

## Lab 9. Tableau Advanced Visualization



In this lab, we will cover the following recipes:

- Lollipop charts
- Sankey diagrams
- Marimekko charts
- Hex-Tile maps
- Waffle charts

### Technical requirements

We will use Tableau 2019.x, and data about mass transit complaints, beer, education, commuter times, and Lord of the Rings.

### Introduction

In this lab, we will learn more chart types that go beyond **Show Me**. We will use external data sources to provide a plotting frame and, in many cases, use advanced calculations to draw our visualizations. In these recipes, we will show how data elements relate to each other or how they are related to the whole.

### Lollipop charts

Lollipop charts get their name from their shape. These charts can offer visual variety. We would typically use lollipops in place of bar charts or dot plots, which also happen to be construction components.

#### Getting ready

We will learn how to successfully combine a bar chart and a dot plot to create a lollipop chart.

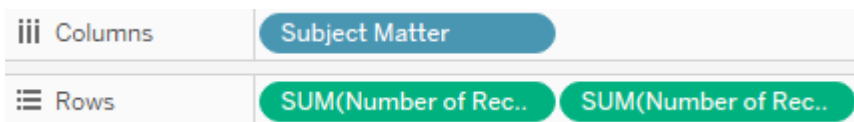
#### How to do it..

We will open the packaged workbook lollipop to work through the recipe. We will examine the NYC mass transit customer complaints by complaint subject matter; use `MTA_Customer_Feedback_Data_Beginning_2014.csv` and `MTA_complaints.twbx`:

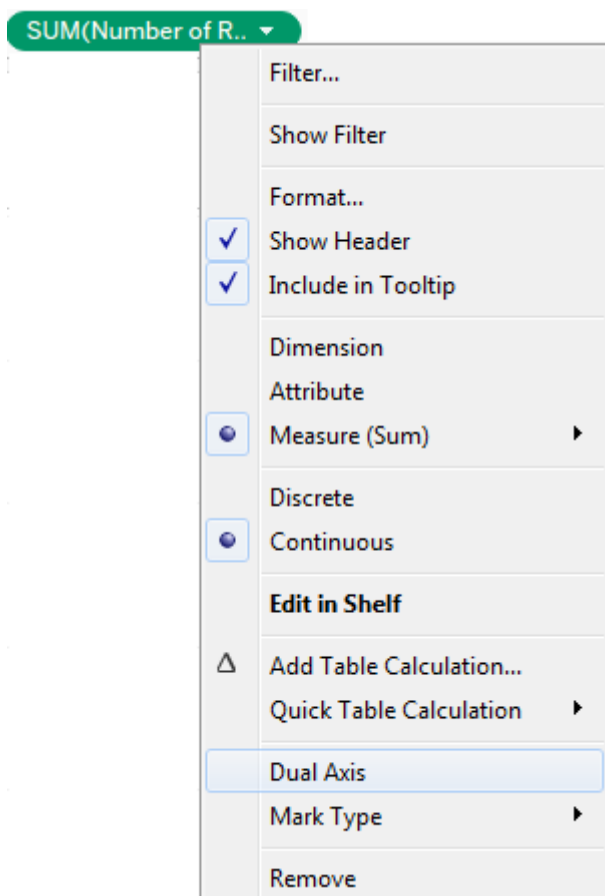
1. Place **Subject Matter** and **Number of Records** on the **Columns** and **Rows** shelves:



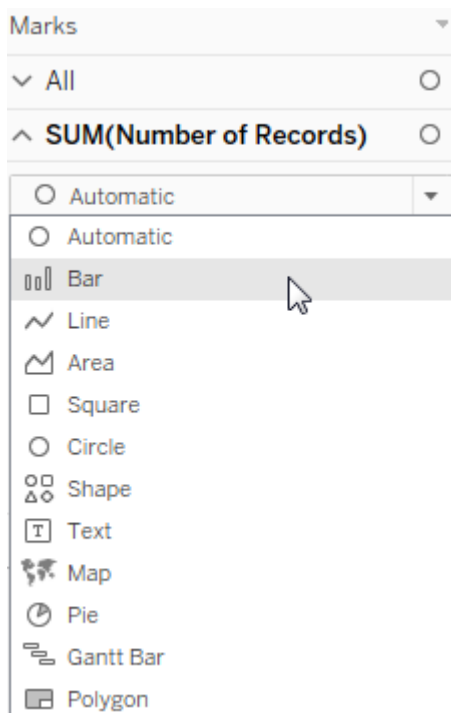
2. Duplicate the **Number of Records** pill on the R\*\*ows\*\*shelf:



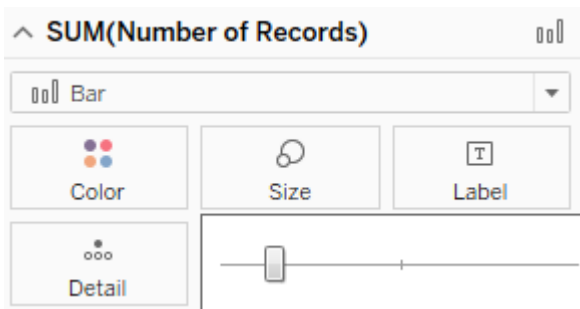
3. Modify the chart to be **Dual Axis**:



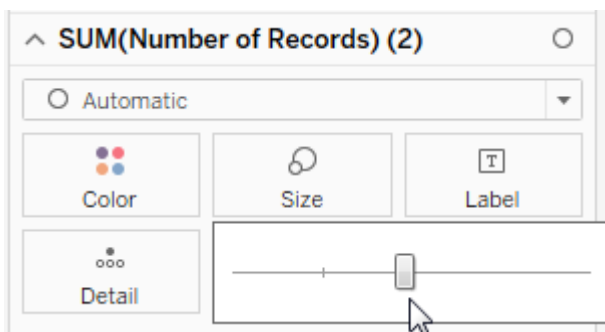
4. Change the mark type of `Sum (Number of Records)` to `Bar` :



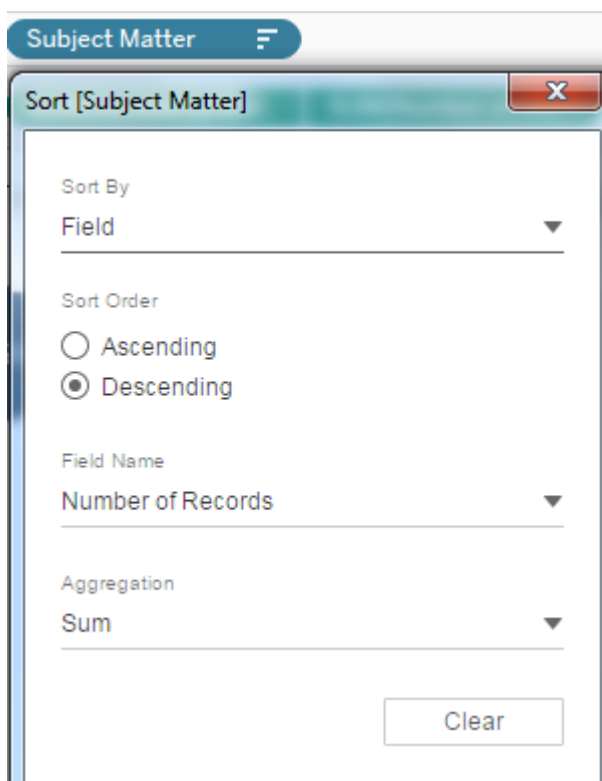
5. Make the **Size** of the bar smaller:



6. Make the **Size** of the circle larger:

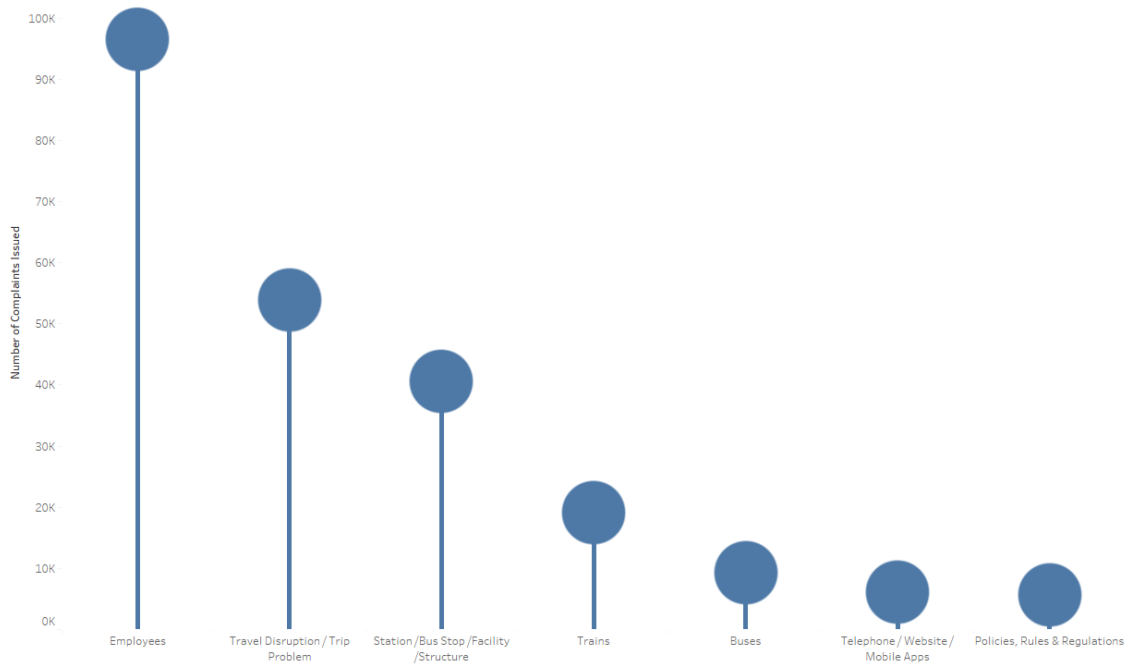


7. Sorting **Subject Matter** in descending order by **Number of Records** :



We now have the following example:

NYC, Mass Transit Customer Complaint Subject Areas



### Note

Because there were many different **Subject Matter**, we filtered **Subject Matter** to the top seven based on the **Number of Complaints** issued.

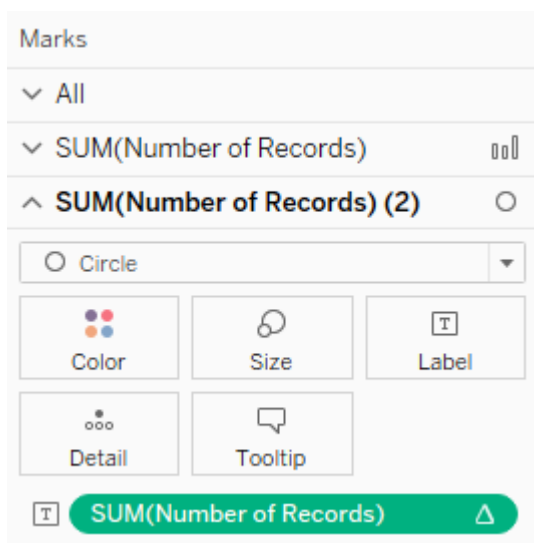
### How it works...

We placed the **Subject Matter** on the **Columns** shelf and **Number of Records** on the **Rows** shelf twice. Next, we made this a **Dual Axis** chart. We modified the mark type to be **Bar** and made the size smaller for the first **Number of Records**. We left the mark type for the second **Number of Records** as **Circle** and increased the **Size**. Finally, we combined a skinny bar chart with a dot plot chart.

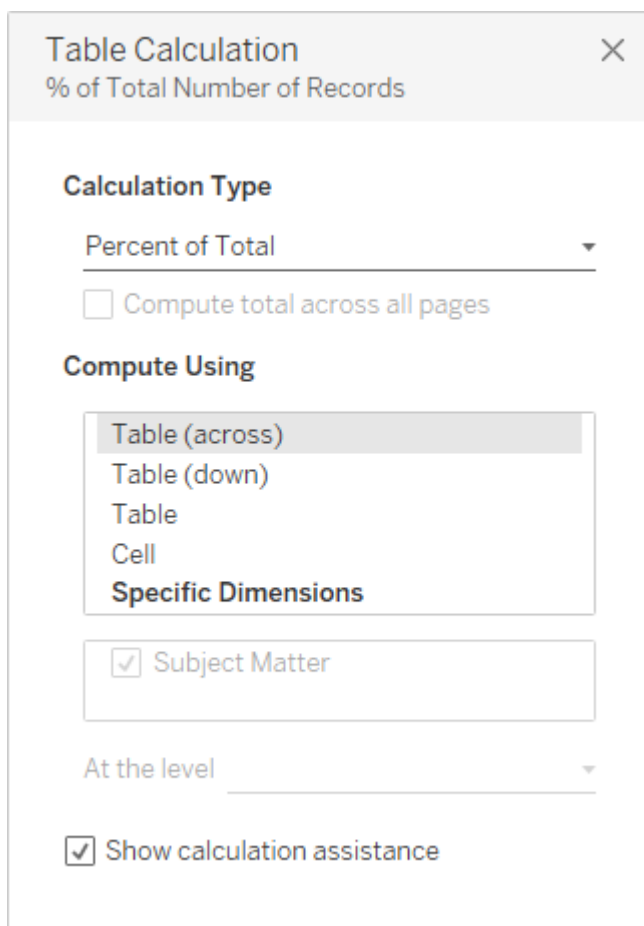
### There's more...

We can place a value in the center of the dot plot:

1. Place **Number of Records** onto the **Marks** card to create **Number of Records 2** :

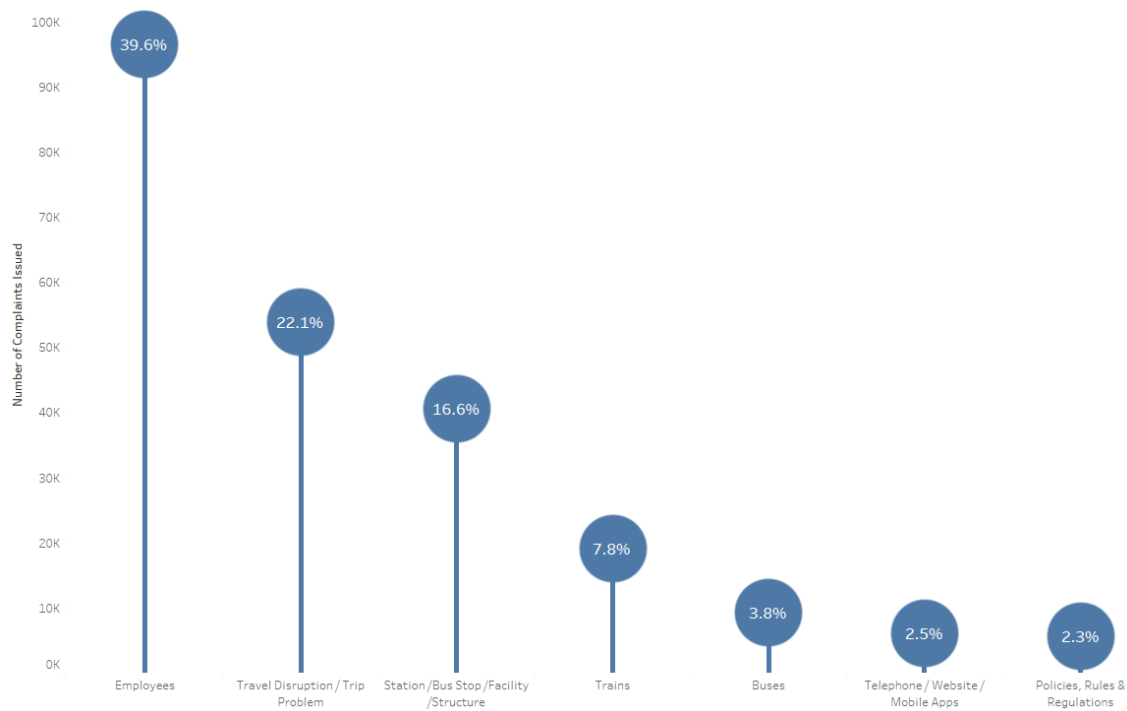


2. Right-click on **SUM(Number of Records)**, choose **Add Table Calculation**, and choose **Percent of Total**.



3. Format the text size, alignment, and color in order to create a similar chart:

NYC, Mass Transit Customer Complaint Subject Areas



## Sankey diagrams

A Sankey diagram is a tool used to show the flow from one dimension to another.

### Getting ready

In this recipe, we will create a Sankey diagram using a web extension provided by Infotopics. Web extensions are a new feature in 2018.2.

### How to do it..

In 2018.2, we can use web extensions to build a Sankey chart. To follow along, use `Beer 2018.2.twbx`, `beers.csv`, and `breweries.csv`.

1. Create the `Select Dimension Left` parameter:

Name:  Comment >>

Properties

Data type:

Current value:

Display format:

Allowable values: ☐ All ☒ List ☐ Range

List of values

Value	Display As
1	City
2	State
3	Style
4	Serving Size
5	Brewery Name
6	Beer Name
Add	

Add from Parameter ▶  
Add from Field ▶  
Paste from Clipboard  
  
Clear All

OK Cancel

2. Create the `Select Dimension Right` parameter:

Name:  Comment >>

Properties

Data type:

Current value:

Display format:

Allowable values: ☐ All ☒ List ☐ Range

List of values

Value	Display As
1	City
2	State
3	Style
4	Serving Size
5	Brewery Name
6	Beer Name
Add	

3. Create the `Dimension Left` parameter:

`Dimension Left`

```
CASE [Select Dimension Left]
WHEN 1 THEN [City]
WHEN 2 THEN [State]
WHEN 3 THEN [Style]
WHEN 4 THEN [Serving Size]
WHEN 5 THEN [Brewery Name]
WHEN 6 THEN [Beer Name]
END
```

4. Create the `Dimension Right` parameter:



### Dimension Right

```
CASE [Select Dimension Right]
WHEN 1 THEN [City]
WHEN 2 THEN [State]
WHEN 3 THEN [Style]
WHEN 4 THEN [Serving Size]
WHEN 5 THEN [Brewery Name]
WHEN 6 THEN [Beer Name]
END
```

5. Create the `Chosen Measure` parameter:

### Chosen Measure

```
[F1]
```

6. Create the `Top Dimension Values` parameter:

**Edit Parameter [Top Dimension Values]**

Name:  Comment >>

**Properties**

Data type:

Current value:

Display format:

Allowable values: ☐ All ☐ List ☒ Range

**Range of values**

☒ Minimum:  Set from Parameter

☒ Maximum:  Set from Field

☒ Step size:

OK Cancel

### Note

The first six steps are not required, but they give the visualization some fun flexibility. The user can select which categories appear on the left-hand side and the right-hand side of the chart. Also, because there are so many values,

which can make the Sankey chart look busy, we've added a top value to filter the dimension values. For this Sankey recipe, we actually only need to create a sheet with two dimensions and one measure.

7. Add **Dimension Right** to the **Columns** shelf and the **Filters** shelf.
8. Add **Dimension Left** to the **Rows** shelf and the **Filters** shelf.
9. Right-click **Dimension Right** and **Dimension Left** in the **Filters** shelf, select **Edit Filter**, and set as follows:

The screenshot shows the 'Edit Filter' dialog box with the 'Top' tab selected. The 'By field:' section is active, showing 'Top' in the first dropdown, 'Top Dimension Values' in the second, and 'Sum' in the third. The 'By formula:' section is empty. The 'Reset', 'OK', 'Cancel', and 'Apply' buttons are at the bottom.

10. Add **Chosen Measure** to **Text** in the **Marks** card.
11. Add a **Percent of Total** option in **Table Calculation**:

Table Calculation

% of Total Chosen Measure

×

Calculation Type

Percent of Total

☐ Compute total across all pages

Compute Using

Table (across)

Table (down)

Table

Cell

Specific Dimensions

☒ Dimension Left
 ☐ Dimension Right

At the level

Sort order Specific Dimensions

☒ Show calculation assistance

The cross-tab should appear as follows:

Pages	Columns	Dimension Right
	Rows	Dimension Left
Filters		
Dimension Left		
Dimension Right		
Marks		
Automatic		
Color	Size	Text
Detail	Tooltip	
SUM(Chosen Measur...		

	Dimension Right									
Dimens ion Left	America n Ambe...	America n Blond...	America n Brow...	America n Doubl...	America n IPA	America n Pale A...	America n Pale ..	America n Porter	Saison / Farmho...	Witbier
CO	9.95%	27.07%	22.52%		22.77%	16.18%	24.56%	6.42%	10.01%	18.19%
WA	15.54%	13.16%	15.22%		18.18%	10.68%	30.86%	33.19%	29.75%	16.55%
CA	7.05%	18.07%	18.37%	9.63%	14.42%	24.44%	3.59%		5.79%	13.89%
NY		20.71%		39.52%	9.22%		19.69%			19.45%
OR	27.19%	3.70%		5.95%	4.75%	16.33%	10.58%			6.89%
WI	11.65%	13.11%	14.91%	1.27%	6.48%	6.84%		21.61%		14.87%
NC	14.91%			22.43%	8.39%	8.19%			32.12%	
PA	8.39%	4.16%		11.77%	3.42%	1.03%		38.79%	22.33%	5.74%
TX	5.33%		13.49%	9.43%	8.77%	3.78%				4.41%
MA			15.49%		3.59%	12.53%	10.72%			

12. Add the sheet to a dashboard, as shown in the following screenshot:

Dashboard

Layout

Device Preview

Size

Desktop Browser (1000 x 800)

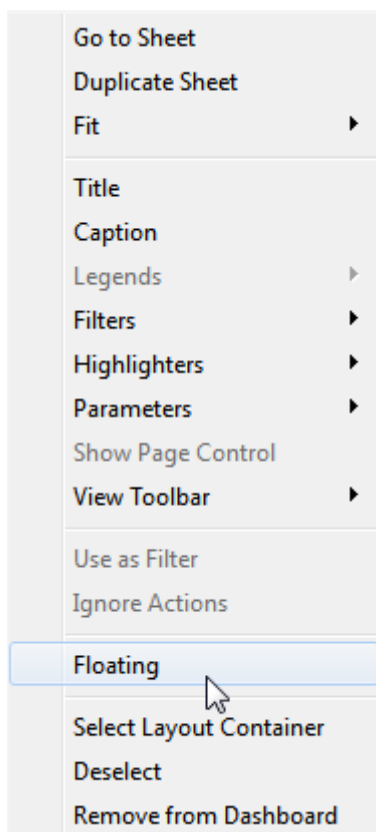
Sheets

Sankey 2018.2

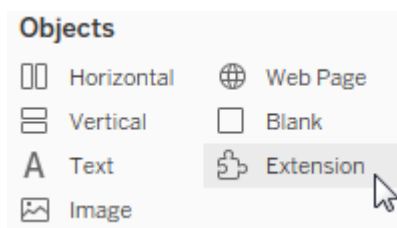
	Dimension Right											
Dimens ion Left	America n Ambe...	America n Blond...	America n Brow...	America n Doubl...	America n IPA	America n Pale A...	America n Pale ...	America n Porter	Saison / Farmho...	Witbier		
CO	9.95%	27.07%	22.52%		22.77%	16.18%	24.56%	6.42%	10.01%	18.19%		
WA	15.54%	13.16%	15.22%		18.18%	10.68%	30.86%	33.19%	29.75%	16.55%		
CA	7.05%	18.07%	18.37%	9.63%	14.42%	24.44%	3.59%		5.79%	13.89%		
NY		20.71%		39.52%	9.22%		19.69%			19.45%		
OR	27.19%	3.70%		5.95%	4.75%	16.33%	10.58%			6.89%		
WI	11.65%	13.11%	14.91%	1.27%	6.48%	6.84%		21.61%		14.87%		
NC	14.91%			22.43%	8.39%	8.19%			32.12%			
PA	8.39%	4.16%		11.77%	3.42%	1.03%		38.79%	22.33%	5.74%		
TX	5.33%		13.49%	9.43%	8.77%	3.78%				4.41%		
MA			15.49%		3.59%	12.53%	10.72%					

13. Make the sheet **Floating** and then minimize it. In this example, we don't want to see the cross-tab.

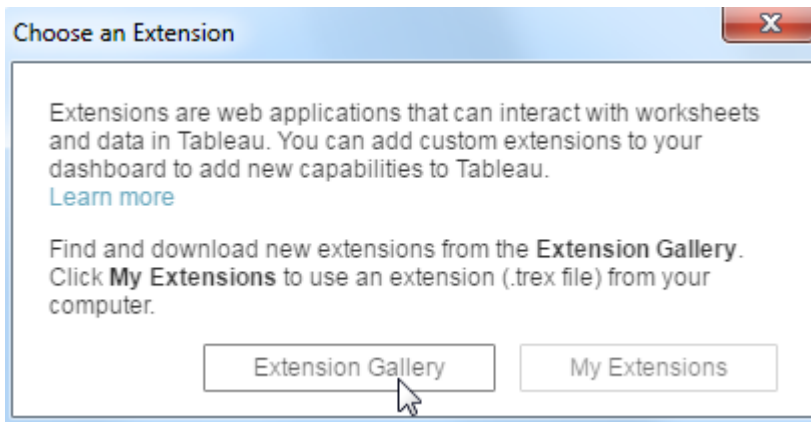
However, the **Show Me More** extension requires it to be on the dashboard in order to work:



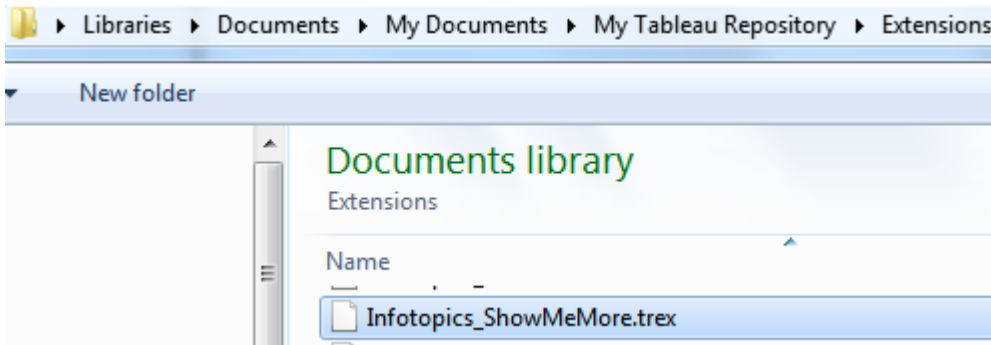
14. Use the **Extension** object:



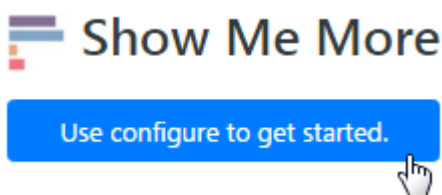
15. When prompted for the first time, download the **Show Me More** web extension by visiting **Extension Gallery** :



16. You can go to **My Extensions** and navigate to your extensions library to choose **Show Me More** :



17. Click on **Use configure to get started.** to begin the configuration:



18. Choose **Sankey Diagram** :

## Show Me More

Do you enjoy show me more? Get access to even more visualization types by buying the full version!


Buy Now

### Select Data Sheet

Select the worksheet you want to retrieve data from. This worksheet will be the input for the visualization you pick in the next step.


