

Fun with Tableau_Basic

Rao Katie Xu

Environment: Win10 64 bit (Parallel Desktop on Mac OSX Yosemite 10.10.5)

Tableau 10.0: <http://www.tableau.com/products/desktop/download>

Project 1:

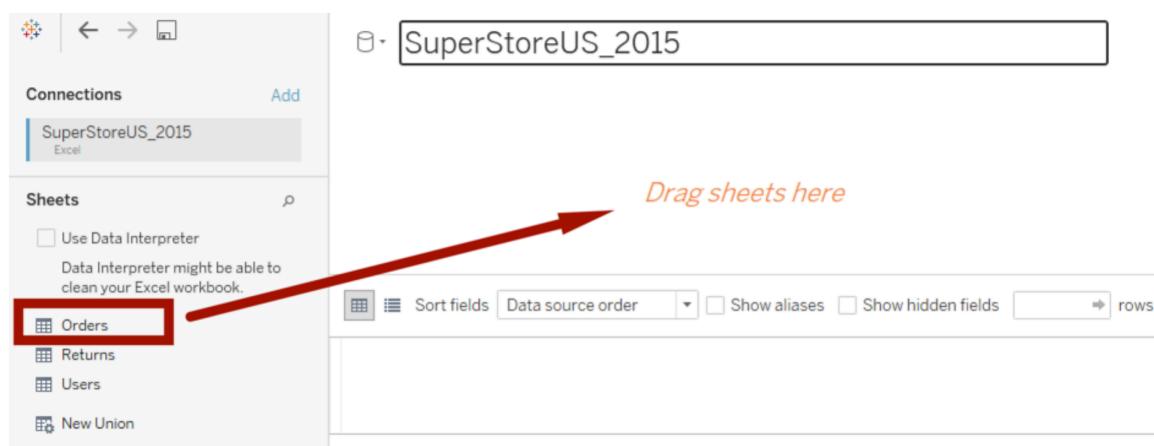
Datasets: <https://www.superdatascience.com/tableau/>

Download: [SuperStoreUS_2015.xlsx](#)

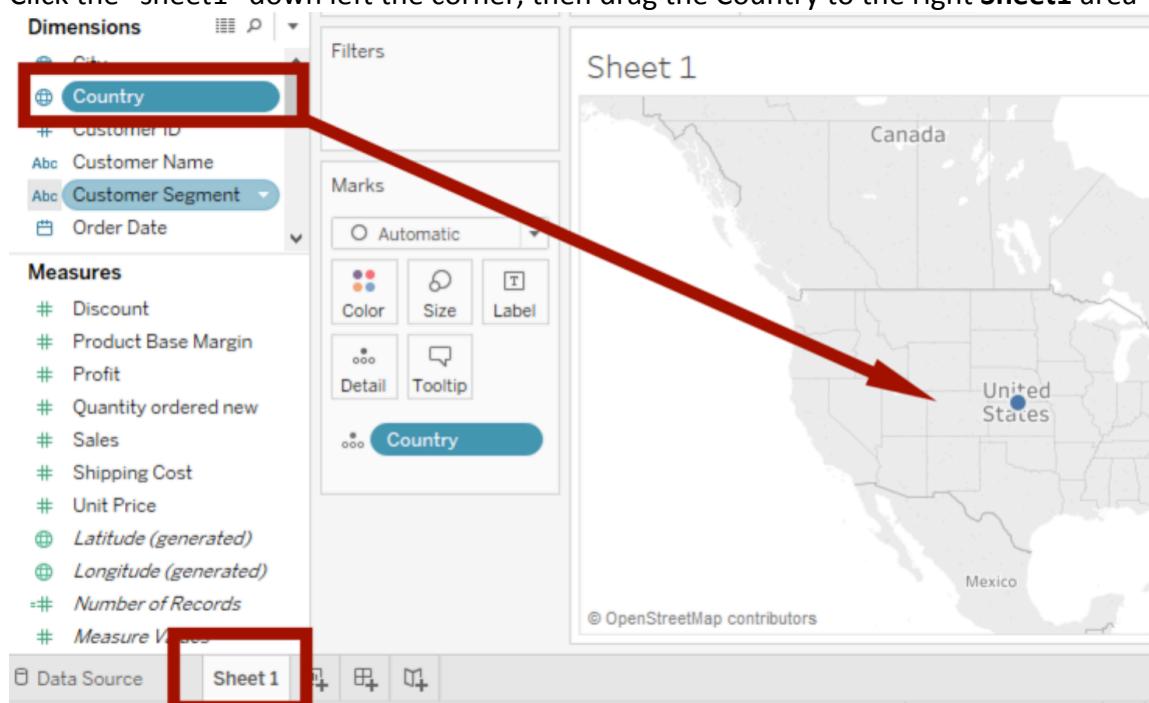
Aim: To find less profit and most profit state and create the state performance

Double Click the [SuperStoreUS_2015.xlsx](#)

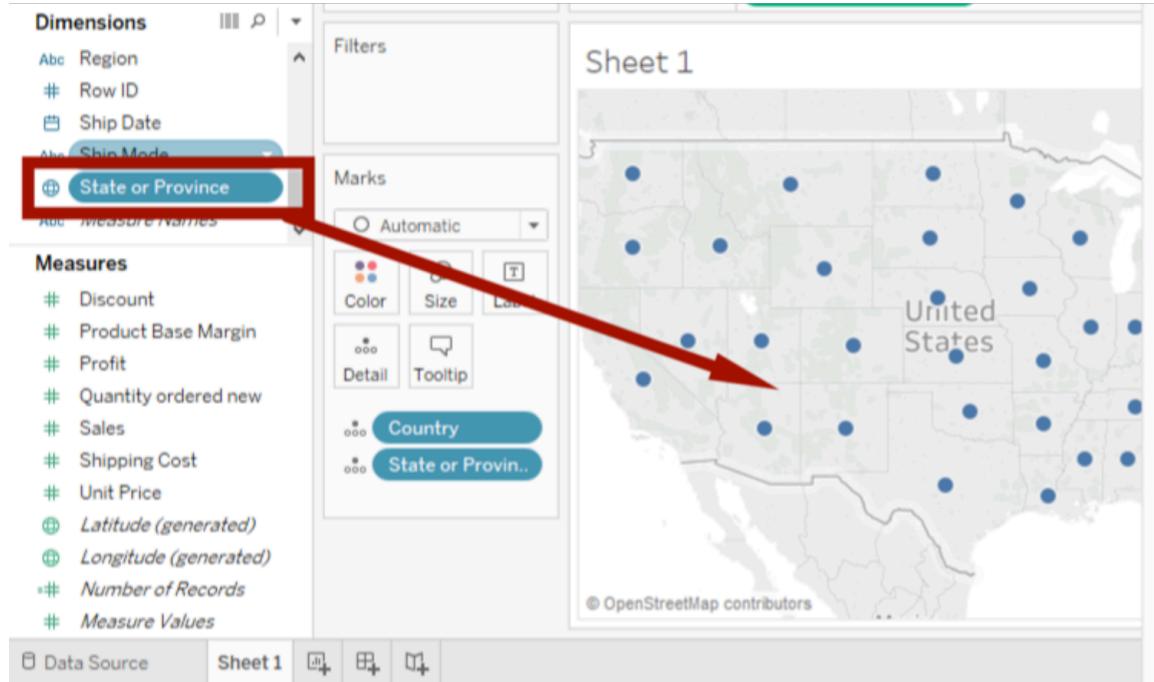
Drag the orders to the “Drag sheets here” area, the detailed data will show up.



