Rakamin Academy

Case Study - Tableau Part 1

by Regita Ardia

PROBLEM

You are a Product Manager exploring data from the Sample Superstore. Do a simple exploration into the data to get the information you need. For increase profits, we must reduce the number of order returned. You have to present regarding any items returned by the buyer.

Make appropriate charts to answer the following questions:

1. Which Category and Ship Mode has the highest number of Order Returns?
2. In what year did the number of Order Returns in July increase from June?
3. Find the percentage of Order Return from United States compared to total Order return in 2015, and what Quarter is the highest Order Return in the US?

DATA PREPARATION

Given the dataset contains Global Superstore Orders records in 2012 - 2015. It has 24 fields and 51290 rows.

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment
40098	CA-2014-AB10015140-41954	11/11/2014	13/11/2014	First Class	AB-100151402	Aaron Bergman	Consumer
26341	IN-2014-JR162107-41675	05/02/2014	07/02/2014	Second Class	JR-162107	Justin Ritter	Corporate
25330	IN-2014-CR127307-41929	17/10/2014	18/10/2014	First Class	CR-127307	Craig Reiter	Consumer

Postal Code	City	State	Country	Region	Market	Product ID	Category	Sub-Category
73120	Oklahoma City	Oklahoma	United States	Central US	USCA	TEC-PH-5816	Technology	Phones
null	Wollongong	New South Wales	Australia	Oceania	Asia Pacific	FUR-CH-5379	Furniture	Chairs
null	Brisbane	Queensland	Australia	Oceania	Asia Pacific	TEC-PH-5356	Technology	Phones

Product Name	Sales	Quantity	Discount	Profit	Shipping Cost	Order Priority
Samsung Convoy 3	221.98	2	0.000000	62.15	40.770	High
Novimex Executive Leather A...	3,709.40	9	0.100000	-288.77	923.630	Critical
Nokia Smart Phone, with Call...	5,175.17	9	0.100000	919.97	915.490	Medium

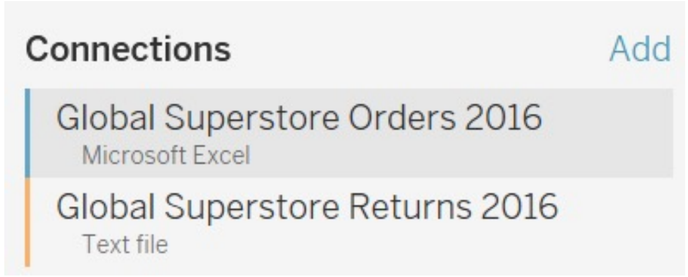
DATA PREPARATION

On the second table, given the dataset contains Global Superstore Returns record in 2012 - 2015. It has 3 fields and 2220 rows.

Abc Global Superstore Returns 2016.csv Returned	Abc Global Superstore Returns 2016.csv Order ID (Global Supers...	Abc Global Superstore Returns 2016.csv Region (Global Supersto...
Yes	IN-2012-PB19210127-41259	Eastern Asia
Yes	IN-2012-PB19210127-41259	Eastern Asia
Yes	IN-2015-JH158207-42140	Oceania
Yes	CA-2014-AS10045140-41727	Southern US
Yes	IN-2014-LC168857-41747	Oceania
Yes	ID-2013-AB1001527-41439	Eastern Asia
Yes	ES-2015-RA1994545-42218	Western Europe

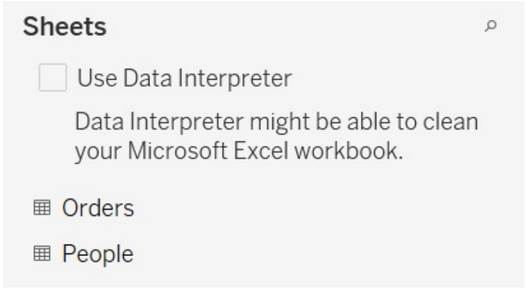
Connections

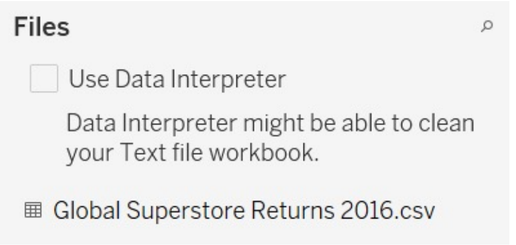
Add all the dataset to Connections menu.



Sheets and Files

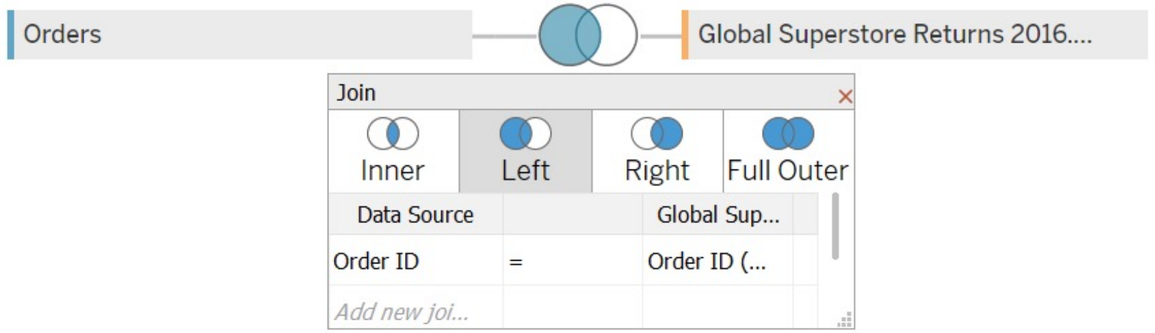
Select Orders sheet and Returns file.



