Tableau Basics: Your first bar chart

1. Tableau can work with various files (excel / csv / access / sas / spss / r / mysql / oracle etc.)

2. All variables from table will be broken into two types into worksheets-

Dimensions: Categorical or qualitative variable; Independent variables

Measures: Numerical values; Dependent variables

3. We can do ordering or sorting the values of a variable by sorting icon

4. Calculated field: Variable created from other variables; Right Click on measures and create calculated fields. Look for the calculated field symbol after creating the new variable.

5. Color: Important for effective communication. Use to convey message.

6. Label: Use it to make it easy to understand the data level in visuals. If sometimes label doesn't appear, Right click-mark label-always show.

7. Format: Format the graph well, keep things consistent.

8. We can export the visuals for producing the image of worksheets or to paste it somewhere else.

Time series, Aggregation, and Filters

1. Humans like unpivoted view, Machine like pivoted view. Take care of it.

2. Live vs Extract; Extract data make a copy before for the analytics, while live data is always open and might change if the data is connected with live database. We'll usually go with extract, live might be slow. After working with extract data, we can return to live after analysis if needed later.

3. Need to always save workbook before we start to work.

4. While working with Time Series, Right Click on the time variable in column or rows, and select the right level of granularity, like unemployment rate vs month (Jan, Feb...) or unemployment rate vs month (Jan 2004, Feb 2004.... Jan2015, Feb2015...). Notice that, this changes the data type color from measure to dimensions or vice versa. Also notice how tableau doing the aggregation based on the right level of granularity.

5. Tableau, In General, assumes highest level of granularity and aggregate the data based on that if no granularity is provided to tableau.

6. Click Analysis >> Aggregate measures; It will aggregate the data at the right granularity level, unticking it will disaggregate the data at the right level. But not on the total or biggest aggregation rather aggregation will happen based on the chart data or other variable we are already using on the worksheet. Wooh!!

7. Introducing the new dimension in the visuals will change the granularity/ aggregation level of

the chart.

8. Variables in the column or rows used in the visuals, in general will be aggregated as sum(X), But right click on the variable in visuals >>measure(sum) and we can always choose other way to aggregate it like average(X), median(X)…, etc.

9. If we want to make our chart change the granularity level based on a dimension or measures without effecting our current visualization state, we can drag the dimension or measures to details tab. Since dragging it in the colors, shapes or other way will change the face of visualization we already have. Details here actually means the level of details or granularity.

10. Highlighting: Click on certain color or shapes in the legend and notice how it highlight the visuals with relevant data. If highlighting doesn't work, please active the highlighting button look like pen, which is in the top of the shape or color legend.

11. For multiple line plot in time series, it's better to convert it into an area chart. Area chart will put all the lines as stacked one on top of another.

12. Add the variable to filters tab when you want to filter down your visualizations based on that variable. After adding to filter, Show filter to worksheets as quick filter. Adjust the filter type you like to see.

Maps, Scatterplots, and Your First Dashboard

1. Joining: Join tables together based on a certain common variable. Inner Join and right join is mostly used. We can control what kind of join we want and based on what field we want to join it.

2. Tableau can create duplicate rows during joining to create a complete table if we have fewer rows in one table than other table.

3. We can create joining based on multiple common field.

4. Geographical hierarchy

5. You can control color intensity based on your variable's value (Red for big, continuum, blue for small)

6. We can apply same filter throughout all worksheet. R-click to filter >> Apply to worksheets >> All using this data source.

7. Create Dashboard: with two or more worksheets.

8. If filter is not coming to the dashboards, then click the downside triangle button of the worksheet you moved to the worksheet and show quick filter.

9. Add Intractability throughout various graphs. For that we need to add action. Tableau have two type of action – filtering and highlighting. Try Dashboard >> Actions >> Add actions. We can do it here or from the downside triangle button of the worksheets added to the dashboard. Click Triangle

button >> use as filters. Inquire the actions you just created by filtering by looking dashboard >> actions.

10. Actions by highlighting: filtering delete all the dots other than relevant one, highlighting keeps all of them but highlight the relevant one. Add highlight by dashboard >> actions >> add actions >> highlight. Highlighting and filtering works different way. When we filter something tableau, only keep that value, show it while removing all the values in the background. For highlighting tableau doesn't run that process in the background.

11. If highlighting doesn't work after adding it, it can be because of same variable from different rows, like if there are two names as Aaron from both Kentucky and NYC where we are trying to make the graph based on state, how tableau will understand which one to highlight? It's because highlight only highlight doesn't calculate the data in the background like filtering. For that we can change the granularity level to the right type. Then highlighting will work.