Click the “sheet1” down left the corner, then drag the Country to the right **Sheet1** area



The drag the State or Province to the map, then the map will show blue dots.



Drag Profit to the **Automatic->Color** and **Automatic ->Label**, we can see the US map covered with color and the profit number. Now we can tell orange means less profit and NC has the least profit with -\$19,428, and CA has the most profit with \$37,422.



Project 2:

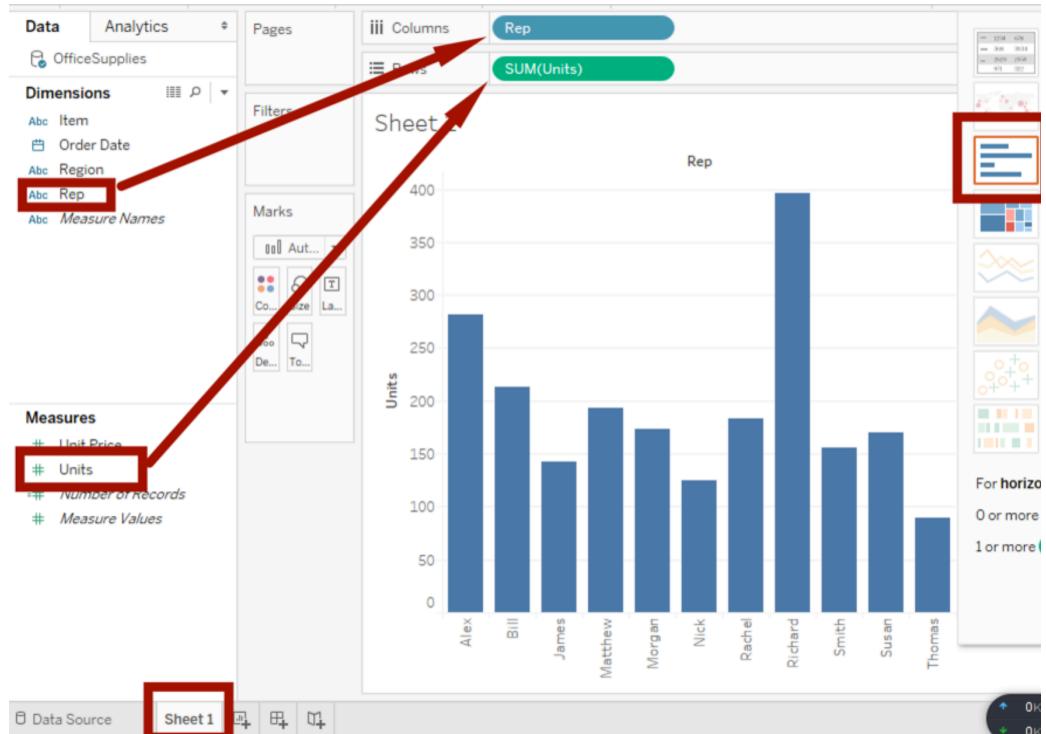
It's EOFY (aka End of Financial Year) and that means time for annual bonuses!

The store operates in three regions and only the top-performing employee in each region qualifies for a bonus. Find out which three employees are eligible to get bonuses for this year. Employees are measured on total sales (\$\$).

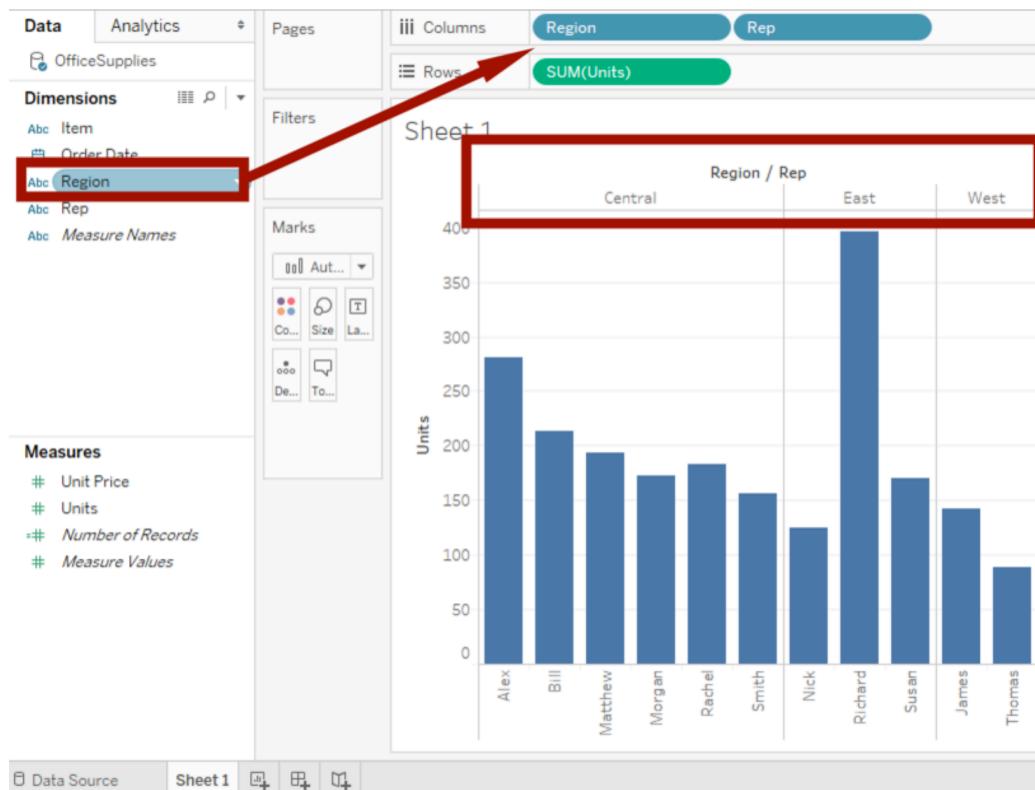
Datasets: <https://www.superdatascience.com/tableau/>

