Introduction to Tableau: Dashboard Creation and Publication

An Applied Case Using Kansas City MO Open Data

Excerpts from

Real Estate Feasibility Analysis: An Analytical Approach

James R. DeLisle, Ph.D.

Revised: Fall, 2020

© All Rights Reserved

Contents

Overview 1
What is Tableau 1
Data Visualization Process 1
Tableau Interface 2
Base Sheet: Tableau Workbooks 2
Tableau Workbook Components 2
Exploring Data 3
Open Data 3
Data Types 4
Exploring the Data Content 5
Tableau Data Connections 8
Connection Overview 8
Connecting to Excel, CSV Files 8
Connecting to KC Open Data 9
Connecting Tableau to Data 9
Setting Up Tableau Sheets 10
Map Data Layers 11
Turning on displays for maps 13
Filtering Data Display 16
Overview 16
Filter by Year: Discrete 16
Make Filter Visible and Change Selection 18
Modifying Filter Views 20
Customized Filter List 20
Adding Labels 21
Adding Colors 22
Controlling Snap to Filter View 22
Managing Sheets 23
Naming Sheets 23
Creating New Sheets 23
Multiple Value Sheets 23

Publish to Tableau Public Server 25
Sharing Files 25
Publishing to Tableau Public Server 25
Share via Link or Embedded Code 27
Options 27
Permanent Link 27
Embed Code 27
Conclusion 28

Overview

What is Tableau

Tableau is a data visualization/analysis tool which includes a spatial component. To start, download Tableau Free Trial and/or purchase. Once Tableau is loaded, you can start working with the default worksheet that opens when you launch Tableau or if you have worked with it, an existing file that you open.

Data Visualization Process

Before launching a Tableau project, it is useful to apply some design thinking to determine what you are trying to accomplish. Exhibit 1 presents a basic planning process that can be adopted to many data visualization processes. Given the relatively low learning curve for Tableau, many users will find that the initial two phases of the project are the most demanding and potentially time consuming. Indeed, the introduction of a new wave of data analytics and visualization tools argues that data scientists will have to pay heightened attention to the Problem Statement phase to ensure that decisions that are made in the data development and processing phases can address the fundamental research question. This is particularly true in light of the explosion in data availability –both open and proprietary data—and the requisite need to turn that data into meaningful information that can be used in decision-making. This need is amplified by the fact that many managers and decision0-makers may not be aware of the breadth and depth of decisions that can be supported by data analytics and data visualization that are transcendent and move beyond dashboards that display descriptive data to interactive dashboards that present predictive analytics that are forward-looking.

Exhibit 1: Data Visualization Process

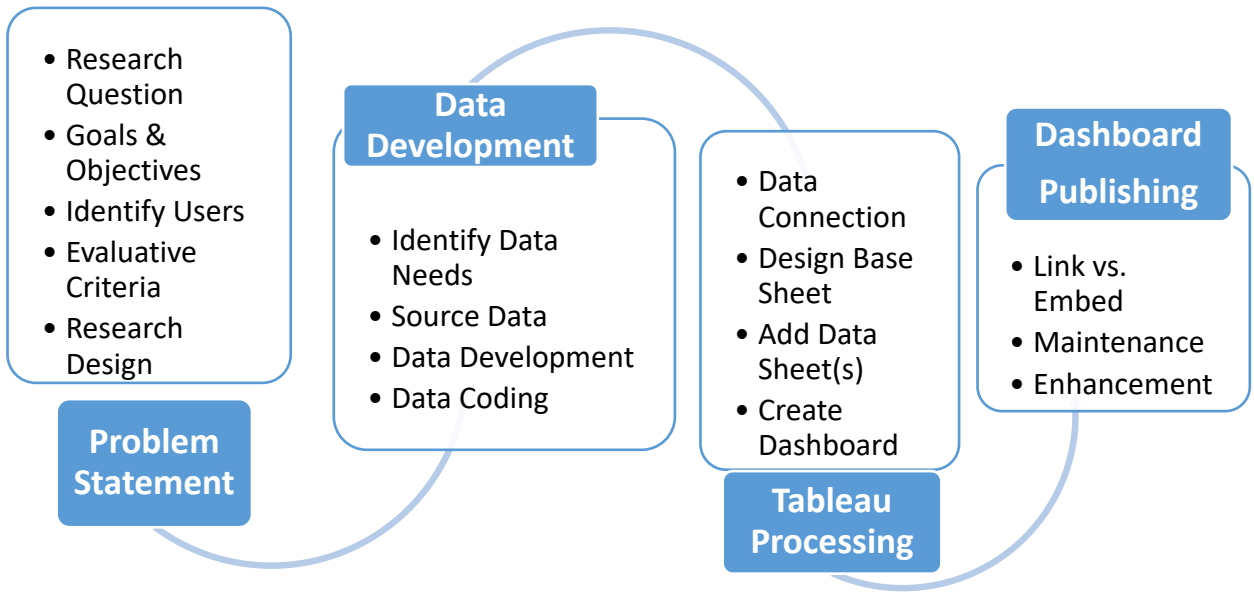


Tableau Interface

Base Sheet: Tableau Workbooks

Tableau files are stored in workbooks which are denoted .twb. Each workbook has several components and can be expanded to include multiple sheets. Exhibit 1 presents the basic layout and highlights some of the key sections.

Exhibit 2: Tableau Workbook Layout

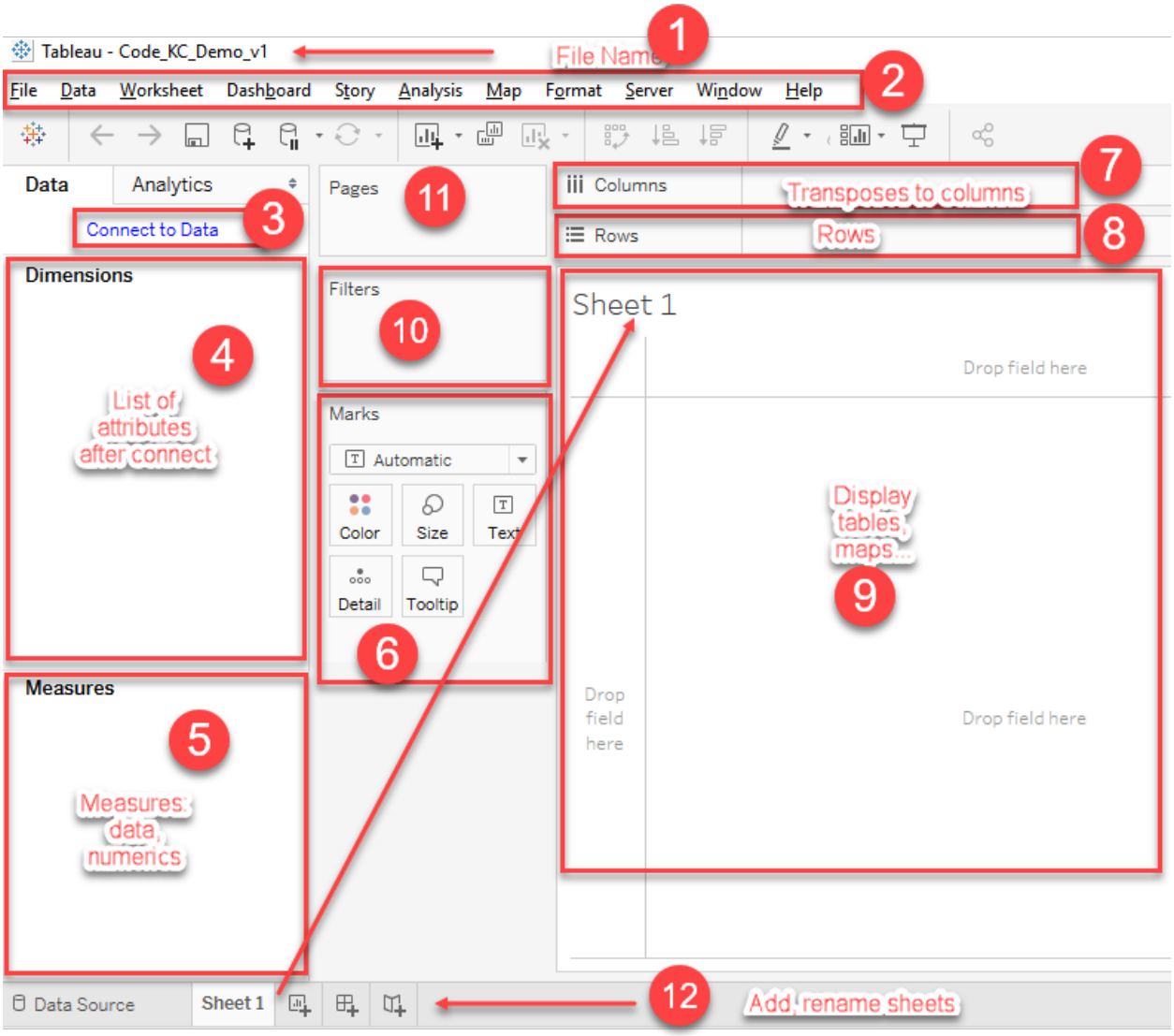


Tableau Workbook Components

1. File Name. This is where you enter your file name which will be your Workbook.twb file. Given the dynamic nature of the product where changes can wipe out progress, it is useful to version number the name.
2. Menu. This menu bar lists the various functions under each heading.

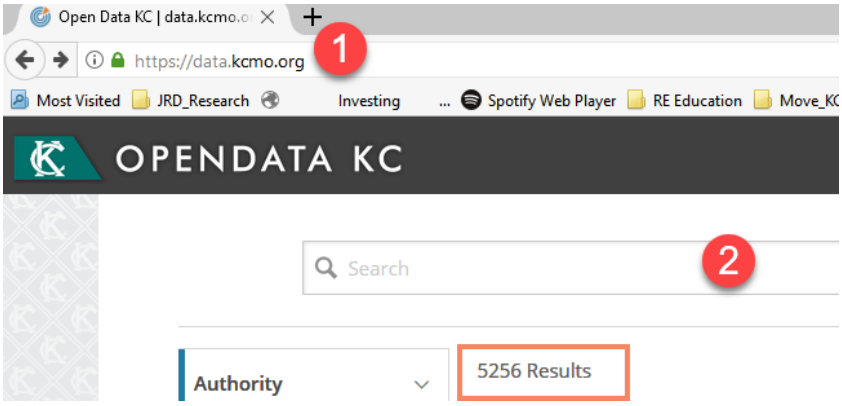
3. Connect to Data. To start, you must first connect to a data source. This could be a spreadsheet, a database, or an open-data connection.
4. Once the data have been "connected" the first Sheet (default named Sheet 1) will auto populate with attributes or qualitative dimensions.
5. The package will auto determine the quantitative variables which include numerics and spatial data (i.e., lat/lon)
6. Once the data are loaded, this section controls what is displayed on the Sheet. It is populated by dropping fields into the bottom of the box.
    a. If spatial data, the Lat/Lon fields should be added along with other data you might want to display on mouse-over or use for filters.
    b. In the Tooltip section, the first several will automatically be displayed on mouse over, but you can edit that to change the contents.
    c. Drag some ID field to the bottom of the Marks box which will then be displayed in the Sheet screen.
7. Drag the Longitude (lon) field to the Column bar; make sure it is in the Measures box and has a globe symbol which denotes spatial data.
8. Drag Latitude (lat) to the Row bar
9. Depending on whether the Sheet is geographic and what data you dragged to Box 6, you should see the results of the data.
    a. At this point you can go back and add new dimensions or measures to Box 6.
    b. You can open the Tooltip icon and edit it to determine what to display on mouseover.
10. If you want to be able to Filter the data (i.e., turn some items on and off) drag the item from Box 6 to 10.
    a. You can edit the filters here, changing wording and/or options
    b. You can make filters visible which will allow them to be used in dashboards.
11. Pages helps you navigate.
12. Sheets can be renamed (right click) and/or added by the Plus signs.
    a. You can add new worksheets
    b. You can also add new dashboards

Exploring Data

Open Data

Kansas City MO is part of the Open Data movement and has made a variety of data available through its Open Data Portal https://data.kcmo.org/ As noted, the OPENDATA KC portal contains 5,256 results.

Exhibit 3: Open Data Portal



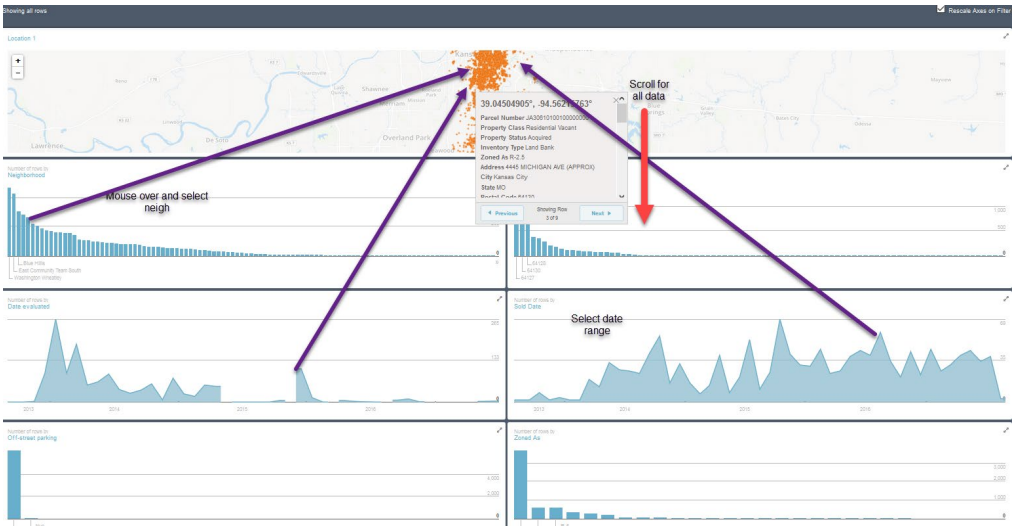
Data Types

These "results" include a range of data in a variety of formats:

Data Lens Pages

These are dashboards that are set up for various datasets. The exhibit presents a Land Bank Datalens with an interactive map and data series noted in the charts.

