## Data Visualization Lab The Tableau Workspace

Nov. 28, 2022

Is there anything you want to ask about Lecture 2?

#### Before we start

Mon 14/11 - 9:00-11:15 Mon 19/12 - 9:00-11:15

Mon 21/11 - 9:00-11:15 Tue 20/12 - 16:45-18:15

Mon 28/11 - 9:00-11:15 Mon 09/01 - 9:00-11:15

Mon 5/12 9:00 11:15 Tue 10/01 - 16:45-18:15

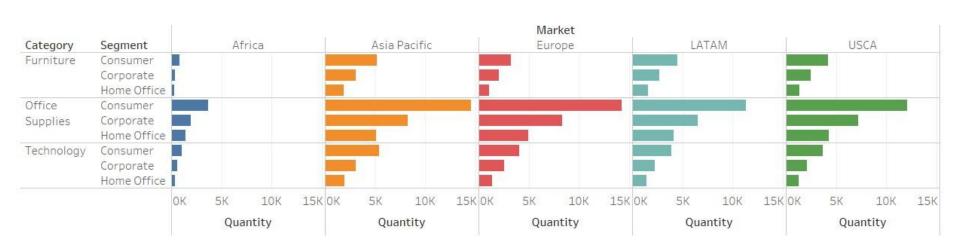
Mon 12/12 - 9:00-11:15 Wed 11/01 - 11:45-14:15

Fri 16/12 - 9:00-10:30

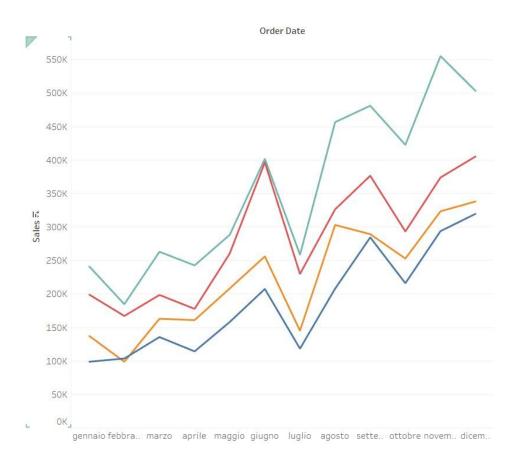
 $\rightarrow$  Next lecture (5/12) is replaced by the Data Management Lab (prof. Avogadro)

# The Tableau Workspace Tableau Sheets

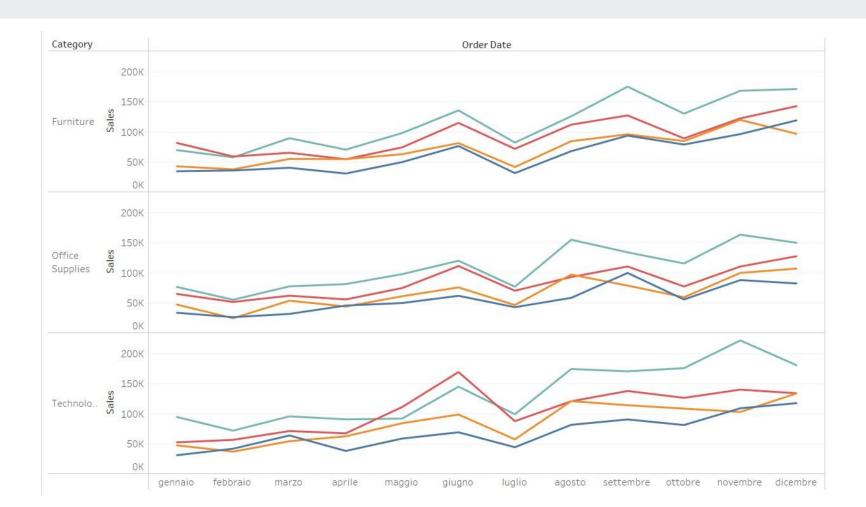
- 1. A chart representing sales divided by category of product sold (**Category**), type of customer (**Segment**) and market (**Market**), in terms of absolute number of items sold (**Quantity**);
- 2. A chart representing sales (Sales) trends (in \$) over the months of the year (MONTH(Order Date)) comparing, on different lines, the years (YEARS(Order Date)) in the dataset;
- 3. A chart that details the preceding by dividing sales by product category;
- 4. A crosstab./table



- 1. A chart representing sales divided by category of product sold (Category), type of customer (Segment) and market (Market), in terms of absolute number of items sold (Quantity);
- 2. A chart representing sales (Sales) trends (in \$) over the months of the year (MONTH(Order Date)) comparing, on different lines, the years (YEARS(Order Date)) in the dataset;
- 3. A chart that details the preceding by dividing sales by product category.
- 4. A crosstab./table



- A chart representing sales divided by category of product sold (Category), type of customer (Segment) and market (Market), in terms of absolute number of items sold (Quantity);
- 2. A chart representing sales (Sales) trends (in \$) over the months of the year (MONTH(Order Date)) comparing, on different lines, the years (YEARS(Order Date)) in the dataset;
- 3. A chart that details the preceding by dividing sales by product category.
- 4. A crosstab./table



- 1. A chart representing sales divided by category of product sold (Category), type of customer (Segment) and market (Market), in terms of absolute number of items sold (Quantity);
- 2. A chart representing sales (Sales) trends (in \$) over the months of the year (MONTH(Order Date)) comparing, on different lines, the years (YEARS(Order Date)) in the dataset;
- A chart that details the preceding by dividing sales by product category.
- 4. A crosstab./table

Now imagine that we want to analyze the **profit trend** on the Crosstab we just created.

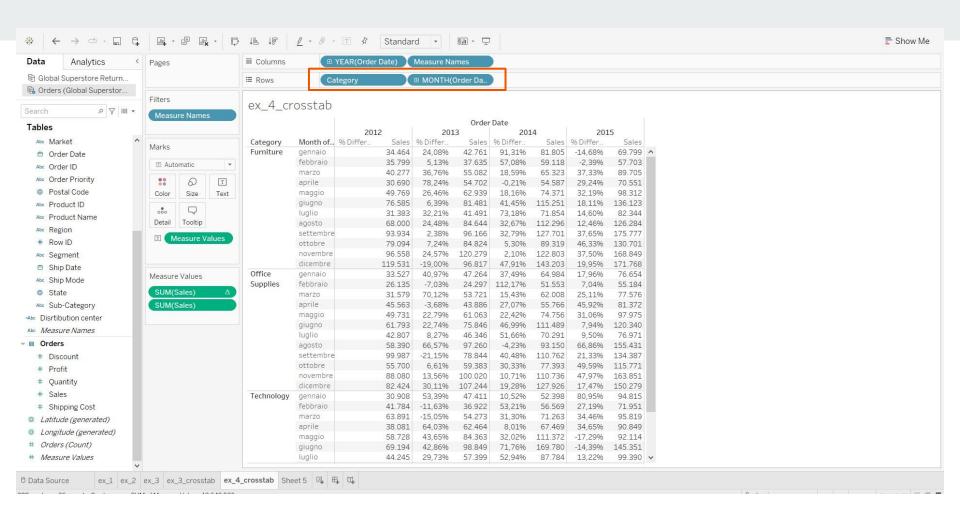
What is the easiest way to do this?

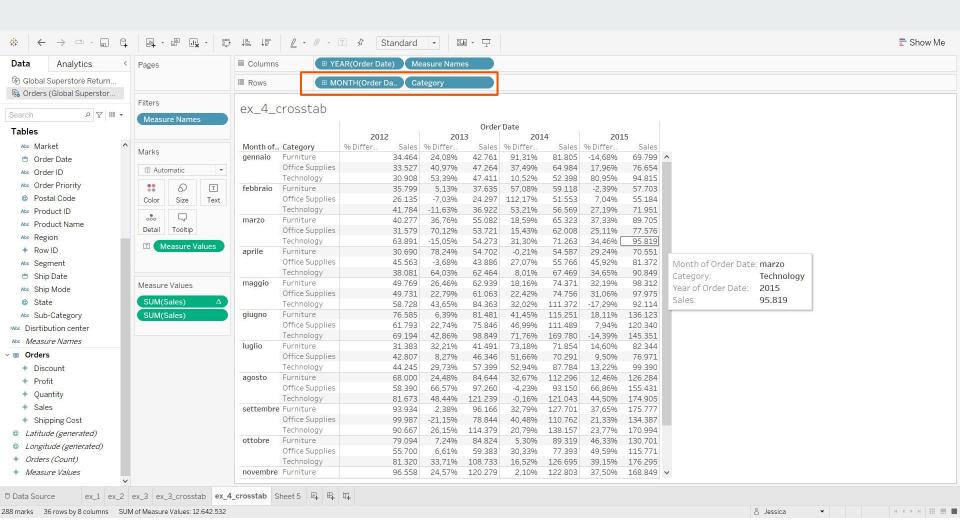
Now imagine that we want to analyze the **profit trend** on the Crosstab we just created.

#### What is the easiest way to do this?

→ Swap the MONTH and YEAR(Order Date) pills so that the YEAR one is on the Columns and the MONTH one on the Rows;

Note that the order in which you put the Category and the MONTH pills on the Rows shelf matters!





Now imagine that we want to analyze the **profit trend** on the Crosstab we just created.

→ Swap the MONTH and YEAR(Order Date) pills so that the YEAR one is on the Columns and the MONTH one on the Rows;

Which of the two versions seems to make more sense to you?

Now imagine that we want to analyze the **profit trend** on the Crosstab we just created.

- Swap the MONTH and YEAR(Order Date) pills so that the YEAR one is on the Columns and the **MONTH** one on the Rows:
- Put the MONTH(Order Date) pill before the Category one on the Rows shelf.

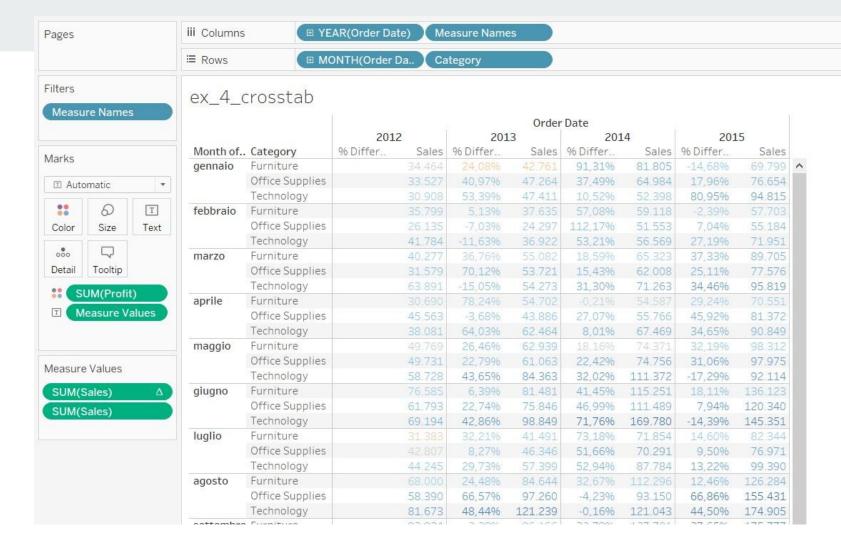
Remember we wanted to analyze a trend in profits.

How can we do that? (spoiler: **(**?)



Now imagine that we want to analyze the **profit trend** on the Crosstab we just created.

- → Swap the MONTH and YEAR(Order Date) pills so that the YEAR one is on the Columns and the MONTH one on the Rows;
- → Put the MONTH(Order Date) pill before the Category one on the Rows shelf;
- → Drag and drop the **Profit** field onto the Color shelf in the Marks Pane

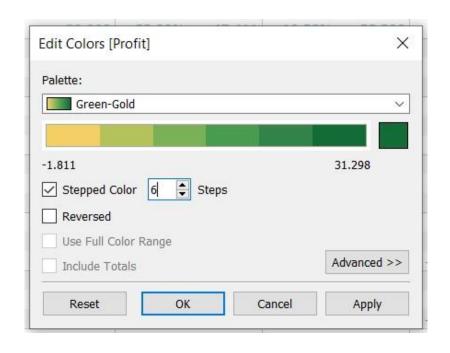


Adding **Profit** to color gives us a clearer understanding of overall trends.

These colors are a bit pale, though, so let's edit how we display this. We'll click on *Color>Edit Colors*.

#### Let's use:

- → Green-gold palette
- → Stepped color (6 steps)



		Order Date							
		2012		2013		2014		2015	
Month of	Category	% Differ	Sales	% Differ	Sales	% Differ	Sales	% Differ	Sales
gennaio	Furniture		4.464	24,08%	42.761	91,31%	81.805	-14,68%	69.799
	Office Supplies	3	3.527	40,97%	47.264	37,49%	64.984	17,96%	76.654
	Technology	3	0.908	53,39%	47.411	10,52%	52.398	80,95%	94.815
febbraio	Furniture	3	5.799	5,13%	37.635	57,08%	59.118	-2,39%	57.703
	Office Supplies	2	6.135	-7,03%	24.297	112,17%	51.553	7,04%	55.184
	Technology	4	1.784	-11,63%	36.922	53,21%	56.569	27,19%	71.951
marzo	Furniture	4	0.277	36,75%	55.082	18,59%	65.323	37,33%	89.705
	Office Supplies	3	1.579	70,12%	53.721	15,43%	62.008	25,11%	77.576
	Technology		3.891	-15,05%	54.273	31,30%	71.263	34,46%	95.819
aprile	Furniture			78,24%	54.702	-0,2196	54.587	29,24%	70.551
	Office Supplies	4	5.563	-3,68%	43.886	27,07%	55.766	45,92%	81.372
	Technology	3	8.081	64,03%	62.464	8,01%	67.469	34,65%	90.849
maggio	Furniture	4	9.769	26,46%	62.939	18,16%	74.371	32,19%	98.312
	Office Supplies	4	9.731	22,79%	61.063	22,42%	74.756	31,06%	97.975
	Technology	5	8.728	43,65%	84.363	32,02%	111.372	-17,29%	92.114

The colors in this table are much clearer, but we can achieve something more impactful.

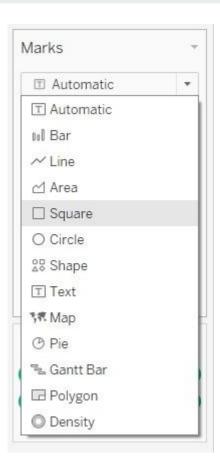
#### How can we do that?

An easy way to do so would be to fill the cells in our table, instead of having the value displayed.

The colors in this table are much clearer, but we can achieve something more impactful.

#### How can we do that?

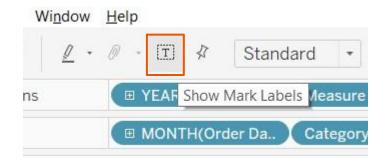
→ Change the mark type to square by clicking on the Automatic dropdown menu in the Marks Pane and select Square;