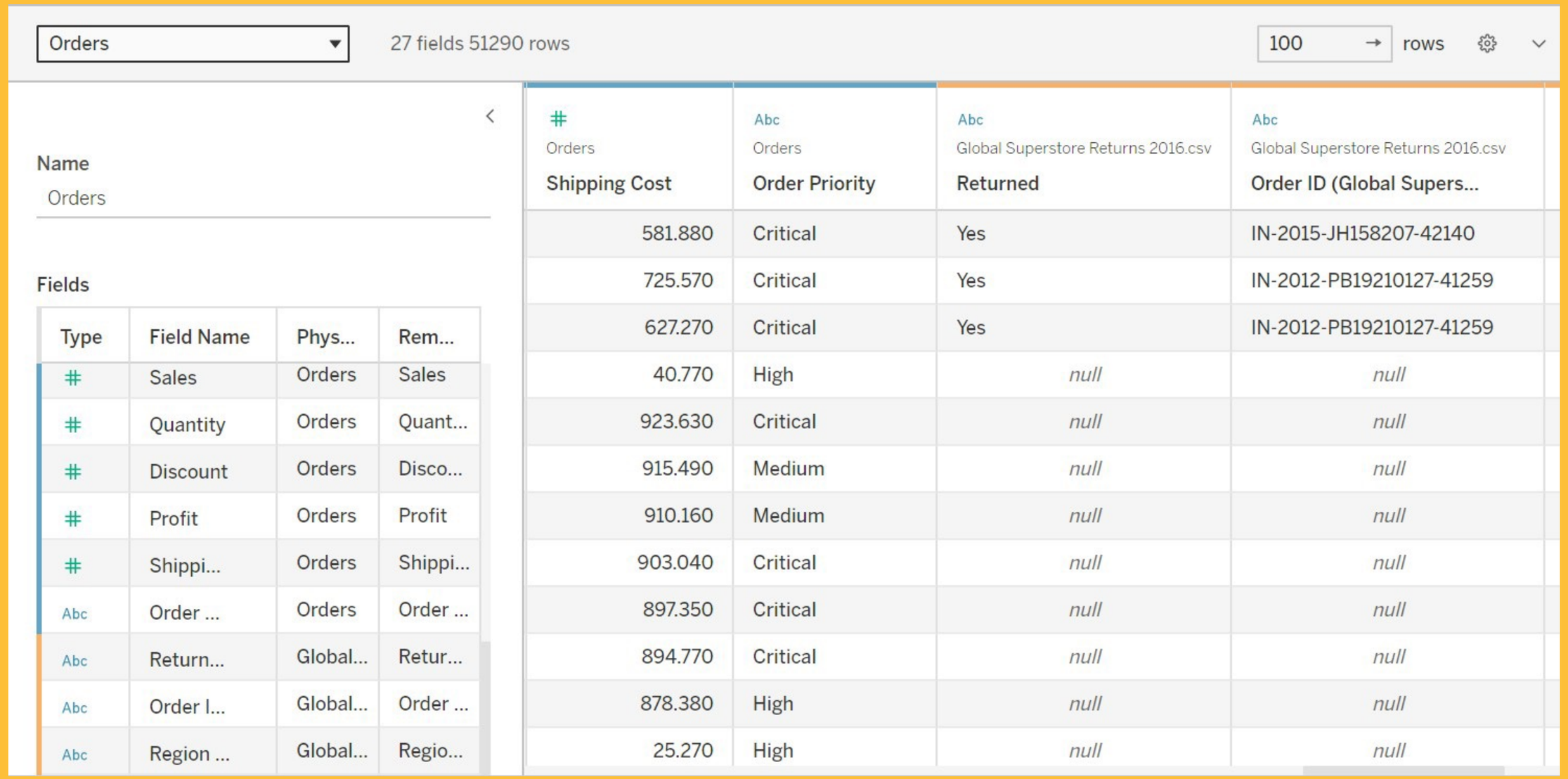


Joins

Choose Left Join, and connect the key collumn using Order ID on both tables.



Data connection is successful. Now it has 27 fields and 51290 rows.

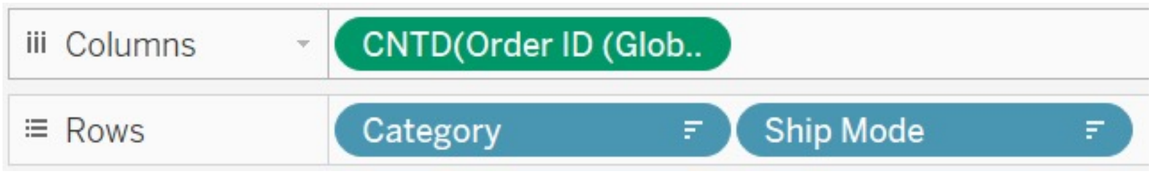


QUESTION 1

Which Category and Ship Mode has the highest number of Order Returns?

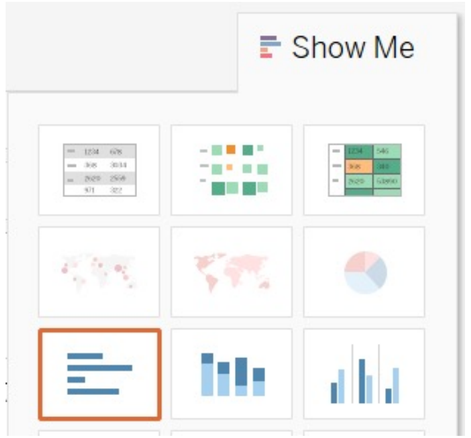
Collumns and Rows

- Add count distinct Order ID from Returns table to collumn
- Add Category and Ship Mode to rows



