MIS 0855 – Data Science (Section 002) – Fall 2015
In-Class Exercise (Day 20-21) – Visualizing Key Performance Indicators

Objective: Create a set of visualizations that enable comparison using Key Performance Indicators using Tableau

Learning Outcomes:

- Create heat maps to visualize key performance indicators.

- Use calculated fields to categorize acceptable and unacceptable performance.

- Visualize those calculate fields using easy-to-read symbols.

In this exercise, you'll be working with a modified and cleaned version of a dataset from the Bureau of Transportation Statistics (see http://www.transtats.bts.gov). The data set contains data about 86,657 airline flights originating from Philadelphia, New York, Phoenix, Pittsburgh, Los Angeles, and Chicago in January, 2014.

You are going to create a set of visualizations to determine the average performance of the airlines with regard to delays and cancellations.

Part 1: Take a look at the data set

1) Go to the Community Site and look for the post with this in-class exercise.

2) Right-click on the link to the data set (OnTime Airline Stats [Jan 2014].xlsx) and save it to your computer. Remember where you saved it!

3) Open the file in Excel and make sure you understand the data (take a look at the Data Dictionary tab too!). Each row is a flight and the departure and arrival delay time is listed along with the time to taxi out to the runway before the plane departs and the time to taxi in from the runway once the plane lands.

    There is also a field indicating whether the flight was cancelled (1 if yes, 0 if no). You'll notice that if the flight was cancelled the rest of the delay data for the row is not available.

4) Click on cell C2 (the AirlineFullName column) and you'll see a VLOOKUP formula:

    =VLOOKUP(B2,'Carrier Lookup'!$A$2:$B$15,2)

    This means the value in B2 is used to find the correct textual description of the source category. The lookup table is in the Carrier Lookup tab ('Carrier Lookup'!$A$2:$B$15,2). Finally, the second column of the lookup table contains the full airline name (2).

5) Click on the Carrier Lookup tab to see the lookup table.

6) Close the file.

Part 2: Start Tableau and open the data file

1) Start Tableau.

2) Click on "Microsoft Excel" under "In a file."

3) Navigate to the location where your data file is stored and select it.

4) You'll see a list of Excel worksheets at the left side of your screen. These are all the sheets contained within the workbook. Drag the "On-Time Statistics" sheet to the workspace.

5) Click "Sheet 1" to Go to Worksheet.


Part 3: Create a heat map to compare cancellations

Our first key performance indicator is flight cancellations. We want to know which airlines, in which cities, cancel the most flights.

1) Drag "Origin City Name" from Dimensions to the Columns shelf.

2) Drag "Airline Full Name" from Dimension to the Rows shelf.

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	Abc	Abc	Abc	Abc		Abc
Alaska Airlines	Abc	Abc		Abc	Abc	
American Airlines	Abc	Abc	Abc	Abc	Abc	Abc
Delta Air Lines	Abc	Abc	Abc	Abc	Abc	Abc
Envoy Air	Abc	Abc	Abc	Abc		Abc

3) Drag "Cancelled" from Measures to the Text icon in the Marks card.

4) Tableau will automatically compute Cancelled as a SUM. Change the measure to average by right-clicking SUM(Cancelled) and selecting Measure/Average.
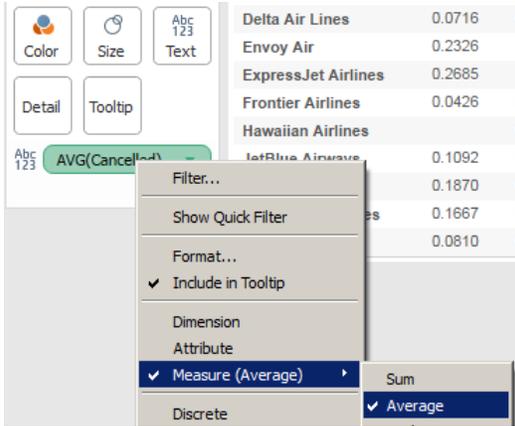
Why average instead of sum? Because the sum will just compute the total number of flights cancelled. We want the percentage of flights cancelled to allow for direct comparison between cities and airlines.