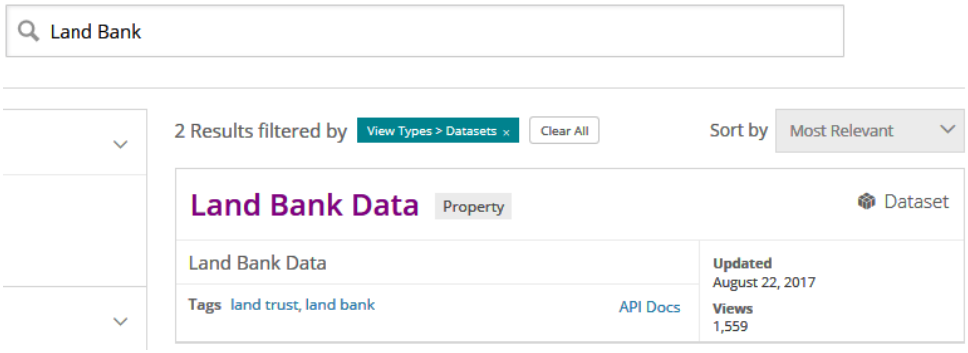
Exhibit 4: Datalens Dashboard for Land Bank



Datasets

Datasets are raw data tables that have been created to track various data items (e.g., 311 Calls, Property Violations, Crime (by year). Clicking on Dataset in the left navigation bar opens a page which starts with a Search box and a listing of most recent data.

Exhibit 5: Land Bank Dataset Search Results



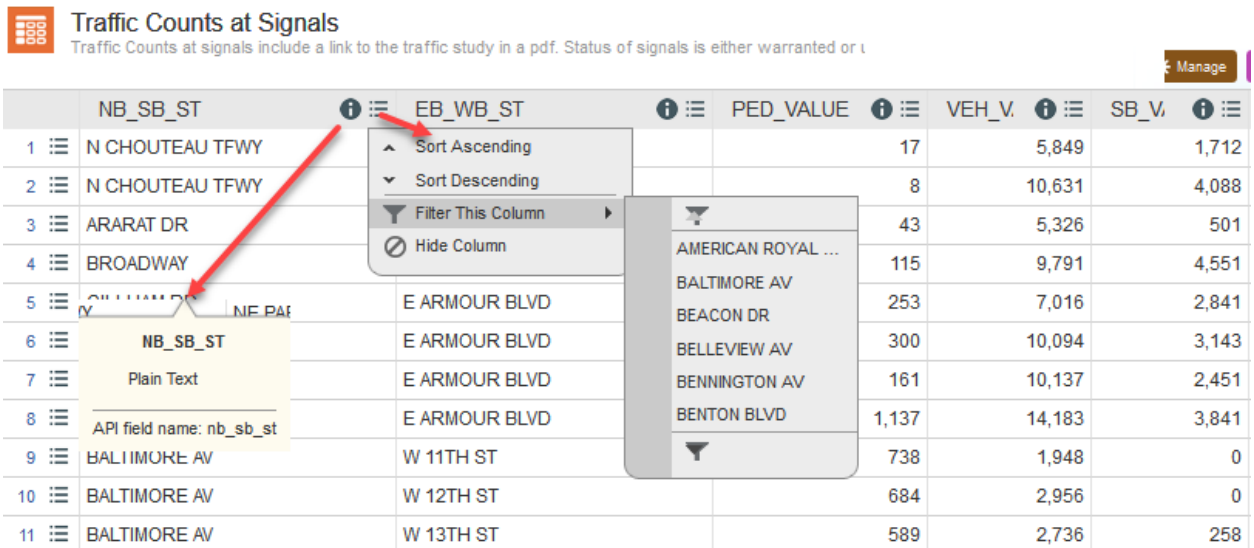
Filtered Views

Filtered views are data tables, maps, or graphs in which fields and/or records have been selected from raw datasets to answer a specific decision or satisfy a certain data requirement. Unfortunately, the source dataset and the filtering rules are not readily available which can pose some problems in determining what the resultant table actually includes.

Exploring the Data Content

Before the data are downloaded or analyzed, it is important to understand the data. As noted in the exhibit of Traffic Counts, the headers indicate the field name and have two icons that can be used to illustrate content. The first indicates the format of the data (i.e., text or number), as well as the API field name. The second provides options for sorting, filtering, and hiding data. The Filtering option provides a preview of the data. In this case, there are several locational components: the North/South Streets, the East/West Streets, and

Exhibit 6: Exploration of Traffic Count Dataset



It is also possible to filter the data by condition or to specify certain fields. This can be helpful when dealing with large datasets and a narrowly defined project.

Scrolling to the right, the balance of the fields can be explored. In this case, two key elements are noteworthy: the lat/lon Coordinates, and the link to a more detailed report. Depending on the application, the inclusion of the lat/lon in a single field may suggest that the data should be downloaded in CSV or Excel format so text functions can be used to strip out the extra characters, and create individual columns. In the case of Tableau, the program automatically parses coordinates to generate lat/lon coordinates.

Exhibit 7: Coordinates

YEAR	Coordinates	Study LINK
2014	(39.162363°, -94.532373°)	http://data.kcmo.org/Traffic/CHOUTEAUPAR-1
2011	(39.154326°, -94.536038°)	http://data.kcmo.org/Traffic/CHOUTEAU33-TF
2010	(39.030856°, -94.50383°)	http://data.kcmo.org/Traffic/ARARATEAST-5-H
2007	(39.063982°, -94.590358°)	http://data.kcmo.org/Traffic/ARMOURBDWY-5
2011	(39.063615°, -94.580995°)	http://data.kcmo.org/Traffic/ARMOURGILLHM

In the case of the Traffic Signals, the hyperlinks open detailed reports for each of the intersections.

Exhibit 8: Detailed Traffic Reports

Kansas City Missouri Public Works
414 East 12th Street
Kansas City, Mo. 64106

W ARMOUR BLVD AND BROADWAY
WEATHER:CLEAR ROAD:DRY
COUNTED BY:AP / JS / DG
3 YEAR COUNT PROGRAM

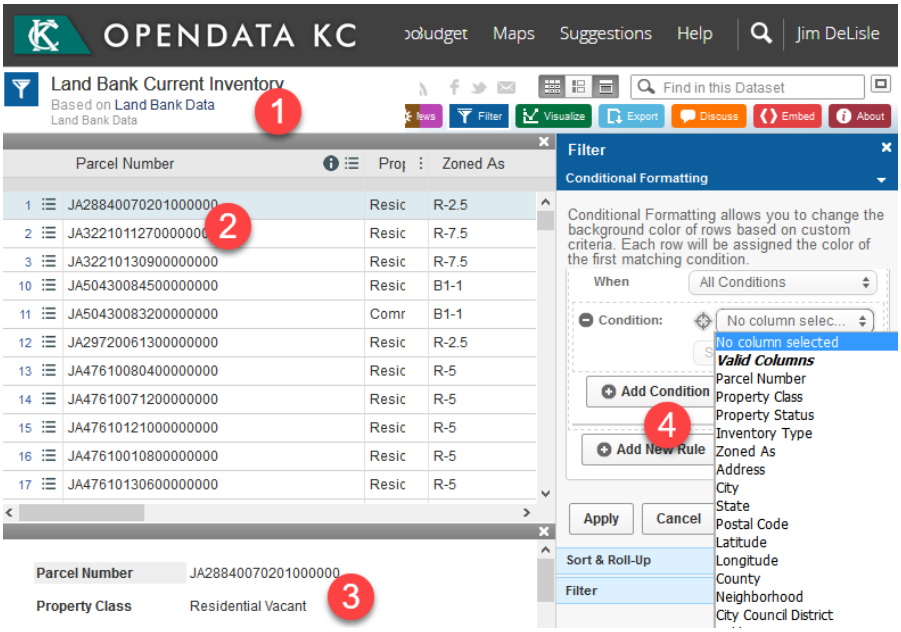
File Name : ARMOURBDWY
Site Code : 30005632
Start Date : 11/15/2007
Page No : 1

Groups Printed- Unshifted

	BROADWAY Southbound					W ARMOUR BLVD Westbound					BROADWAY Northbound					DRIVEWAY (ONE-WAY EB) Eastbound							
Start Time	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Right	Thru	Left	Peds	App. Total	Exclu. Total	Inclu. Total	Int. Total
07:00 AM	0	110	17	0	127	22	0	14	0	36	17	119	0	0	136	0	0	0	0	0	0	299	299
07:15 AM	0	113	29	3	142	45	0	28	1	73	12	126	0	0	138	0	0	0	0	0	4	353	357
07:30 AM	0	126	27	1	153	37	0	45	4	82	22	167	0	1	189	3	0	0	1	3	7	427	434
07:45 AM	0	175	38	1	213	66	0	41	1	107	28	205	0	0	233	0	0	0	0	0	2	553	555
Total	0	524	111	5	635	170	0	128	6	298	79	617	0	1	696	3	0	0	1	3	13	1632	1645
08:00 AM	0	187	48	1	235	50	0	28	0	78	19	167	0	0	186	0	0	0	0	0	1	499	500
08:15 AM	0	198	32	2	230	28	0	19	5	47	20	186	0	0	206	0	0	0	0	0	7	483	490
08:30 AM	0	131	19	0	150	35	0	24	3	59	21	116	0	0	137	0	0	0	0	0	3	346	349
08:45 AM	0	148	34	5	182	26	0	27	3	53	21	100	0	0	121	0	0	0	0	0	8	356	364
Total	0	664	133	8	797	139	0	98	11	237	81	569	0	0	650	0	0	0	0	0	19	1684	1703

If you filter views, it's important to document the filter as well as the date and source of data. This is especially important if the work is to be updated or the research/analysis will be replicated. Once the dataset has been selected, click on the link and it opens the dataset overview. This page includes: About this Dataset (i.e., date of update, number of views, number of downloads, the number of rows (records) and columns (fields), a listing of the fields, and a table preview). At this point it is important to explore the data to understand the content, format, coding, and other elements. The option you select depends on the nature of the data, your needs, and the analytical tools you will use to process or explore the data. If the data are going to be processed (i.e., scrubbed, filtered, transformed, or appended) CSV might be a good option. If the data are acceptable as is, they can be pulled into Tableau directly using the OData connection.

Exhibit 9: Filter Options



In the upper right, you can also find a set of options to access the data including: View, Download, API Connect, and … which leads to the option to access via OData.

Exhibit 10: Open Data Access Options

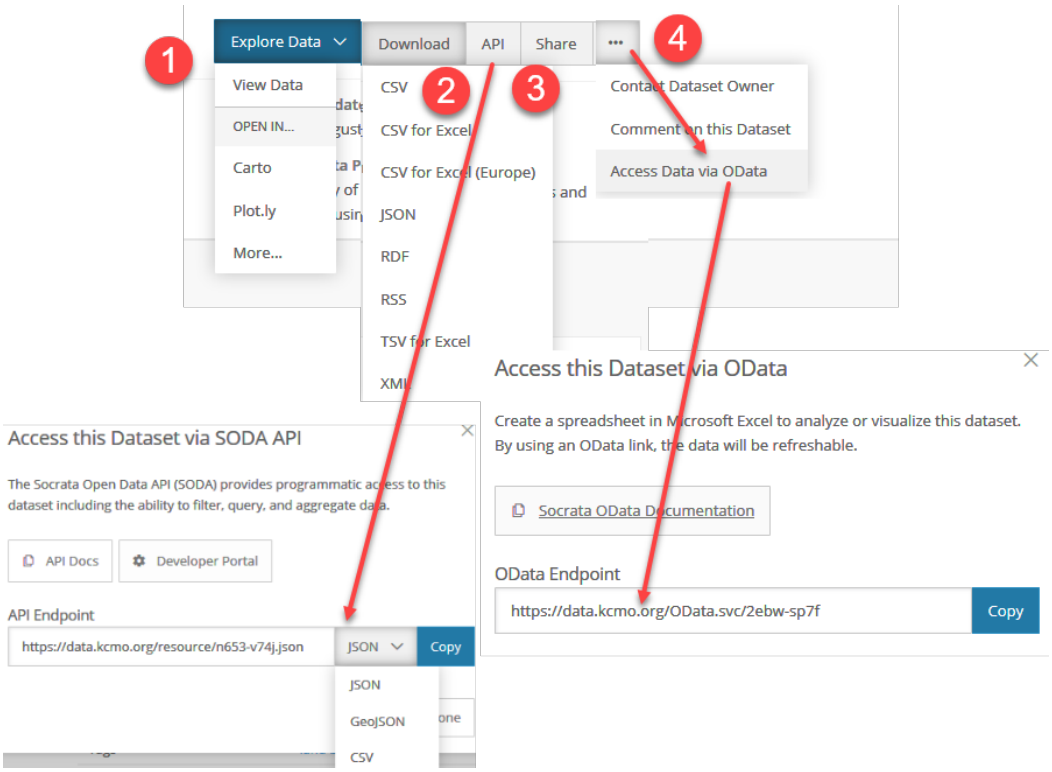
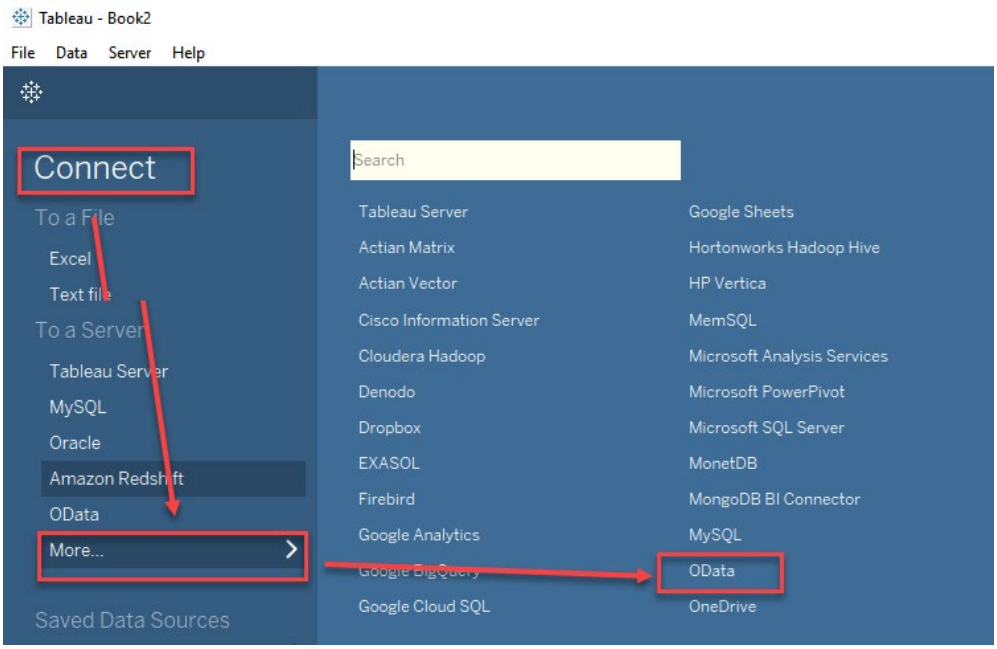


Tableau Data Connections

Connection Overview

To use Tableau you must feed it from an external source. To begin, Open Tableau and Create a New File and select Connect. As noted, there are a variety of formats that can be taken including Connect to a File, or Connect to a Server. Regardless of how you connect it is important to develop understand the data, both in terms of its contents, and format. Data that Tableau can process will be laid out in a tabular format, consisting of a set of rows and columns. In general, the columns will consist of fields with the first row the names, while the rows will be a set of records. The cell contents will be populated with either values, text, or spatial objects depending on the type of data it represents. It is important that the analyst develops a solid understanding of the data to ensure that it is not used inappropriately which can result in misleading results.