Download: [OfficeSupplies.csv](#)

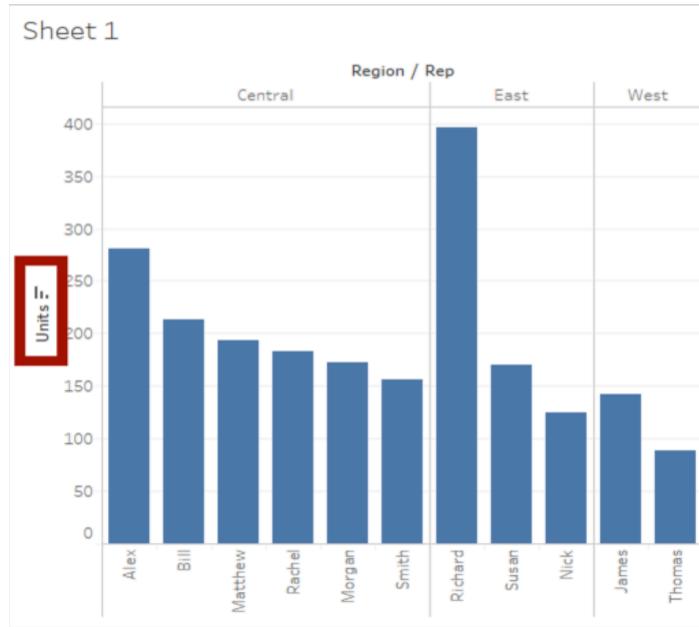
csv file is actually text file, so choose **open->text File**, and open OfficeSupplies.csv file on the **sheet1** label, drag **Rep** to **Columns**, drag **Units** to **Rows**, now we can see the bar graph, so we know Richard has the best sales.



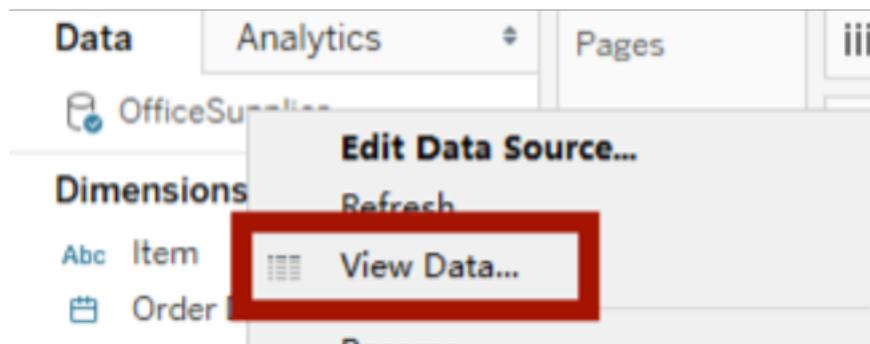
Now we drag **Region** and drop before the **Rep** on Columns. Now on the top of bar graph, it shows the region.



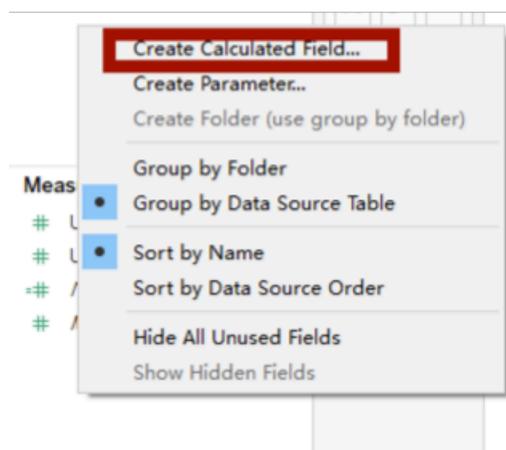
If we click the Units on the y-axis label, you will find the bar has been sorted.



However, we need the total sales (\$) = Unit Price * Units which is not in the table, then how to calculate in the tableau? First, we try to open the data file by right click the **OfficeSupplies**->**View Data** to open the data file. Yes, there are no total sales. Then we need to calculate.



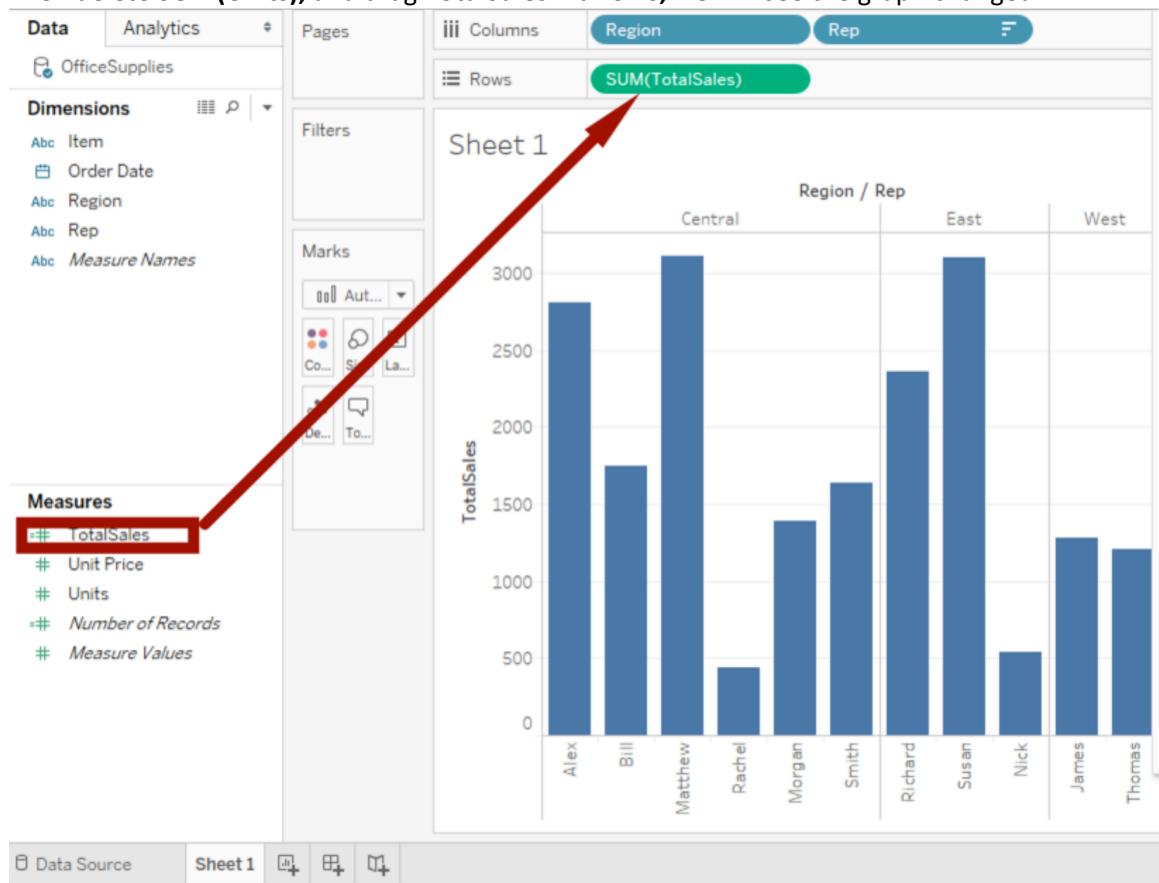
Right click the **Measure** panel, choose **Create Calculated Field**, and put **TotalSales** in the Text Box, and put **Units * Unit Price**, click **OK**.



