Visualization Using Tableau
Converting Data into Information

By Mark Friedman, University of Miami

Sign up for a free student version of Tableau Desktop

Download 14-day free trial of Tableau Desktop

https://www.tableau.com/products/desktop/download
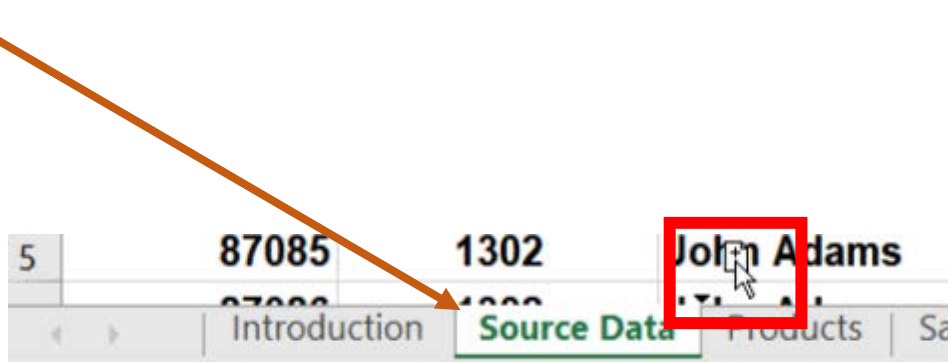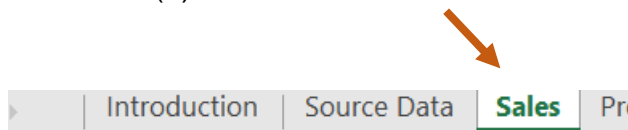
Your supervisor has forwarded to you an Excel file and wants you to create several visualizations. Recurring projects are usually cleaned and transformed using specialized ETL (Extract, Transform, and Load) software such as Tableau Prep Builder, Alteryx, or Power BI. However, this is a onetime project which contains predominantly clean data, which you will correct in Excel. To provide an audit trail, copy the tab "Source Data" to a new tab "Sales".

Open the file Cuddly_Stuffed_Animals.xlsx in Excel, and select the "Source Data" sheet.

Click on the tab "Source Data", and then press the Ctrl key. Move slightly to the right, release the mouse, and then release the Ctrl key.

| 5 | 87085 | 1302 | John Adams |

Introduction | **Source Data** | Products | Sa

Introduction | **Source Data** | **Source Data (2)** | Products

Rename the "Source Data (2)" tab "Sales".

Introduction | Source Data | **Sales** | Pro

Part 1: Excel, Clean the Data

Before connecting the Excel worksheet to Tableau, any inconsistently recorded data must be rectified. For example, there are only four valid entries for "Salesperson". If a spelling error or an abbreviation exists in the data, such as "G. Washington", it must be corrected. The data verification can easily be accomplished by using either Pivot Tables or Filters.

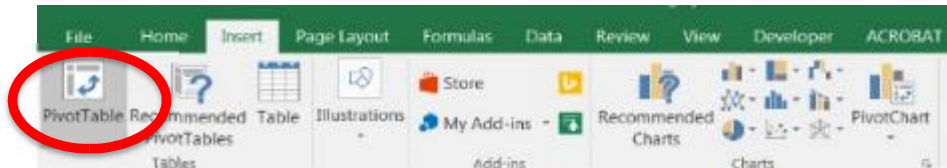| Code	Salesperson
1301	George Washington
1302	John Adams
1303	Thomas Jefferson
1304	James Madison

Verifying the "Salesperson" field using a Pivot Table.

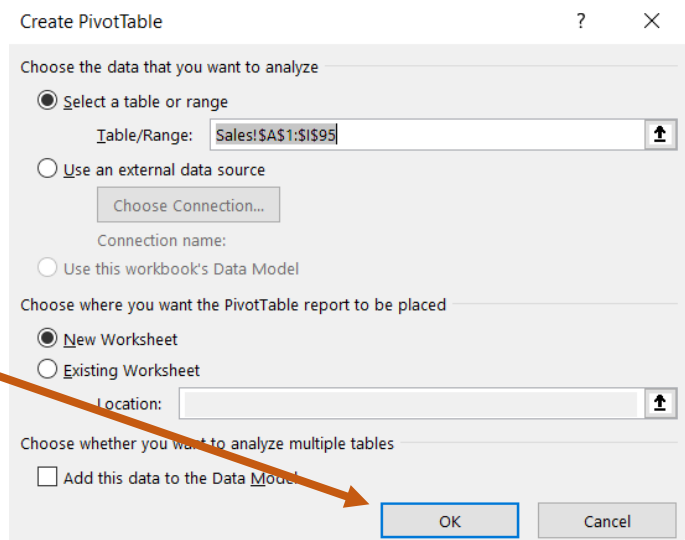Confirm that the file Cuddly_Stuffed_Animals.xlsx in Excel is open and that the "Sales" sheet is selected.

On the Sales sheet, select any cell with data, for example, cell C7.

	A	B	C	D	E	F	G	H	I
1	Sales Order #	Salesperson #	Salesperson	Title	Region	State	Product	# of units	Sales
2	87082	1301	George Washington	Sales Associate I	South	MS	Stuffed Pig	480	$12,000.00
3	87083	1302	John Adams	Senior Sales Associate	Northeast	VT	Stuffed Horse	1,344	$32,256.00
4	87084	1304	James Madison	Sales Associate I	Northeast	NH	Stuffed Panda	684	$15,732.00
5	87085	1302	John Adams	Senior Sales Associate	West	AK	Stuffed Pig	1,140	$28,500.00
6	87086	1302	John Adams	Senior Sales Associate	Northeast	CT	Stuffed Pig	1,224	$30,600.00
7	87087	1301	George Washington	Sales Associate I	Northeast	NJ	Stuffed Horse	468	$11,232.00
8	87088	1301	George Washington	Sales Associate I	Northeast	ME	Stuffed Bear	1,548	$30,960.00

Introduction | Source Data | **Sales** | Products | Salesperson | Region

On the Insert tab, select "Pivot Table".

By default, the "Table/Range" is defined based on the number of rows and columns in the data set, and the placement for the Pivot Table will be a "New Worksheet". Select "OK" for the default options.

Create PivotTable

Choose the data that you want to analyze
- ⦿ Select a table or range
  - Table/Range: Sales!$A$1:$I$95
- ○ Use an external data source
  - Choose Connection...
  - Connection name:
  - ○ Use this workbook's Data Model

Choose where you want the PivotTable report to be placed
- ⦿ New Worksheet
- ○ Existing Worksheet
  - Location:

Choose whether you want to analyze multiple tables
- ☐ Add this data to the Data Model

OK | Cancel

Add "Salesperson" to the "Rows" section by selecting the field and dragging it.
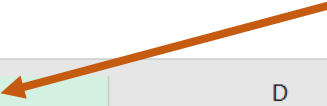
Drag "Sales" to the "Values" section by selecting the field and dragging it.
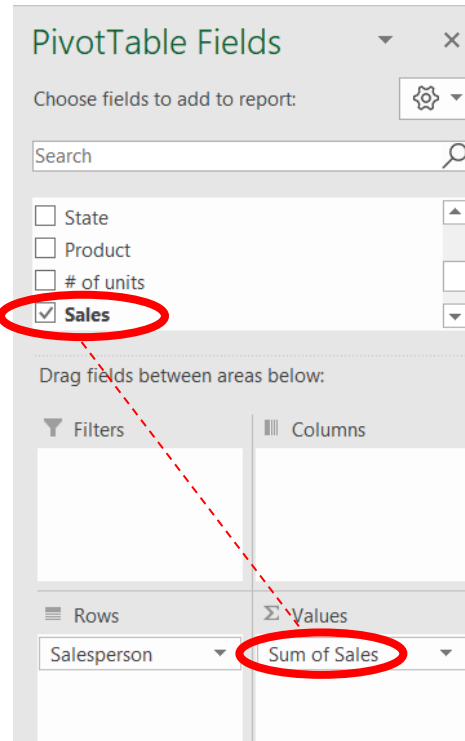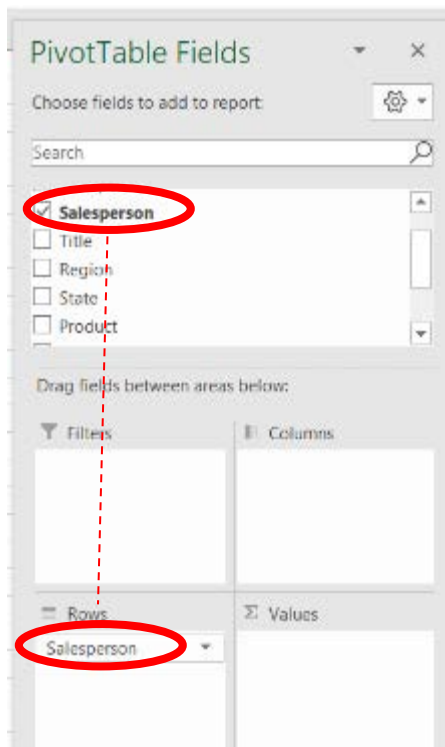
Note: In the data, Tom Jefferson appears instead of Thomas Jefferson

Row Labels	Sum of Sales
George Washington	471648
James Madison	720936
John Adams	341352
Thomas Jefferson	302100
Tom Jefferson	69900
Grand Total	**1905936**

Return to the Sales sheet, and select "Column C" by left-clicking on the "C".

	A	B	C	D	E
1	Sales Order #	Salesperson #	Salesperson	Title	Regior
2	87082	1301	George Washington	Sales Associate I	South
3	87083	1302	John Adams	Senior Sales Associate	Northea
4	87084	1304	James Madison	Sales Associate I	Northea
5	87085	1302	John Adams	Senior Sales Associate	West

Visualization Using Tableau

With "Home" selected
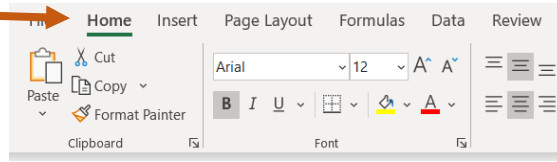
Select "Find & Select"
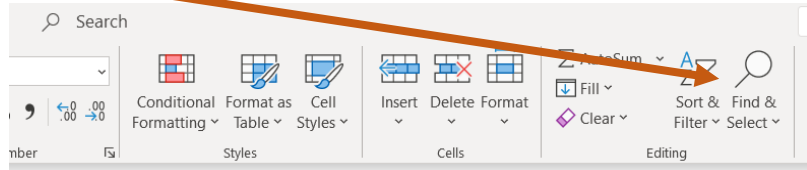
and then "Replace".

Complete the dialog box.
  "Find what:" Tom
  "Replace with:" Thomas
  Select: "Replace All"

Find and Replace
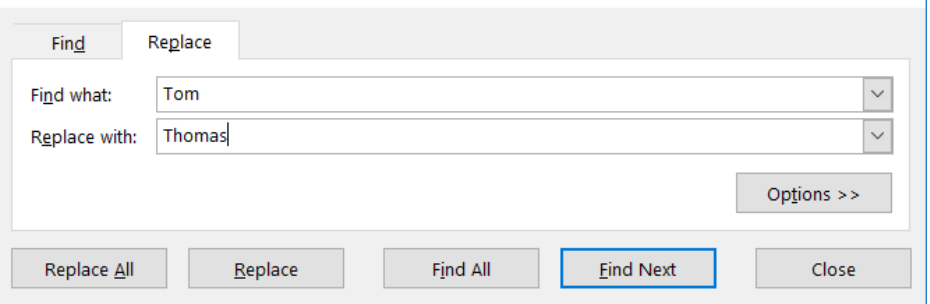
Find	Replace

Find what: Tom

Replace with: Thomas

Options >>

| Replace All | Replace | Find All | Find Next | Close |

Select "OK".

Microsoft Excel

All done. We made 3 replacements.

OK

"Close" the "Find and Replace" dialog box.

Find and Replace ?

Find	Replace

Find what: Tom

Replace with: Thomas

Options >>

| Replace All | Replace | Find All | Find Next | Close |
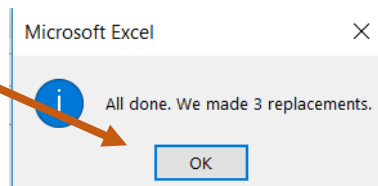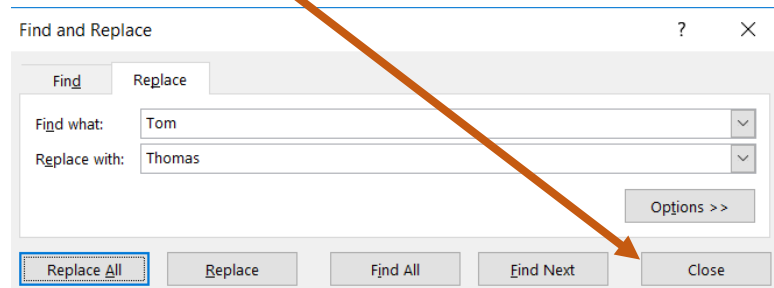
Return to the sheet containing the Pivot Table, and the "Salesperson" field will not be correctly displayed until the Pivot Table is refreshed.

	A
1	
2	
3	**Row Labels**
4	George Washington
5	James Madison
6	John Adams
7	Thomas Jefferson
8	Tom Jefferson
9	**Grand Total**

Select any cell inside the Pivot Table, for example, cell A5.
    Select "Analyze".
    Select "Refresh".
    Select "Refresh All".

The "Salesperson" field now contains only valid entry data.

Clean the Data Using a Filter

Continue to verify that the Excel data is consistently recorded before connecting the Excel worksheet to Tableau. For example, there are only four valid entries for the field Region.

Number	Region
1	Midwest
2	Northeast
3	South
4	West

A spelling error or a missed coding in the field, such as "Western", must be corrected. The data verification can easily be accomplished by using a Filter.

Return to the "Sales" sheet.

Select Cells "A1" through "I1".

Select: "Data".
Select: "Filter".

Select the dropdown arrow to the right of the field "Region".
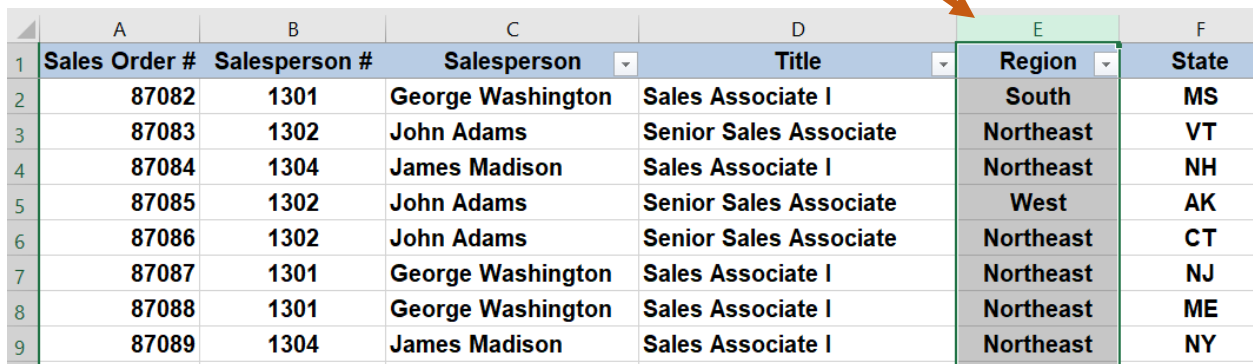
Note: "Western" appears in the data in addition to "West".

