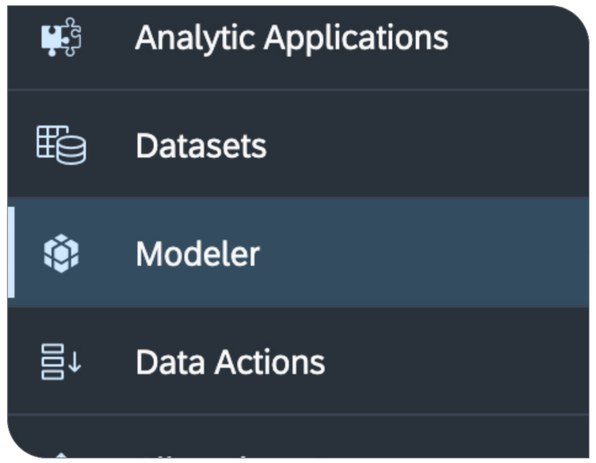
Task Guide: Data Analysis

Import the data into SAC: Sales

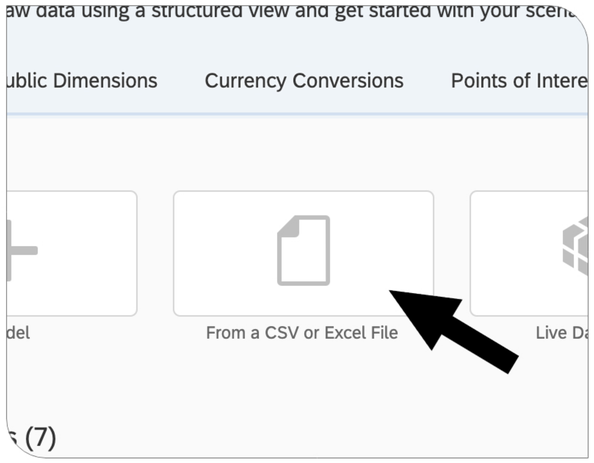
Building your model (data import & wrangling) is the lengthiest task you have.

NB!: Everything before building your story has to be done in 1 session, or you are likely going to lose your progress. Once you've created and saved your model (until page 14, incl.) you can safely take a break.

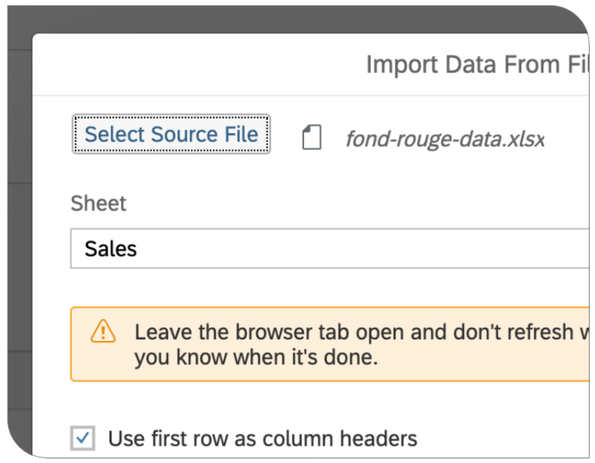
1. Go to the Modeler (in the main menu).



2. Select import from a file.



3. Upload the Sales sheet.



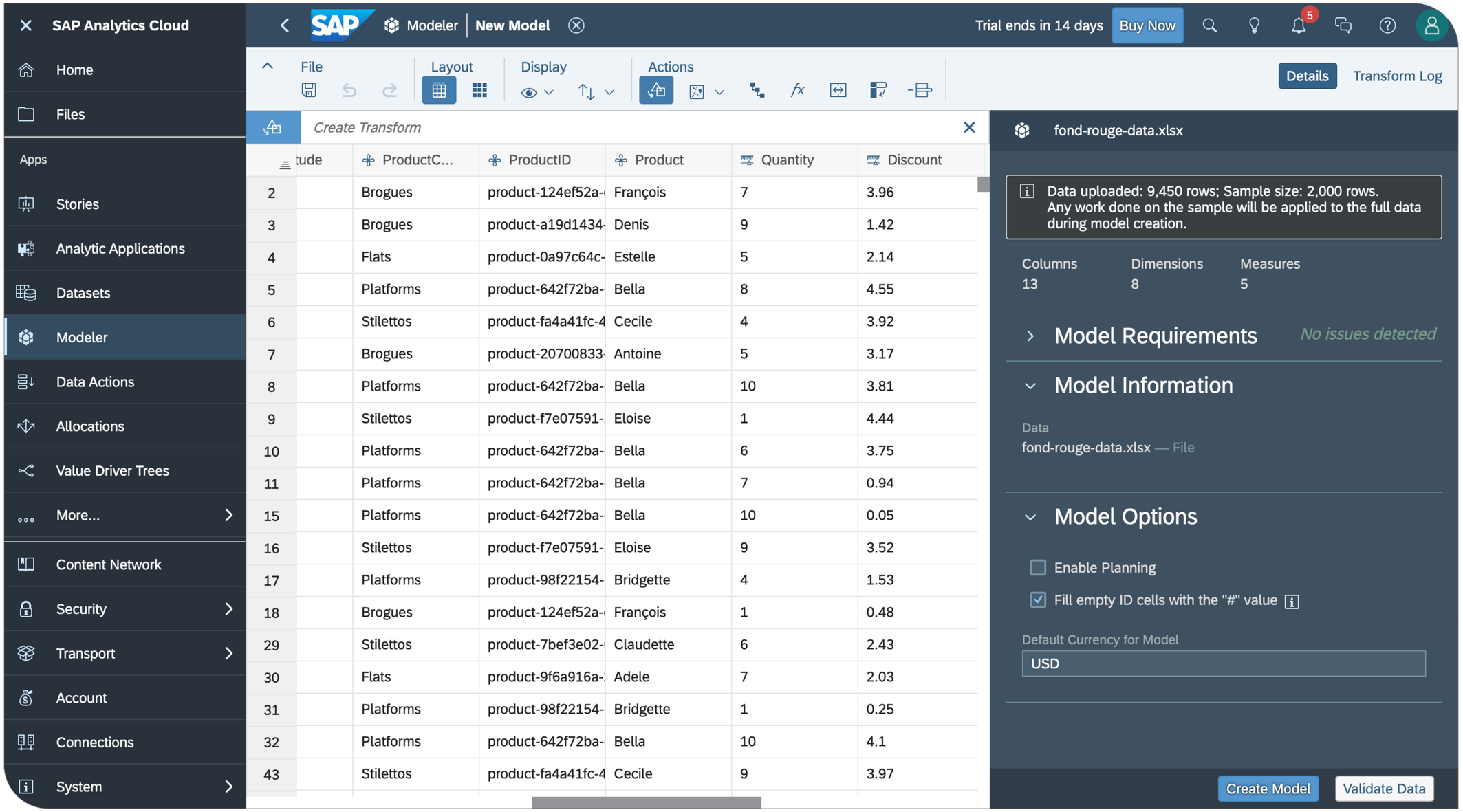
Once you click Import the data will start importing into SAC.

Sometimes it is imported as Draft Data - in this case you will see a popover show up and there you will be able to track the progress of the upload. When the upload progress completes just click the Draft Data item to open the Modeler (see next page for reference on how the Modeler looks).

More often than not - you will directly get into the Modeler (without the need to look for a Draft Data item to click on).

Import the data into SAC: Sales

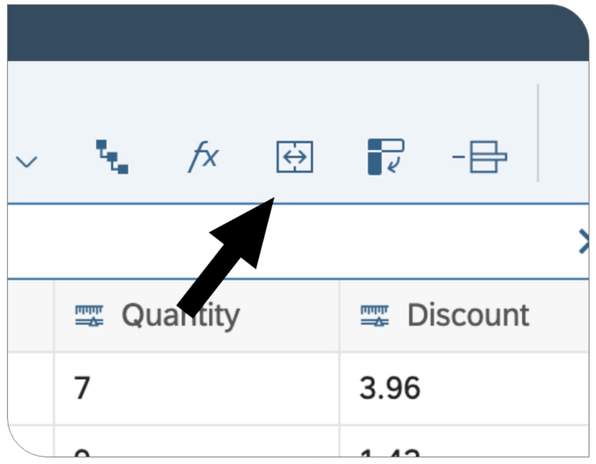
After the import of Sales is done, you should see something similar to this.



	ude	ProductC...	ProductID	Product	Quantity	Discount
2		Brogues	product-124ef52a-	François	7	3.96
3		Brogues	product-a19d1434-	Denis	9	1.42
4		Flats	product-0a97c64c-	Estelle	5	2.14
5		Platforms	product-642f72ba-	Bella	8	4.55
6		Stilettos	product-fa4a41fc-4	Cecile	4	3.92
7		Brogues	product-20700833-	Antoine	5	3.17
8		Platforms	product-642f72ba-	Bella	10	3.81
9		Stilettos	product-f7e07591-	Eloise	1	4.44
10		Platforms	product-642f72ba-	Bella	6	3.75
11		Platforms	product-642f72ba-	Bella	7	0.94
15		Platforms	product-642f72ba-	Bella	10	0.05
16		Stilettos	product-f7e07591-	Eloise	9	3.52
17		Platforms	product-98f22154-	Bridgette	4	1.53
18		Brogues	product-124ef52a-	François	1	0.48
29		Stilettos	product-7bef3e02-	Claudette	6	2.43
30		Flats	product-9f6a916a-	Adele	7	2.03
31		Platforms	product-98f22154-	Bridgette	1	0.25
32		Platforms	product-642f72ba-	Bella	10	4.1
43		Stilettos	product-fa4a41fc-4	Cecile	9	3.97

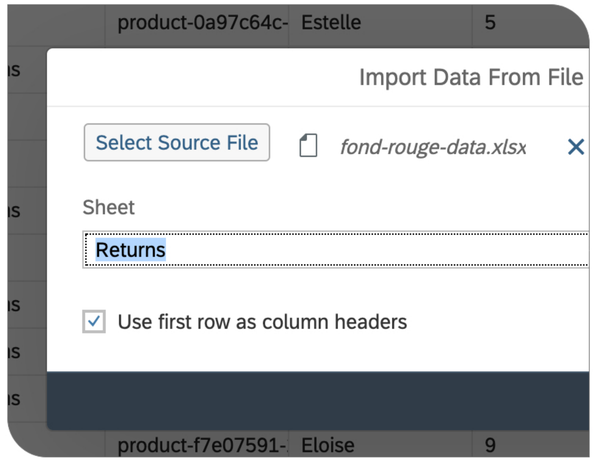
Import the data into SAC: Returns

1. Click on the Combine Data button in the toolbar.



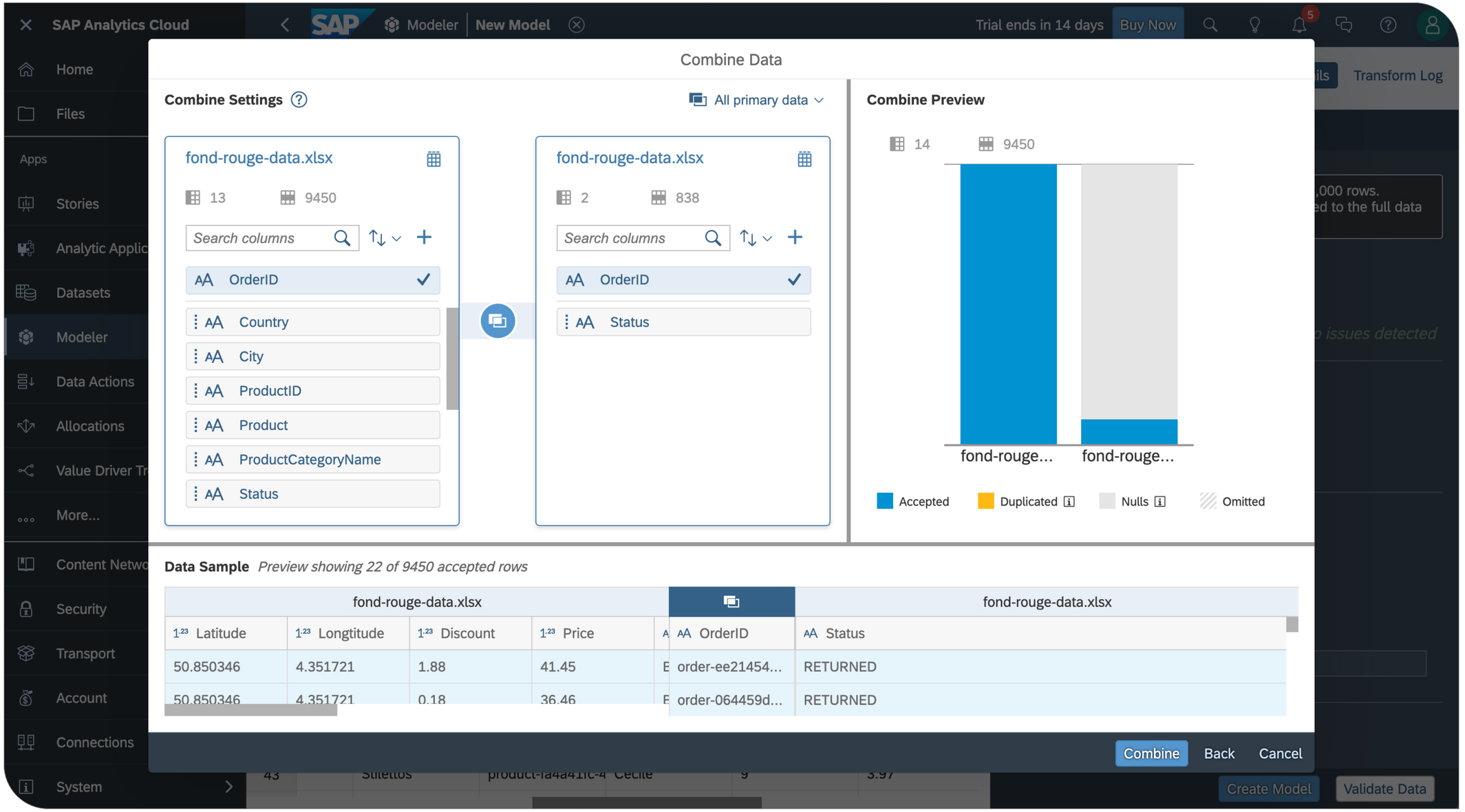
2. Select import from a file.

3. Upload the Returns sheet.



Import the data into SAC: Returns

Make sure to combine the data you've already imported (Sales) with the Returns, using the OrderID column. As you can see bellow - only the OrderID is selected on both sides. Now click the Combine button (bottom-right) to run the data combination.

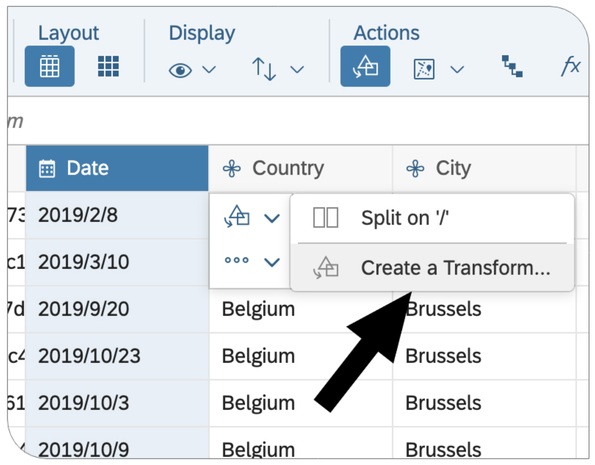


Import the data into SAC: Sentiment

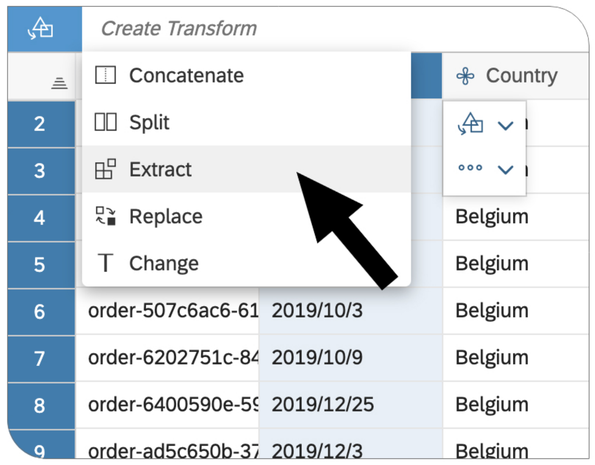
To prepare for importing the Sentiment data, we need to create a couple of new columns to be used as key columns during the data combination. We need to create a Year_Month column through a smart transform of the Date column; and Location_ID through a transform of the Country and City columns.

Let's start with Year_Month.

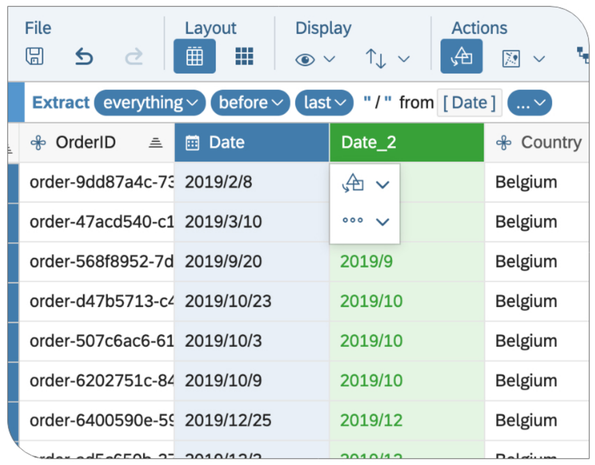
1. Select the Date column and click on "Create a Transform…" in the menu.



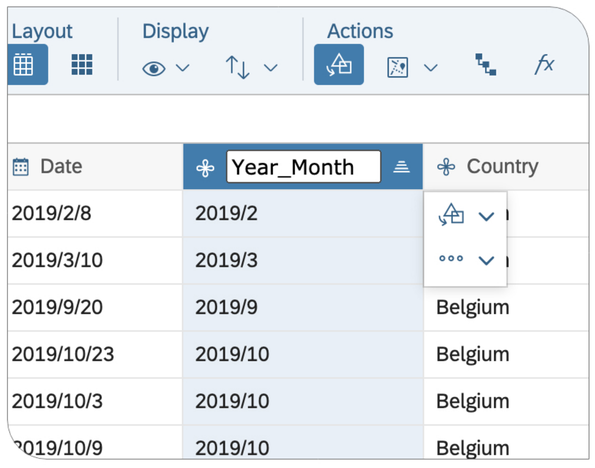
2. Select the Extract transform.



3. Extract everything before the last '/'.



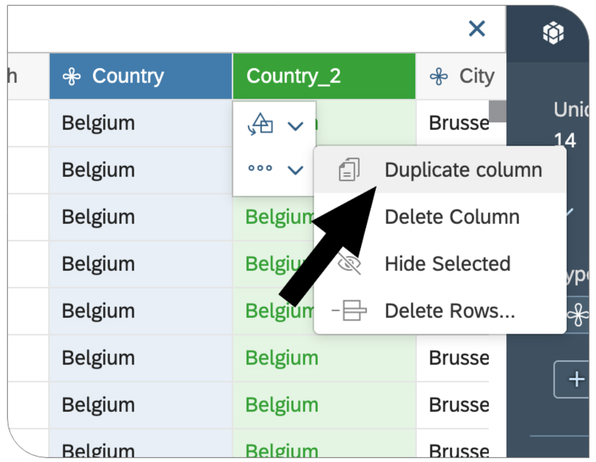
4. Rename the new column to Year_Month (Press Enter to save).



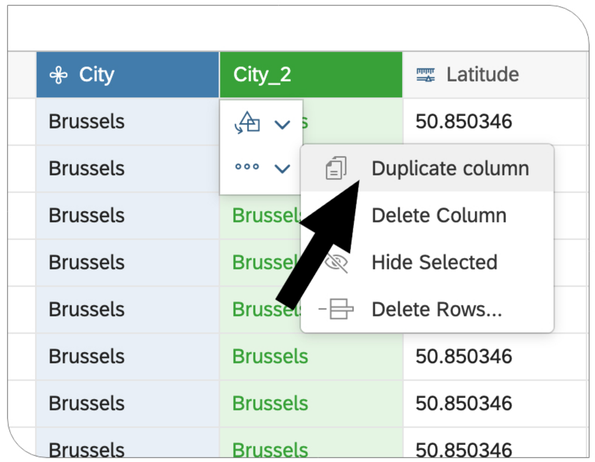
Import the data into SAC: Sentiment

In the Sentiment data we have a Location_ID column with data looking like this: “Belgium, Brussels”.
Here are the steps to create a Location_ID column in the data we’ve imported so far, so we can use it to later match the Sentiment’s Location_ID.