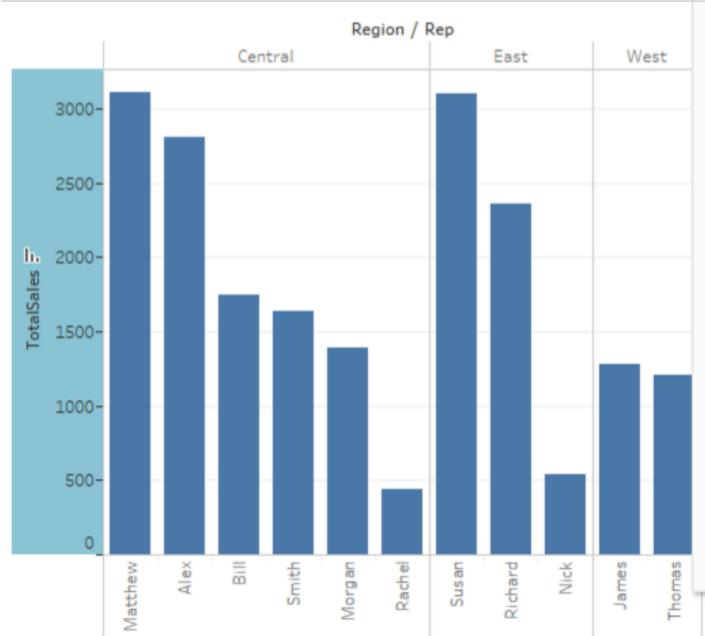
Then we will find in **Measures** panel, # **TotalSales** shows up.

The screenshot shows the Tableau interface with the 'Measures' panel open. In the top-left corner, there is a search bar containing 'TotalSales'. Below it, a calculation 'Units * [Unit Price]' is defined. A note at the bottom says 'The calculation is valid.' There are 'Apply' and 'OK' buttons at the bottom right. To the right of the search bar, the 'Measures' panel lists several items: TotalSales (highlighted with a red box), Unit Price, Units, Number of Records, and Measure Values.

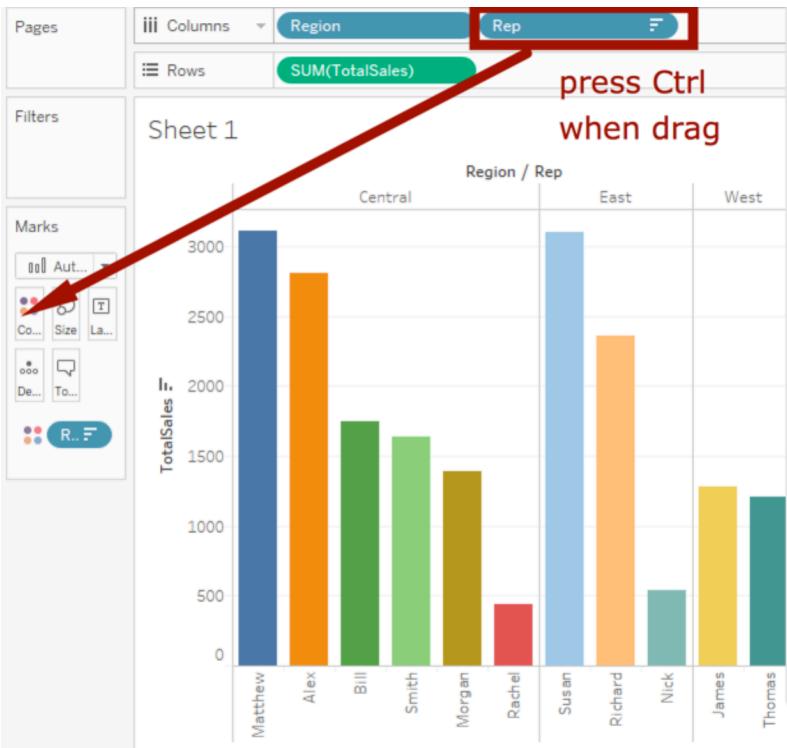
Then delete **SUM(Units)**, and drag **TotalSales** int **Rows**, we will see the graph changed.



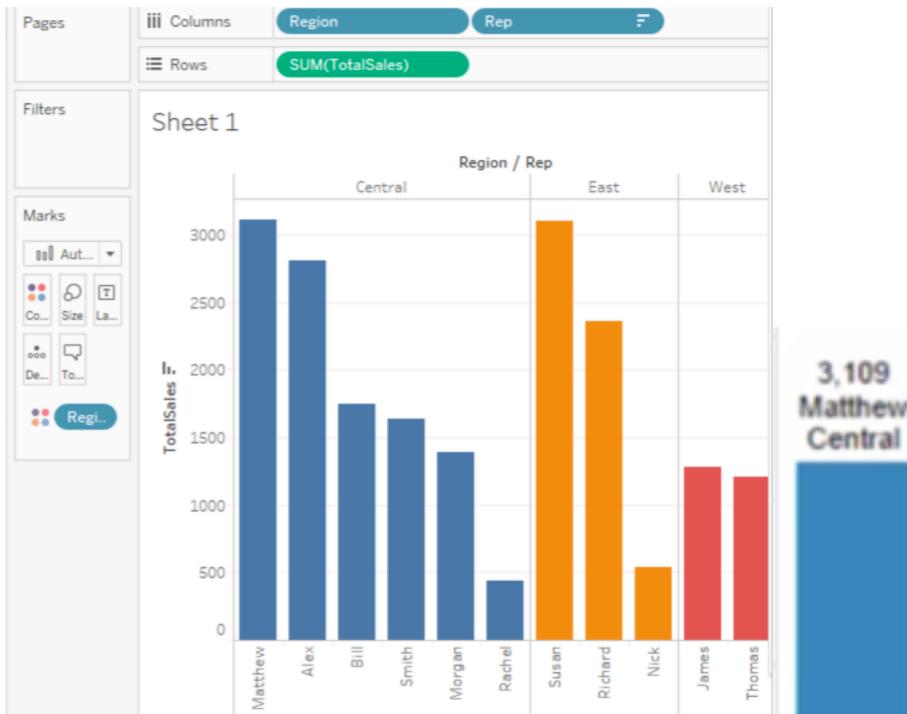
Sort **TotalSales** by click the y-axis label.