Sankey 2018.2

### Select a visualization




**Sankey Diagram**  
sankey diagram uses:

2 Dimension(s) 1 Measure(s)

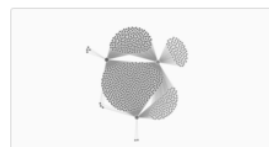


**Radar Chart**  
radar chart uses:

2 Dimension(s) 1 Measure(s)



**Horizon Chart**  
Get access to horizon chart by purchasing the full version.



**Network Diagram**  
Get access to network diagram by purchasing the full version.

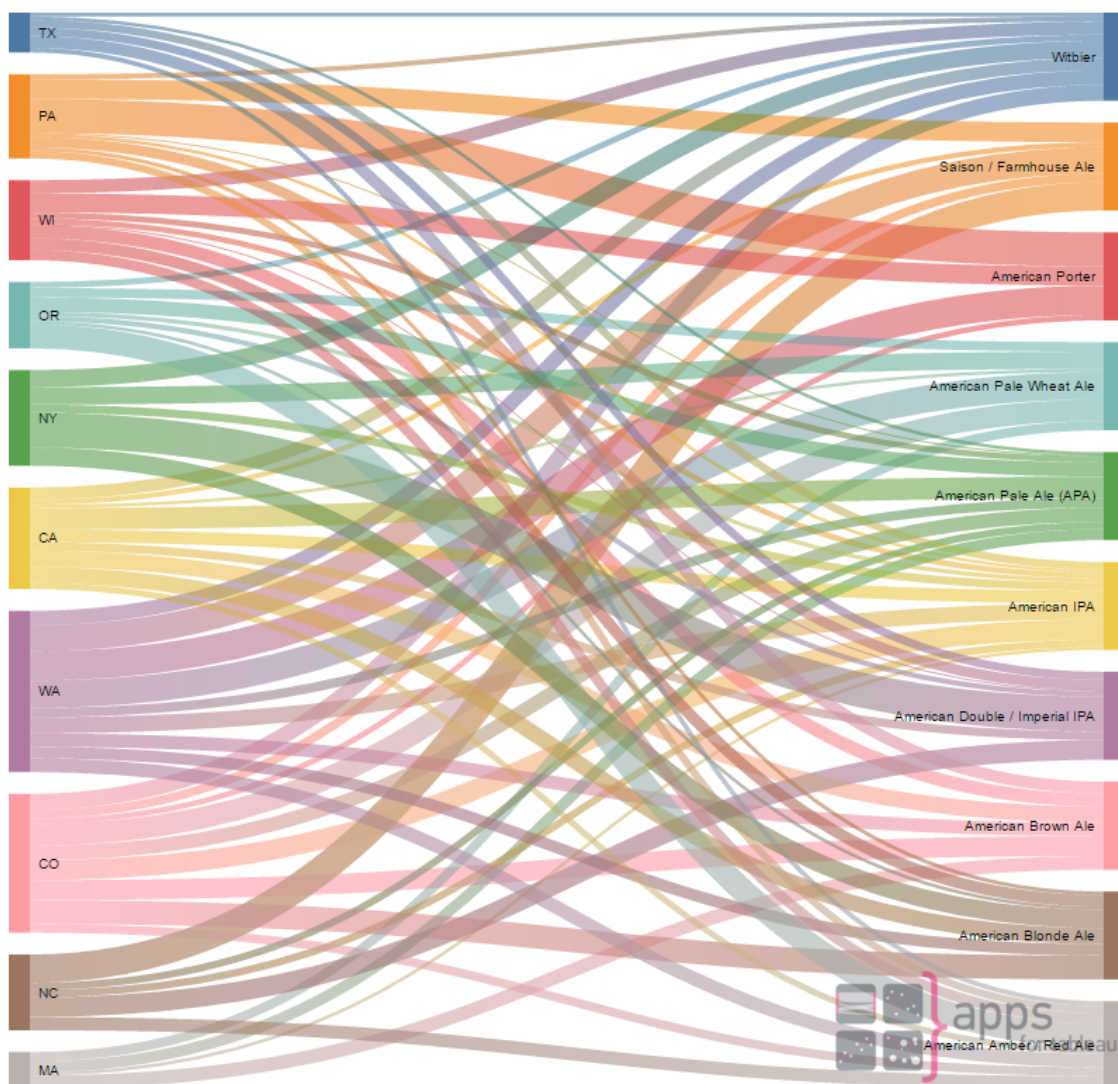
19. Select the dimensions and measures as shown in the following screenshot:

## Customize Sankey Diagram

Level 1	Dimension Left	⌵
Level 2	Dimension Right	⌵
Need more dimensions? Buy show me more and unlock unlimited dimensions for this visualization.		
Measure	SUM(Chosen Measure)	⌵
Format	None	⌵

Color by: ☒ unique value per dimension ☐ dimension value

We can see the final visualization in the following screenshot:



## How it works...

We created a cross-tab with two dimensions and one measure. We then added this to a dashboard. Because we only wanted to see the Sankey chart and the extension requires that the cross-tab be on the dashboard, we "hid it" by making it very small and floating. Using the extension object, we chose the **Show Me More** extension. We configured it for **Sankey Diagram**.

## See also

There are several ways to create a Sankey chart published on various blogs. Ian Balwin's post on *[Information Lab]* is excellent because it is flexible and does not require outside data prep. However, there are many calculations, which can make it hard to follow. But once it's set, there is a great deal of flexibility because of how the dimensions and measures have been created. The following directions are taken from his blog post at <https://www.theinformationlab.co.uk/2018/03/09/build-sankey-diagram-tableau-without-data-prep-beforehand/>.

Use `Beer.twbx`, `beers.csv`, and `breweries.csv` to work through this example:

1. Create the `Select Dimension Left` parameter or use the same one from the previous recipe:

Name:  Comment >>

Properties

Data type:

Current value:

Display format:

Allowable values: ☐ All ☒ List ☐ Range

List of values

Value	Display As
1	City
2	State
3	Style
4	Serving Size
5	Brewery Name
6	Beer Name
Add	

Add from Parameter ▶  
Add from Field ▶  
Paste from Clipboard  
  
Clear All

OK Cancel

2. Create the `Select Dimension Right` parameter or use the same one from the previous recipe:



Name:  Comment >>

Properties

Data type:

Current value:

Display format:

Allowable values: ☐ All ☒ List ☐ Range

List of values

Value	Display As
1	City
2	State
3	Style
4	Serving Size
5	Brewery Name
6	Beer Name
Add	

3. Create the `Dimension Left` parameter or use the same one from the previous recipe:

`Dimension Left`

```
CASE [Select Dimension Left]
WHEN 1 THEN [City]
WHEN 2 THEN [State]
WHEN 3 THEN [Style]
WHEN 4 THEN [Serving Size]
WHEN 5 THEN [Brewery Name]
WHEN 6 THEN [Beer Name]
END
```

4. Create the `Dimension Right` parameter or use the same one from the previous recipe:

### Dimension Right

```
CASE [Select Dimension Right]
WHEN 1 THEN [City]
WHEN 2 THEN [State]
WHEN 3 THEN [Style]
WHEN 4 THEN [Serving Size]
WHEN 5 THEN [Brewery Name]
WHEN 6 THEN [Beer Name]
END
```

5. Create the `Chosen Measure` parameter or use the same one from the previous recipe:

### Chosen Measure

```
[F1]
```

6. Create the `Top Dimension Values` parameter or use the same one from the previous recipe:

**Edit Parameter [Top Dimension Values]**

Name:  Comment >>

**Properties**

Data type:

Current value:

Display format:

Allowable values: ☐ All ☐ List ☒ Range

**Range of values**

☒ Minimum:  Set from Parameter

☒ Maximum:  Set from Field

☒ Step size:

OK Cancel

7. Add `Dimension Right` and `Dimension Left` to the `Filters` shelf.

8. Right-click `Dimension Right` and `Dimension Left` in the `Filters` shelf, select `Edit Filter`, and set it as follows:

General Wildcard Condition Top

☐ None

☒ By field:

Top Top Dimension Values by

Chosen Measure Sum

☐ By formula:

Top 10 by

Reset OK Cancel Apply

9. Create the `Path Frame` parameter:

Path Frame

---

```
IF [Chosen Measure] = {FIXED [Dimension Left]: MIN([Chosen Measure])} THEN 0 ELSE 97 END
```

10. Create a `Path Index` parameter:

Path Index

---

Results are computed along Path Frame (bin).

```
INDEX ()
```

11. Create the `T` parameter:

T

```
IF [Path Index] < 50
THEN (([Path Index]-1)%49)/4-6
ELSE 12 - (([Path Index]-1)%49)/4-6
END
```

12. Create the Sigmoid function:

Sigmoid

$1/(1+\text{EXP}(1)^{-[T]})$

13. Create the Sankey Arm Size parameter:

Sankey Arm Size

Totals summarize values from Dimension Left, Dimension Right.  
 $\text{SUM}([\text{Chosen Measure}])/\text{TOTAL}(\text{SUM}([\text{Chosen Measure}]))$

14. Create all the following calculations for the top of the Sankey Arm:

- Max Position Left
  - $\text{RUNNING\_SUM}([\text{Sankey Arm Size}])$
- Max Position Left Wrap
  - $\text{WINDOW\_SUM}([\text{Max Position Left}])$
- Max Position Right
  - $\text{RUNNING\_SUM}([\text{Sankey Arm Size}])$
- Max Position Right Wrap
  - $\text{WINDOW\_SUM}([\text{Max Position Right}])$

15. Create all of the following calculations for the bottom of the Sankey Arm:

- Max for Min Position Left
  - $\text{RUNNING\_SUM}([\text{Sankey Arm Size}])$
- Min Position Left
  - $\text{RUNNING\_SUM}([\text{Max for Min Position Left}]) - [\text{Sankey Arm Size}]$
- Min Position Left Wrap
  - $\text{WINDOW\_SUM}([\text{Min Position Left}])$

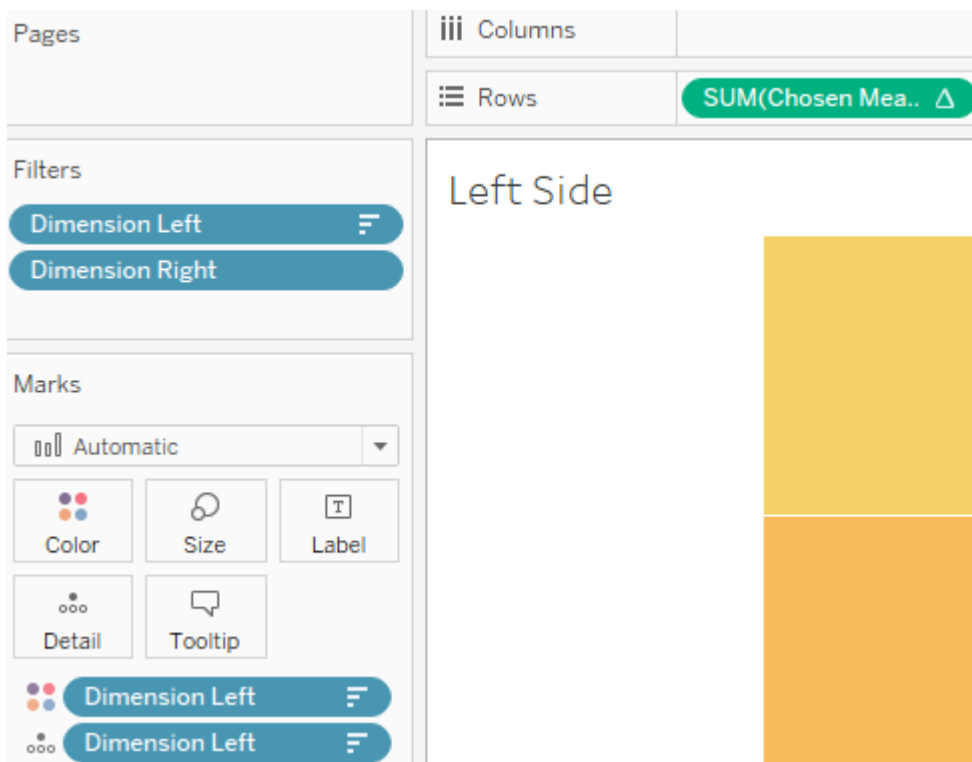
- Max for Min Position Right
  - `RUNNING_SUM([Sankey Arm Size])`
- Min Position Right
  - `RUNNING_SUM([Max for Min Position Right])-[Sankey Arm Size]`
- Min Position Right Wrap
  - `WINDOW_SUM([Min Position 2])`

16. Create the `Sankey Polygons` calculation as follows:

`Sankey Polygons`

```
IF [Path Index] > 49
THEN [Max Position Left Wrap]+([Max Position Right Wrap]-[Max Position Left Wrap])*[Sigmoid]
ELSE [Min Position Left Wrap]+([Min Position Right Wrap]-[Min Position Left Wrap])*[Sigmoid]
END
```

17. Create the `Left Side` sheet:



18. Apply a `Percent of Total` option in `Table Calculation` to `Chosen Measure` :

Table Calculation

×

% of Total Chosen Measure

Calculation Type

Percent of Total

▼

☐ Compute total across all pages

Compute Using

Table (across)

Table (down)

Table

Cell

Specific Dimensions

☒ Dimension Left

At the level

▼

☒ Show calculation assistance

19. Create the **Right Side** sheet:



20. Apply a **Percent of Total** option under **Table Calculation** to **Chosen Measure** :

Table Calculation

×

% of Total Chosen Measure

Calculation Type

Percent of Total

☐ Compute total across all pages

Compute Using

Table (across)

Table (down)

Table

Cell

Specific Dimensions

☒ Dimension Right

At the level

☒ Show calculation assistance

21. Create `INDEX ()` by double-clicking in the **Marks** card and typing `INDEX ()` :

Marks

Automatic

Color

Size

Label

Detail

Tooltip

Dimension Right

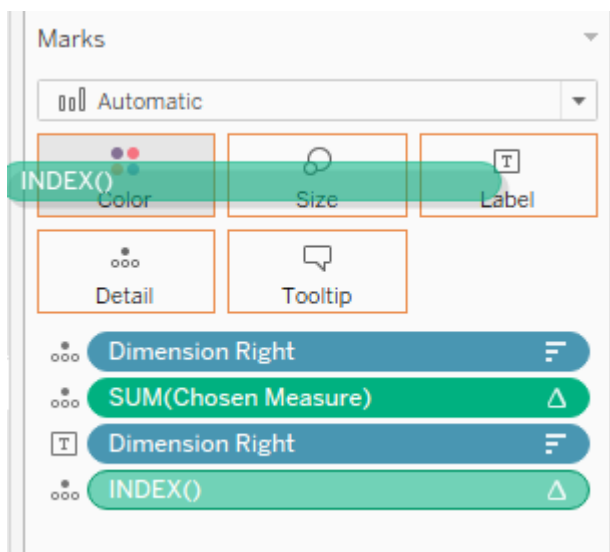
SUM(Chosen Measure)

Dimension Right

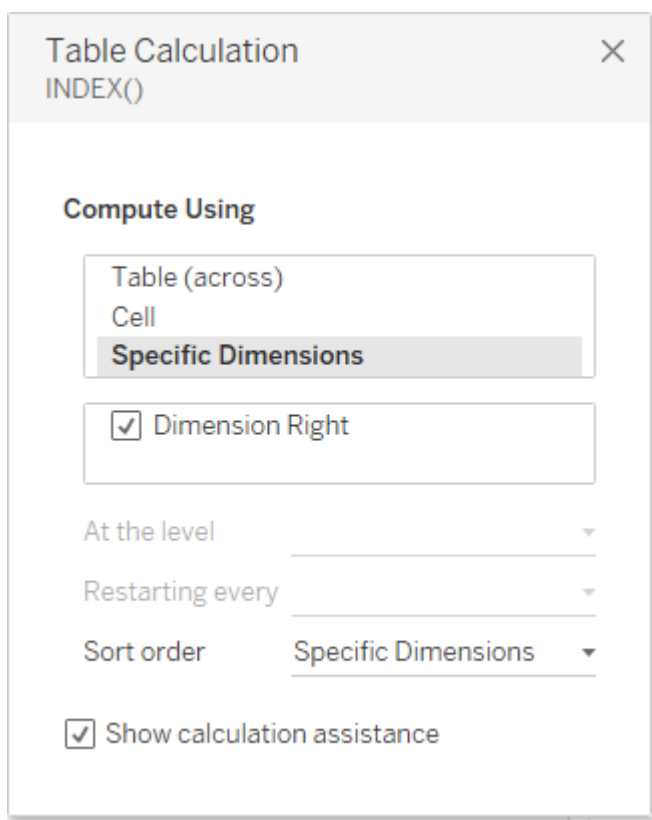
INDEX ()