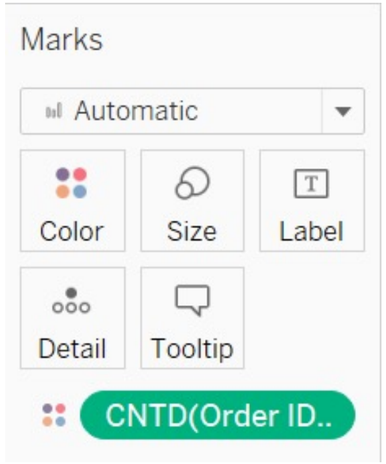
Chart

Use horizontal bars chart for categorical series



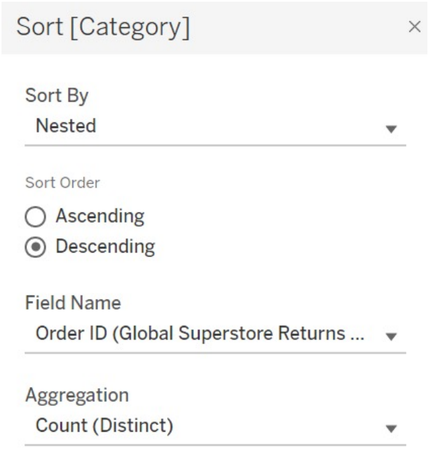
Marks

Add count distinct Order ID from Returns table to Color

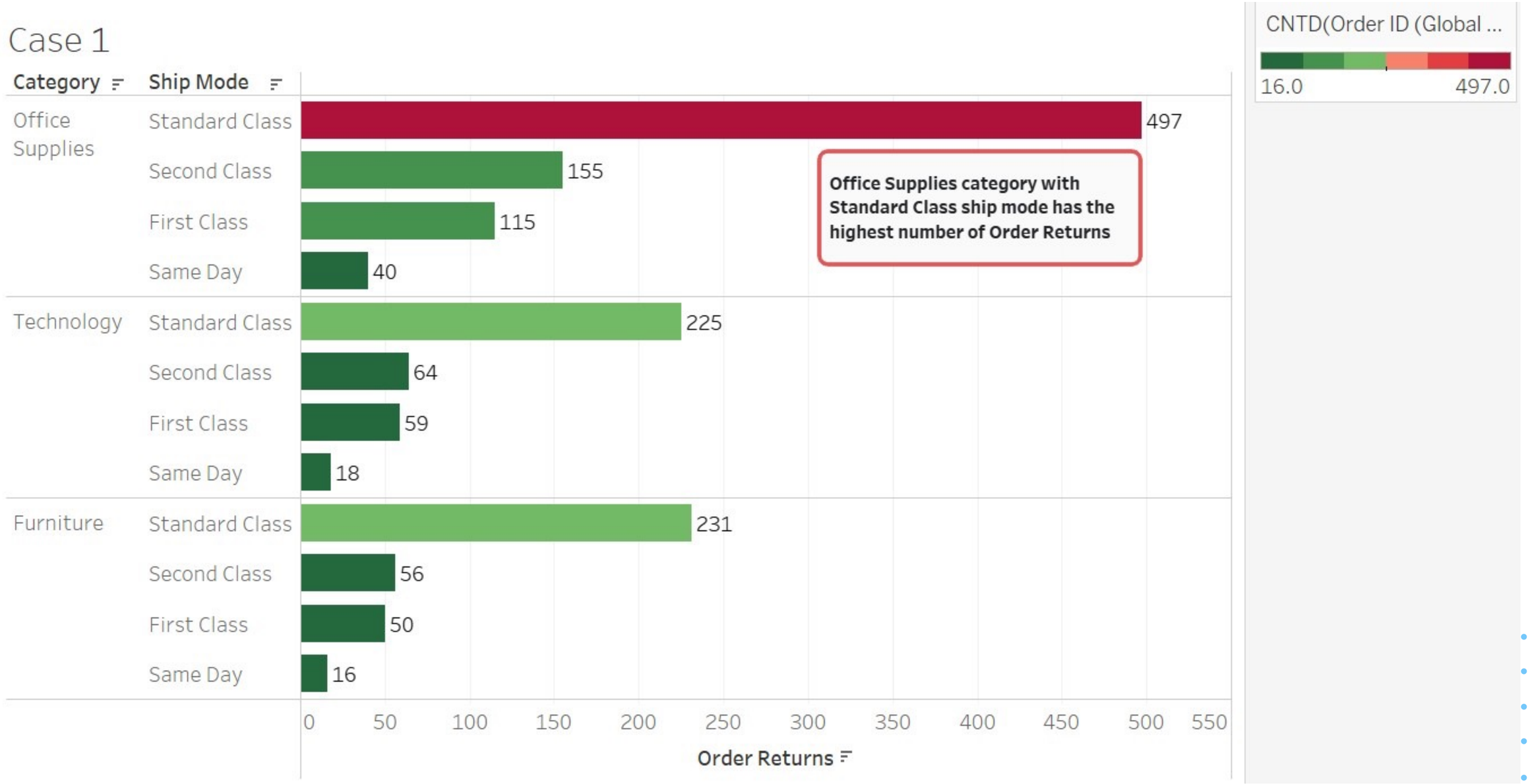


Sort

Sort Category field and Ship Mode field by count distinct Order ID from Returns table with descending order



RESULT

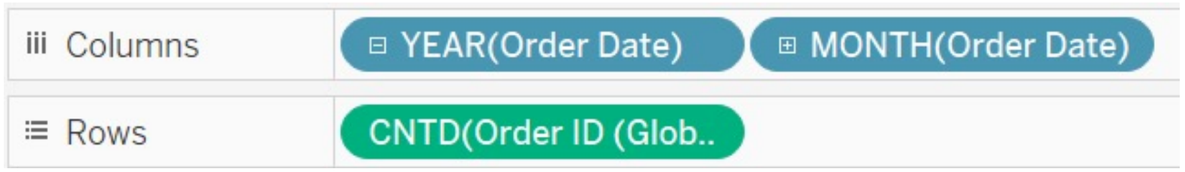

Case 1
Category
Ship Mode
Office Supplies
Standard Class
497
Second Class
155
First Class
115
Same Day
40
Technology
Standard Class
225
Second Class
64
First Class
59
Same Day
18
Furniture
Standard Class
231
Second Class
56
First Class
50
Same Day
16
0
50
100
150
200
250
300
350
400
450
500
550
Order Returns
Office Supplies category with Standard Class ship mode has the highest number of Order Returns
CNTD(Order ID (Global ...
16.0
497.0

QUESTION 2

In what year did the number of Order Returns in July increase from June?

Collumns and Rows

- Add Order Date extracted by Year and Month to collumn
- Add count distinct Order ID from Returns table to row

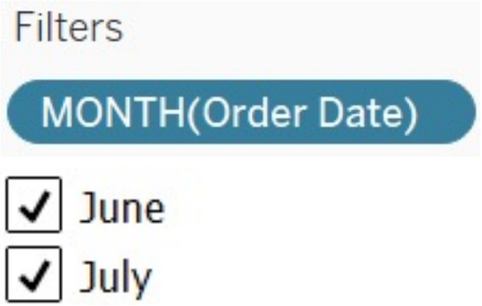


Chart

Use line chart for time series

Filters

Add filter by Month of Order Date, choose June and Julyonly



RESULT


Order Date
2012
2013
2014
2015
Order Returns
40
35
30
25
20
15
10
5
0
14
10
13
19
33
15
38
16
June
July
The number of Order Returns in July increased from June occurred in 2013

QUESTION 3

Find the percentage of Order Return from United States compared to total Order return in 2015, and what Quarter is the highest Order Return in the US?