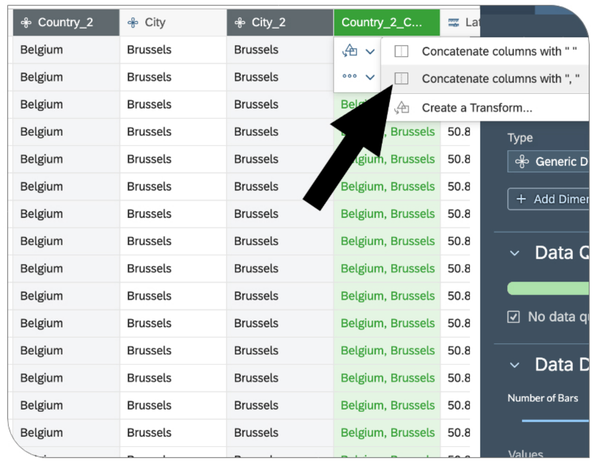
1. Duplicate Country.



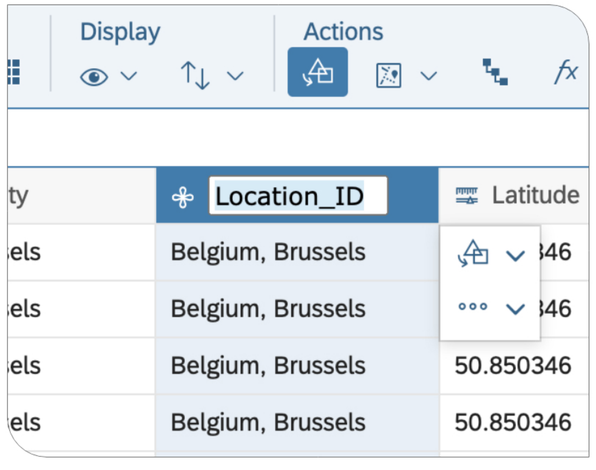
2. Duplicate City.



3. Select Country_2 -> add City_2 to the selection (holding Cmd/Ctrl) -> “Concatenate … with ‘, ’”.

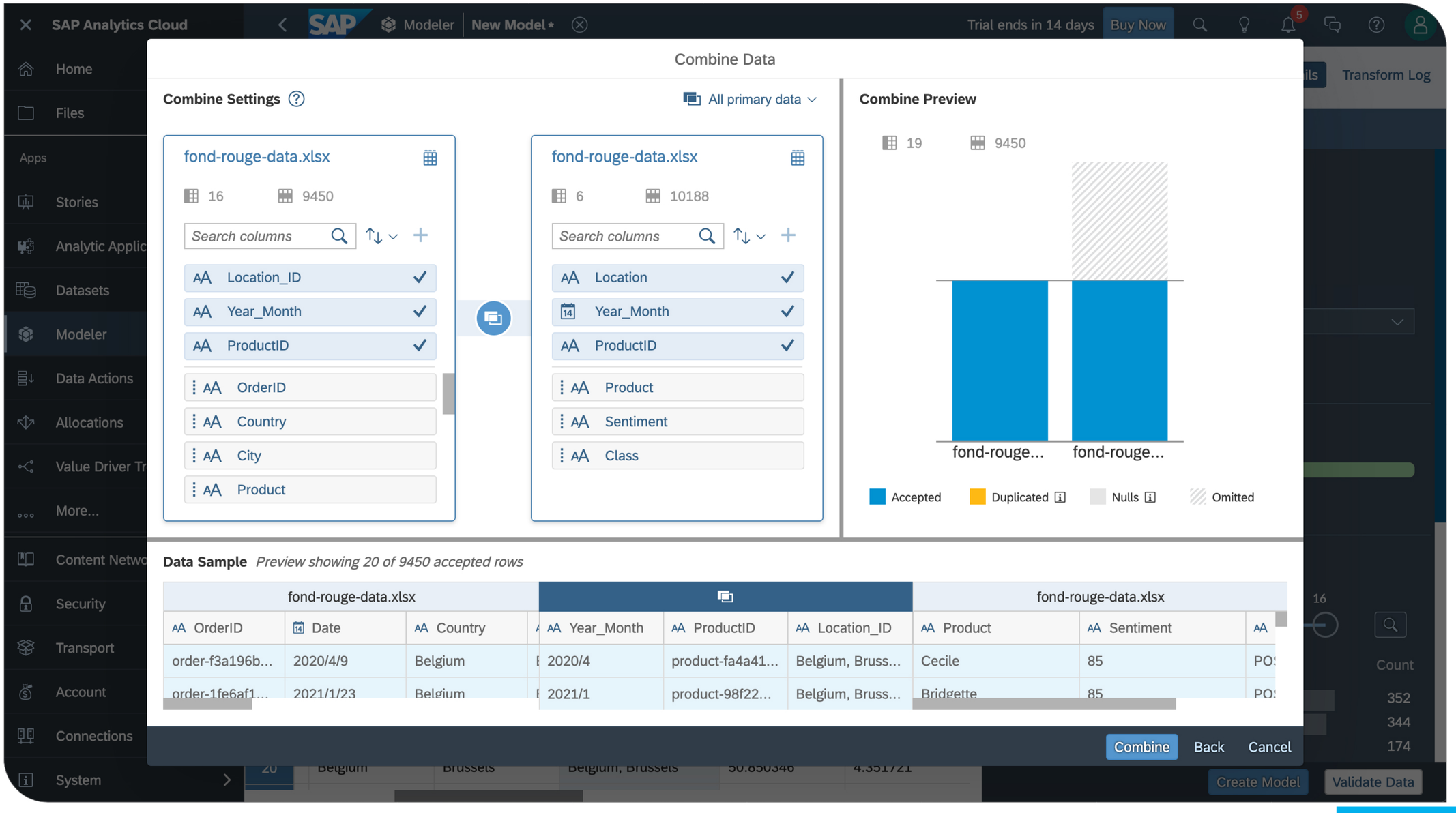


4. Rename the new column to Location_ID (Press Enter to save).



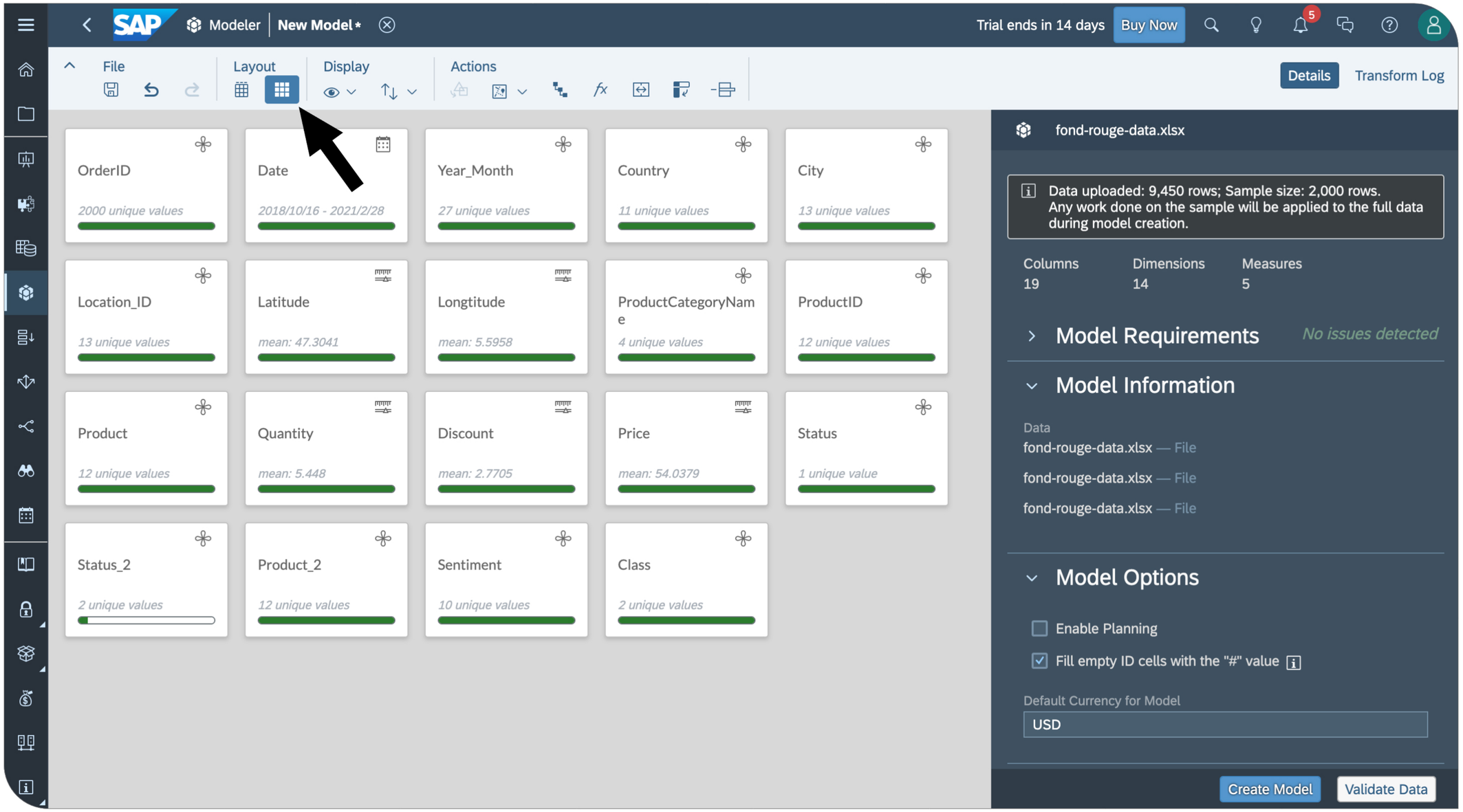
Import the data into SAC: Sentiment

Having all the needed identification columns - it's time to import the final sheet of data - Sentiment. Here is how to map the data. **Make sure the selected column names and their order is the same on both sides!**



Data Wrangling

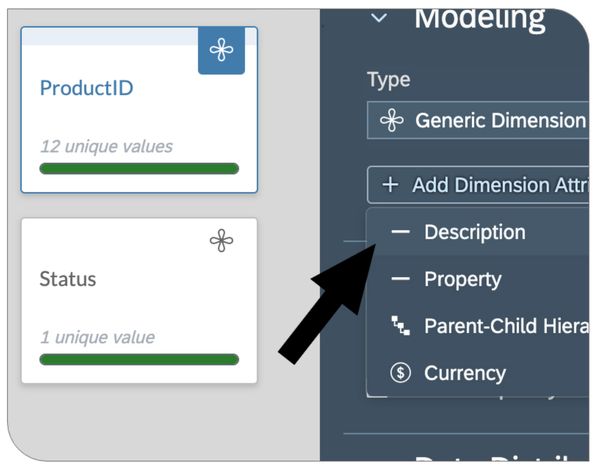
We have all the data we need. Now it’s time to clean it up. Let’s first look at the data in the Card View.