From the graph, we can tell in Central area, Mathew has the best sales, East area is Susan, West is James. Now we try to make the graph more colorful. Press **Ctrl** when drag **Rep** to left Color panel. Now it is prettier. (If not press **Ctrl**, the bar will become another bar graph)



Now try to make region colorful. How to see the sales number? Now need to add labels and details. Click **Label**, the sales number will show on the top, also, we can drag **Rep** and **Region** from **Dimensions Panel** to the **Label**, we can see TotalSales, Rep name and region name are all on the top. Here the bar is too small so it won't show all the data detail.

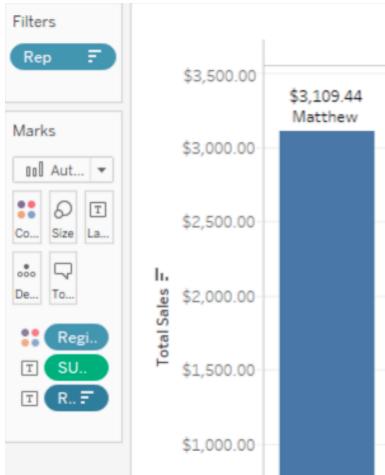


But we want to add \$ before the number, right click **Marks** panel->**Region**, choose **Format**, a new panel will show on the left with all formatting. Click the **Rows->SUM(TotalSales)**, then we can change **Numbers** to **Currency**, then \$ shows up.

The screenshot shows the "Format SUM(TotalSales)" panel open on the left, with the "Numbers" dropdown set to "\$123,456....". A context menu is open over the bar for "Matthew" in the "Central" region, with the "Format..." option highlighted.

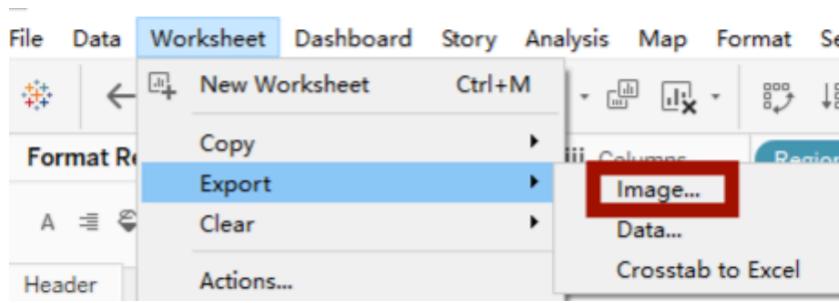
Region / Rep	Rep	TotalSales
Central	Matthew	\$3,109
Central	Alex	\$2,850
Central	Bill	\$1,750
Central	Smith	\$1,641
Central	Morgan	\$1,388
Central	Rachel	\$438
East	Susan	\$3,102
East	Richard	\$2,363
East	Nick	\$537
West	James	\$1,284
West	Thomas	\$1,203

If choose **Currency(Custom)**, we can adjust to UK or digits. Also, the formatting can help change word direction.

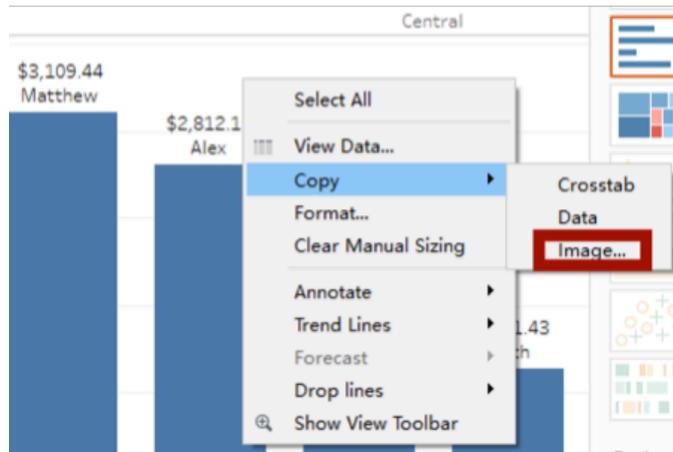


Now it is time to export the result.

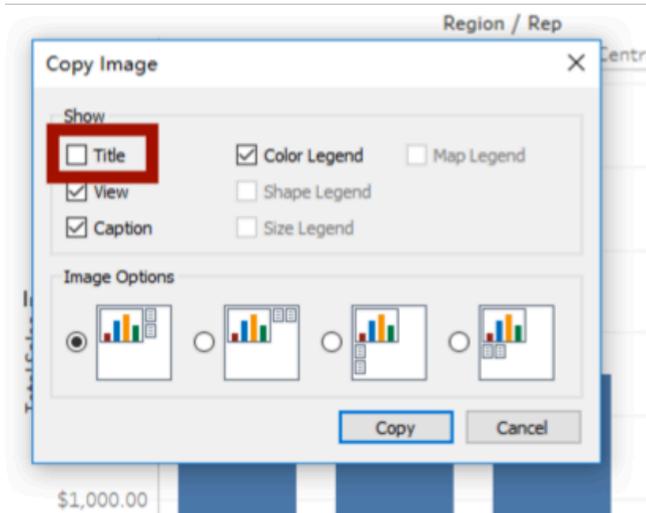
Plan A: Worksheet->Export->Image



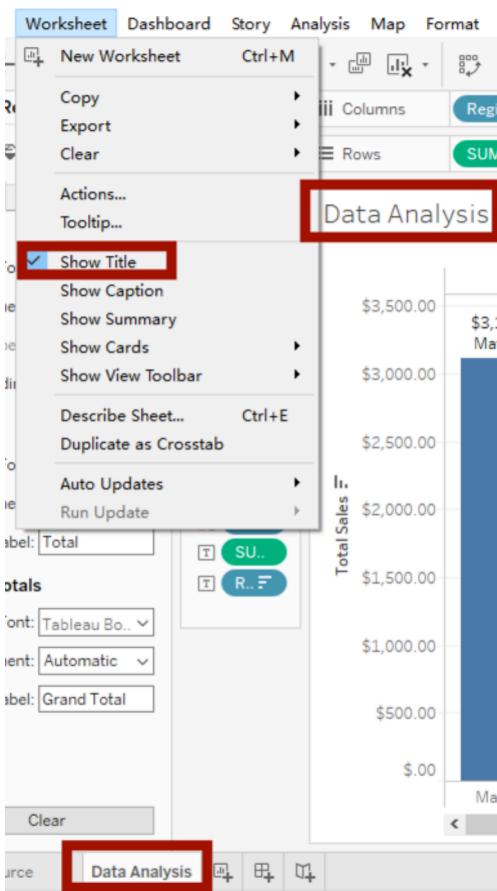
Plan B: right click the graph, Copy -> Image



Once open a Word file, we can paste the image. However, we found sheet1 on the top, so we need to **uncheck the Title** when copying Image. Next, how to add the title?