Collumns and Rows

- Add Order Date extracted by Quarter to collumn
- Add count distinct Order ID from Returns table to row

Chart

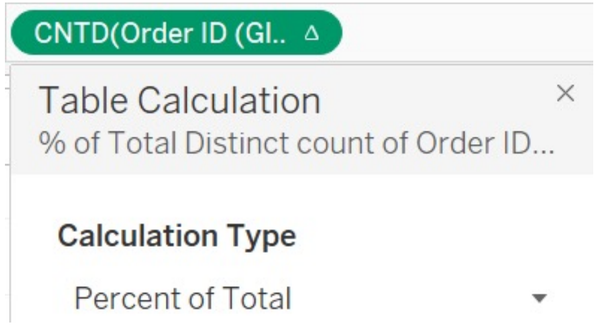
Use line chart for time series

Filters

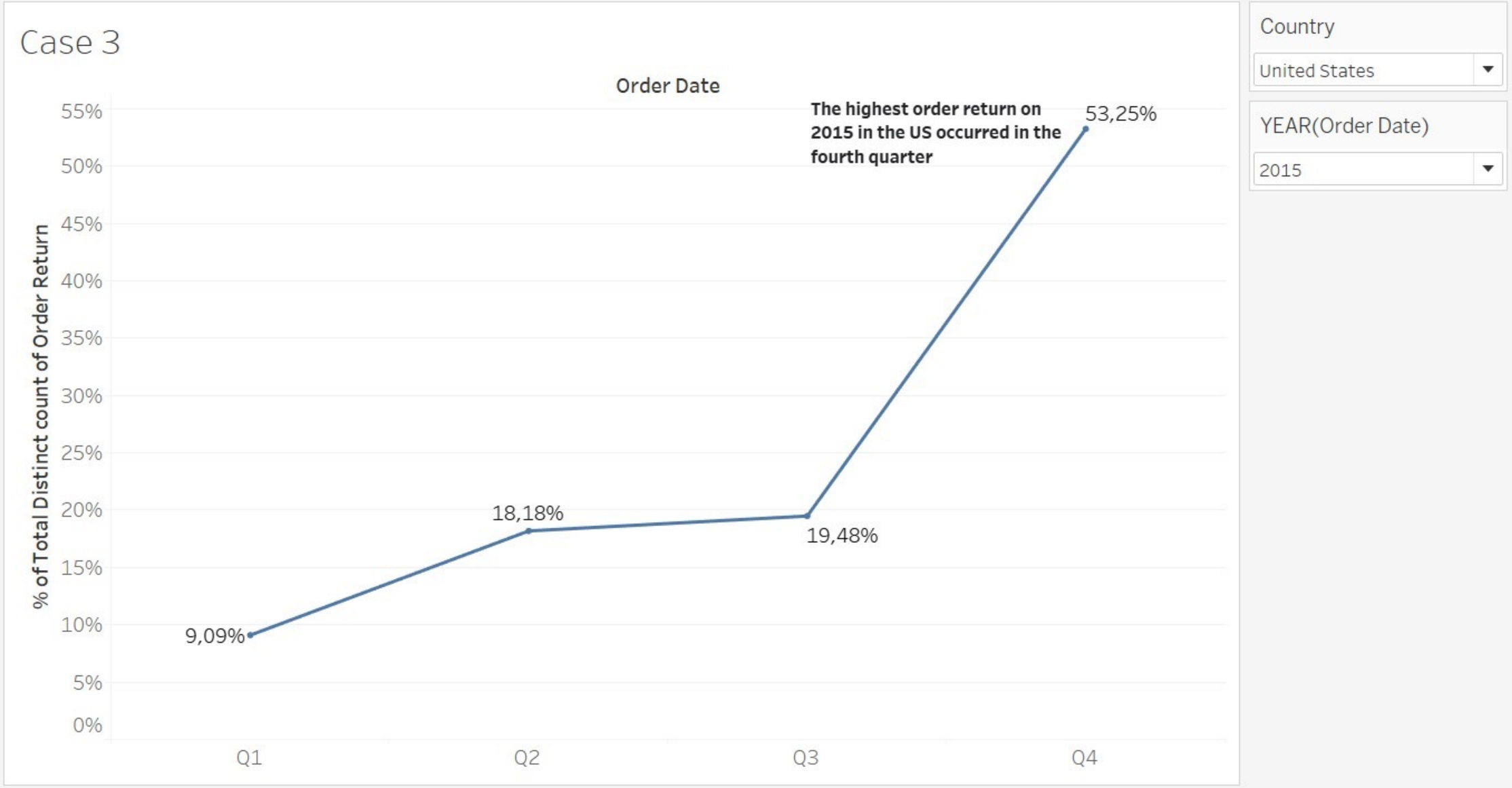
- Add Year of Order Date, choose 2015
- Add Country, choose United States

Table Calculation

Add quick table calculation on count distinct Order ID from Returns table, choose Percentof Total



RESULT


Case 3
Order Date
% of Total Distinct count of Order Return
55%
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
The highest order return on 2015 in the US occurred in the fourth quarter
53,25%
18,18%
19,48%
9,09%
Q1
Q2
Q3
Q4
Country
United States
YEAR(Order Date)
2015

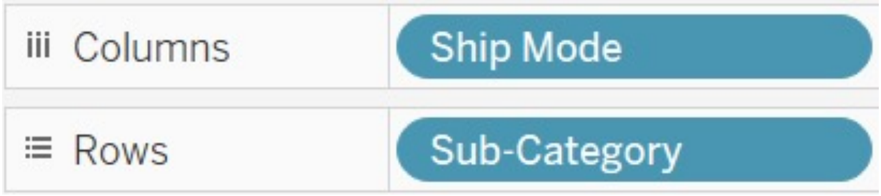
ADDITIONAL QUESTIONS

1

Create a Crosstab that displays the Sub Category and Ship Mode. Choose the following Sub-category and Shipmode field that has the highest number of Order Returns!

