**Chapter 1 – No Lab**

**Chapter 2 – No Lab**

Although the pages in the book have all the labs included, it is written in a format which is very small print and also is written in paragraph format, rather than step by step. This lab manual is provided for those students who wish to follow the lab concepts “step by “step”. Each Chapter’s labs are labeled with page numbers, timing and data sources used.

Next, we will get our files to be used in the labs:

1. Go to github.com/onlc-classes
2. Find the XTBD-Data Visualizations link and click on it
3. Click the green “clone or download” button and save the .zip file to your desktop
4. Once it downloads, double click the .zip file and when it opens, copy and paste the “Workbooks” folder onto the desktop.
5. ALL Solutions are listed separately in the “Workbooks” folder, under each chapter. You will see completed .twbx files for all labs. You can use this to check your work, or to copy calculations, etc.

**Chapter 3 – How Much and How Many?**

**Lab Time: Approximately 45 Minutes (Pages 33-50)**

Data Sources:

* CDWT\_ch3\_DSNY.xlsx – (all about Garbage in 5 NYC boroughs – recycle vs. regular garbage)
* CDWT\_ch3\_RatSightings.xlsx

Completed Workbook(s):

* CDWT\_ch3\_DSNY.twbx
* CDWT\_ch3\_RatSightings.twbx

LAB: Starts on Page 33 and takes you to page 50. – You will start with the simplest of charts but “spin” them in different ways to get to the point of How Much and How Many!

**Pages 33 - 34**

**How Much? - The Bar Chart**

1. Start a new workbook.
2. Use the CDWT\_ch3\_DSNY.xlsx data source located in the Chapter 3 folder.
3. Click the Borough Dimension THEN hold down your CTRL key and also click the Refuse Tons Collected Measure.

Look at your Show Me! Tableau has determined the best type of charts for your two fields.

1. Take one of the suggestions (Pie Chart). Can you get a good idea of how much and how many (comparisons of Boroughs?) The answer is NO.
2. Although boring, next choose the Horizontal Bar Chart.
3. Next, use the “Sort and Swap” options on the Toolbar. Check page 37 so that you can see different options. Does this more readily tell “How Much and How Many”? YES!
4. However, we would like to be sure we understand the analytics behind this sort, and we would like to perform a “custom” sort:
   1. Turn to Page 38.
   2. Select the Borough Pill.
   3. Choose SORT.
   4. Sort by the field of Refuse Tons Collected and the aggregate of Sum – then try some different options in order to see differences in the data!
5. Rename the tab at the bottom to Bar.
6. Save the .twbx file onto the desktop as Chapter 3 – How Much and How Many.

**Pages 39 - 42**

**How Much? - The Dot Chart**

1. Duplicate the Bar sheet. (Right click the Bar tab and choose Duplicate)
2. Rename the tab at the bottom to Dot Chart.
3. Swap the pills so that the Borough pill is in columns and the Refuse Tons Collected pill is in Rows.
4. Go to the Show Me tool and choose the Circle View. Notice that the result is somewhat like the bar chart, in the sense that each borough is separate. If you hover over each circle, you can see the Refuse Tons Collected for each borough.
5. To resolve the overlapped circles, on Page 39, drag the Borough pill to columns and Refuse Tons Collected to rows.
6. Click the Swap Rows and Columns button to rotate the view from vertical to horizontal. Leave the view horizontal.
7. Play with the size button to get bigger circles. Swap the pills again, the circles change again. Click Undo.
8. Go to the Marks shelf and change the chart type to Circle. Notice the filled in circles. They are all blue.
9. Drag the Borough pill from the Data Pane to the Color shelf. Click the Color shelf and put a border around the circles. Notice the nice visual effect.
10. Lastly, you will build a formatted dot chart similar to page 42. The lines won’t be exactly like they are in the book as we would have to build a Dual Axis chart to accomplish this visualization (See note below).

To do this, start on page 41:

* Be sure the Marks card chart type is circles (and size them bigger)
* Change the view to Entire View
* Click Format menu and choose Lines
* Click the Rows tab and activate Grid Lines.

NOTE: At the bottom of page 41, they explain that the actual figure you see on Page 42, was done with “Dual Axis”. So for now, just know that your chart will have gridlines below the circles. To see the solution that is pictured on Page 42, you can open the SOLUTION in the Workbooks folder called CDWT\_ch3\_DSNY.twbx. It is called “dot chart” on the tab. You will see the dual axis as the reason that your chart does not look exactly the same as the picture on page 42.

**Pages 44 - 45**

**How Many? – Finding Errors with Number of Records**

1. Create a new worksheet. Name it Number of Records.
2. Create a new data source using CDWT\_ch3\_RatSightings.xlsx.
3. Drag the Borough dimension to Rows.
4. Drag the Number of Records measure to the Text shelf.
5. Notice that the five boroughs in NYC are accounted for, but there is also an “unspecified” borough.
6. To find this error, we will need to change the Rows pill from Borough to City in order to see which city IN the boroughs is reporting this error.
7. When you do this, notice that Brooklyn appears TWICE. ONLY… the second Brooklyn is spelled wrong. And, it is the reason your Borough information has one specified error. We are not going to fix this error at this time, but in the “real world” it would have to be fixed.
8. Clear the sheet.
9. Move Number of Records measure to Columns.
10. Move Borough Dimension to Rows.
11. Change the View to Entire View.
12. Go to the Marks card and change the chart type to Circle.
13. Compare your view to Page 46. Notice the “unspecified” in the Y Axis.

**Pages 46 - 50**

**How Many of How Much? - Histograms**

1. Create a new worksheet. Name it Histogram.
2. Change to the DSNY\_Collection\_Tonnages data source.
3. Click the Refuse Tons Collected measure.
4. On the Show Me tool, click the Histogram option. Notice you get a new dimension called Refuse Tons

Collected (bin). A bin is a range of data or “buckets” of data.

1. Click the Refuse Tons Collected (bin) pill in the Data Pane and choose the Edit menu option. The default bin size is set to 444. The bin size can be changed to make the most sense for your data.
2. Change the bin size to a value of 100. Notice the gaps in the data.
3. Click the Refuse Tons Collected (bin) in Columns and change from Continuous to Discrete.
4. Click the Refuse Tons Collected (bin) in Columns (again) and remove the check box from Show Missing Values. The gaps will now be gone.
5. Lastly, you will build the chart on Page 50. There is a real push these days to use muted colors such as grey and black. This affords those with color challenges to see the same view as someone without them.
   1. Clear the sheet.
   2. Drag the Refuse Tons Collected (bin) to Columns.
   3. Edit the Refuse Tons Collected (bin) pill to be 500.
   4. Drag the Refuse Tons Collected measure to Rows.
   5. Change the Refuse Tons Collected measure in Rows to CNTD (Count Distinct).
   6. Click the Color shelf to pick a grey color and use a border for the outside.

**Chapter 4 – Ratios and Rates**

**Lab Time: Approximately 45 Minutes (Pages 52-66)**

Data Sources:

* CDWT\_ch4\_DSNY.xlsx
* CDWT\_ch4\_NYCDistrictPop.xlsx

Completed Workbook(s):

* CDWT\_ch4\_DSNY.twbx

LAB: Starts on Page 52 and takes you to page 66

**Pages 52 - 54**

**Ratios**

1. Start a new workbook. Save it as Chapter 4 – Ratios and Rates.
2. Use the CDWT\_ch4\_DSNY.xlsx data source located in the Chapter 4 folder.
3. Click Sheet 1 in your workbook and name it Ratio 1.
4. Create two calculated fields: (type them or copy them from this document):

Name: Recyclable Tons Collected

Calculation: [MGP Tons Collected] + [Paper Tons Collected]

Name: Recycle to Refuse Ratio

Calculation: SUM([Recyclable Tons Collected]) / SUM([Refuse Tons Collected])

**Pages 54 – 55**

**Ratios - Bar Chart**

1. Drag the new Recycle to Refuse Ratio measure to Columns.
2. Drag the Borough dimension to Rows.
3. Drag a second Borough dimension to Color shelf.
4. Add a Descending Sort to Borough.
5. Check the results with the top of page 55.

**Pages 55 – 56**

**Ratios - Clearly Define w/Additional Formatting**

1. Move the Borough pill to Columns (so it is in front of Recycle to Refuse Ratio).
2. Drag the Community District measure up into the dimension area to convert it to a dimension.
3. Drag the Community District dimension to Rows. Check your work with the bottom of Page 55.
4. On Page 56 – you will add a label to show the exact ratio value. CTRL DRAG the Recycle To Refuse Ratio measure to the Label shelf in the Marks card. You will see the percentage.
5. Format these numbers if desired. Click the pill in the Marks card, choose Format. In the format screen on left, ensure you are on the Pane tab. Click the Default Numbers and choose Percentage. Choose 1 decimal point.

**Pages 57 – 58**

**Ratios - Highlight Table**

1. Create a new worksheet and name it Ratio 2.
2. Drag the Community District dimension to Columns.
3. Drag the Borough dimension to Rows.
4. Drag Recycle To Refuse Ratio measure to the Color shelf.
5. Drag the Recycle To Refuse Ratio measure to the Label shelf.
6. Change the Chart Type to Square in the Marks card.
7. Check your work with the bottom of page 57 (Change colors to “green” pallet to exactly match the book if desired).
8. Click the Swap button on the toolbar to produce the chart on Page 58 and check your work.

**Pages 58 – 59**

**Ratios - Bars Chart w/Combined Field**

1. Create a new worksheet and name it Ratio 3.
2. CTRL Click the Borough and Community District fields. Right click one of them. Choose Create. Click Combined Field.
3. Ensure that Borough is the first field and Community District is the second field in the Combined Field.
4. Drag the new combined field to Rows.
5. Drag the Recycle To Refuse Ratio measure to Columns.
6. Drag the Borough dimension to the Color shelf.
7. Use the sort tool to apply a Descending order.
8. Look for patterns in the data. Districts with higher or lower ratios. Compare your work to page 59.

**Pages 60 – 61**

**Ratios using Rank (Index)**

1. Duplicate the Ratio 3 Tab. Rename to Rank 1.
2. Create a new calculated field (type them or copy them from this document)::

Name: Rank

Calculation: Index()

DON’T CLICK OK YET

1. Click the blue link that says Default Table Calculation.
2. Choose Borough & Community District (Combined) for the Compute Using option. Leave all other fields as default.
3. Click OK twice.
4. Right click the new Rank field and choose Convert to Discrete.
5. Drag your new Rank pill to the left of the Borough Combined field on Rows.
6. Check your work with page 61.

**Pages 61 - 62**

**Ratio using Rank Table Calculation**

1. Duplicate the Rank 1 tab and rename it to Rank 2.
2. Remove the Rank pill from Rows.
3. Add the Recycle to Refuse Ratio pill to Rows.
4. Choose the down arrow. Select Quick Table Calculation and choose Rank.
5. Change the pill to Discrete.
6. Move the Recycle to Refuse Ratio to the left of the combined pill in Rows.
7. Right click Recycle to Refuse Ratio and choose Compute Using.
8. Click Borough Community District (Combined).

**Pages 62 – 63**

**Ratio as a Discrete Pill**

1. Duplicate the Rank 2 tab and rename it to Rank 3.
2. Add the Recycle To Refuse Ratio measure to Rows to the right of the Borough & Community District Combined Field.
3. Right click the Recycle To Refuse Ratio pill and click Discrete.
4. Notice the rank field which helps to number the rows in order from largest to smallest.
5. Format Recycle To Refuse Ratio pill to display 3 decimals.

**Pages 63 – 66**

**Rates – Blending Data Sources**

1. Create a new worksheet and name it Rates.
2. Click the Data menu and select New Data Source.
3. Select CDWT\_ch4\_NYCDistrictPop.xlsx in the Chapter 4 folder.

Tableau will take you to the main data source tab to build the data you want to use.

1. Double click 2010 Population.
2. Click the Rates tab in your workbook.
3. Create a new calculation (page 65):

Name: Refuse Per Person (you will use fields from both data sources):

* 1. Ensure you are on the DSNY\_Collection\_Tonnages data source.
  2. Type in 2000 \* SUM([Refuse Tons Collected]) / SUM(
  3. Type Population and select from the 2010 Population (CDWT\_ch4\_NYCDistrictPop) data source in the drop down.
  4. Your finished calculation will read:

2000 \* SUM([Refuse Tons Collected]) / SUM([2010 Population (CDWT\_ch4\_NYCDistrictPop)].[Population])

1. Drag your new Refuse Per Person measure to Columns. An error will display stating that you need a relationship between the two fields.
2. Click OK to clear the error.
3. Drag Borough to Color shelf.
4. Drag combined field to Rows and apply a Descending sort.
5. Drag Rank to the left of the combined field on Rows.
6. Drag Refuse Per Person to the right of Combined field in Rows. Change to Discrete.
7. Check your work with page 66 – it may look a bit different. But look at the benefits of this data on your own chart - Rank, combined field and Refuse per Person.

**Chapter 5 – Proportions and Percentages**

**Lab Time: Approximately 45-60 Minutes (Pages 71-85)**

Data Sources:

* CDWT\_ch5\_2012NYStats.xlsx
* CDWT\_ch5\_AL\_HomeRuns\_2011.xlsx

Completed Workbook(s):

* CDWT\_ch5\_2012NYStats.twbx
* CDWT\_ch5\_AL\_HomeRuns\_2011.twbx

LAB: Starts on Page 71 and takes you to page 85 – All about baseball. It might be helpful to double click the excel files and have a look at the data before beginning. A few helpful terms:

BA – Batting Average

HR – Home Run

AB – At Bat

RBI – Run Batted In

DH – Designated Hitter

POS – Position

**Pages 70 – 81**

**Part-to-Whole**

1. Create a new workbook.
2. Create the Excel data source using CDWT\_ch5\_2012NYStats.xlsx in the Chapter 5 folder. Use the 2012 Team Player Stats sheet.
3. Click the Data menu and select New Data Source.
4. Select CDWT\_ch5\_AL\_HomeRuns\_2011.xlsx in the Chapter 5 folder.
5. Save the workbook as Chapter 5 – Proportions and Percentages.

**Pages 70 – 71**

**At least one “At Bat” per Yankee**

1. Create a new tab and name it Bat Avg.
2. Select the 2012 Team Player Stats data source.
3. Drag BA measure to Columns.
4. Drag Name dimension to Rows.
5. Sort in Descending order.
6. Check your work on Page 71.

**Pages 71 – 74**

**Filters/Quick Filters**

1. Continue with your Bat Avg sheet.
2. Right click the AB Pill in the Data Pane and choose Show Filter.
3. Play with the slider to show that people can change the amount of Batting Average. Leave it 100 and 683 for the range. On Page 73, they call this a quick filter, but it is no longer called this in Tableau.
4. Go to the Marks card and change the chart type to circle.
5. Create a new calculation called Rank (as you did in a previous lab for a different data source). Name it Rank and the calculation is: Index().
6. Drag Rank pill to Rows. Change it to Discrete and move it to the left of the Name pill.
7. Drag Pos pill to the Rows.
8. Drag BA pill to the Rows. Change it to Discrete.
9. Check your work with page 74.

**Pages 74 – 77**

**Table Calculations using Home Runs**

1. Create a new tab and call it Home Runs. Continue to use the 2012 Team Player Stats data source.
2. Drag HR to Columns.
3. Drag Rank to Rows and change it to Discrete.
4. Drag Name dimension to the right of it.
5. Sort in Descending order.
6. Change Compute Using on the Rank pill to Name.
7. Create a filter using the HR measure to show “At Least” one home run.
8. Check your work on top of page 75.
9. Add the HR measure to the Label shelf on the Marks card.
10. Click the Analysis menu, select Totals and click Show Column Grand Totals.
11. Check your work at the bottom of page 75.
12. Right click the HR pill (on the Label). Add a Quick Table Calculation : Percent of Total.
13. Drag HR pill from Label on Marks card to Rows. Change to Discrete.
14. Format HR pill on Rows to only one decimal.
15. Check your work on page 77.
16. CTRL drag the HR pill on Rows to Color. Feel free to re-color with greys or muted colors, or the color of your choice.

**Pages 79 – 81**

**Proportions as Waterfalls Using Gantt**

1. Create a new tab and name it Gantt. Continue to use the 2012 Team Player Stats data source.
2. Drag Name to Columns.
3. Go to the Marks card and change the type of chart to Gantt Bar.
4. Drag HR measure to Rows. Add a Quick Table calculation for Running Total.
5. Drag the HR pill to the Filters shelf.
6. Change the filter to At Least = 1.
7. Create a new calculated field:
   1. Name: HR (rev)
   2. Formula: [HR] \* -1 (read page 81 for some info on this)
8. Drag HR(rev) to Size shelf.
9. Sort the Name pill in Descending order by the HR (Sum) field.
10. Click Analysis menu. Choose Totals. Click Show Row Grand Totals.
11. Check your work by looking at Page 81.

**Pages 82 – 85**

**Current-to-Historical**

**The Bullet Graph**

1. Create a new tab and name it Bullet.
2. Select the AL Team Homeruns data source.
3. CTRL Click City, 2011 HR and 2012 HR.
4. Click the Bullet graph on the Show Me panel.
5. Right click the X-axis and Swap the Reference Line Fields.

We want to see how the 2012 home run totals compared to the previous year. See Page 83 to check your work.

1. Adding reference lines:
   1. Click the Analytics tab on the Data Pane.
   2. Drag the Reference Line onto the view and choose Table.
      1. Leave the field as is Sum(2012 HR) with Average. Change the line color to a color of choice and make it a bit thicker.
   3. Click and drag a second Reference Line onto the view and choose Table.
      1. Click Distribution.
      2. Notice the 60%, 80% of average on the Value computation.

**Chapter 6 – Mean and Median**

**Lab Time: Approximately 20-30 Minutes (Pages 92-95)**

Data Sources:

* CDWT\_ch6\_2012SalariesMLS.xlsx
* CDWT\_ch6\_2012StatsMLB.xlsx

Completed Workbook(s):

* CDWT\_ch6\_2012SalariesMLS.twbx
* CDWT\_ch6\_2012StatsMLB.twbx

The data is all about baseball. It might be helpful to double click the excel files and have a look at the data before beginning.

A few helpful terms:

BA – Batting Average

HR – Home Run

AB – At Bat

RBI – Run Batted In

DH – Designated Hitter

POS - Position

Get both of the above data sources into Tableau. You will need to click the first table in each data source for use in your workbook.

1. Create a new workbook.
2. Create the Excel data source using CDWT\_ch6\_2012StatsMLB.xlsx in the Chapter 6 folder. Select Sheet1 sheet.
3. Click the Data menu and select New Data Source. (Only need to add this if Non-Normal distribution is going to be created. Page 96-97)
4. Select CDWT\_ch6\_2012SalariesMLS.xlsx in the Chapter 6 folder. Select the 2012 MLS Salaries sheet. (Only need to add this if Non-Normal distribution is going to be created. Page 96-97)
5. Save the workbook as Chapter 6 – Mean and Median.

**Pages 88 – 90**

**Normal Distribution**

1. Select the Sheet1-2012StatsMLB data source.
2. Name the new tab Histogram.
3. Click RBI in the data pane.
4. Click Histogram in the Show Me panel.
5. Check results on page 90.

**Pages 90 - 91**

**The Box and Whisker Plot – Show Me**

1. Create a new tab and name it Box Plot-Show Me. Continue to use the Sheet1-2012StatsMLB data source.
2. CTRL Click the Player and Pos dimensions and the RBI measure.
3. Click Box and Whisker Plots button on the Show Me panel.
4. Check the results on page 91

**Pages 92 - 95**

**The Box and Whisker Plot – Build From Scratch**

1. Create a new tab and name it Box Plot-Build From Scratch. Continue to use the Sheet1-2012StatsMLB data source.
2. Drag the POS dimension to Columns.
3. Drag the RBI measure to Rows.
4. Drag the Player dimension to the Detail.
5. Drag Pos dimension to Filters. Uncheck DH.
6. Change the chart type in the Marks shelf to Circle.
7. Color the circles with light grey and a border of black.
8. Check your work on page 92.
9. Change the chart type to Line.
10. Drag the Player dimension you had on Detail to the Path shelf that is now active when Line was selected.
11. Add three reference lines to the chart as follows:
    1. Right click the Y-axis and select Add Reference Line
       1. Type: Distribution
       2. Scope: Per Cell
       3. Computation: Quantiles
       4. Label: None
       5. Click OK
    2. Right click the Y-axis and select Add Reference Line
       1. Type: Line
       2. Scope: Per Cell
       3. Line: SUM(RBI) – Minimum
       4. Label: None
       5. Click OK
    3. Right click the Y-axis and Add Reference Line
       1. Type: Line
       2. Scope: Per Cell
       3. Line: SUM(RBI) – Maximum
       4. Label: None
       5. Click OK
12. Right Click the RBI measure and drag to Rows to the right of the existing Sum(RBI). Choose AVG(RBI). You will now have two charts, one on top of the other.
13. Click the AVG(RBI) pill in Rows. Change the chart type in the Marks card to Circle.
14. Drag the Player pill out of the Marks card (to remove it).
15. Right click the AVG(RBI) pill on Rows and choose Dual Axis.
16. Right click the new Y-axis (to the right of your chart) and choose Synchronize Axis.
17. Click the AVG(RBI) pill on Rows. Click the Color shelf. Click Edit Colors. Choose Seattle Grays in the Palette drop down. Change both data values to new grey colors.

NOTE: Read through the rest of the chapter on your own to inform yourself about outliers.

**Chapter 7 – Variation and Uncertainty**

**Lab Time: Approximately 15-20 Minutes (Pages 108-111)**

NOTE: The instructor will use the solution for the Uncertainty portion of the lab to explain the “chess club” solution. It is too much work to create in a training session. So once you complete pages 108-111, open the SOLUTION (Chapter 7 lab, ConfidenceIntervals.twbx). The instructor will work through this with you to gain an understanding of “confidence” once different variation is present. The lab pages also discuss outliers, which we will cover in chapter 8.

Data Sources:

* CDWT\_ch7\_Eathquakes.xlsx

Completed Workbook(s):

* CDWT\_ch7\_Eathquakes.twbx

**Pages 108 - 111**

**Control Charts**

1. Create a new workbook.
2. Create the Excel data source using CDWT\_ch7\_Earthquakes.xlsx in the Chapter 7 folder.
3. Save the workbook as Chapter 7 – Variation and Uncertainty.
4. Create a new tab and call it Control.
5. Right click the Date & Time dimension. Select Change Data Type. Select Date & Time.
6. Drag Date & Time to Columns.
7. Drag Number of Records to Rows.
8. Select Fit Width view.
9. Right click the Y-axis and select Add Reference Line as follows:
   1. Type: Distribution
   2. Scope: Per Pane
   3. Computation Value: -3, 3 Standard Deviation
   4. Formatting: Red, Dotted Line – NO FILL
10. Right click the Y-axis again and select Add Reference Line as follows:
    1. Type: Line
    2. Scope: Per Pane
    3. Line Value: SUM(Number of Records) with Average
    4. Line Format: Blue Dotted
11. Check your work on page 109
12. Change the Year(Date & Time) pill to a continuous Month (the second month on the list).
13. Check your work on page 110.
14. Drag a second Number of Records measure to Rows. Change the chart type in the Marks card to Circle.
15. Right click the second Number of Records measure in Rows and choose Dual Axis.
16. Change the Month(Date & Time) pill back to the first Year value in the list. Check your work on the top of page 111.

**Chapter 8 – Multiple Quantities**

**Lab Time: Approximately 30-45 Minutes (Pages 127-148)**

Lab is fantastic and ends with a background image on a scatter plot. This lab works with Hockey Data! Have fun!

Data Sources:

* CDWT\_ch8\_NHLTop100.xlsx

Completed Workbook(s):

* CDWT\_ch8\_ NHLTop100.twbx

Image:

* wg.jpg (Wayne Gretzky)

A few helpful terms:

G – Goals

A – Assists

GP – Games Played

P - Points

**Page 127**

**Scatter Plots**

1. Create a new workbook.
2. Create the Excel data source using CDWT\_ch8\_NHLTop100.xlsx in the Chapter 8 folder.
3. Save the workbook as Chapter 8 – Multiple Quantities.
4. Create a new tab and name it Scatter 1.
5. Drag G to Columns (Goals).
6. Drag A to Rows (Assists).
7. Drag Player to Detail shelf.
8. Check your work on Page 127

**Pages 128 - 130**

**Scatter Plots**

1. Duplicate the Scatter 1 sheet.
2. Rename to Scatter 2.
3. Drag GP to Color shelf (Games Played).
4. Drag P to Size shelf (Points).
5. Change color if desired.
6. Check your work on page 128.
7. Change chart type to Circle on Marks card.
8. Change color to desired color – book depicts Orange-Blue-White Diverging.
9. Click the border tool in the Color shelf and put a black border around the circles.
10. Change the size a bit to make the circles bigger.
11. Check your work on page 129.
12. Drag the Player pill from Detail shelf to Label shelf in the Marks card.
13. Check your work on page 130.
14. Notice that Wayne Gretzky is an “outlier”.

**Tooltips (Pages 131 - 132):** Hover over data marks in your completed chart to see the tooltip. If there is anything you would like to change, click the tooltip shelf on the Marks card to update.

**Annotations (Page 133):** Any data mark can be annotated by performing a right the data mark, selecting Annotate and clicking Mark. The contents of the annotation can be altered as desired.

**Filters (Pages 134 – 135):** You can show filters for any data. Right click the pill and select Show Filter on any fields you desire (Goals, Players, Assists). This allows the visualization to be interactive for the end user. You can remove filters when needed by performing a right click the pill in the Filters shelf and selecting Remove.

**Pages 136 - 137**

**Background Image**

1. Continue on Scatter 2.
2. Click Map, Background Image. Choose the Top 100 (CDWT\_ch8\_NHLTop100) data source.
3. Click Add Image.
4. Click Browse button next to File or URL.
5. Navigate to the Workbooks folder and Chapter 8. Select the wg.jpg file and click Open.
6. Look at page 137 for confirmation on how to fill out the fields, but here they are in detail:
   1. X Field: Select G data field, change the RIGHT measurement to 1,000
   2. Y Field: Select A data field, change the TOP measurement to 2,200
   3. Drag the Washout slider to the right to lighten the picture (you can always edit this later)
7. The result? A beautiful chart with a washed out background picture of Wayne Gretzky!
8. Right click the chart and choose Trend lines. Click Show Trend Lines. There is only one trend line.
9. Add POS to the Color shelf. Tableau now adds a trend line for each position.

**Pages 137 - 140**

**Stacked Bar Chart**

1. Create a new tab and name it Stacked Bar.
2. Create 3 Calculated Fields as follows: (Shortcut: Copy and paste and change numerator and name)
   1. Goals Per Game: Sum([G]) / Sum([GP])
   2. Assists Per Game: Sum([A]) / Sum([GP])
   3. Points Per Game: Sum([P]) / Sum([GP])
3. Drag Player to Rows.
4. Drag Measure Values from the Measures section to Columns, creating bars for every measure.
5. Drag Measure Names from the Dimensions section to Color shelf in the Marks card.
6. CTRL click to select ALL BUT Assists Per Game and Goals per Game from the Measure Values shelf. Drag all selected pills off of the Measure Values shelf to remove from the view canvas.
7. If desired, change the colors by editing the two colors.
8. Click the Player pill and choose Sort. Apply a Descending order by Points Per Game.
9. Check Page 140 for your work.

**Pages 141 - 145**

**Regression and Trend Lines**

1. Create a new tab and name it Regression and Trend Lines.
2. Drag Shots to Columns.
3. Drag G to Rows (Goals).
4. Drag Player to the Detail shelf in the Marks card.
5. Drag Pos to the Color shelf in the Marks card.
6. Drag Player to the Label shelf in the Marks card.
7. Check your work on page 142.
8. Click the 16 nulls and select Filter Data.
9. Right click anywhere in the view canvas and select Trend Lines. Click Show Trend Lines.
10. Check your work on page 143.
11. Right click anywhere in the view canvas and select Trend Lines. Click Edit Trend Lines.
12. Uncheck the Allow a trend line per color option.
13. Check the Force y-intercept to zero option. NOTE: If there is not shot, there cannot be a goal.
14. Check your work on the top of page 144.
15. Hover on the trend line to view the slope, R-squared and P-value detail.
16. Compare to view on bottom of page 144.
17. Right click the trend line and select Describe Trend Model to see all statistics to support the trend model.

**Pages 146 - 148**

**Quadrant Chart**

1. Create a new tab and name it Scatter 3.
2. Drag Shots measure to Columns.
3. Drag G measure to Rows.
4. Drag Pos to Color shelf. Change chart type to Circle on Marks card. Apply a black border on circles.
5. Drag Player to Label shelf.
6. Add Two Reference Lines as follows:
   1. Reference on Y-Axis: Right click Y-axis and choose Add Reference Line
   2. Leave defaults and click OK
   3. Reference on X-Axis: Right click X-axis and choose Add Reference Line
   4. Leave defaults and click OK
7. Add an Area Annotation to each quadrant as displayed on Page 148.

**Chapter 9 – Changes over Time**

**Lab Time: Approximately 20 minutes (Pages 151-168)**

Data Sources:

* CDWT\_ch9\_StrikeoutsMLB.xlsx (sheet 1)

Completed Workbook(s):

* CDWT \_ch9\_StrikeoutsMLB.twbx

Helpful Terms:

SO – Strikeout

TMS – Teams in League Each Year

**Pages 151 - 154**

**Basic Line Chart**

1. Create a new workbook.
2. Create the Excel data source using CDWT\_ch9\_StrikeoutsMLB.xlsx in the Chapter 9 folder. Add Sheet1 to your data source.
3. Save the workbook as Chapter 9 – Changes Over Time.
4. Create a new tab and name it Line Chart.
5. Drag the Year dimension to Columns.
6. Drag the SO to Rows.
7. Check your work with Page 152.
8. Click the Analytics tab on the Data Pane. Drag a Linear Trend Line onto the view.
9. Check your work with Page 153.

**Pages 154 - 157**

**Dual Axis with Trend**

1. Create a new tab and name it Dual Trend.
2. Drag Year to Columns.
3. Drag SO to Rows.
4. Create new Calculated Field:

Name: Pitchers Per Team

Formula: Sum([#Pitch]) / Sum([Tms])

1. Drag Pitchers per Team onto the row, next to SO. There are now two charts.
2. Right click Pitches per Team and click Dual Axis.
3. Click the color legend and change the colors if desired.
4. Check your work on page 157.

**Pages 158 - 161**

**Connected Scatterplots**

1. Get a new tab and name it Connected Scatter.
2. Drag Pitches per Team to Columns.
3. Drag SO to Rows.
4. Drag Year to the Color shelf.
5. Drag a second Year pill to the Label shelf.
6. Edit the colors and put a border around the circles.
7. Change the chart type to Circle.
8. Right click the view and show the Trend Line.
9. Check your work on page 159.
10. Duplicate the sheet and name it Connected Scatter 2.
11. Change the chart type to Line on the Marks card.
12. Drag the Year pill to the Path shelf in the Marks card.
13. Pages 160-161 show you various formatting options that can be applied to this line chart. Feel free to do any desired formatting.

**Pages 162 - 165**

**Date Field Type**

1. Add a new data source using CDWT\_ch9\_NYCRatSightings.xlsx.
2. Create a new tab and name it Date Field Type.
3. Drag the Created Date dimension to Columns.
4. Drag the Number of Records measure to Rows.
5. Change the Created Date pill on Columns to be Continuous.

The line chart changed a bit, but notice that the pill is now green and there is not a hierarchical choice.

1. Change the Created Date pill back to Discrete.

The plus sign is available again to drill down on the dates.

1. Click the plus sign to drill down to quarter.
2. Check page 163 (on the bottom) for the first breakdown.
3. Continue to click the plus sign to break the chart down through all layers of the date hierarchy.
4. Drag the Borough field to the Color shelf. The breakdown now shows Borough, per day. Remove all of the date pills except the Month pill.
5. Change the Month date field back to Continuous.
6. Check the results on the top of page 165.

**Pages 166 - 171**

**The Timeline**

**Open the CDWT\_ch9\_PresidentialGantt.twbx to review a complex timeline using presidential data. The Presidential Chart contains filters that provide interactivity by the end user to choose a specific timeframe and/or political party to specify the information. If you would like to work on this lab from page 169-171 after class, feel free to do so by following the steps in the book.**

**Pages 171 - 179**

**The Slopegraph**

**Open the CDWT\_ch9\_Slopegraphs.twbx to review a complex slope graph using soccer club data. The Slope graph has complicated calculated fields, but you can use the techniques laid out in this chapter to create the visualization, if desired. The Slope Graph dashboard also contains filters that provide interactivity by the end user to choose a specific soccer club and/or stat to specify the information. If you would like to work on this lab from page 171-179 after class, feel free to do so by following the steps in the book.**

**Chapter 10 – Maps and Location**

**Lab Time: Approximately 25-35 Minutes (Pages 183-200)**

Data Sources:

* CDWT\_ch2\_NYBoroughs.xlsx (first table) (this is not a mistake, they are purposely using the data from chapter 2)
* World Indicators Saved Data Source using the Data by country sheet
* Sample – World Bank Indicators.xlsx in the Chapter 10 folder

Completed Workbook(s):

* CDWT \_ch10\_NYBoroughs.twbx
* CDWT\_ch10\_GlobalInternetUsage.twbx

**Pages 183**

**Symbol Maps – Three Different Ways To Create**

1. Create a new workbook.
2. Create a new data source using World Indicators Saved Data Source.
3. Create another new Excel data source using CDWT\_ch2\_NYBoroughs.xlsx in the Chapter 2 folder. Add Sheet1 to your data source.
4. Create another new Excel data source using Sample – World Bank Indicators.xlsx in the Chapter 10 folder. Add Data by country to your data source.
5. Save the workbook as Chapter 10 – Maps and Location.
6. Create a new tab and call it Symbol Map.
7. Select the CDWT\_ch2\_NYBoroughs.xlsx data source.
8. Double click the State pill.
9. CLEAR the sheet.
10. Click once on the State pill and click the map type with the red boarder on the Show Me panel.
11. CLEAR the sheet.
12. Drag the Latitude (generated) pill to Rows.
13. Drag the Longitude (generated) pill to Columns.
14. Drag State pill to the Detail shelf in the Marks card.

Tableau is super smart and will create the same map for all three approaches.

**Page 183**

**Symbol Maps - Encoding**

1. Create a new tab and call it Encoding.
2. Use the Sheet 1 CDWT\_ch2\_NYBoroughs data source.
3. Drag Longitude (generated) to Columns.
4. Drag Latitude (generated) to Rows.
5. Drag Population measure to the Size shelf in the Marks card.
6. Drag County dimension to the Detail shelf in the Marks card. Notice the State is also placed on the Detail shelf too. This is because State is at the top of the hierarchy in this data source.
7. Drag County dimension to the Label
8. Drag the State and County pills to the Label shelf in the Marks card.
9. Resize the circles larger using the Size shelf scroll bar so you can see how the population is more for some boroughs than others.
10. Drag County dimension to the Color shelf in the Marks card.
11. Check your work on page 186.

**Pages 186 - 187**

**Symbol Maps With Many Data Marks**

1. Create a new tab and name it Multiple Data Marks.
2. Select the World Indicators data source.
3. Double click the Country pill. A world map will display.
4. Drag Internet Usage pill to the Size shelf in the Marks card.
5. Drag Region to the Color shelf in the Marks card.
6. Check your work on page 187.

**Pages 190 - 192**

**Filled Map**

1. Create a new tab and name it Filled Map.
2. Select the World Indicators data source.
3. Double click the Country pill.
4. Click the drop down arrow in the Marks card and select Map OR click the second map icon on the second row in the Show Me panel.
5. Drag Internet Usage measure to the Color shelf in the Marks card.
6. Check your work on page 192.
7. Change the color on the Color Legend as desired.

**Pages 186 - 190**

**Filled Map With Rates**

1. Create a new tab and name it Filled Map With Rates.
2. Select the Data by country (Sample – World Bank Indicators) data source.
3. Create a calculated field:

Name: Internet Users

Formula: [Business: Internet users (per 100 people)] \* [Population: Total (count)] / 100

1. Double click the Country Name pill. Change to a filled map using one of the options demonstrated above.
2. Drag the Internet Users calculation pill to the Color shelf. Notice the coloring that tableau uses for a measure.
3. Change the color in the Color legend to “Orange--Blue- White-Diverging. Click Reversed.
4. Check your work on page 193. The colors will be slightly different as the color palettes have been enhanced in the recent version of Tableau.
5. Edit the colors on the map by clicking the Color legend.
6. Click the Advanced button.
7. Check Start and enter 0. Check End and enter 1B. Check Center and it should fill in automatically with 500M.
8. Click the Map Menu option and choose Map Layers. Try Light, Normal and Dark for changes in the background if desired.
9. Check results as the map will go through various stages as the book does from pages 193-196.

**Pages 197-200**

**Dual Axis Map**

1. Duplicate the Filled Map tab and name it Dual Axis Map.
2. Drag a second Latitude (generated) pill onto the row, to the right of the first one. This will create two identical maps.
3. Click the first instance of Latitude (generated) and remove the Internet Usage pill from the Color shelf.
4. Drag Region to the Color shelf in the Marks card.
5. Click the Color shelf and drag the opacity to the left to give a lighter, washed out color.
6. Click the second instance of Latitude (generated) and change the drop down to a Circle.
7. Change the color of the circles to grey.
8. Right click the second instance of Latitude (generated) pill in Rows and choose Dual Axis. Both maps will come together into one.

The final product has a filled map showing the regions with circles representing Internet Usage. It is a bit different from the results on Page 198-200, but the concept of a dual map is the same.

**Chapter 11 – Advanced Maps**

**Lab Time: Approximately 35-45 Minutes (Pages 204-223)**

Data Sources: We will not use ALL the data sources in this chapter. We will use:

* CDWT\_ch11\_NFLTeams.xlsx
* CDWT\_ch11\_2012Hurricanes.xlsx
* CDWT\_ch11\_USVotingDemographics.xlsx

Completed Workbook(s):

* CDWT\_ch11\_NFLTeams.twbx
* CDWT\_ch11\_2012Hurricanes.twbx
* CDWT\_ch11\_USVotingDemographics.twbx

**Pages 204-208**

**Maps and Shapes**

1. Create a new workbook.
2. Create a new Excel data source using CDWT\_ch11\_NFLTeams.xlsx in the Chapter 11 folder. Add the NFL Stadiums to the data source.
3. Create another new Excel data source using CDWT\_ch11\_2012Hurricanes.xlsx in the Chapter 11 folder. Add Sheet1 to your data source.
4. Create another new Excel data source using CDWT\_ch11\_USVotingDemographics.xlsx in the Chapter 11 folder. Add Sheet1 to your data source.
5. Save the workbook as Chapter 11 – Advanced Maps.
6. Create a new tab and name it NFL Teams.
7. CDWT\_ch11\_NFLTeams.xlsx
8. Double click the State dimension.
9. Drag Team(s) dimension to the Label shelf in the Marks card.

You will now see a blue dot in each city with an NFL team.

NOTE: To see the full data in the data source, click the tiny grid to the right of the word Dimension in the Data Pane. The points that are being made on page 205 regarding this data are not accurate. It is stated that some Team information does not include Latitude and Longitude information. Please ignore this information as Tableau will automatically fill in the Latitude and Longitude when a geographic pill is double clicked.

1. Close the grid.
2. Change the chart type on the Marks card to Shape.
3. Drag the Team(s) dimension to the Shape card. Each team now has a different symbol. Additionally, the Team(s) dimension can be dragged to the Color shelf in the Marks card to change the color of each shape.
4. Check the results on page 208 for each option.

**Pages 208 - 210**

**Custom Maps and Shapes**

1. Create a new tab and name it Custom NFL Stadiums. Continue to use the NFL Stadiums data source.
2. Double click the City dimension to get a default map. It will display as a Symbol map w/blue circles for the locations where the NFL teams reside.
3. Drag the Team(s) dimension to the label.
4. Change the Chart type in the Marks card to Shapes.
5. Drag Roof Type dimension to the Shapes card in the Marks card. Each team city gets a generic symbol for their Roof Type: Domed, Open, Retractable.
6. Follow these steps to customize the Roof Type symbols:
   1. Minimize Tableau.
   2. Double click the Workbooks folder on your desktop
   3. Double click the Chapter 11 folder. Notice there is a Stadiums folder in this directory. It contains the three shapes for each of the Roof Types.
   4. Right click the Stadiums folder and select Copy.
   5. Click the Documents link on the left. Double click My Tableau Repository. Double click Shapes.
   6. Right click under the folder listing. Click Paste.
   7. Close the windows explorer and maximize Tableau once again.
   8. Click the Shapes shelf. Click the drop down arrow under Select Shape Palette. The Stadiums palette is still not listed.
   9. Click the Reload Shapes button. Click the drop down arrow under Select Shape Palette. Click Stadiums.
   10. Associate each type of roof with the corresponding shape:
       1. Click Domed and choose the first “domed” shape
       2. