### Tableau Lecture 4: Filters and Calculations

#### **Agenda**

* Charts
  + Stacked Bar chart
  + Scatter plot
  + Treemap
  + Combined axis chart
  + Dual-axis chart
* Filters and Calculations
  + Types of Filter
  + Order of Operations

#### Dataset :[**sample superstore**](https://docs.google.com/spreadsheets/d/1Dim0dtvQBWWvrqRRxTt0WLAI6VOJs_oe/edit#gid=2115366062)

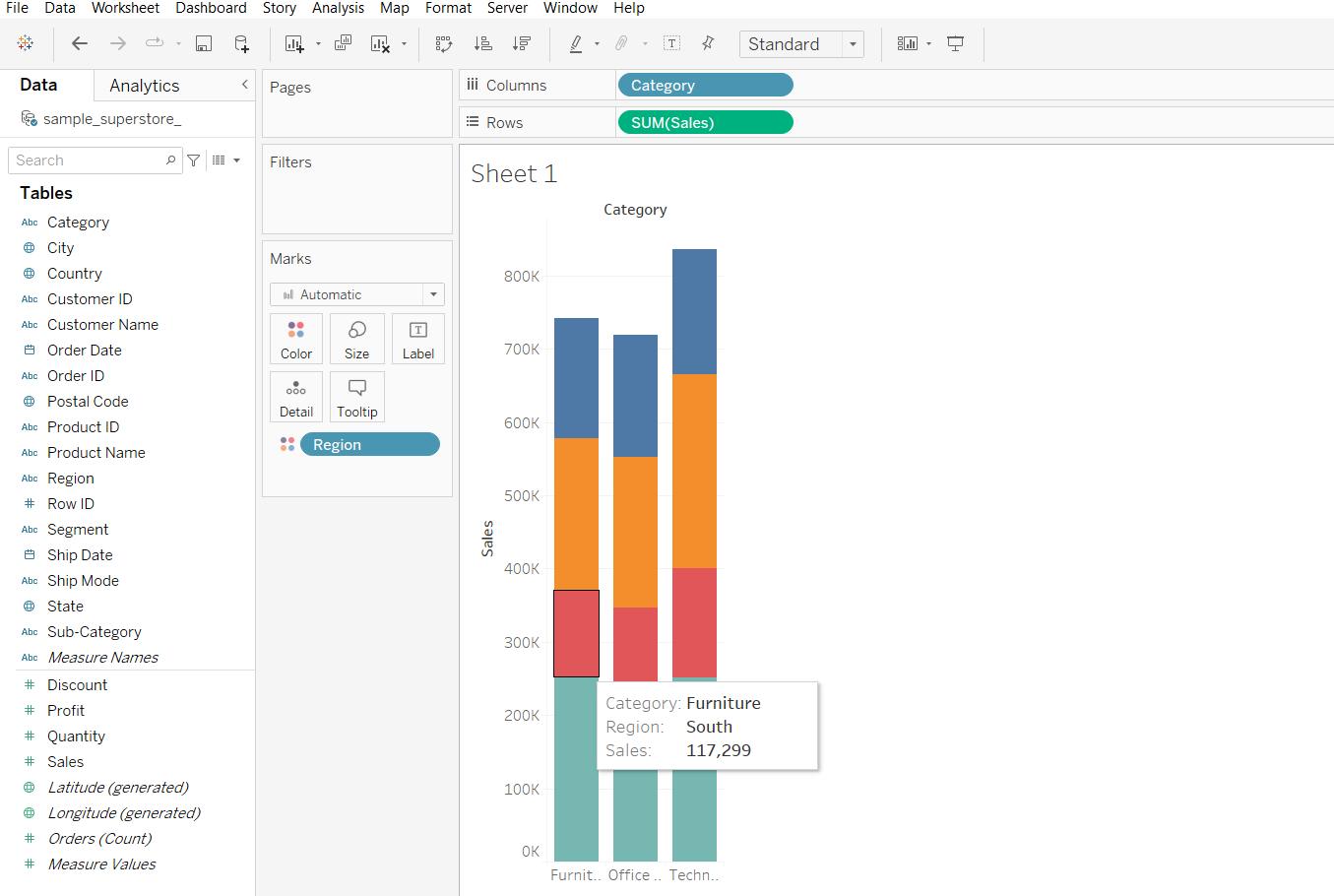
### **Stacked Bar Chart**

* It is a simple bar chart with segmented bars.
* It requires
  + 1 or more dimension and
  + 1 or more measures.
* When you select dimension into the color marks chart what you get is a stacked bar chart.

#### **Business problem 1:**

Display total sales for each category by region.

* Drag Category to Columns
* Drag Sales to Rows
* Drag Region to Color in Marks card



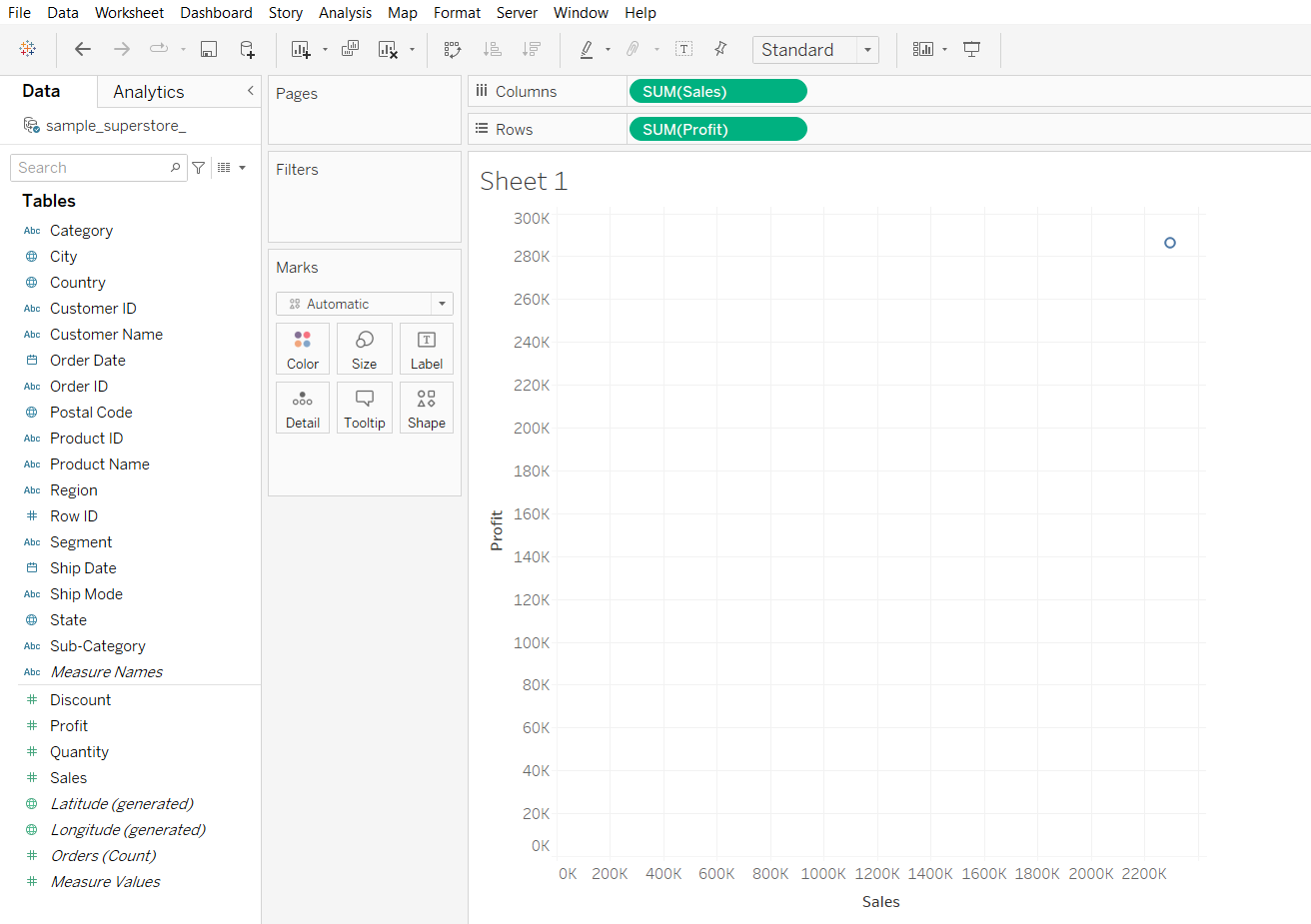
**Scatter Plot**

* It displays a data point at their respective intersection of both the measures
* We can create a scatter plot by adding at least one measure on row shelf and one measure on column shelf

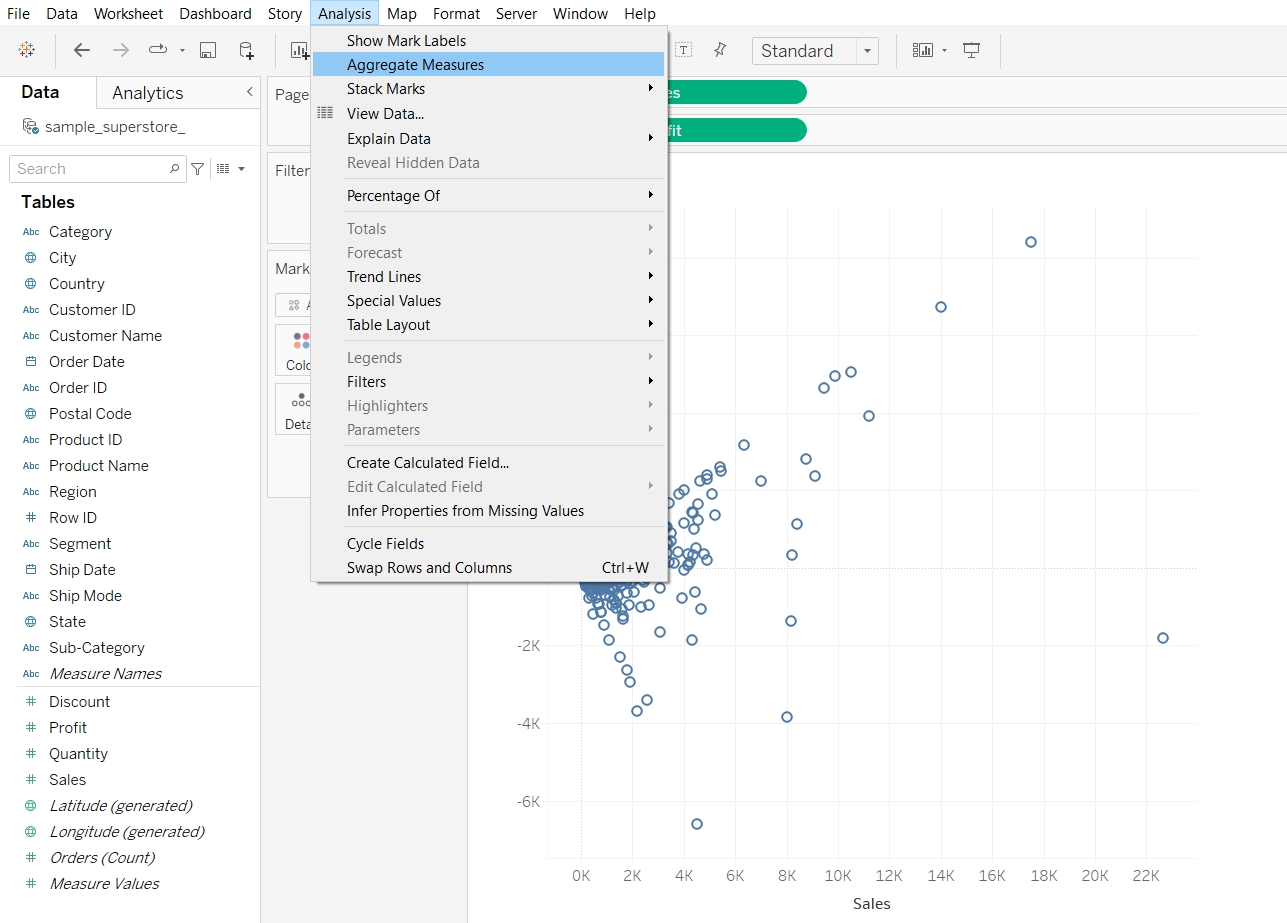
#### **Business problem 2:**

Find the order id that has the highest sales and highest profit values.