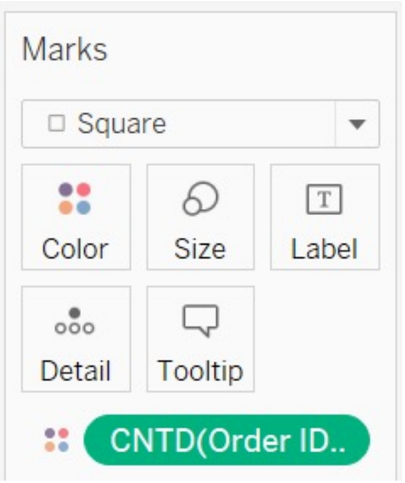
Collumns and Rows

- Add Ship Mode to collumn
- Add Sub-Category to row



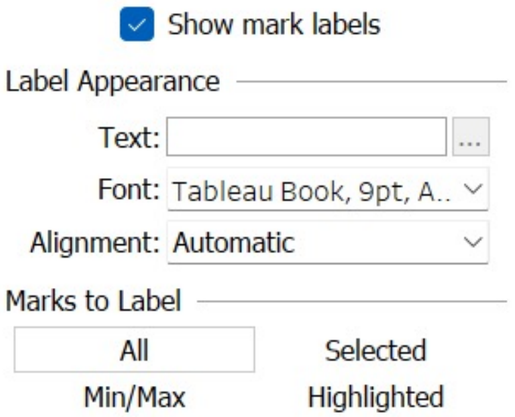
Marks

- Choose Square marks
- Add count distinct Order ID from Returns table to Color



Labels

Tick the "Show mark labels"



RESULT

	Ship Mode			
Sub-Catego..	First Class	Same Day	Second Class	Standard Class
Accessories	15	7	18	89
Appliances	3	2	11	43
Art	25	10	31	121
Binders	33	12	50	144
Bookcases	19	3	15	62
Chairs	11	9	25	87
Copiers	17	6	12	59
Envelopes	15	7	11	58
Fasteners	9	3	19	67
Furnishings	15	5	14	90
Labels	28	4	26	69
Machines	15	3	11	30
Paper	15	10	25	82
Phones	21	2	27	83
Storage	24	14	37	111
Supplies	13	7	23	53
Tables	8	3	6	21

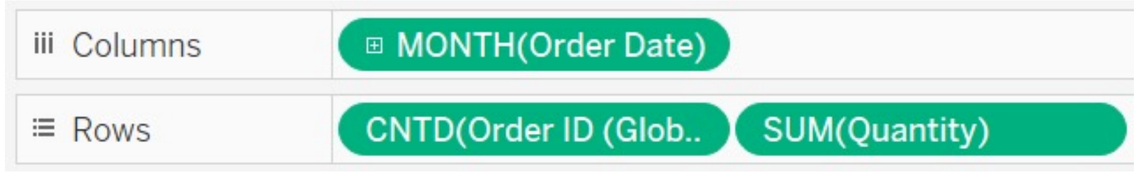
Binders sub-category with Standard Class ship mode has the highest number of Order Returns

2

Make a Line Chart with Order Date and Order Return Amount, how many the Order Return Amount and Quantity in July 2013?

Collumns and Rows

- Add Order Date extracted by Month to collumn
- Add count distinct Order ID from Returns table and sum of Quantity to row