22. Apply `INDEX ()` to the\*\* Color \*\* card:

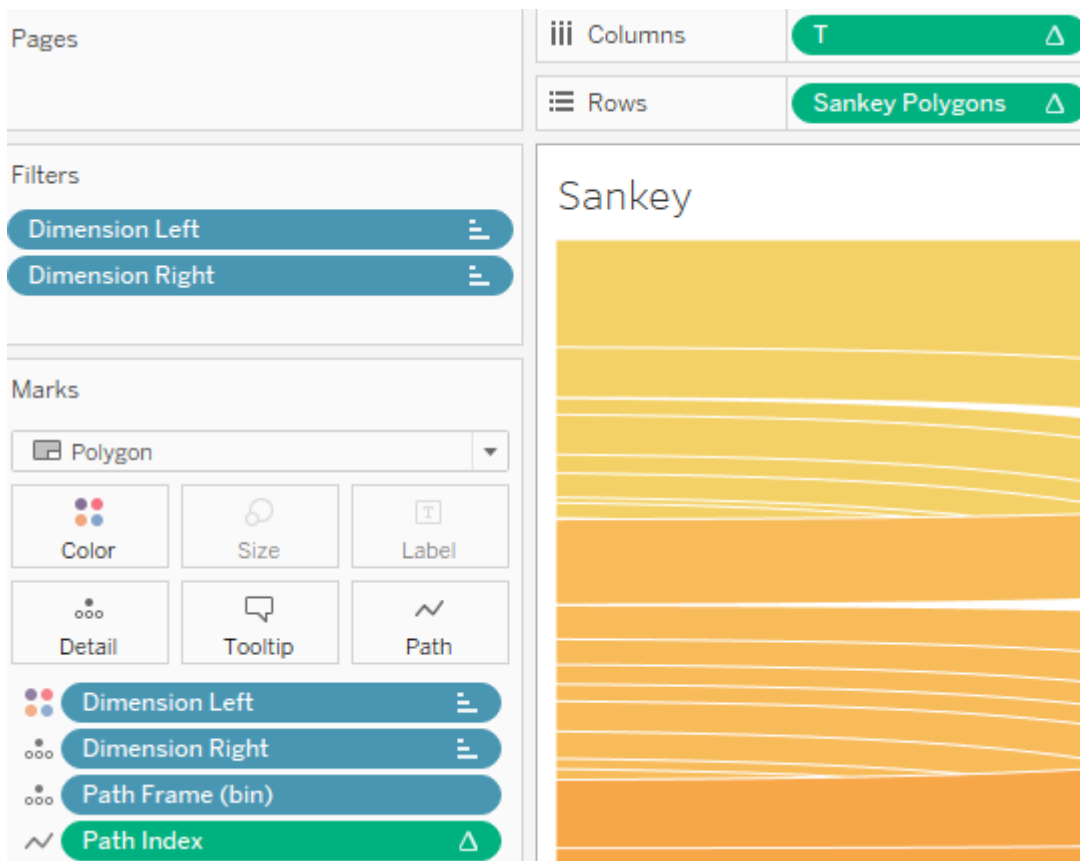




23. Apply a table calculation to **INDEX()** , as shown in the following screenshot:



24. Create the\*\* Sankey \*\* sheet:



25. Table Calculation for Sankey Polygons :

- These are all the table calculations we will need to configure for the **Sankey Polygons** :

Table Calculation

Sankey Polygons

✕

### Nested Calculations

Path Index

C

Path Index

Max Position Left Wrap

Max Position Left

Sankey Arm Size

Max Position Right Wrap

Max Position Right

Min Position Left Wrap

Min Position Left

Max for Min Position Left

Min Position Right Wrap

Min Position Right

Max for Min Position Right

Sort order

Specific Dimensions

☒ Show calculation assistance

Table Calculation

Sankey Polygons

✕

### Nested Calculations

Path Index ▾

### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Path Frame (bin)

☐ Dimension Left

☐ Dimension Right

At the level ▾

Restarting every ▾

Sort order Specific Dimensions ▾

☒ Show calculation assistance

26. ◦ Path Index :

Table Calculation

Sankey Polygons

✕

### Nested Calculations

Max Position Left Wrap ▾

☐ Null if there are not enough values

### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Path Frame (bin)

☐ Dimension Left

☐ Dimension Right

At the level ▾

Restarting every ▾

Sort order      Specific Dimensions ▾

☒ Show calculation assistance

27.    ◦    **Max Position Left Wrap :**

Table Calculation

Sankey Polygons

✕

### Nested Calculations

Max Position Left

### Compute Using

Table (across)  
Cell  
**Specific Dimensions**

☒ Dimension Left  
☒ Dimension Right  
☐ Path Frame (bin)

At the level **Deepest**

Restarting every **None**

Sort order **Specific Dimensions**

☒ Show calculation assistance

28. ◦ **Max Position Left** :

Table Calculation

Sankey Polygons

×

### Nested Calculations

Sankey Arm Size

☐ Compute total across all pages

### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Path Frame (bin)

☒ Dimension Left

☒ Dimension Right

At the level **Deepest**

Restarting every **None**

Sort order **Specific Dimensions**

☒ Show calculation assistance

29. ◦ **Sankey Arm Size** :

Table Calculation

Sankey Polygons

×

### Nested Calculations

Max Position Right Wrap ▾

☐ Null if there are not enough values

### Compute Using

Table (across)  
Cell  
**Specific Dimensions**

☒ Path Frame (bin)  
☐ Dimension Left  
☐ Dimension Right

At the level ▾

Restarting every ▾

Sort order Specific Dimensions ▾

☒ Show calculation assistance

30. ○ Max Position Right Wrap :



Table Calculation

Sankey Polygons

×

Nested Calculations

Max Position Right

▼

Compute Using

Table (across)

Cell

Specific Dimensions

☒ Dimension Right

☒ Dimension Left

☐ Path Frame (bin)

At the level

Deepest

▼

Restarting every

None

▼

Sort order

Specific Dimensions

▼

☒ Show calculation assistance

31. ◦ Max Position Right :

Table Calculation

Sankey Polygons

×

### Nested Calculations

Min Position Left Wrap ▾

☐ Null if there are not enough values

### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Path Frame (bin)

☐ Dimension Left

☐ Dimension Right

At the level ▾

Restarting every ▾

Sort order Specific Dimensions ▾

☒ Show calculation assistance

32. ◦ **Min Position Left Wrap** :

Table Calculation

Sankey Polygons

×

### Nested Calculations

Min Position Left

### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Dimension Left

☒ Dimension Right

☐ Path Frame (bin)

At the level **Deepest**

Restarting every **None**

Sort order **Specific Dimensions**

☒ Show calculation assistance

33. ◦ **Min Position Left** :

Table Calculation

Sankey Polygons

×

Nested Calculations

Max for Min Position Left

▼

Compute Using

Table (across)

Cell

Specific Dimensions

☐ Dimension Left

☐ Dimension Right

☐ Path Frame (bin)

At the level

▼

Restarting every

▼

☒ Show calculation assistance

34. ◦ Max for Min Position Left :

## Table Calculation ✕

### Sankey Polygons

#### Nested Calculations

Min Position Right Wrap ▼

☐ Null if there are not enough values

#### Compute Using

Table (across)

Cell

**Specific Dimensions**

☒ Path Frame (bin)

☐ Dimension Left

☐ Dimension Right

At the level ▼

Restarting every ▼

Sort order Specific Dimensions ▼

☒ Show calculation assistance

35. ◦ **Min Position Right Wrap :**

Table Calculation

Sankey Polygons

✕

Nested Calculations

Min Position Right

▼

Compute Using

Table (across)

Cell

Specific Dimensions

☒ Dimension Right

☒ Dimension Left

☐ Path Frame (bin)

At the level

Deepest

▼

Restarting every

None

▼

Sort order

Specific Dimensions

▼

☒ Show calculation assistance

36. ◦ **Min Position Right** :

Table Calculation

Sankey Polygons

✕

Nested Calculations

Max for Min Position Left

▼

Compute Using

Table (across)

Cell

Specific Dimensions

☐ Dimension Left

☐ Dimension Right

☐ Path Frame (bin)

At the level

Restarting every

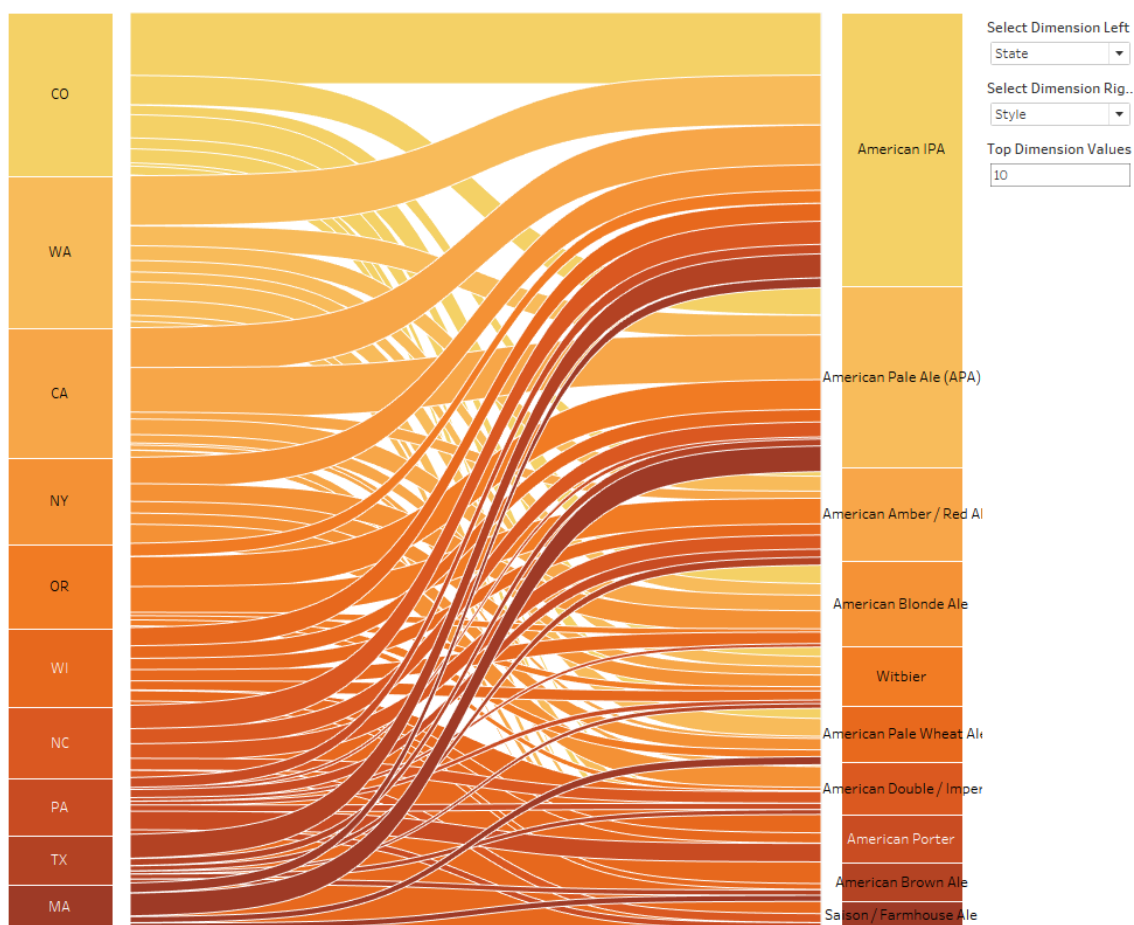
▼

▼

☒ Show calculation assistance

37.   ◦   **Max for Min Position Right :**

You will see the final dashboard in the following screenshot:



## Marimekko charts

Marimekko charts go by many names: mekko, mosaic, or matrix, to name a few. It is a two-dimensional stacked chart. They are used to analyze data composition or distribution across two variables at once. Each axis represents 100%.

### Getting ready

In this recipe, we use complex calculations to build the Marimekko chart.

### How to do it..

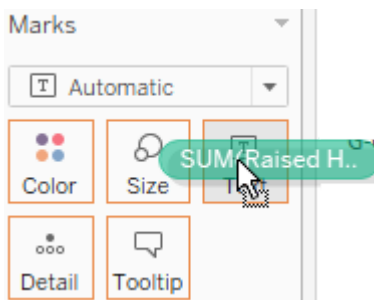
To follow along, open the Marimekko packaged workbook: `Mekko.twbx` and `xAPI-Edu-Data.csv`. We are going to compare gender, parent satisfaction with school, and student participation by raised hands.

### Note

It is recommended you start with a text table to get the calculations correct.