* Creating a scatter chart what we see below is the total sales and total profit.
  + Drag Profit to Rows
  + Drag Sales to Columns



* Disaggregating the measures to plot every row values in the plot and solving the business problem.
  + Click on Analysis in Menu bar
  + Uncheck the Aggregate Measures option



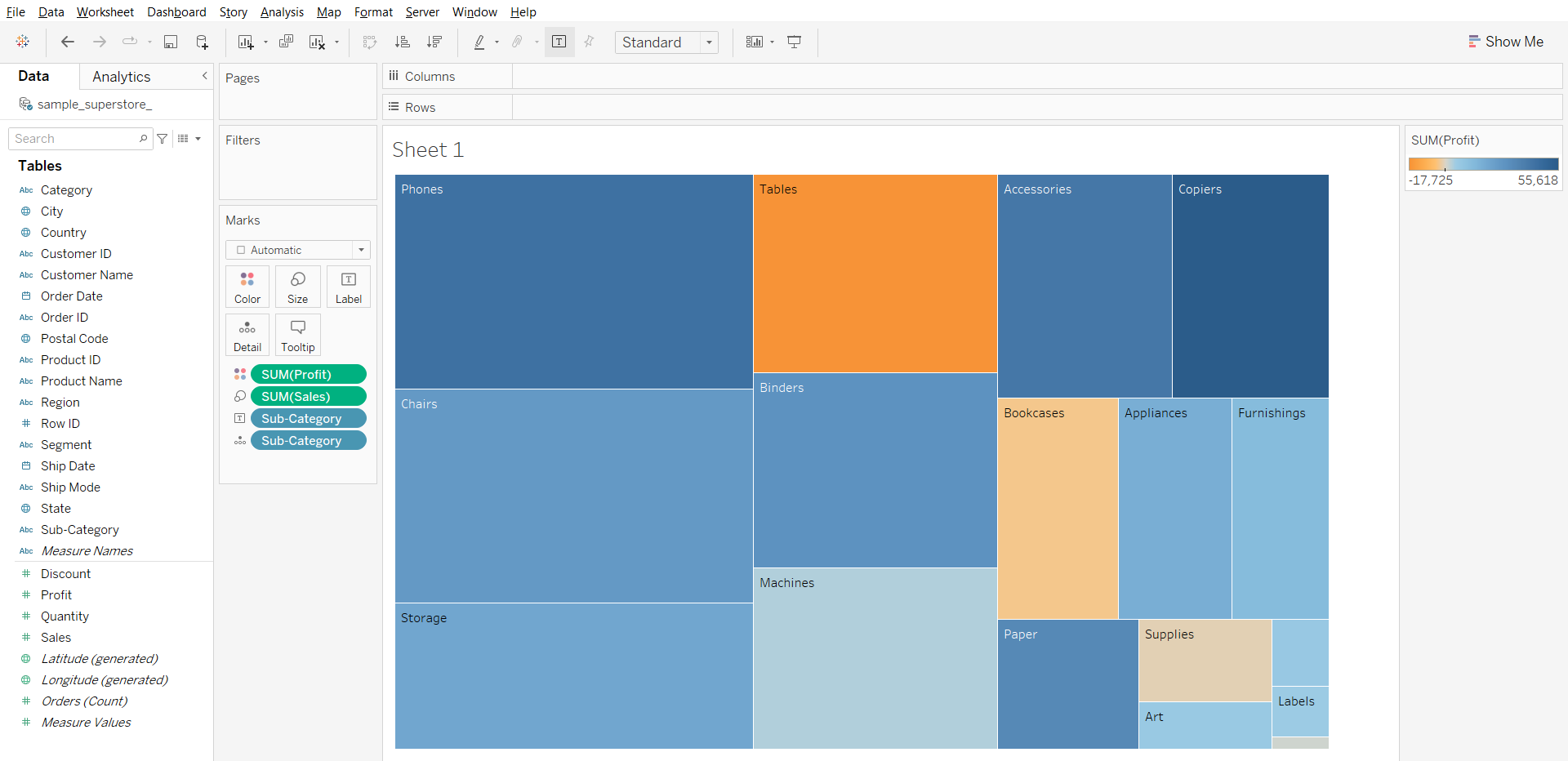
**Tree Map**

* It is a visualization that nests rectangles in hierarchies
* It uses dimension to define the structure of the tree map and measure to define the size or color of each of the rectangle
* It requires one or more dimensions and one or 2 measures

#### **Business problem 3:**

Find the subcategory that has the highest sales and highest profit.

* Drag sub-category to Detail in Marks Card
* Drag sub-category to Label in Marks card
* Drag Sales to Size in Marks card
* Drag Profit to Color in Marks Card

l

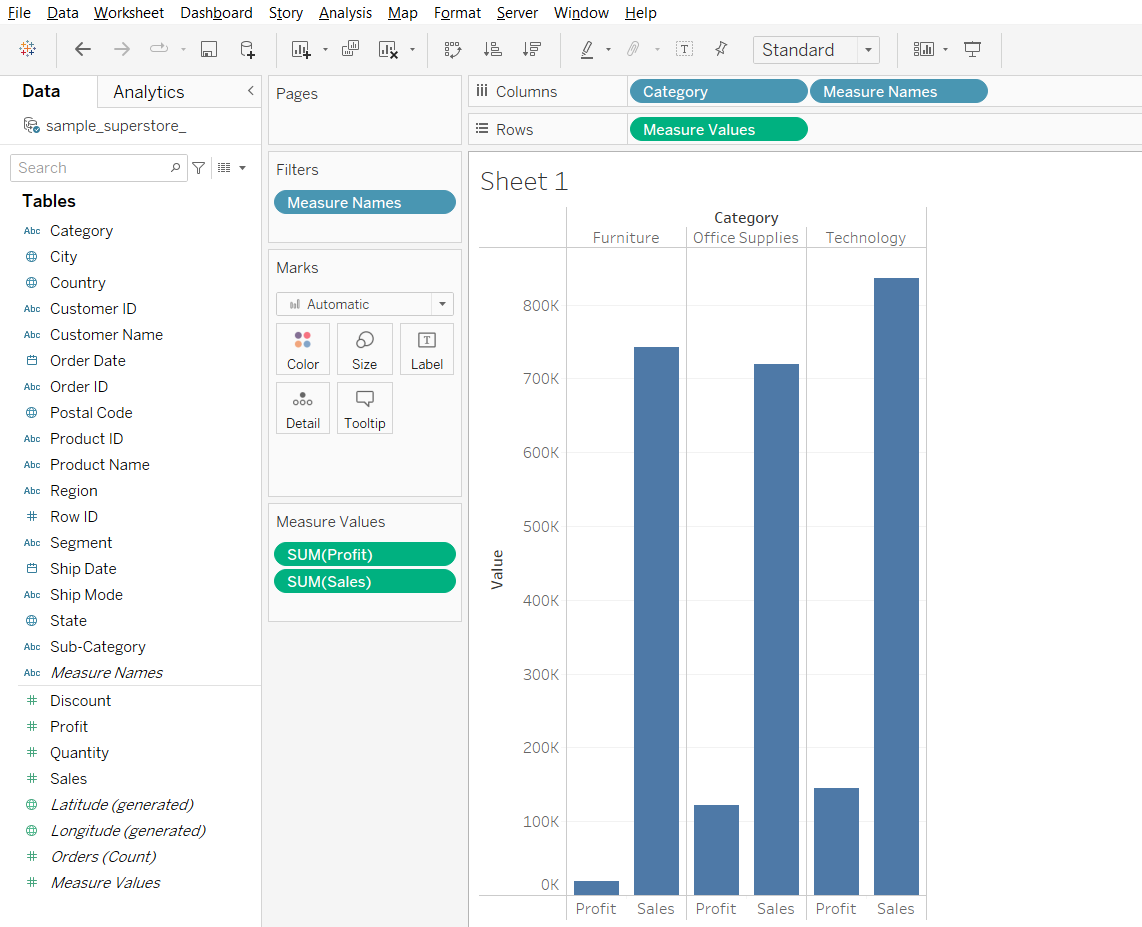
### **Combined Axis Chart**

* Also called as Blended axis or shared axis chart
* Measures share a single axis so that all the marks are shown in a single pane
* They are used to compare measures of same data type and same scale
* Only single Mark card is present and we can compare more than 2 measure
* To create a combined axis chart drag the measure on the left side of the already existing axis
* It requires one or more dimension and 2 or more measures

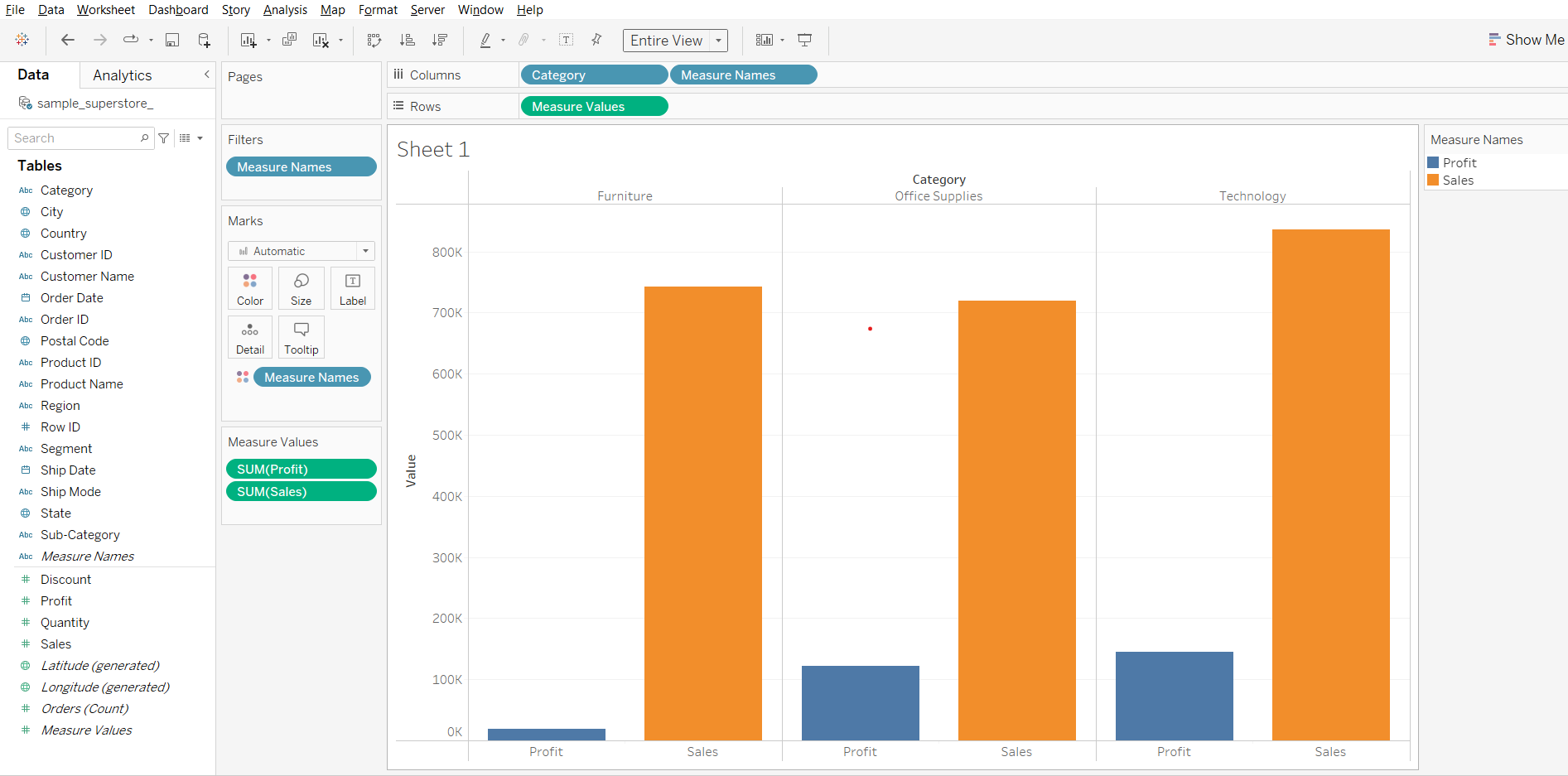
#### **Business problem 4:**

Compare sales and profit for each category in a single plot.