Select "Cancel".

Part 1: Clean the Data in Excel using a Filter

Select "Column E" by left-clicking on the "E".
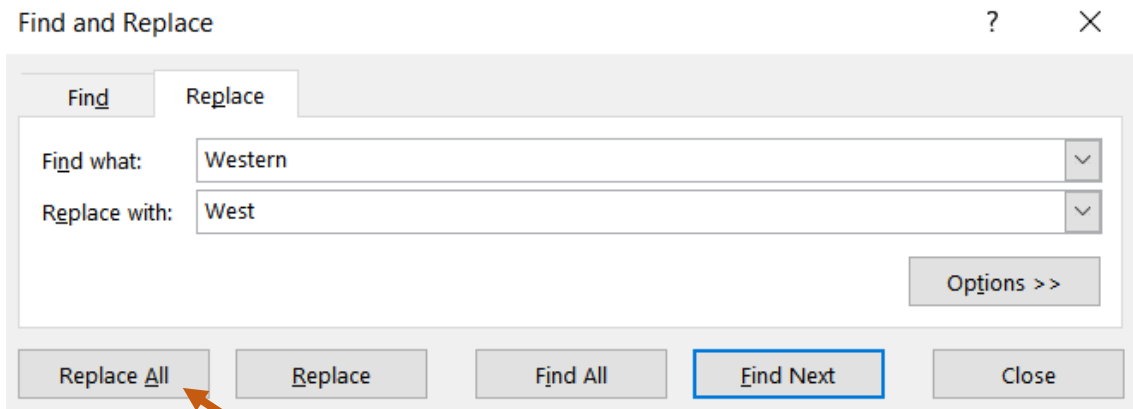
	A	B	C	D	E	F
1	Sales Order #	Salesperson #	Salesperson	Title	Region	State
2	87082	1301	George Washington	Sales Associate I	South	MS
3	87083	1302	John Adams	Senior Sales Associate	Northeast	VT
4	87084	1304	James Madison	Sales Associate I	Northeast	NH
5	87085	1302	John Adams	Senior Sales Associate	West	AK
6	87086	1302	John Adams	Senior Sales Associate	Northeast	CT
7	87087	1301	George Washington	Sales Associate I	Northeast	NJ
8	87088	1301	George Washington	Sales Associate I	Northeast	ME
9	87089	1304	James Madison	Sales Associate I	Northeast	NY

Select "Home" followed by "Find & Select".
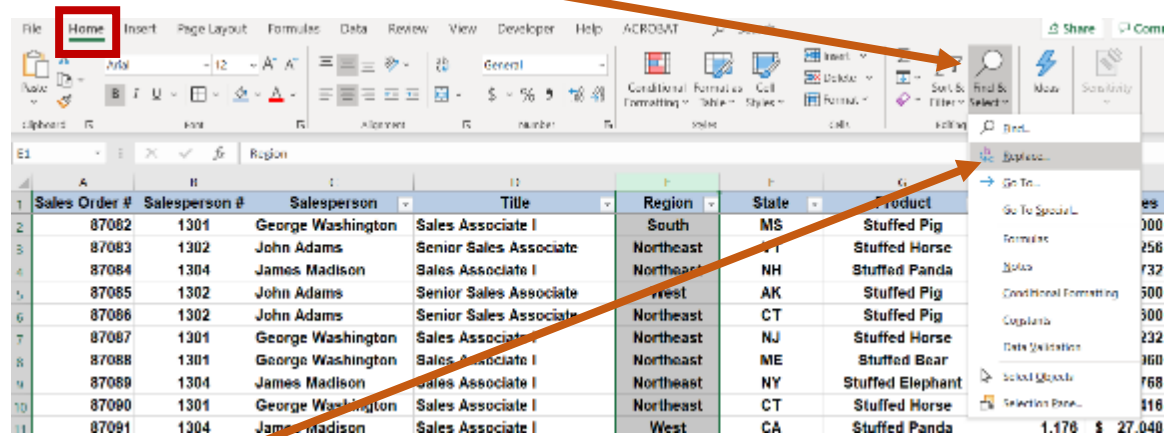
Select "Replace".

Complete the dialog box.

"Find what:" Western
"Replace with:" West"
Select: "Replace All"

Part 1: Clean the Data in Excel using a Filter

Select "OK".

> Microsoft Excel   ✕
>
> ⓘ   All done. We made 3 replacements.
>
> OK

"Close" the "Find and Replace" dialog box.

> Find and Replace   ?   ✕
>
> Find   Replace
>
> Find what: Western
>
> Replace with: West
>
>   Options >>
>
> Replace All   Replace   Find All   Find Next   Close

Select the dropdown arrow to the right of the field "Region".

	A	B	C	D	E	F	G	H
1	Sales Order #	Salesperson #	Salesperson	Title	Region	State	Product	# of units
2	87082	1301	George Washington	Sales A		MS	Stuffed Pig	48
3	87083	1302	John Adams	Senior !		VT	Stuffed Horse	1,34
4	87084	1304	James Madison	Sales A		NH	Stuffed Panda	68
5	87085	1302	John Adams	Senior !		AK	Stuffed Pig	1,14
6	87086	1302	John Adams	Senior !		CT	Stuffed Pig	1,22
7	87087	1301	George Washington	Sales A		NJ	Stuffed Horse	48
8	87088	1301	George Washington	Sales A		ME	Stuffed Bear	1,54
9	87089	1304	James Madison	Sales A		NY	Stuffed Elephant	44
10	87090	1301	George Washington	Sales A		CT	Stuffed Horse	68
11	87091	1304	James Madison	Sales A		CA	Stuffed Panda	1,17
12	87092	1303	Thomas Jefferson	Senior !		PA	Stuffed Lamb	36
13	87093	1304	James Madison	Sales A		WA	Stuffed Lamb	97
14	87094	1301	George Washington	Sales A		IL	Stuffed Lamb	1,16
15	87095	1304	James Madison	Sales A		ND	Stuffed Pig	32
16	87096	1302	John Adams	Senior !		RI	Stuffed Pig	64
17	87097	1304	James Madison	Sales A		ID	Stuffed Pig	1,03

Verify that there are only four regions used in the data.

Select "Cancel".

Verify the data in the field Product.

	A
1	**Product**
2	Stuffed Bear
3	Stuffed Elephant
4	Stuffed Horse
5	Stuffed Lamb
6	Stuffed Panda
7	Stuffed Pig

Select the dropdown arrow to the right of the field "Products".

	A	B	C	D	E	F	G	#
1	Sales Order #	Salesperson #	Salesperson	Title	Region	State	Product	
2	87082	1301	George Washington	Sales Associate I	South	MS		
3	87083	1302	John Adams	Senior Sales Associate	Northeast	VT		
4	87084	1304	James Madison	Sales Associate I	Northeast	NH		
5	87085	1302	John Adams	Senior Sales Associate	West	AK		
6	87086	1302	John Adams	Senior Sales Associate	Northeast	CT		
7	87087	1301	George Washington	Sales Associate I	Northeast	NJ		
8	87088	1301	George Washington	Sales Associate I	Northeast	ME		
9	87089	1304	James Madison	Sales Associate I	Northeast	NY		
10	87090	1301	George Washington	Sales Associate I	Northeast	CT		
11	87091	1304	James Madison	Sales Associate I	West	CA		
12	87092	1303	Thomas Jefferson	Senior Sales Associate	Northeast	PA		
13	87093	1304	James Madison	Sales Associate I	West	WA		
14	87094	1301	George Washington	Sales Associate I	Midwest	IL		

No corrections are needed as the six products used in the data are in the Product table.

Select "Cancel".

Select Data > "Filter" to remove the filter.
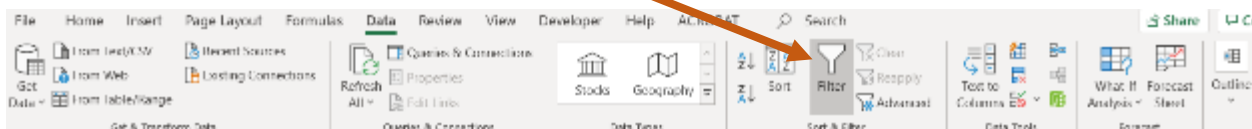
The last row of data is on line 95. There is a header and 94 rows of data.

	A	B	C	D	E	F
1	Sales Order #	Salesperson #	Salesperson	Title	Region	State
92	87172	1302	John Adams	Senior Sales Associate	Midwest	SD
93	87173	1304	James Madison	Sales Associate I	Northeast	NJ
94	87174	1304	James Madison	Sales Associate I	Northeast	NY
95	87175	1303	Thomas Jefferson	Senior Sales Associate	Northeast	MA
96						

Save, close the file and exit Excel.

Part 1: Clean the Data in Excel using a Filter

Part 2: Tableau

Connect the Data.

Open Tableau.

Select Microsoft Excel.
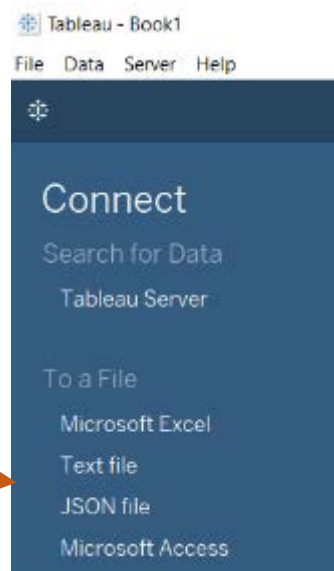
Select the file name.

Select "Open".
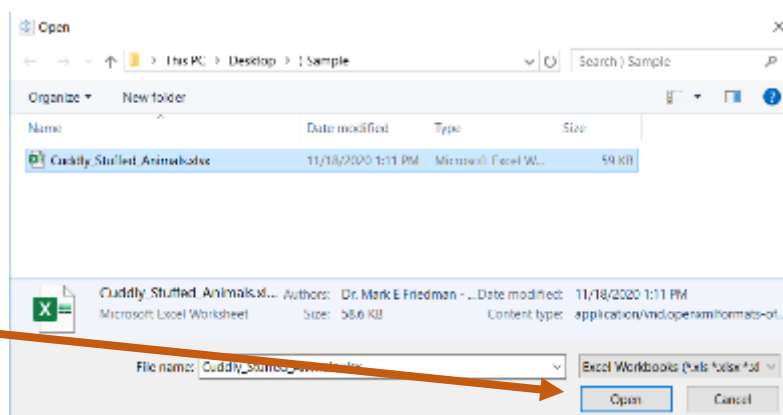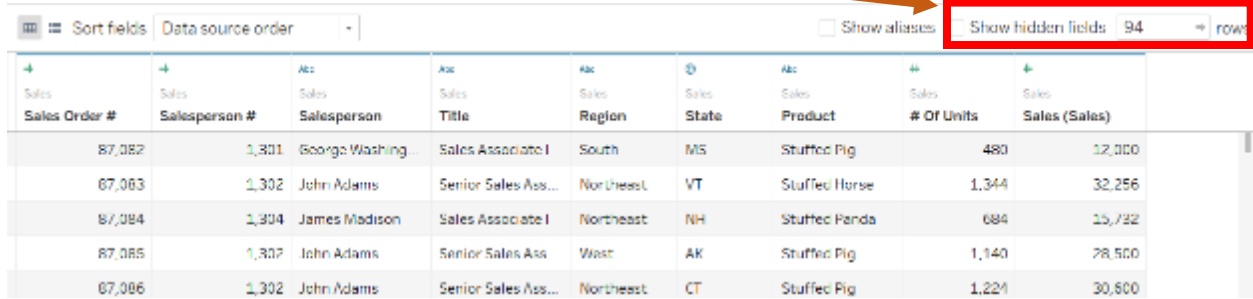
Double click on the sheet "Sales" that contains the data.
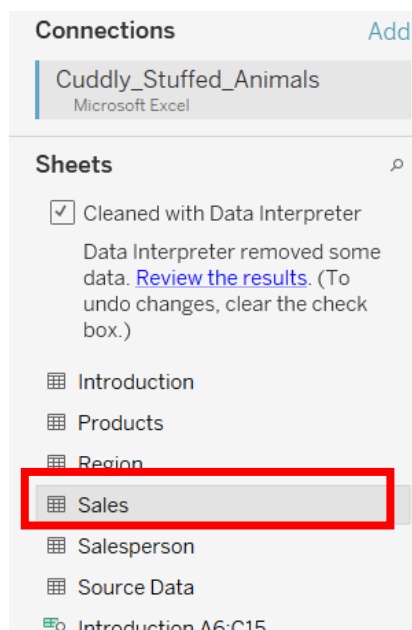
The 94 rows are connected.

Sales Order #	Salesperson #	Salesperson	Title	Region	State	Product	# Of Units	Sales (Sales)
87,082	1,301	George Washing...	Sales Associate I	South	MS	Stuffed Pig	480	12,000
87,083	1,302	John Adams	Senior Sales Ass...	Northeast	VT	Stuffed Horse	1,344	32,256
87,084	1,304	James Madison	Sales Associate I	Northeast	NH	Stuffed Panda	684	15,732
87,085	1,302	John Adams	Senior Sales Ass	West	AK	Stuffed Pig	1,140	28,500
87,086	1,302	John Adams	Senior Sales Ass...	Northeast	CT	Stuffed Pig	1,224	30,600

Prepare a visualization of the "Number of Units Sold by Region".