Joining and Blending, Data, PLUS: Dual Axis Charts

1. Joins: Inner, Right, Left, Full

2. Duplicates in joins: for multiple match during the joining, some of rows might be duplicated.

3. Joining on multiple fields: Based on different granularity level of different tables, we might have to join data on multiple fields for correct joining results.

4. Joining vs Blending: If two table have very different granularity, then joining them we might lose a lot of useful information or data. Like the granularity in the table A+B the date and in C it's month. If we join them based on date/month and category, we loose item, order id and so on. In this case we will use blending. If the detailed data is not important for us, we can use the join. Also, if our data source is different like one csv, another excel file, we can't join them, blending is the way to go in that case.

5. Blending: For blending we need to upload the data twice. In the worksheet we can see two data connections. If we take a look at both data at our worksheets, we can see sometimes tableau already pick up some blending based on some common variable from the two table which shows as a chain button. If we want to manually blend, we can make the name of the variables from two table as same, so that tableau can understand where it has to blend. To manually do it we can go Data >> edit relationships and add the relationships of the variables we want to blend. Blending is actually a smart Left join. If blending is not participating actively in the visuals, blends will not show active status. We can make it active by clicking on the chain button.

6. Dual Axis Chart: Adding two charts in a same graph and common axis. R-click on the graph and click dual axis. The graph will become a dual axis chart. After that we must need to R-click on the

axis and synchronize the axis to make the chart right.

7. Creating Calculated Fields in a Blend: We can create a calculated field which need variables from different tables. Be careful about the aggregation level of the variables while writing new calculated field.

Table Calculations, Advanced Dashboards, Story

1. If our variable doesn't give us any geographical roles automatically, we can set it up manually. R-click on the variable and set the desired geographical role.

2. We sometimes have to assign our state or city or provinces to specific country manually if it cannot be shown on the map automatically.

3. Tableau sometimes automatically generate variables, like number of records. Sometimes it might be useful.

4. Table Calculations: Make the number to percentage, average or so on for the variable. R-click on the variable in visuals and add the table calculation or quick table calculation.

5. Create bins to control your size of your bar chart/histogram. For that R-click on the variable you want to create bin and then create bin with your required dimension. Notice that if the variable was in measures, it will be in dimensions after adding bins.

6. Control bin size by creating parameter: Right click on measures > > create parameter. Give it a name, type and size etc. Look at the worksheet and you will find the parameter already been created. R-click on it and show parameter. But if we change parameter slider nothing will happen because this parameter to any variable. Now if we R-click on the bin variable we created and edit it on the size of bins we can choose the parameter already been created. And doing so will link our parameter to the bin.

7. Tree Map: No tree, all boxes.

8. Create Storyline and add the insights

Advanced Data Preparation

1. Understand pivoted and unpivoted data

2. Unpivoted data is good for tableau, because all the dimensions/variable have a separate column

3. Use Data Interpreter to do the initial formatting of the raw data (review the results when you use the data interpreter)

4. Select the columns you want to pivot and R-click >> Pivot to pivot the data.

5. Split the data: One column to multiple column. Click on the dropdown of the column and select split. Customize your splitting it as you need the data.

6. Metadata Grid: Useful when we have a lot of columns, so that we don't have to scroll a lot during working. We can work on the columns as a whole on the metadata grid.

7. Understand table down, table across etc. options while creating the table calculations.

8. In the geographical map, if we have something unknown, we can click on it and fix the names manually to be recognized by tableau.

What's new in Tableau 10

1. Custom territory: create by using group. select the state or cities you want to make in a group and right click and click on the attach button to create group.

2. You can drag the custom geographical territory to the geographical hierarchy.

3. Same city can be in different states, so if you directly drag city to worksheet tableau might not recognize some of them, Rather drag the states, and drill down to cities. Voila!!

4. R-click on the variable and click show highlighter to highlight some specific variable or to highlight them.

5. Clustering: Go to analytics tab and drag the cluster on the worksheet, it will apply k-means clustering on the visuals.

6. Cross-database joins: choose data blending for if we have different database formats like csv, xlsx or so.

7. Domain knowledge is important.

8. Modeling with clusters: You can add new variables to the clustering, R-click on cluster variable >> edit and add the new variable for using that variable to the new clustering.

9. Go analytics tab add drag trendline to the visuals to get the data trend.

10. Save your cluster to carry out the analysis based on that cluster. For that just drag the cluster from worksheet to dimensions.

KEEP CALM AND USE DATA WISELY