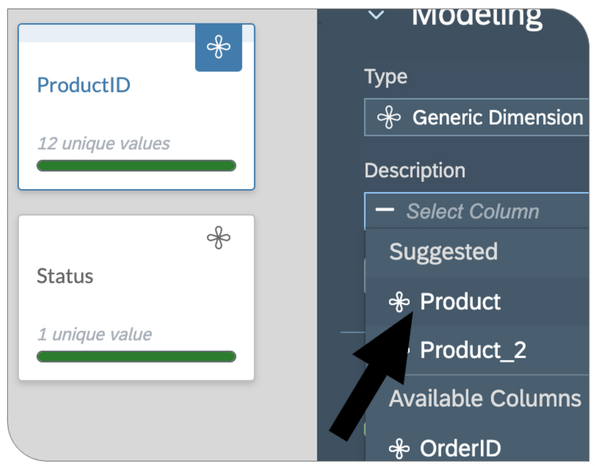
Data Wrangling: ID Descriptions

Let’s start by assigning Description properties to ID columns that have one. In Fond Rouge’s data, only ProductID has a description.

1. Select the ProductID card and click Add Dimension Attributes -> Description.



2. In the Description dropdown select the Product column.



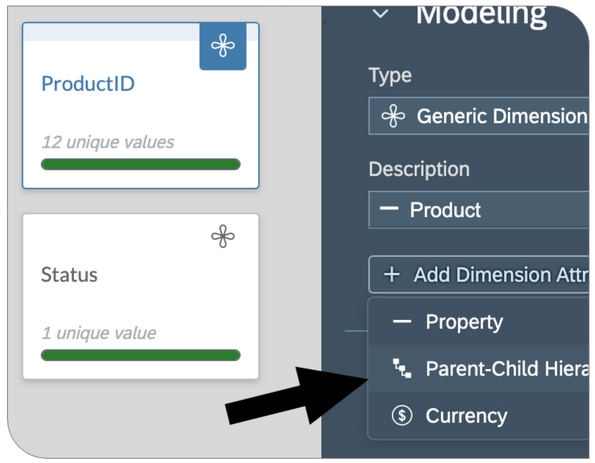
Now ProductID is connected to Product via description attribute.

Data Wrangling: Hierarchies

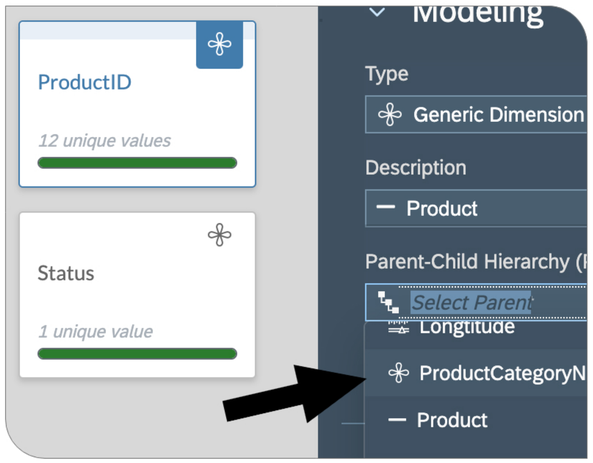
We have 2 pairs of columns that can be connected in a parent-child hierarchy:
- ProductCategoryName -> ProductID
- Country -> City

Here is how to express that relationship between the first pair and you can do the 2nd pair on your own.

1. Select ProductID and click Add Dimension Attributes -> Parent-Child Hierarchy.



2. Choose ProductCategoryName from the dropdown.



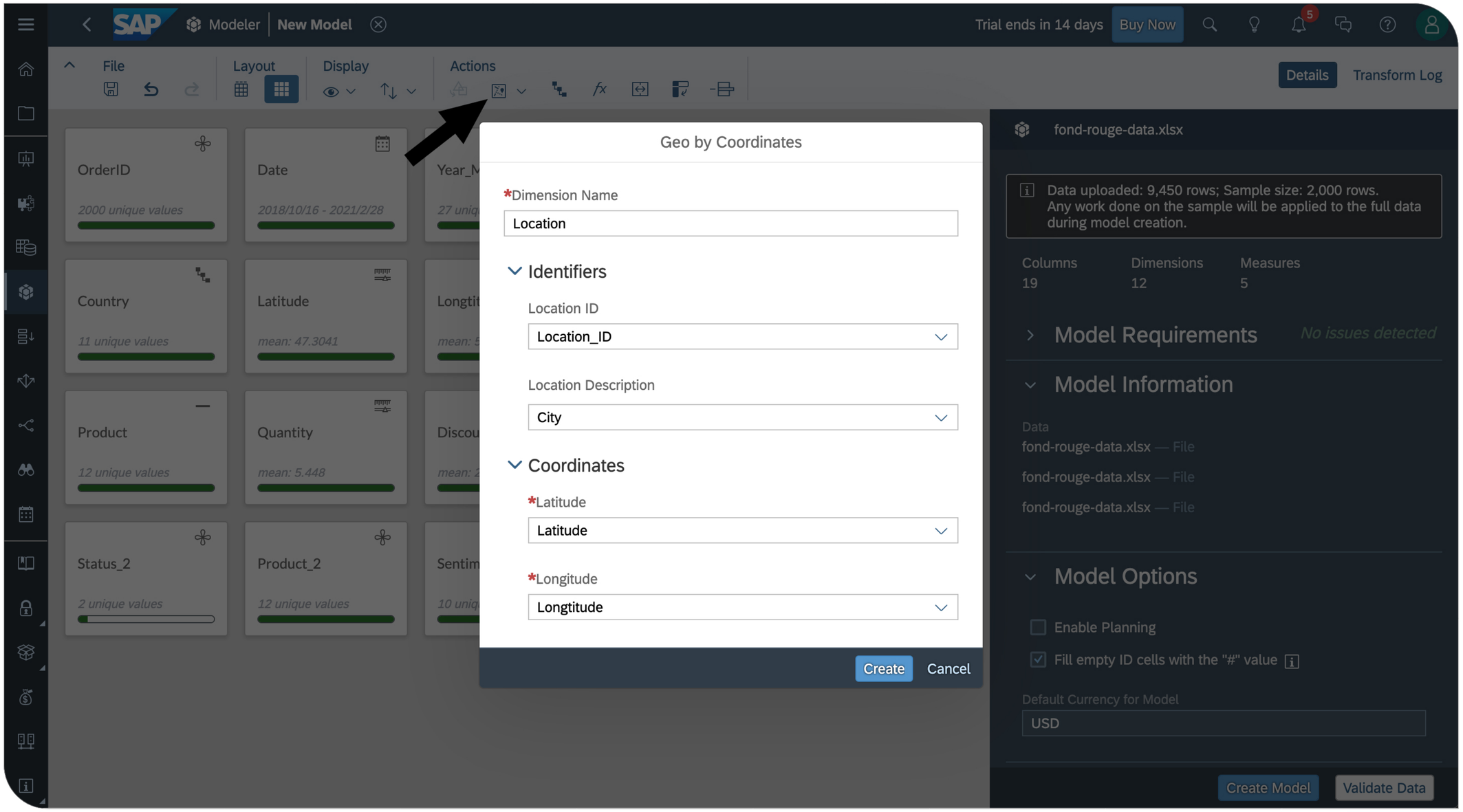
Repeat the same process for the Country -> City.

NB!: expressing the connection is done on the child column (i.e. City points to the Parent - the Country).

Data Wrangling: Geo Enrichment

Having the Latitude and Longitude columns allows us to geo enrich our model and later be able to plot data on a map.

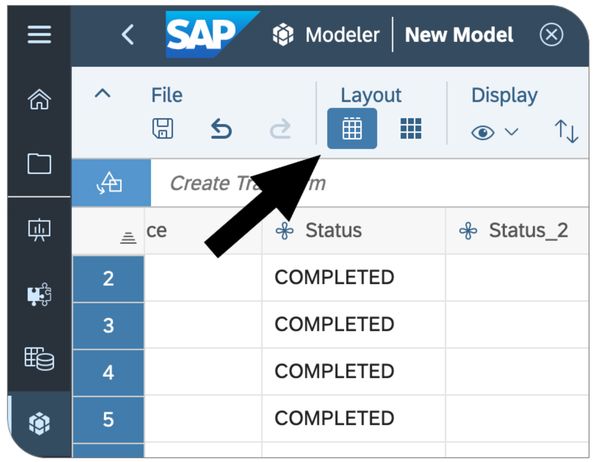
From the Actions toolbar - select Geo Enrich By -> Coordinates and fill out the form like this and click Create.



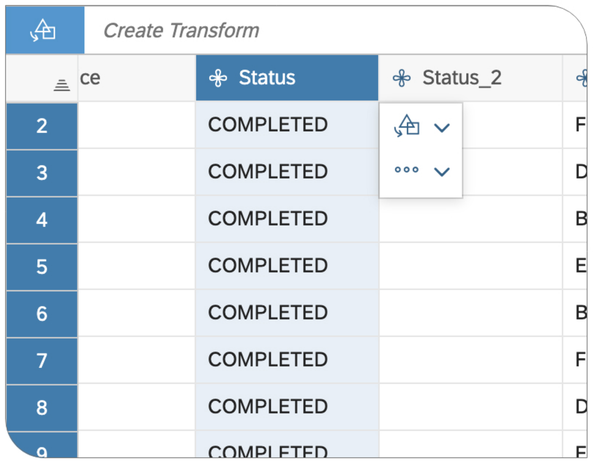
Data Wrangling: Merge Columns

At this point, there should be 2 Status columns (Status and Status_2). We have to find a way to merge the 2nd into the first, so only the returned orders change their Status to RETURNED. For this job we will use the **Replace Smart Transform** with a **Where clause**.