Double click on "Sheet 1", and rename it "Graph 1".
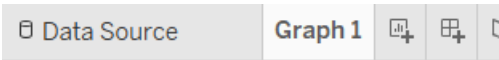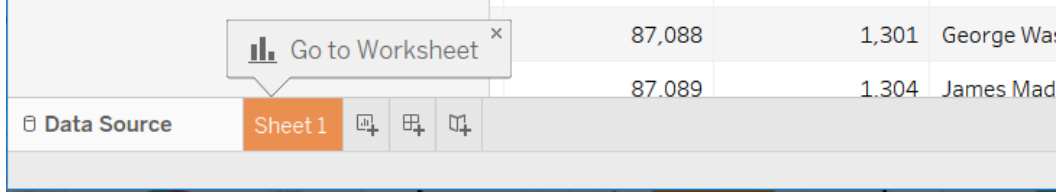
Select "Region", and drag it to "Rows".
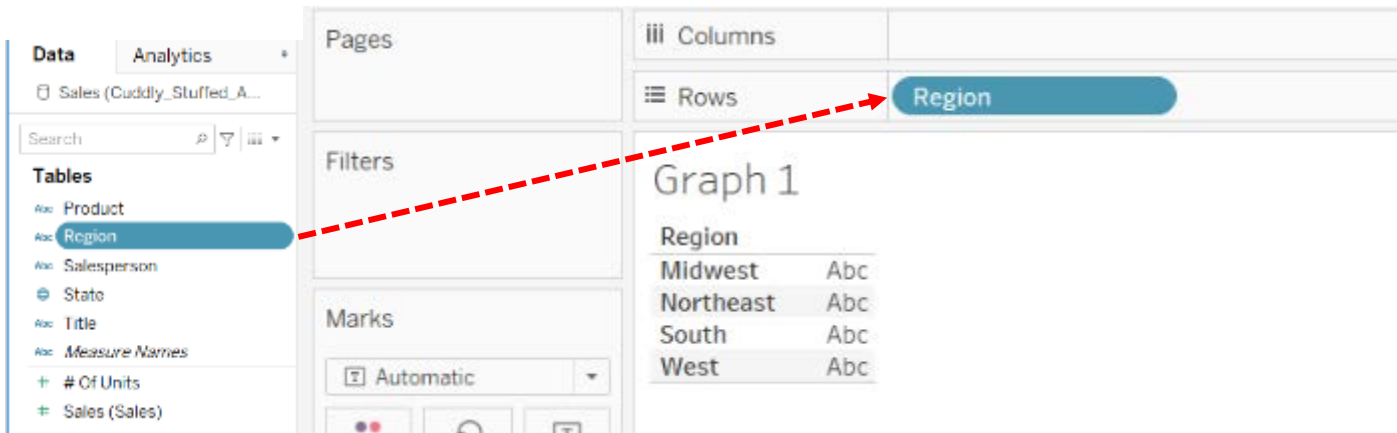
Select "# of Units", and drag it to the "Abc" column.
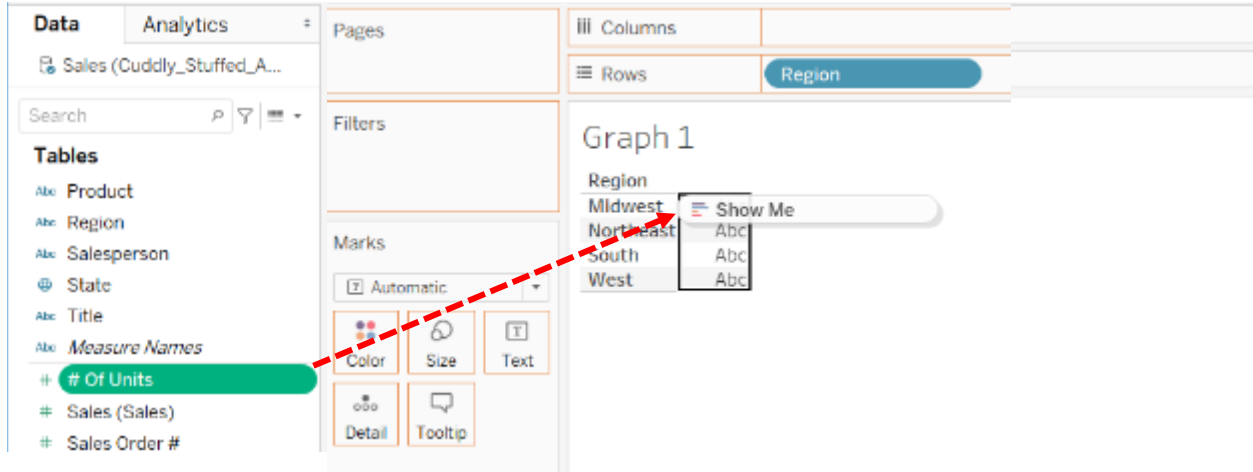
Select "Pie Graph".
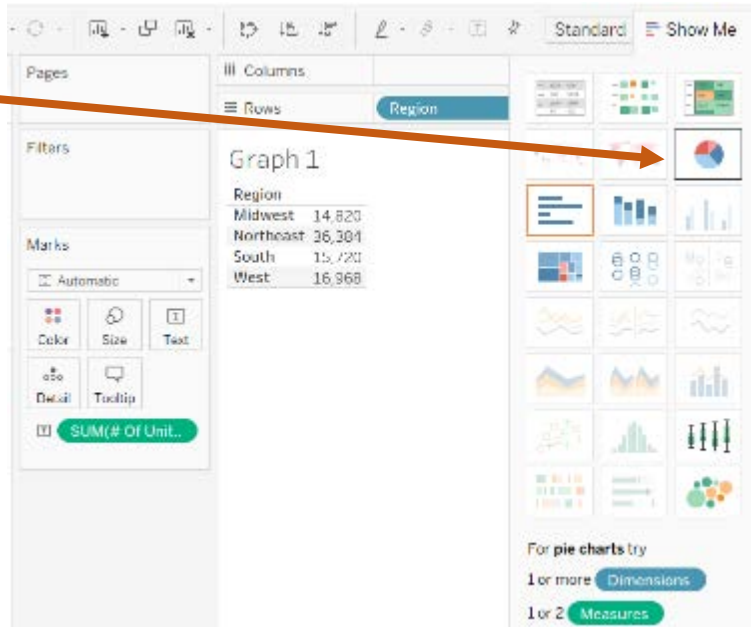
Change the "Fit" to "Entire View".
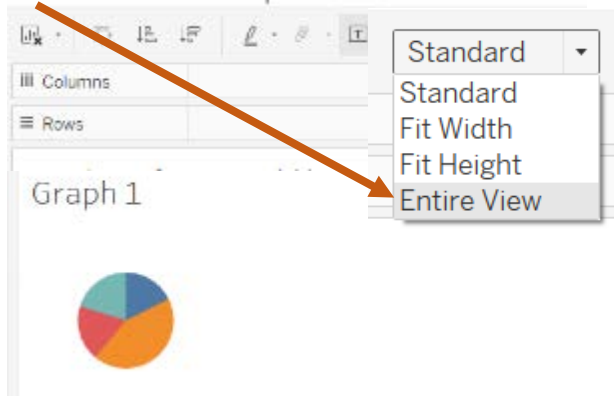
Drag "Region" to "Label".
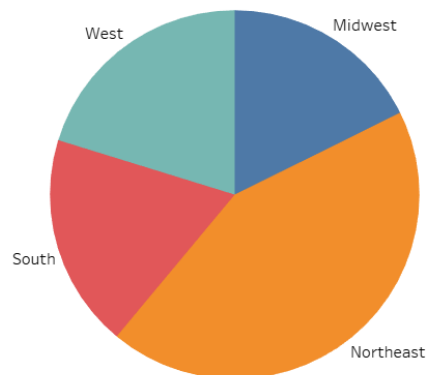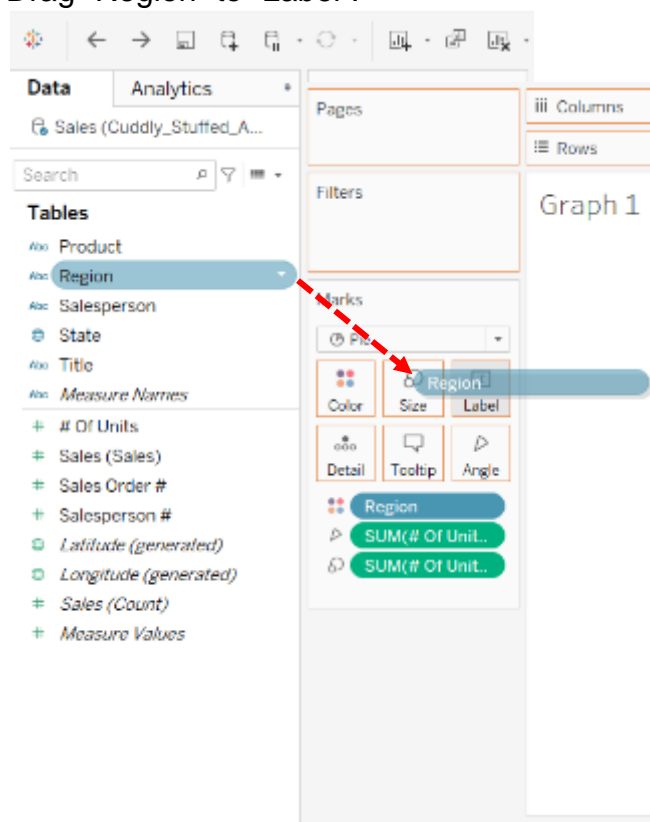
Drag "# of Units" to "Label".
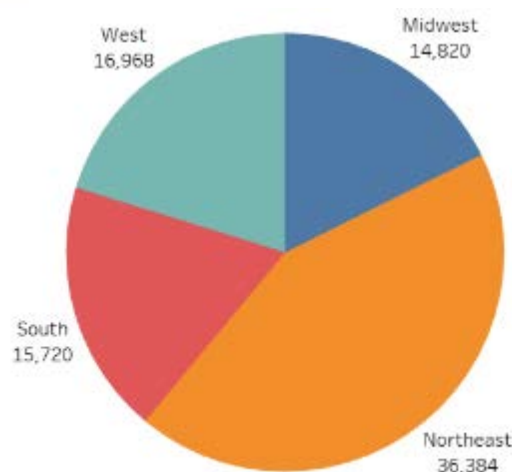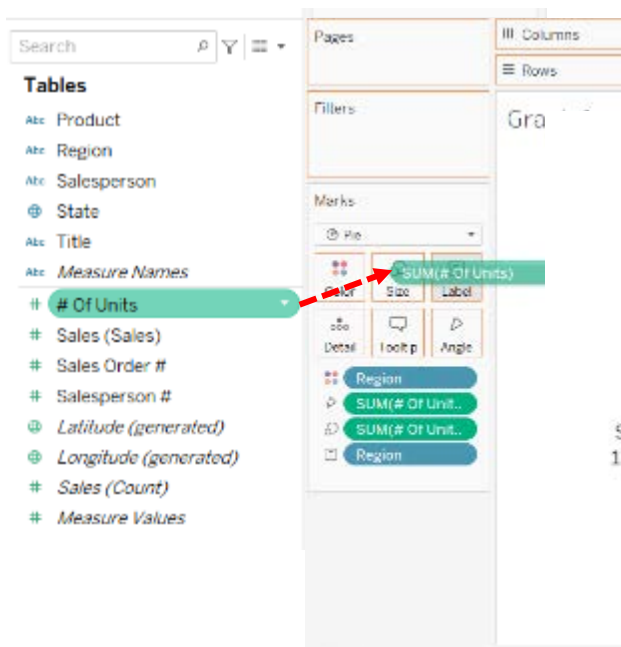
Now it is easy to visualize that Northeast sold the most units, 36,384, and that Midwest sold the least number, 14,820.

Filters	Graph 1
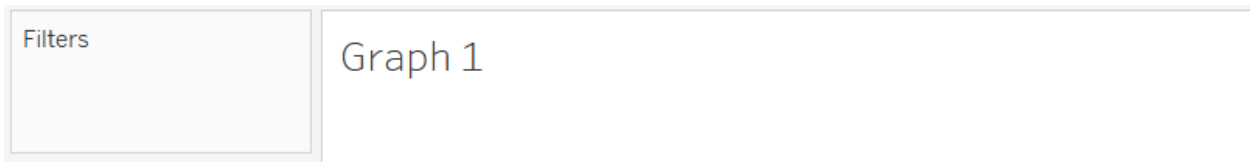
Double click on the Title of the graph, "Graph 1", and change it to "Number of Units Sold by Region". Do **NOT** change the Sheet Name on the tab.
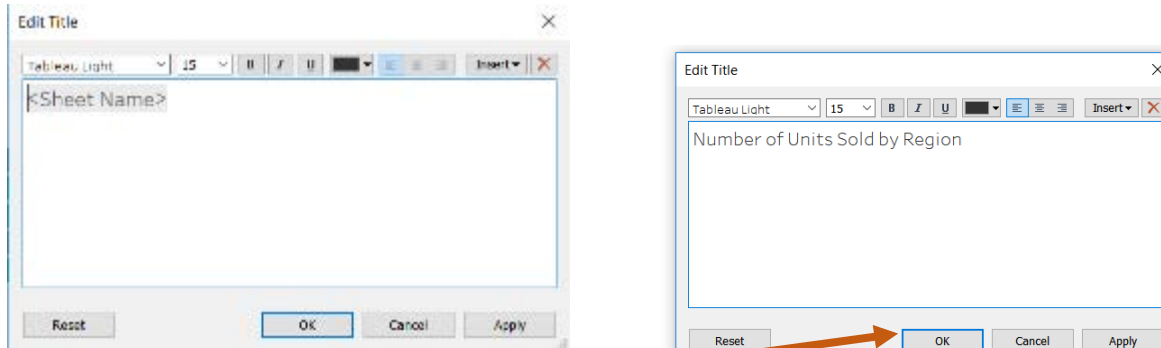
Select OK.

Graph 1 is now completed.

Preparing a visualization of the "Sales by Region".