* Drag Category to Columns
* Drag Sales to Rows
* Drag Profit to vertical axis of the chart



* Drag Measure Names to Color in Marks Card
* Change Fit from Standard to Entire View in toolbar



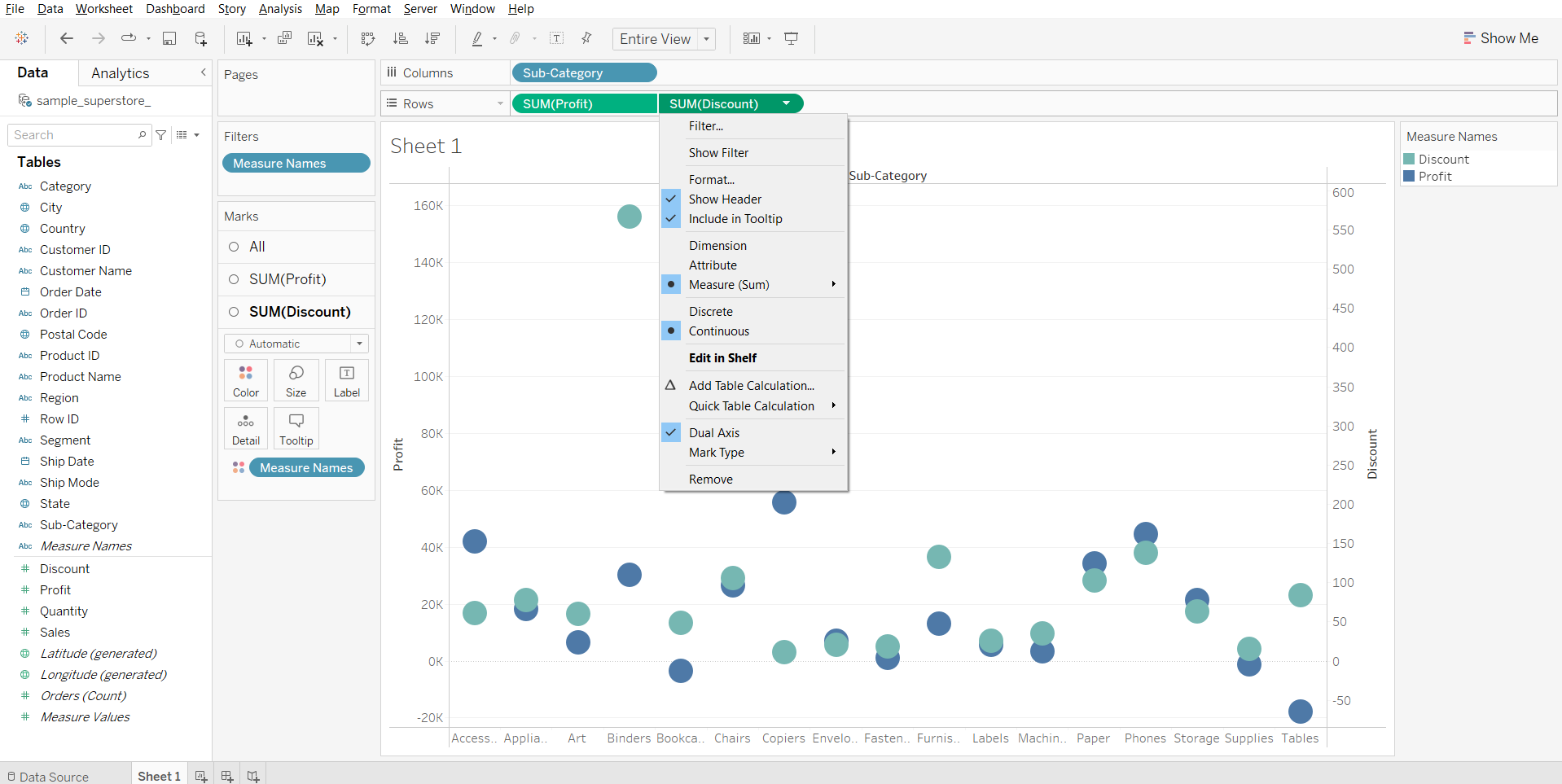
**Dual Axes Charts**

* You can compare multiple measures using dual axes, which are two independent axes that are layered on top of each other.
* Dual axes are useful for analyzing two measures with different scales.
* Creates multiple marks
* Can’t compare more than 2 measures
* To create a dual axis chart we can drag 2nd measure to right side of the plot to create another axis or right click on the 2nd measure and select dual axis
* It requires 2 measures and at least one dimension

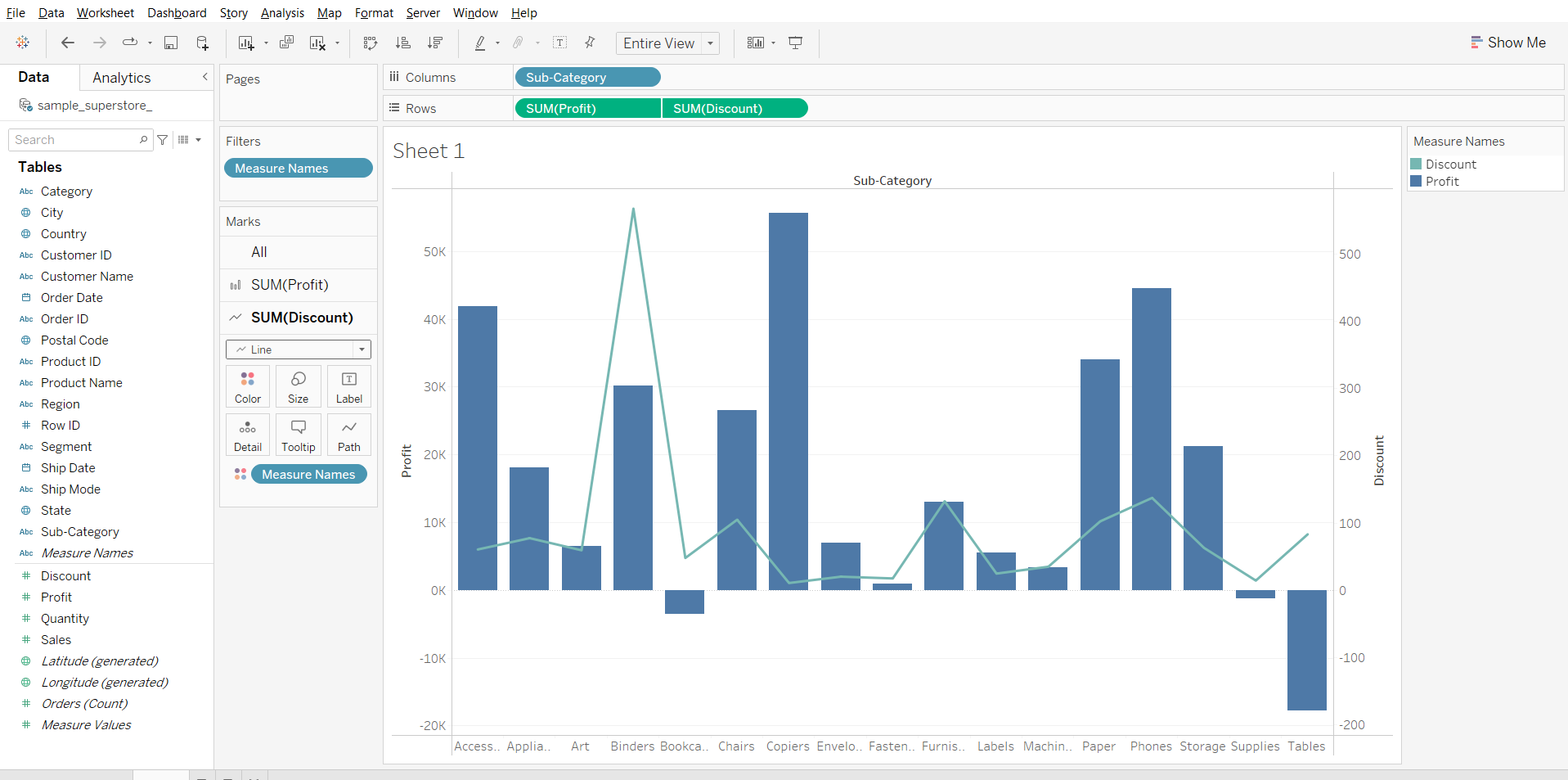
#### **Business problem 5:**

Find the sub-category that has the lowest discount but highest profit.

* Drag Sub-Category to Columns
* Drag Profit to rows
* Drag Discount to Rows
* Open the Discount field dropdown and choose Dual Axis



* In Marks card, change plot type of Profit to Bar
* In Marks card, change plot type of Discount to Line



**Filters**

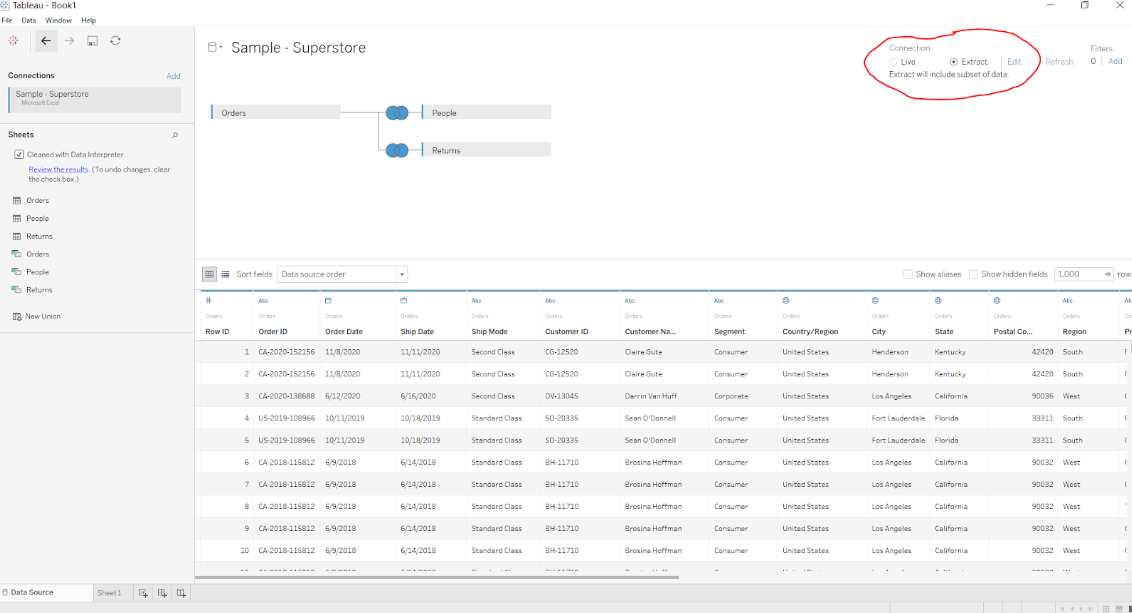
* Filter removes some scope of data from a data set.
* Filters are very helpful to create dashboards in Tableau.
* Filters can help to minimize the size of data sets for efficient use, eliminate irrelevant dimension elements, clean up underlying data, set date ranges and measures as required, simplify and organize data, etc.