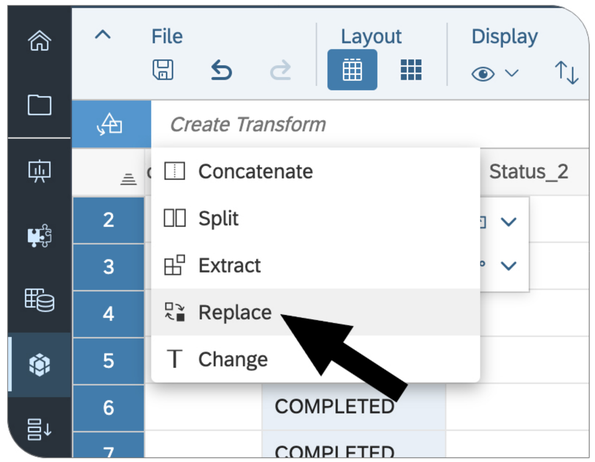
1. Go back to Table View.



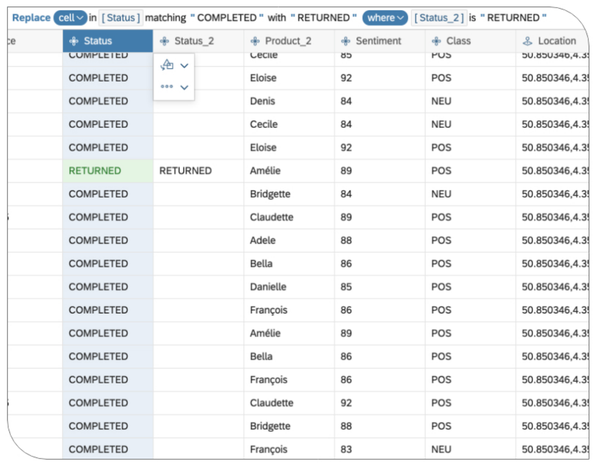
2. Select the main Status column.



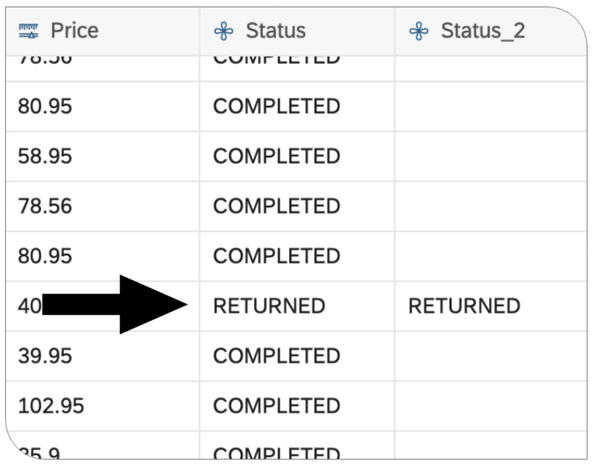
3. From the menu: Create a Transform -> Replace.



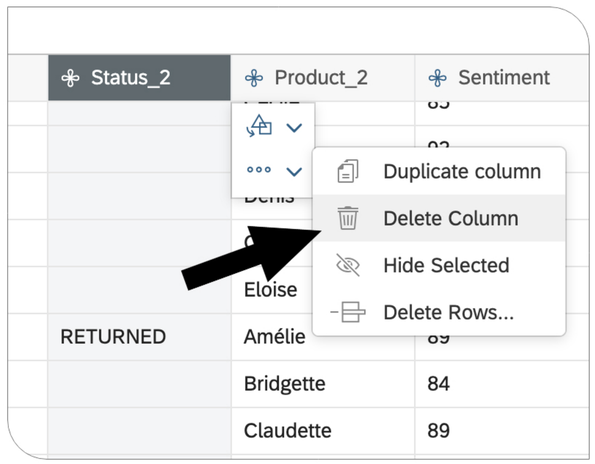
4. Write the following query (Zoom-in to read the top line). If the full line isn't visible - close the main menu.



5. Press Enter/Return to execute and afterwards ensure only the returned orders changed Status.



6. Delete the Status_2 column.

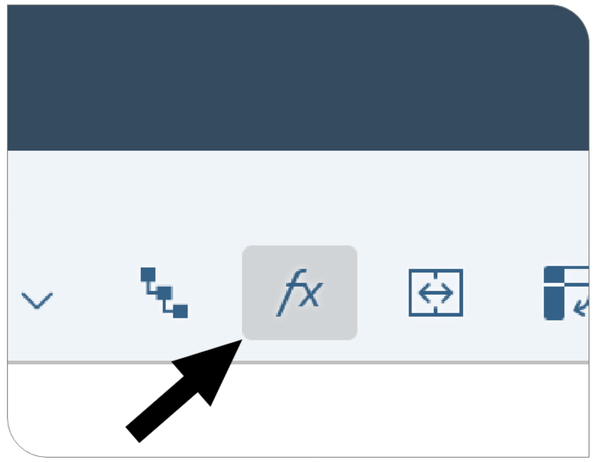


The other column you would want to delete is Product_2.

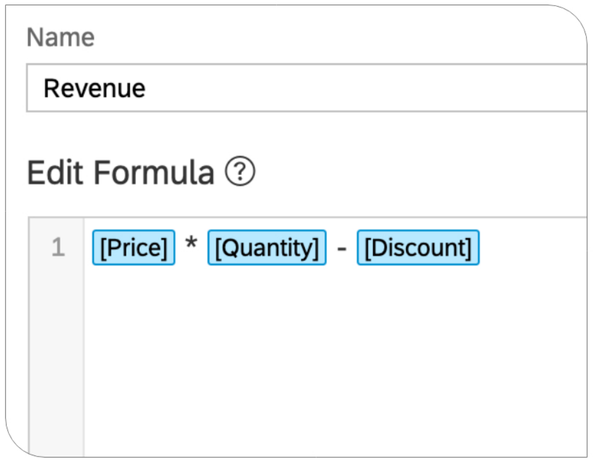
Data Wrangling: Computed Columns

We need to create a couple of columns that get derived from the rest of the data. Namely, we need to add Revenue & Refund. As you type the name of columns you get suggestions for autocompletion.

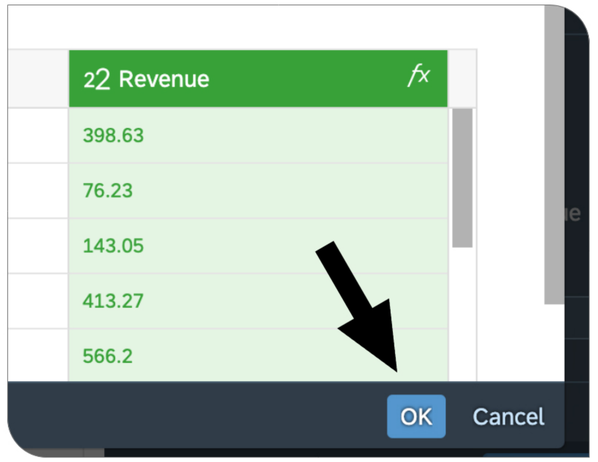
1. [Revenue] Click on the Calculated Columns tool.



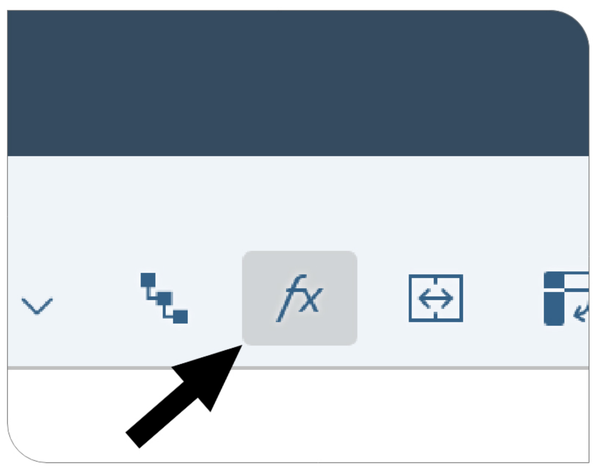
2. Fill the name of column and write down this formula.



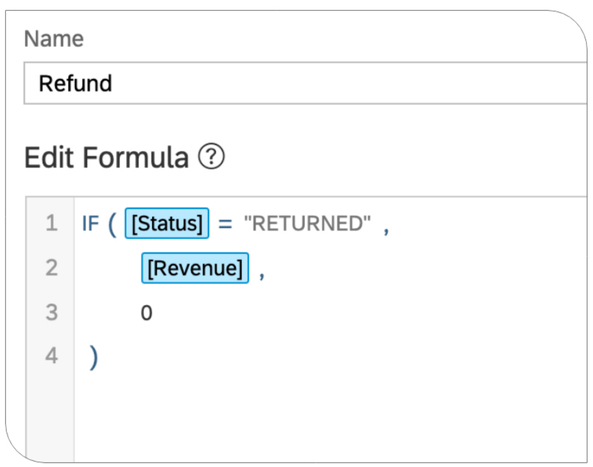
3. Press OK to save the column.



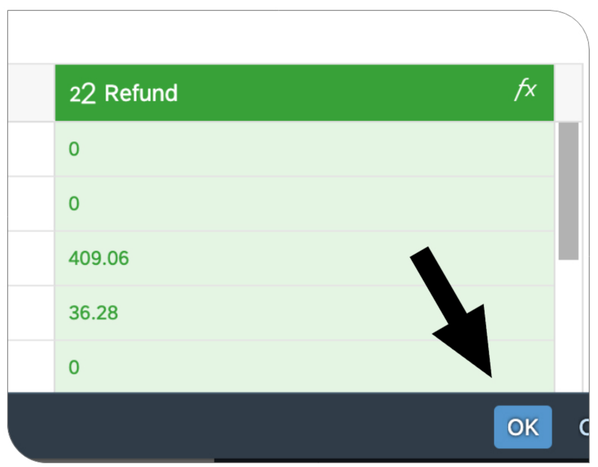
4. [Refund] Click again on the Calculated Columns tool.



