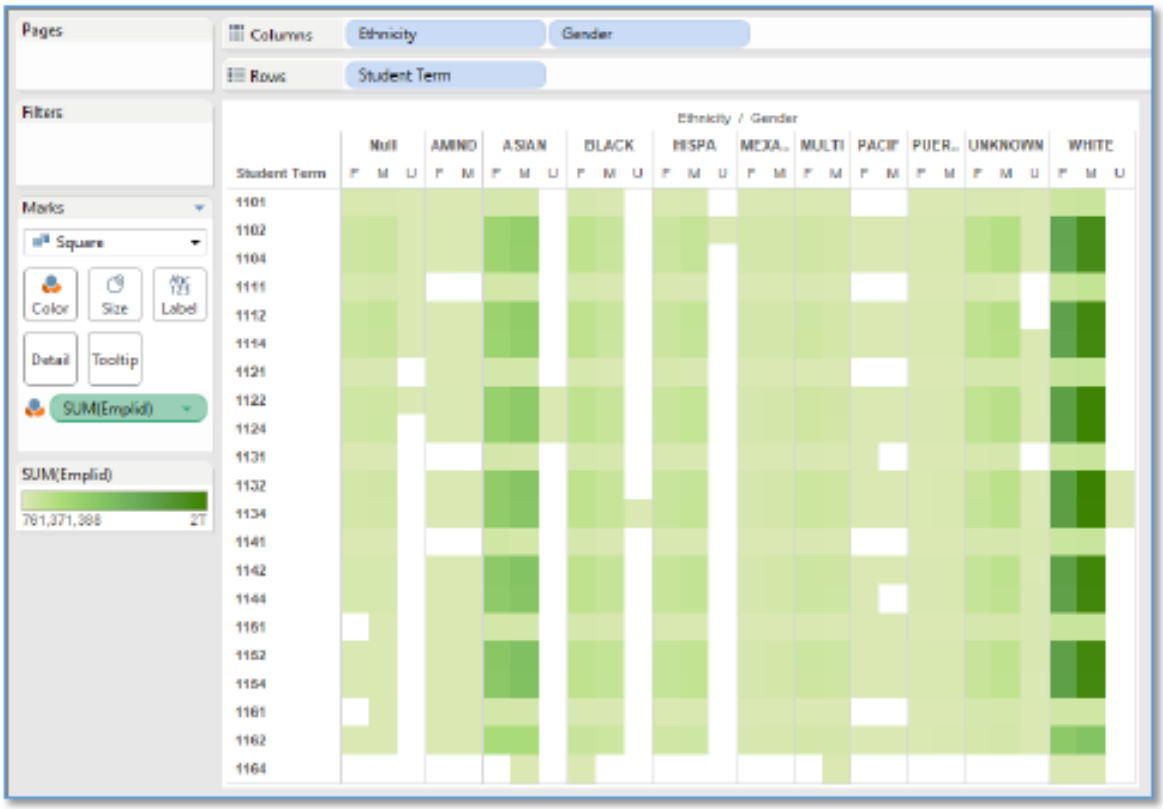
5. Optimize the view format.

- On the Format menu select Cell Size and then Bigger to increase the overall cell size.
- Increase the column width to a preferred size. Using the Ctrl key and the right arrow will make the column wider, and the Ctrl key and left arrow will make the column narrower.

6. Change the Emplid measure from automatically aggregating the data as a summation to counting by hovering your mouse over the SUM(Emplid) and clicking on the down arrow. Select Measure(Sum) and choose Count(Distinct).



This Heat Map shows, over time, via the Y axis, how many Students attended, by Ethnicity and further by Gender. It is easy to identify the most and least populous Ethnicity and Gender, and notice differences through time periods.



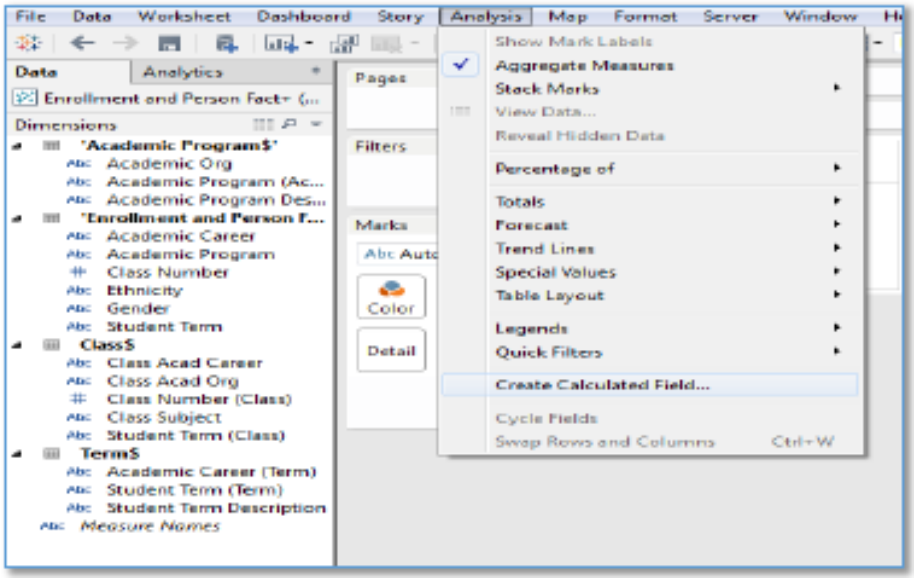
Donut Charts

A donut chart is a variation of a pie chart where the center is blank allowing for additional information about the data as a whole to be included. There are several instances when you might want to consider using a donut chart. With similarities to a standard pie chart, these include when:

- You have one or more data series you want to plot.
- None of the values you are depicting are negative.
- None of the values you are tracking is zero.
- You don't have more than seven categories per data series.
- The categories represent parts of the whole in each ring of the donut chart.

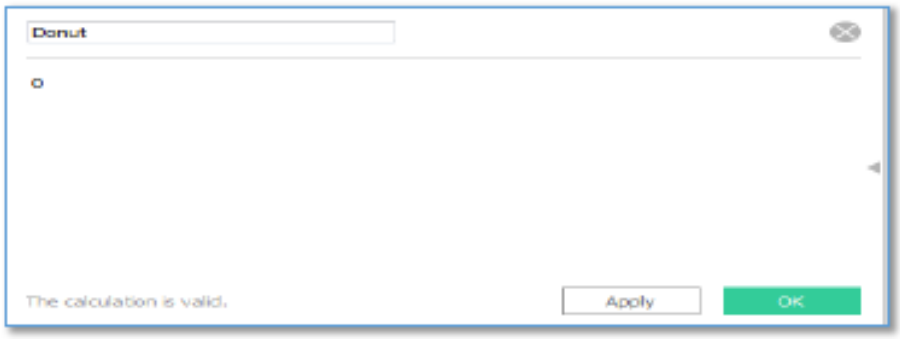
Let's create one now to give you some experience using the server data source Enrollment Person and Fact +.