### **Types of Filters**

#### **1. Extract filters**

* Extract filters in Tableau are used to extract a small subset of data from the original data source.
* Tableau then creates a local copy of the data set that is to be stored in the repository.
* These methods reduce Tableau queries.
* The data size can be further reduced by applying the measure or dimension filter to the extract as required.
* Unlike Tableau Desktop(paid) which supports both extract and live connections, Tableau public, by default, only supports extract.

**Important Note:**Tableau Public always works with extracts, and therefore does not show the options ‘Live’ and ‘Extract’.  
Here is a screenshot of Tableau Desktop Professional that supports both extract and live connections.



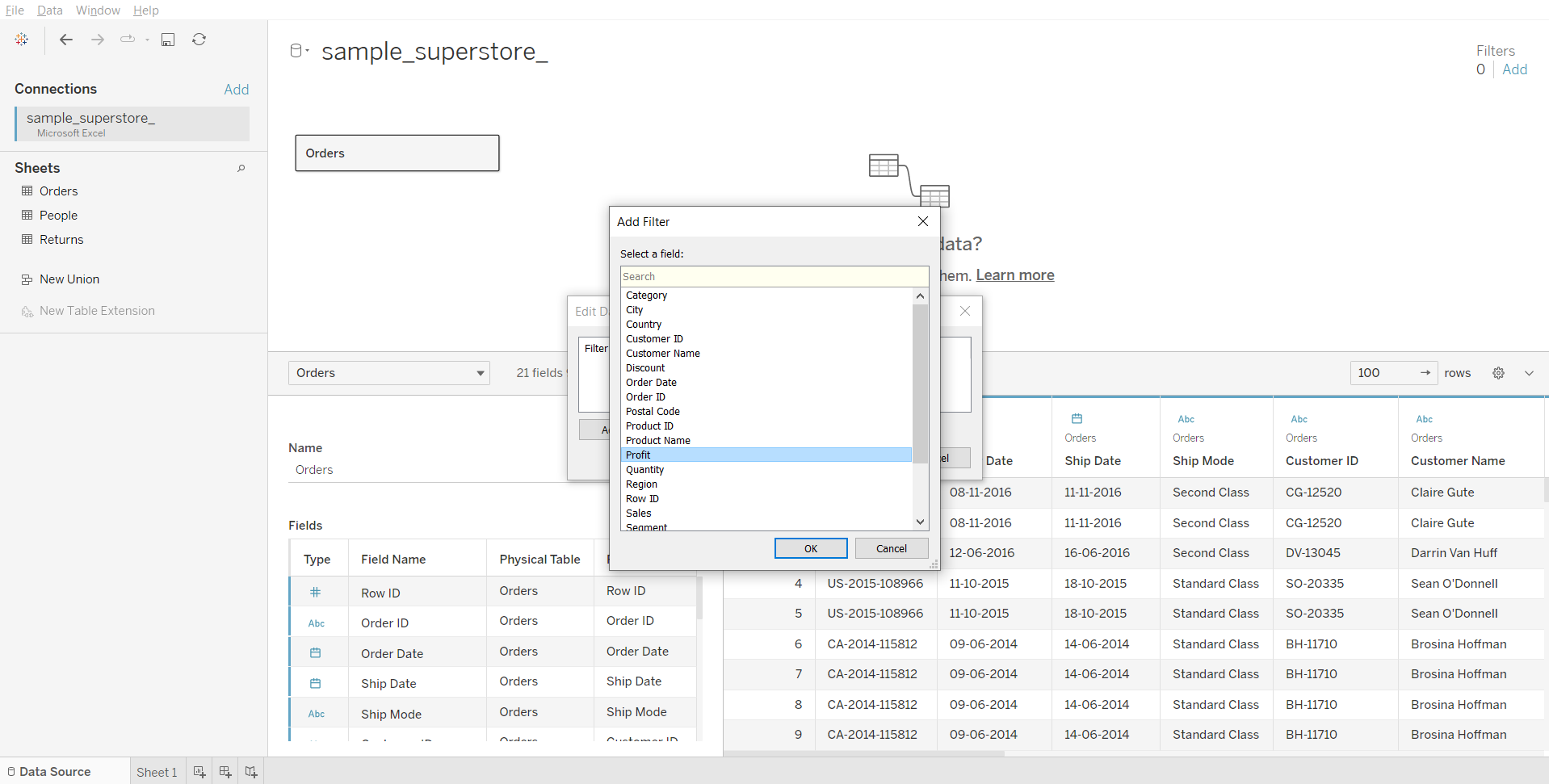
#### **2. Data Source filters**

* Data source filters in Tableau are mainly used to restrict sensitive data from viewers and reduce data feeds.
* Viewers can, however, have certain access rights to view the underlying data.
* Data source filters allow the direct application to source data.
* One important thing to mention is that the extract filter and the data source filter are not linked, and if you happen to go back to a live connection, the data source filter will remain intact.

#### **Business Problem 6 :**

Show orders which made profit >= 1K.

* Go to the Data Source tab
* Click on Add Filter on the top right.
* Click on Add…
* Select Profit from the list of fields.
* Set the minimum value to 1000.



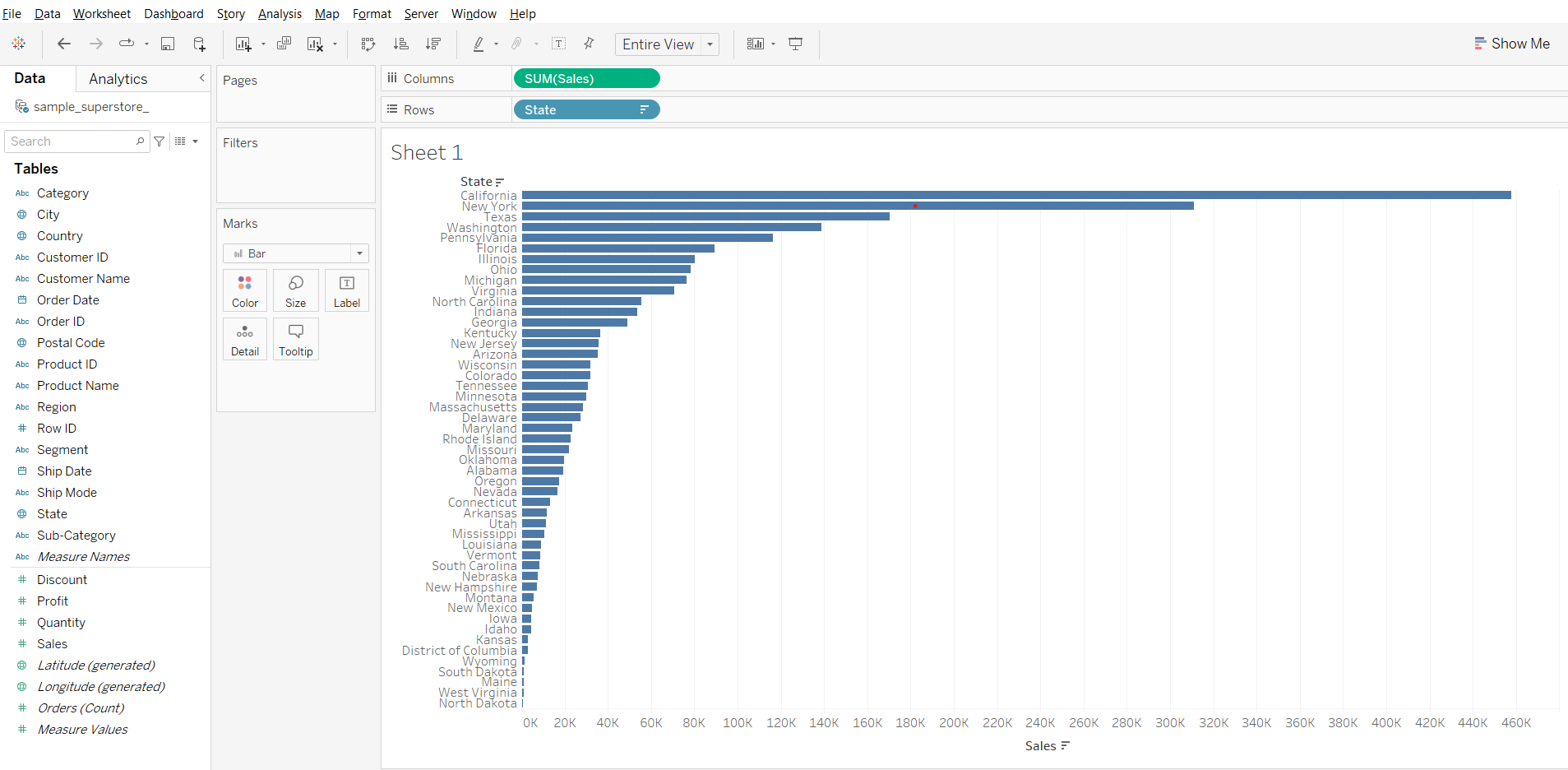
#### **3. Context filters**

* By default, all filters that you set in Tableau are computed independently. That is, each filter accesses all rows in your data source without regard to other filters. However, you can set one or more categorical filters as context filters for the view.
* You can think of a context filter as being an independent filter. Any other filters that you set are defined as dependent filters because they process only the data that passes through the context filter.
* The context filter adds an actionable context to data analysis, but if the data is not reduced enough, the cost of computing can be very high.

