Tableau, Day #2

# Calculated fields, combined fields, rank fields, joins

From *Communicating Data with Tableau,* chapter 4: *Ratios and Rates*.

From the opening screen, choose *Connect | To a file | Microsoft Excel.* Connect to the NYC Trash Excel file.

Go to Sheet 1.

We are going to look at some ratios, but to do so, we need to create some new variables.

# Create the *Recyclable Tons Collected* calculated field.

Right-click on the *Data* area on the left side of Tableau. Select *Create Calculated Field*.

In the dialog box, replace *Calculation1* with *RecyclableTonsCollected.* The *RecyclableTonsCollected* will be the sum of the *PaperTonsCollected* and the *MGPTonsCollected* (Metal, Glass, Plastic).

Tab/click in the blank area below the field name. Drag *PaperTonsCollected* to the blank area. It will appear with square brackets around it. Add a plus sign to its right. Drag *MGPTonsCollected* to the right of the plus sign.

Click on OK.

# Create the *Recycle to Refuse Ratio* field.

Right-click on the *Data* area on the left side of Tableau. Select *Create Calculated Field.*

In the dialog box, replace *Calculation1* with *Recycle to Refuse Ratio*. This will be the sum of the *RecyclableTonsCollected* divided by the *RefuseTonsCollected*.

Tab/click in the blank area below the field name. Drag *RecyclableTonscollected* to the blank area. It will appear with square brackets around it. Add a slash to its right. Drag *RefuseTonsCollected* to the right of the slash.

As it stands now, this will compute a ratio for each row in our data. However, we want to find the ratio for the entire borough, so we need to sum the *RecyclableTonsCollected* and sum the RefuseTonsCollected. Make the following modifications to your formula:

SUM([RecyclableTonsCollected]) / SUM([Refuse Tons Collected])

Click on OK.

# Create a bar chart visualization

Drag *Borough* to the *Rows* shelf:

Text

Description automatically generated with medium confidence

Drag *RecyclableToRefuseRatio* to the *Columns* shelf:

Graphical user interface, application

Description automatically generated

Sort the bars in descending order:

.

This will put Manhattan at the top and the Bronx at the bottom:

Chart

Description automatically generated with medium confidence

Change the colors of the bars by dragging the *Borough* field to the *Color* shelf. NOTE: Do NOT drag it from the *Rows* shelf. Drag it from the fields list on the left. Our chart now looks like this:

Chart, timeline, bar chart

Description automatically generated

This compares the boroughs, but there is a lot of variation from one Community District to another.

*CommunityDistrict* is a number so Tableau has put it in the *Measures* group, but we don't want to add it; it is used to categorize, so drag it up to the *Dimensions* group (with *Borough, Month,* and *Measure Names*):

Graphical user interface, text, application

Description automatically generated

Drag *Borough* from the *Rows* shelf to the *Columns* shelf and put it to the left of *AGG(RecyclableToRefuseRatio)*.

Drag *CommunityDistrict* to the *Rows* shelf. Our chart now looks like this:

Chart, funnel chart

Description automatically generated

## Add numbers

Drag *Recyclable to Refuse Ratio* to the *Label* card in the *Marks* group:

Graphical user interface, application

Description automatically generated

This will cause numbers to be displayed at the end of each bar.

Chart, bar chart

Description automatically generated

## Format the numbers

In the *Marks* group, click on the down-arrow on the *AGG(Recyclable to Refuse Ratio)* pill and choose *Format…*:

Graphical user interface, text, application

Description automatically generated

In the *Format AGG* window on the left side of the screen, click on the *Pane* tab, then in the *Default* group, click on the down-arrow on the right edge of the *Numbers:* box:

Graphical user interface, application

Description automatically generated

Click on *Percentage* and reduce the number of decimal places to 0. Close the *Format AGG* window on the left side. to restore the *Dimensions* and *Measures* window.

Graphical user interface, application, Word

Description automatically generated

# A Highlight table

Create a new sheet by clicking on the *Add Sheet* button at the bottom:

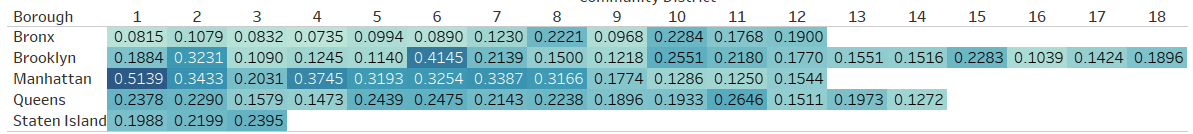
It should be named *Sheet 2* by default.

Select all of the following (ctrl+click): *Borough, Community District*, and *Recyclable to Refuse Ratio*. Then click on *Show Me* and create a *Highlight Table* (first row, last item):

Graphical user interface

Description automatically generated

You get a very wide table:



If you want to flip the rows and columns to get a very tall table, click on the *Transpose* button:



Your result will look like this:

Table

Description automatically generated

## Change the default colors

We can change the default colors. Click on the *Color* card in the *Marks* group. Click on *Edit Colors*.

Graphical user interface, application

Description automatically generated

If you select *Stepped Color,* you will get a fixed number of colors rather than a continuous palette of many colors. You can select a palette of *Green,* turned *Stepped Color* on, with 5 *Steps.*

Graphical user interface

Description automatically generated with low confidence

After selecting *Green,* click on *OK*:

Graphical user interface

Description automatically generated with medium confidence

The table now looks like this:

Table

Description automatically generated

## List ALL Community Districts in descending order of **Recyclable to Refuse Ratio**.

Go back to sheet 1.

We want all of our community districts to be listed in the bar chart in descending order.

We need to get all of the 59 community districts in descending order.

The solution is to create a *combined field* that is the *Borough* and the *Community District* combined (concatenated):

1. Click on *Borough* (in the fields list on the left0, and control-click on *Community District*.
2. Click on the down-arrow on the *Borough* pill (because we want *Borough* to be to the left of *Community District* in our combined field). From the menu, click on *Create*. Then click on *Combined Field*.

Graphical user interface, application

Description automatically generated

Drag the combined field to the *Rows* shelf. Remove the existing *Community District* pill from the *Rows* shelf. Remove the *Borough* pill from the *Columns* shelf. You will get the following:

Chart, bar chart

Description automatically generated

Click on the *Sort Descending* button on the toolbar:



Your chart will now look like this:

Chart, bar chart

Description automatically generated

# Adding Rank

We have 59 rows. To find the relative rank of a given borough/community district combination, we don't want to have to count by hand. We will add a rank field.

Right-click on the *Measures* and *Dimensions* box on the left side and choose *Create a calculated field*:

Graphical user interface, text, application, chat or text message

Description automatically generated

Set the name to *Rank.* Click on the small arrow on the right side of the window:

Graphical user interface, text, application, Word

Description automatically generated

A list of functions will appear:

Graphical user interface, application

Description automatically generated

Scroll through the function names on the right side and double-click on *Index()*.

Graphical user interface, application

Description automatically generated

As soon as the message "The calculation is valid" appears, click on the *Default Table Calculation* hyperlink.

Change the *Compute Using* drop-down (it should currently say *Automatic*) to *Borough & Community District (Combined)*:

Graphical user interface, text, application

Description automatically generated

After clicking OK twice, we should see a new field called *Rank* in the *Measures* group. Before we can add this to the chart, we have to make it discrete. Right-click on the *Rank* field and select *Convert to Discrete*.

Graphical user interface, text, application

Description automatically generated

Drag *Rank* to the left of *Borough & Community District (Combined)* on the *Rows* shelf.

It will appear at the left of the *Borough & Community District (Combined)* on the chart, like this:

Chart, bar chart

Description automatically generated

# Add an average line

Click on the Analytics tab on the left side of the screen:

Text

Description automatically generated with medium confidence

Click on *Average Line* and drag it to the chart area. WITHOUT LETTING GO OF THE MOUSE BUTTON, click on *Table* from the list of popup icons.

Graphical user interface, application, PowerPoint

Description automatically generated

Chart, bar chart

Description automatically generated

# Filtering

If you only want to see part of the data, you can use a filter. Click on the *Data* tab on the top left. Drag the *Borough* field to the *Filters* card:

Graphical user interface, application

Description automatically generated

Select the fields that you want and deselect the fields you don't want. In the dialog box below, select Manhattan and Brooklyn:

Graphical user interface, application

Description automatically generated

The resulting chart looks like this:

Chart, bar chart

Description automatically generated

# Adding Annotations

Right-click on the chart. Choose *Annotate*. Then choose the type of annotation you want to make:

Graphical user interface, application, table, Excel

Description automatically generated

Add your annotation:

Graphical user interface, text, application

Description automatically generated