To see why taking the average computes the percentage of cancelled flights, remember that the value for cancelled is either 1 (meaning the flight was cancelled) or 0 (meaning it was not). So think about the case where there are four flights – two are cancelled, two are not – the average would be $\frac{(1+1+0+0)}{4} = \frac{2}{4} = .5$, or 50%.

5) You will now see a table of cancellation percentages by airline by city.

			Origin City Name			
Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelp hia, PA	Phoenix, AZ	Pittsburg h, PA
AirTran Airways	0.1073	0.0323	0.0922	0.1170		0.0307
Alaska Airlines	0.0000	0.0000		0.0323	0.0000	
American Airlines	0.0889	0.0216	0.0766	0.0569	0.0295	0.0265
Delta Air Lines	0.0716	0.0124	0.0729	0.0731	0.0206	0.0586
Envoy Air	0.2326	0.0912	0.1621	0.3333		0.1563
ExpressJet Airlines	0.2685		0.2340	0.2632	0.5000	0.1825
Frontier Airlines	0.0426	0.0179	0.0303		0.0074	
Hawaiian Airlines		0.0000	0.0000		0.0000	
JetBlue Airways	0.1092	0.1053	0.1281	0.0970	0.1356	0.0842
SkyWest Airlines	0.1870	0.0186			0.0075	0.1429
Southwest Airlines	0.1667	0.0283	0.1028	0.0977	0.0175	0.0466
United Airlines	0.0810	0.0233	0.0672	0.0505	0.0153	0.0190

6) Let's change these values to be easily readable percentages. Go to the Format menu and select Font…
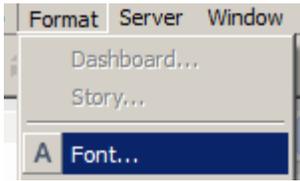
7) A "Format Font" tab will appear on the right. Select the Fields menu and choose AVG(Cancelled).
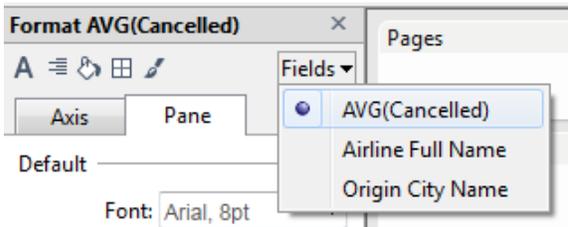
8) Under the Default section, choose Percentage from the Numbers menu. You'll now see the numbers formatted as percentages.


| | Origin City Name | | | | | |
Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelp hia, PA	Phoenix, AZ	Pittsburg h, PA
AirTran Airways	10.73%	3.23%	9.22%	11.70%		3.07%
Alaska Airlines	0.00%	0.00%		3.23%	0.00%	
American Airlines	8.89%	2.16%	7.66%	5.69%	2.95%	2.65%
Delta Air Lines	7.16%	1.24%	7.29%	7.31%	2.06%	5.86%
Envoy Air	23.26%	9.12%	16.21%	33.33%		15.63%
ExpressJet Airlines	26.85%		23.40%	26.32%	50.00%	18.25%
Frontier Airlines	4.26%	1.79%	3.03%		0.74%	
Hawaiian Airlines		0.00%	0.00%		0.00%	
JetBlue Airways	10.92%	10.53%	12.81%	9.70%	13.56%	8.42%
SkyWest Airlines	18.70%	1.86%			0.75%	14.29%
Southwest Airlines	16.67%	2.83%	10.28%	9.77%	1.75%	4.66%
United Airlines	8.10%	2.33%	6.72%	5.05%	1.53%	1.90%

9) But this is still not easy to compare. So under "Show Me," click on the heat map:
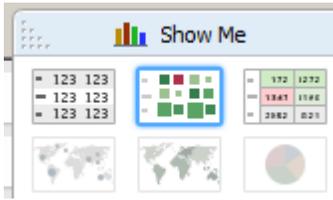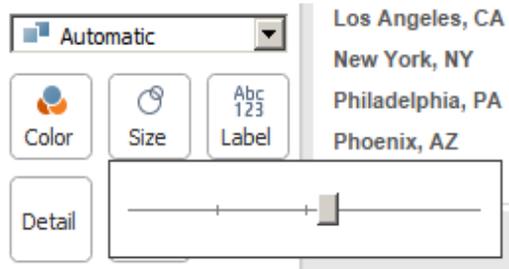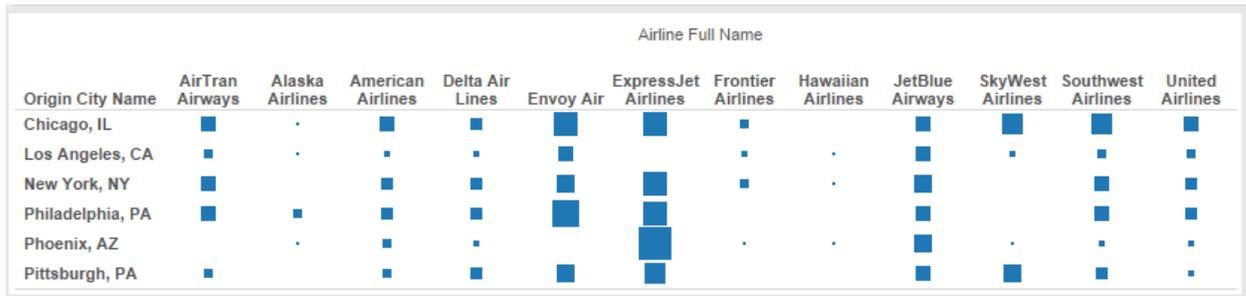

10) You'll see this:
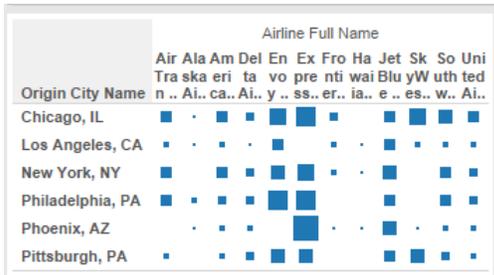
Which is still difficult to see. So go to the Format menu and select Cell Size/Wider. Do it again if it's still not easy to see.
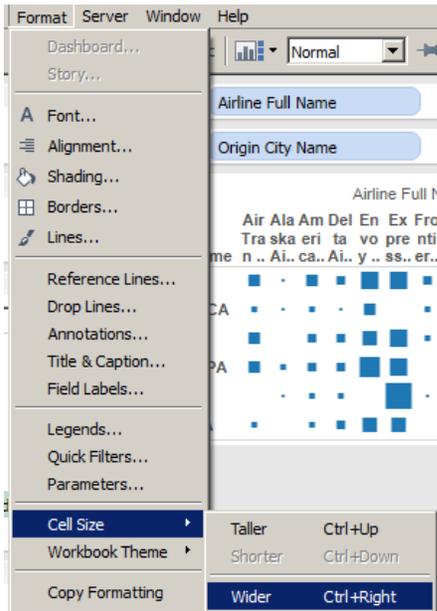
11) Now change the size of the icons by clicking on Size under the Marks card and moving the slider a little to the right of the midpoint.

Your chart should now look like this:

| | Airline Full Name | | | | | | | | | | | |
Origin City Name	AirTran Airways	Alaska Airlines	American Airlines	Delta Air Lines	Envoy Air	ExpressJet Airlines	Frontier Airlines	Hawaiian Airlines	JetBlue Airways	SkyWest Airlines	Southwest Airlines	United Airlines
Chicago, IL	■	·	■	■	■	■	■		■	■	■	■
Los Angeles, CA	■	·	■	■	■		■	·	■	·	■	■
New York, NY	■		■	■	■	■	■	·	■		■	■
Philadelphia, PA	■	■	■	■	■	■			■		■	■
Phoenix, AZ		·	■	■		■	·	·	■	·	■	■
Pittsburgh, PA	■		■	■	■	■			■	■	■	■

You can see that ExpressJet and Envoy Air have some serious cancellation issues!

Run your mouse over the icon representing ExpressJet flights originating in Phoenix and you'll see that the cancellation rate is 50%!
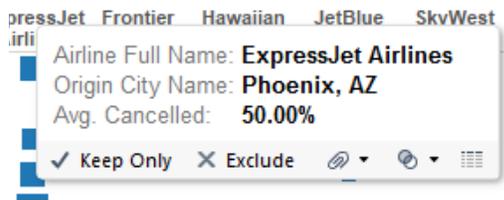
12) Name the worksheet "Airline Cancellations by City" and save the workbook.

Part 4: Create a KPI visualization of airports with significant departure delays by airline.

We want to create an easy-to-read visualization where we can quickly pick out which airline/city combinations have significant departure delays. We'll define a significant departure delay as greater than 15 minutes. So for an airline to be considered on time, their flights must have an average delay of 15 minutes or less.

Looking at a lot of numbers or text can be hard to do, so we'll make something very simple and graphical, but still gives us access to a lot of information at once.

1) Create a new worksheet by clicking on the [icon] icon or selecting New Worksheet from the Worksheet menu.

2) Drag "Origin City Name" from Dimensions to the Columns shelf.

3) Drag "Airline Full Name" from Dimension to the Rows shelf.

4) Now we need to create a calculated field that will differentiate between an average departure delay of greater than, or less than or equal to, 15 minutes.

5) Go to the Analysis menu and select Create Calculated Field.

6) Name the field "DepartureDelayKPI" and use the following formula:

IIF(AVG([Dep Delay Minutes])<=15,"Satisfactory","Unsatisfactory")
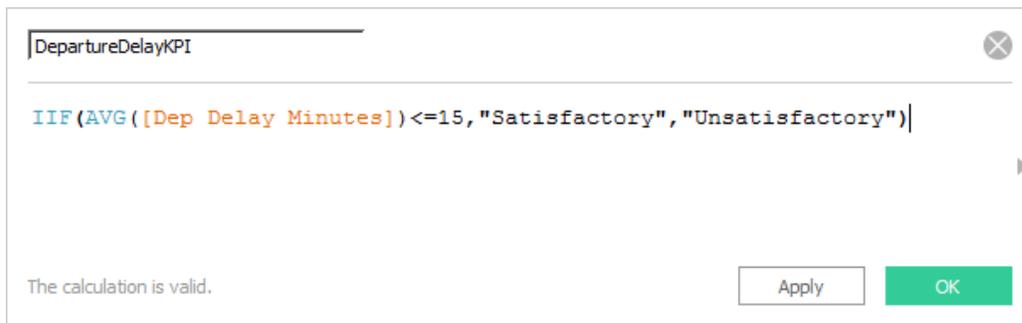
This formula will create our Key Performance Indicator. DepartureDelayKPI will carry one of two values: "Satisfactory" if the average departure delay is less than or equal to 15 minutes, and "Unsatisfactory" if the average departure delay is greater than 15 minutes.
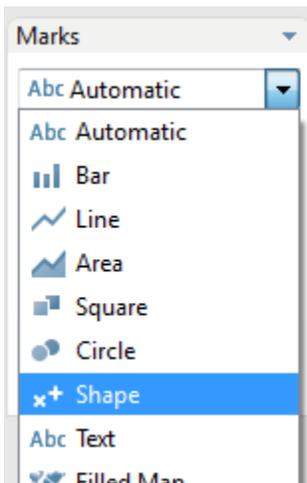
7) Click "OK" to create the new calculated field.

8) Drag the new DepartureDelayKPI field from Measures to the Text icon in the Marks box.

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelp hia, PA	Phoenix, AZ	Pittsburg h, PA
			Origin City Name			
AirTran Airways	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..		Satisfact..
Alaska Airlines	Unsatisfa..	Satisfact..		Unsatisfa..	Satisfact..	
American Airlines	Unsatisfa..	Satisfact..	Unsatisfa..	Unsatisfa..	Satisfact..	Satisfact..
Delta Air Lines	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..
Envoy Air	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..		Unsatisfa..
ExpressJet Airlines	Unsatisfa..		Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..
Frontier Airlines	Unsatisfa..	Unsatisfa..	Unsatisfa..		Unsatisfa..	
Hawaiian Airlines		Satisfact..	Unsatisfa..		Satisfact..	
JetBlue Airways	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Satisfact..
SkyWest Airlines	Unsatisfa..	Satisfact..			Satisfact..	Unsatisfa..
Southwest Airlines	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..	Unsatisfa..
United Airlines	Unsatisfa..	Satisfact..	Satisfact..	Unsatisfa..	Satisfact..	Satisfact..

This works, but it's hard to read all those "Satisfactory" and "Unsatisfactory" labels. So we'll use icons instead of text.

9) Using the menu under the Marks card, select Shape

Marks ▼

Abc Automatic ▾

- Abc Automatic
- ıl Bar
- ⟋ Line
- ⊿ Area
- ⊪ Square
- ⬤ Circle
- x⁺ Shape
- Abc Text
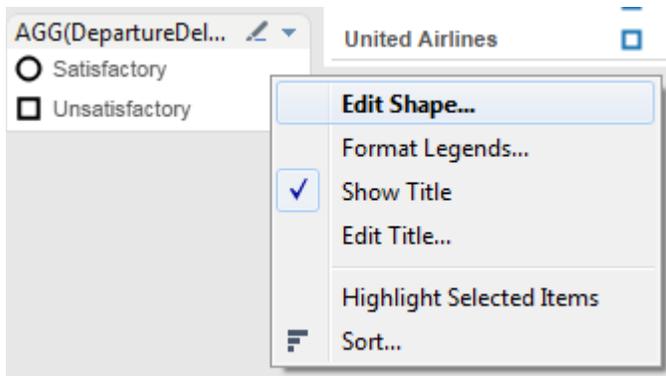- ⊙ Filled Map

10) Drag AGG(DepartureDelayKPI) from its location under the Marks card to the Shape icon.
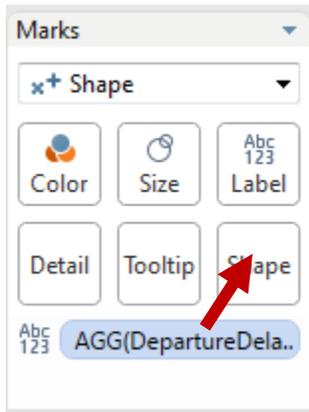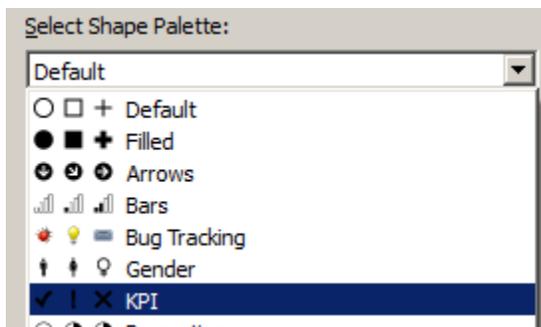

You'll now see circles (for Satisfactory) and squares (for unsatisfactory). But we can make the icons even clearer!

11) Right-click inside the AGG(DepartureDelayKPI) card – it's right below the Marks card. Select "Edit Shape…"

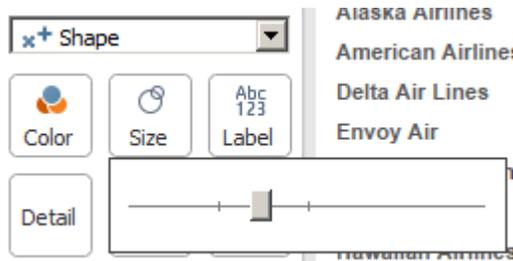12) Use the menu under "Select Shape Pallete" to select KPI.

13) Click once on the ◯ Satisfactory data item and then click once on the green check (✔).

14) Click once on the ☐ Unsatisfactory data item and then click once on the red X (✖).

15) Click "OK" to accept the changes and you'll now see checks and Xs.

16) Click the Size icon under the Marks card and use the slider to increase the size of the checks and Xs until they are easy to see.

17) It should now look like this:

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	✖	✖	✖	✖		✔
Alaska Airlines	✖	✔		✖	✔	
American Airlines	✖	✔	✖	✖	✔	✔
Delta Air Lines	✖	✖	✖	✖	✖	✖
Envoy Air	✖	✖	✖	✖		✖
ExpressJet Airlines	✖		✖	✖	✖	✖
Frontier Airlines	✖	✖	✖		✖	
Hawaiian Airlines		✔	✖		✔	
JetBlue Airways	✖	✖	✖	✖	✖	✔
SkyWest Airlines	✖	✔			✔	✖
Southwest Airlines	✖	✖	✖	✖	✖	✖
United Airlines	✖	✔	✔	✖	✔	✔

18) Run your mouse over the AirTran entry for Chicago. You'll see the tooltip indicates it is Unsatisfactory, but not how bad it really is.
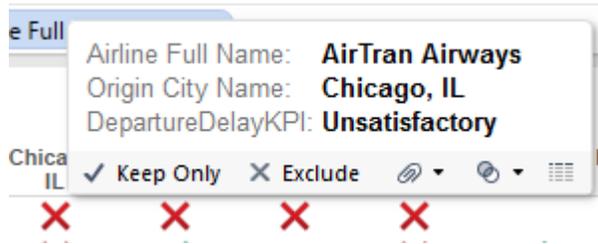
So now let's modify the tooltip to give more information.

19) Drag "Dep Delay Minutes" from Measures to the Tooltip icon under the Marks card.
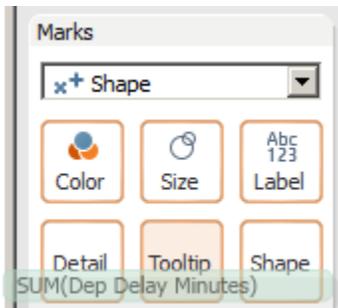
20) Right-click on SUM(Dep Delay Minutes) and select Measure/Average.

21) Now run your mouse again over the AirTran entry for Chicago.
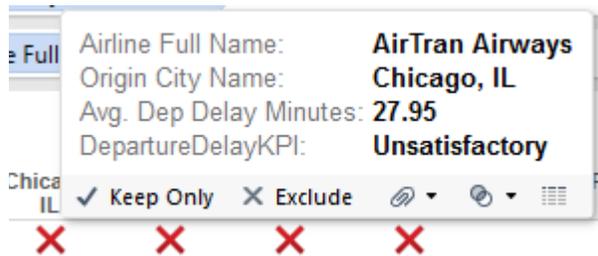
You now see that AirTrain flights out of Chicago experience, on average, delays of about 28 minutes!

22) Name the worksheet "Departure Delays" and save the workbook again so you don't lose your work!!

Part 5: Create an Airline Delay Dashboard

1) Create a new dashboard by clicking on the [icon] icon or selecting New Dashboard from the Dashboard menu.

2) Drag the "Airline Cancellations by City" worksheet onto the Dashboard canvas (the blank space on the right side of the window).

3) Drag the "Departure Delays" worksheet onto the Dashboard canvas, placing it under the "Airline Cancellation by City" visualization that is already there.

4) We don't need the legend for DepartureDelayKPI. It should be pretty obvious what the symbols mean. Delete the legend for DepartureDelayKPI by clicking once on it and then clicking the X.

5) However, we should explain a little more about what is being shown in this chart. Double-click on the words "Departure Delays" chart title. In the dialog enter the following text:
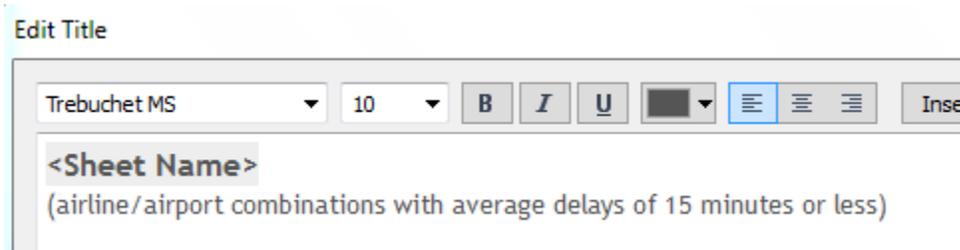
When you type the text it will be in the default font (Trebuchet MS, 12 point, bold). So highlight the text you entered and change the size to 10 point and unbold it.

6) Click OK and you will see this:

Airline Cancellations by City

Airline Full Name

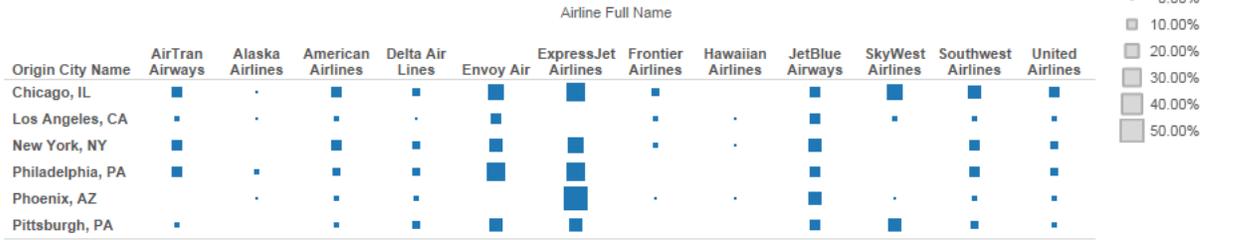| Origin City Name	AirTran Airways	Alaska Airlines	American Airlines	Delta Air Lines	Envoy Air	ExpressJet Airlines	Frontier Airlines	Hawaiian Airlines	JetBlue Airways	SkyWest Airlines	Southwest Airlines	United Airlines
Chicago, IL	■	·	■	■	■	■	■		■	■	■	■
Los Angeles, CA	▪	·	▪	·	■		▪	·	■	▪	▪	▪
New York, NY	■		■	■	■	■	▪	·	■		■	■
Philadelphia, PA	■	▪	■	■	■	■			■		■	■
Phoenix, AZ		·	▪	▪		■	·	·	■	·	▪	▪
Pittsburgh, PA	▪		▪	■	■	■			■	■	▪	▪

Avg. Cancelled
- 0.00%
- 10.00%
- 20.00%
- 30.00%
- 40.00%
- 50.00%

Departure Delays
(airline/airport combinations with average delays of 15 minutes or less)

Origin City Name

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	✗	✗	✗	✗		✓
Alaska Airlines	✗	✓		✗	✓	
American Airlines	✗	✓	✗	✗	✓	✓
Delta Air Lines	✗	✗	✗	✗	✗	✗
Envoy Air	✗	✗	✗	✗		✗
ExpressJet Airlines	✗		✗	✗	✗	✗
Frontier Airlines	✗	✗	✗		✗	
Hawaiian Airlines		✓	✗		✓	
JetBlue Airways	✗	✗	✗	✗	✗	✓
SkyWest Airlines	✗	✓			✓	✗
Southwest Airlines	✗	✗	✗	✗	✗	✗
United Airlines	✗	✓	✓	✗	✓	✓

7) Name the dashboard tab "Airline Delay Dashboard." And, of course, save your work!

Part 6: Try it on your own!

1) Create a new heatmap visualization that shows the average departure delay by city name and by airport.

   HINT: Use the Origin City Name and the Origin fields!

   Your chart should look something like this:

Origin City Name	Origin	
Chicago, IL	MDW	▮
	ORD	▮
Los Angeles, CA	LAX	▪
New York, NY	JFK	▮
	LGA	▮
Philadelphia, PA	PHL	▮
Phoenix, AZ	PHX	▪
Pittsburgh, PA	PIT	▮
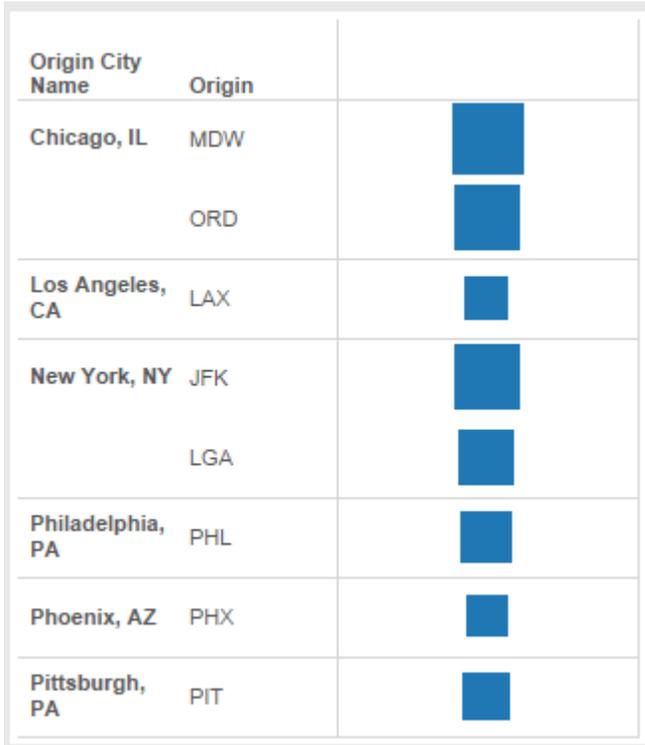
2) Create a visualization that shows airline/airport combinations with excessive taxi times (greater than 20 minutes).

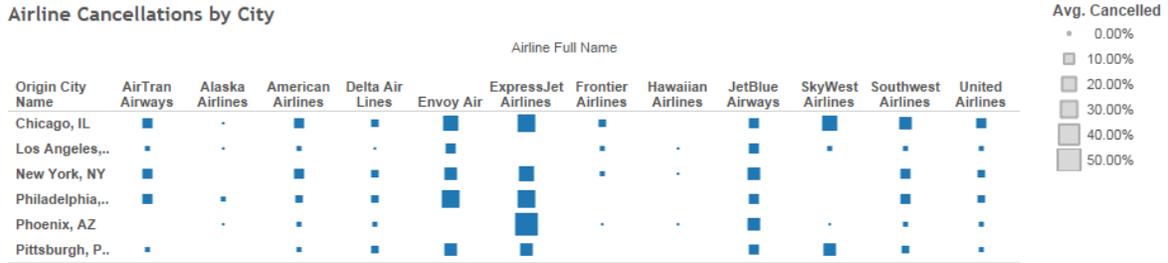| | Origin City Name | | | | | |
Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	✓	✓	✓	✗		✓
Alaska Airlines	✗	✓		✗	✓	
American Airlines	✓	✓	✗	✗	✓	✗
Delta Air Lines	✗	✓	✗	✗	✓	✗
Envoy Air	✗	✓	✗	✗		✓
ExpressJet Airlines	✓		✗	✗	✓	✓
Frontier Airlines	✗	✓	✗		✓	
Hawaiian Airlines		✓	✗		✓	
JetBlue Airways	✗	✓	✗	✗	✓	✓
SkyWest Airlines	✗	✓			✓	✓
Southwest Airlines	✓	✓	✓	✓	✓	✓
United Airlines	✓	✓	✗	✗	✓	✗

Here are some guidelines:

- You will need to create a new calculated field to compute the taxi time KPI. It will look a lot like the IIF formula we used back in Page 7!

- "Taxi Out" is the time it takes to taxi out to the runway (not "Taxi In").

- Use the same icons as we used before (green check, red X).

- Add the average taxi time to the tooltip.

3) Add the "taxi time" KPI visualization to the Airline Delay Dashboard you created earlier in the exercise.

Airline Cancellations by City

Avg. Cancelled
- 0.00%
- 10.00%
- 20.00%
- 30.00%
- 40.00%
- 50.00%

Departure Delays
(airline/airport combinations with average delays of 15 minutes or less)

Origin City Name

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	✗	✗	✗	✗		✓
Alaska Airlines	✗	✓		✗	✓	
American Airlines	✗	✓	✗	✗	✓	✓
Delta Air Lines	✗	✗	✗	✗	✗	✗
Envoy Air	✗	✗	✗	✗		✗
ExpressJet Airlines	✗		✗	✗	✗	✗
Frontier Airlines	✗	✗	✗		✗	
Hawaiian Airlines		✓	✗		✓	
JetBlue Airways	✗	✗	✗	✗	✗	✓
SkyWest Airlines	✗	✓			✓	✗
Southwest Airlines	✗	✗	✗	✗	✗	✗
United Airlines	✗	✓	✓	✗	✓	✓

Taxi Times
(airline/airport combinations with average taxing to takeoff of 20 minutes or less)

Origin City Name

Airline Full Name	Chicago, IL	Los Angeles, CA	New York, NY	Philadelphia, PA	Phoenix, AZ	Pittsburgh, PA
AirTran Airways	✓	✓	✓	✗		✓
Alaska Airlines	✗	✓		✗	✓	
American Airlines	✓	✓	✗	✗	✓	✗
Delta Air Lines	✗	✓	✗	✗	✓	✗
Envoy Air	✗	✓	✗	✗		✓
ExpressJet Airli..	✓		✗	✗	✓	✓
Frontier Airlines	✗	✓	✗		✓	
Hawaiian Airlines		✓	✗		✓	
JetBlue Airways	✗	✓	✗	✗	✓	✓
SkyWest Airlines	✗	✓			✓	✓
Southwest Airlin..	✓	✓	✓	✓	✓	✓
United Airlines	✓	✓	✗	✗	✓	✗