1. On the Analysis menu, select Create Calculated Field.



Alternatively, with your mouse over the white space in the Data Pane, right click and select Create Calculated Field.

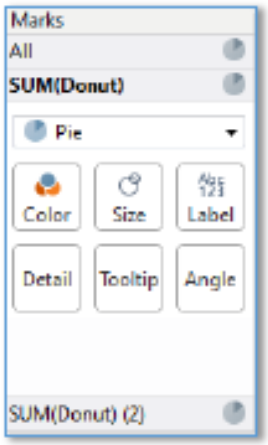
2. Name your new field Donut and type a 0 (zero) in the function box and click OK.



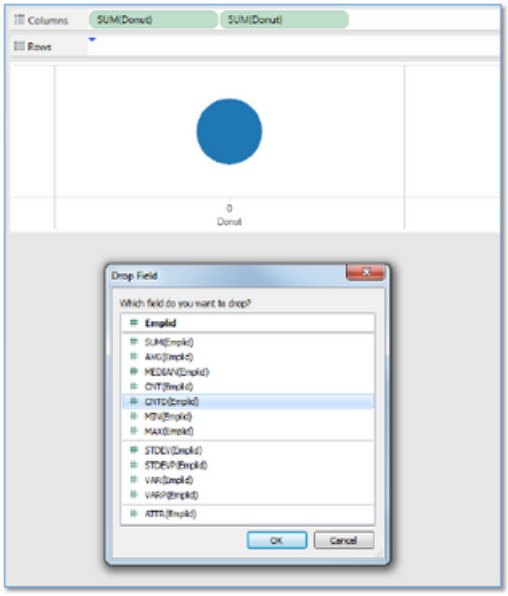
3. You'll now see Donut listed as a new Measure. Drag this new field to Columns two times. You will see two columns, each with Donut as the detail.

Change the Marks Type from Automatic to Pie.

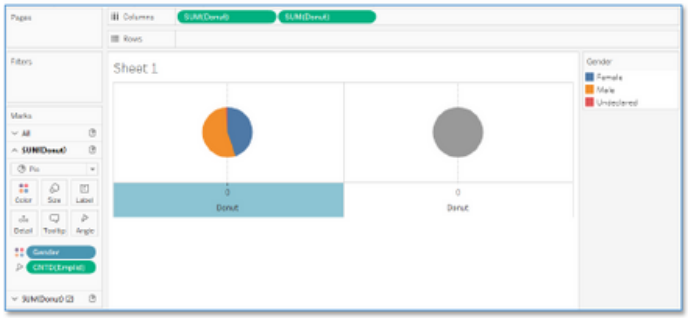
4. Click the SUM(Donut) Marks Shelf



5. Right-click Emplid and drag it to Angle. Select count distinct or CNTD(Emplid). Click OK.

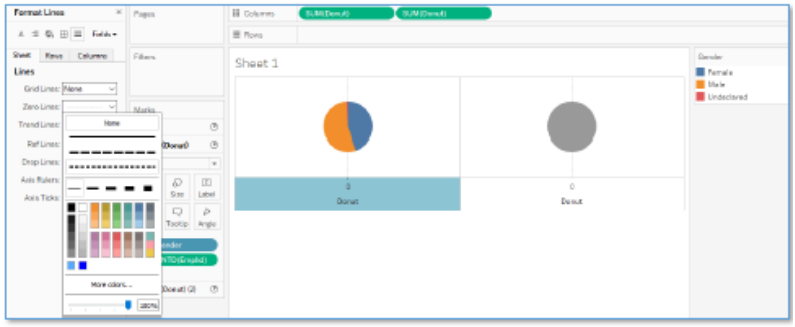


6. Drag Gender to Color.

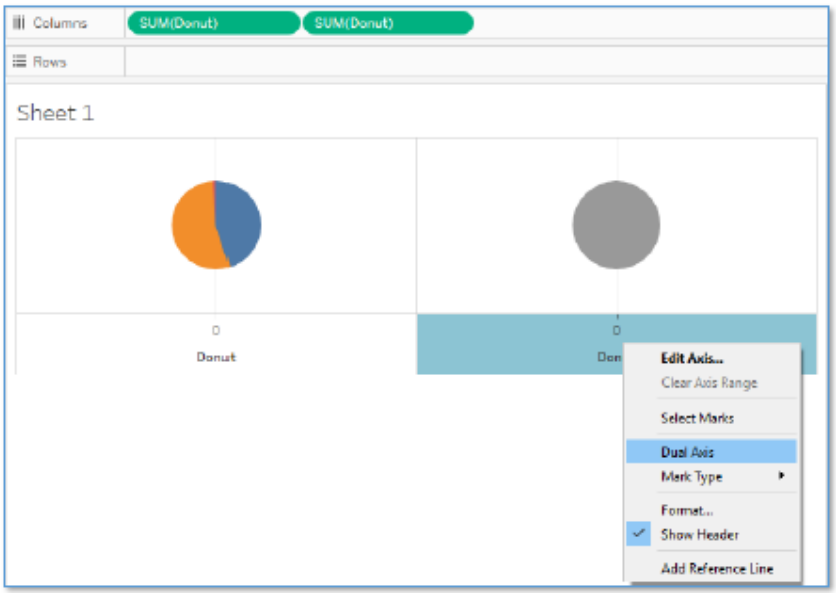


7. Right-click the Axis for Donut and select Format.

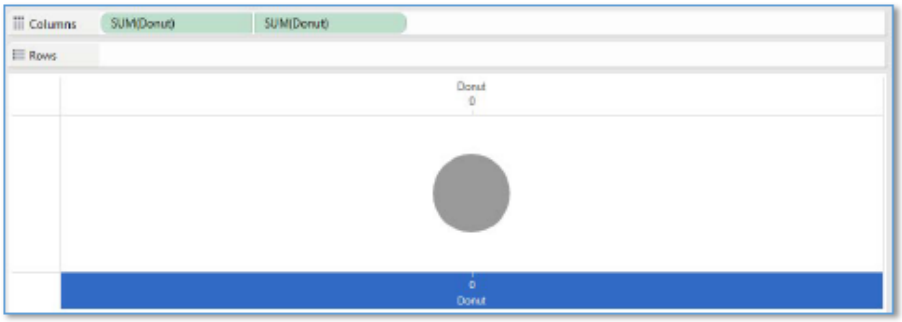
8. Click the graduated lines (show to the right here) and set Zero Lines to None.



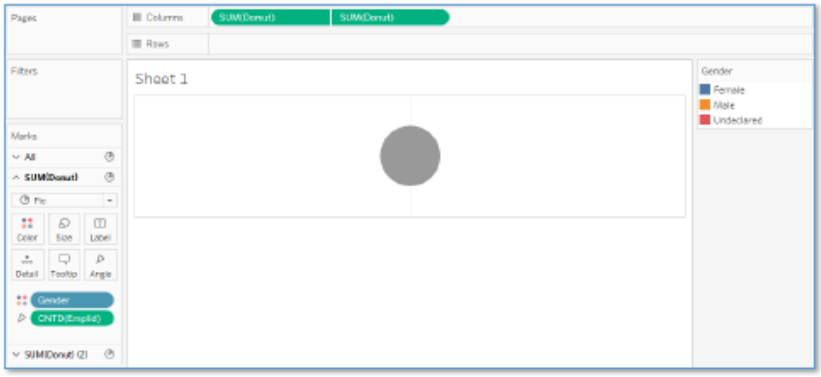
9. Right-click the axis of the grey pie chart and select Dual Axis.



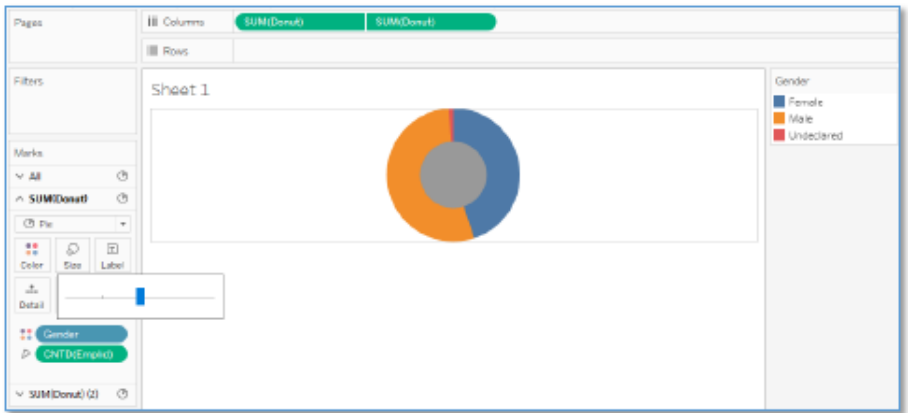
As a result, your two pie charts will lay over the top of one another.



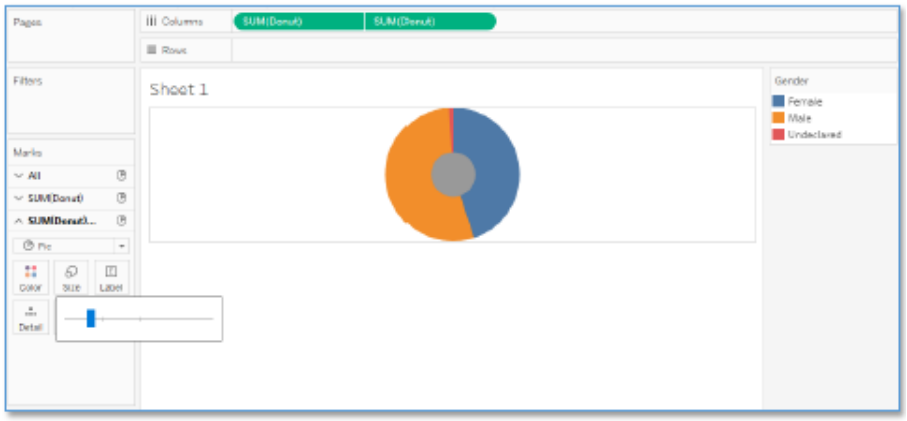
10. Right-click SUM(Donut) in Columns. Notice the Show Header has a check mark next to it. Click this to remove the header. You chart should now look like this.



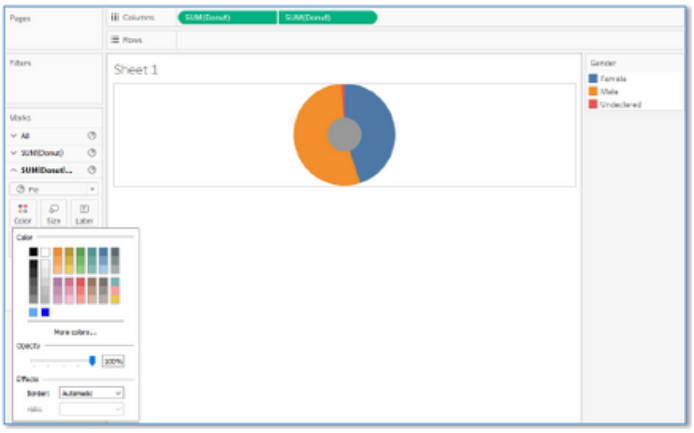
11. Click Size of the SUM(Donut) in the Marks Shelf. Click and hold down the size bar sliding it to the right.



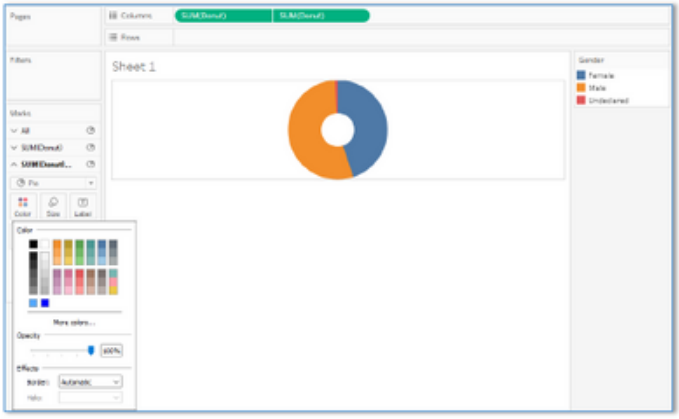
As a result, the Donut chart will be larger and appear from behind the grey pie chart.
You can optionally change the size of the grey pie chart, SUM(Donut) (2), as well.



12. You can change the color of the inner circle by clicking Color for the Marks shelf of SUM(Donut) (2).

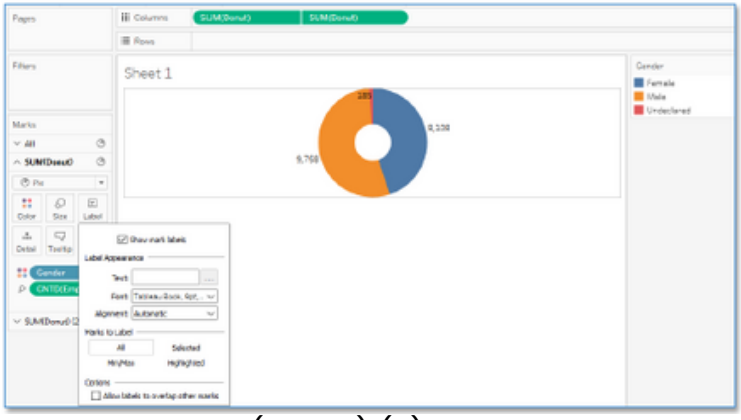


13. Select White and notice the change in your chart.

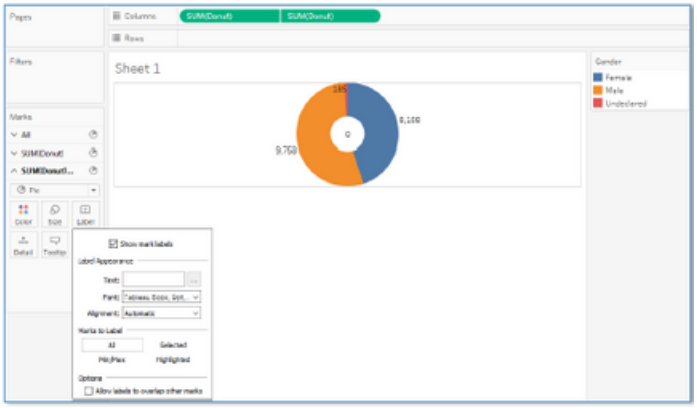


14. Next you can adjust your formatting and add labels. As mentioned earlier, you can adjust the size of either chart by clicking on the Size Mark.

You can add labels with the Label Mark. First click the main chart, SUM(Donut). Click Label and check the box to Show mark labels.



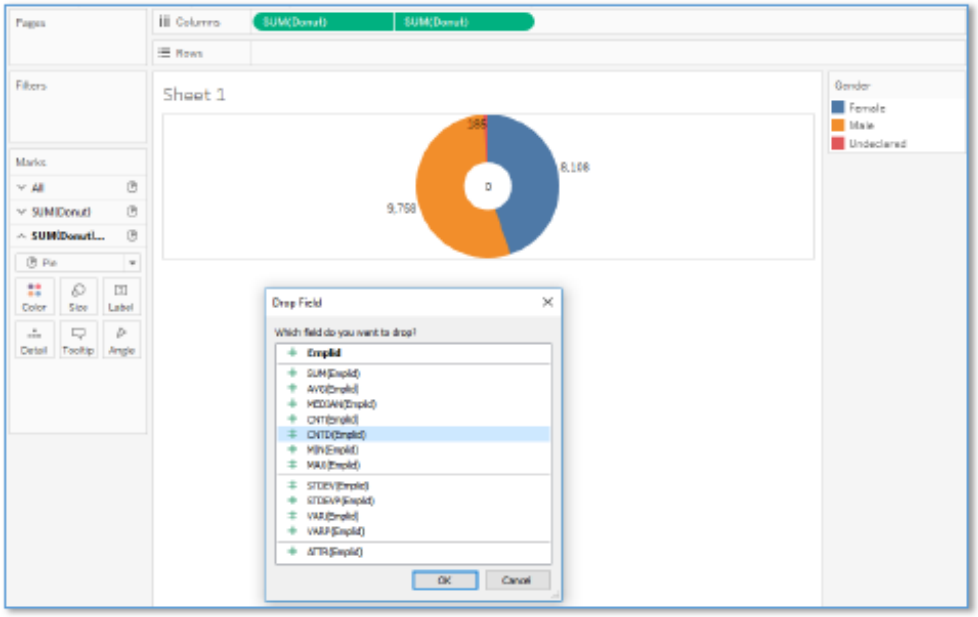
15. Select the inner chart, SUM(Donut) (2). Click Label and check the box to Show mark labels.



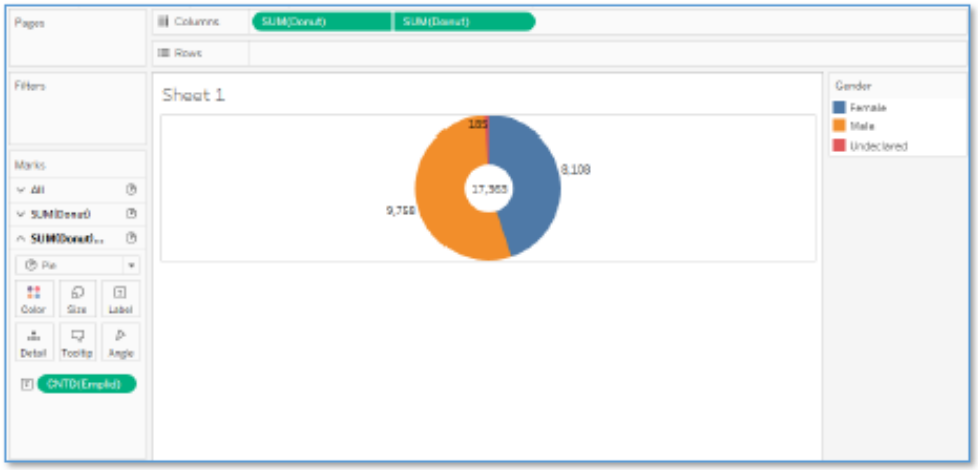
In order to include the Count of Emplid, you'll need to connect the data to the chart. To do this:

16. Right-click Emplid and drag it to the Label Mark of SUM(Donut) (2).

17. As before, select count distinct or CNTD(Emplid) and click OK.



You will now see the total count of Emplid in the center of your Donut chart.

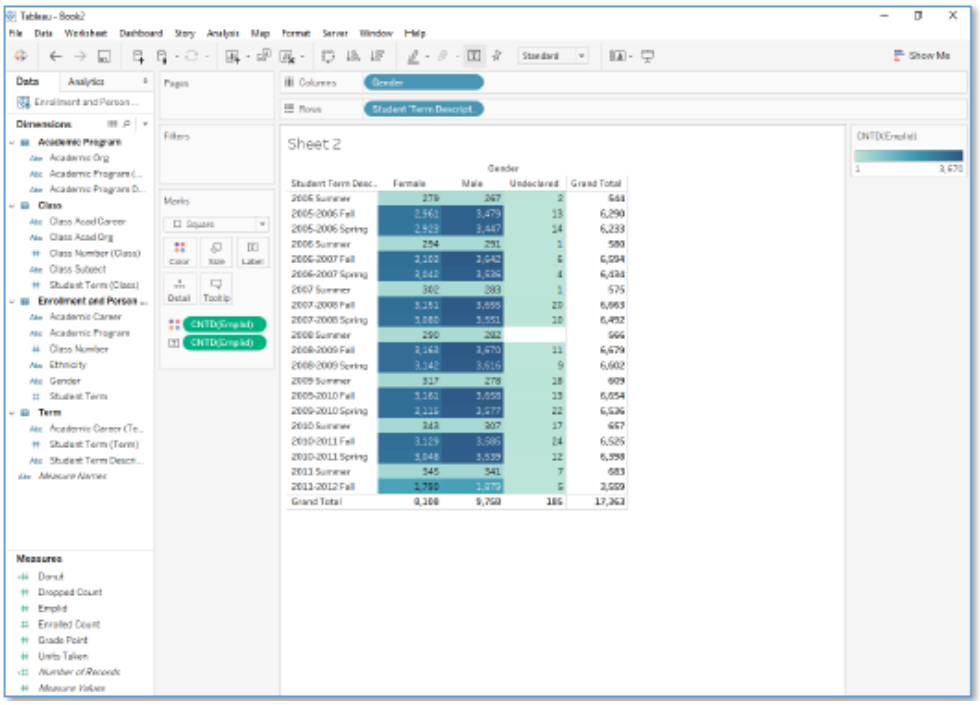


Dashboards – Connecting Your Worksheets to One Another

Once you've created several visualizations, you will often want to incorporate them with one another to provide a richer experience for your report consumers. Tableau allows you to do this through the creation of Dashboards and Stories. We'll create a Dashboard here so you can gain some experience.

Using the Enrollment and Person Fact + data source from the Tableau Server, we'll combine a crosstab with the donut chart we just created.

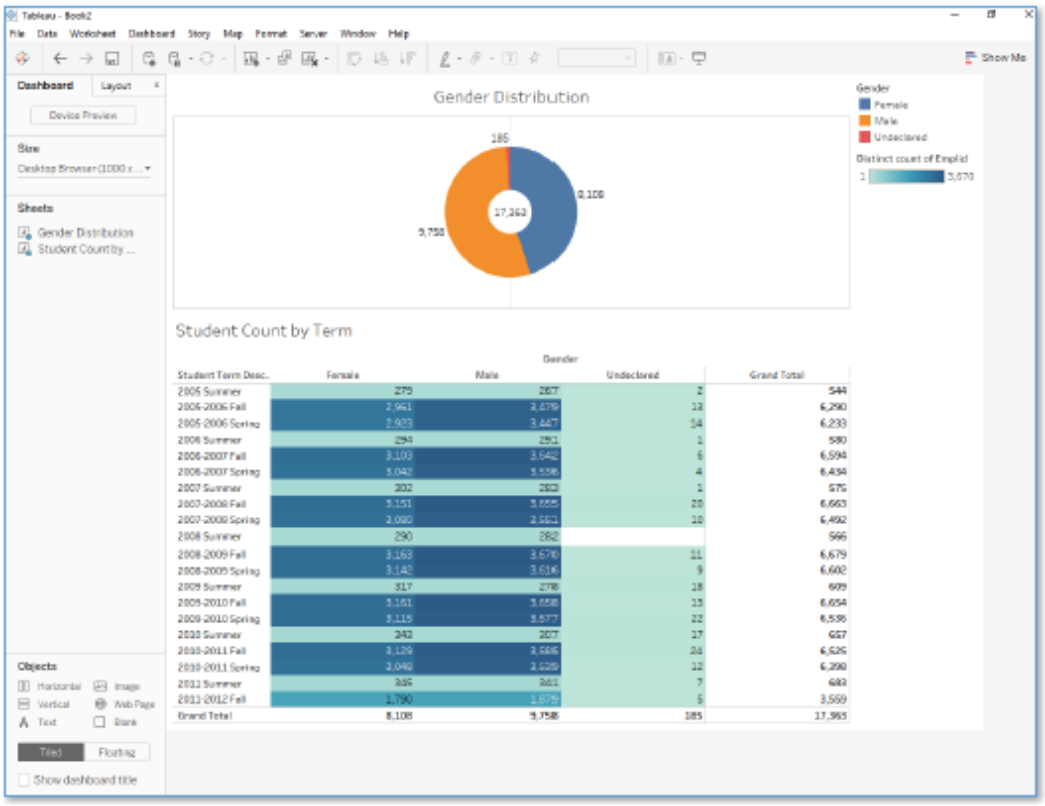
1. Drawing from our previous examples, start by creating the crosstab shown in the below image.



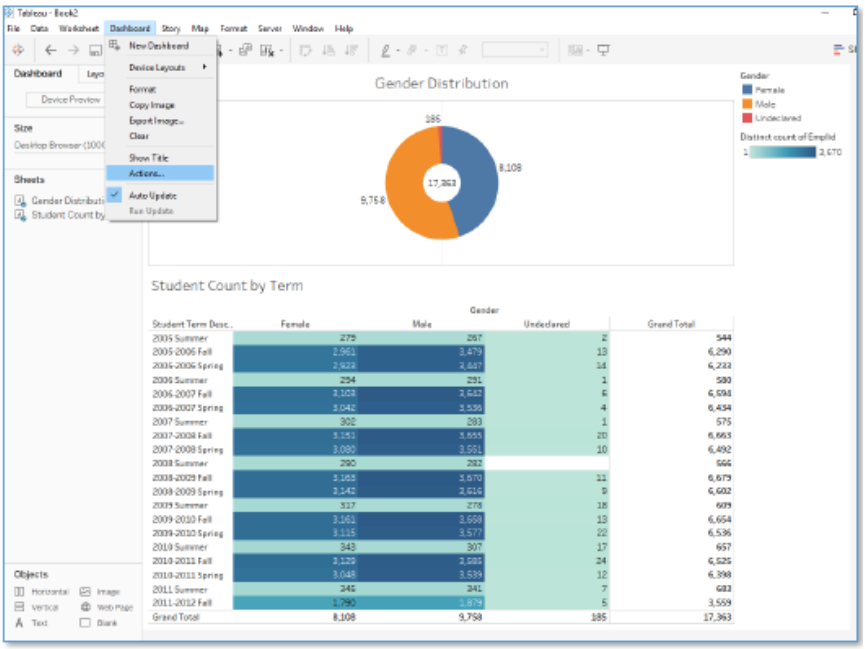
Make sure to set the Mark Type, Color, and Add Totals. (Hint: use the Analysis menu for totals.)

2. Click the New Dashboard icon (shown to the right) to create a new blank dashboard.

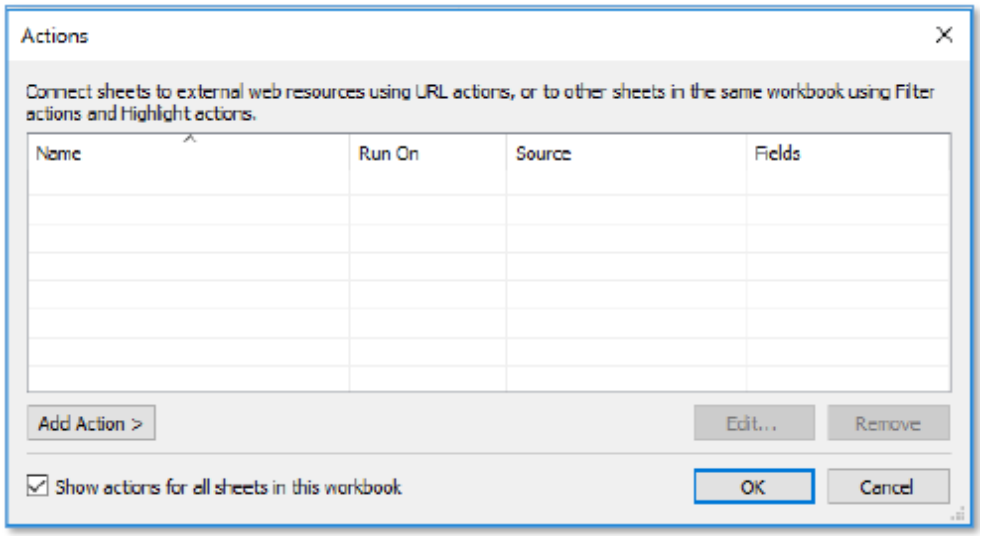
3. Bring your Donut Chart and Crosstab onto the dashboard as shown below.



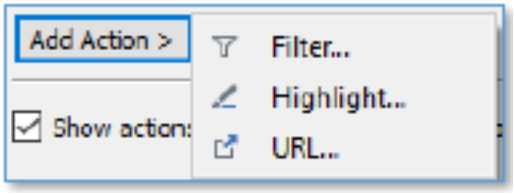
4. Once on the Dashboard, you can add actions that will tie your visualizations together. To do this, on the Dashboard menu, select Actions...



The Actions dialog box will open.



Click Add Action and select Filter.



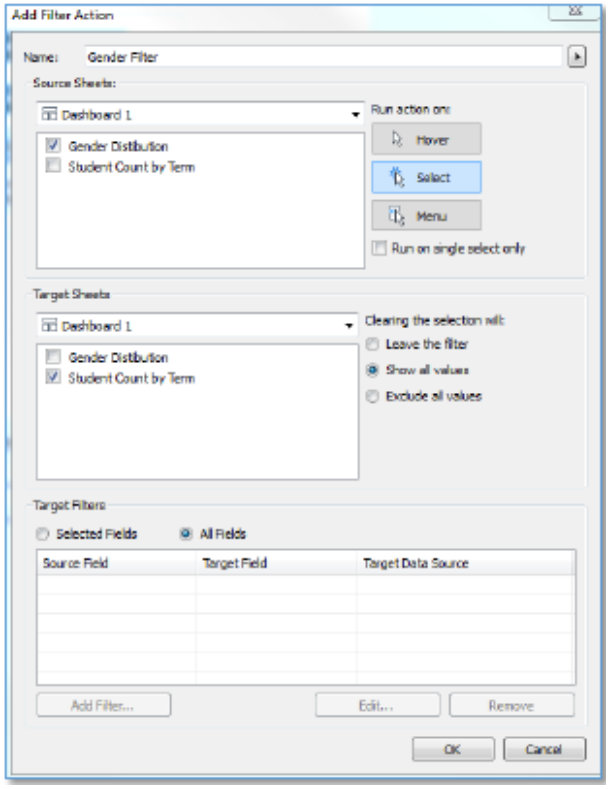
5. In the Add Filter Action dialog box set Name to Gender Filter.
6. Change Run action on to Select.

7. The Source Sheet will be the Gender Distribution donut chart, so deselect the Student Count by Term crosstab.

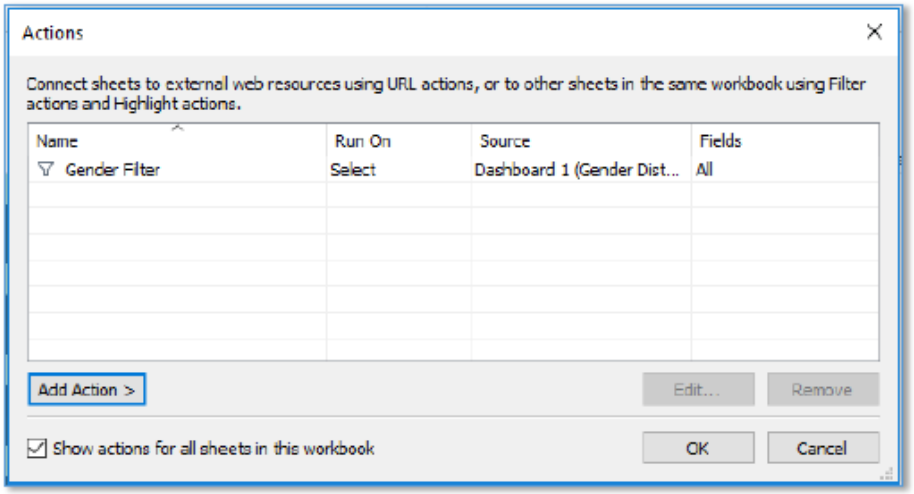
8. The Target Sheet will be the Student Count by Term crosstab so deselect the Gender Distribution donut chart.

9. Reset Clearing of Selection to Show all values.

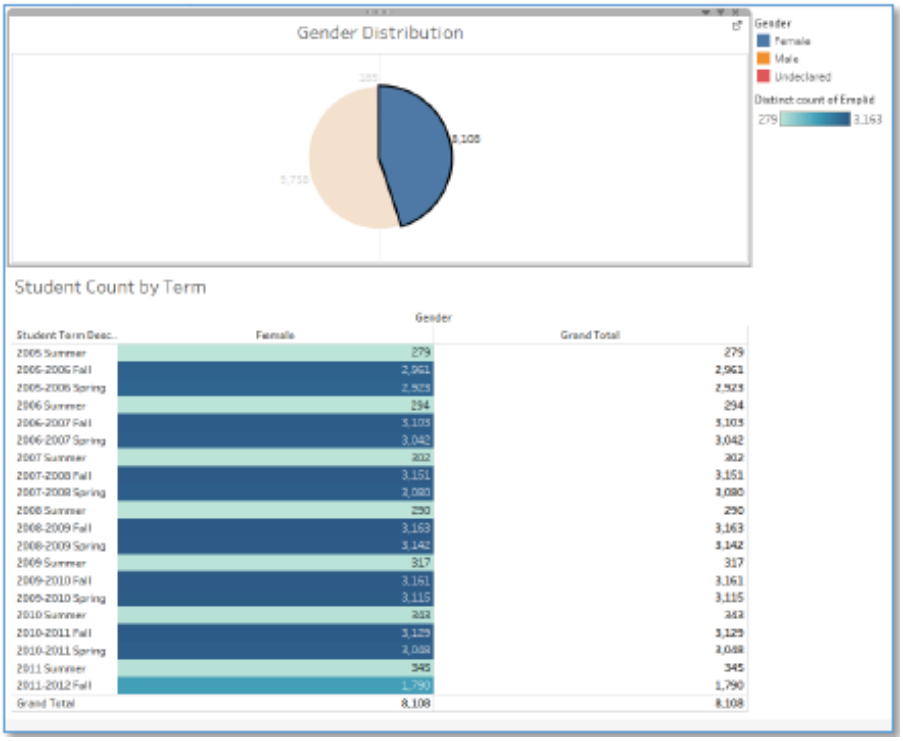
When complete your selections should appear like this. Click OK to confirm.



