

Project 1

Sets for Top N and Others

If you collect large sets of data that you want to visualize, you might find that limiting the amount of information displayed to an important subset of records helps you work with and answer questions about the data more effectively.

This article describes how to create an interactive view that separates your customers into two dynamic groups:

- The top N customers
- All other customers

The view includes a control that your users can adjust to change the number of customers included in the top customers group. When they change the number, the view updates accordingly.

Sets and supported data sources

The method described in this article for creating a view of the top customers uses the In/Out functionality of sets.

Sets were introduced with Tableau Desktop version 8.0.

For live connections, the In/Out functionality requires a relational or multidimensional data source.

If you use a file-based data source, such as a Microsoft Excel workbook or text file, you can take an extract with which you can create sets.

Step 1: Create the parameter

1. In Tableau Desktop, open a new workbook and connect to the **Sample-Superstore** data source.
2. Open a new worksheet.
3. In the Data pane, click the drop-down arrow in the upper right corner and select **Create Parameter**.
4. In the Create Parameter dialog box, do the following:
 - In the **Name** text box, type **Top Customers 2**.
 - For **Data type**, select **Integer**.
 - For **Current value**, type **5**.
 - For **Allowable values**, click **Range**.
 - Under Range of values, do the following:
 - Click **Minimum** and type **5**.
 - Click **Maximum** and type **20**.
 - Click **Step size** and type **5**.


This parameter will be used, in combination with the top N set you will create in the next step, to quickly adjust the top N value in the view.