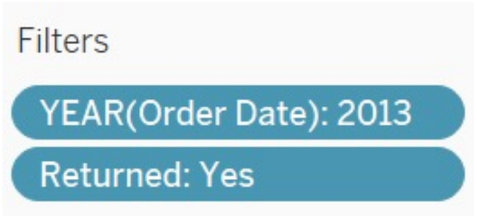


Chart

Use line chart for time series

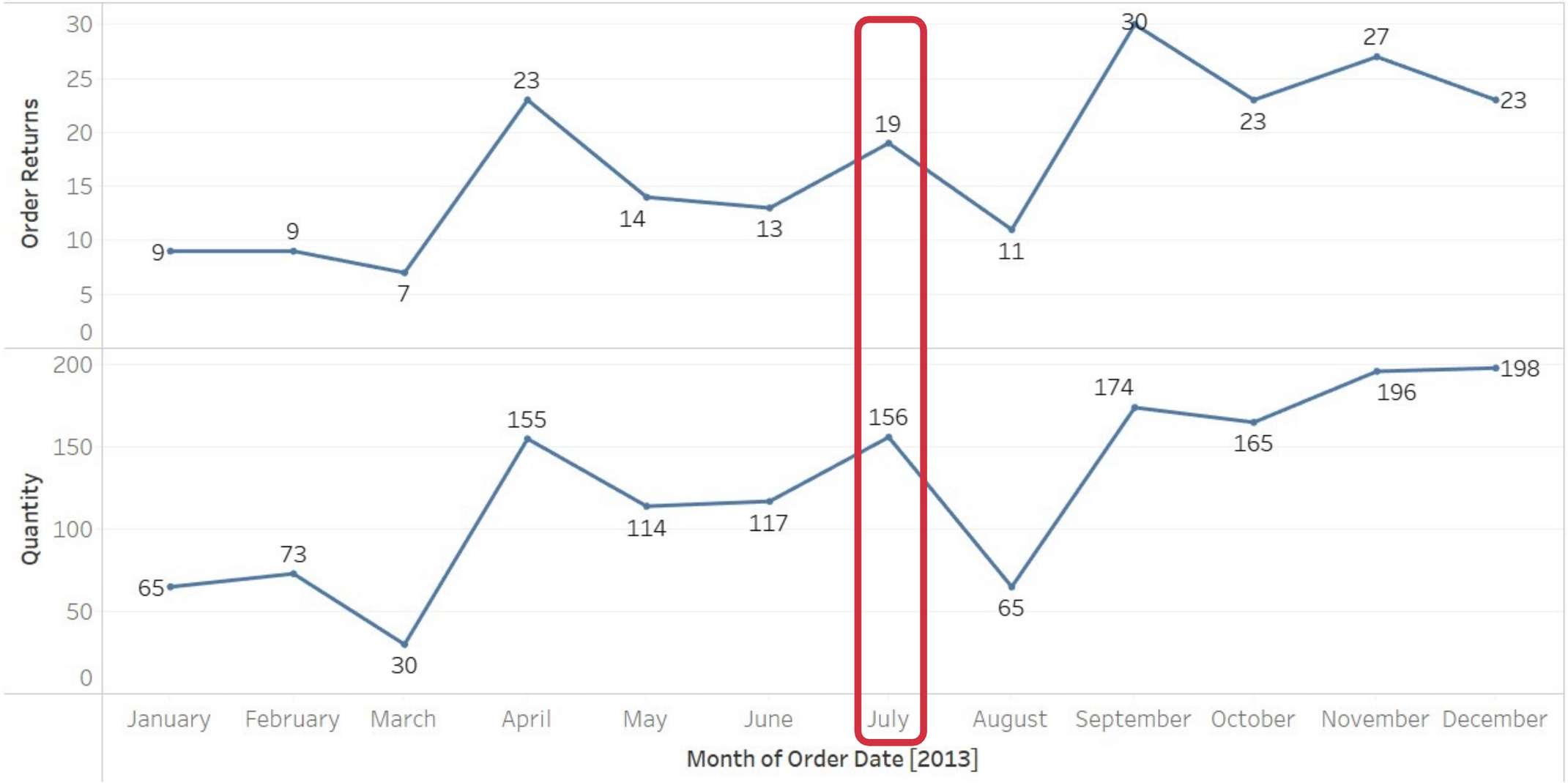
Filters

- Add Year of Order Date, choose 2013
- Add Returned, choose "Yes"



RESULT

There are 19 Order Returns and 156 Quantity Returned on July 2013



3

What is the Total Accumulated Order Return throughout 2015?

Collumns and Rows

- Add Order Date extracted by Month to collumn
- Add count distinct Order ID from Returns table to row

Chart

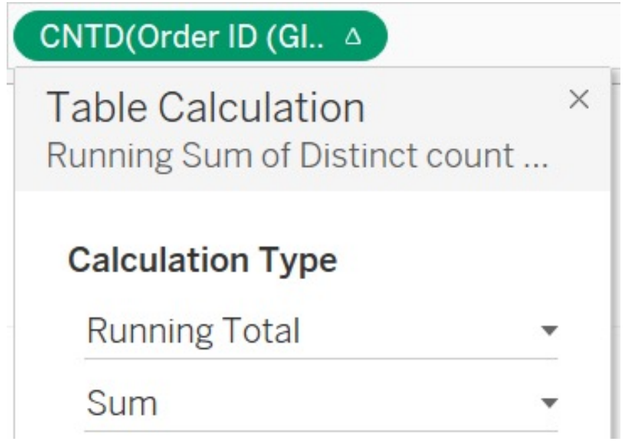
Use line chart for time series

Filters

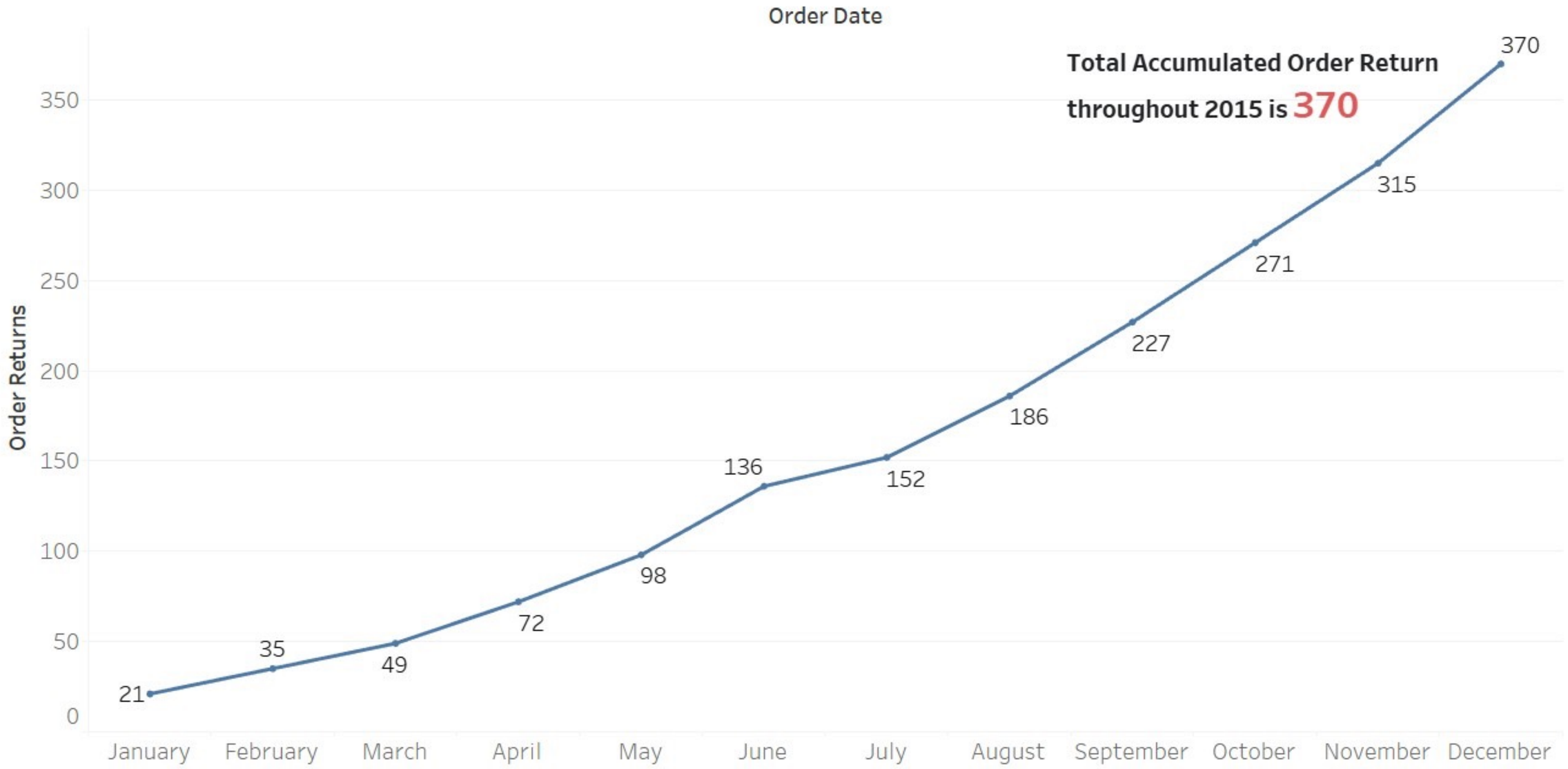
Add Year of Order Date, choose 2015

Table Calculation

Add quick table calculation on count distinct Order ID from Returns table, choose Running Total