Your new filter will now be shown in the Actions dialog box.



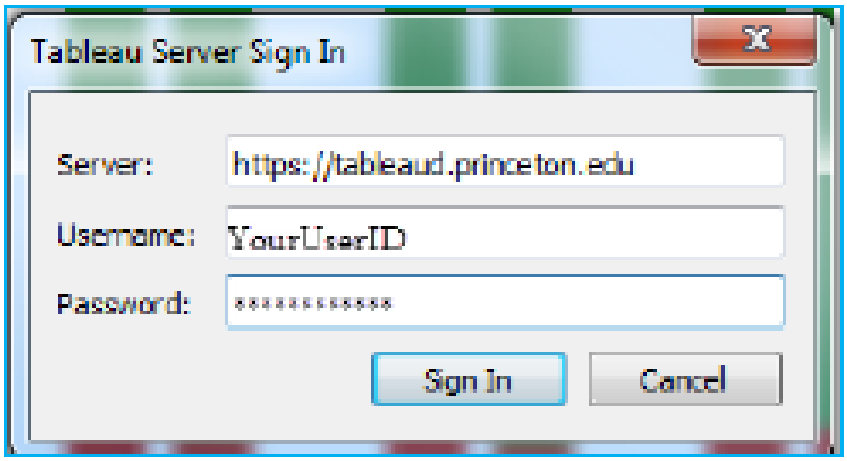
10. Click OK to confirm.
With the filter in place now, clicking on any portion of the donut chart will reflect changes in the crosstab.



Publishing to Tableau Server

Once you've created your visualization and are ready to share it with others, you'll want to publish to the Tableau Server. This will make the Worksheet, Dashboard, or Story available to those with appropriately defined access on your Tableau site.

1. On the Server menu, select Publish Workbook.
2. The first thing you'll be asked to do is to sign into the server. The server we'll be using for training is https://tableaud.princeton.edu. You will enter your username (netid) and password.



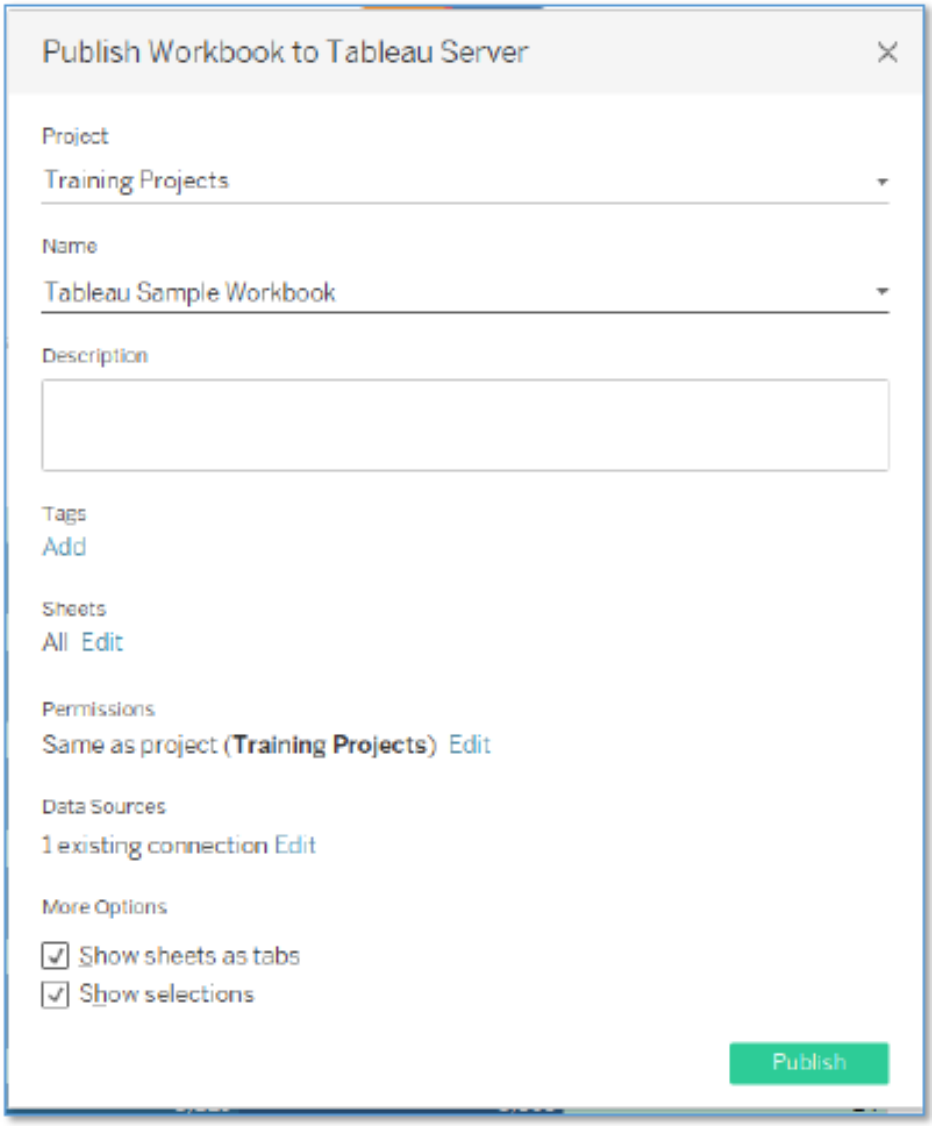
If you have access to more than one site on the server, after providing your username and password you'll be asked to select a site. We'll be using Tableau Training for this course.

3. Next you'll select the Project for your visualization. We'll use Training Projects for this class. You can also optionally customize the name of your workbook, set a description, and add tags.

As the publisher you may also customize the permissions for your work. However, we do recommend leaving the default settings. The site administrator will have final control over how the individual permissions are set to content published to your server site.