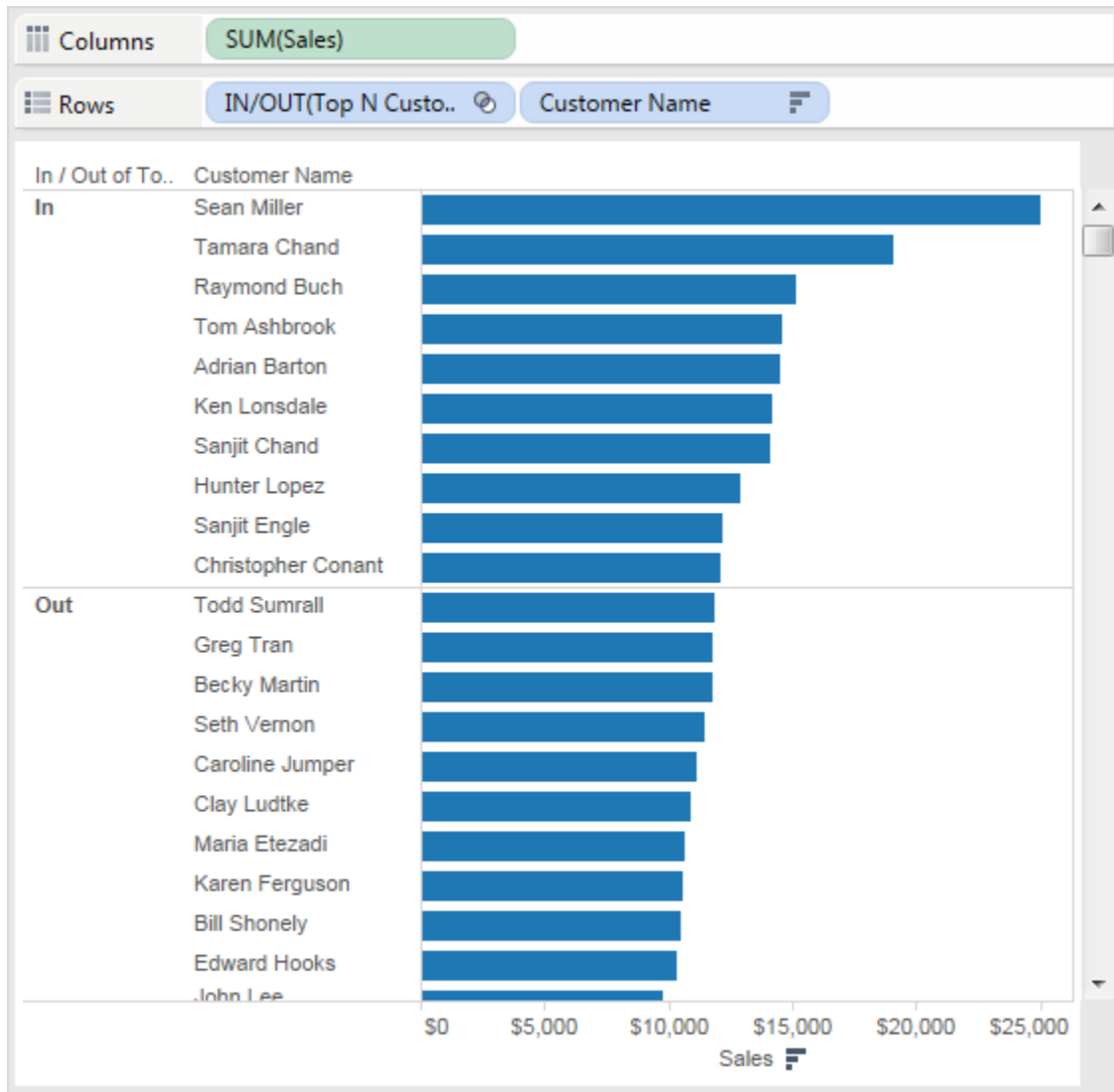
Step 2: Create the top N customers set

1. In the Data pane, right-click the Customer Name dimension and select **Create > Set**.
2. In the Create Set dialog box that opens, do the following:
 - In the Name text box, type **Top N Customers by Sales**.
 - Click the **Top** tab.
 - Select **By Field**.
 - From the field drop-down list (Category), select **Sales**.
 - From the aggregation drop-down list, select **Sum**.
 - When finished, click **OK**.

Step 3: Set up the view

1. From the Data pane, under Sets, drag **Top N Customers by Sales** to the Rows shelf.

2. Drag the Customer Name dimension to the Rows shelf, positioning it to the right of the set.
3. Drag the Sales measure to the Columns shelf.
4. On the toolbar, click the Sort Descending button  to make sure that the set is working.