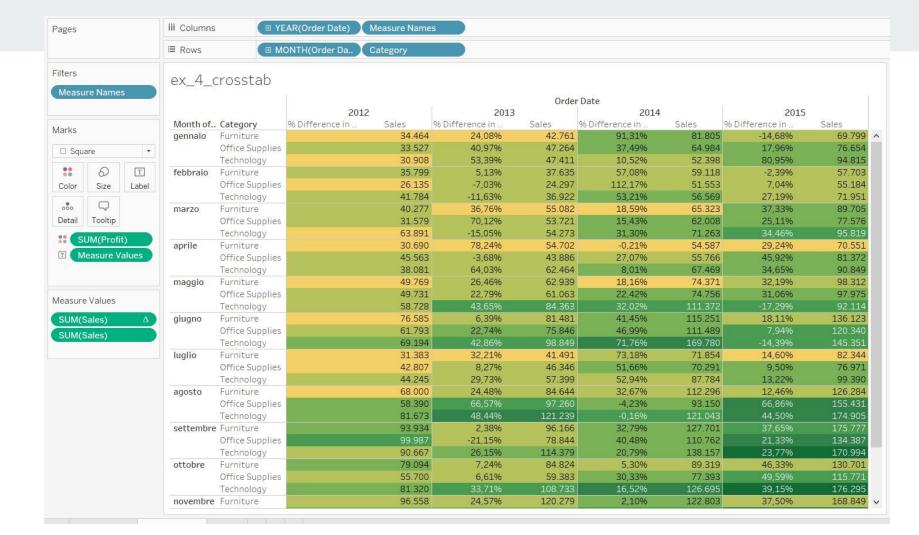


The colors in this table are much clearer, but we can achieve something more impactful.

#### How can we do that?

- → Change the mark type to square by clicking on the **Automatic** dropdown menu in the Marks Pane and select **Square**;
- Turn on mark labels by using the **Show Mark**Labels button on the toolbar.





## **Dynamic highlighting**

What we have here is a comprehensive <u>outlook of profits</u> throughout the years and months in our dataset.

Suppose we want to compare the <u>trends within the same Category</u> without modifying the current view.

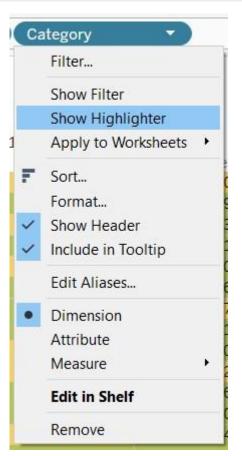
How do we do that?

## Dynamic highlighting

Go over to the Column shelf, open the dropdown menu on the **Category** pill and click on **Show Highlighter**.

A small window (**Highlight Category**) will pop up on the rightmost side of the view displaying a filter.

Now hovering over the categories in the **Highlight Category** box can be a quick way of assessing profits and comparing them category-by-category.







#### A point on data

#### Why is dynamic highlighting useful?

Remember that we noticed some sort of <u>sales seasonality</u> in the previous charts, so now we are interested in understanding whether different categories lead to different profits in different times of the year.

If you only consider one year, let's say 2015, and hover over the categories in the **Highlight Category** pane, you can clearly compare sales and profits in different months and categories.

#### Remember that:

- → Color denotes **Profits**
- → <u>Numbers</u> denote **Sales**

Month of	Category
febbraio	Technology
marzo	Furniture
	Office Supplies
	Technology
aprile	Furniture
	Office Supplies
	Technology
maggio	Furniture
	Office Supplies
	Technology
giugno	Furniture
	Office Supplies
	Technology
luglio	Furniture
	Office Supplies
	Technology
agosto	Furniture
	Office Supplies
	Technology
settembre	Furniture
	Office Supplies
	Technology
ottobre	Furniture
	Office Supplies
	Technology
novembre	Furniture
	Office Supplies
	Technology
dicembre	Furniture
	Office Supplies
	Technology

	2015	
	% Difference in	Sales
3	37,33%	89.705
7	29,24%	70.551
		90.849
L	32,19%	98.312
2		92.114
	18,11%	136.123
1		
ļ	14,60%	82.344
l		99.390
5	12,46%	126.284
2	44,50%	174 905
	37,65%	175.777
1		170.994
)	46,33%	130.701
		176.295
3	37,50%	168.849
		222.579
3	19,95%	171.768
		150.279
	34,82%	181.097

2015	
% Difference in	Sales
27,1996	71.951
25,11%	77.576
29,24%	70,551
45,92%	81.372
32,1996	98.312
31,06%	97.975
18,1196	136.123
7,94%	120.340
-14,39%	145.351
14,6096	82.344
9,50%	76.971
12,4696	125,284
66,86%	155.431
	174.905
37,65%	175.777
21,33%	134.387
	17/0.394
45,33%	130.701
49,59%	115.771
	1/6.295
37,50%	168.849
47,97%	163.851
	4445/9
70,0010	150.279
17,47%	
34,94%	181.097

2015		
% Difference in	Sales	
27,19%	71.951	-
25,1195	77.576	
34,46%	95.819	
45,9296	81.372	
34,65%	90.849	
31,06%	97.975	
-17,29%	92.114	
7.94%	120 340	
-14,39%	145.351	
14,6096		
9,50%	76.971	
13,22%	99.390	
66,86%	155,431	
44,50%	174.905	
23,77%	170.994	
39,15%	176.295	
58,47%	222.579	
19,95%	171.768	
34,82%	181.097	

Furniture

Office supplies

Technology

## The Show me tool

## Know what you want, don't know how to do it

Now that we have explored all the data about profits and we have noticed that <u>furniture profits are apparently bad during fall</u>, we can start to make hypotheses about the reason why this happened.

For example, we may wonder *is this happening across all stores*? We can therefore analyze this phenomenon from many perspectives, one of which is <u>regionality</u>.

→ How can we do that? (hint: **■**)

### Know what you want, don't know how to do it

Now that we have explored all the data about profits and we have noticed that <u>furniture profits are</u> <u>apparently bad during fall</u>, we can start to make hypotheses about the reason why this happened.