5. Fill the name of column and write down this formula.



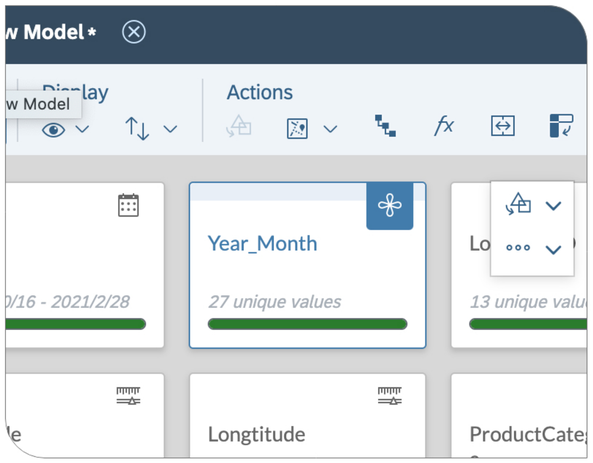
6. Press OK to save the column.



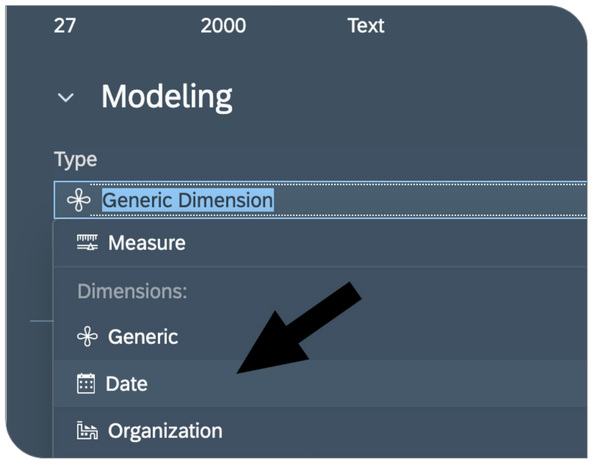
Data Wrangling: Column Types

Most of the time SAC will do a good-enough job at guessing the right types of your columns, but sometimes you will have to manually assign whether something is a measure or a dimension and what's the exact type of it. For example - the Year_Month is actually a Date and not a Generic Dimension. Here is how to change it. After you go back to card view:

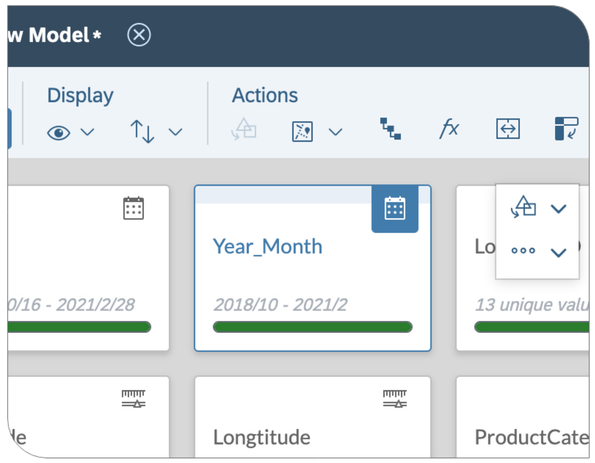
1. Select the Year_Month card.



2. In the Designer popover select the Date Type.



3. The icon of the card changes while the data quality bar stays green.



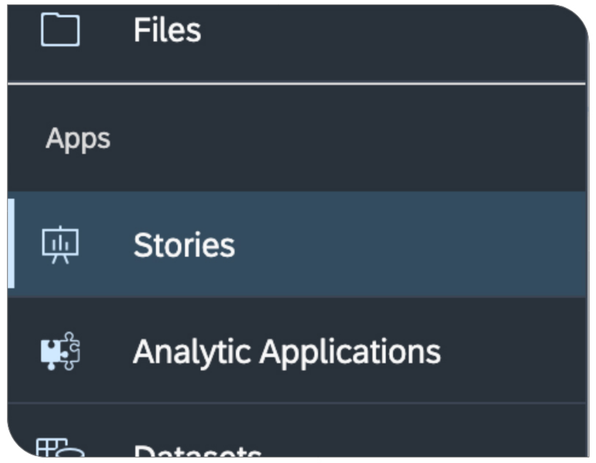
Here are the other columns needing a change of type:
- Latitude & Longtitude: Should be (Generic) dimensions not measures.
- Sentiment: Should be a measure.

Now you are ready to create the Model! Press Create Model (there may be a 2nd confirmation you need to give for the creation of the model) and if there aren't any issues with your data - you will choose a location for your new model and you are done with this long first phase! You can name It something as simple as "Model". Make sure to remember where in your Files you've saved it, as we will need to later find it and link it to the SAC Story we are going to build.

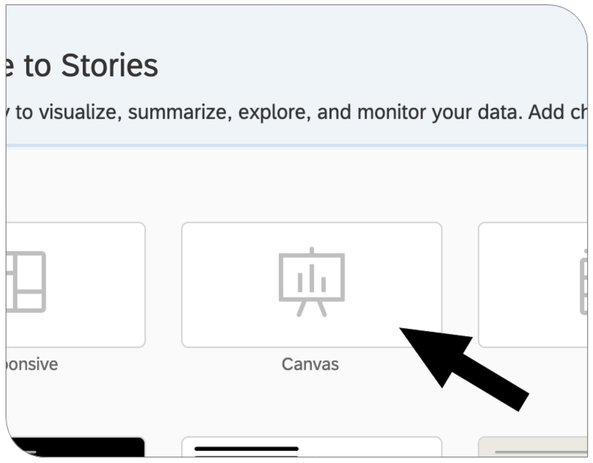
Building a Story

Let’s create a story for Fond Rouge’s data analysis.

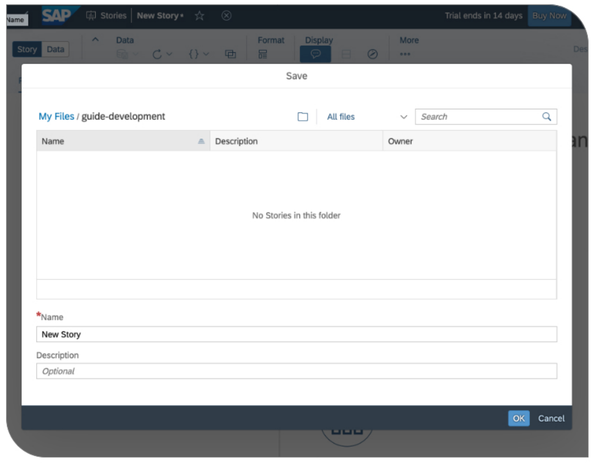
1. Go to Stories in the main menu of SAC (far left).



2. Select the Canvas template.



3. Ctrl/Cmd + S to save the story. Place the story in the same directory you put your model in.

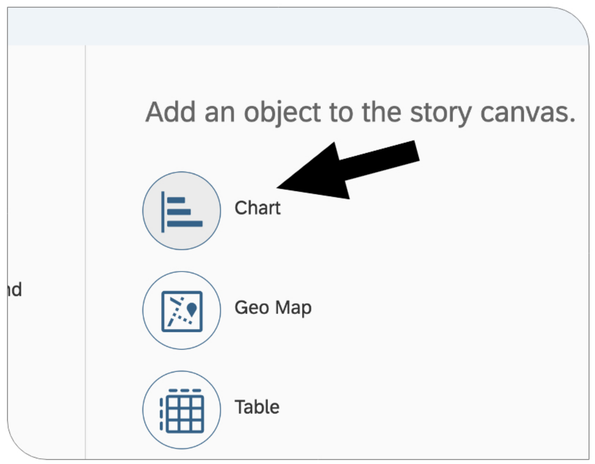


When choosing a location for the story, the model will not be visible in the dialog because it’s from a different type - it’s Model and not a Story. Now we can add our first chart and link the story to the model we created earlier.

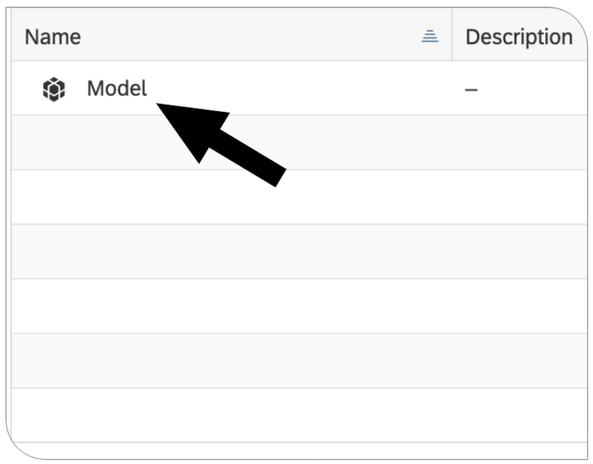
Building a Story: Chart [Column & Line]

The first chart we will create is: **[Column & Line] Revenue & Refund per country.**

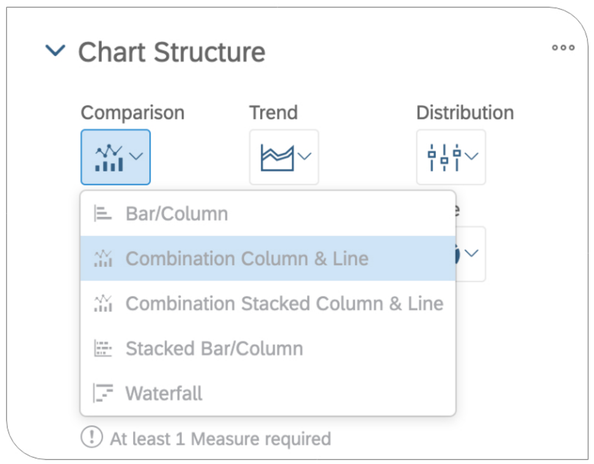
1. Add chart to the canvas.



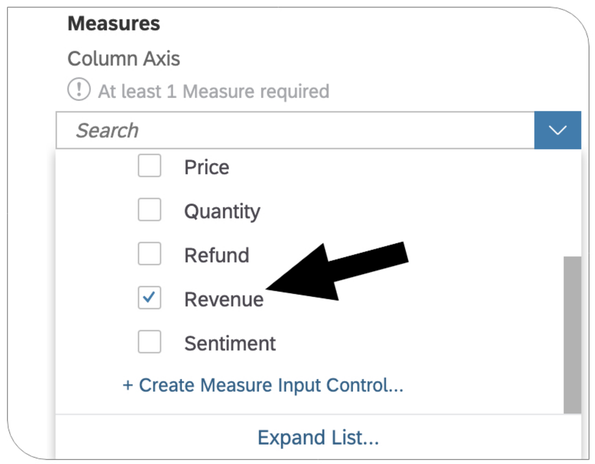
2. Browse-to & select Model.