RESULT


Order Date
Total Accumulated Order Return throughout 2015 is 370
Order Returns
350
300
250
200
150
100
50
0
21
35
49
72
98
136
152
186
227
271
315
370
January
February
March
April
May
June
July
August
September
October
November
December

4

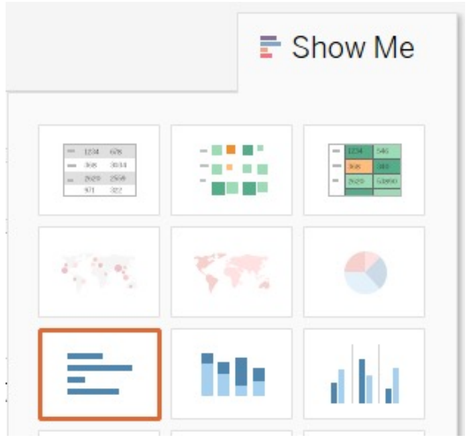
Make a Bar Chart that displays the Country and Order Return Amount. Show the top 5 Countries with the highest number of Returned orders in 2015!

Collumns and Rows

- Add count distinct Order ID from Returns table to collumn
- Add Country to row

Chart

Use horizontal bars chart for categorical series

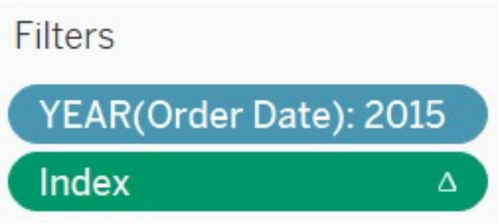


Calculated Field

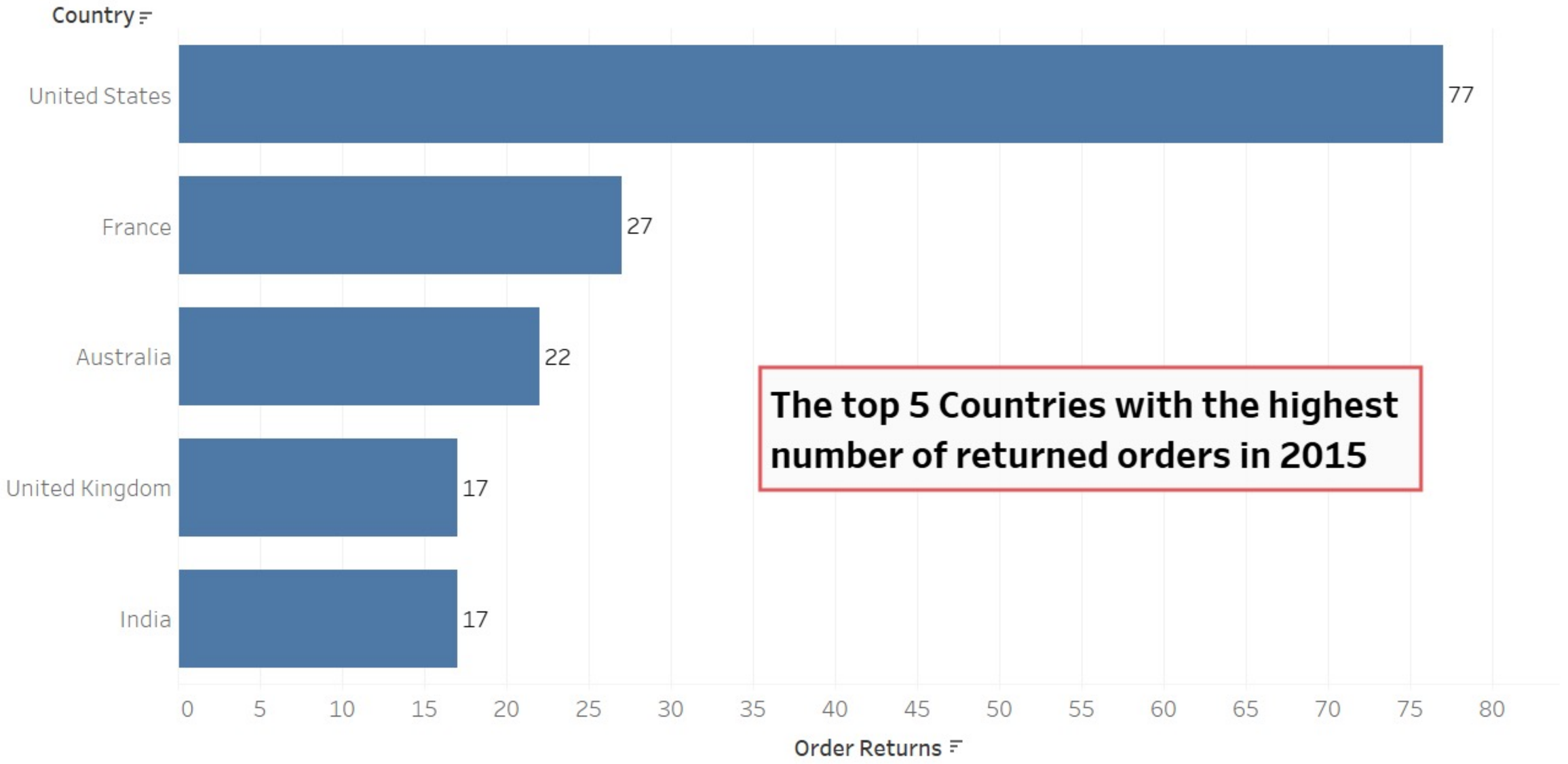
Create a calculated field named "Index" to help us filtering with Top function, write INDEX() function to return the index of filtered result.