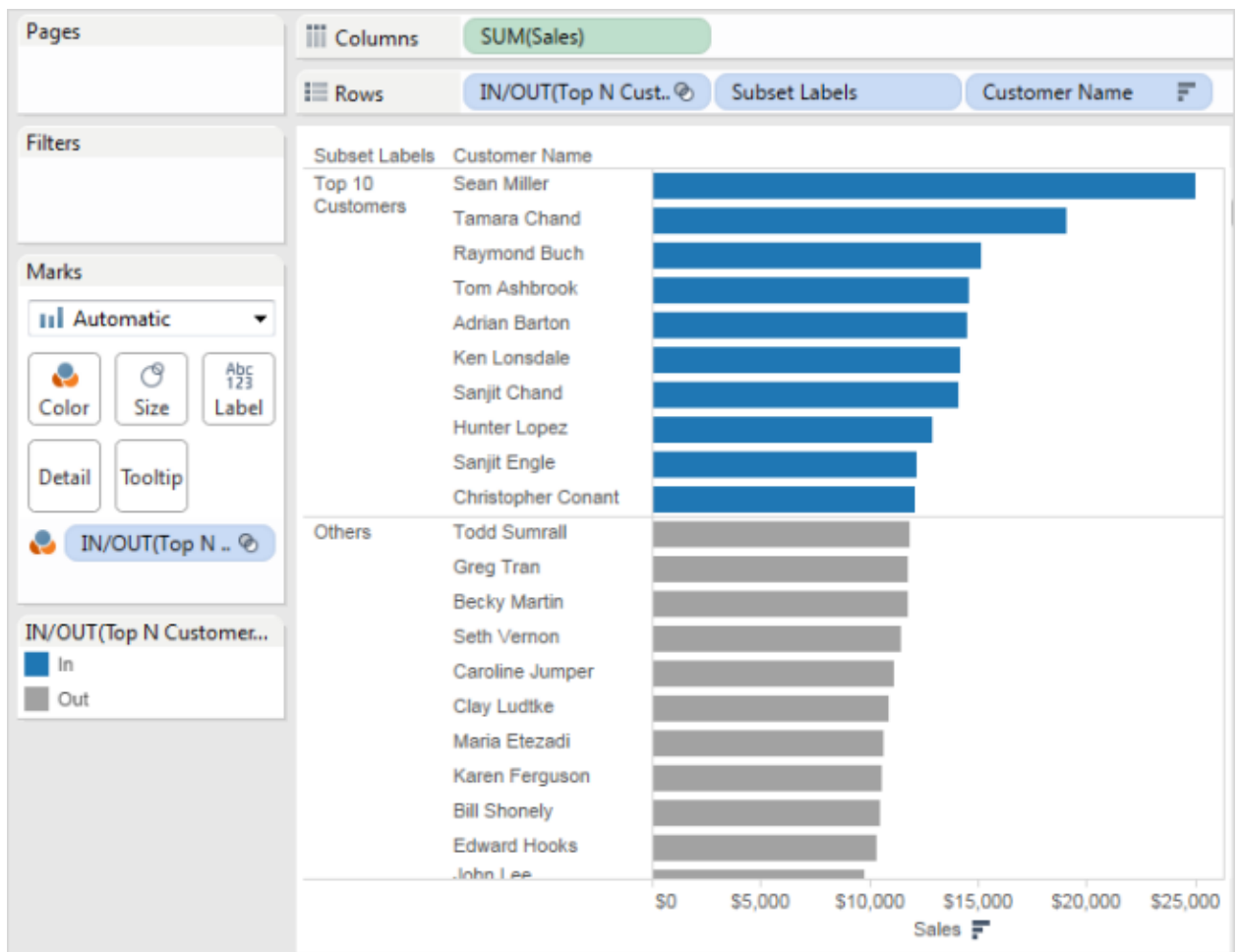
5. In the Data pane, right-click the Top N Customers by Sales set, and then click Create Calculated Field.
6. In the Calculated Field dialog box that opens, complete the following steps:
 - In the Name text box, type Subset Labels.
 - In the Formula text box, type the following formula to create dynamic labels for the customers in the set:

```

IF [Top N Customers by Sales]
THEN "Top " + str([Top Customers 2]) + " Customers"
ELSE "Others"
END

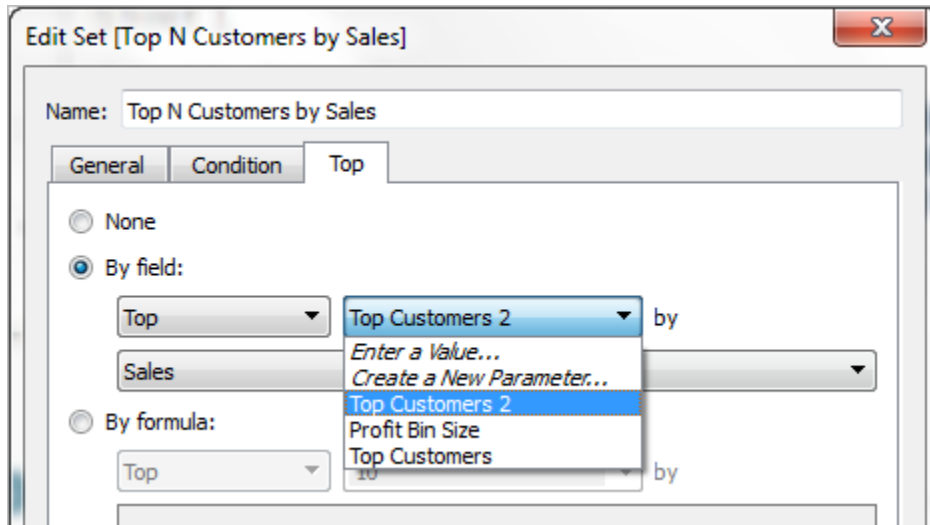
```

- When finished, click OK.
 - 7. From the Data pane, drag Subset Labels to the Rows shelf, placing it between the Top N set and the Customer Name dimension.
 - 8. On the Rows shelf, right-click the IN/OUT(Top N Customers by Sales) set, and then clear Show Header.
 - 9. This hides the In/Out labels while retaining the sort order so that your top N subset always appears at the top of the view.
- From the Data pane, drag the Top N Customers by Sales set to Color on the Marks card.



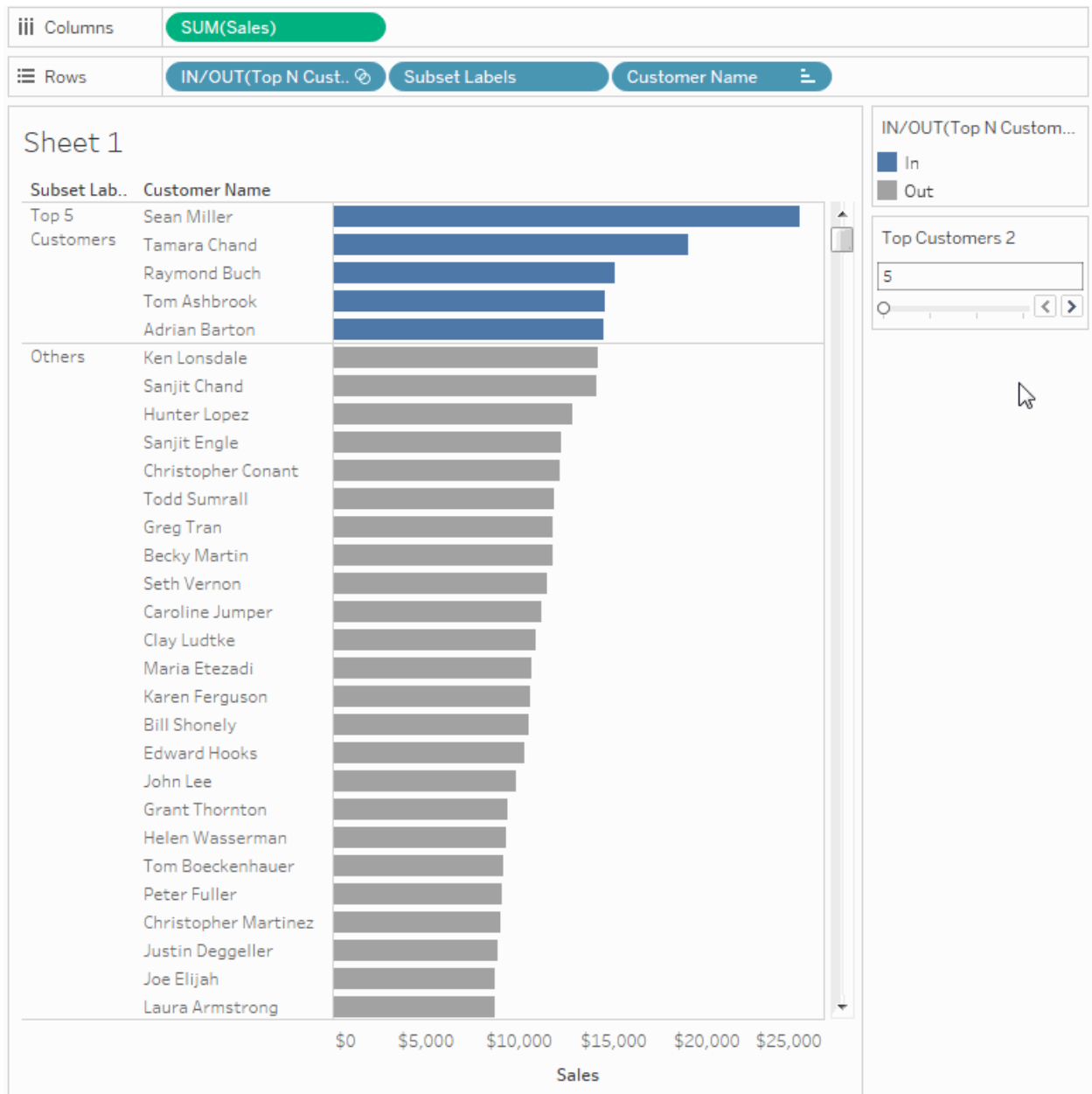
Step 4: Combine the Top N set with a dynamic parameter

1. In the Data pane, right-click Top N Customers by Sales, and then select Edit Set.
2. In the Edit Set dialog box, do the following:
 - Select the Top tab.
 - Click the value drop-down menu, and select the Top Customers 2 parameter.
 - Click OK.



3. This links the Top N Customers by Sales set to the Top Customers 2 dynamic parameter, instead of to a static list of 10. This parameter will be used in combination with the Top N Customers by Sales set, to adjust the top N value in the view. In the Data pane, under Parameters, right-click the Top Customers 2 parameter, and select Show Parameter. You can control the top N value by using the Top Customers 2 parameter

control that appears in the view.



Additional tips for improving the view's functionality

Here are some additional steps you can take to give your viewers more flexibility in displaying the customer subsets.

1. From the Data pane drop-down menu, select Create Parameter.

2. In the Create Parameter dialog box that opens, do the following:
 - For Name, type Expand or Collapse.
 - For Data type, select String.
 - For Allowable values, select List.
 - In the List of values, type the values Expand and Collapse.

Create Parameter

Name: Comment >>

Properties

Data type:

Current value:

Value when workbook opens:

Display format:

Allowable values: ☐ All ☒ List ☐ Range

List of values

Value	Display As
Expand	Expand
Collapse	Collapse
Add	

☒ Fixed

Add values from ▶

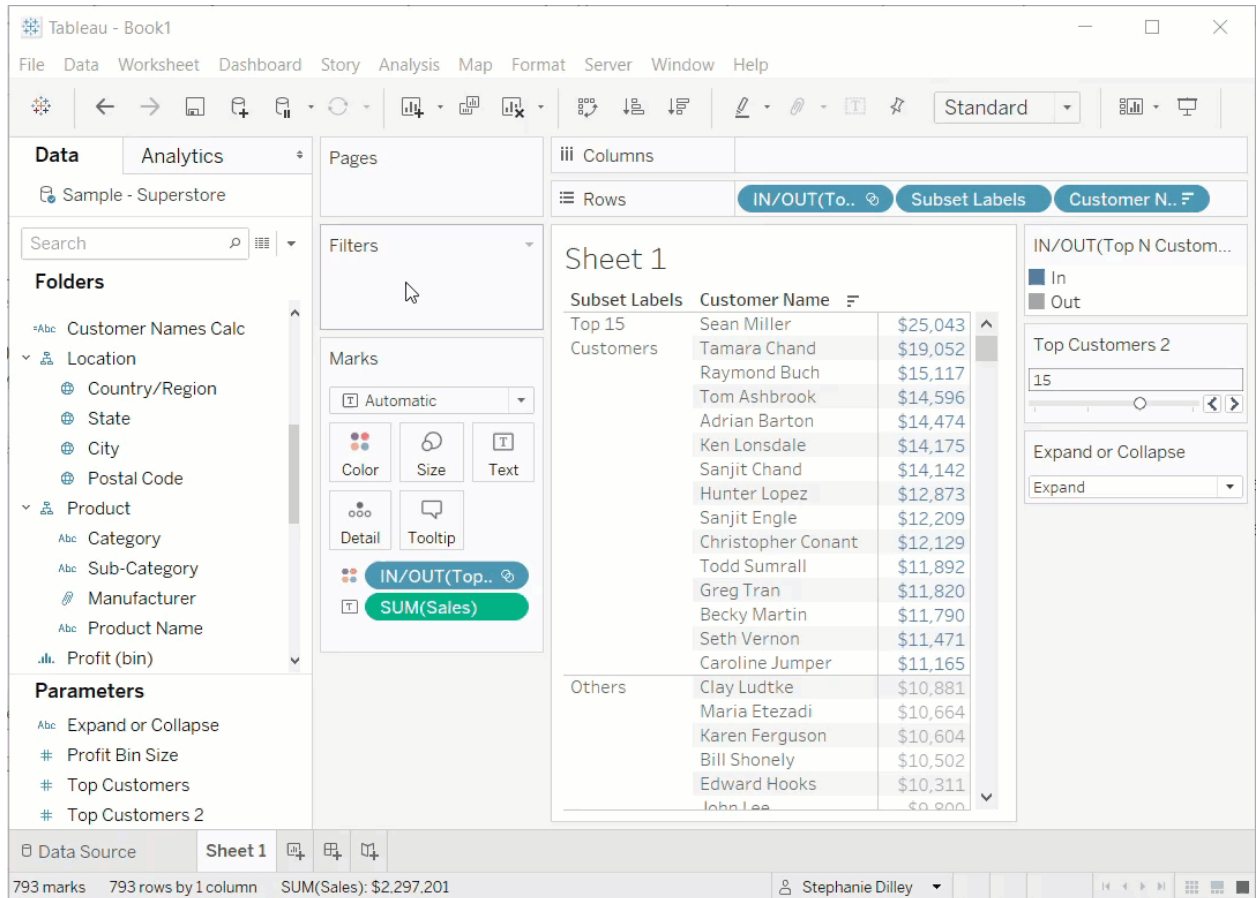
☐ When workbook opens

None ▼

- When finished, click OK.
3. Select Analysis > Create Calculated Field.
 4. In the Create Calculated Field dialog box that opens, do the following to create a calculation that uses the parameter you just created. This calculation enables viewers to specify how to view customers in the Others subset:
 - For Name, enter Customer Names Calc.
 - In the formula box, type the following formula, and then click OK:


```
IF [Expand or Collapse]="Collapse" THEN
  IF [Top N Customers by Sales]
  THEN [Customer Name]