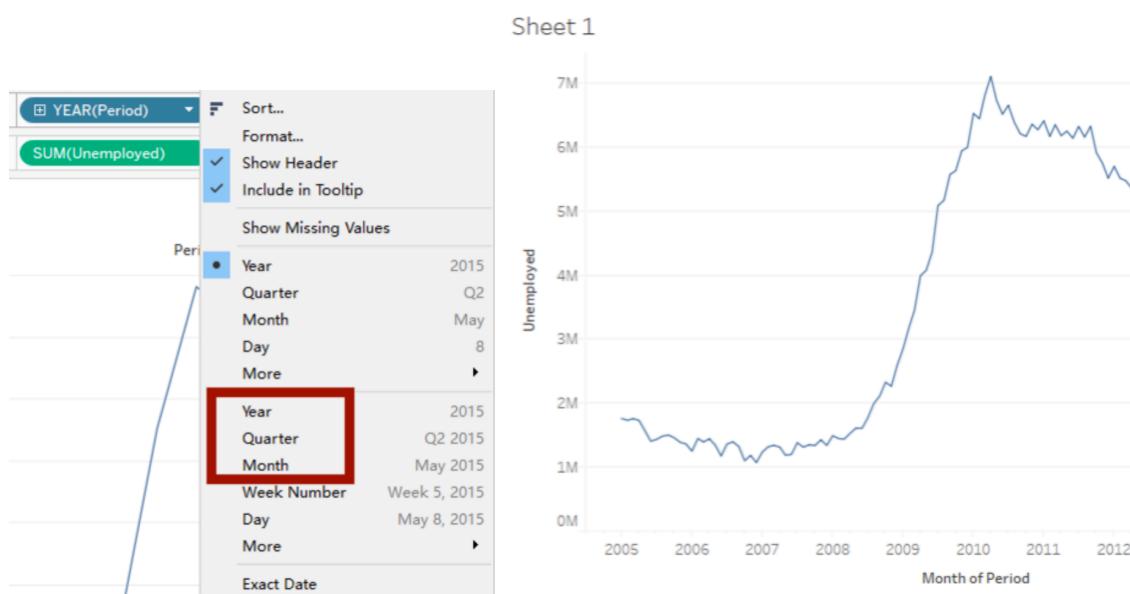
Go change **sheet1** to **Data Analysis, Worksheet** -> **Show Title**, Check **Title** when copy Image, the image will have a title.



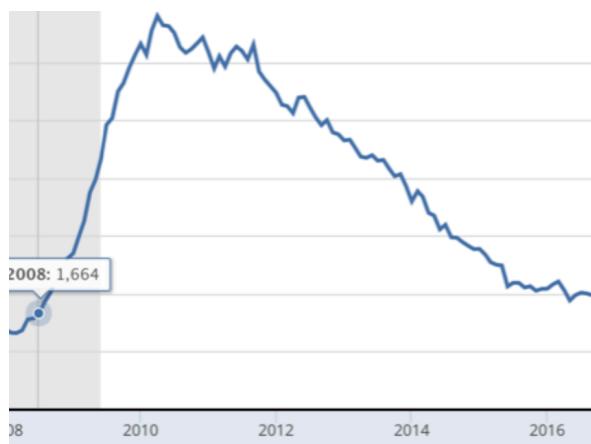
<http://public.tableau.com/profile/kirill.eremenko#!/>
for more tutorials.

Project 3:
Download: <https://www.superdatascience.com/tableau/>
[Long Term Unemployment Statistics.xlsx](#)

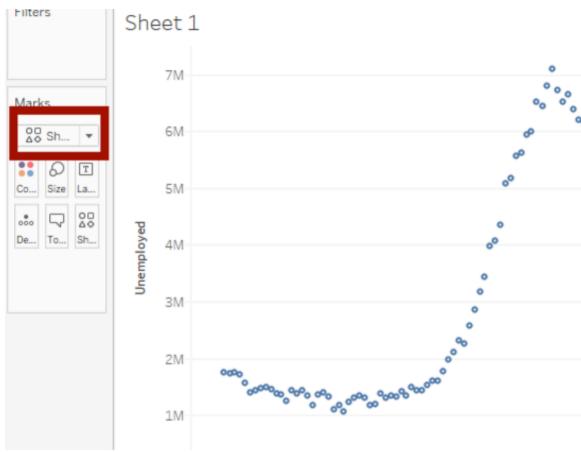
Drag **Measures->Unemployed** to Rows, drag **Dimensions-> Period** to Columns. Now we notice they are two-repeated Year/Quarter/Month/Day. What's the difference? If you choose the first Month, then x-axis will show only Jan-Dec 12 spots, however, we want to show all 12 months in each year, then we need to choose the second Month. Now we have all the months.



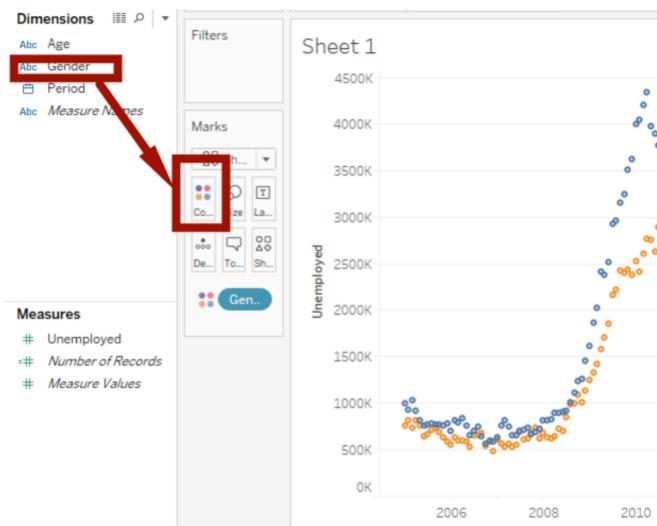
go search unemployed 27 weeks fred in st louis, <https://fred.stlouisfed.org/series/UEMP27OV>, it is very similar which means the data is real unemployed data.