Click on new worksheet, and rename "Sheet 2" as "Graph 2".
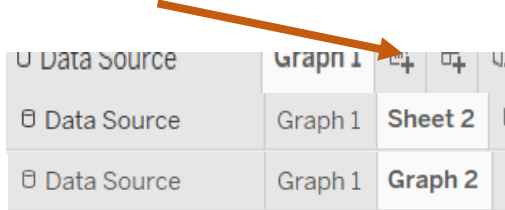
Select "Region", and drag it to "Rows".
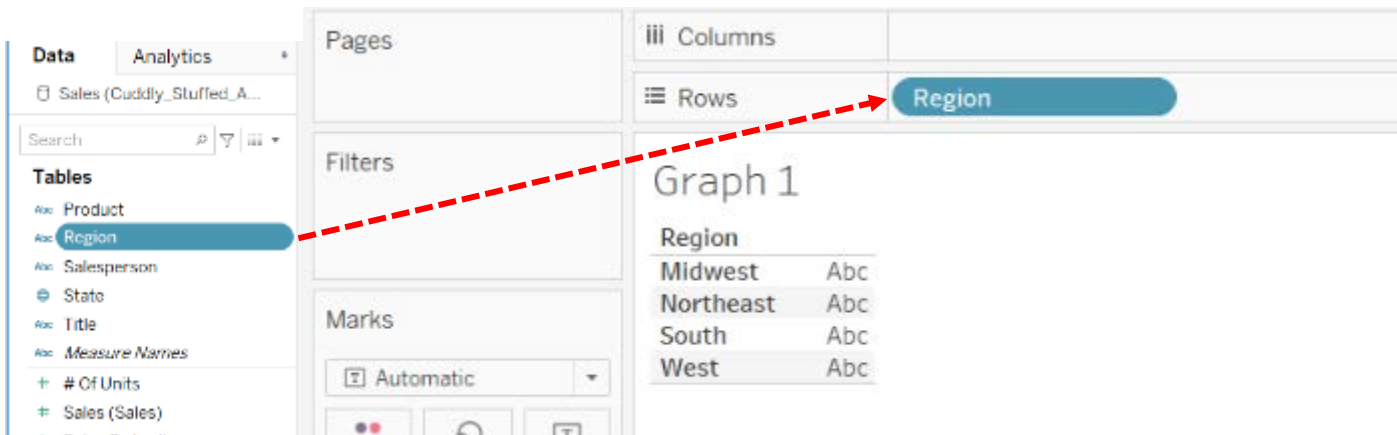
Select "Sales", and drag it to the "Abc" column.
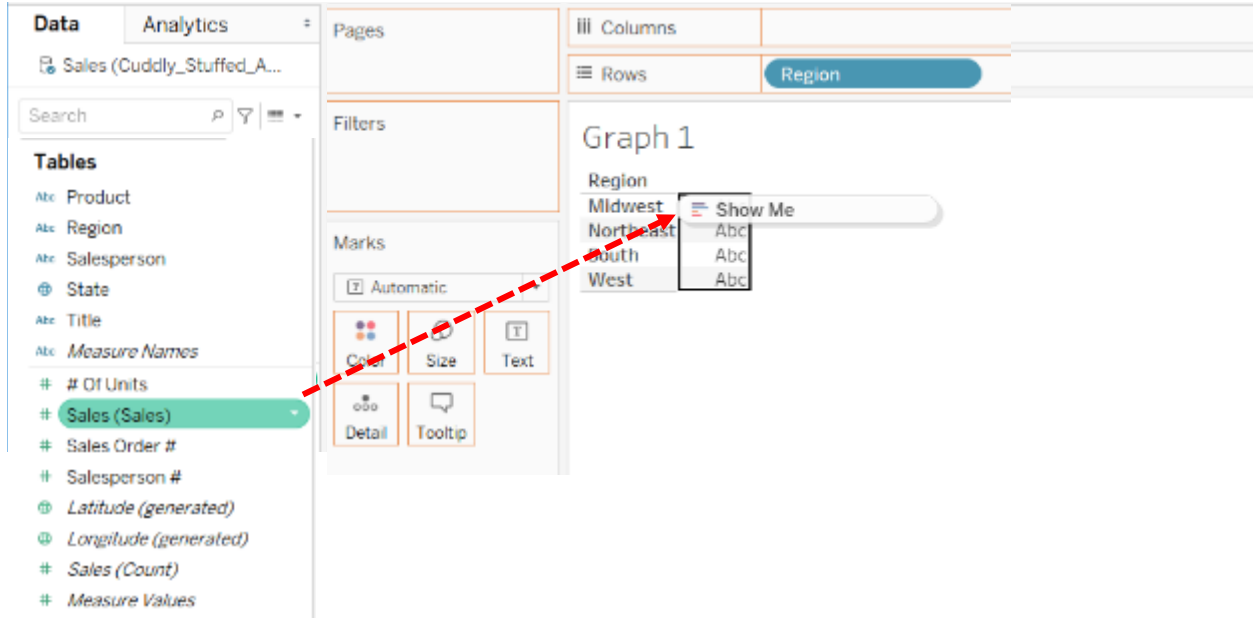
Select the Horizontal Bar Graph.
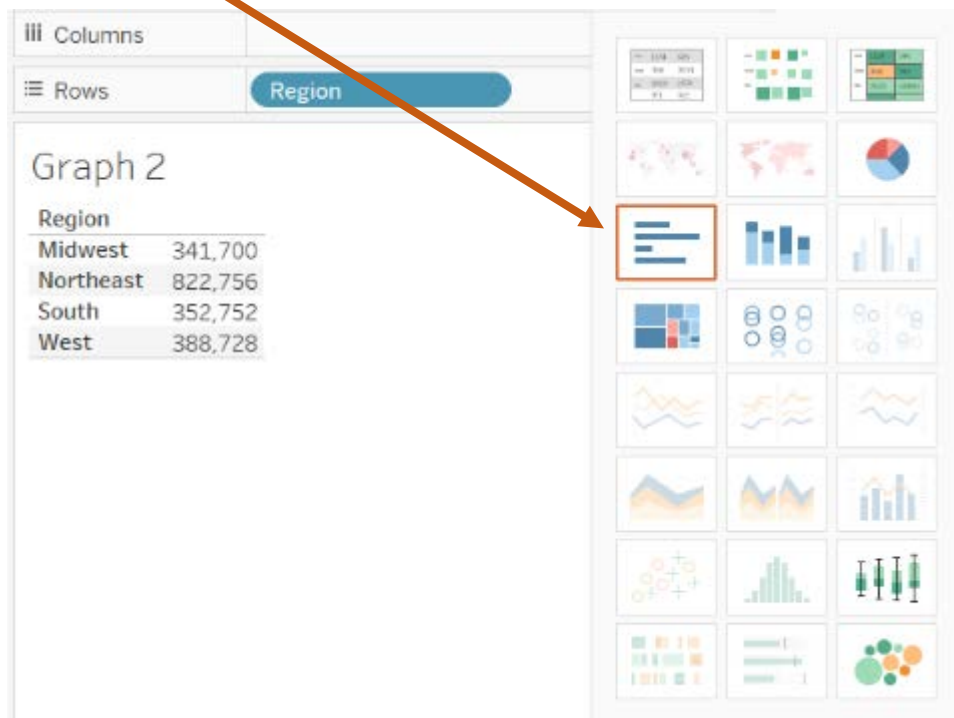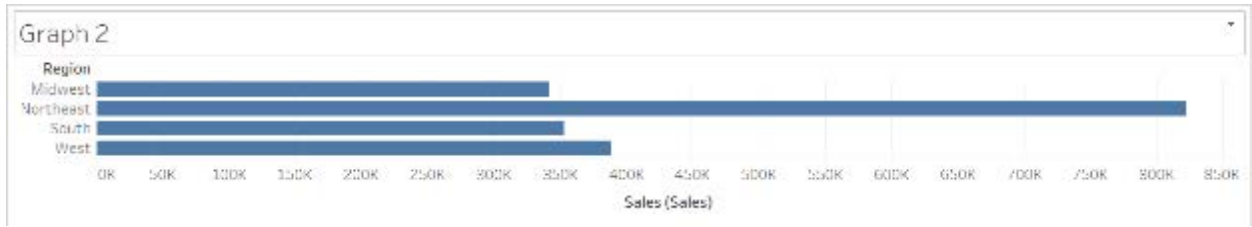
Select the "Fit" dialog box, and select "Fit Height".
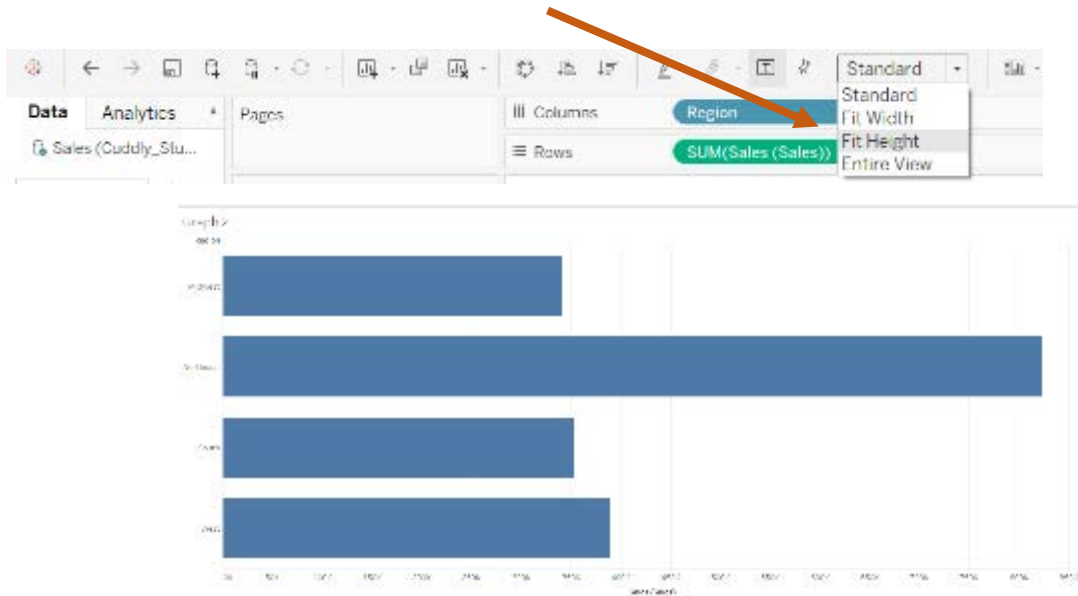
Select "Sales", and drag it to "Label".
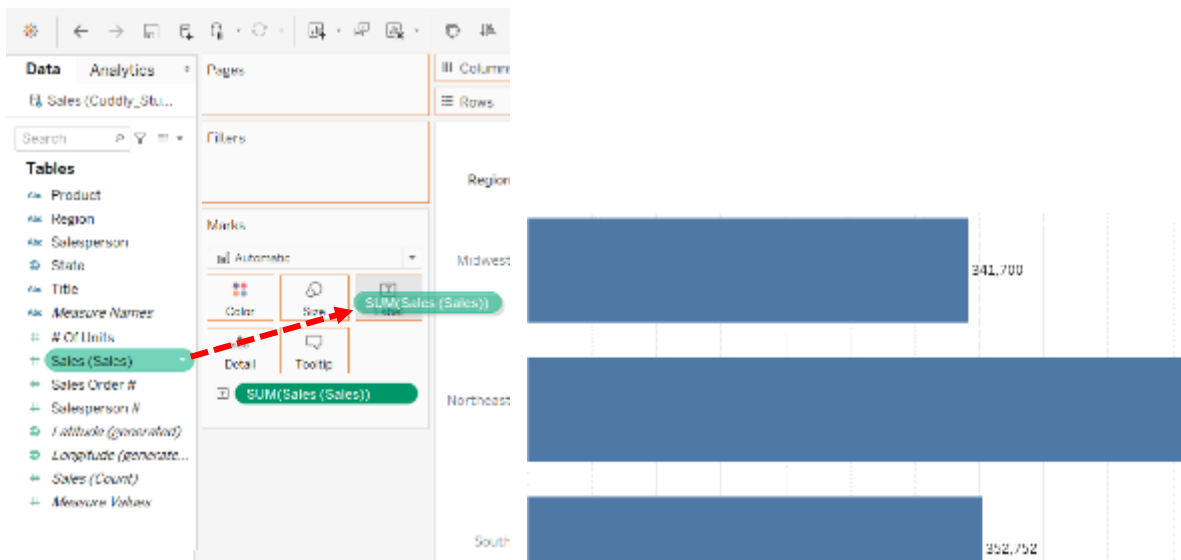
Add dollar signs to the labels by clicking on the drop-down on the right side of "Sum(Sales(Sales))".

Select "Format".
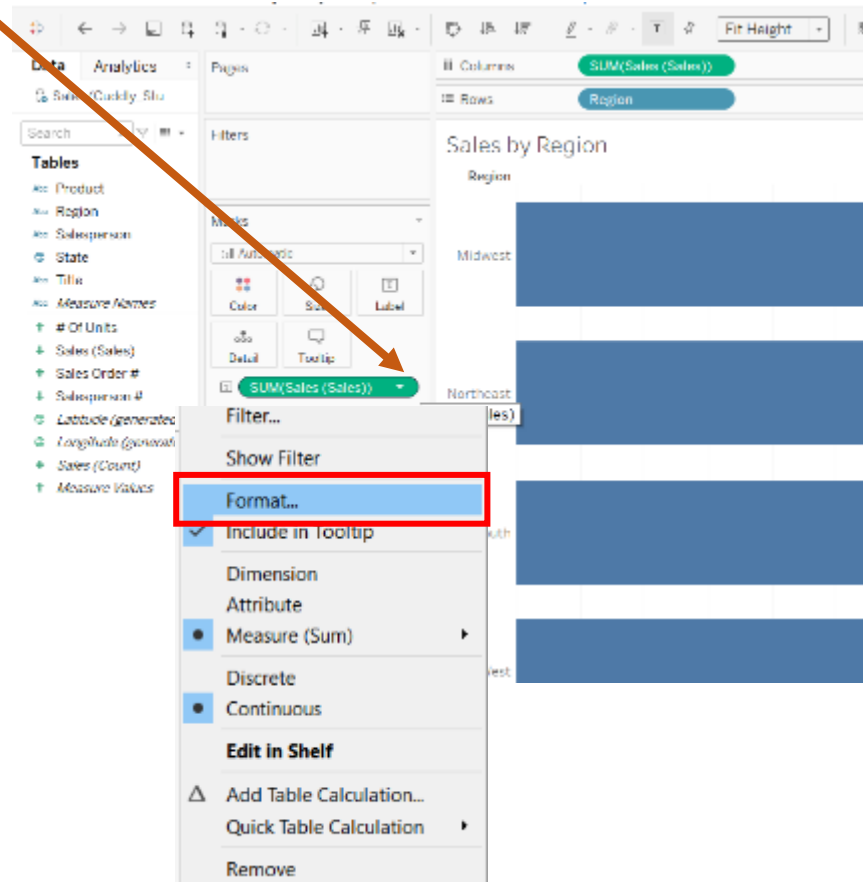
Click on the drop down on the right side of "Default", "Numbers".
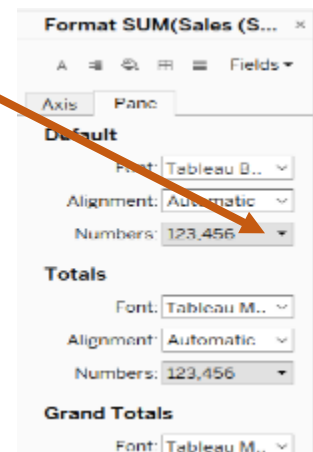
Select Currency (Custom), and change the number of decimal places to 0 decimal places.
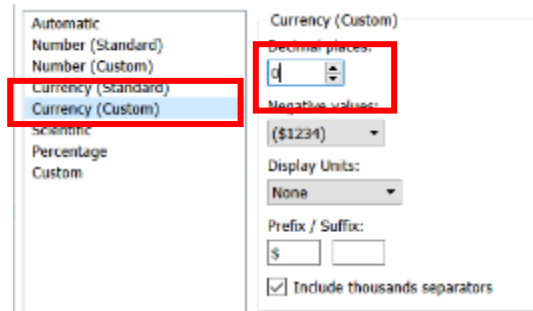
Close the Format window by clicking on the "x".

Graph 2

Region

Midwest $341,700

Northeast $822,756

South $352,752

West $388,728

0K 50K 100K 150K 200K 250K 300K 350K 400K 450K 500K 550K 600K 650K 700K 750K 800K 850K

Sales (Sales)
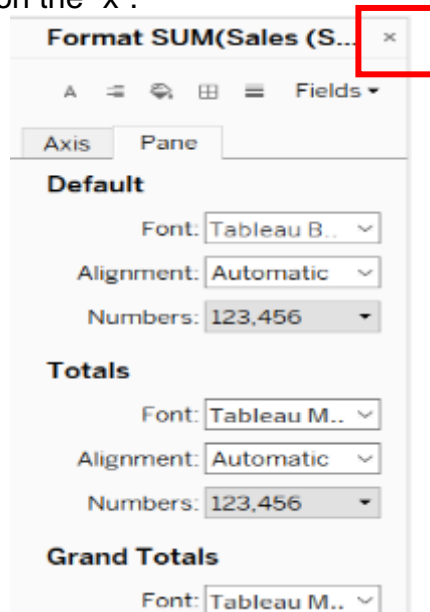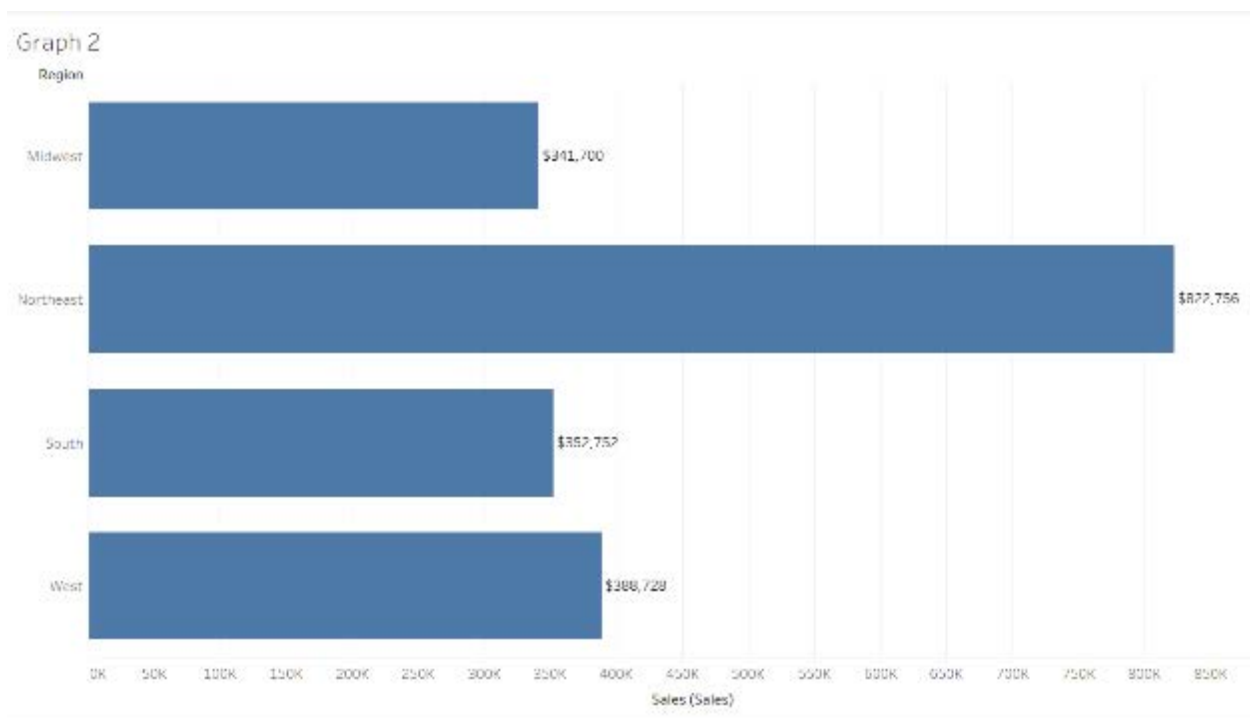
Now it is easier to visualize, Northeast had the highest sales, $822,756, and Midwest the lowest, $341,700.