Exhibit 11: Alternative Types of Data Connections



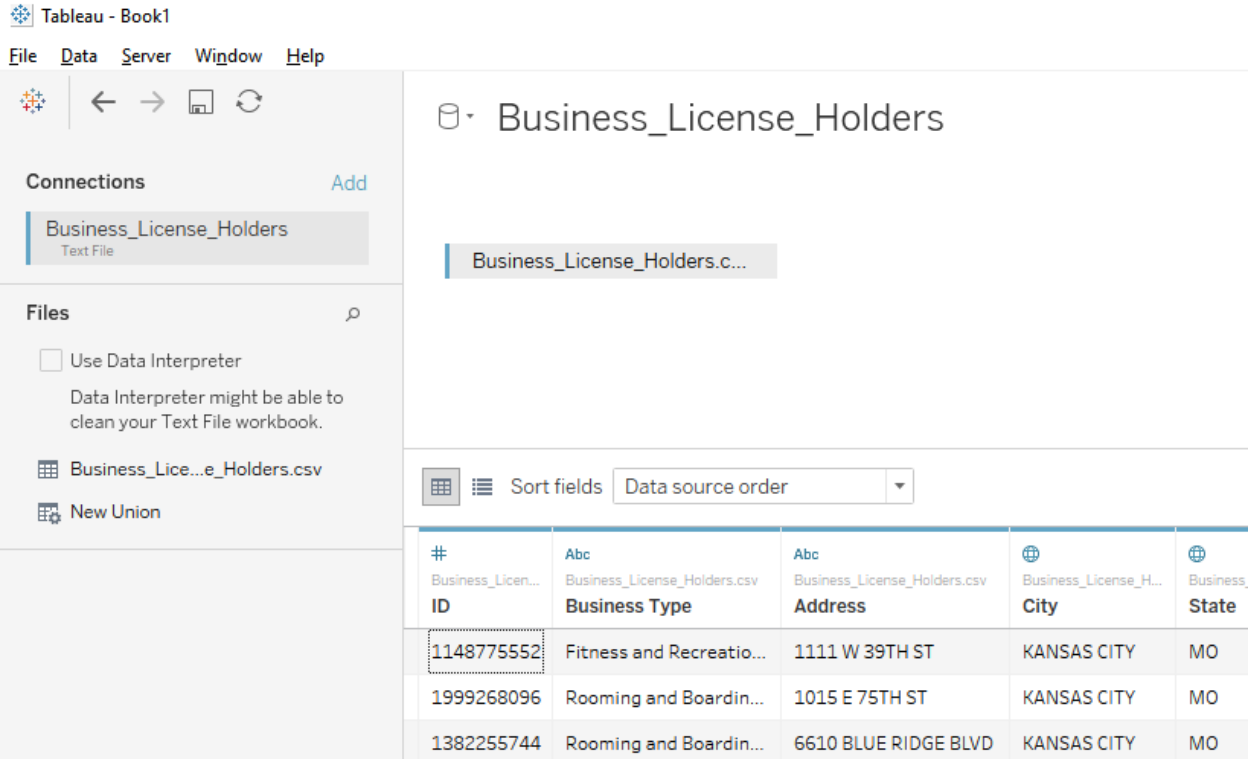
Connecting to Excel, CSV Files

In many cases you will connect to an Excel or “Comma Delimited” (CSV) file. When using Excel, if it is a multi-sheet file, you will have to identify the Worksheet to connect to in the workbook. Since CSV files to not support multiple sheets, the data connection is straightforward.

If the data are downloaded to CSV or Excel, they can be explored before feeding into Tableau. For example, by using Pivot tables, categories of data can be explored. Additionally, new fields can be generated such as calculation of PED/Vehicle ratios or calculation of Quintiles/Quartiles which indicate whether the counts are high or low compared to all of the records.

Exhibit 9 shows the connection to a Business License Holders CSV file that was downloaded from the Open Data portal.

Exhibit 12: Business License CSV Connection



Connecting to KC Open Data

Connecting Tableau to Data

In this case, we are connecting to the Open Data

Paste OData Endpoint copied from Open Data portal. Once pasted in the Server box, use the enter key to activate the connection. You do not need a key to complete the connection. Merely Copy (CTRL-C) and Endpoint from the Open Data Portal, toggle back to Tableau and Paste (CTRL-V). Then, press enter and Tableau will attempt to extract the data into your workbook.

Exhibit 13: OData Endpoint for Land Bank

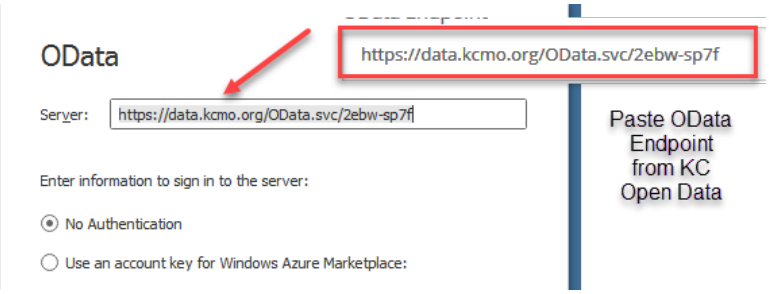
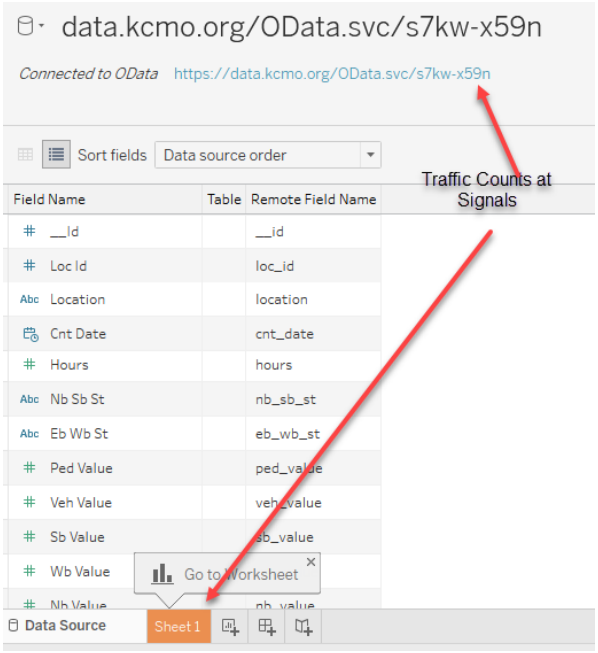


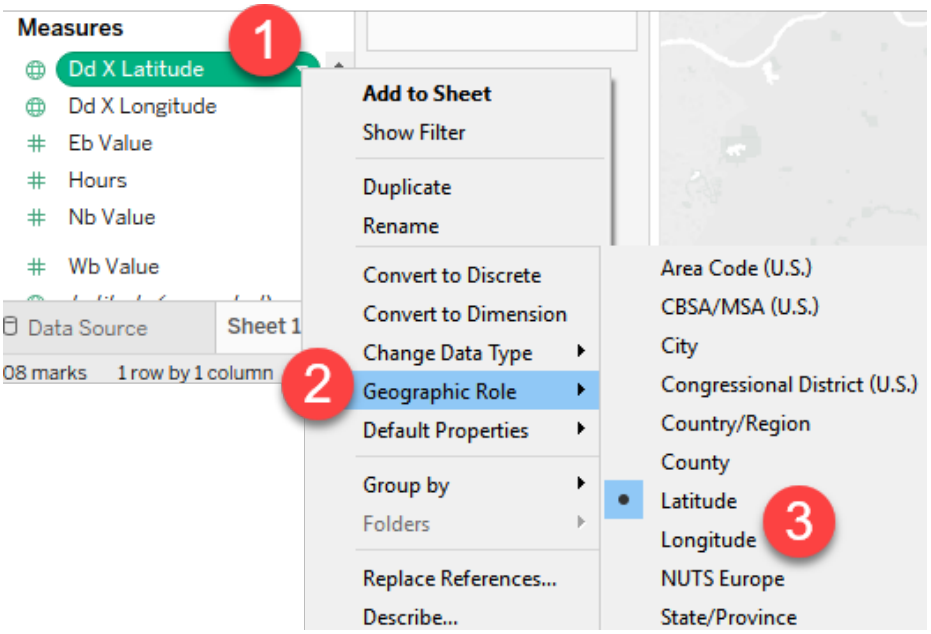
Exhibit 14: Summary of Data Source Variables



Setting Up Tableau Sheets

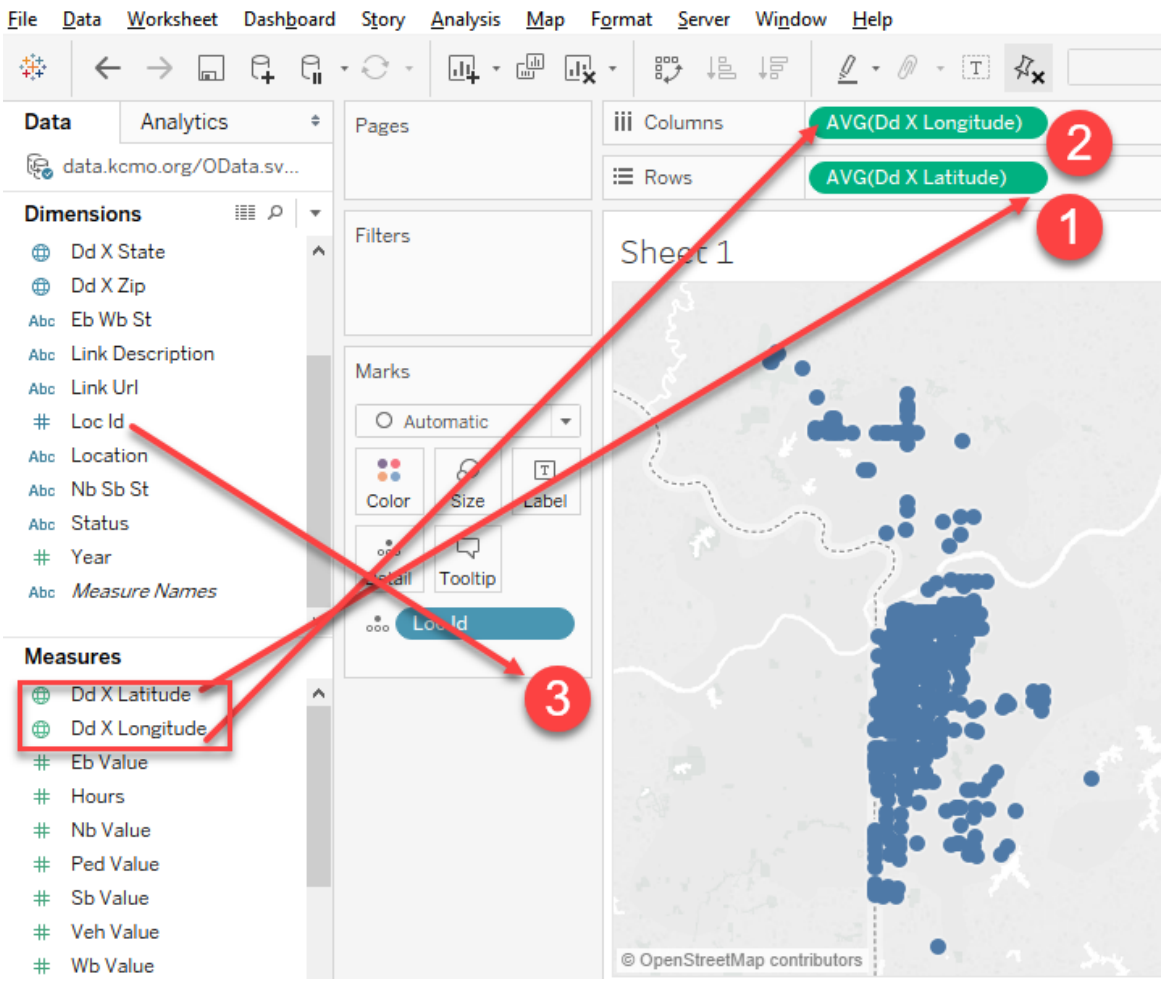
Once the connection has been made (or data imported from files) the Data Source will be highlighted. Go to Sheet 1. As noted, Tableau uses a matrix or tabular format. If the data has the lat/lon coordinates, the columns/rows will be displayed as lon/lat with a geographic reference. To check, the fields should be in the Measures (quantitative) section and have the Globe icon. If they are lat/lons but do not display the icon, they can be converted.

Exhibit 15: Verification of Geographic Role



Once the data have been connected to and you open Sheet 1, the base map can be set up. This is a three-step process. First, drag the Lat to the Rows field. Second, drag the Lon to the column field. In some cases, the Drag and Drop may not work. If so, click on the Row field to anchor the cursor, and then double-click on Lat under measures and repeat for the Lon. Third, drag a nominal variable in the Marks section which should populate the map.

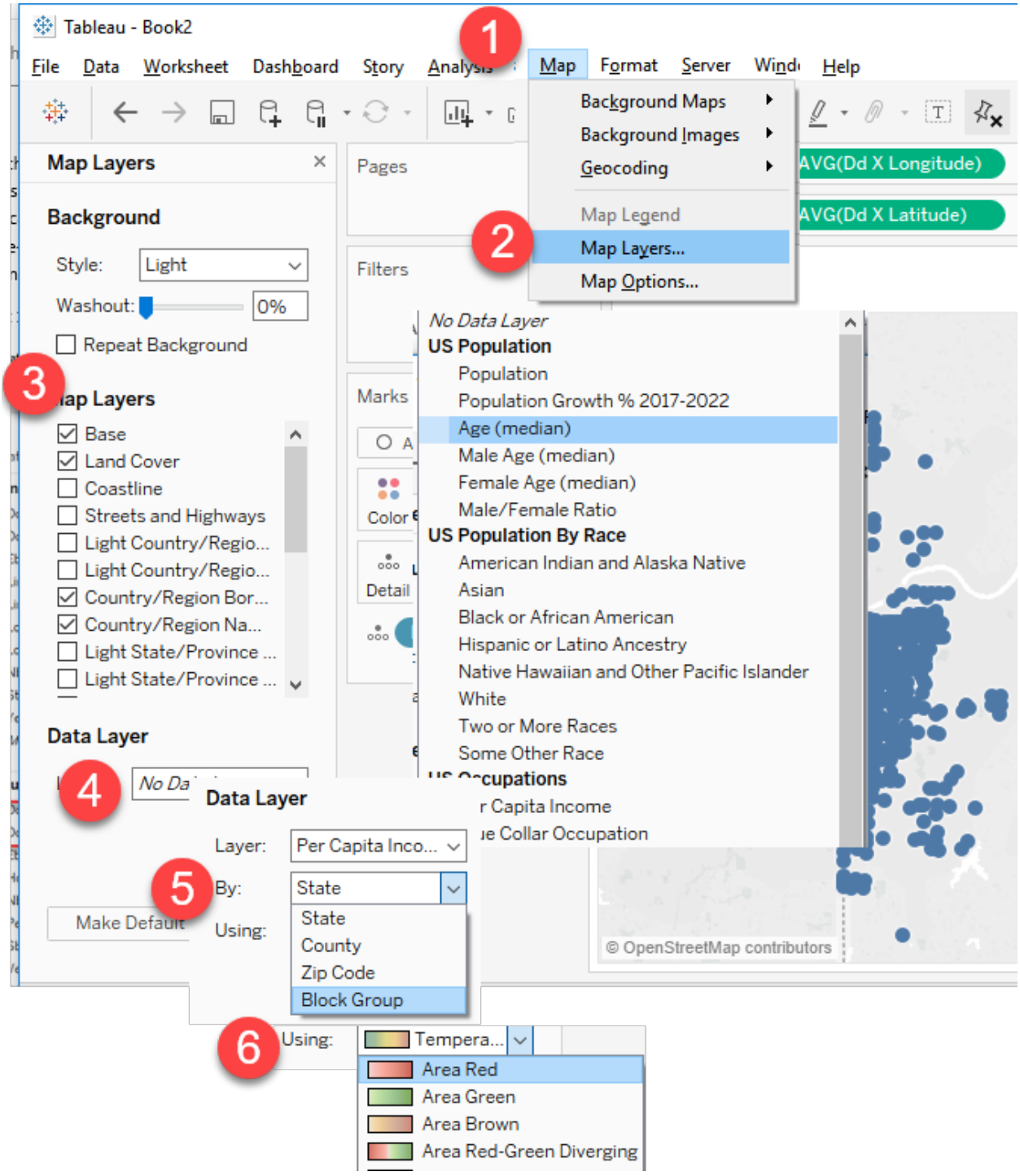
Exhibit 16: Initial Setup for Base Map



Map Data Layers

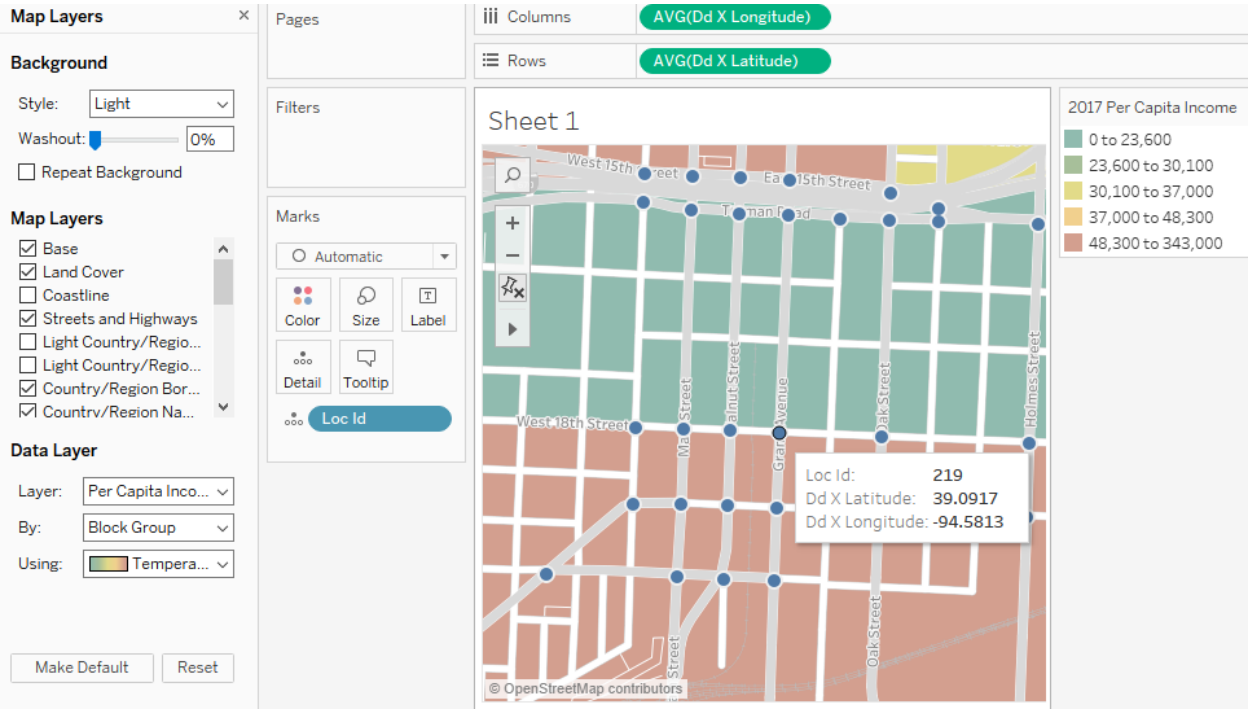
Once the data points are displayed, you can change the data layers in the background: To do this, click on the Map Layers and select the map features (e.g., Streets and Highways). Now, select a Data Layer, indicate the Geography, and select colors with the Using tab.

Exhibit 17: Map Layers



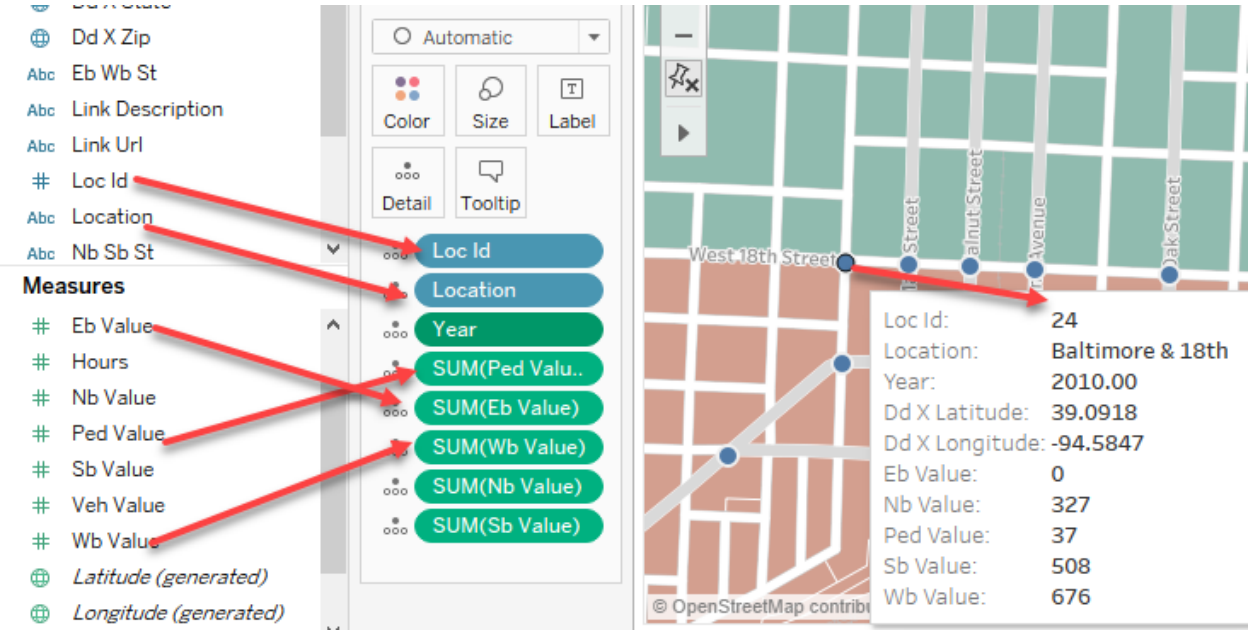
Once the Map layers have been selected, the Base Map displays the results. By zooming in, the roads can be viewed. The Legend is automatically generated which in this cases, shows Per Capita Income. When a user mouses over one of the datapoints, the Fields noted in the Marks section and the lat/lon will be highlighted.

Exhibit 18: Traffic Count Zoom



Turning on displays for maps

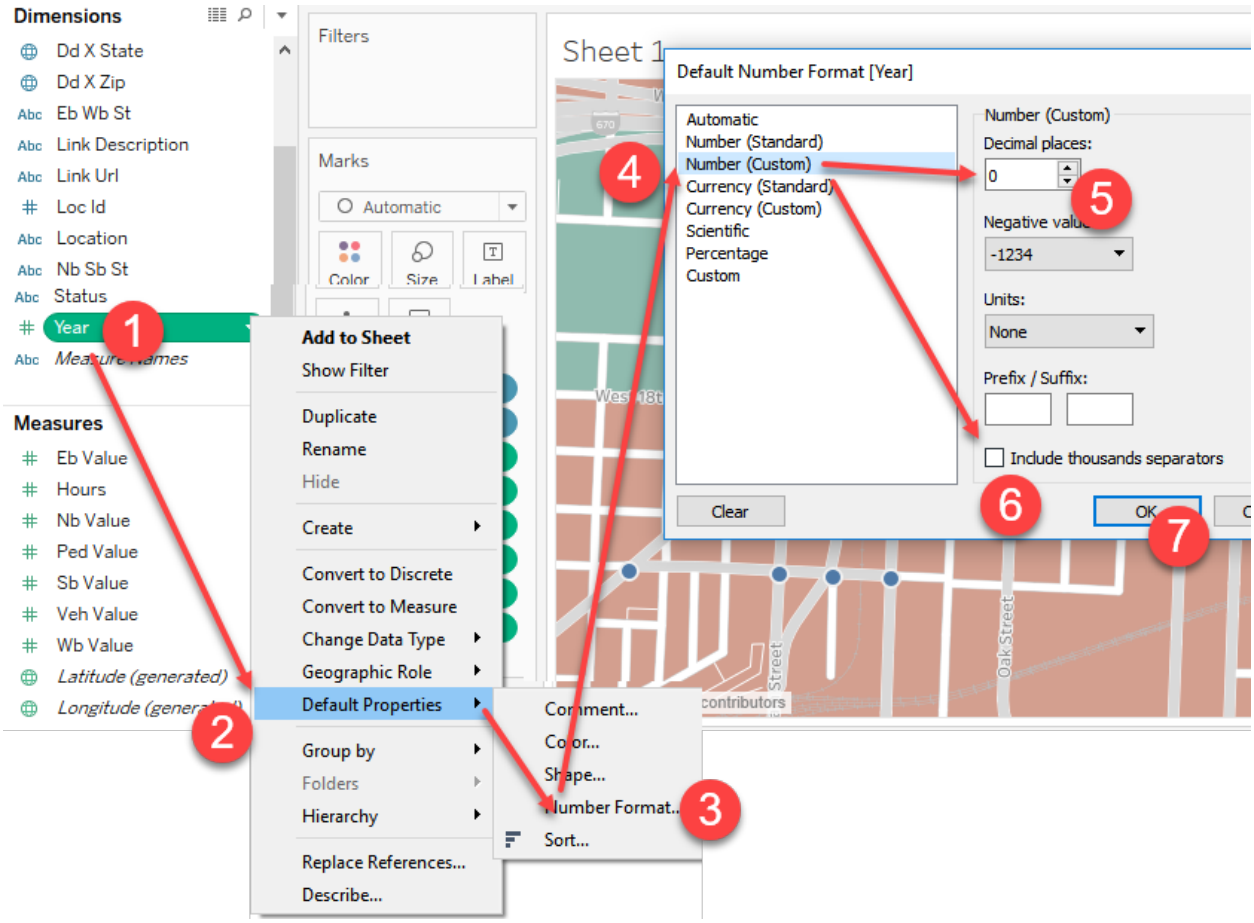
The default setup for the display variables includes the Lat/Lon, as well as any variable you drag to the Marks section. To add, simply drag variables to the Marks and then mouse-over to see results.



Change Number Format

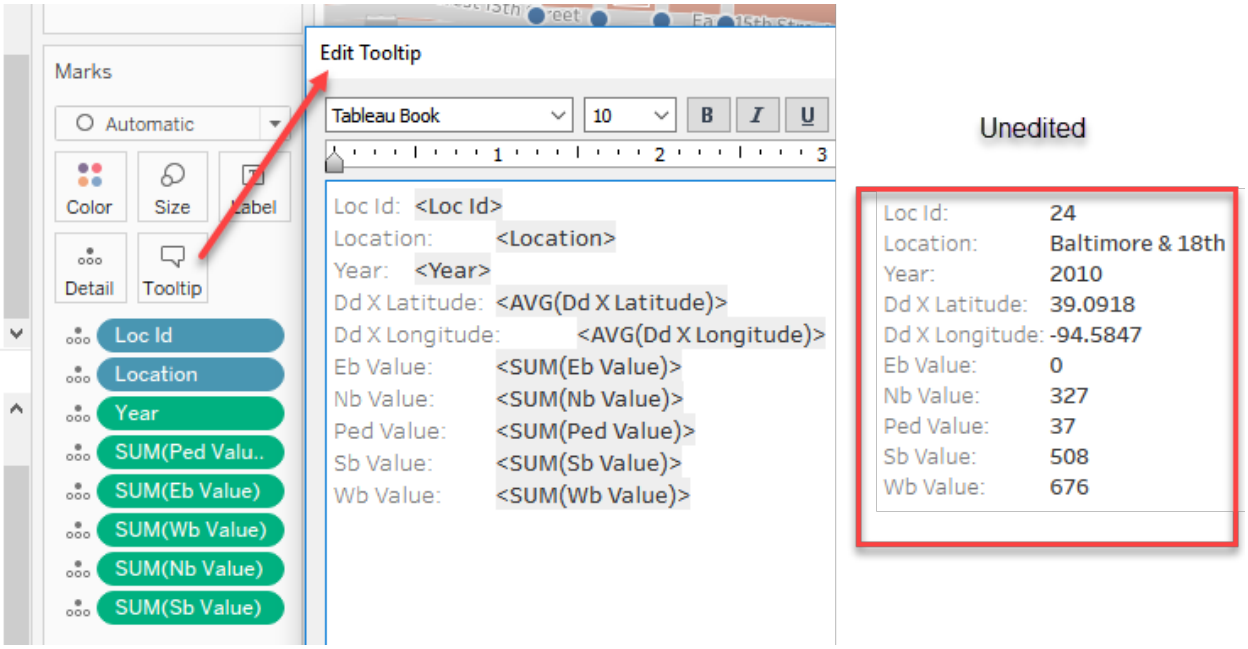
As noted, the Year displays as 2010.00 which is the default setting for numbers. To change it to 2010, change the default properties as noted, the decimals should be changed to 0 and the thousands indicators (i.e., 2,100) checked off.

Exhibit 19: Changing Number Defaults



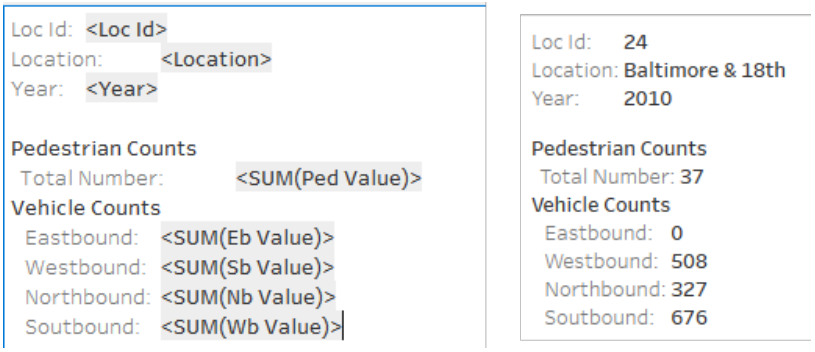
To change the Display options, edit the Tooltip from the default presented in the Exhibit. This is done by clicking on the Tooltip icon, then Editing the results. As noted, the left displays the Field Name, and the fight in <BRACKETS> the values. In this case, the lat/lon are not needed for consumers, and the Pedestrian and Vehicle counts are in alphabetical order.

Exhibit 20: Default Tooltip Values and Results



To edit the Tooltip, items can be deleted (i.e., lat/lon) and field names can be changed. Blank lines can be added, and the order can be changed. The easiest way to change the order is to highlight one of the lines, Copy, Delete and Paste in the desired order.

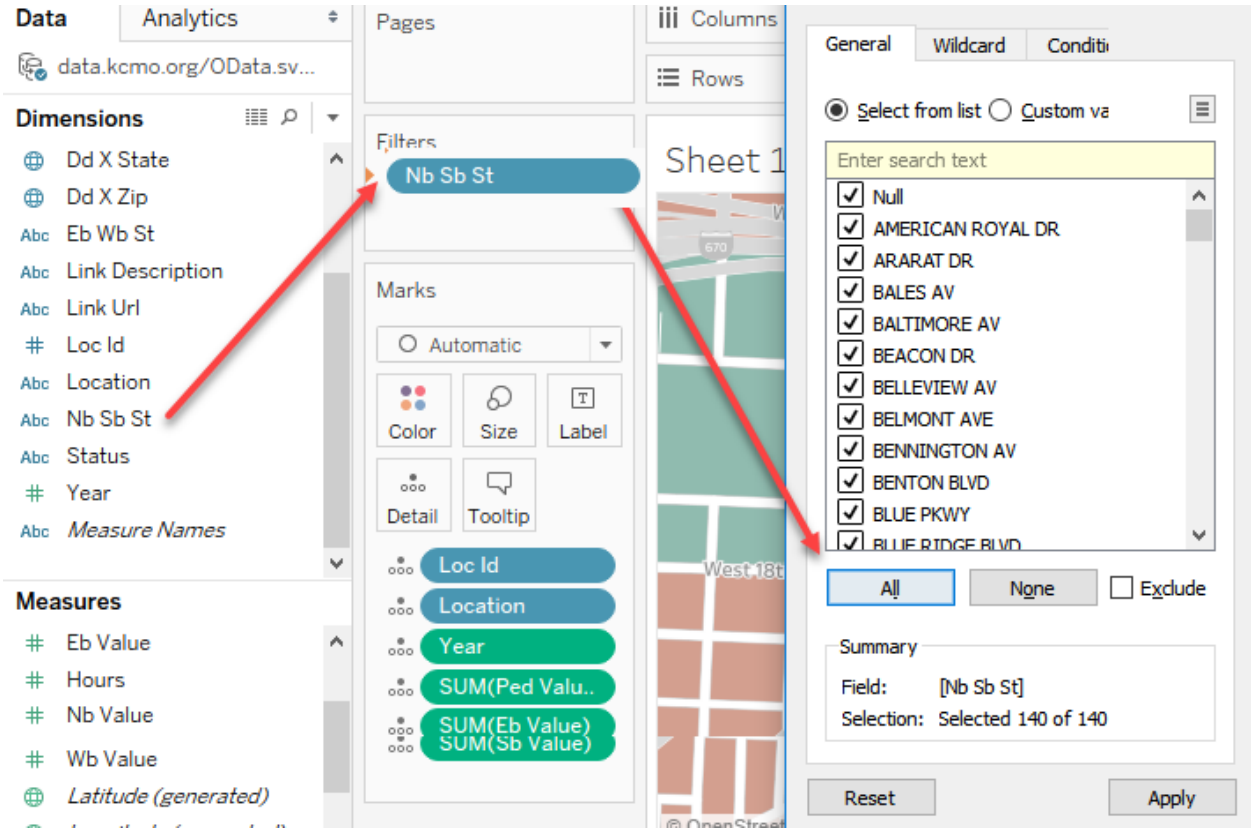
Exhibit 21: Edited Tooltip



Filtering Data Display

Overview

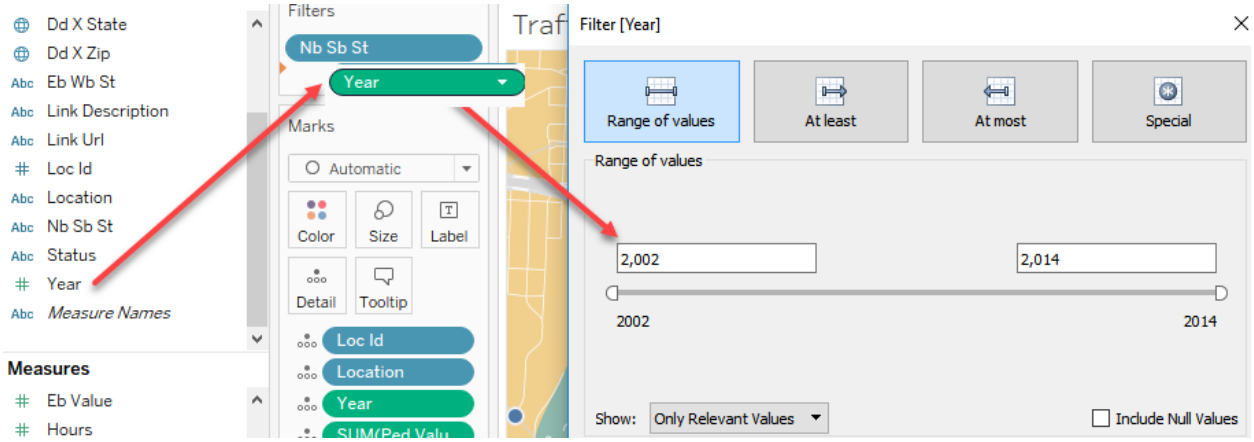
To create an interactive Dashboard, you can use filters that can be toggled off and on. In this case, the filter is set to Northbound/Southbound Streets. This is done by dragging the variable to the Filter Box and then selecting options from the list. As noted, there are 140 unique records.



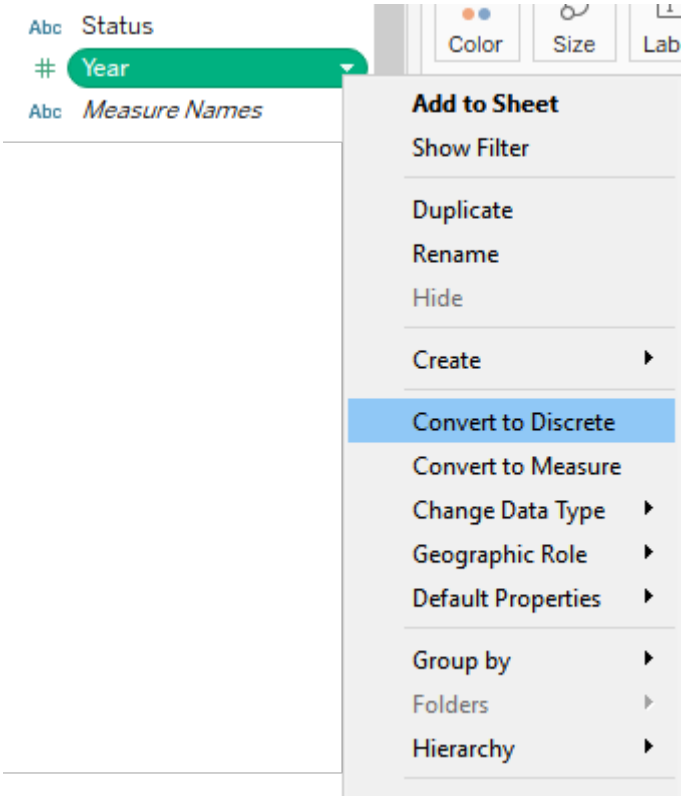
Filter by Year: Discrete

In this case, assume you want to filter by year. If you use the same drag and drop, the results may not let you filter by year (e.g., 2010, 2011) but come in as ranges. This occurs due to a data format issue. That is the data are formatted as continuous rather than discrete (i.e., whole numbers). However, the continuous feature is also useful as in the case where you want users to select houses that fall within a cost range, or do not exceed a maximum value.