  ELSE "Others" END
ELSE [Customer Name] END
```

5. In the Data pane, under Parameters, right-click the Expand or Collapse parameter, and select Show Parameter.
6. From the Columns shelf, drag the SUM(Sales) measure to Label on the Marks card.
7. From the Data pane, drag Customer Names Calc directly on top of the Customer Name field on the Rows shelf, so that it replaces it.



Now you can use the Expand or Collapse parameter control to see the list of names in the top N customers and the remaining customers rolled up into a

single Others entry.

Data | Analytics | Pages | Columns

Sample - Superstore

Search

Folders

- Customer
 - Customer Name
 - Segment
- Order
 - Order Date
 - Order ID
 - Ship Date
 - Ship Mode
- Customer Names Calc
- Location
 - Country/Region
 - State
 - City
 - Postal Code
- Product

Parameters

- Expand or Collapse
- Profit Bin Size
- Top Customers
- Top Customers 2

Filters

Marks

Automatic

Color Size Text

Detail Tooltip

IN/OUT(Top N Custom...)

SUM(Sales)

Sheet 1

Subset Labels Customer Names Calc

Top 15	Adrian Barton	\$14,474
Customers	Becky Martin	\$11,790
	Caroline Jumper	\$11,165
	Christopher Conant	\$12,129
	Greg Tran	\$11,820
	Hunter Lopez	\$12,873
	Ken Lonsdale	\$14,175
	Raymond Buch	\$15,117
	Sanjit Chand	\$14,142
	Sanjit Engle	\$12,209
	Sean Miller	\$25,043
	Seth Vernon	\$11,471
	Tamara Chand	\$19,052
	Todd Sumrall	\$11,892
	Tom Ashbrook	\$14,596
Others	Aaron Bergman	\$886
	Aaron Hawkins	\$1,745
	Aaron Smayling	\$3,051
	Adam Bellavance	\$7,756
	Adam Hart	\$3,250
	Adam Shillingsburg	\$3,255
	Adrian Hane	\$1,736
	Adrian Shami	\$59
	Aimee Bixby	\$967
	Alan Barnes	\$1,114
	Alan Dominguez	\$6,107

IN/OUT(Top N Custom...)

In Out

Top Customers 2

15

Expand or Collapse

Expand

Data Source | Sheet 1