The graph may look better if flipped. Select "Swap Rows and Columns".
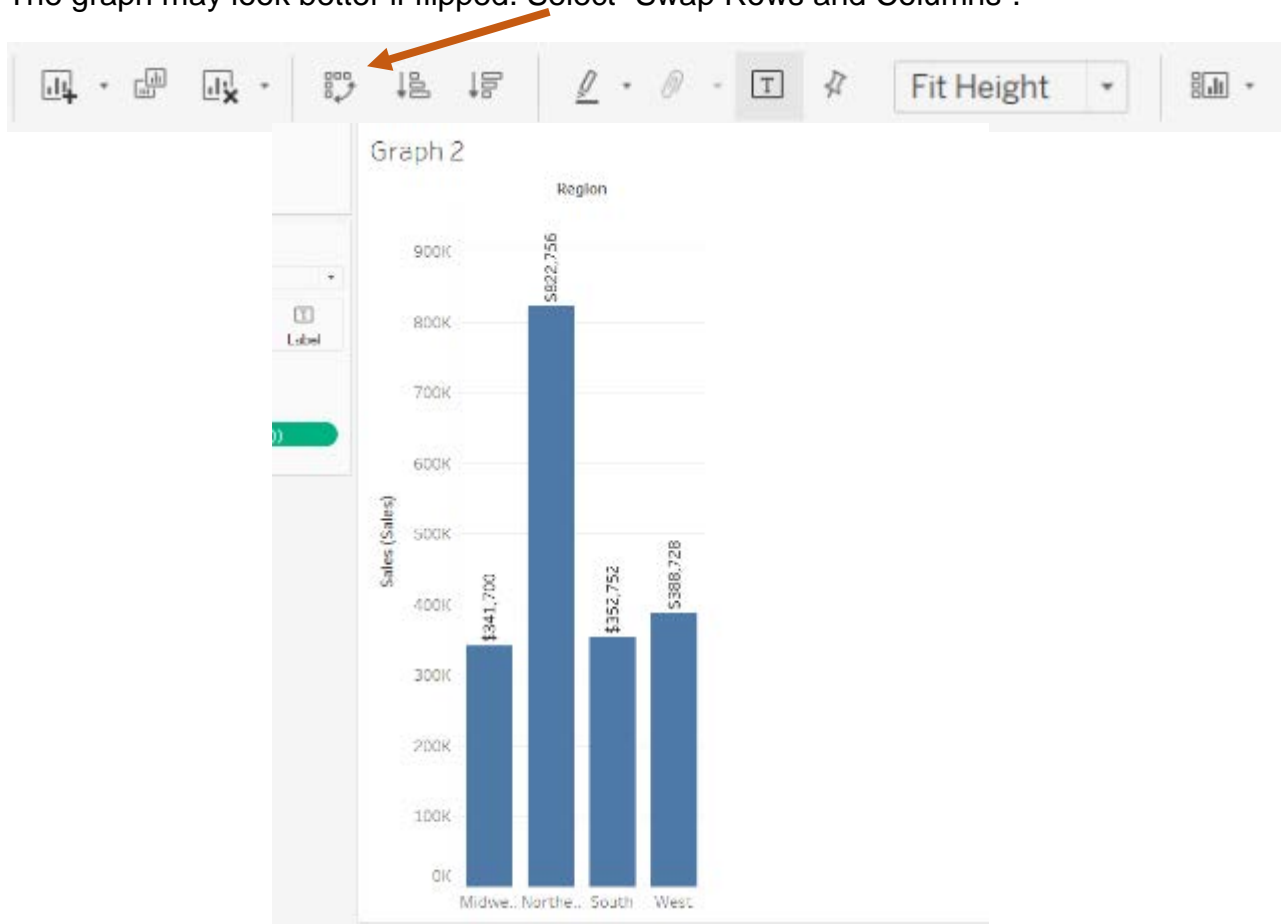
Change the "Fit" dialog from "Fit Height" to "Fit Width".
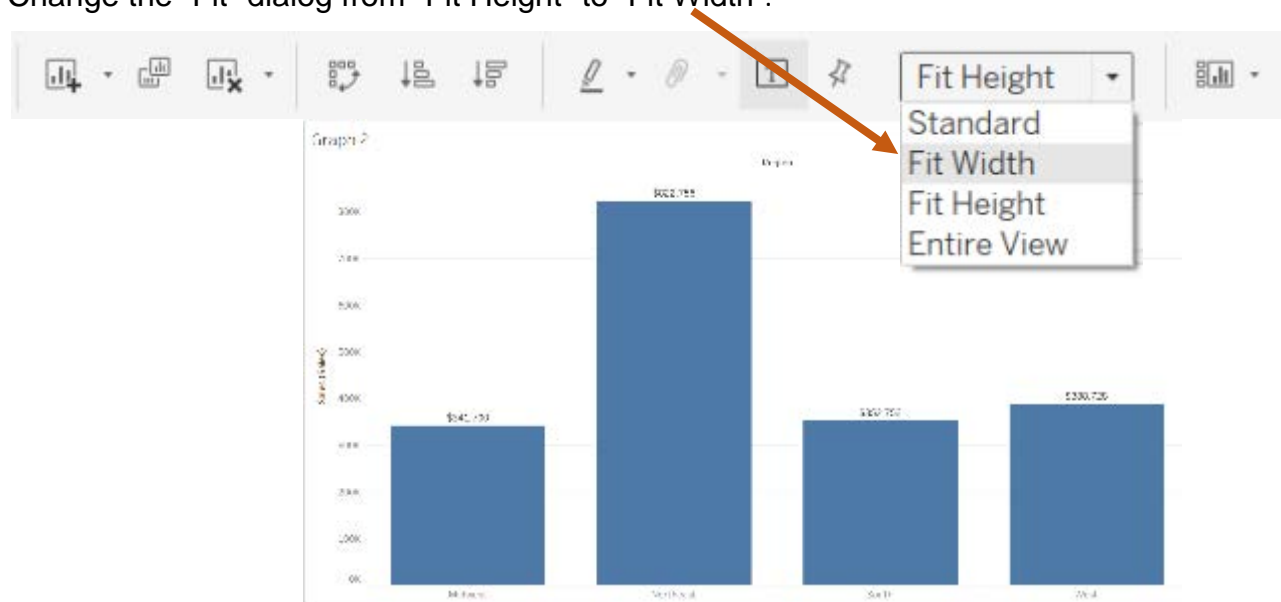
Double click on the Title of the graph, "Graph 2", and change it to "Sales by Region". Do **NOT** change the Sheet Name on the tab.
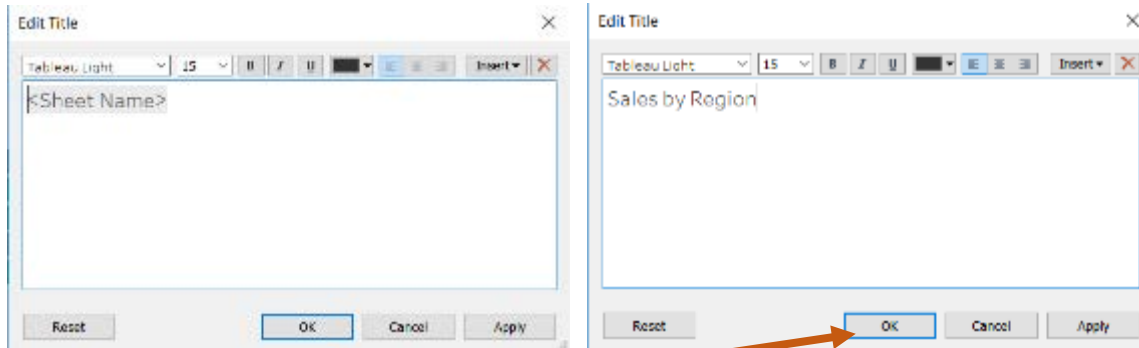
Select OK.

You must try different alternatives until you find a visualization that clearly communicates the story.

Prepare a visualization by adding "Product" to the graph.