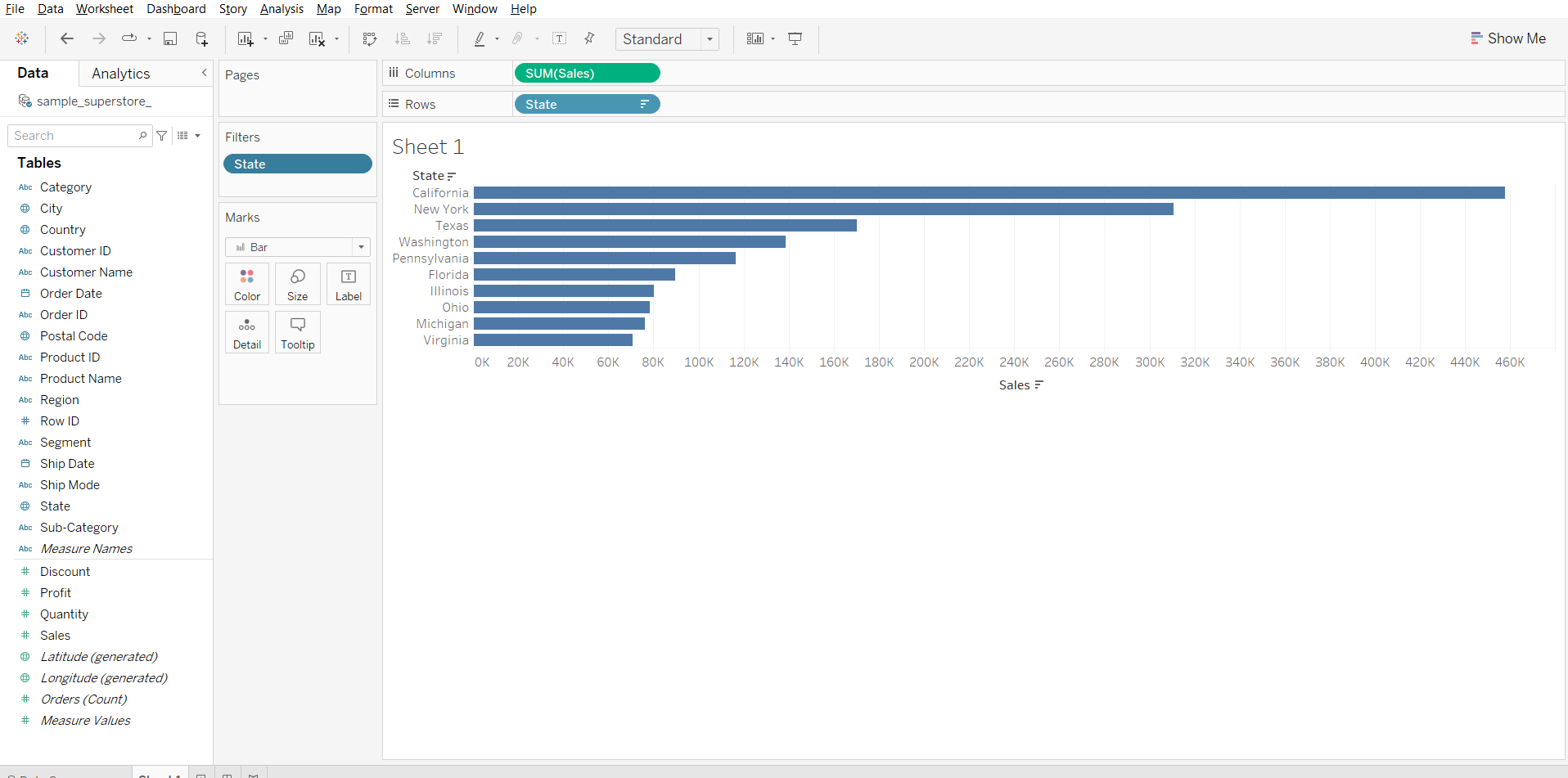
#### **Business Problem 7 :**

Find the top ten states in the West region with the highest sales.

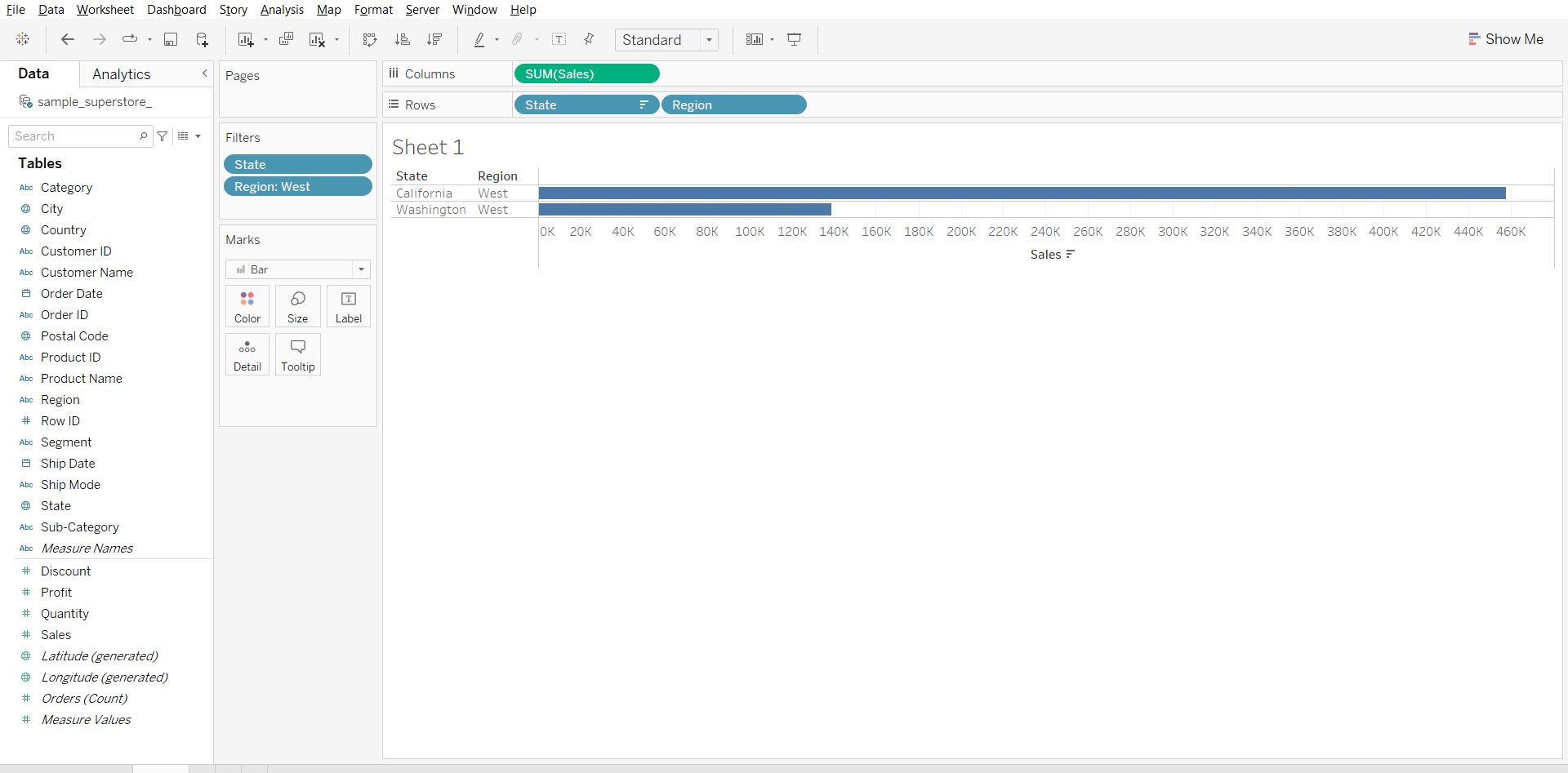
* Drag State to Rows
* Drag Sales to Columns
* Click on the sort-descending button in toolbar



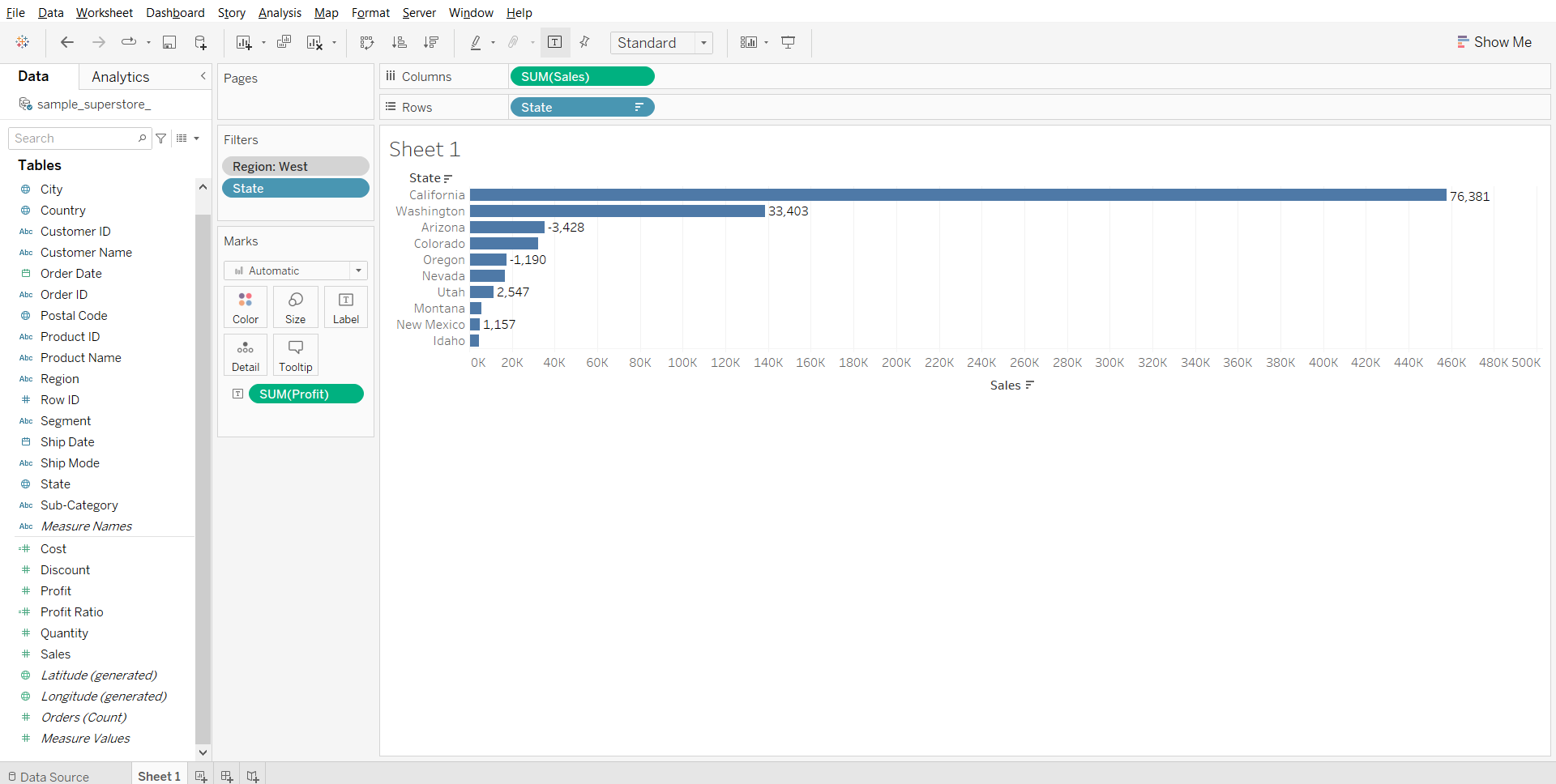
* Drag State to Filters card
* Click on the Top tab of Filter dialog box
* Choose Top 10 by Sales field



* Drag Region to Rows
* Drag Region to Filters card
* Deselect all regions, then select West



* Right click on Region in Filters card, select Add to Context.



**Explanation:**

1. First we filter out top 10 states by sales
2. Next when we apply filter for region=West
3. We only see 2 entries that instead of 10 that is because top 10 filter gets executed first and then the region filter is executed if you see from the 10 states only 2 states belong to west region
4. Now in order to change the order of operation we add region filter to context filter this ensures that first our region filter gets executed and then the top 10 by sales is executed.

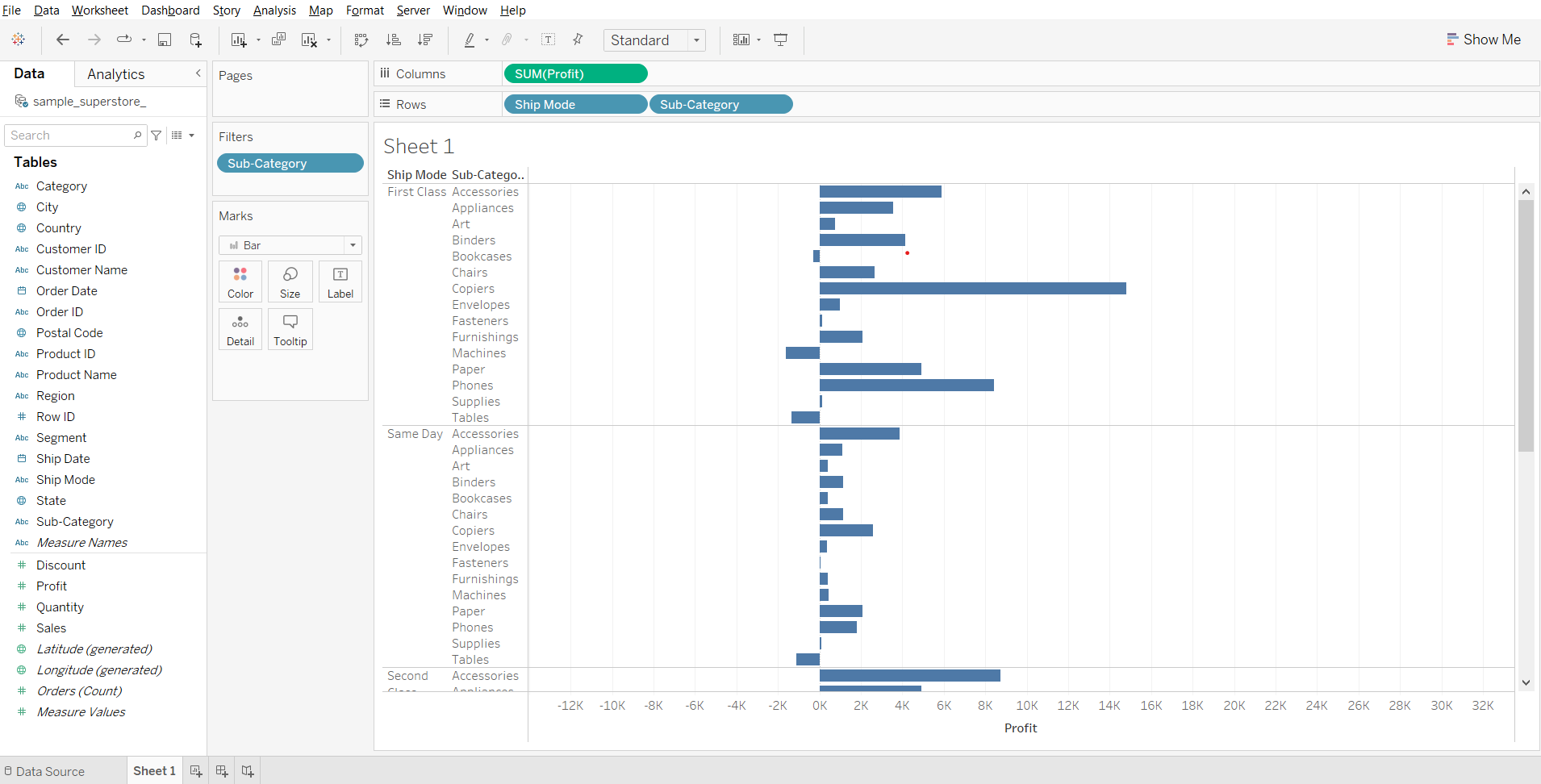
#### **4. Filters on dimensions (whether on the Filters shelf or in filter cards in the view)**

* Dimension filters in Tableau are non-aggregated filters. The dimensions that are used are mostly blue pills. Blue pills correspond to discrete data.
* If there are many dimensions, one can search for them. Dimension filter provides four options, General, Wildcard, Condition, and Top/Bottom. You can pick up any of the four options to select the right data or remove the unwanted data.
* One can create their own formula as well and then use it in the Condition filter and the Top/Bottom filter for data selection. They provide a channel to measure to get the required data.

#### **Business Problem 8:**

Show ship mode and subcategories wrt profit where subcategories Labels and Storage are excluded.

* Drag Ship Mode to Rows
* Drag Sub-Category to Rows
* Drag Profit to Columns
* Drag Sub-Categories to Filters card
* Deselect all boxes and select Labels and Storage
* Check the exclude box and click OK.



#### **5. Filters on measures (whether on the Filters shelf or in filter cards in the view)**

* Using a Measure filter in Tableau allows for various operations and aggregate functions such as sum, median, avg, standard deviation, etc.
* Aggregated filters are always applied after non-aggregated filters, no matter what the order is on the Filters pane.
* The filters are applied to Measure fields consisting of quantitative data.

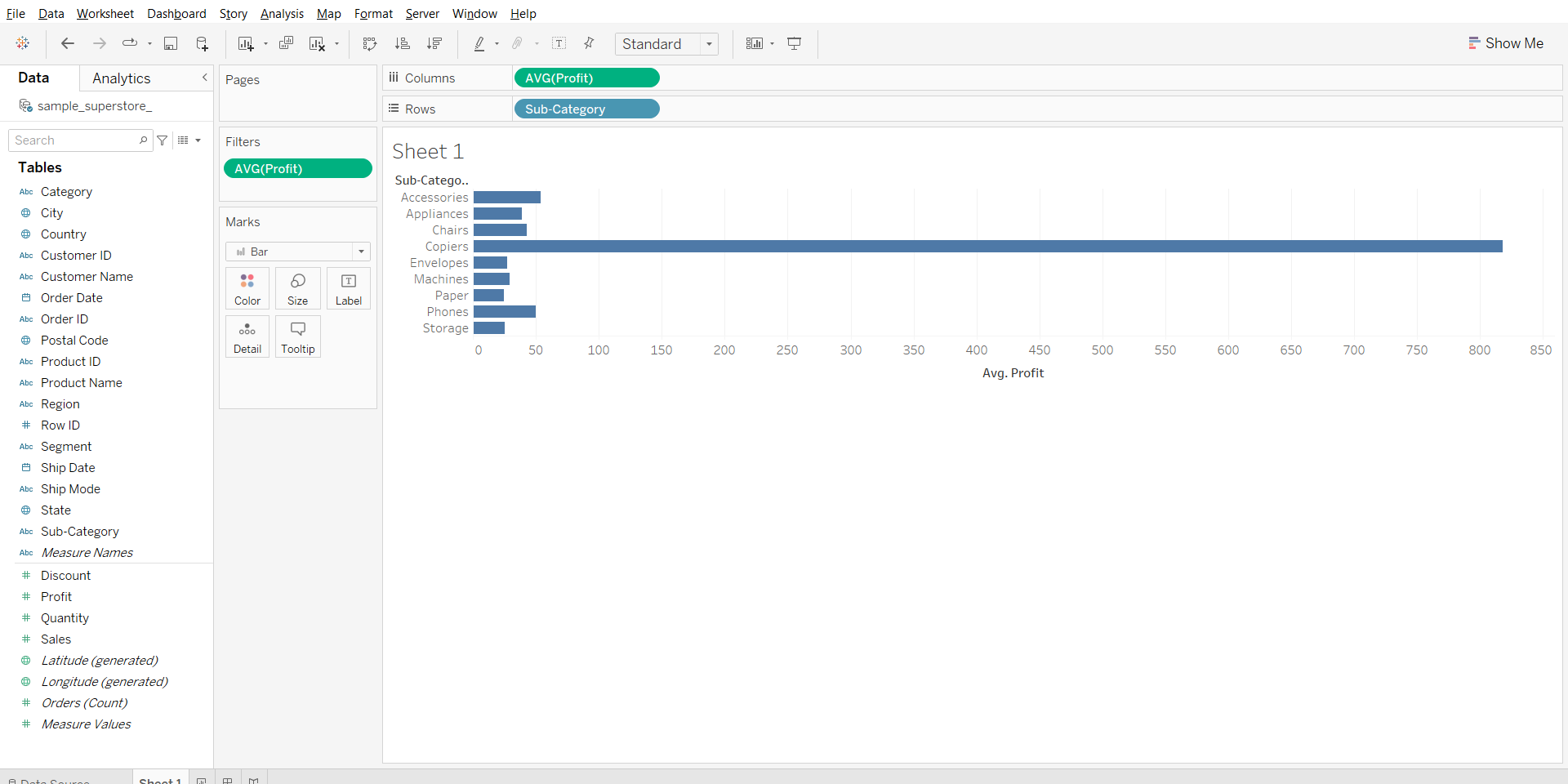
In a subsequent dialog box, you will get four types of filters:

1. Range: Select the range of values to include in the result
2. At least: Select the minimum value of a measure
3. At most: Select the maximum value of a measure
4. Special: Select null or non-null values

#### **Business Problem 9 :**

Show only the subcategories whose average profit is greater than 20.

* Drag Profit to Columns
* From the Profit drop down change measure to AVG.
* Drag Sub-Category to Rows
* From the Profit dropdown open Filter dialog box
* Change the minimum value to 20 and click OK.

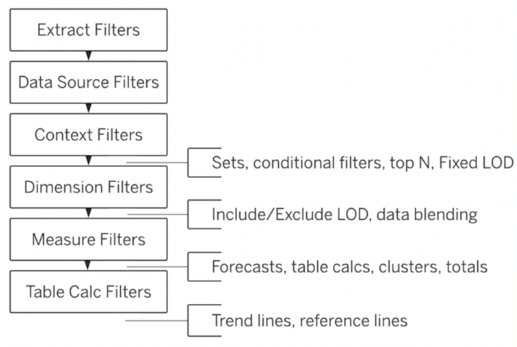


**Filtering Order of Operations**

Tableau performs actions on our view in a very specific order; this is called the Order of Operations. It helps you avoid filter conflicts and achieve efficiency with your dashboard.

Filters are executed in the following order:

1. Extract filters
2. Data source filters
3. Context filters
4. Filters on dimensions (whether on the Filters shelf or in filter cards in the view)
5. Filters on measures (whether on the Filters shelf or in filter cards in the view)
6. The Table Calculation filter



### **Calculated Fields**

* Calculated fields allow you to create new data from data that already exists in your data source.
* When you create a calculated field, you are essentially creating a new field (or column) in your data source, the values or members of which are determined by a calculation that you control.
* This new calculated field is saved to your data source in Tableau, and can be used to create more robust visualizations.
* But don’t worry: your original data remains untouched.

You can use calculated fields for many, many reasons. Some examples might include:

* To segment data
* To convert the data type of a field, such as converting a string to a date.
* To aggregate data
* To filter results
* To calculate ratios

### **Formatting Calculations in Tableau**

There are four basic components to calculations in Tableau:

* Functions - Statements used to transform the values or members in a field.
* Fields - Dimensions or measures (columns) from your data source.
* Operators - Symbols that denote an operation.
* Literal expressions - Constant values that are represented “as is”, such as “Profitable” and “Unprofitable”

**Note:** The syntax (how to format them to work in Tableau) of each of these components will be discussed in class

### **Operators in Tableau**

We have mainly 3 types of operators in Tableau.

* Arithmetic operator
  + Few types of arithmetic operator are addition,subtraction,division,multiplication,modulo
  + Example-profit+sales,profit/sales,
* Comparison operator
  + Few types of comparison operator are “==,>,<,>=,<=,!=”
  + Example-sales>100k,profit<=1k
* Logical operator
  + Few types of logical operator are “And,OR,Not”
  + Example-profit=100 AND sales=1000,Not(sales=profit)

Order of precedence of Operators : [Link](https://help.tableau.com/current/pro/desktop/en-us/functions_operators.htm)

### **Tableau Functions**

Tableau supports many functions for use in Tableau calculations.