Exhibit 22: Continuous Variable Filter Options

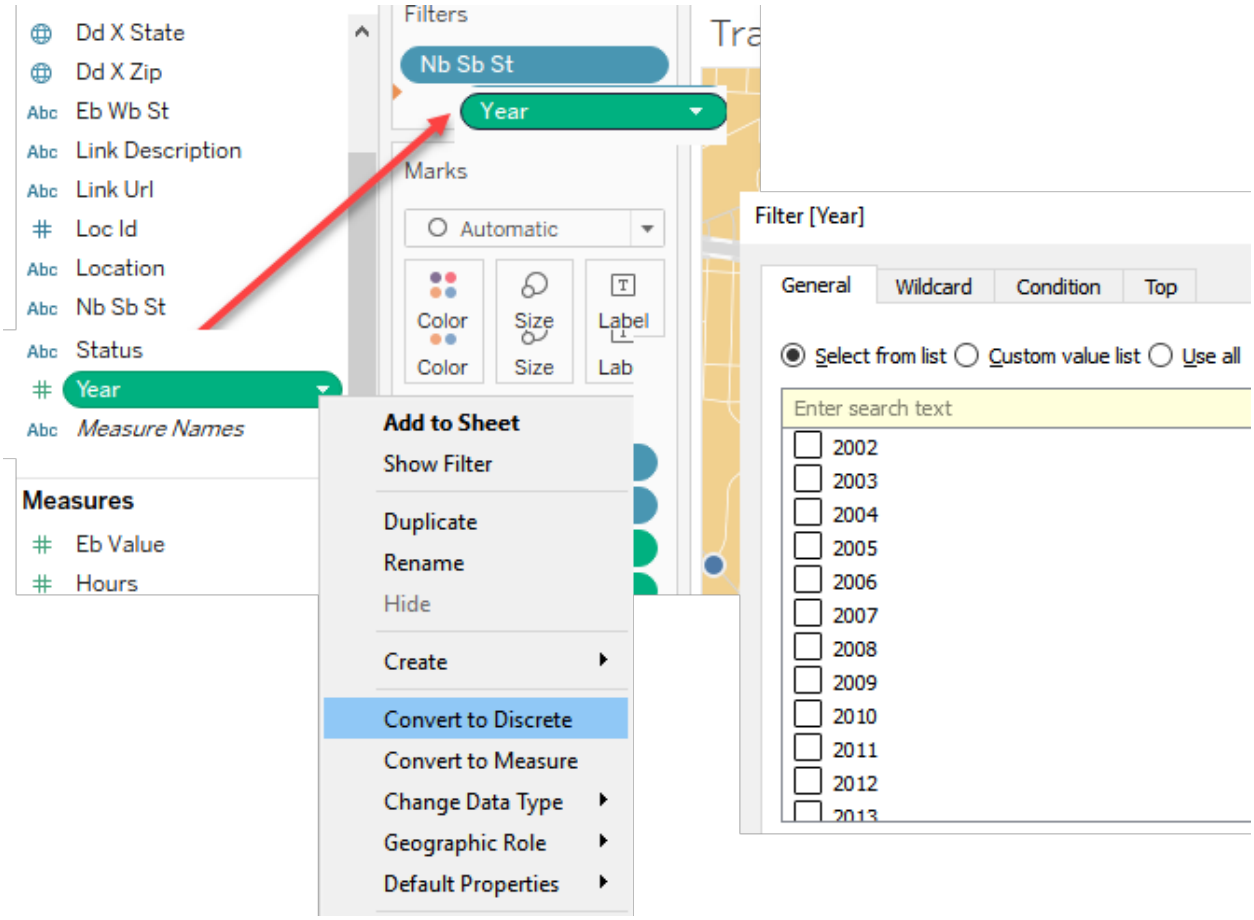


To use specific years or numbers to filter when they are Continuous, the data type must be changed. This is achieved by converting the value to a discrete field and then dragging the variable to the filter.



To change from Continuous to Discrete and vice versa, right click on the Year and select Covert to Discrete; note this is a toggle so if the data was discrete Convert to Continuous would appear. Now Select All and the user can filter by year of data update.

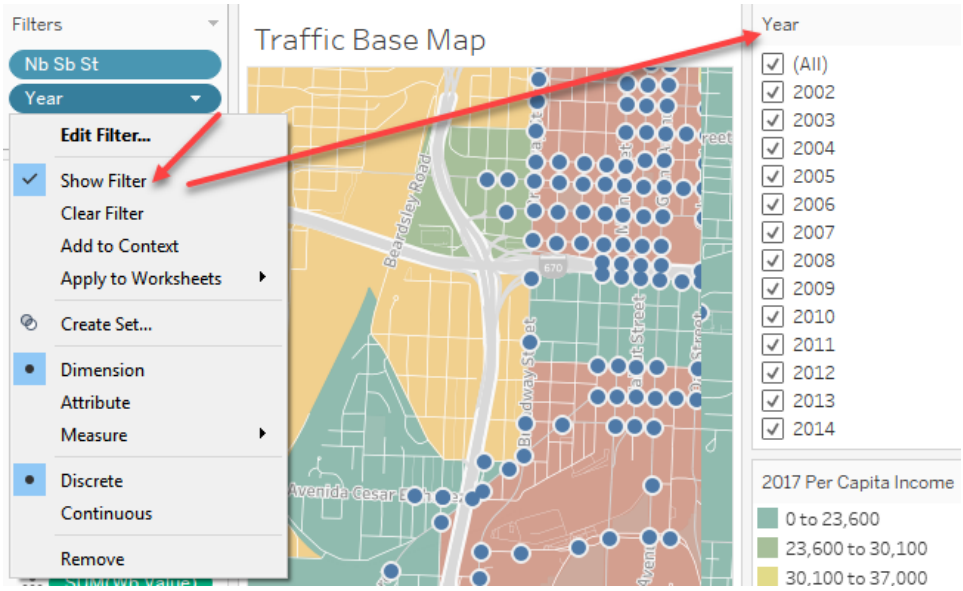
Exhibit 23: Discrete Filters



Make Filter Visible and Change Selection

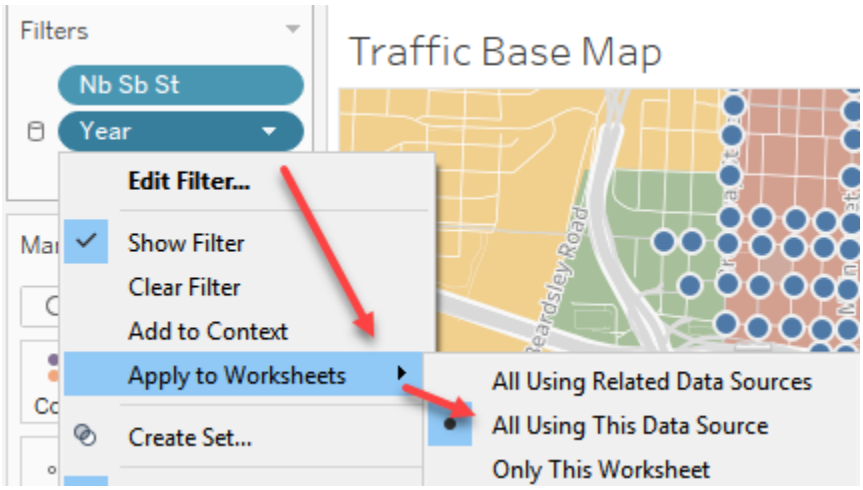
The next step is to activate the filter in the map. This is done by using the dropdown to Select Show filter. As noted, the Filter will appear in the upper right of the map.

Exhibit 24: Showing Filters



A couple of refinements can be added now. First, the Filter can be applied to individual sheets selected by the user, or to all sheets. Again, this option is selected by using the dropdown and indicating the desired behavior. Since we only have one sheet set up at this point, it might not appear to make a difference. However, if you are adding sheets with other information and/or creating a dashboard, it is an important feature. If All Using is selected, when the filter is selected only the records satisfying those values will be displayed. That is, if you have a table, the rows and columns along with the values will be displayed while the others will be suppressed. If the user clicks on NONE, the sheet will disappear, and then reappear as filter criteria are expanded. Since filters can be layered –multiple filters added—the behavior should be thought through. In most cases, having them apply to All is a good default option.

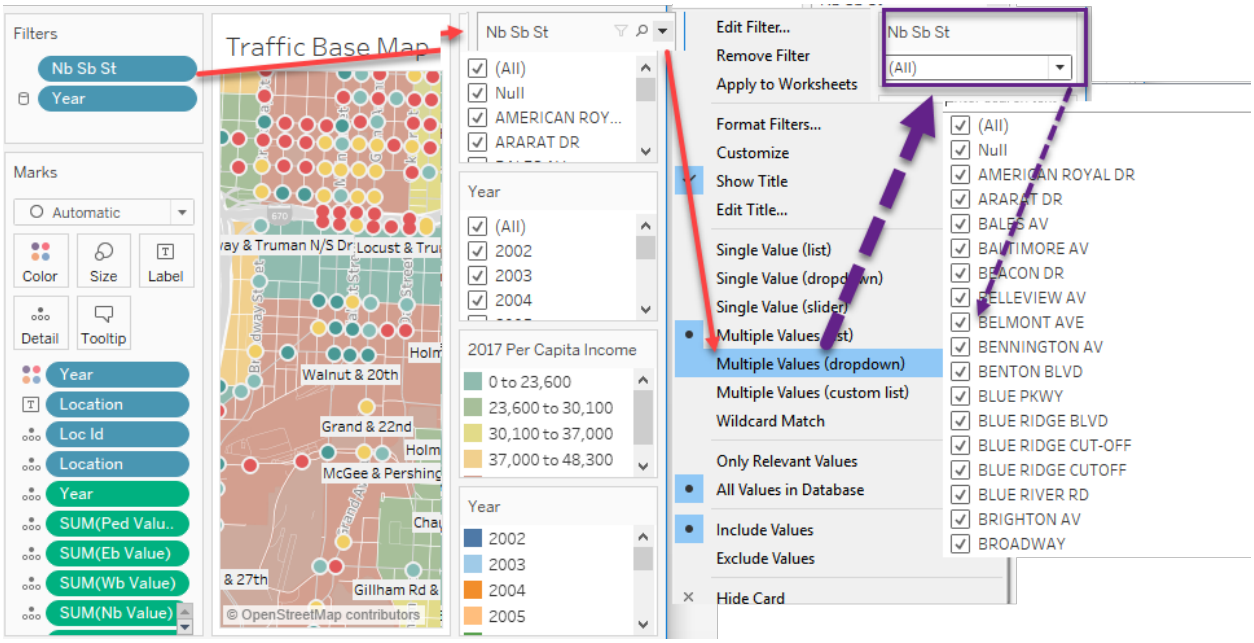
Exhibit 25: Applying Filters



Modifying Filter Views

Once the Make Visible option is selected, the filter will be added to the right of the map/display. As noted the default is a list of the values which the user can scroll down and select. If the list is long or there are multiple filters, this can be tedious. Similarly, if the user wants to walk through a set of filter values, the scroll may be a barrier. As an option, the filter can be changed to a Multiple Item Dropdown. As noted, the display seen by the user will have the variable name, the All default, and the ability to expand the entire list of options and make individual selections.

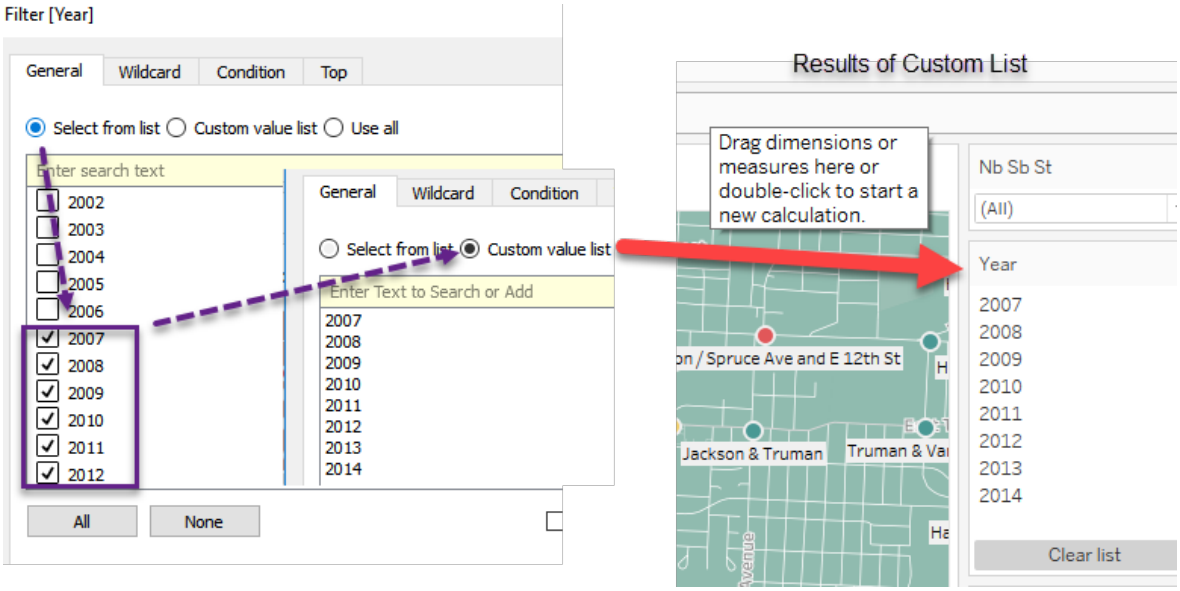
Exhibit 26: Filter Options



Customized Filter List

In some cases you may want to restrict the number of options that are offered through the filters. This can be done by selecting a subset of options. In this case, the data extend back to 2002 but the designer only wants to display the past 10 years since the older data could be misleading. This is a multi-step process: select desired values and select Custom Value option.

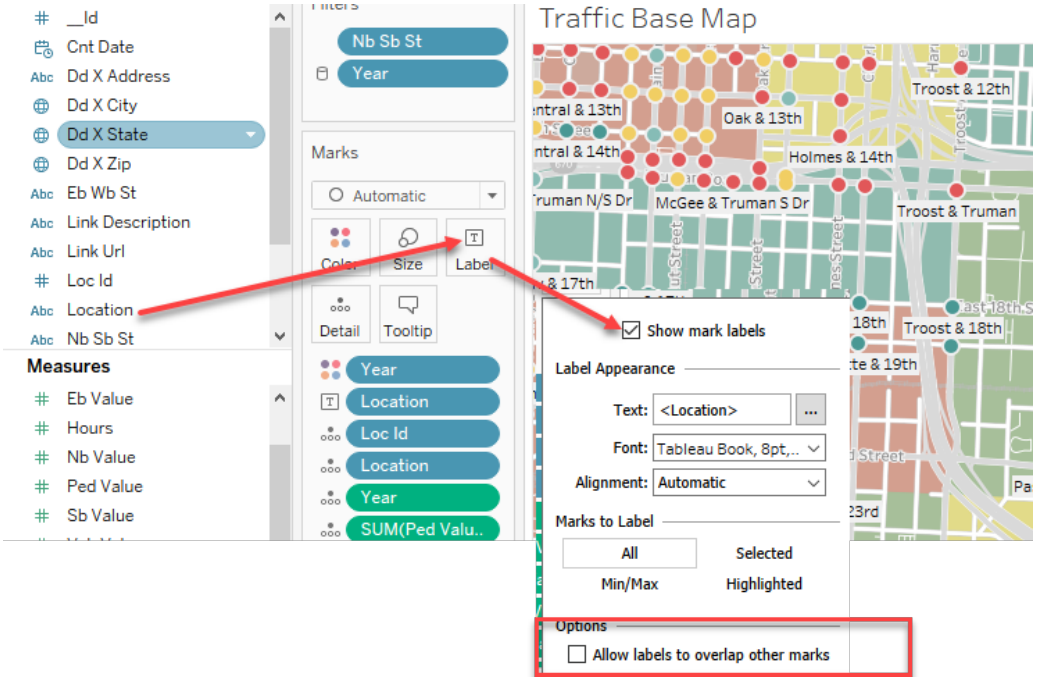
Exhibit 27: Customized Filter List



Adding Labels

Although you may have a background map that displays street names, since the labels will not activate and be legible until a user zooms in, Labels can be turned on. This is done by dragging a Field to the Label Icon, and then selecting the font and behavior. In the default case, the labels will not overlap but will pop up when zoomed in sufficiently. This can be preferable where there are many records or the field is long. To change, the Allow labels to overlap can be selected.