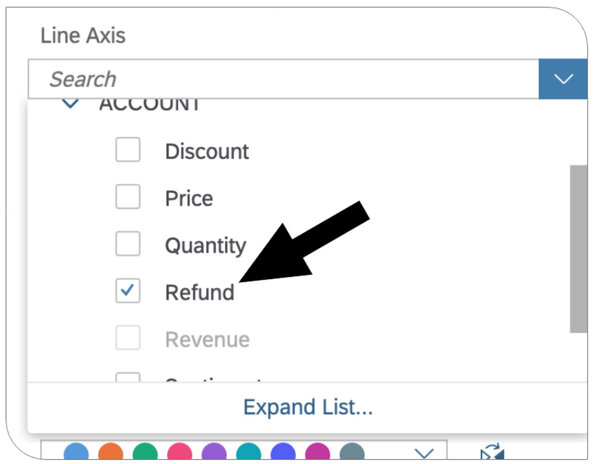
3. Select the Column & Line Comparison.



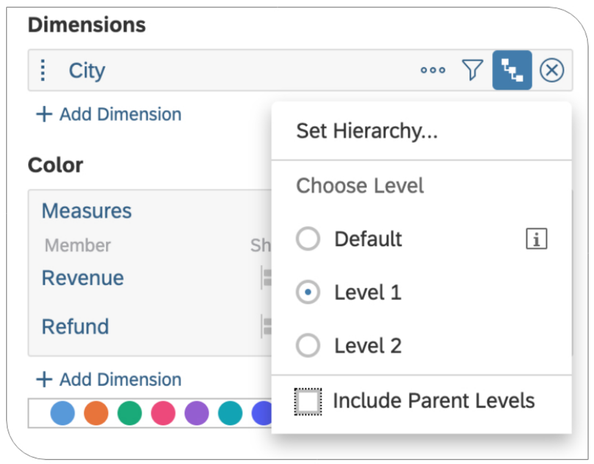
4. Select Revenue as a Measure for the bars.



5. Select Refund as a Measure for the line.

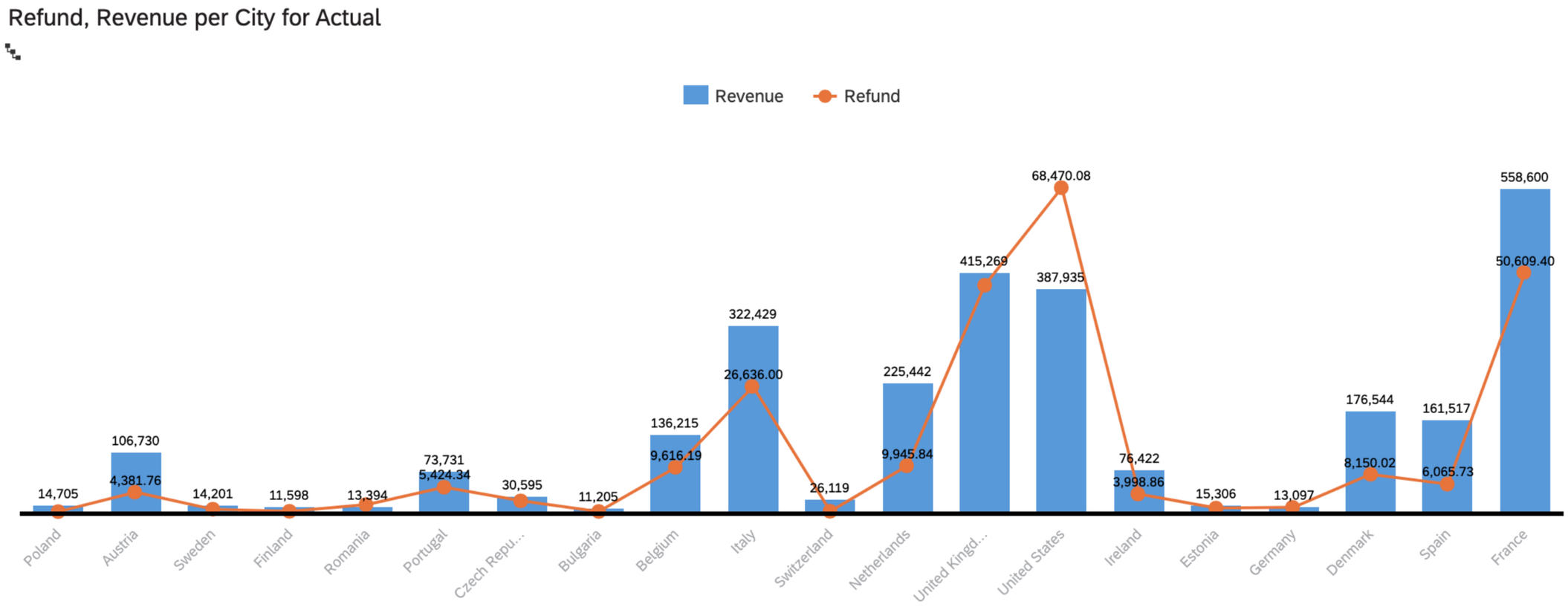


6. Select City as Dimension and ensure Level 1 (Country) is selected.



Building a Story: Chart [Column & Line]

The (unsorted) chart can look different in terms of country order, but you should see the same data for each country:

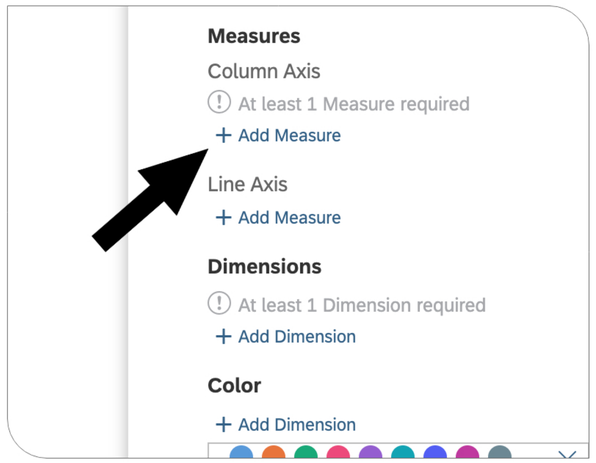


Remember to often press the Save button in the toolbar (or Ctrl/Cmd+S) to preserve your progress.

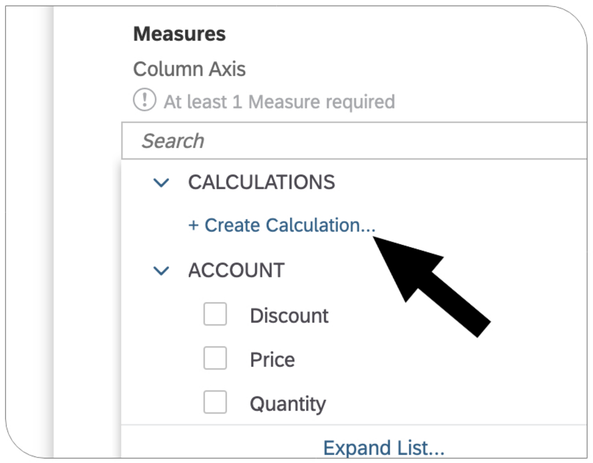
Building a Story: Calculated Measures

There are 2 calculated measures you need to create that will be used for some of the charts to be built - the **Refund/Revenue ratio** and **Avg. Sentiment**. Let's create both in a single chart - **[Column & Line] Refund/Revenue & Avg Sentiment per US city**. From the Toolbar -> Insert -> Chart (If it's not visible - find it under the last toolbar item named "More") -> select the Column & Line chart type once again.

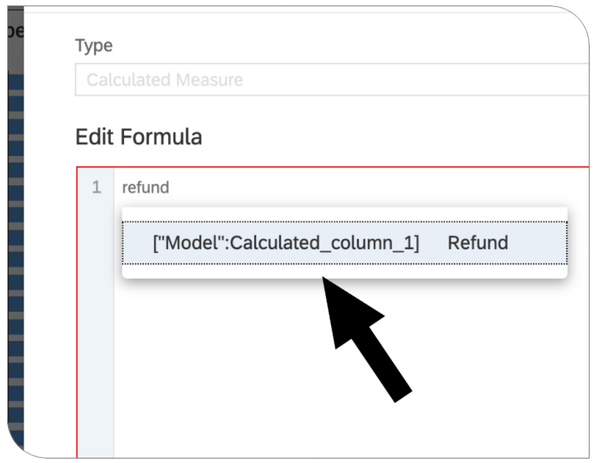
1. [Refund/Revenue] Add Column Axis Measure.



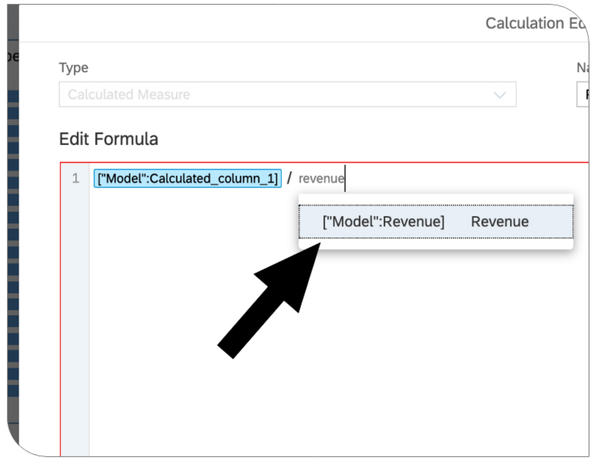
2. Click "+ Create Calculation".



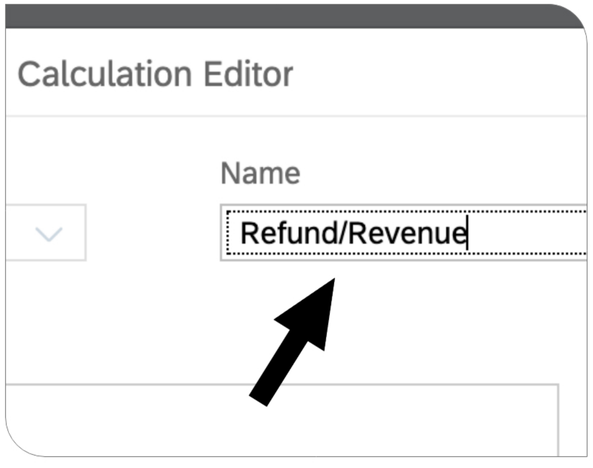
3. Type "refund" in the text area -> select the suggested item in the autocomplete menu that shows up.