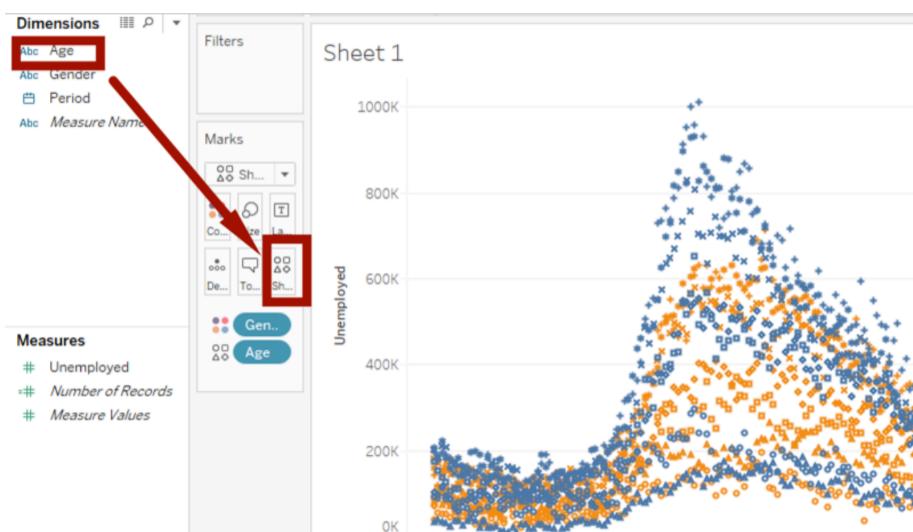
When choose the **Marks->Shape**, then the curve will change to dot, then adjust the size will change the circle's size.



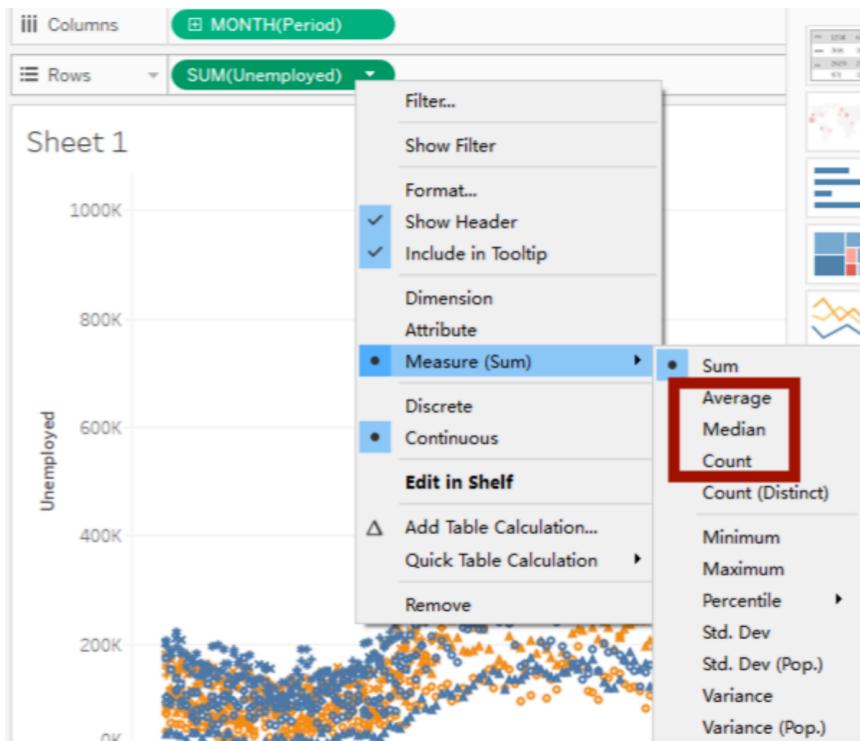
Then drag the **Gender** to the **Color**, we can see the gender data show up on the graph.



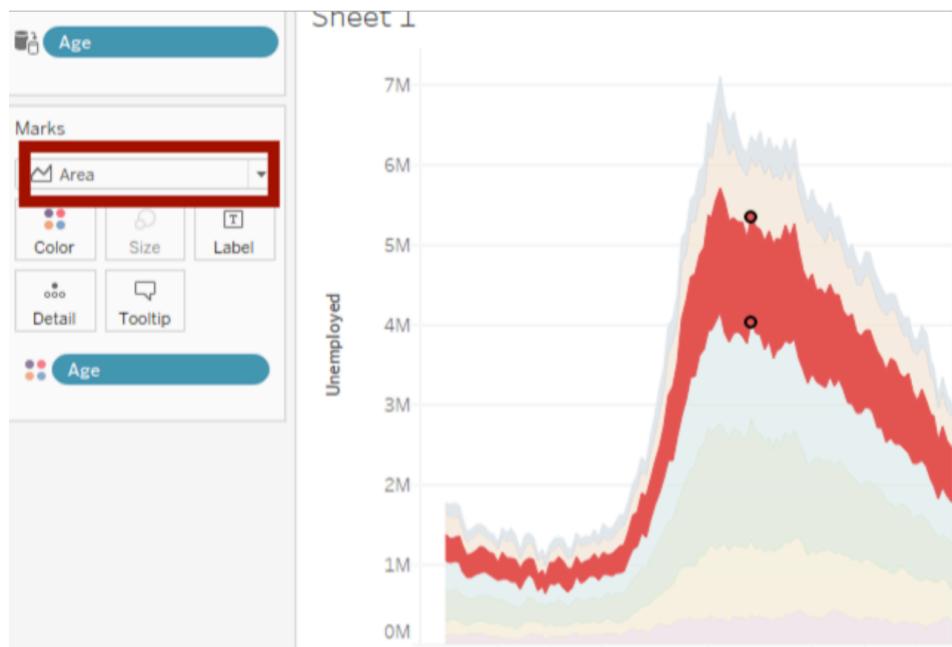
Then drag **Age** to the **shape**, we can see all different ages on the graph.



Now **SUM(Unemployed)** can do more than just **Sum**, such as **Average**, **Median**, **Count**



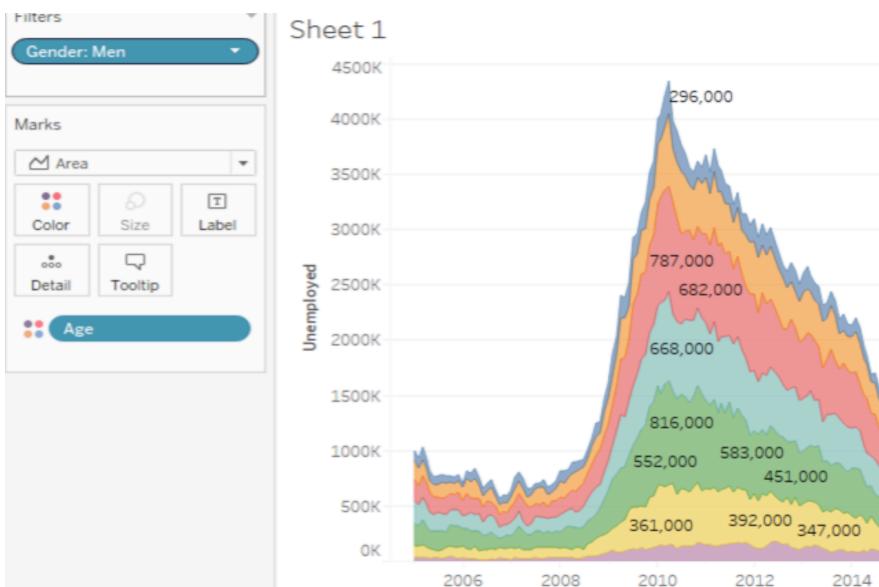
We can drag **Gender** to **Details** for aggregations. Change graph to **Area**, when clicking the area, this area will be highlighted.