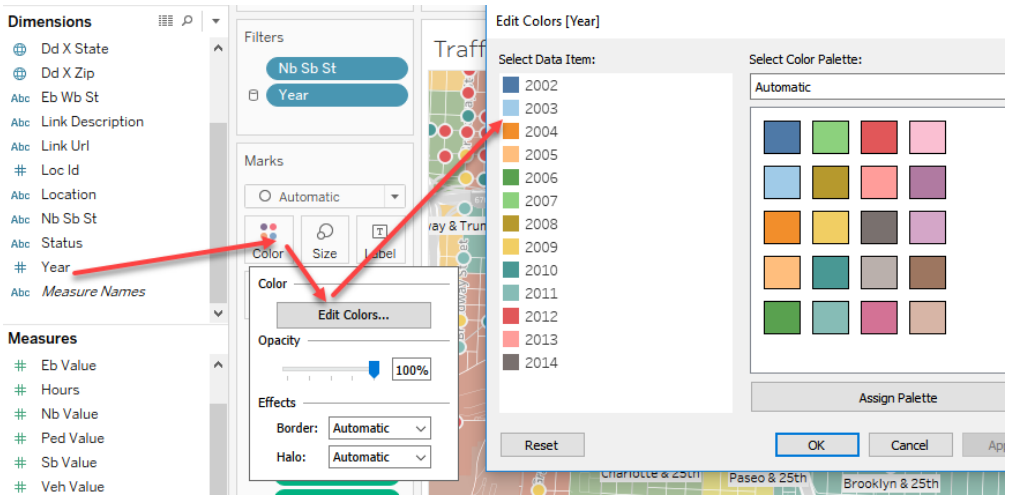
Exhibit 28: Record Labels



Adding Colors

To make it easier to see patterns, the records can be color-coded using one of the fields. If the data are Eiscrete, the colors will show up as patterns. If the data are Continuous, the results will be a continuum of light to dark with a given color.

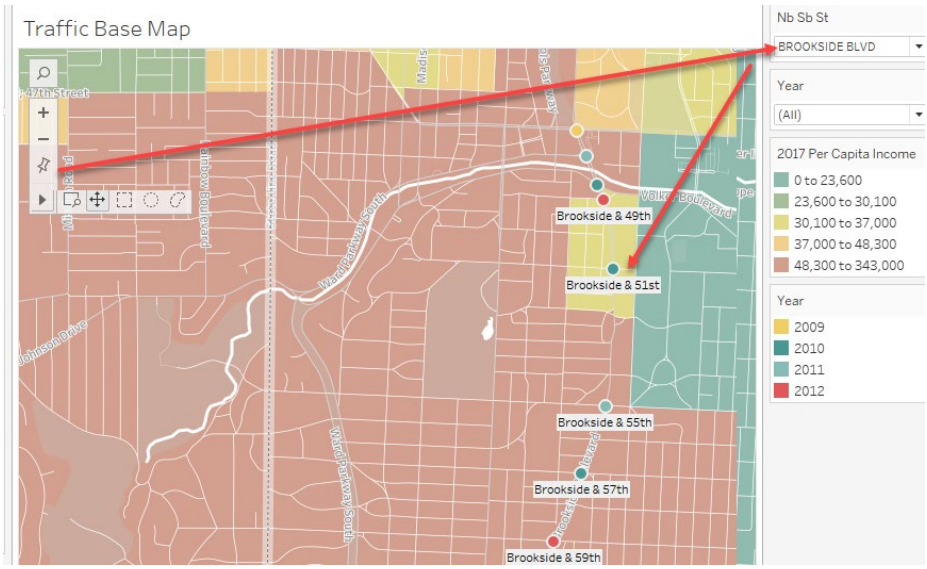
Exhibit 29: Adding Scale Colors



Controlling Snap to Filter View

In the upper left of the map are several icons for designing the dashboard and offering options to users. If the push-pin is selected, the map will pan/snap to the geographical area that displays the filtered data. Without this feature turned on, the user will have to pan and zoom to find the results. In this case, BROOKSIDE BLVD is highlighted.

Exhibit 30: Pan to Pushpin Filter Results

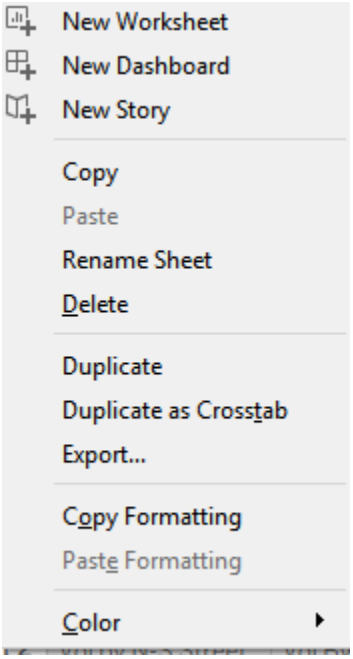


Managing Sheets

Naming Sheets

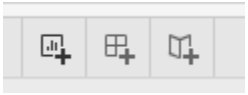
To this point we have focused on the Map sheet. While an important component, other types of data visualization may be helpful. In addition adding sheets provides more input options for the dashboard. To organize sheets, it is useful to name them. This is done by Right-Clicking on the Sheet name tab and selecting the Rename option.

Exhibit 31: Managing Sheets



Creating New Sheets

New sheets can be added by using the optional tabs, at the bottom of the current sheet, or right-clicking and choosing New.



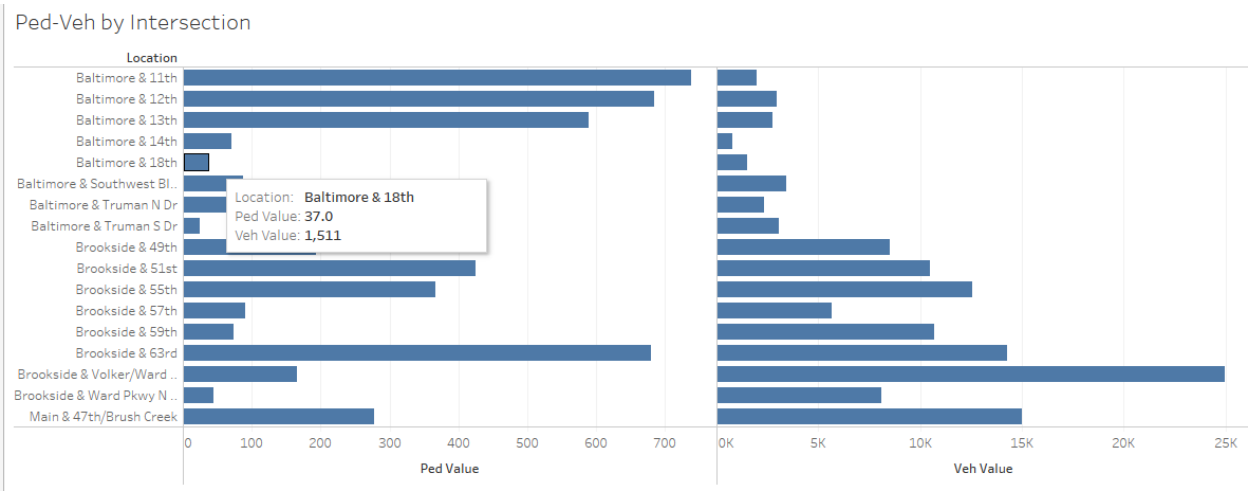
Duplicate Sheets

Sheets can be duplicated from the right-clicking and choosing Duplicate. This can be handy when fine-tuning a map and wanting to avoid destroying what has been created. The designer can then delete sheets to focus on the revised or enhanced sheet or go back to the default.

Multiple Value Sheets

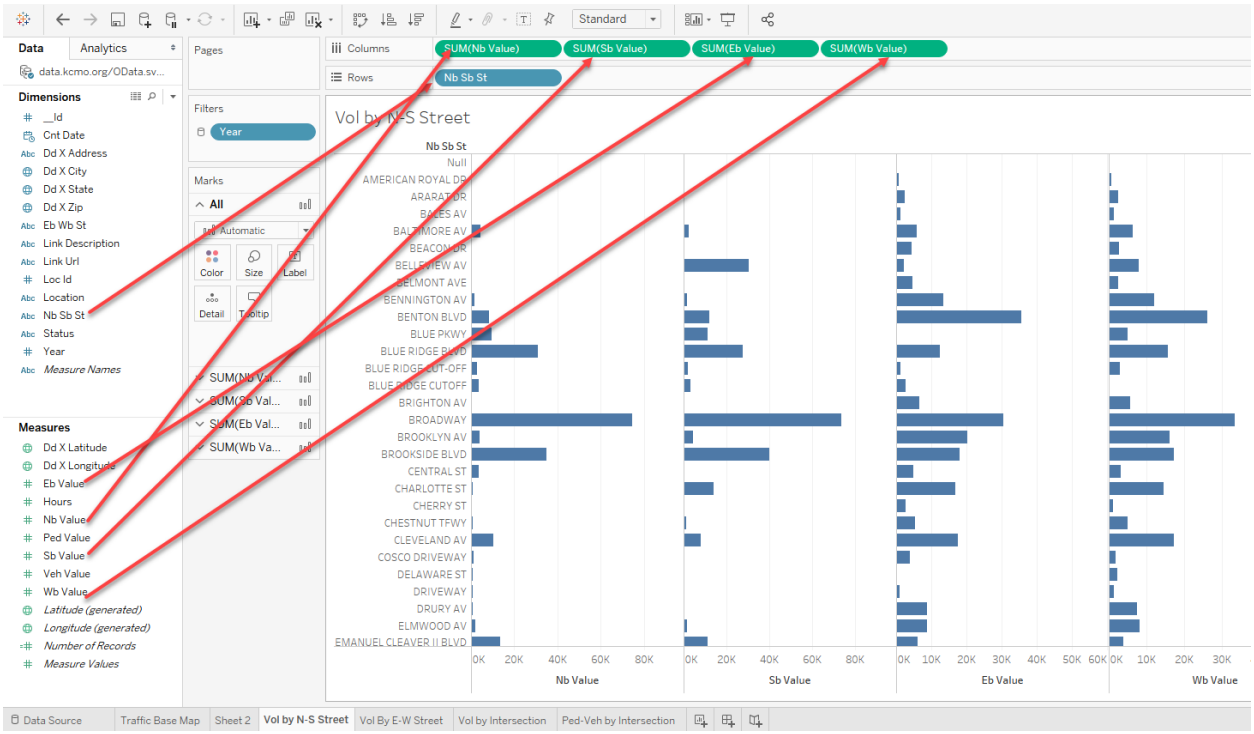
In some cases, the designer may want to compare records across several fields. This can help visualize patterns and compare and contrast results. In the exhibit, the rows are the North-South Streets in Kansas City. The columns contain the vehicular traffic volumes by direction (i.e., North, South, East and West). The scale is set automatically and is floats by volume for that direction. In this exhibit, the Pedestrian and Vehicular counts for two filtered streets is displayed: Baltimore and Brookside.

Exhibit 32: Filtered Multi-variable Display



Data can also be mixed to tell a story about an intersection. For example, a designer may want to help users visualize the level of pedestrian traffic at an intersection to determine whether there is demand for the retail component of a mixed use project. In addition they may be concerned about traffic volume to get a sense of how "pedestrian-friendly" an intersection might be, as well as how much traffic goes by that might be drawn into a site. As noted in the detailed report section on page 6, in addition to the Open Data in the table to which Tableau is connected, there are more detailed traffic reports available. In particular, these reports indicate volume by time of day. This information may help in determining whether a restaurant should focus on breakfast and lunch, or whether dinner and evening might fair better. Similarly, it could help an analyst determine which side of the street a drive-in should be located so customers do not have to fight traffic. A good example of this would be a coffee shop or a "take and bake" food service where the destination of the vehicle (i.e., to work, to home) might be an important factor.

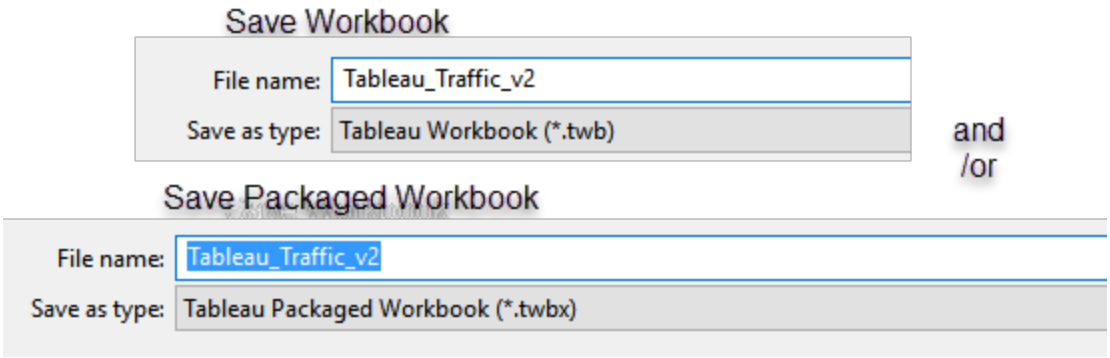
Exhibit 33: Traffic Volume by N-S Street by Direction



Save Workbook and Packaged Workbook

It is recommended that designers save Workbooks along with way with version numbers. This can help when changes are inadvertently made that cannot be "undone" after a file is closed. There are two ways to save the file. The first is "Save to Workbook" which includes the set-up and the data connection. When another user opens the Workbook, the connection will have to be refreshed. The second is "Save to Packaged Workbook." With this option, the file will be bigger but the data will be embedded for quick access.