They have been organized category-wise below :

* [Number functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_number.htm) : Number functions allow you to perform computations on the data values in your fields. Number functions can only be used with fields that contain numerical values.
* [String functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_string.htm) : String functions allow you to manipulate string data (i.e. data made of text).
* [Date functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_date.htm) : Date functions allow you to manipulate dates in your data source.
* [Type conversion](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_typeconversion.htm) : Type conversion functions allow you to convert fields from one data type to another.
* [Logical functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_logical.htm) : Logical calculations allow you to determine if a certain condition is true or false (boolean logic).
* [Aggregate functions](https://help.tableau.com/current/pro/desktop/en-us/calculations_calculatedfields_aggregate_create.htm) : Aggregate functions allow you to summarize or change the granularity of your data.
* [User functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_user.htm) : User functions can be used to create user filters or row-level security(RLS) filters that affect visualizations published to Tableau Server or Tableau Online, so that only certain people can see your visualization. Not available in Tableau Public.
* [Table calculations functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_tablecalculation.htm): Table calculation functions allow you to perform computations on values in a table.
* [Pass-Through functions (RAWSQL)](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_passthrough.htm) : These RAWSQL pass-through functions can be used to send SQL expressions directly to the database, without first being interpreted by Tableau. If you have custom database functions that Tableau doesn’t know about, you can use these pass-through functions to call these custom functions. Not available in Tableau Public. Link1 Link2
* [Spatial functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_spatial.htm) : Spatial functions allow you to perform advanced spatial analysis and combine spatial files with data in other formats like text files or spreadsheets.
* [Additional functions](https://help.tableau.com/current/pro/desktop/en-us/functions_functions_additional.htm) :
  + Regular Expressions
  + Hadoop Hive Specific Functions
  + Google BigQuery Specific Functions

### **Types of Calculations**

There are three main types of calculations you can use to create calculated fields in Tableau:

* Basic expressions : They allow you to transform values or members at the :
  + Data source level of detail (a row-level calculation) or
  + At the visualization level of detail (an aggregate calculation).
* Level of Detail (LOD) expressions :
  + They too allow you to compute values at the data source and the visualization levels.
  + LOD expressions give you even more control (compared to basic expressions) on the level of granularity you want to compute.
  + They can be performed at a more granular level (INCLUDE), a less granular level (EXCLUDE), or an entirely independent level (FIXED).
* Table calculations : They allow you to transform values at the level of detail of the visualization only and are added to measures.

### **Basic expressions**

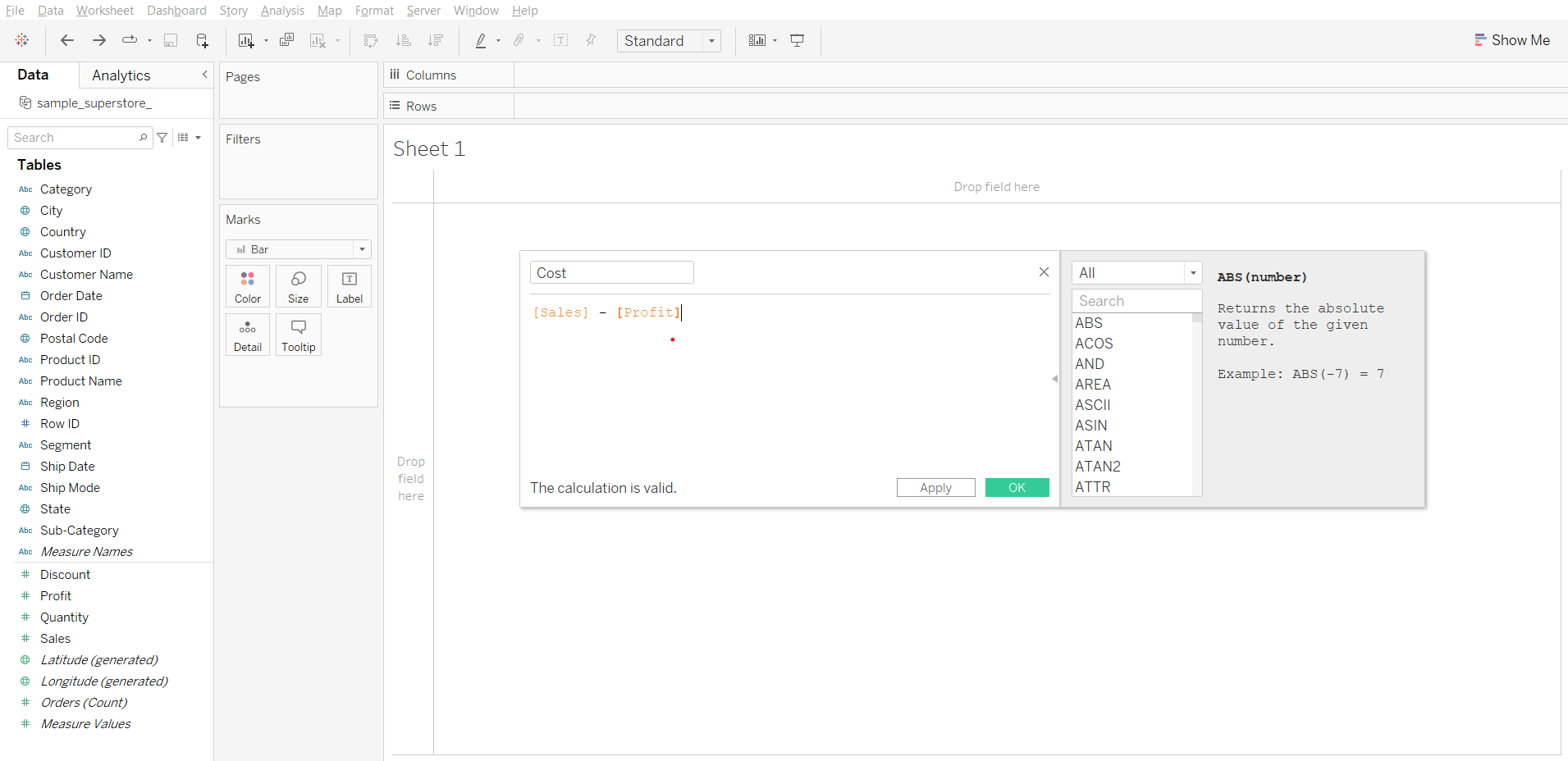
**Row level calculations:**

Calculations are performed for every row of underlying data.

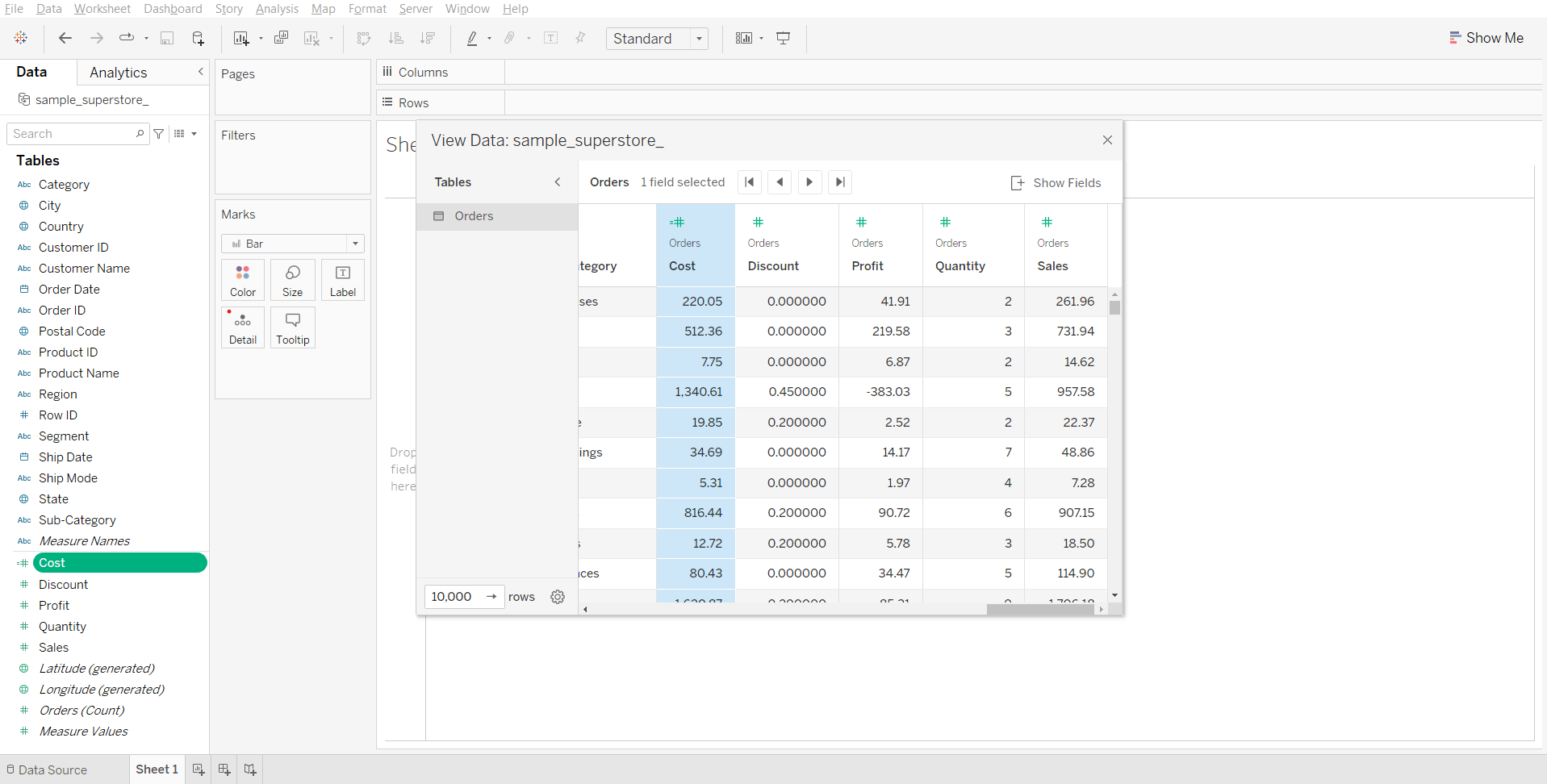
#### **Business problem 10:**

Find average cost of each product sub-categories (cost=Sales-Profit)

* Click on the drop down at top of left pane
* Choose Create Calculated Field…
* Enter the formula “[Sales] - [Profit]”
* Rename the field as Cost and click ok.