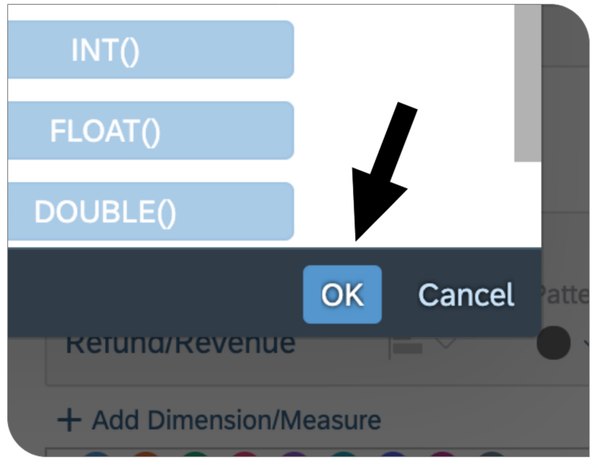
4. Type "/ revenue" and select the suggested item.



5. Name the Measure "Refund/ Revenue".



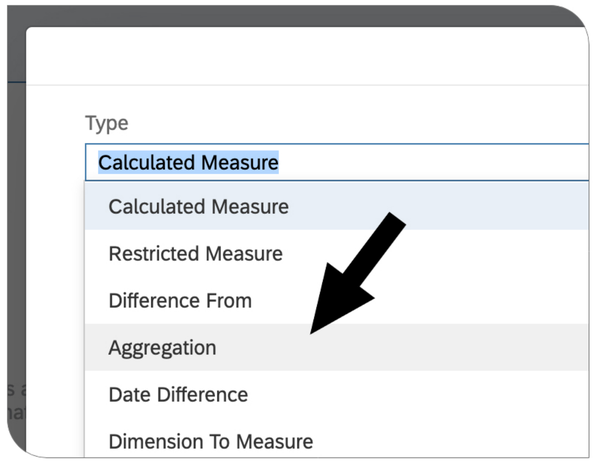
6. Press OK to save the measure.



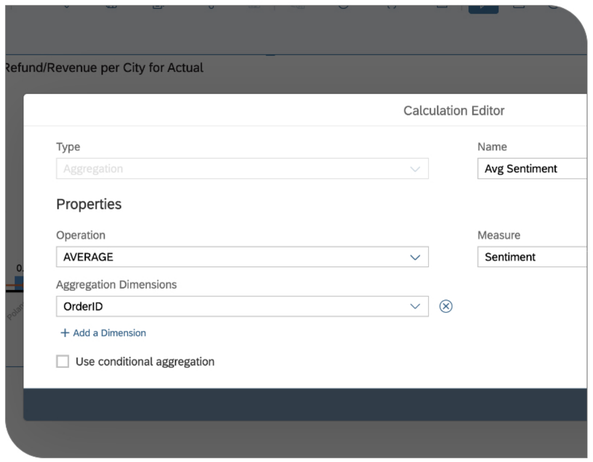
Building a Story: Calculated Measures

Now let's add the **Avg. Sentiment** and filter by location (US Cities only).

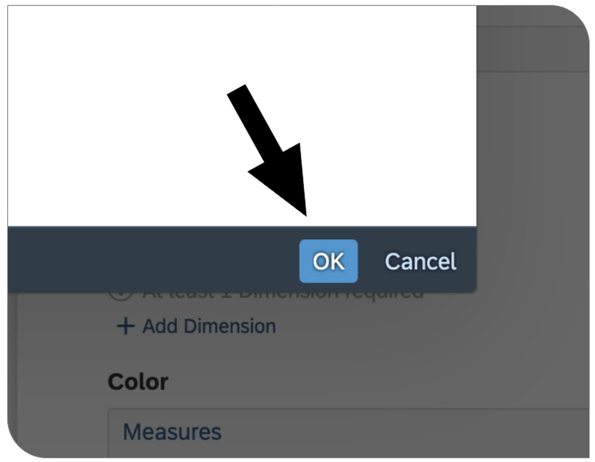
1. [Avg. Sentiment] Add Line Axis Measure -> + Create Calculation. Change the Type to Aggregation.



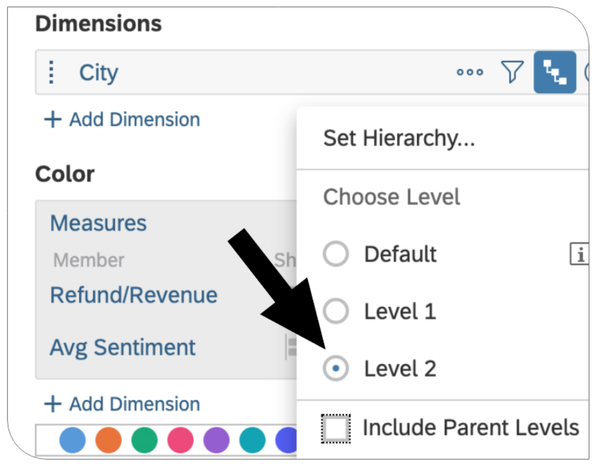
2. Fill out the form as follows. (Zoom-in to read all values)



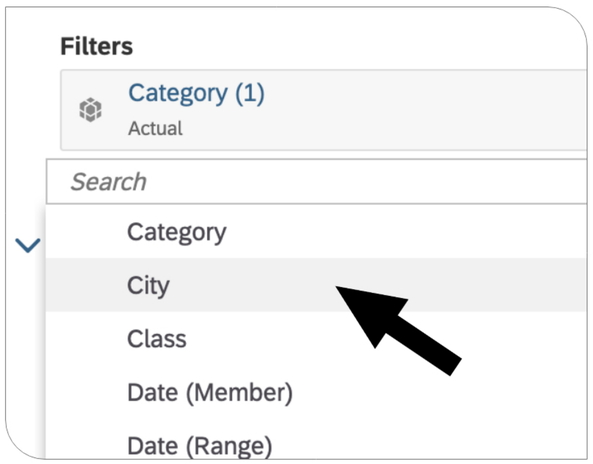
3. Press OK to Save the Measure.



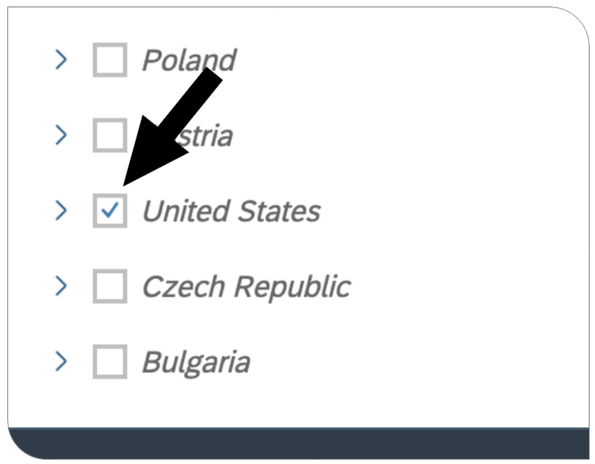
4. Set City as Dimension and select Level 2 of the hierarchy to be used.



5. Add new Filter (towards the bottom of the panel) and select City from the dropdown.

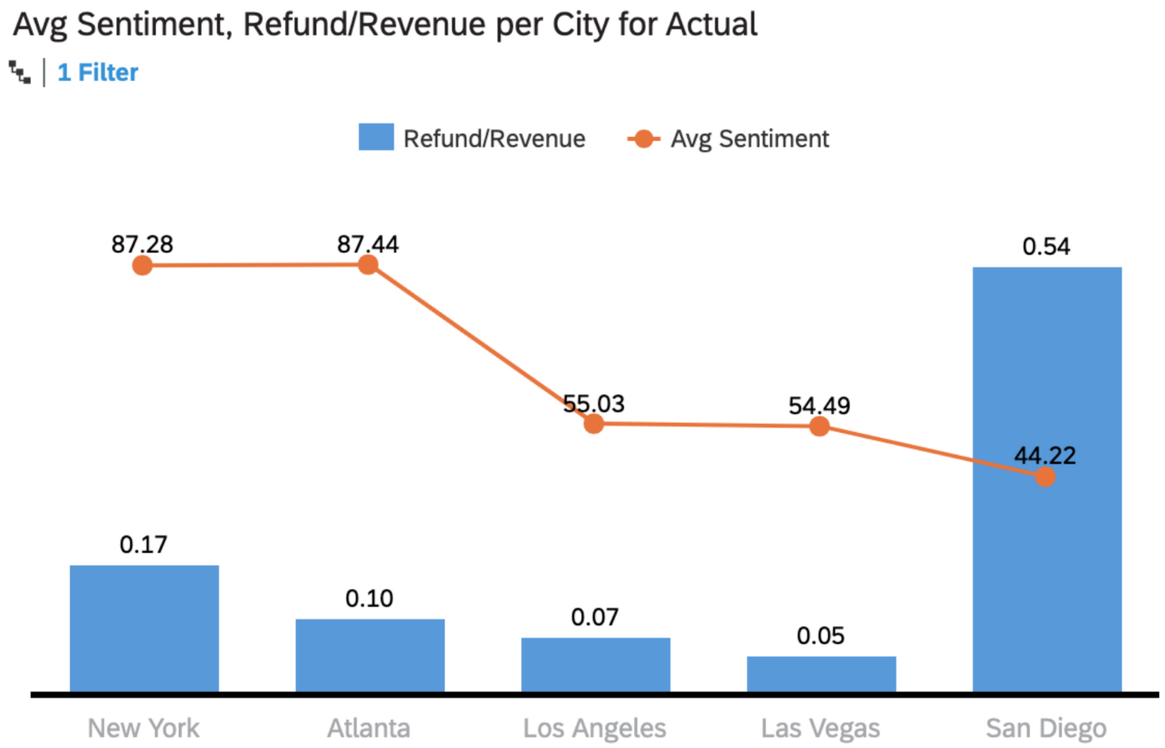


6. Check the United States in the list and click the OK button (bottom right).



Building a Story: Aggregates

The resulting chart should look similar to this.



Now you are ready to build the rest of the charts, independently from this guide. Have fun!