Exhibit 34: Workbook Save Options



Publish to Tableau Public Server

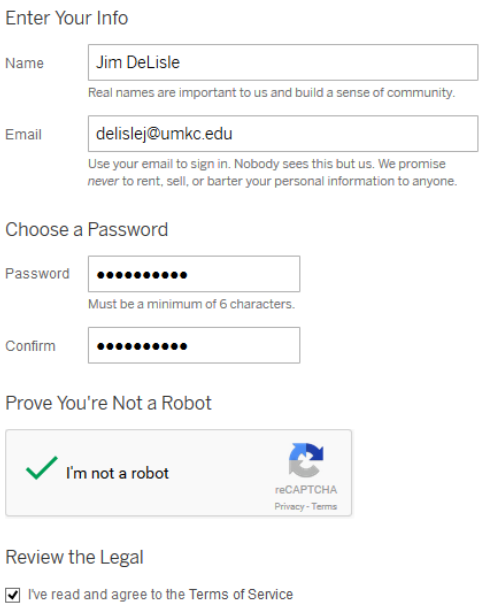
Sharing Files

There are several options for sharing Tableau Dashboards. If users have Tableau, the Workbook or Packaged Workbook can be sent or posted. In some cases, users may want to open a Tableau Online account. That allows selective sharing to other Tableau users, and has options for restricting downloads and modification or opening options. If users do not have Tableau or if the designer wants to share the results, the Workbook can be posted to the public Tableau Server. The author will have the option to indicate which sheets are accessible, and whether the data and/or Workbook can be downloaded by all users.

Publishing to Tableau Public Server

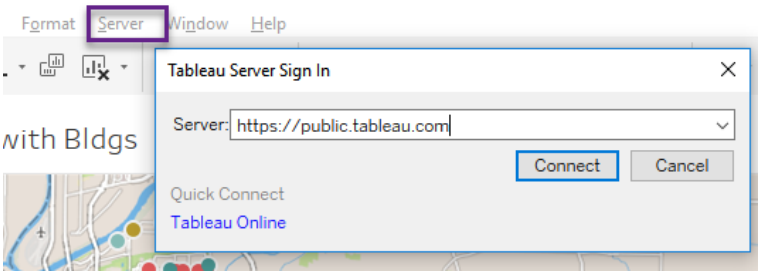
To make your dashboards available to others without Tableau, you can publish them to the Public server. You can set the permissions and then embed the link in a webpage or send the URL to interested parties. Note that your file will be accessible to the public and they can download, modify, etc.

Exhibit 35: Create a Public Profile



To publish to the Public Server, log into the Tableau Public server from within Tableau.

Exhibit 36: Log Into Tableau Server



You will be prompted to Log in to Your Account so you will be the author of record.

Exhibit 37: Log in to Your Account

Edit Details to control various options for viewers of the visualization.

Exhibit 38: Edit Details of Tableau

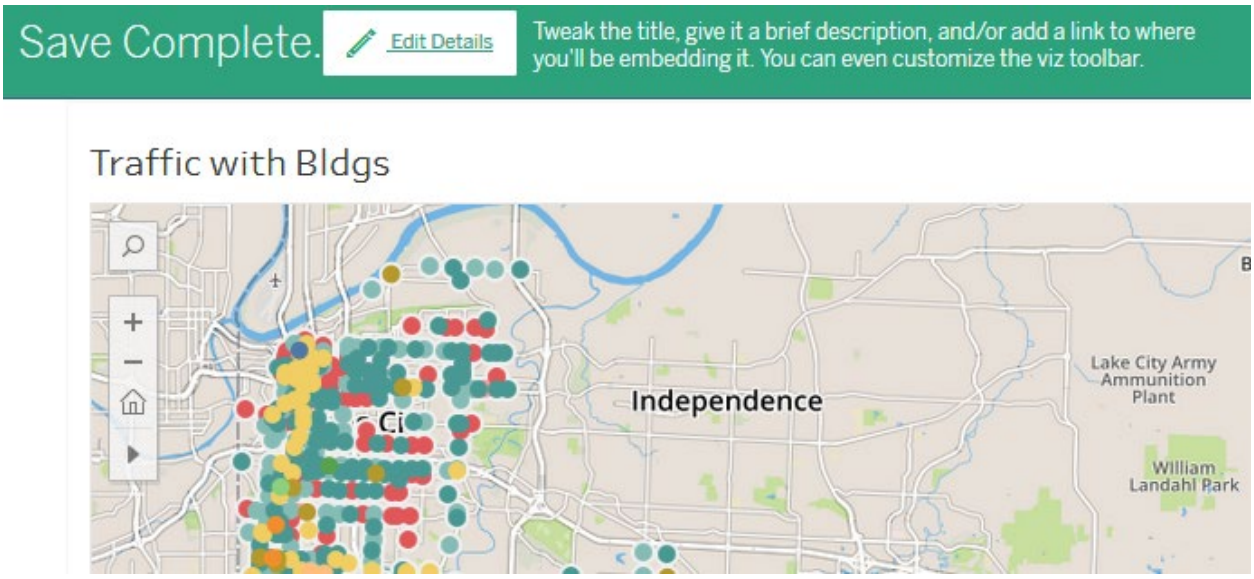
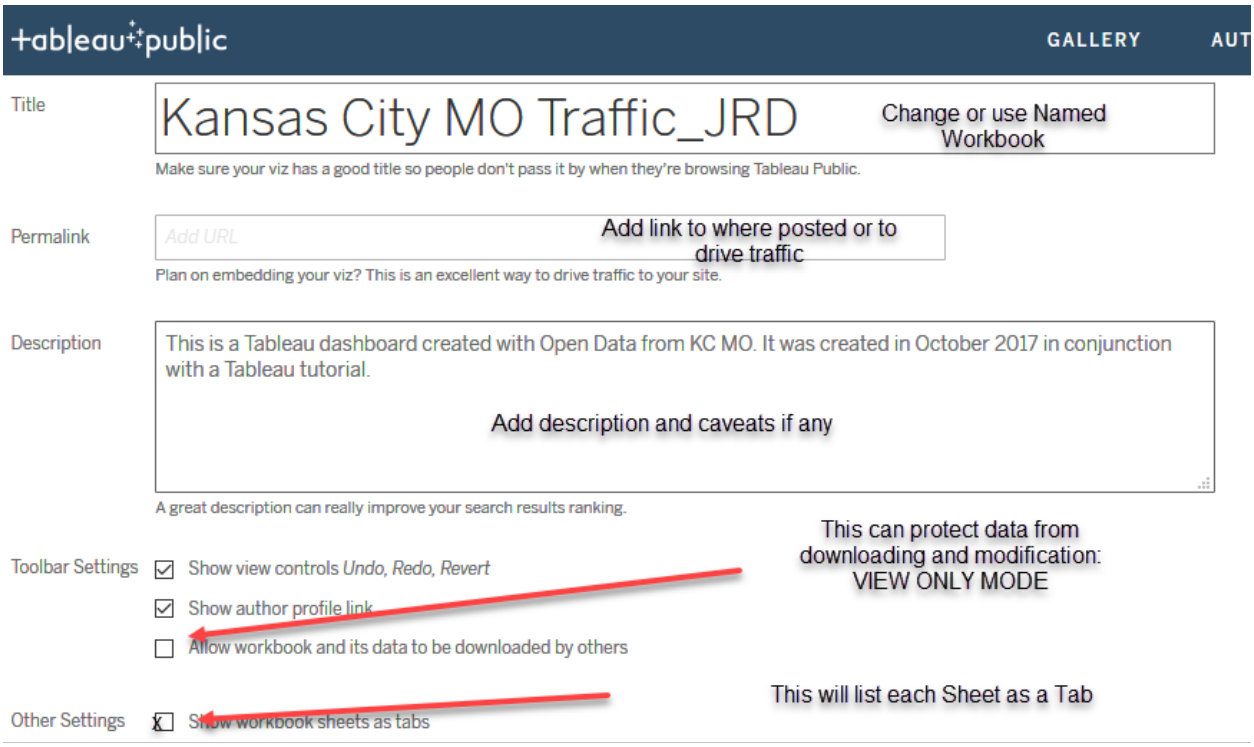


Exhibit 39: Edit Details



Share via Link or Embedded Code

Options

At this point, sharing options can be selected including creation of a link that can be shared, to embed code that can be placed in a website.

Permanent Link

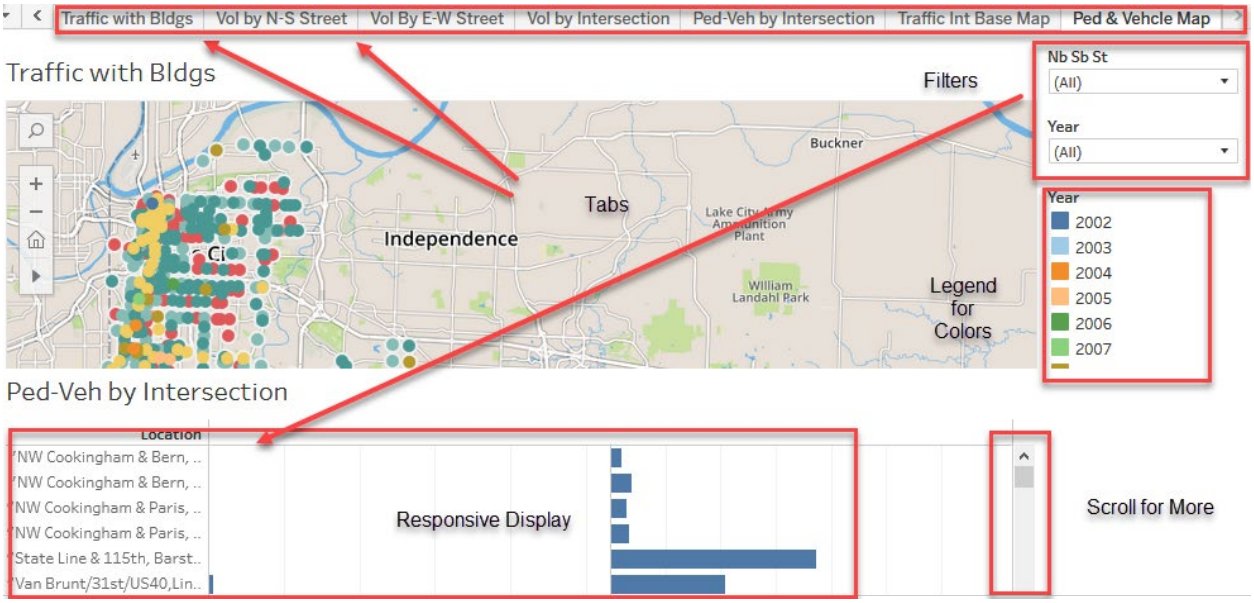
https://public.tableau.com/views/KansasCItyMOTraffic_JRD/PedVehcleMap?:embed=y&:display_count=yes&publish=yes

Embed Code

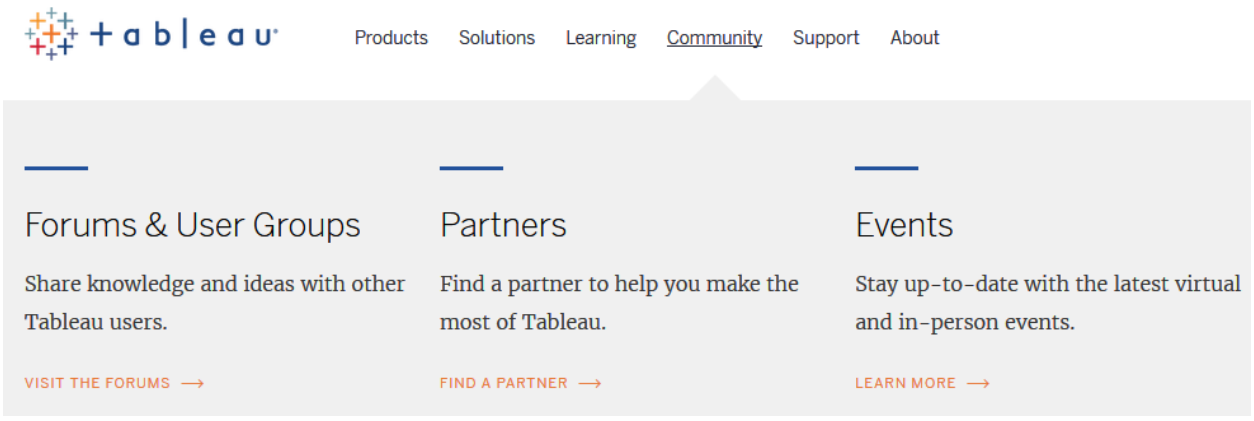
As an alternative to creating a link, the author can copy automatically generated code that can be embedded in a website.

The actual visualization will depend on what the author publishes. In this case, the 8 sheets are presented as separate tabs that the user can navigate through to explore the data and make decisions.

Exhibit 40: Published Tableau Visualization



Conclusion

The objective of this tutorial was to introduce students and other readers to some of the funcationality of Tableau. This was accomplished by documenting the creation of a Tableau Dashboard and to publication of the resultant file on a step-by-step basis. The case study draws on Open Data from Kansas City MO using the OData connection. When the data are refreshed on the Open Data portal, assuming the same structure, the visualization can be automatically updated.

It should be noted that Tableau has many more features and applications than explored in this introductory tutorial. Interested users can download a free trial. Tableau also makes fully functional copies available for free to students and instructors under its “Tableau for Teaching” program. In addition to using the package during their studies, students can request a free extension beyond completion of a course. Users are also encourage to explore Tableau’s website to and Community to get a better understanding of its many features and applications.