The visualization "Sales by Region" is a great place to start. Select "Duplicate".
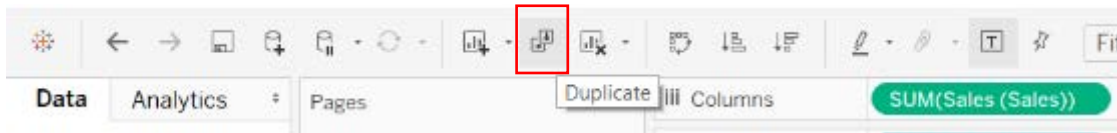
A duplicate worksheet Titled "Graph 2 (2)" is created.
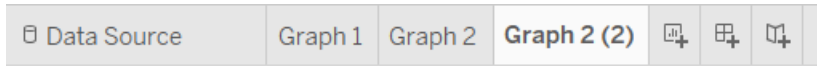
Rename the new worksheet "Graph 3".
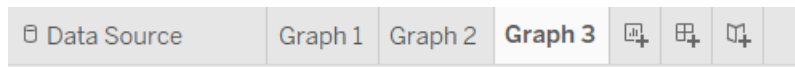
Select "Product", and drag it to the right of "Region" in the "Columns".
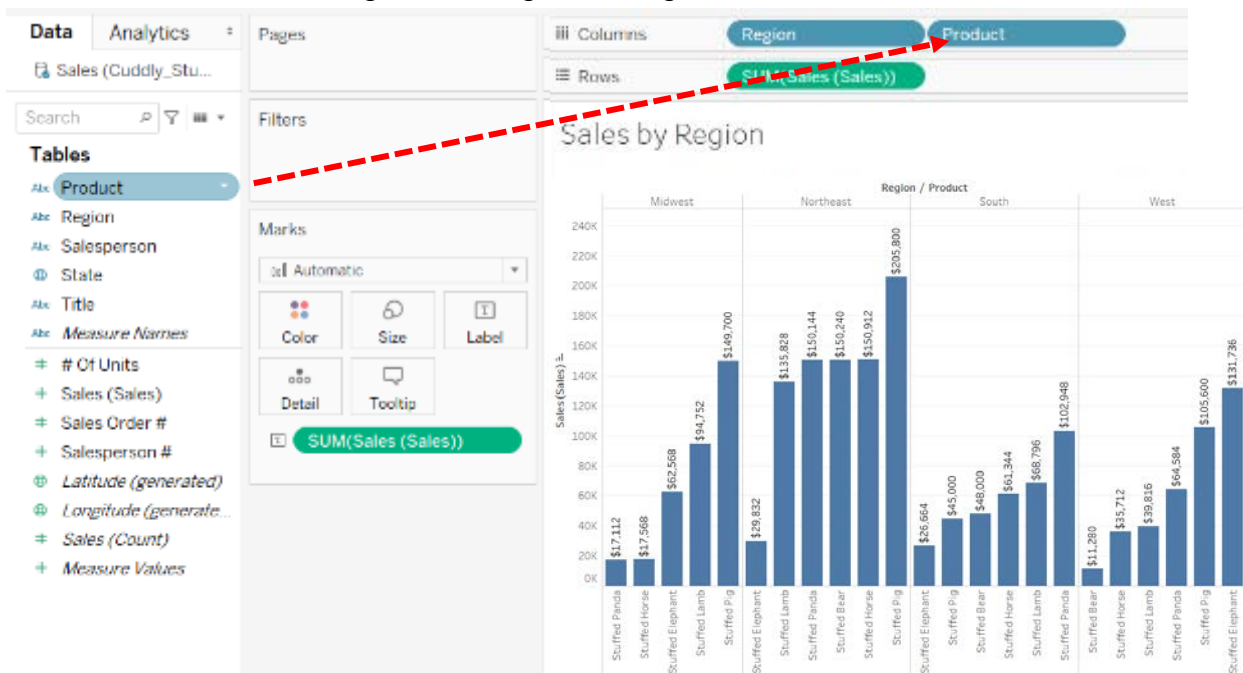
That looks very busy, so switch the "Row" and "Columns".
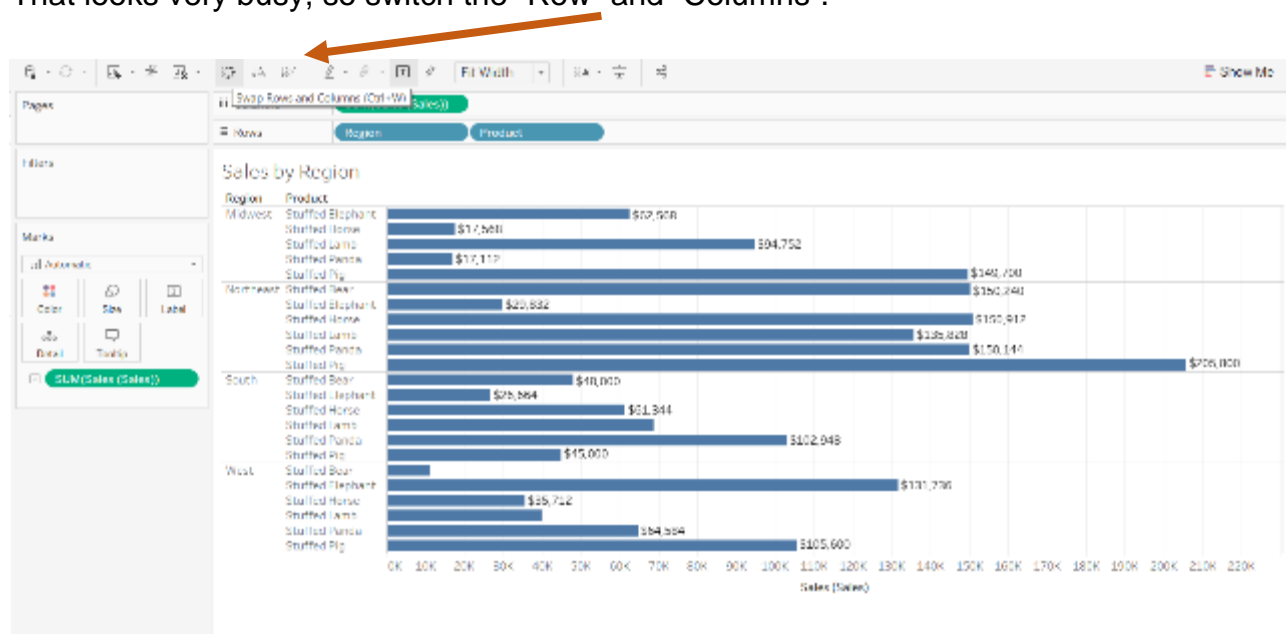
Change the "Fit" to "Fit Height".
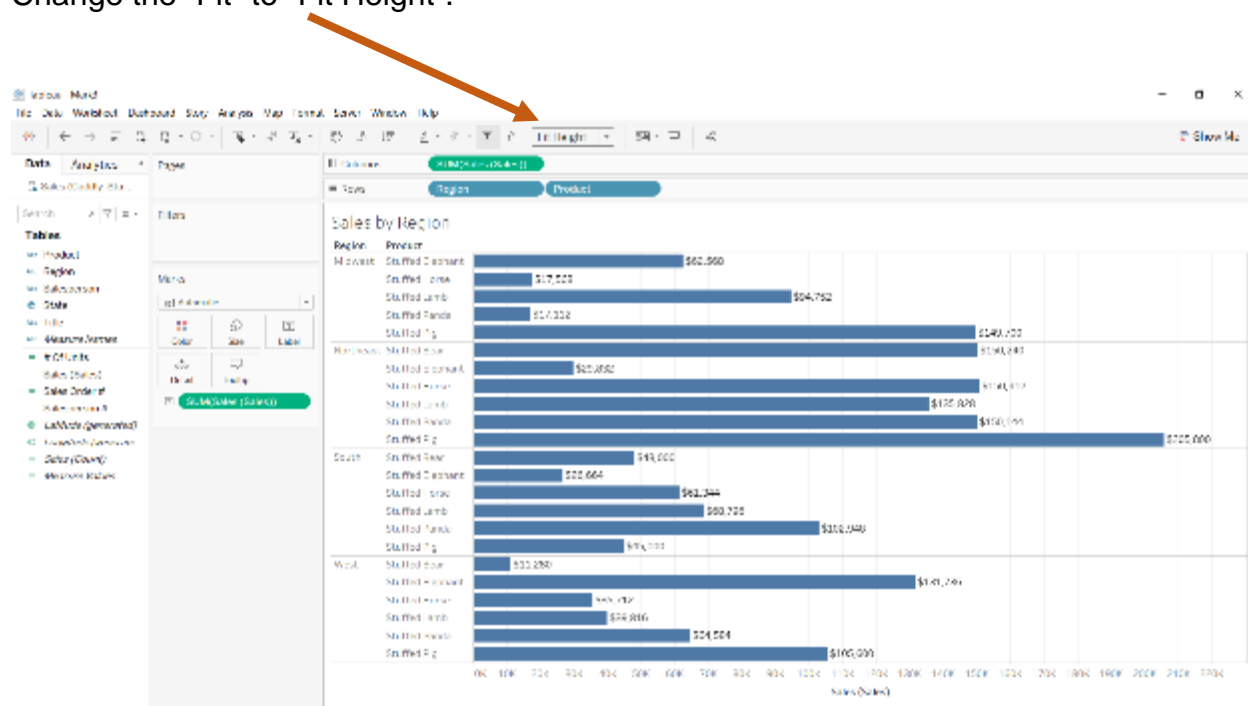
Change the visualization to "Region" within "Product", by dragging "Product" to the left of "Region" on the "Rows".
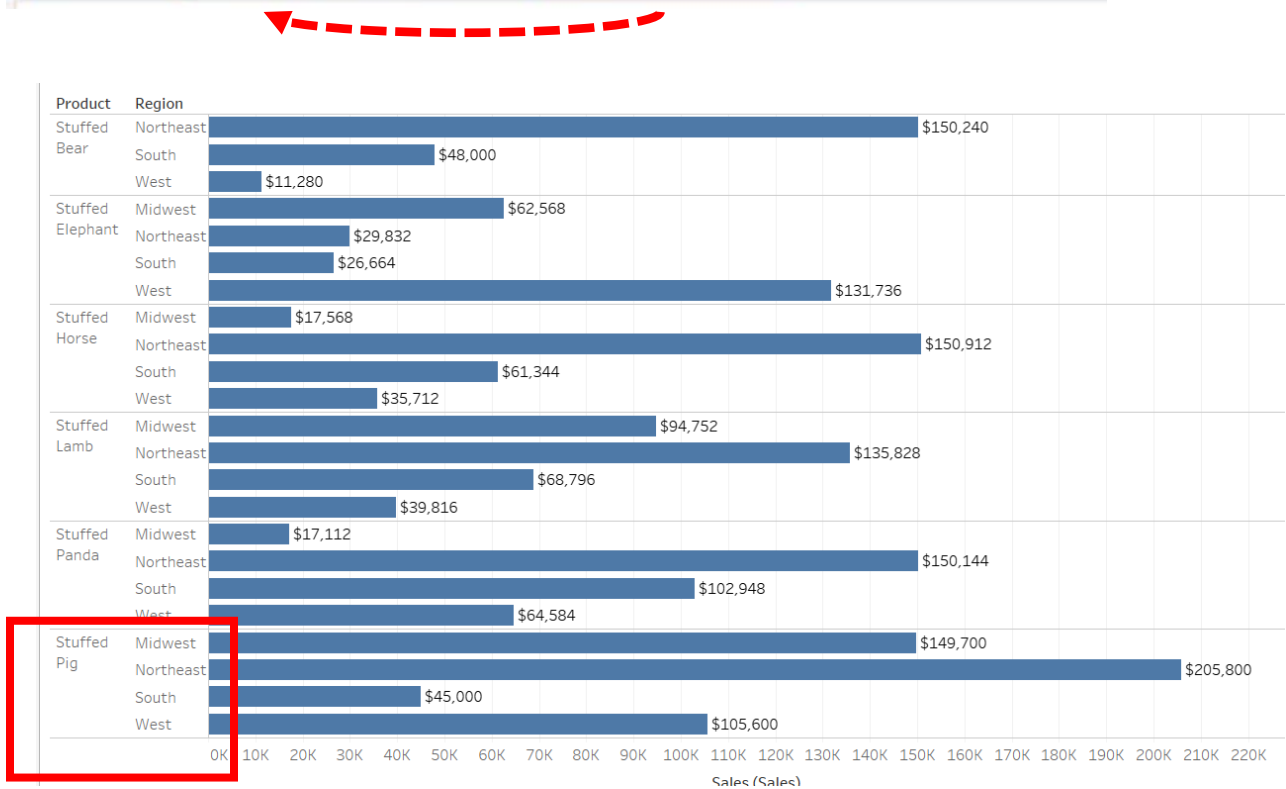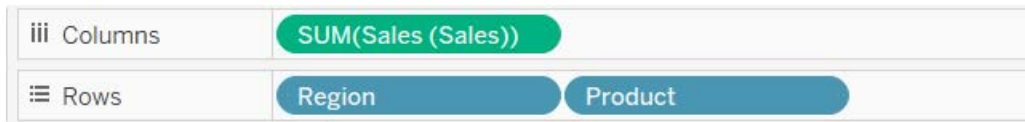
The Sales are thus presented alphabetically by "Region" within "Product".

Select "Sort Ascending".
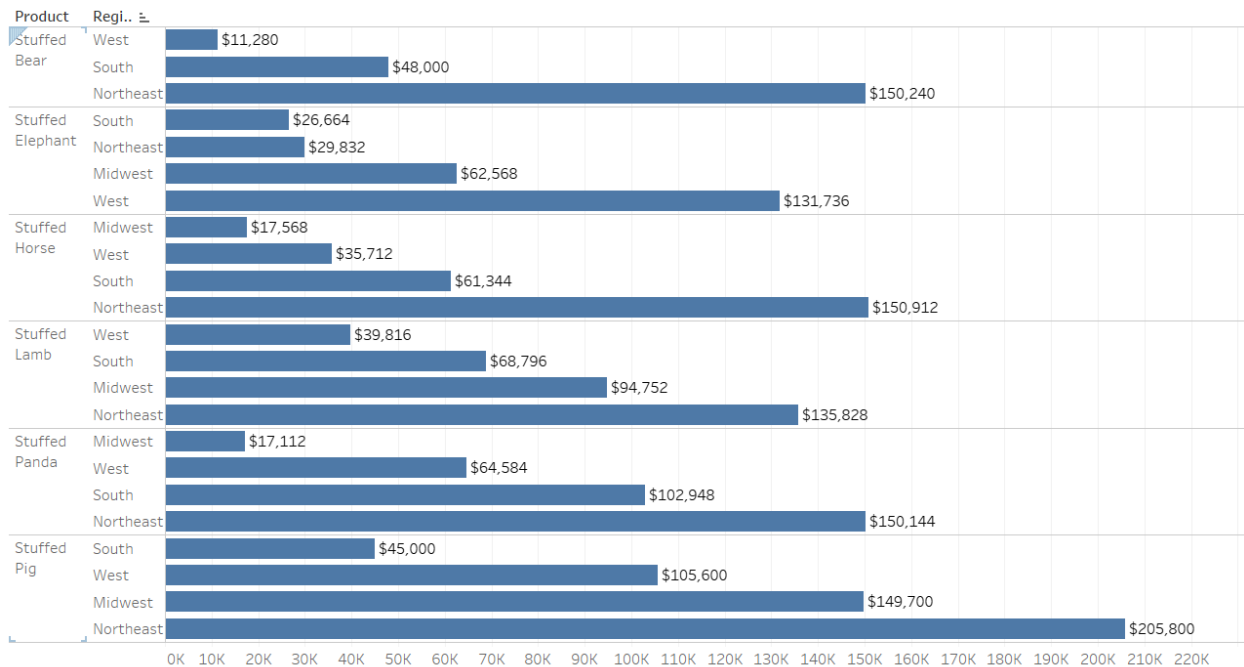
Now it is easier to visualize: the Stuffed Pig had the highest sales in dollars in the Northeast and the lowest in the South region.

Double click on the title, and change "Sales by Region" to "Sales by Product within Region" and select "OK". Do **NOT** change the Sheet Name on the tab.
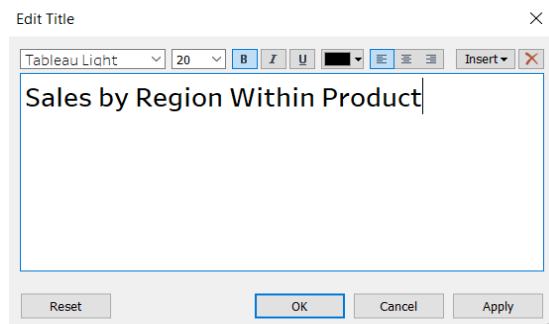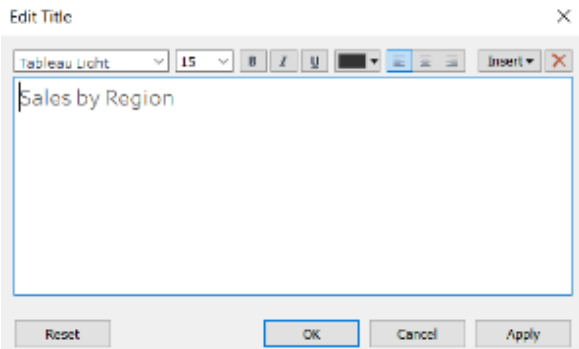
Prepare a visualization of the "Sales by State".

Select "New Worksheet", and rename the sheet "Graph 4".
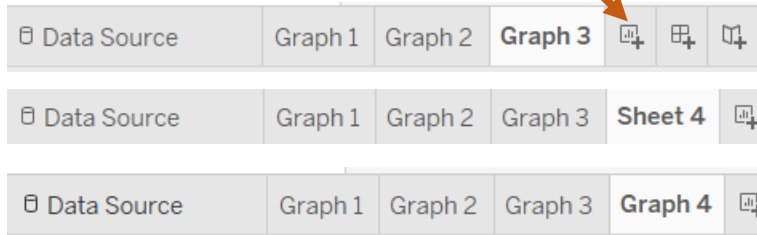
Drag "State" to "Rows".
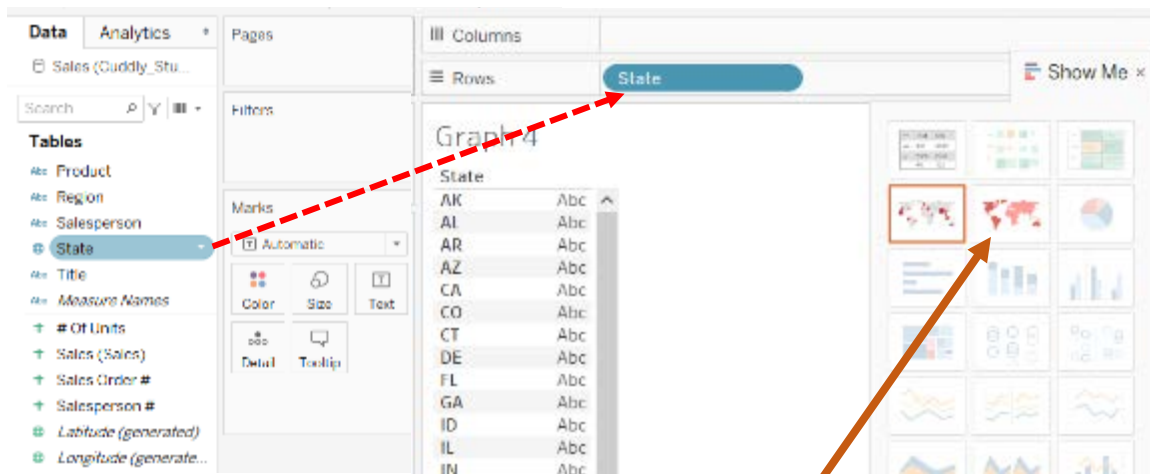
Select the "Filled Map" from the "Show Me" section.