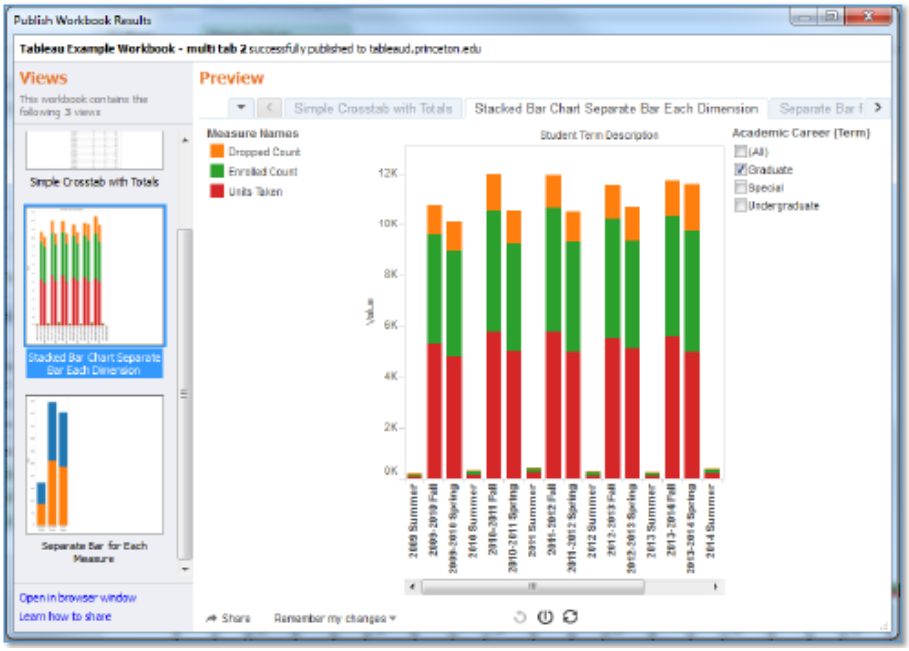
If your workbook has more than one worksheet, dashboard, or story, you should select any or all that you would like to include on the server.

You can also customize the appearance of your visualization by indicating whether to show sheets as tabs, show selection or include external files such as images you may have included in your work.

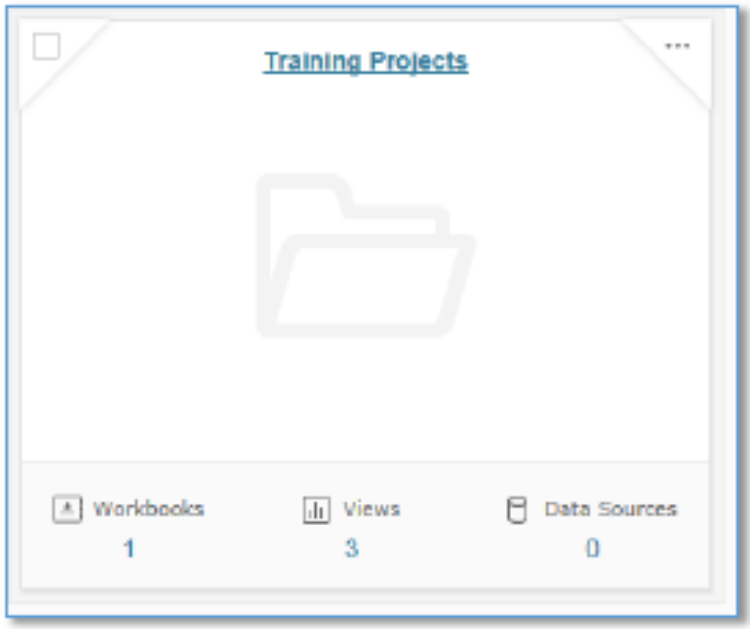


You can also set scheduling when publishing your work, though data extracts will be managed by your server administrator.

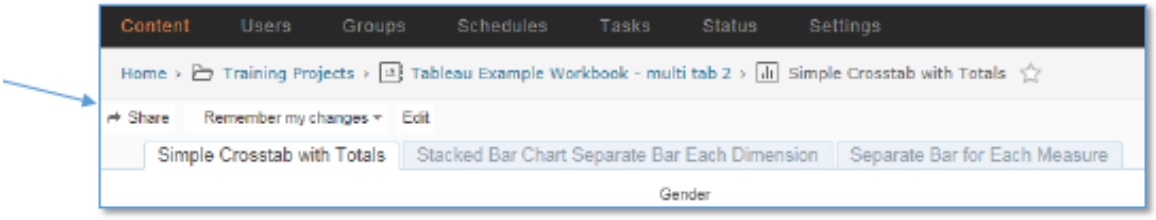
Once published, you may preview a sample of how your visualization will appear on the server.



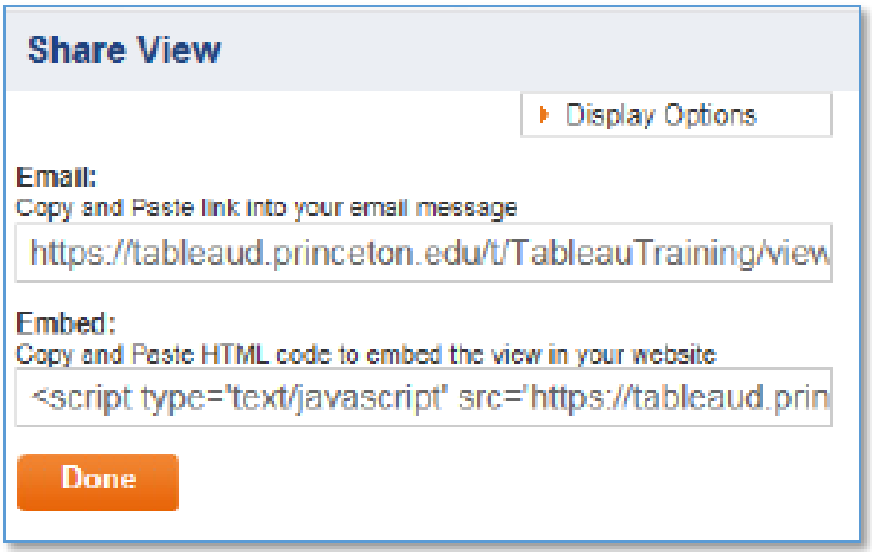
When you log into the server site, you will now see your work in the Project you selected.



4. Once posted to the server, you have the option to share your work either by sending a URL that will direct people to the content on the server or by embedding code in webpages. To do this, when viewing your visualization on the server, click Share.



5. Then select the appropriate link or embed code.



Using our Development or Production server will require users to be authenticated prior to accessing the views. If for any reason you need open, unauthenticated access you will need to work on the Princeton Tableau Public server.

THANK YOU

Akshay Kamble