Filters

- Add Year of Order Date, choose 2015
- Add Index, set the range of value from 1 to 5



RESULT



The top 5 Countries with the highest number of returned orders in 2015

5

Create a Calculated Field that calculates Time to Ship (Order Date - Ship Date). Show the top 5 countries with the longest average Time to Ship in 2015!

Calculated Field

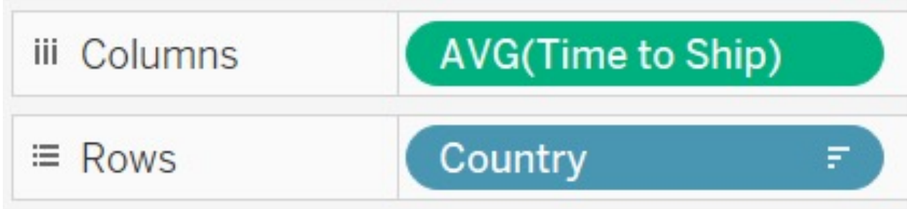
Create a calculated field named "Time to Ship", use DATEDIFF() function to return the time of the order has been shipped.

Time to Ship

```
DATEDIFF('day',[Order Date],[Ship Date])
```

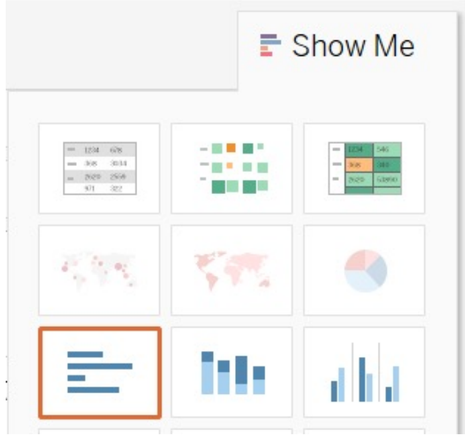
Collumns and Rows

- Add average of Time to Ship to collumn
- Add Country to row



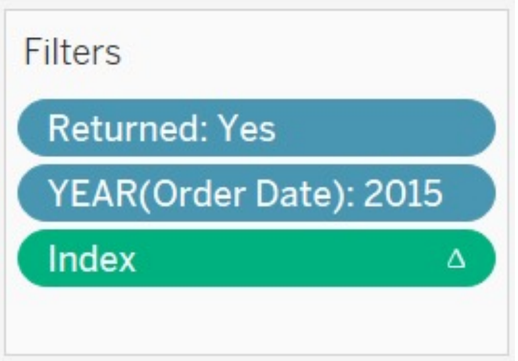
Chart

Use horizontal bars chart for categorical series

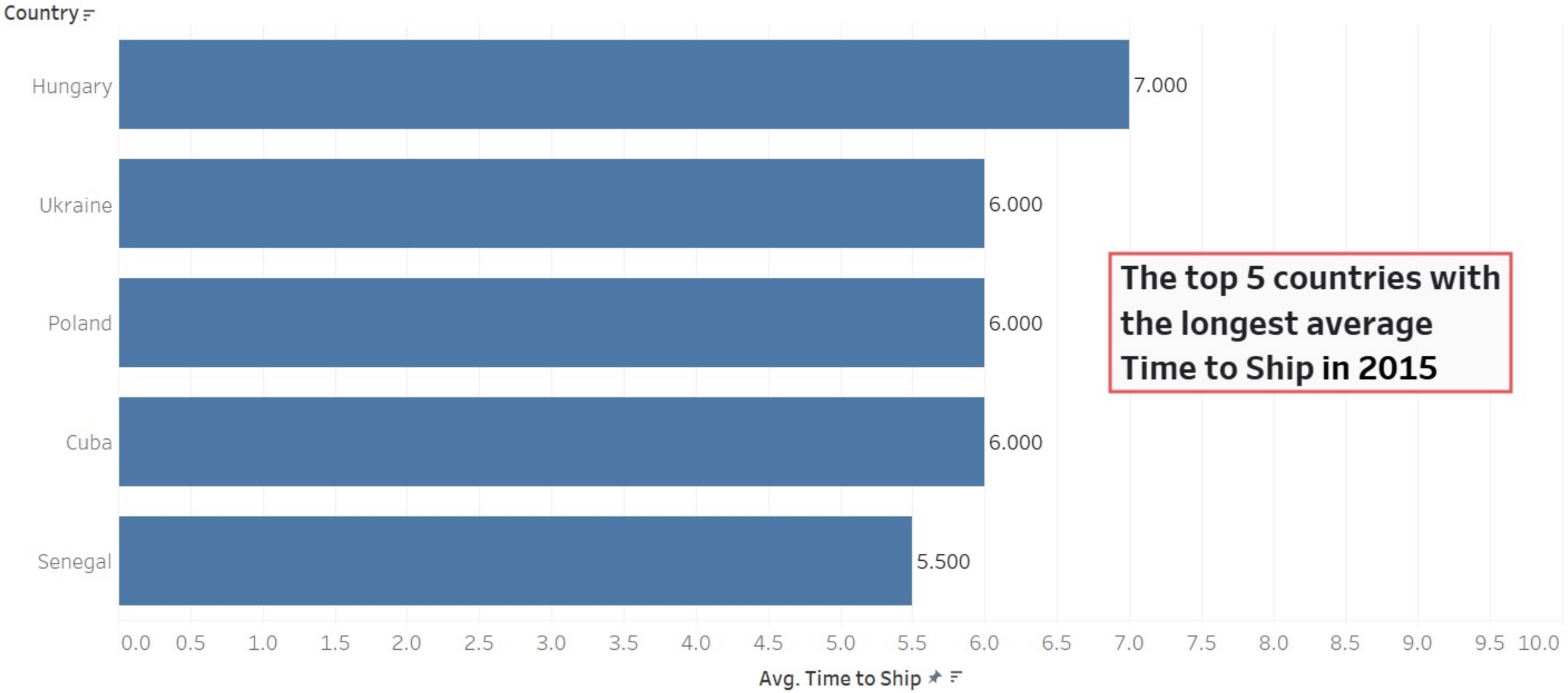


Filters

- Add Returned, choose "Yes"
- Add Year of Order Date, choose 2015
- Add Index, set the range of value from 1 to 5



RESULT



The top 5 countries with the longest average Time to Ship in 2015