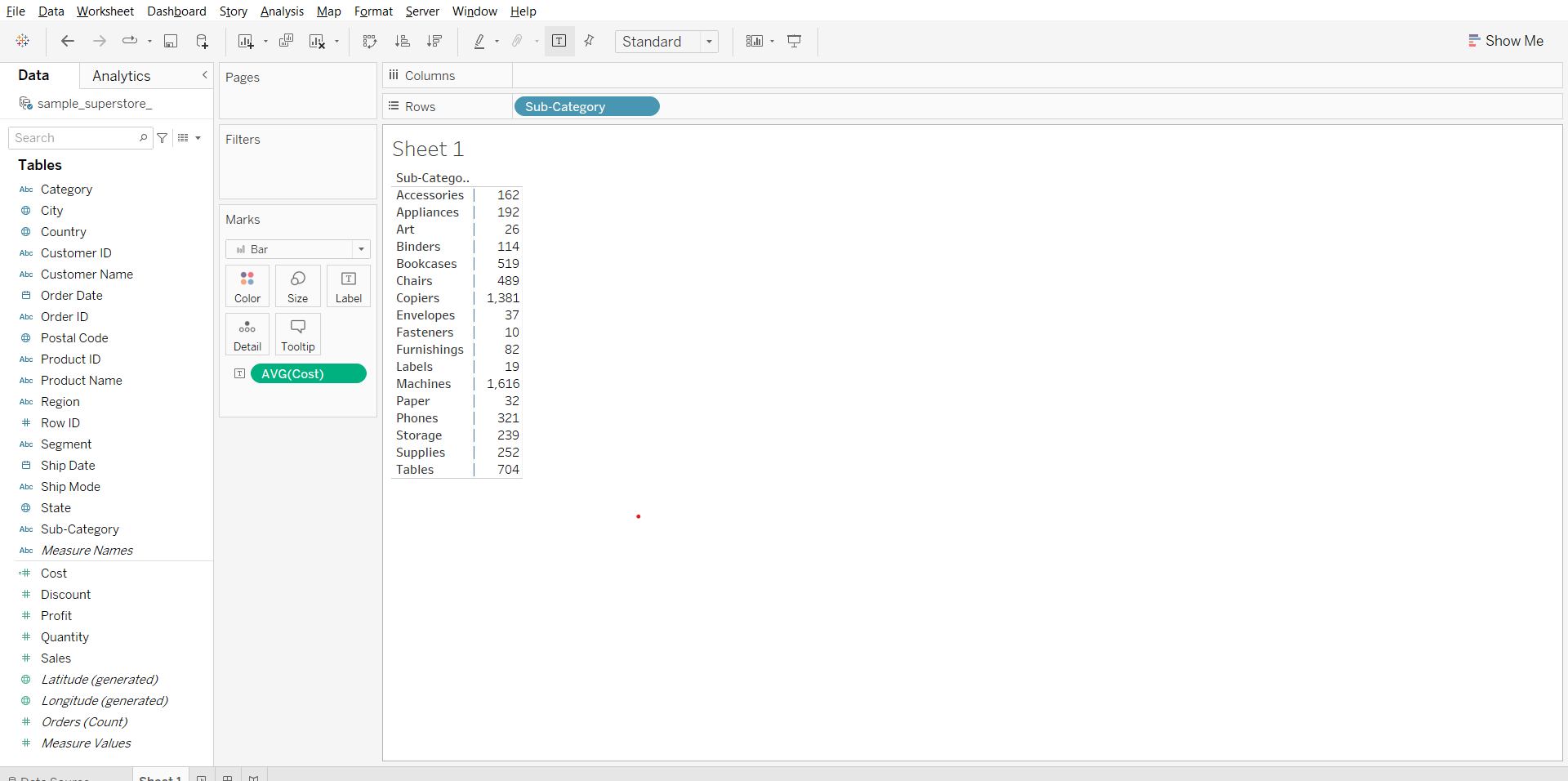
* We can verify that it is done for each row by clicking on view data (near the search bar in left pane) we can see a new column added “cost”



**Note:** Subtraction arithmetic operator is used in this calculation.

#### **Solving the business problem:**

* Drag Sub-Category to Rows
* Drag Cost to Label in Marks card
* Change the measure of Cost to AVG using its dropdown menu

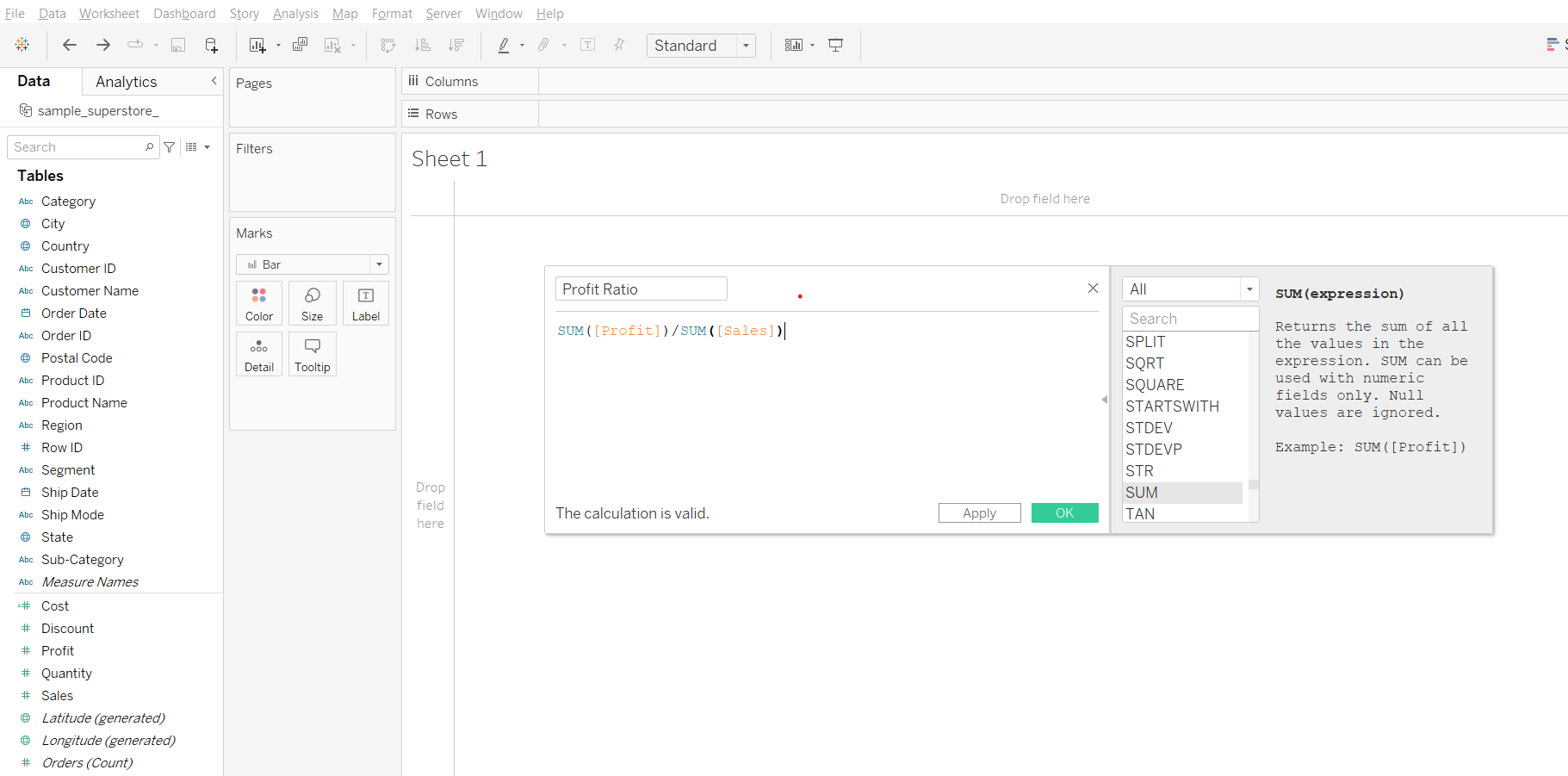


**Aggregation calculations:**Calculations are performed at an aggregate level which is defined by the dimension used in the view.

#### **Business problem 11**

Find profit ratio of each product sub-category. Profit ratio = profit / sales

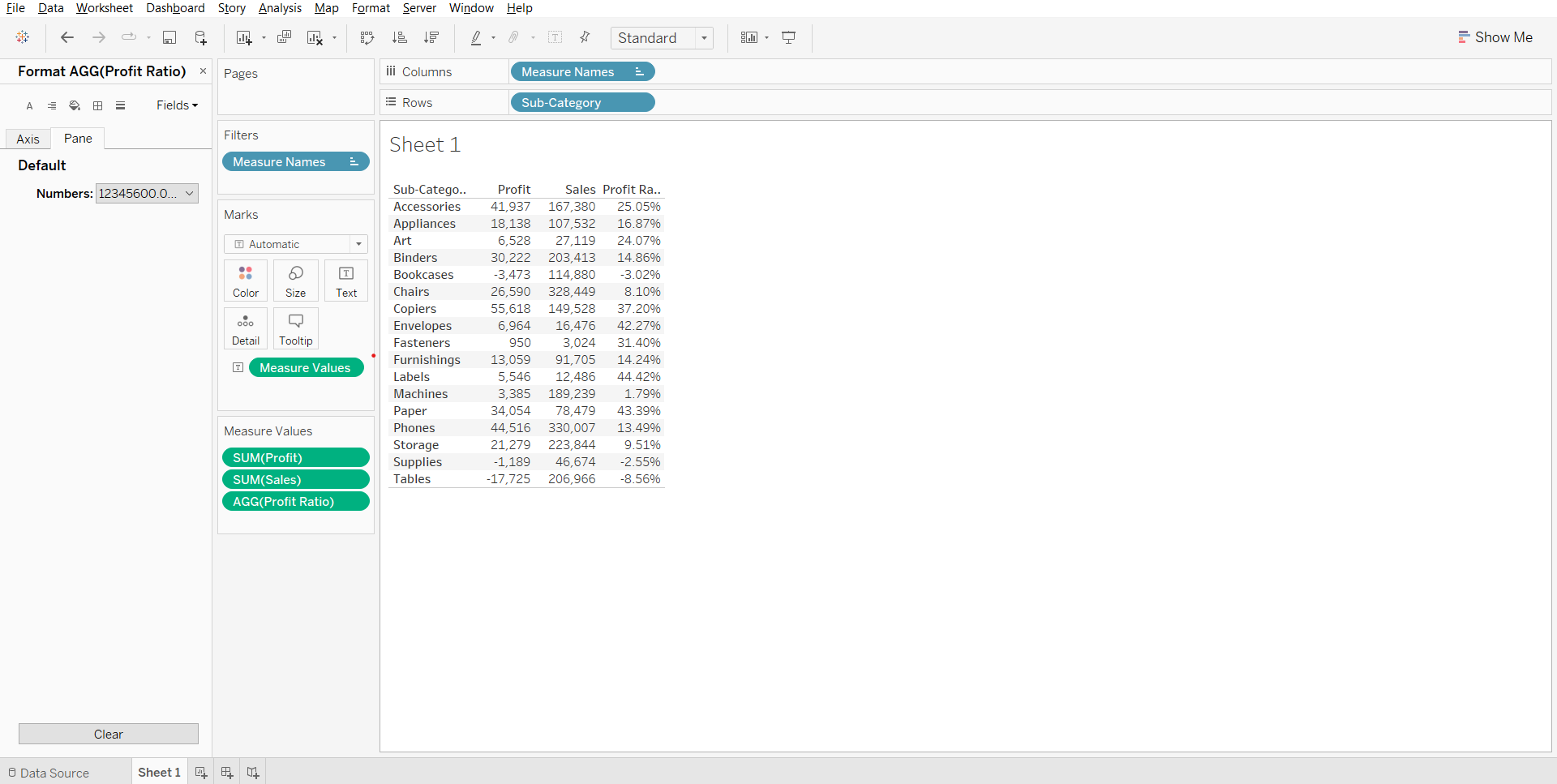
* Open the calculated field dialog box
* Enter the formula - SUM([Profit])/SUM([Sales])
* Name the new field as Profit Ratio and click ok.



**Note:** Division arithmetic operator is used in this calculation.

#### **Solving the business problem:**

* Drag Sub-Category to Rows
* Drag Profit to the labels column next to Sub-Category in the sheet
* Drag Sales to the Profit column in the sheet
* Drag Profit Ratio to the Sales column
* From the Measure Values card change the format of Profit Ratio to Percentage



**Note:**

* Notice AGG in pill profit ratio refer here the SUM is performed 1st based on the dimension sub category (i.e. sum of profit and sum of sales for each sub category) and then the division occurs.
* Now if we change the dimension from sub category to category the calculation will be done based on it.