For example, we may wonder *is this happening across all stores*? We can therefore analyze this phenomenon from many perspectives, one of which is <u>regionality</u>.

→ How can we do that? (hint: **■**)

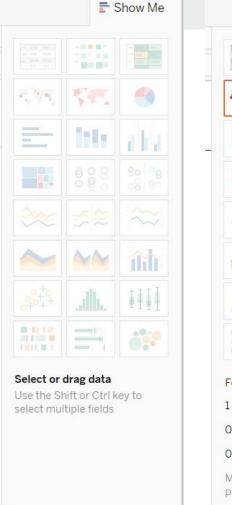
Feel lost? I feel you 😪

#### The Show Me tool

TP provides you with a <u>one-click option</u> to identify a proper chart type for your data. It comes on handy when you know what you want to do but you don't know how to do it.

The **Show Me** on the rightmost part of the screen contains a list of common chart types that can help you start your analysis.

The chart suggestions in Show Me are not necessarily the best ones, still they represent a great starting point





#### The Show Me tool

Which dimensions and measures do we need in order to visualize sales regionality? 🔔

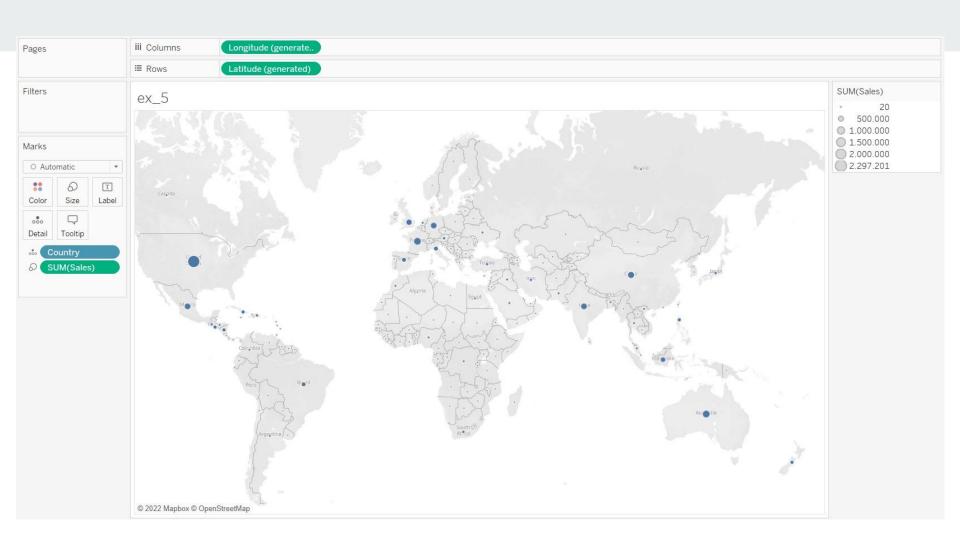
#### The Show Me tool

Which dimensions and measures do we need in order to visualize sales regionality?

- → Sales;
- → Country.

Hold the **Ctrl** button on your keyboard and select the two data fields in the Data Pane.

You will see that the Show Me tab is populated with chart types to choose among. We will pick **symbol maps** to start off.



Symbol maps are fully customizable; we will perform the following operations:

- → Refine the visualization with additional dimensions;
- → Adjust dot size;
- → Adjust transparency;
- → Adjust borders;
- → Change color palette.

Before we start, have a look at the Rows/Columns shelves.

Notice that neither the **Longitude** nor the **Latitude** measures are present in the dataset. TP automatically creates them based on:

- → Measures/dimensions you select;
- → Type of chart you choose.



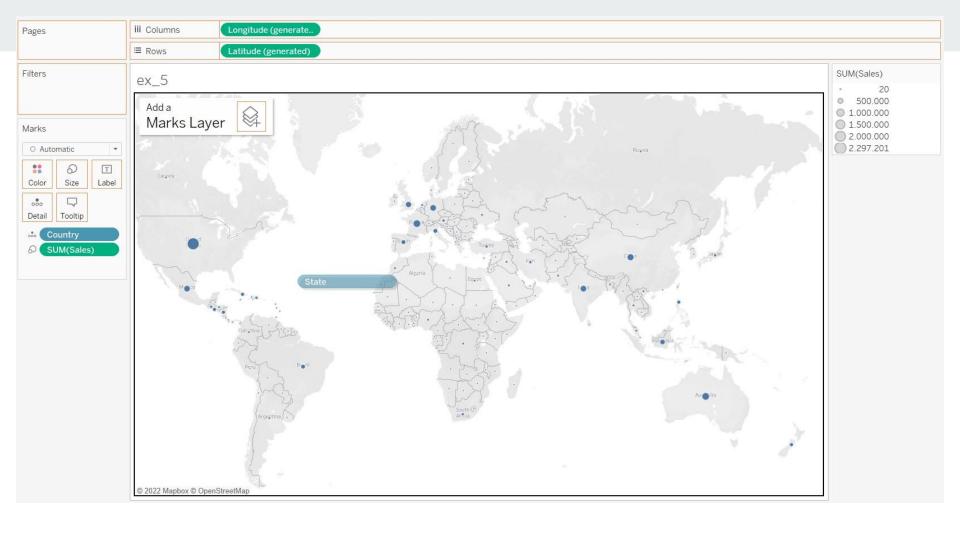
#### Refine the visualization with additional dimensions:

What does the symbol map show? What is the chart representing?