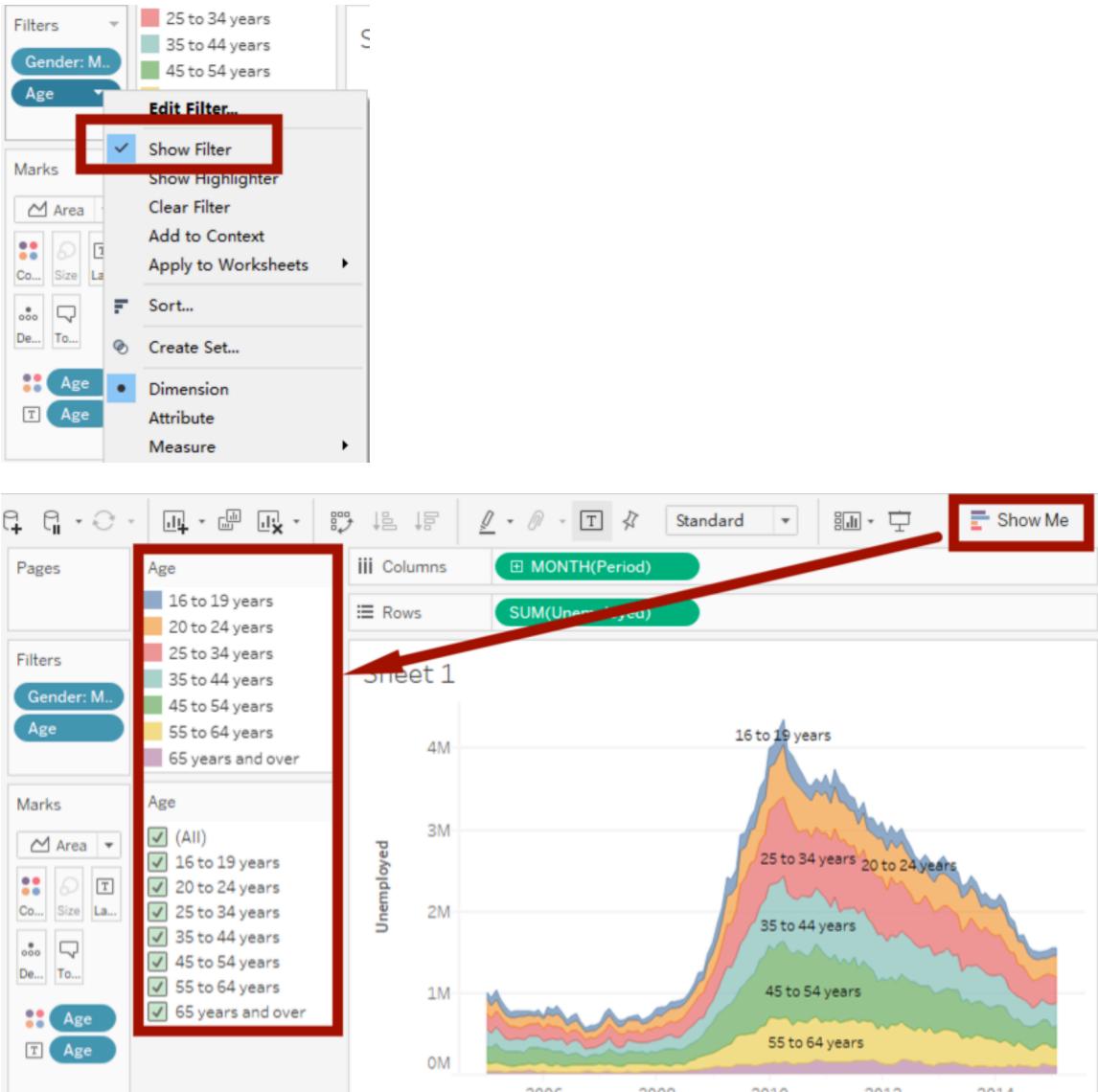
What if I only want to see the men's data? Just drag the **Gender** to the **Filters** panel, a window will pop up, then choose **Men**.

The screenshot shows the Tableau Data pane on the left and a 'Filter [Gender]' dialog on the right. In the Data pane, under 'Dimensions', 'Abc Age', 'Abc Gender' (highlighted with a red box), and 'Period' are listed. Under 'Measures', '# Unemployed', '# Number of Records', and '# Measure Values' are listed. The 'Filters' section is highlighted with a red box. The 'Filter [Gender]' dialog has 'Select from list' selected. Under 'Enter search text', 'Men' is checked (highlighted with a red box) and 'Women' is unchecked. Buttons at the bottom include 'All', 'None', and 'Exclude'. A summary at the bottom states 'Field: [Gender]' and 'Selection: Selected 1 of 2 values'.

Now the graph only shows the men's data.



However, if there are so many data like age ranges. Is any way to be convenient? Yes. Right click the **Age**, choose **Show Filter**, then you can see a panel with different ages with colors. It's originally hiding behind the **Show me** on the right side of the window. Just click the **Show me**, you can see the filter panels. Just check or uncheck the age ranges to find out the changes of the graph.

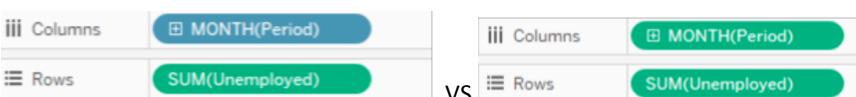


Which of these is a valid reason to use a Data Extract over a Live Connection to the Dataset in Tableau?

ANS: Your Data is constantly being updated and you want to work with a static file when building your visual (you will later return to the live connection when the visual is ready)

What is the difference between the Blue Month() and the Green Month() in Tableau?

Blue is a dimension and Green is a measure; Blue ignores higher periods such as year and treats month as a category - just like Gender; Green creates a proper timeline



The Blue Month() is actually just treats month as a category.

Green Month() represents all months in each year(a timeline).

Measure vs Dimension: <http://www.evolytics.com/blog/tableau-fundamentals-dimension-vs-measure/>

Generally, the measure is the number; the dimension is what you “slice and dice” the number by.