1. Add the dimensions of interest to the **Rows** shelf. We are going to add **Grade ID**, **Gender**, and **Parentschool Satisfaction**.
2. Add the **Raised Hands** to the **Text** shelf in the **Marks** card:





3. Add the **Percent of Total** option in **Table Calculation** , and under **Compute Using** select **Parentschool Satisfaction** :

Table Calculation
×

% of Total Raised Hands

**Calculation Type**

Percent of Total
▼

☐ Compute total across all pages

**Compute Using**

Table (across)
Table (down)
Table
Pane (down)
Pane
Cell
**Specific Dimensions**

☒ Parentschool Satisfaction
☐ Gender
☐ Grade ID

At the level ▼

Sort order **Specific Dimensions** ▼

☒ Show calculation assistance

4. Double-click on\*\* Measure Values \*\*:



5. Add the **Raised Hands** to the **Measure Values** card:



6. Create calculation Raised Hands per Column :

### Raised Hands per Column

```
{EXCLUDE [Parentschool Satisfaction]:SUM([Raised Hands])}
```

7. Add Raised Hands per Column to the Measure Values card:



8. Create Calculation # of Raised Hands, Grade, Gender ID, Parentschool Satisfaction and add it to the Measures Value card. This calculation creates somewhat of a running total summary for grades and genders. This value will be the [x] axis:

Results are computed along Table (across).

```
//If it's the first row in partition
```

```
IF FIRST()==0 THEN
```

```
    //return this value
```

```
    MIN([Raised Hands per Column])
```

```
//check if this grade is NOT the same as the previous one
```

```
ELSEIF MIN([Grade ID]) != LOOKUP(MIN([Grade ID]),-1) THEN
```

```
    //Add the previous value of raised hands per column to this one
```

```
    PREVIOUS_VALUE(0) + MIN([Raised Hands per Column])
```

```
//check if gender is NOT the same as the previous one
```

```
ELSEIF MIN([Gender]) != LOOKUP(MIN([Gender]),-1) THEN
```

```
    //add the previous value of raised hands per column to this one
```

```
    PREVIOUS_VALUE(0) + MIN([Raised Hands per Column])
```

```
ELSE
```

```
    //it's the same grade and gender, show the same raised hands value
```

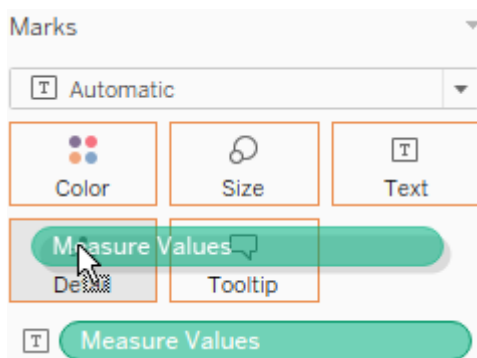
```
    PREVIOUS_VALUE(0)
```

```
END
```

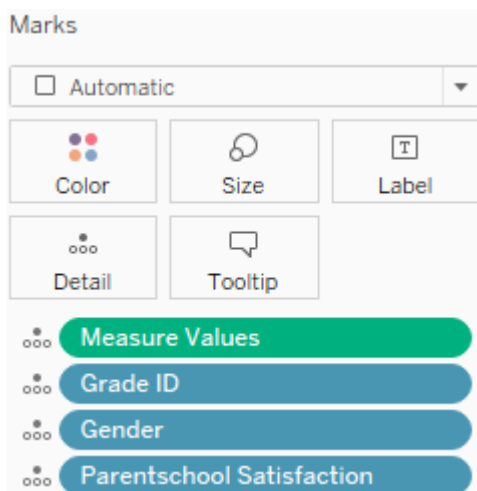
9. We've completed the calculation work in order to make the visualization. Your grid should look as follows:

Grade ID	Gender	Parentschool Satisfaction	Raised Hands	% of Total Raised Hands along Parents..	Raised Hands per Column	# of Raised Hands, Grade, Gender ID, Pa..
G-02	F	Bad	499	20.18%	2,473	2,473
		Good	1,974	79.82%	2,473	2,473
	M	Bad	1,072	37.00%	2,897	5,370
		Good	1,825	63.00%	2,897	5,370
G-04	F	Bad	272	24.46%	1,112	6,482
		Good	840	75.54%	1,112	6,482
	M	Bad	345	30.50%	1,131	7,613
		Good	786	69.50%	1,131	7,613

10. Move **Measure Values from Text** to **Detail** :



11. Move **Grade ID** , **Gender** , and **Parentschool Satisfaction** to **Detail** :



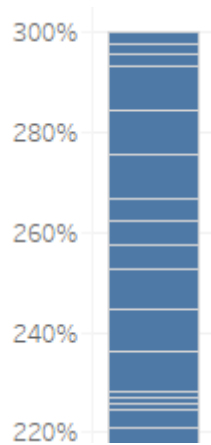
12. Move **Sum(Raised Hands)** with the **Percent of Total** calculation to the **Rows** shelf:



13. Remove **Measure Names** from the **Columns** shelf:

Columns	
Rows	SUM(Raised Hand.. Δ)

14. We have a stacked bar at this step:



15. Move # of Raised Hands, Grade, Gender ID, Parentschool Satisfaction to the Columns shelf:

Columns	# of Raised Hands.. Δ
Rows	SUM(Raised Hand.. Δ)

16. Change the Scatter Plot mark type to Bar :

Marks

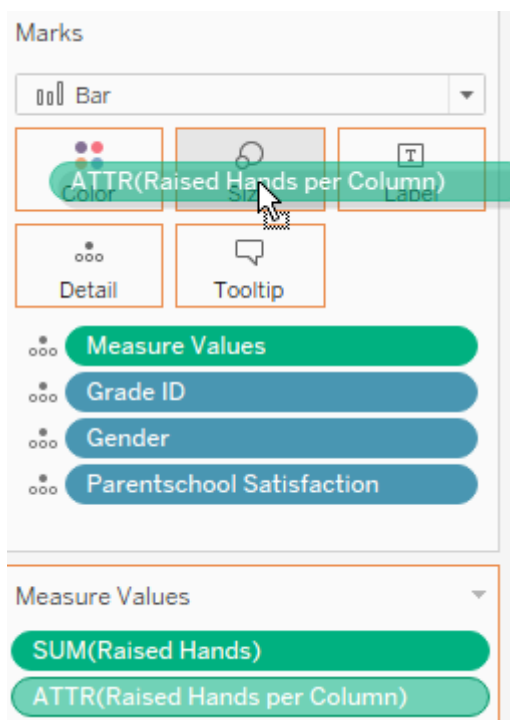
Automatic

Automatic

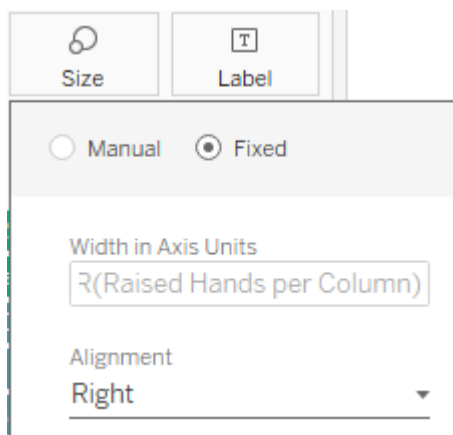
Bar

Line

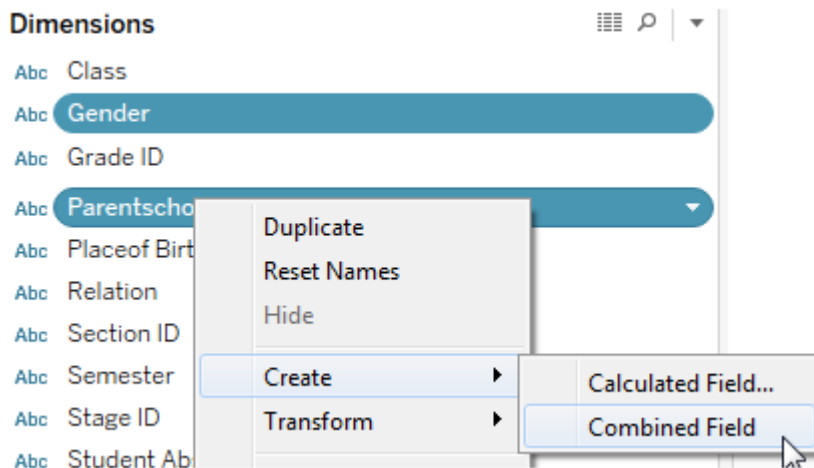
17. Move Raised Hands per Column from Measure Values to Size :



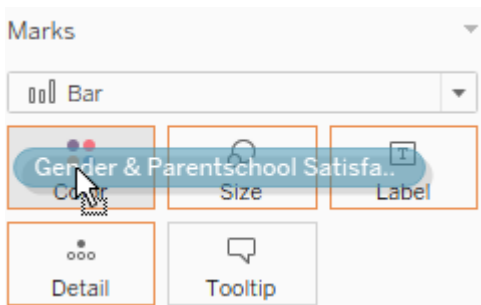
18. Set **Size** to **Fixed** and **Alignment** to **Right** :



19. Create a **Gender** and **Parentschool satisfaction** combined field, as shown in the following screenshot:

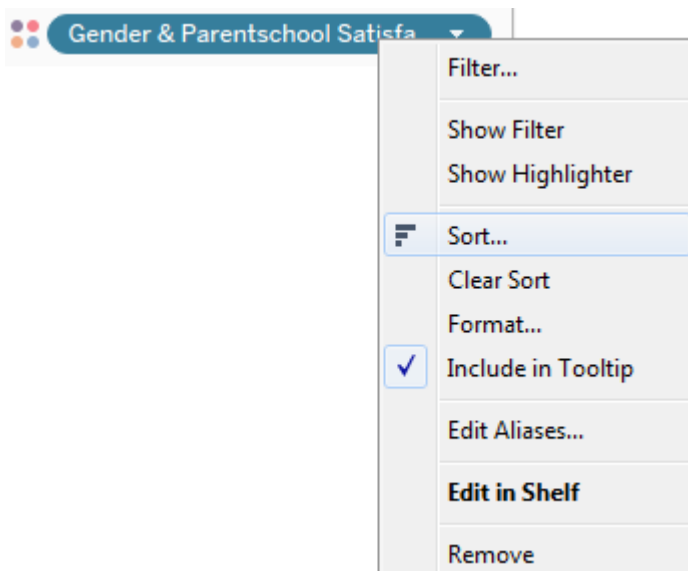


20. Add the combined field to **Color** :



21. Manually sort by **Gender & Parentschool Satisfaction** :

- Right-click on this dimension in the marks card and select **Sort** :



- Choose **Manual Sort** and make it appear as follows:

Sort [Gender & Parentschool Satisfaction (...]

Sort By

Manual sort

Males w/ Unhappy Parents

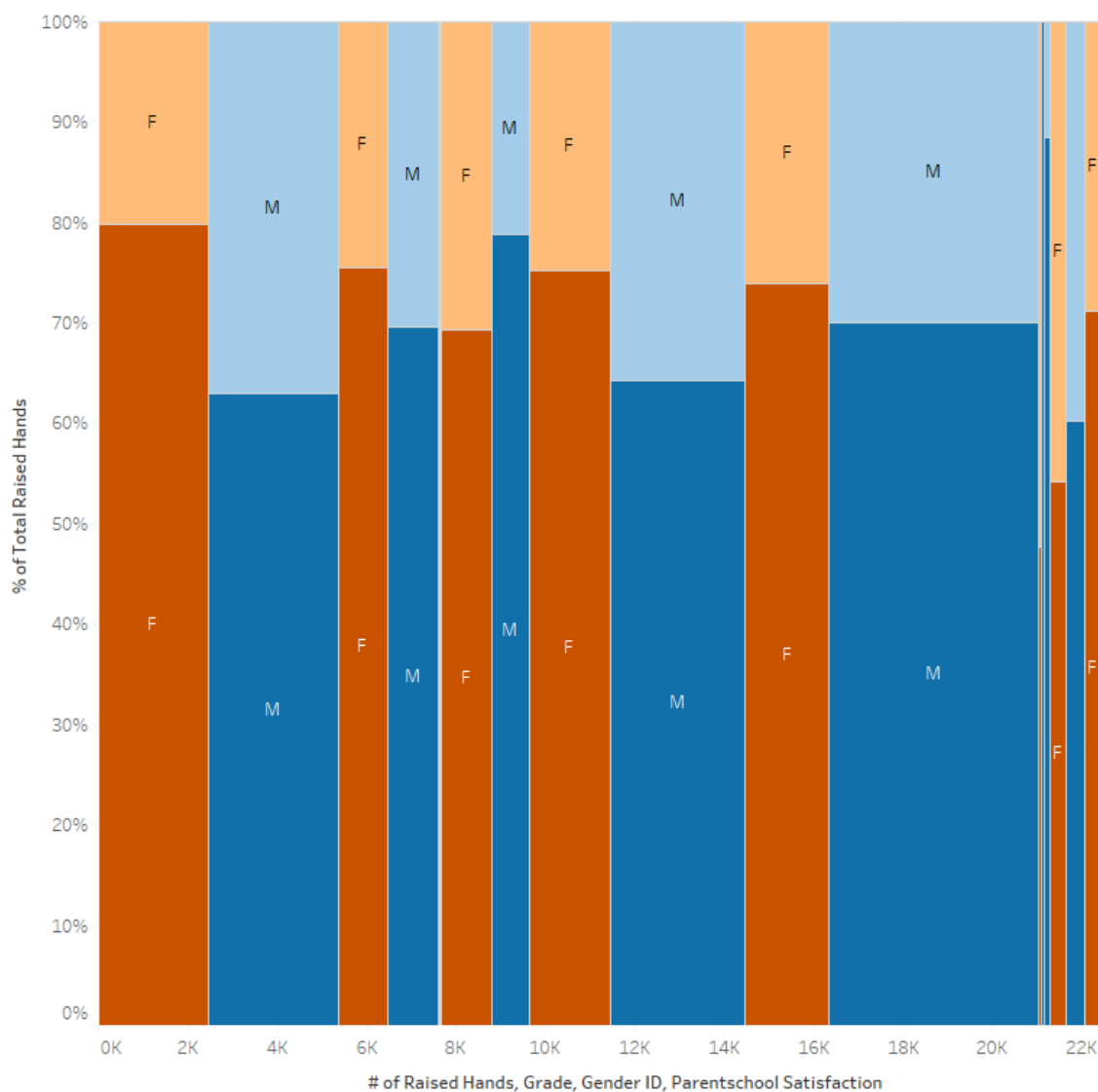
Males w/ Happy Parents

Females w/ Unhappy Parents

Females w/ Happy Parents

Clear

22. This should produce the following chart. Second, seventh, and eighth graders raise their hands more in class than other grades. Parents with students who raise their hands in class were generally happier with the school. High school students do not raise their hands in class, as shown in the following screenshot:



## How it works...

First, we created a text table to work through our calculations. We then added the dimensions of interest to the **Rows** shelf and the measure of interest to the **Text** shelf.

For column height, we created `Percent of Total for Raised Hands`, which we compute by using `Parentschool Satisfaction`. This allows us to see 100% for each **Gender** and **Grade** combination, which is our column height. We add this calculation to the **Measure Values** card.

In order to get our column widths, we created the `Raised Hands per Column` calculation, which calculates the sum for all hand raises for each combination of **Gender** and **Grade**. We then add the calculation to the **Measure Values** card.

Next, we created `# of Raised Hands, Grade, Gender ID, Parentschool Satisfaction` to help order the columns of each grade and gender correctly along the x-axis. This calculation is building a running total based on raised hands per column.



It does this math by going through the following checks:

- Is it the first row in the partition then return this value?
- Has the grade changed then add this value and the previous value?
- Has the gender changed? If it has, add this value and the previous value; otherwise we have to return the previous value. Then we began to create the visualization. We moved\*\* `Measure Values` \*\* from `Text` to `Detail` ; move `Grade ID` , `Gender` , and

`Parentschool Satisfaction` to `Detail` . We moved `Sum(Raised Hands)` with the percent of total calculation to the `Rows` shelf, and removed `Measure Names` from the `Columns` shelf. At this point, we have a stacked bar.