- → It shows **Sales** in <u>each country</u> over a global map;
- → The <u>size of the dots</u> represents the **total amount of sales**, bigger dots simply mean higher sales.

We know that we have much more precise information on the origin of sales than the **Country**. For example, we know the **State** and the **City** in which they were made.

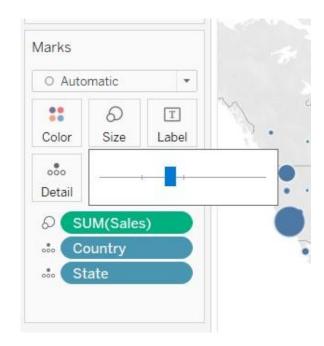
Let's start with the **State** data field; drag and drop in onto the canvas!



#### Adjust dot size:

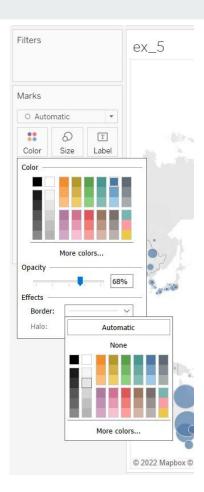
Increasing the size is useful when you want to display quantities that are really small in a dataset with huge variance in data or datasets in which the highest and the lowest values belong to different orders of magnitude.

To increase the dot size, head over to the **Size** shelf in the Marks pane and use the slider to set the proper size for your dots.



### Adjust transparency, add borders:

Move to the **Color** shelf in the Marks Pane and adjust dot opacity (or transparency) and add some border to the dots so that they stand out better in the overall visualization.



### **Change color palette:**

Now, let's bring back the **Profits** measure.

How can we do that?

### **Change color palette:**

Now, let's bring back the **Profits** measure.

#### How can we do that?

Similar to what we did earlier, we can use color. In this case:

- → The <u>color</u> will denote the sum of **Profits** in each State;
- → The <u>size of the dot</u> will represent the total **Sales** in each State.

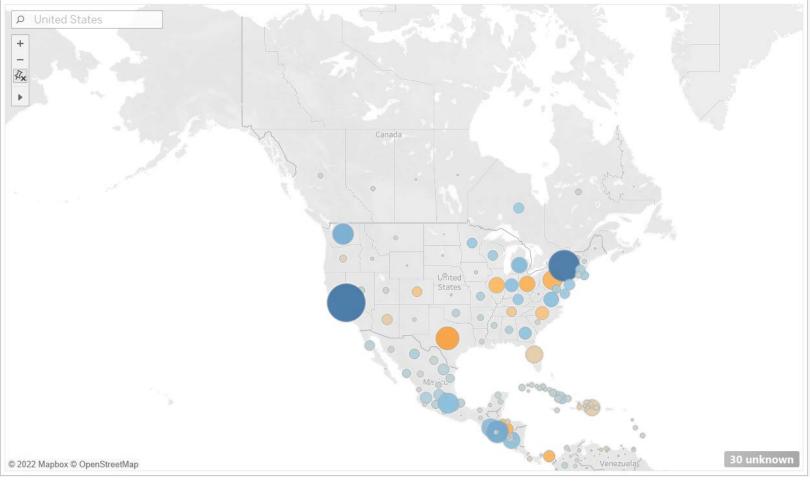
Let's drag and drop the **Profit** field onto the **Color** shelf in the Marks Pane.

Maps on Tableau have some tools that allow you to <u>interact</u> with them during data analysis.

The **Search tool**, for example, allows you to perform geographic search, by simply typing the name of the State or city that you want to look for (whether they are present in the dataset or not). TP will zoom in to highlight the area you searched for.









Now imagine that we are a big company and we have just witnessed a <u>dip in sales in July</u>; we may want to investigate the reasons behind this dip in sales and we have two hypotheses:

- → The fall in sales is due to our <u>actions as a company</u>;
- → The fall in sales is due to <u>seasonality</u>.

We know from previous analyses that seasonality is a factor that impacts on our sales and profits.

How can we use the symbol map to test whether there is a seasonality effect in sales?

We might want to divide the globe, thus the data, by hemispheres, but we do not have this in our dataset.

We can use the **Lasso Selection tool** to select data directly on the map and create a new group of data. It will be a rough division of data, but it can give you useful insights on your data.

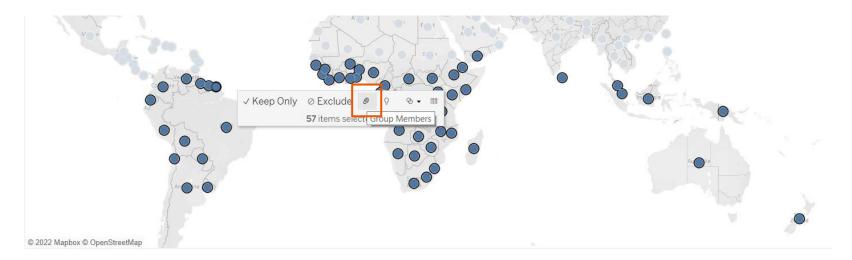
Remove everything from the Marks pane except for the **Country** dimension to make the view clearear.