Drag "Sales" to "Colors".
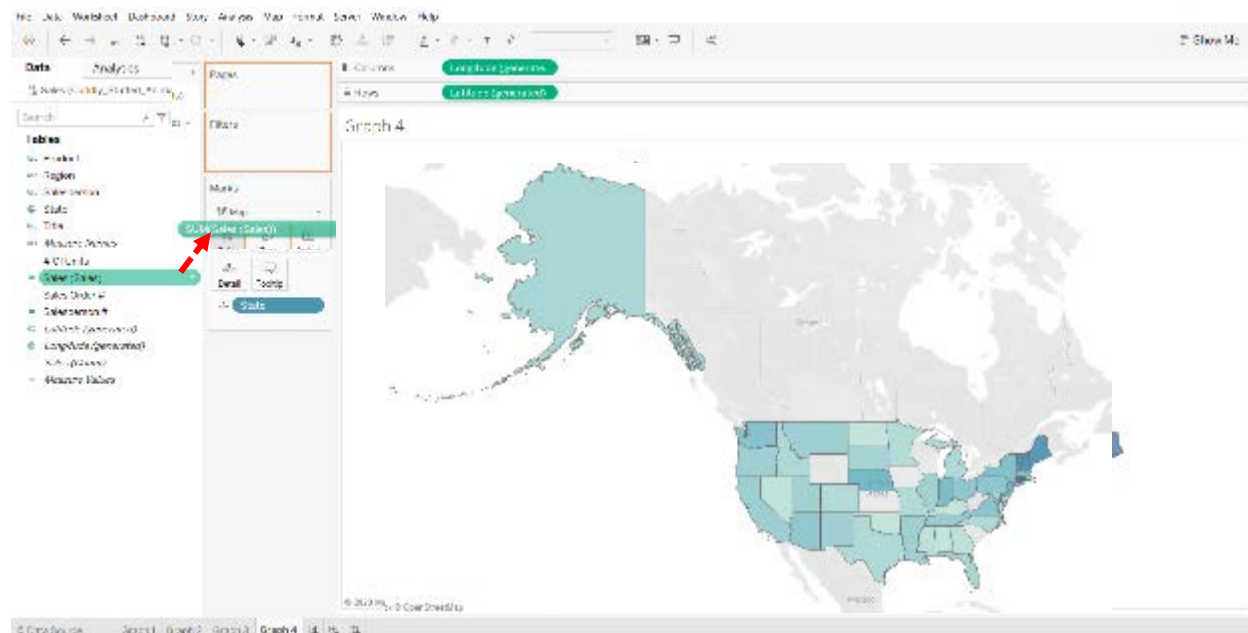
Select "Color".

Select "Edit Colors".
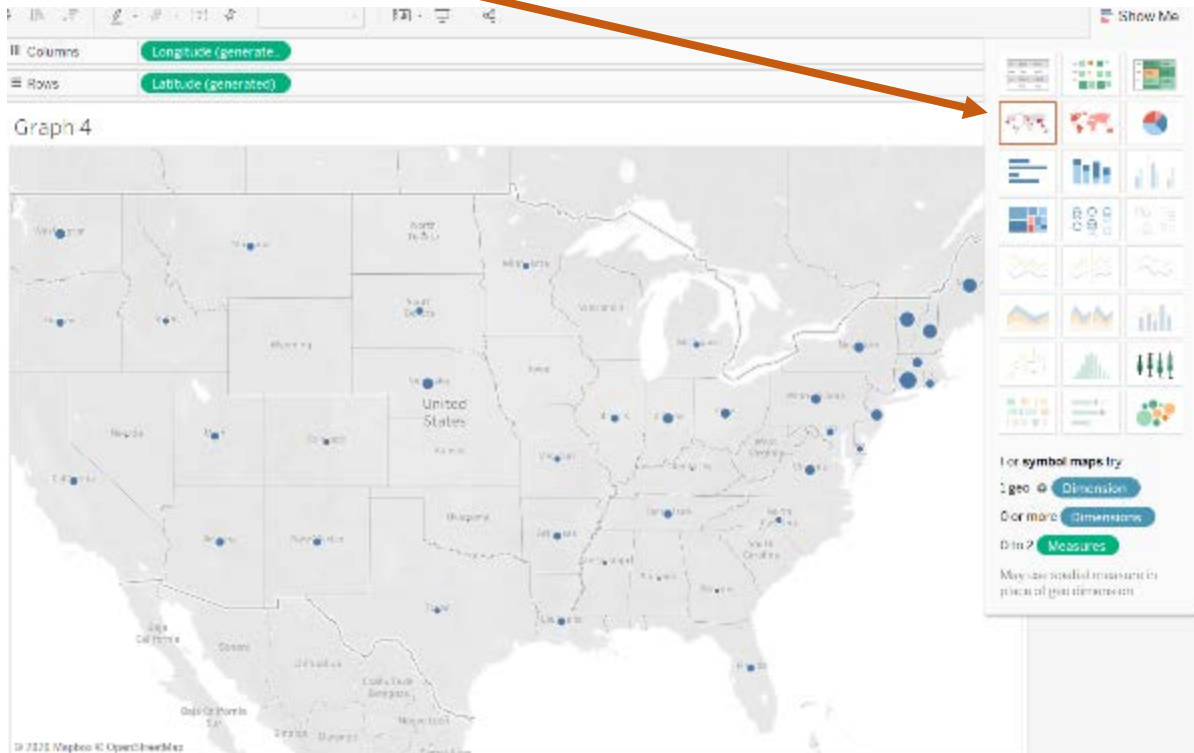
Select "Red-Black Diverging".

Select "OK"
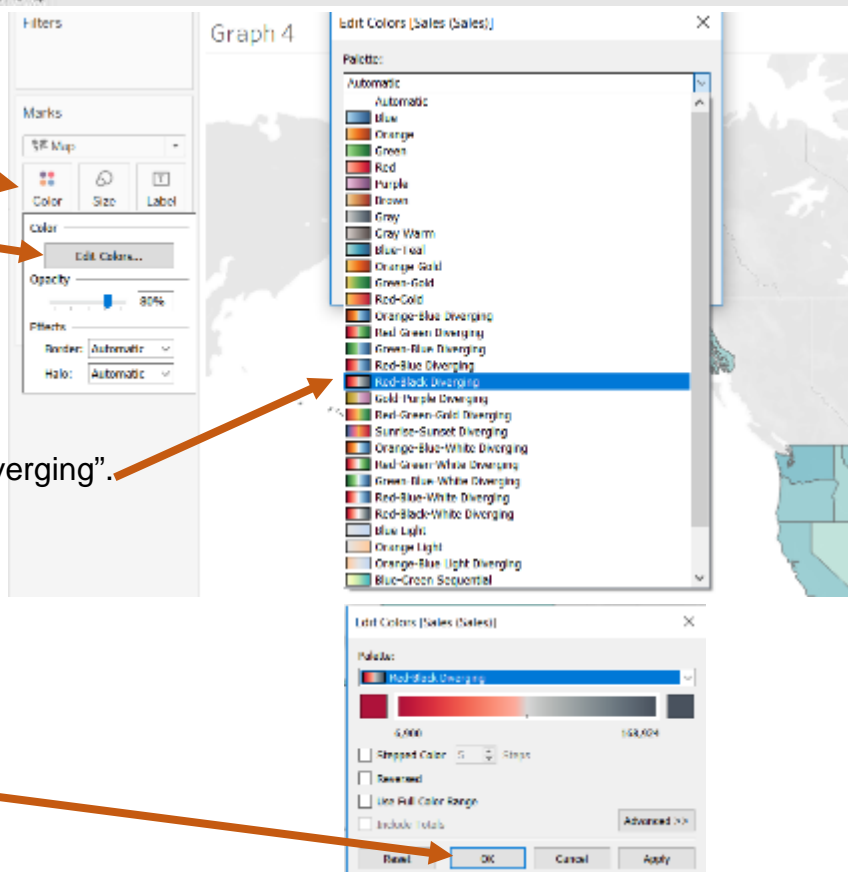
Click on Alaska and select "Exclude".

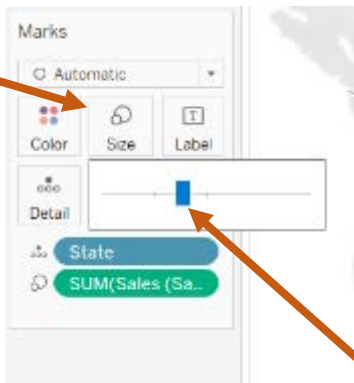When you hover over the darkest-colored state, you will observe that CT has the most sales.
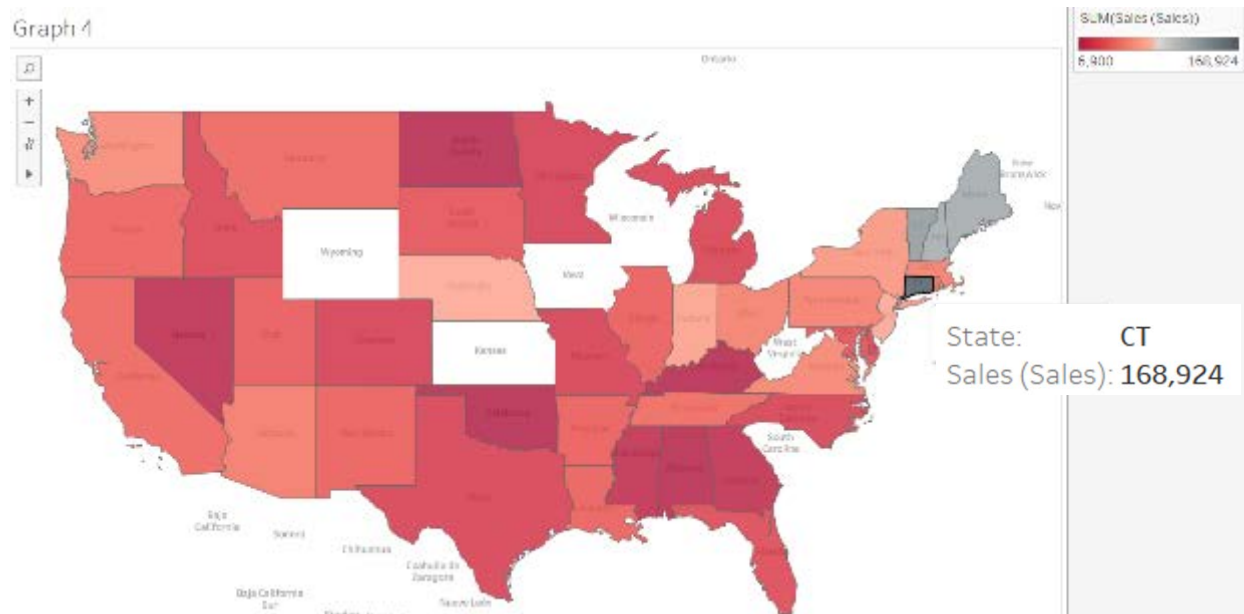
Let us look at the alternative map.
Select "Symbol Map".

Select "Size".

Move the indicator to the right to make the circles larger.

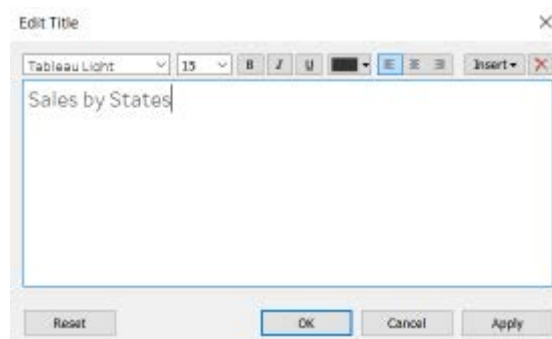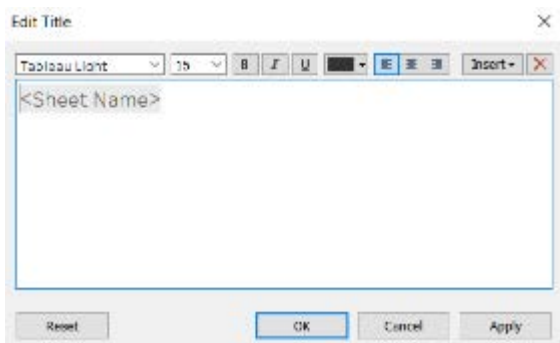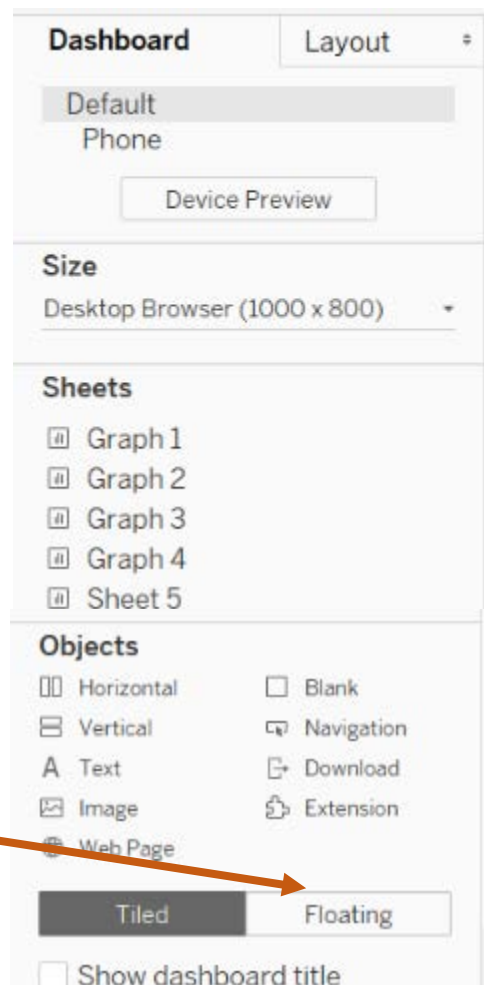This is a meaningful visualization.

Double click on the Title of the graph and change it to "Sales by States", and select OK. Do **NOT** change the Sheet Name on the tab.

Create a "Dashboard".

Select "New Dashboard".

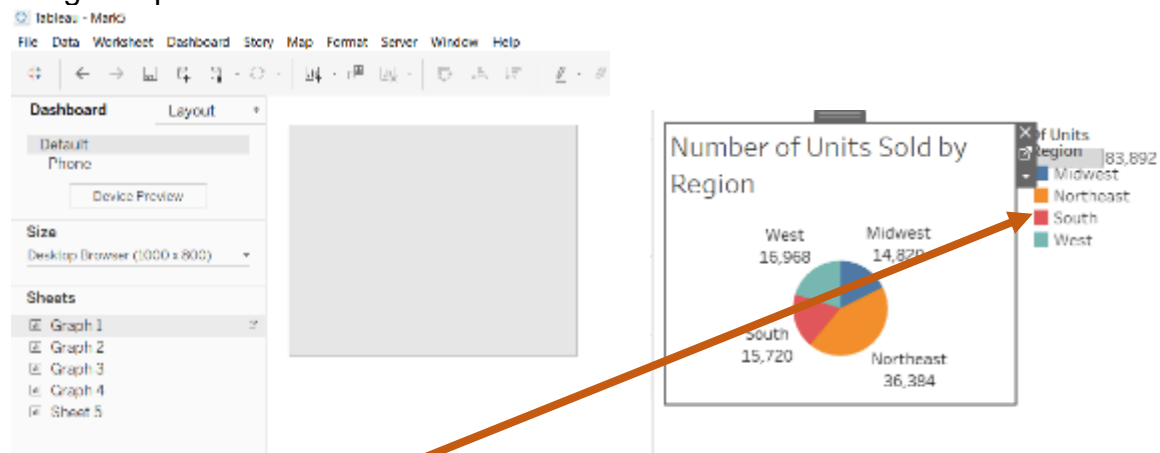| ⊟ Data Source	Graph 1	Graph 2	Graph 3	**Graph 4**	Sheet 5
42 marks	1 row by 1 column	SUM(Sales (Sales)): 1,877.436		New Dashboard	

Dashboard   Layout ⬦

Default
  Phone

Device Preview

Size
Desktop Browser (1000 x 800) ▾

Sheets
▦ Graph 1
▦ Graph 2
▦ Graph 3
▦ Graph 4
▦ Sheet 5

Objects
▯▯ Horizontal   ▢ Blank
⊟ Vertical   🖵 Navigation
A Text   ⬚ Download
🖾 Image   ⬡ Extension
⊕ Web Page

Tiled	Floating

☐ Show dashboard title

Change the mode to "Floating".
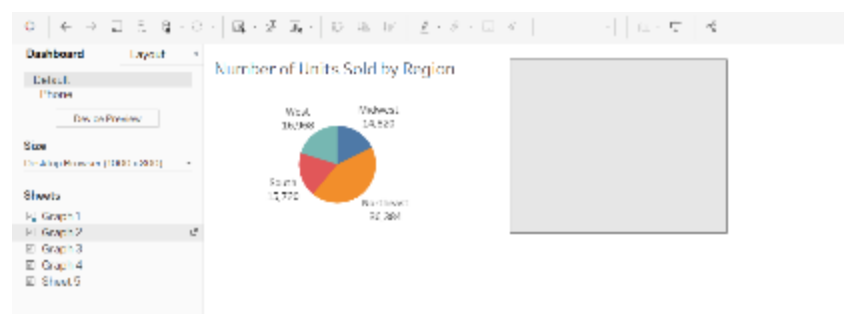
Drag "Graph 1" into the "Dashboard".

remove the legend and size "Graph 1".
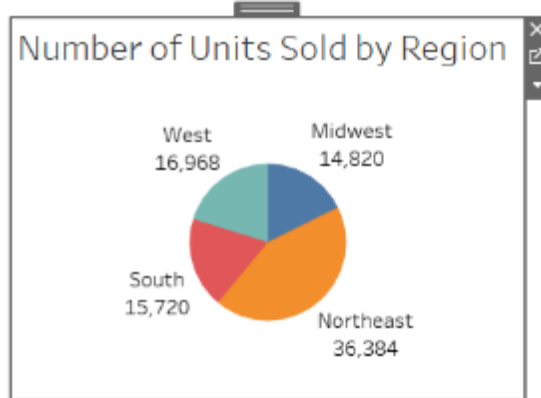
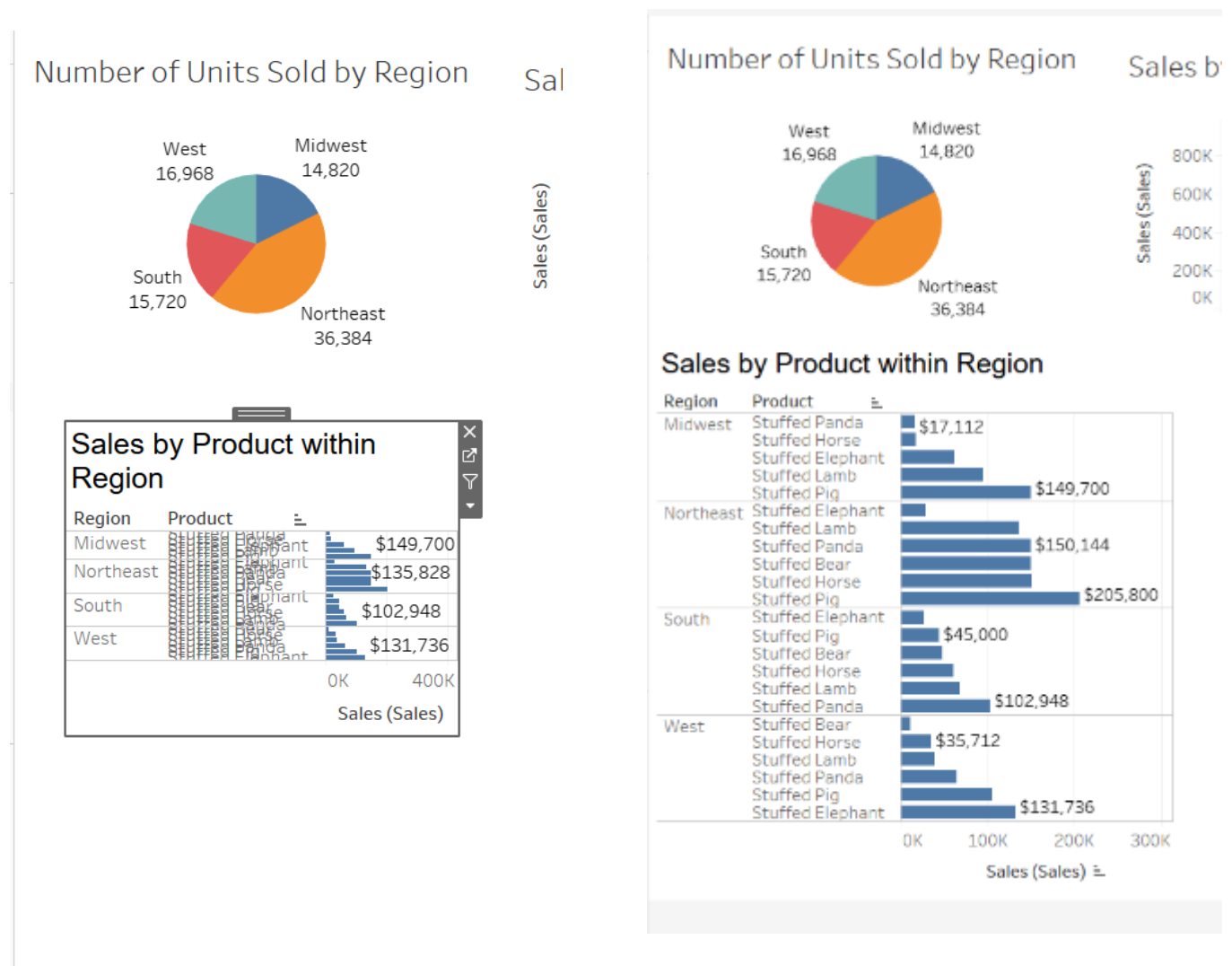
Drag "Graph 2" into the dashboard to the right of "Graph 1"
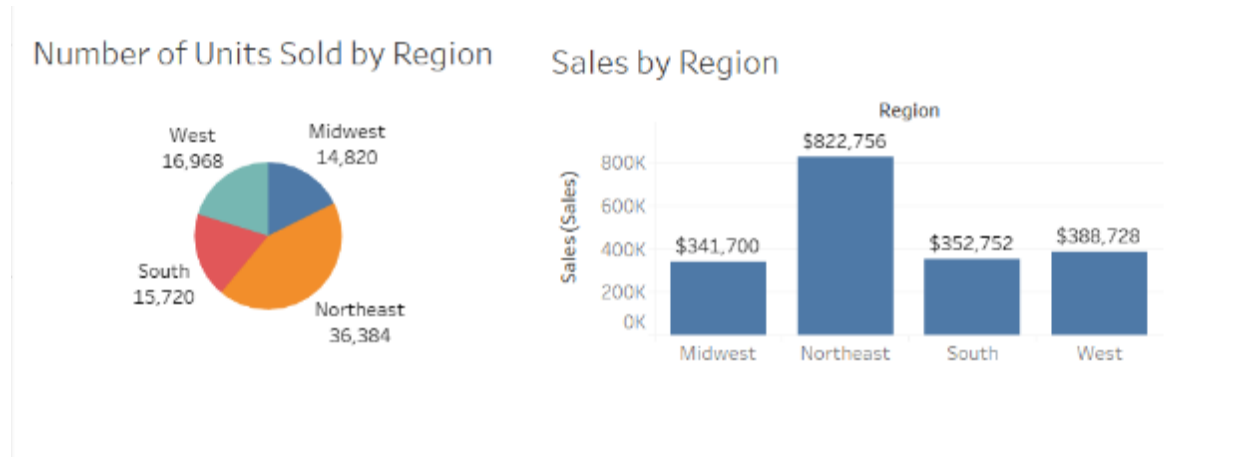
Size Graph 2

Drag "Graph 3" into the dashboard under "Graph 1" and size "Graph 3"

Drag "Graph 4" into the dashboard to the right of "Graph 3" and size "Graph 4".
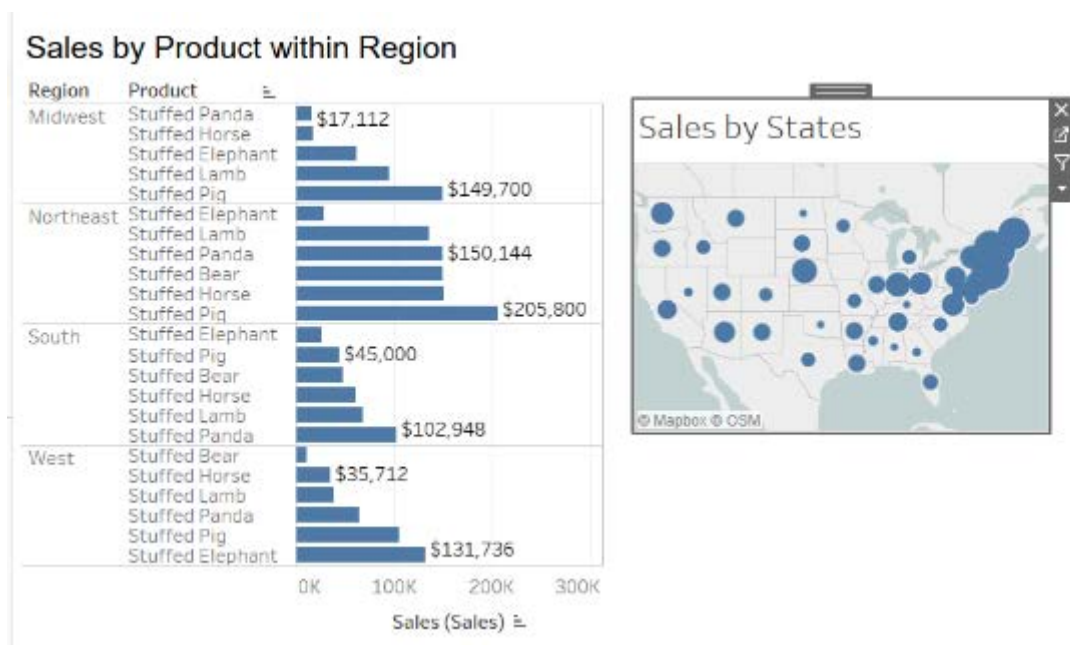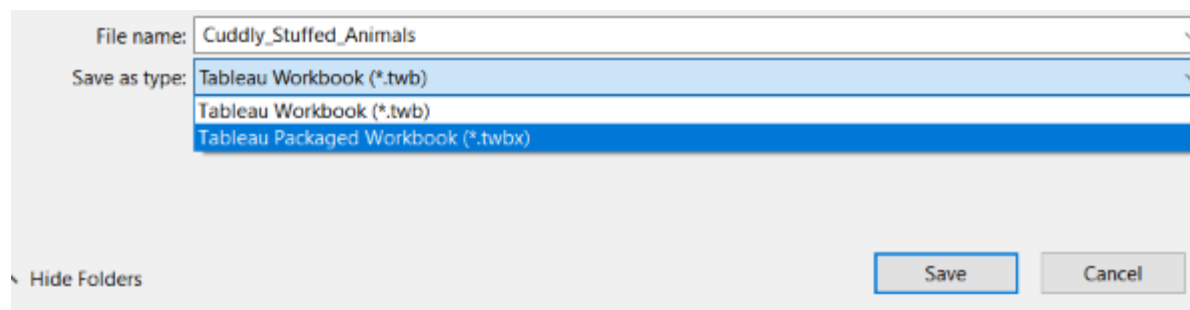
The Dashboard is completed
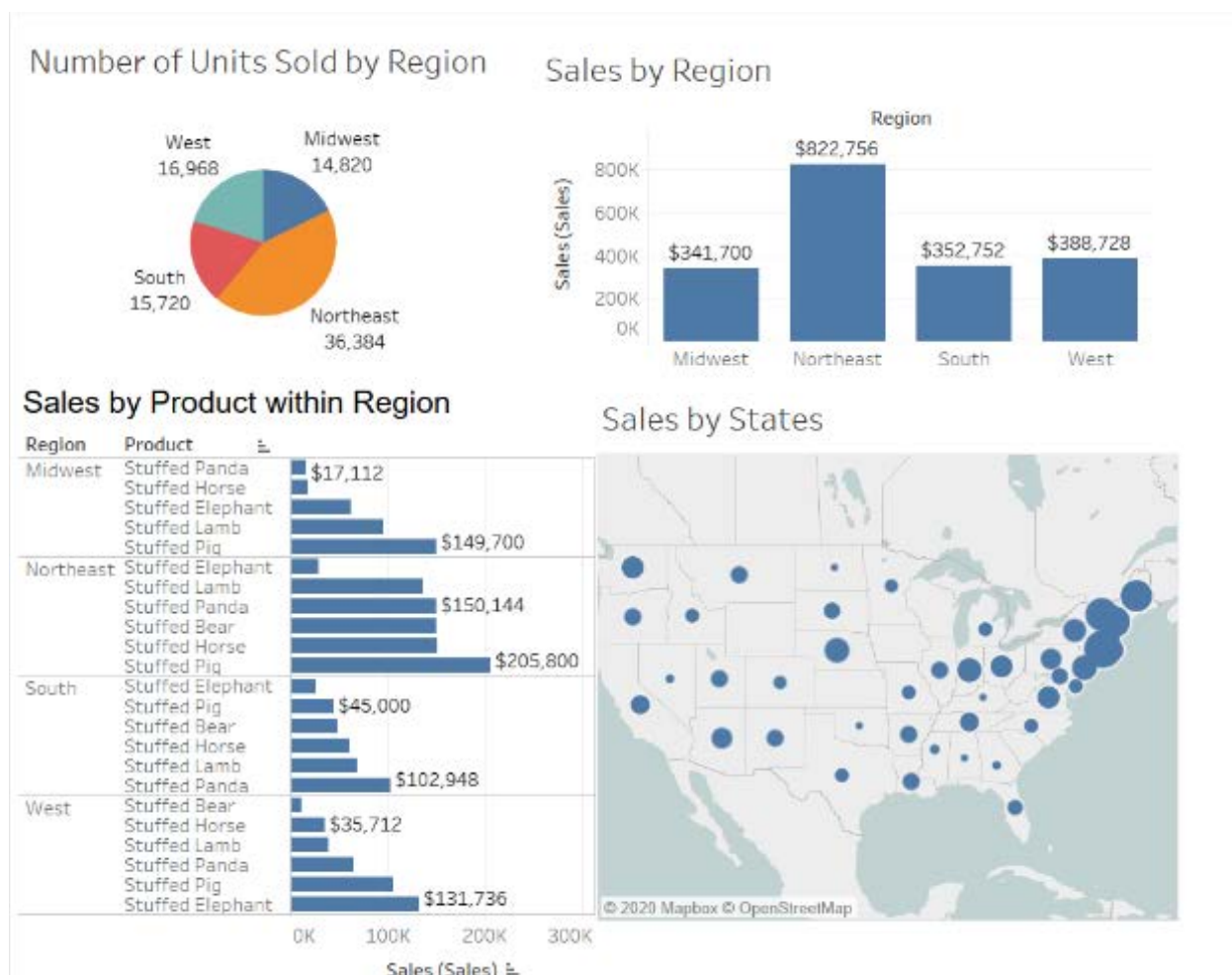
Select File>Save As

Save As a Tableau Packaged Workbook (*.twbx) so that the data is connected to the Tableau Visualization.