We moved `# of Raised Hands` , `Grade` , `Gender ID` , `Parentschool Satisfaction` to the `Columns` shelf. Next, we changed the `Scatter Plot mark type` to `Bar` . To get the column widths, we move `Raised Hands per Column` from `Measure Values` to `Size` and set it to `Fixed` and `Alignment` to\*\* `Right` \*\*.

In order to color our visualization, we created a `Gender` and `Parentschool Satisfaction` combined field that we add to `Color` . Finally, to make `Gender` more visible, we added it to `Label` .

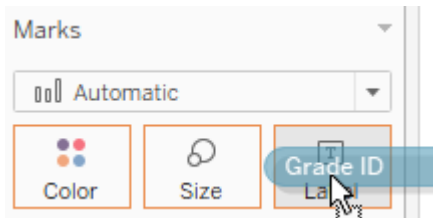
### There's more...

We can create a header visualization to use in a dashboard so that the grades are labelled better:

1. Add `Raised Hands` to the `Column` shelf:

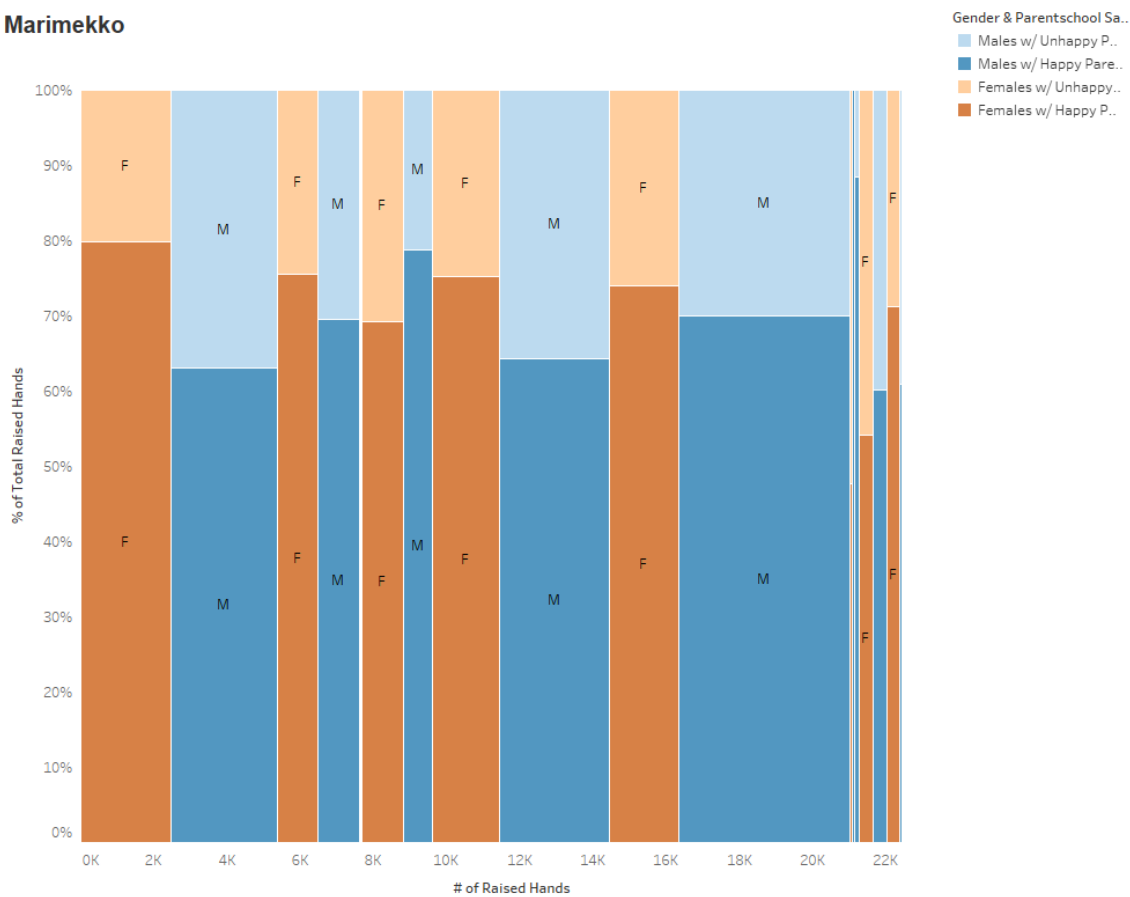


2. Add `Grade ID` to `Label` :

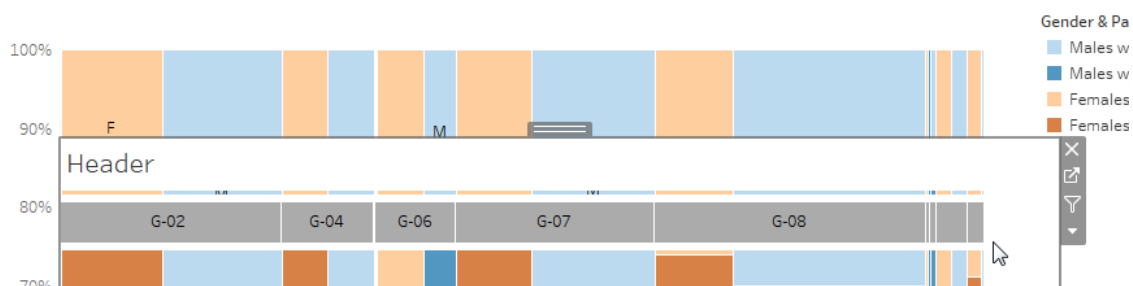


3. Add the `Marimekko` chart to a dashboard:

## Marimekko

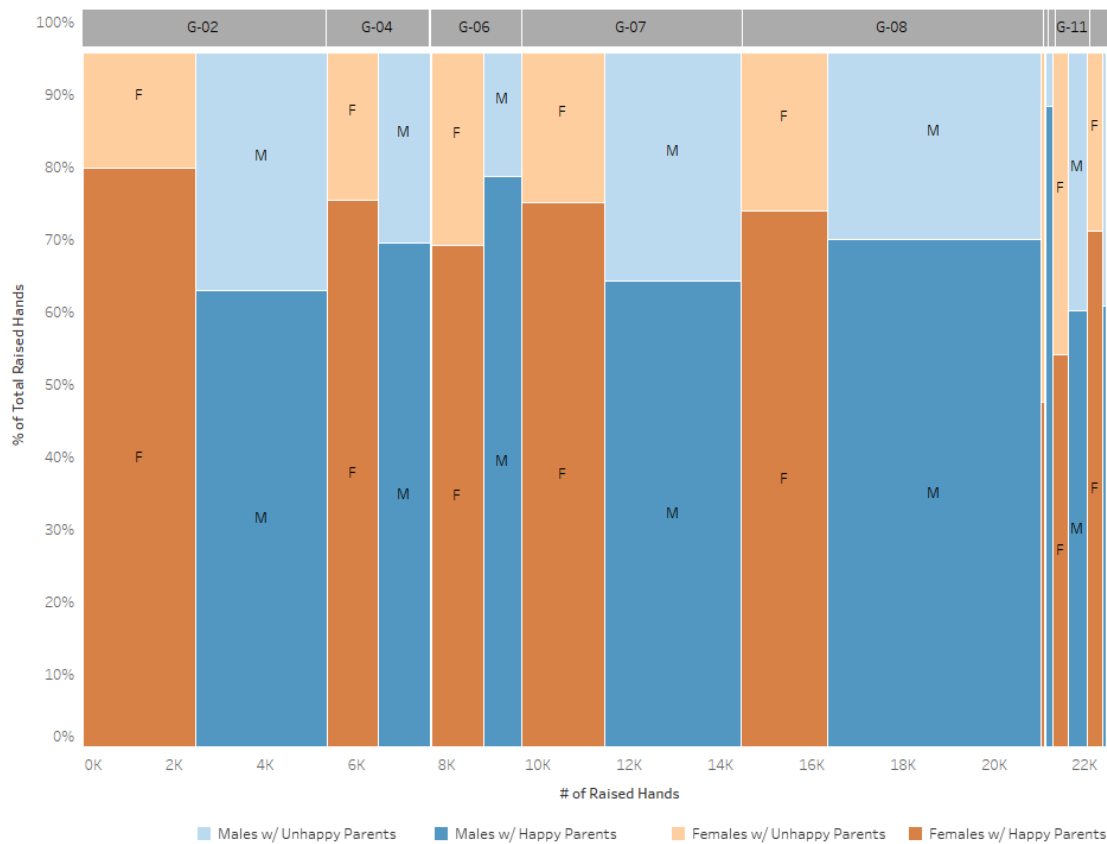


4. Add the **Header** visualization to the dashboard as a floating object and resize it:



5. After some formatting and resizing, we can view our final visualization in the following screenshot:

How does student classroom engagement, parent satisfaction, and gender relate?



## Hex-Tile maps

Here, you will learn how to eliminate the visual perceptions that occur due to different sizes of different states or countries and focus on showing more actionable trends.

### Getting ready

In this recipe, we will create a Hex-Tile map using a scatter plot as a base.

### How to do it..

Follow along using the `Hexmap.twbx`, `hexmap_plots.xlsx`, and `Data USA - Map of Commuting Alone over 30 Minutes by State.csv` datasets:

1. Choose the `hexmap_plot.xlsx` file.
2. Join `Data USA - Map of Commuting Alone over 30 Minutes by State` to the `hexmap_plot` data using the State and Geo Name columns as the join condition:

**Connections** [Add](#)

- Data USA - Ma...utes by State  
Text file
- hexmap\_plots  
Microsoft Excel

**Sheets** [p](#)

☐ Use Data Interpreter  
Data Interpreter might be able to clean your Microsoft Excel workbook.

Sheet1

New Union

**Join**

Inner Left Right Full Outer

Data Source		Sheet1
Geo Name	=	State
<a href="#">Add new join clause</a>		

Sort fields: Data source order

#	Abc	Abc	#	#	#	Abc	
Data USA - ...	Data USA - Map of Comm...	Data USA - Ma...	Data USA - Map of Commuting...	Sheet1	Sheet1	Sheet1	
Year	Geo Name	Geo	Long_Commute_....	Row	Column	State	Abbreviation
2017	Alaska	04000US02	0.162000	0	0.5000	Alaska	AK

3. Add **Column** to the **Columns** shelf, add **Row** to the **Rows** shelf, and use **AVG** for the aggregation:

**Columns** **AVG(Column)**

**Rows** **AVG(Row)**

4. Use **Abbreviation** as the\*\* Label \*\*:

**Marks**

Automatic

Color Size Label

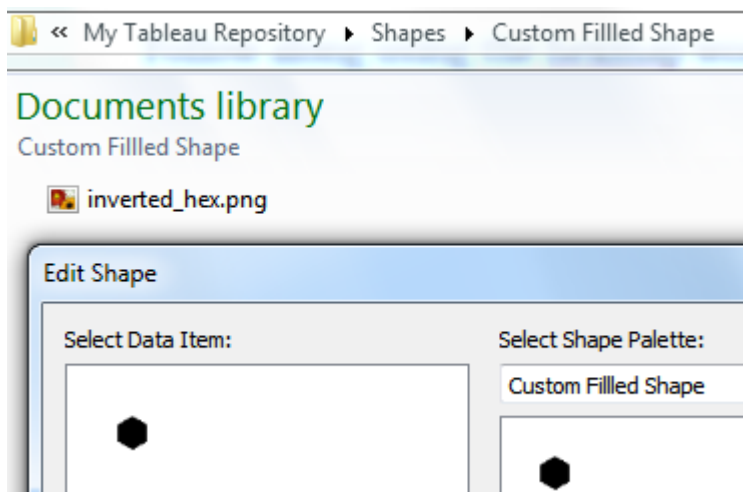
Detail Tooltip Shape

**Abbreviation**

5. Edit the row axis and reverse the scale:

General		Tick Marks	
<b>Range</b>			
<input checked="" type="radio"/> Automatic		<input checked="" type="checkbox"/> Include zero	
<input type="radio"/> Uniform axis range for all rows or columns			
<input type="radio"/> Independent axis ranges for each row or column			
<input type="radio"/> Fixed			
Automatic ▼	Automatic ▼		
-0.461196106	8.461196106		
<b>Scale</b>			
<input checked="" type="checkbox"/> Reversed			
<input type="checkbox"/> Logarithmic			
<b>Axis Titles</b>			
Title			
Avg. Row			
Subtitle			
		<input checked="" type="checkbox"/> Automatic	
<div>Reset</div>			

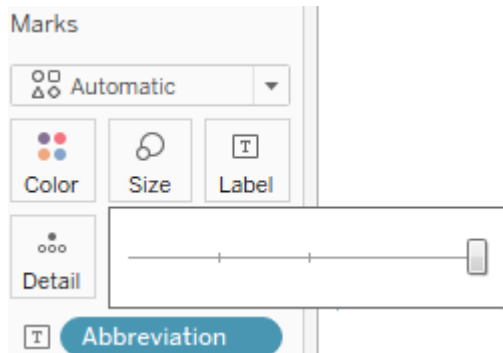
6. Use the `inverted_hex.png` image as a custom shape:



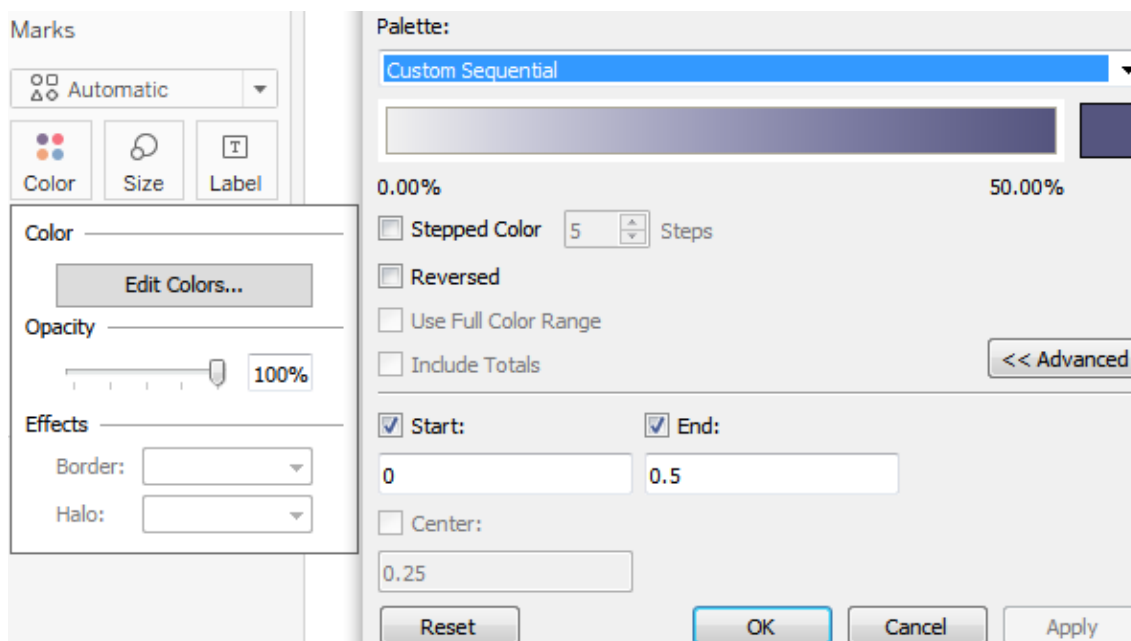
**Note**

Save the `inverted_hex.png` image to your `Tableau Repository | Shapes | Custom Filled Shape` folder.

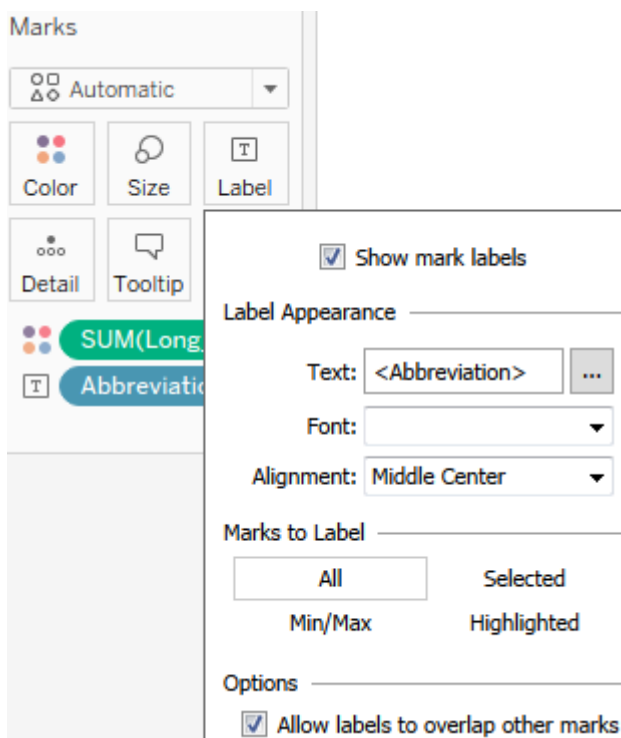
7. Adjust the **Size** :



8. Color and shade the tiles by a measure in your data file. In our visualization, we're using `Longest_Commute_Driving_Alone` :



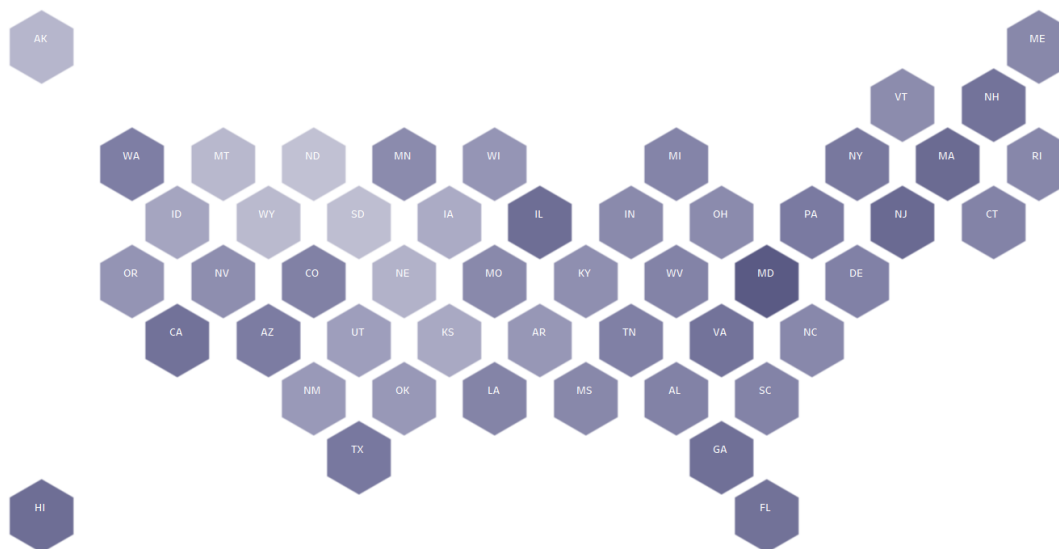
9. Add state labels, by putting\*\* `Abbreviation` \*\* on **Label** :



After some formatting, we can see the final

visualization in the following screenshot:

Which state has the **longest commute** in the United States?



## How it works...

The key to the *[Hex-Tile maps]* recipe was using the `hexmap_plot` data and the tile image. The column and row values position the tiles so they were arranged in a similar, relative place as each state. The custom shape helped tie the visualization together in a compact way.

## Waffle charts

Waffle charts have almost similar use cases to Donut charts. They are used to show how items contribute to a whole. They are best used when comparing only a few categories.

### Getting ready

In this recipe, we will create a Waffle chart using a text table as the foundation.

### How to do it..

To follow along, open the Waffle chart packaged workbook. In this example, we are looking at word counts for each Lord of the Rings character by race:

1. Use the `Waffle frame` Excel sheet:
  - This Excel sheet is 100 rows representing each percentage point. Because we want to create a frame of 100 squares in a 10 x 10 frame, we have columns called `Rows` , `Columns` , and `Percentage` . Each row and column has a value of 1 through 10, repeating. This creates a 10 row



x 10 column square. The first 20 are displayed in the following screenshot:

# Sheet1 (Waffle Frame)

Sheet1

Sort fields

Data source order

# Sheet1 Rows	# Sheet1 Columns	# Sheet1 Percentage
1	1	1%
1	2	2%
1	3	3%
1	4	4%
1	5	5%
1	6	6%
1	7	7%
1	8	8%
1	9	9%
1	10	10%
2	1	11%
2	2	12%
2	3	13%
2	4	14%
2	5	15%

2. Add **Columns** to the **Columns** shelf and **Rows** to the **Rows** shelf. Because we want these values to be grouped in discrete buckets, we have set these to discrete. This can be achieved by right-clicking and choosing discrete, as opposed to continuous:

Columns	Columns
Rows	Rows

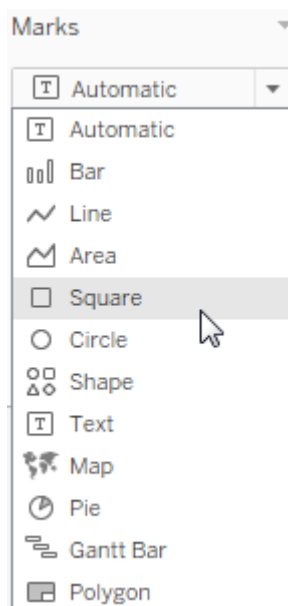
3. Add percentage to **Label1** to see the layout:

Rows		Columns									
		1	2	3	4	5	6	7	8	9	10
10		91%	92%	93%	94%	95%	96%	97%	98%	99%	100%
9		81%	82%	83%	84%	85%	86%	87%	88%	89%	90%
8		71%	72%	73%	74%	75%	76%	77%	78%	79%	80%
7		61%	62%	63%	64%	65%	66%	67%	68%	69%	70%
6		51%	52%	53%	54%	55%	56%	57%	58%	59%	60%
5		41%	42%	43%	44%	45%	46%	47%	48%	49%	50%
4		31%	32%	33%	34%	35%	36%	37%	38%	39%	40%
3		21%	22%	23%	24%	25%	26%	27%	28%	29%	30%
2		11%	12%	13%	14%	15%	16%	17%	18%	19%	20%
1		1%	2%	3%	4%	5%	6%	7%	8%	9%	10%

4. Order the **Rows** column in descending order:

Rows		Columns						
		1	2	3	4	5	6	7
10		91%	92%	93%	94%	95%	96%	97%
9		81%	82%	83%	84%	85%	86%	87%

5. Remove the percentage from **Label1** and change the mark type to **Square** :



6. Create the actual share of words spoken by the Hobbits. Go the `WordsByCharacter` dataset and create a `Hobbit` calculation based on `Race = Hobbit` for percent of total spoken words:

Hobbit

WordsByCharacter

```
SUM(IF [Race] = "Hobbit" THEN [Words] END)/SUM([Words])
```

7. Create a true/false calculation to indicate whether the `Hobbit` is greater than or equal to each square in the Waffle chart. The `Hobbit Percentage` calculation allows us to color each square:

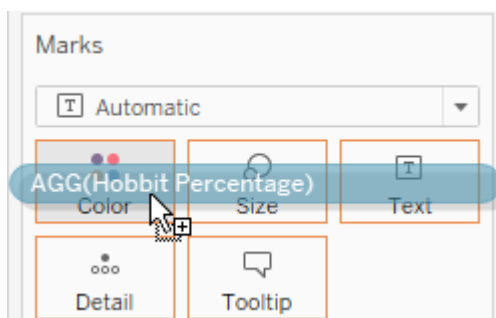
Hobbit Percentage

Sheet1 (Waffle Frame)

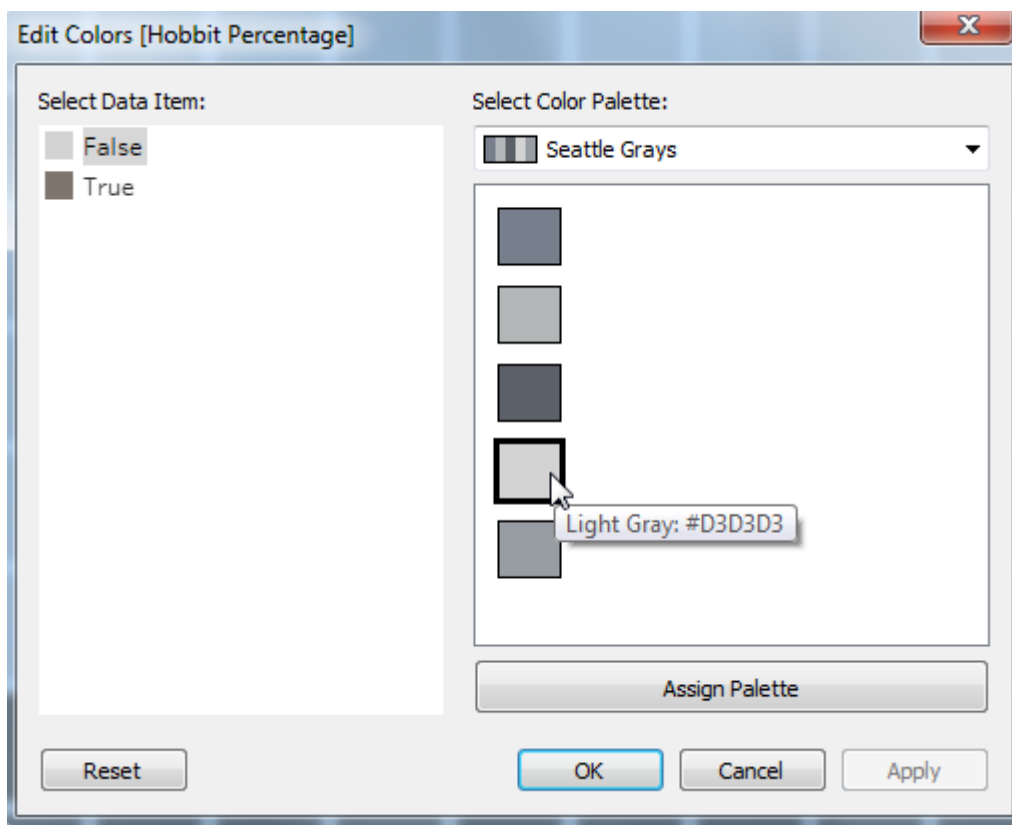
```
[WordsByCharacter].[Hobbit] >= sum([Percentage])
```

8. Apply `Hobbit Percentage` to

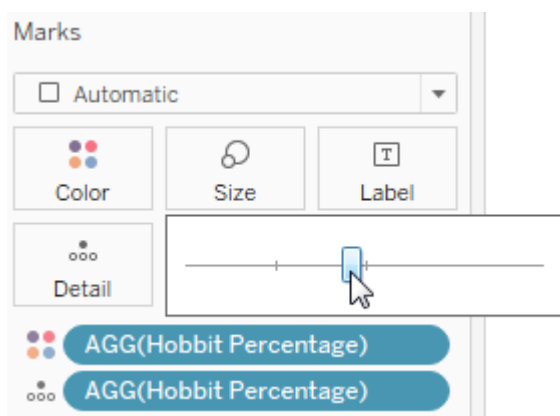
Color :



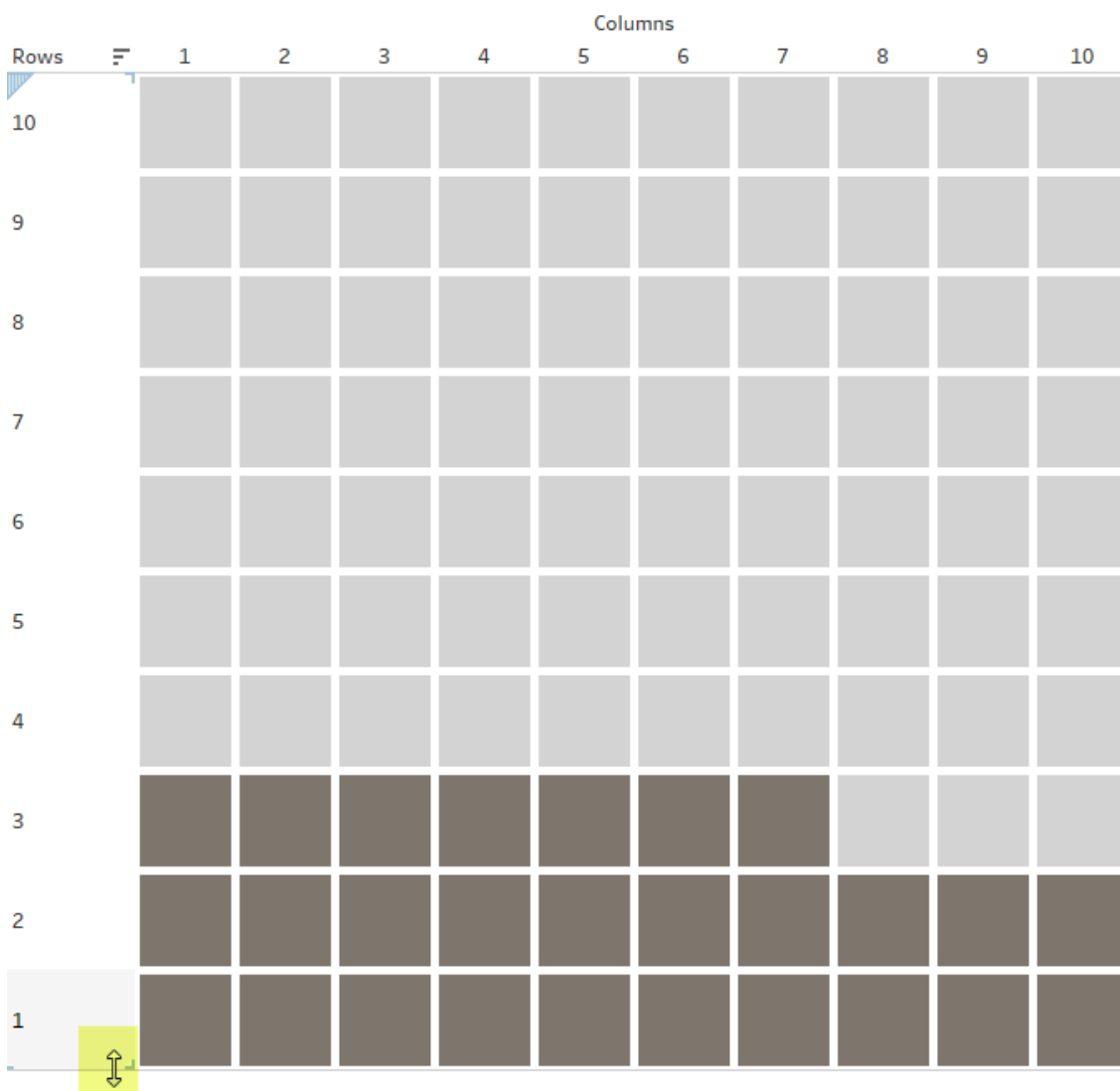
9. Adjust the color so that false is faint and subtle:



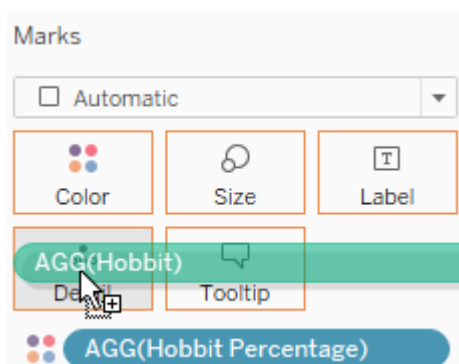
10. Adjust the mark size as follows:



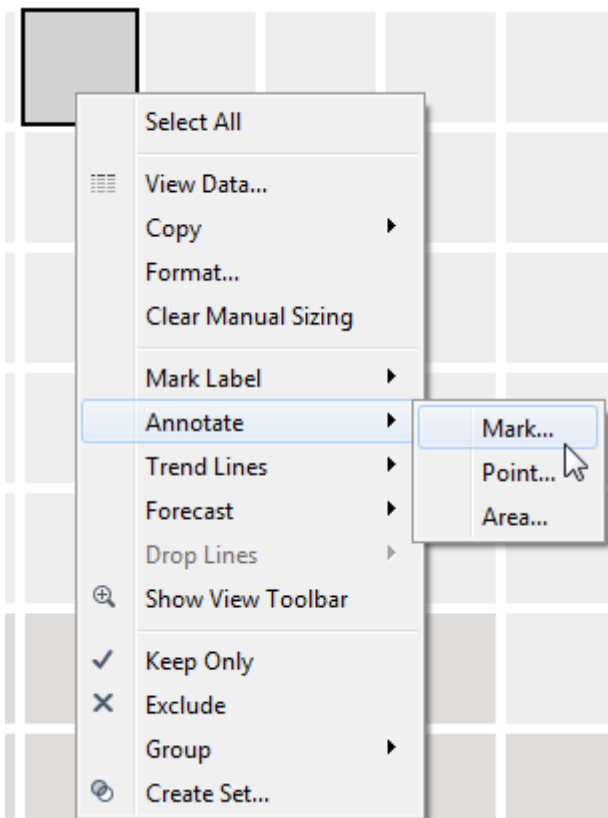
11. Adjust the chart size by manually adjusting the columns and rows:



12. Add **Hobbit** to details, so we can annotate the chart:

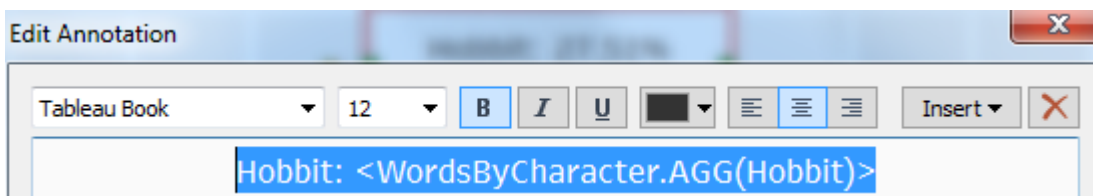


13. Annotate the Waffle chart as follows:

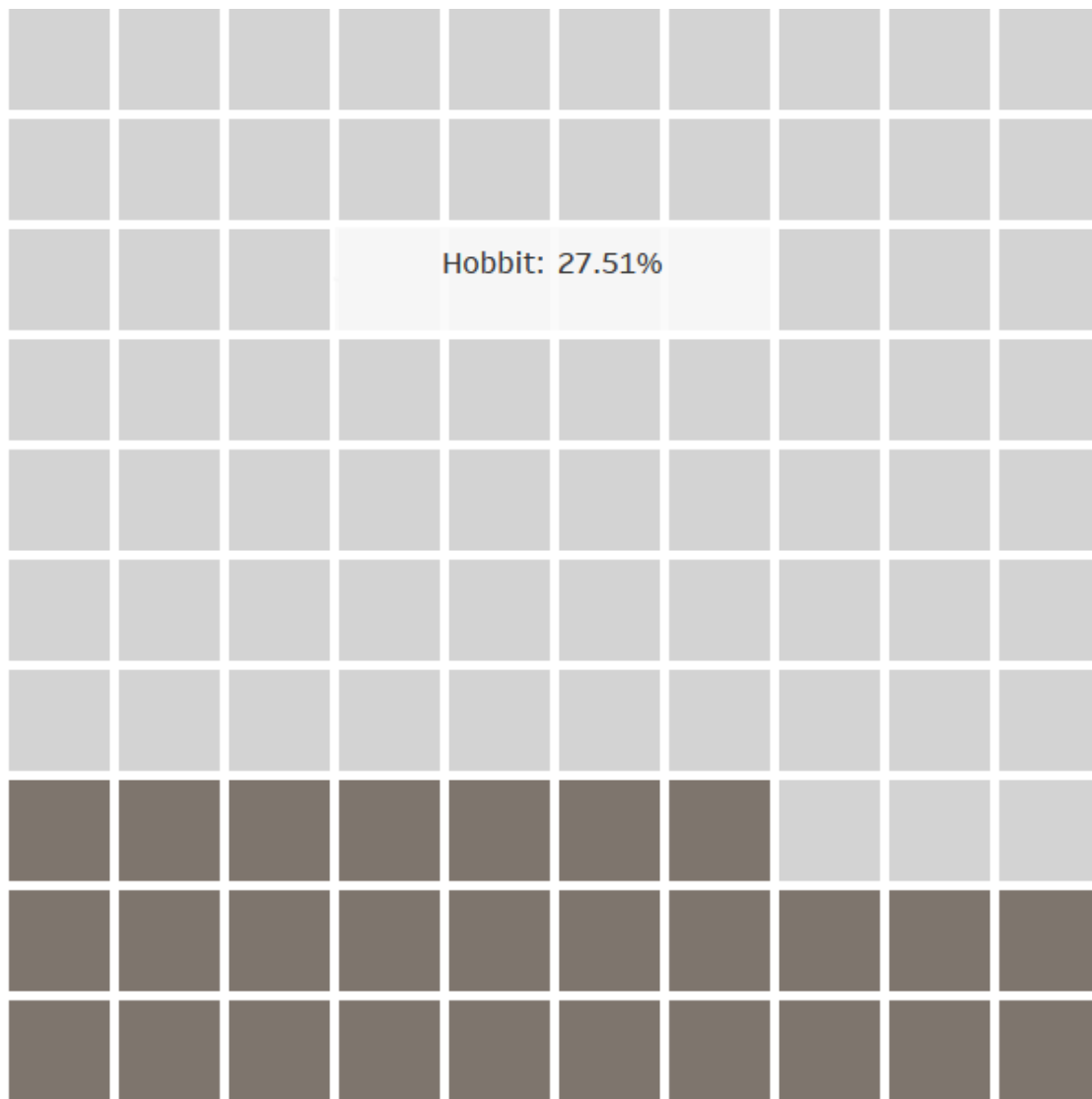


14. Modify the annotation as shown in the following

screenshot:



15. After formatting and hiding headers, the final Waffle chart should look like this:



**How it works...**

We used the Excel 10 row x 10 column grid to represent a 100% grid, and put columns and rows on their respective shelves. We then changed the mark type to **Square** . We colored the tiles by creating percent of total for each category, in this case for race. In order to emphasize each category, we checked whether the value is less than the percentage value in our Waffle frame, which drives the color. We also changed the size of each square to get a nice waffle shape. Finally, we annotated the Waffle chart to make it easier to read.

**There's more...**

We can create more complicated calculations to represent more than one category in each Waffle chart. In a new sheet, starting with columns and rows in their respective shelves, continue with the following steps:

- 1. Create a calculation that will color every square in the Waffle chart depending on what percentage of words came from which movie:

Film Percentage

WordsByCharacter

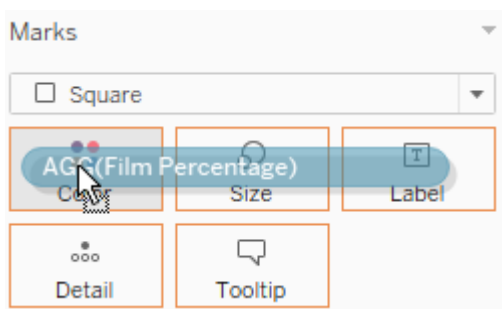
```
IF AVG([Sheet1 (Waffle Frame)].[Rows]*10)-(10-[Sheet1 (Waffle Frame)].[Columns])
  <= ROUND([Film, Fellowship]*100) THEN "The Fellowship Of Ring"

ELSEIF AVG([Sheet1 (Waffle Frame)].[Rows]*10)-(10-[Sheet1 (Waffle Frame)].[Columns])
  <= ROUND([Film, Fellowship] + [Film, Towers]*100) THEN "The Two Towers"

ELSE "The Return Of The King"

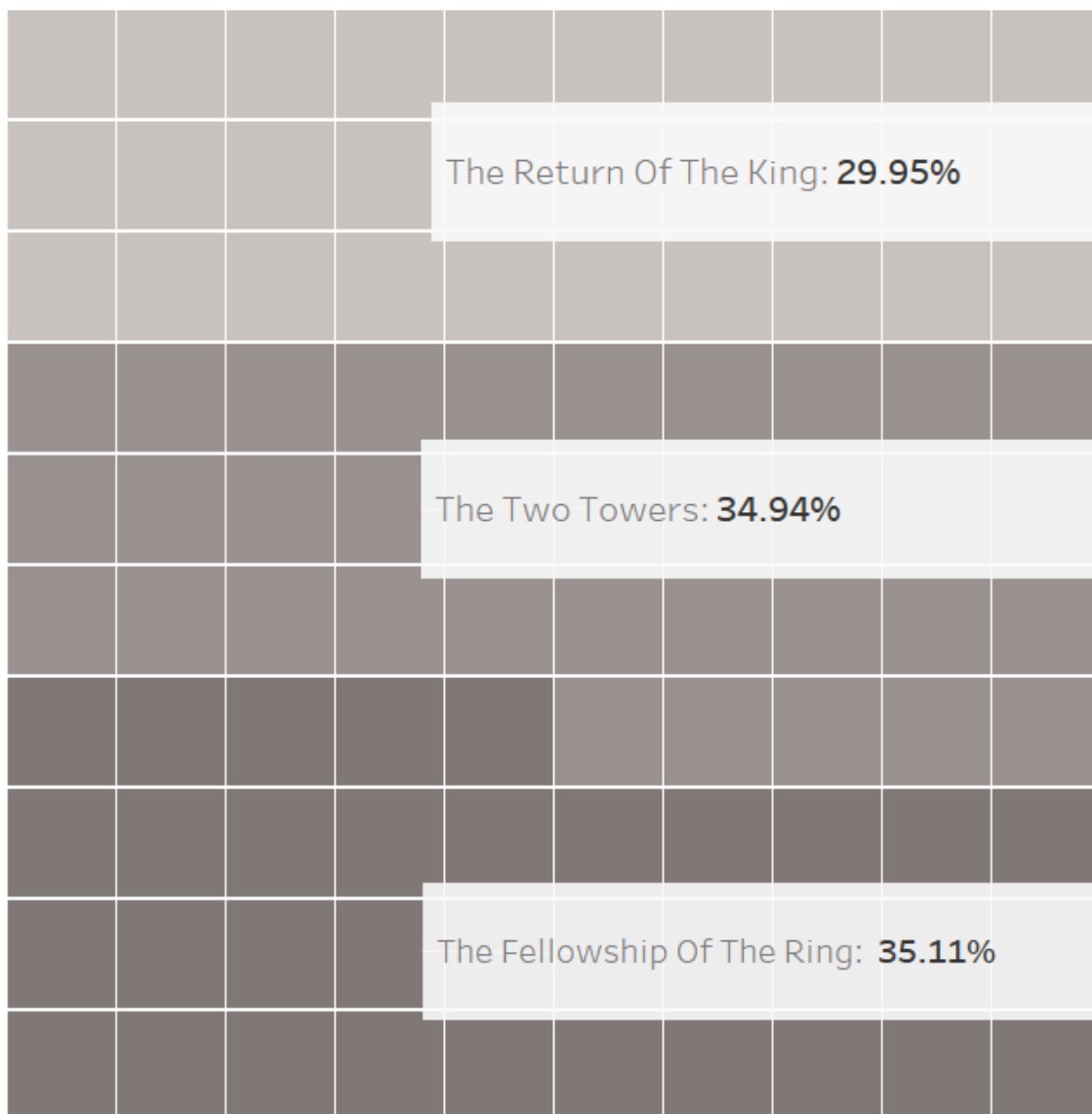
END
```

2. Add this calculation to **Color** :



3. Annotate the chart:





An example of a dashboard, after some cleanup and formatting, can be seen in the following screenshot:

What characters had the most to say in the Lord of the Rings Trilogy?  
Which trilogy had the largest percentage of words?

Film

(All)