🥊 If you are familiar with Photoshop, the lasso tool works exactly the same



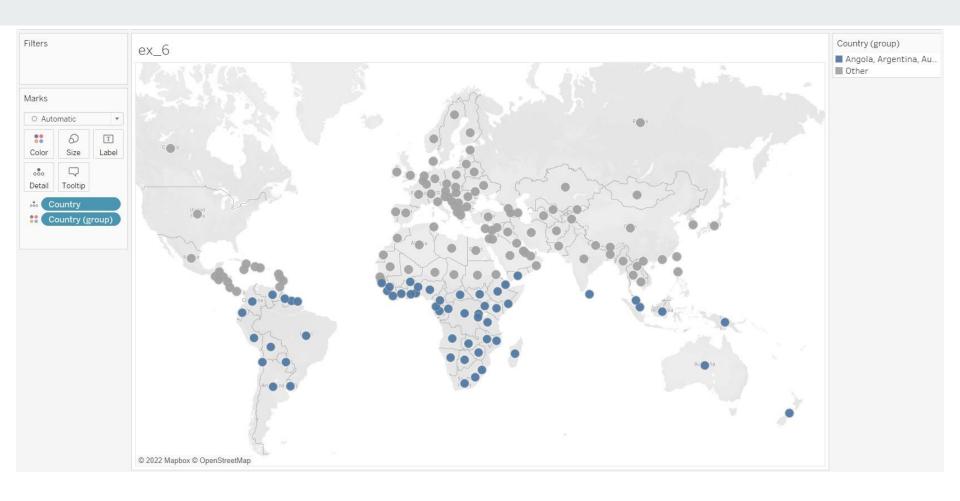
Drag the Lasso tool over the southern region of the map and create a new group of data by clicking on the paper clip  $\oslash$  icon on the tooltip.



Drag the Lasso tool over the southern region of the map and create a new group of data by clicking on the paper clip  $\oslash$  icon on the tooltip.

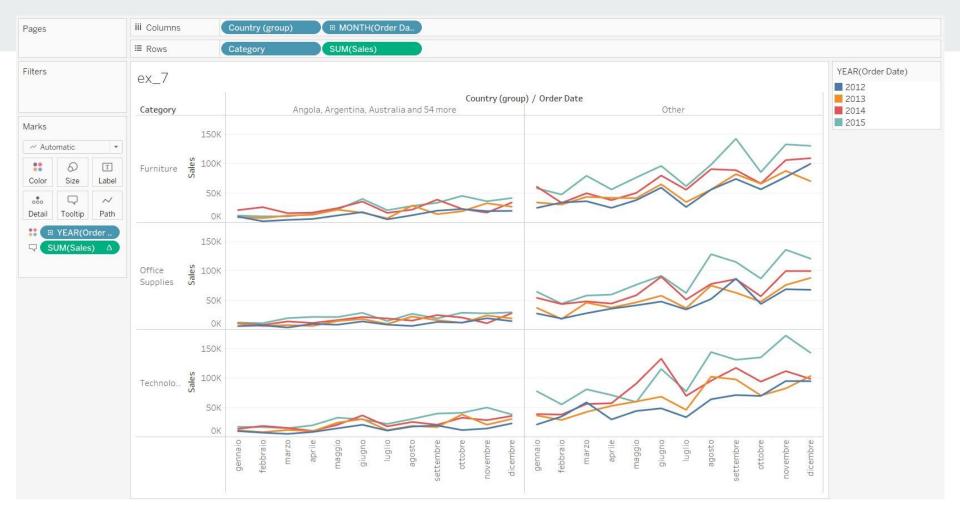
A new data field (Country (group)) will appear in the Data pane.

The new data field is automatically added to the **Color** shelf in the Marks pane.



We have just isolated data about **Sales** and **Profits** in the southern hemisphere, we can compare the <u>sales</u> <u>seasonality</u> chart we have done before to an equivalent chart with our new data.

Let's duplicate the sales seasonality chart sheet and drag and drop the **Country (group)** data field onto the columns shelf (before **MONTH(Order date)**)

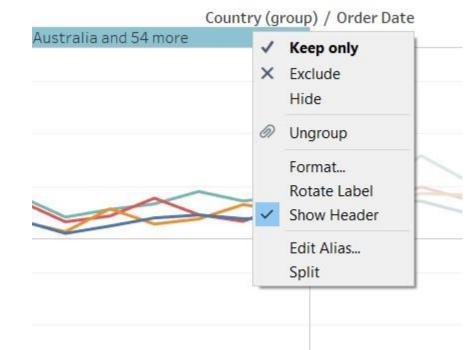


What can we conclude about the seasonality effect from the chart we just created? Is the sales seasonality hypothesis true?

What can we conclude about the seasonality effect from the chart we just created? Is the sales seasonality hypothesis true?

Actually we don't know yet ...

If we want to see how sales go only in the southern hemisphere, we can right click on the new group column title, and select **Keep only** from the dropdown menu.



Notice that when you choose to **Keep only** one dimension/measure on either Row or Column shelves, the dimension on which you operate (which is **Country (group)** in this case) appears on the **Filters** shelf above the Mark pane.



What can we conclude about the seasonality effect from the chart we just created? Is the sales seasonality hypothesis true?

Actually, no, the sales seasonality effect doesn't seem to hold in this case.

Why?



Now, remember the initial hypotheses in our analysis?

- → The fall in sales is due to our actions as a company;
- The fall in sales is due to seasonality.

Also, we had concluded that Furniture was the worst performing category in terms of profits.

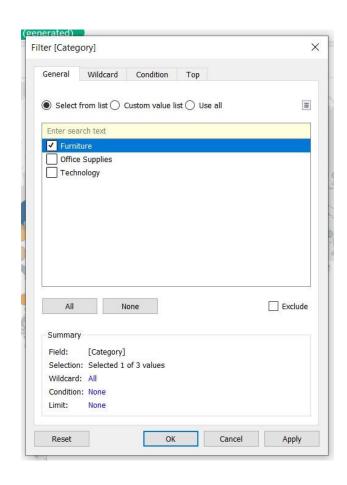
Now, we want to see how *Furniture* sales and profits go over the entire globe on the symbol map we created before.

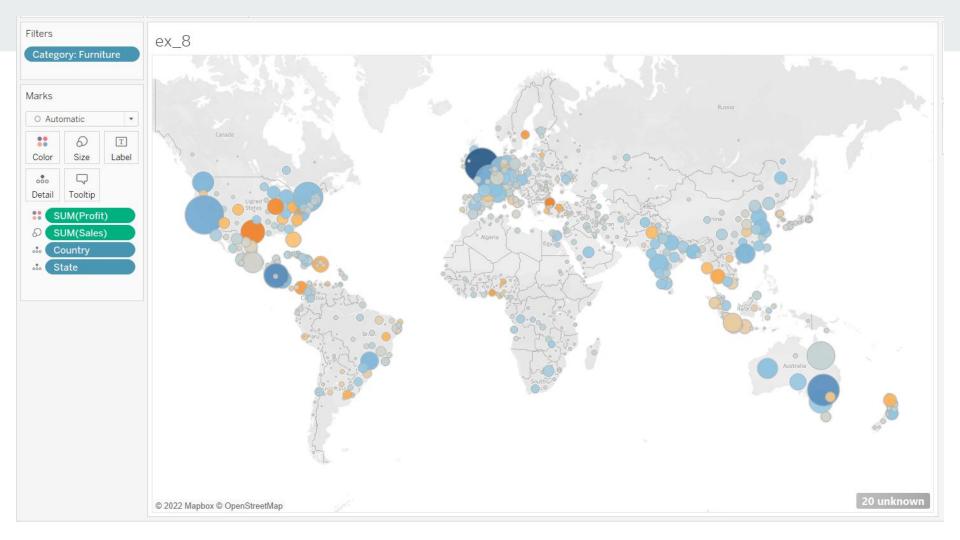
How can we do that? (hint: look at the slide title)

#### We need to:

- → Duplicate the symbol map without hemisphere subdivision;
- → Add Category to the Filters shelf;
- → Select Furniture in the popup window.

You can play around with all the Filter options in this window





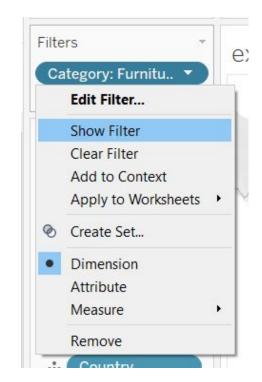
The resulting chart is not very different from the previous; you'll notice that:

- → The colors of the dots have slightly changed to adapt to the Furniture profits;
- → The <u>size of the dots</u> has slightly changed to adapt to the Furniture **sales**.

Notice that if you add continuous numerical measures (e.g. Sales) to the Filters pane, you will also be able to dynamically select the range of data to display

Filters in TP are <u>not static</u>; in fact, you can manipulate your filters by simply **right clicking** onto the **Category: Furniture** pill in the **Filters** pane and select **Show Filter**.

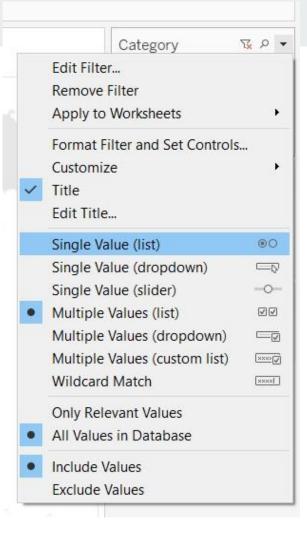
A new pane will appear on the right of your canvas, showing the dynamic **Category** filter.





You can edit the type of filter to show; the default setting is **Multiple Values (list)**, which allows you to choose to display a either a single category at a time or a combination of categories together.

Change the filter type to **Single Value (list)**, which enables you to either display one category at a time or all the categories aggregated.



# 15 minute break

Answer the Wooclap when you come back:



#### Up next:

- → 10 minutes to answer the Woodlap
- → Lessons learned and recap
- → 15-minute Q&A at the end of the lesson

Play around with what we've seen today with different data sources (or start working on your projects)

Additional free data sources are available here

# Recap and lessons learned

# Recap

- → Working with crosstabs:
  - ◆ Table formatting;
  - Dynamic highlighting.
- → Show me tool;
- → Symbol maps:
  - Customization;
  - Data grouping
  - Filtering.

# **Takeaway**

Link to the solution workbook for the lesson: PTBD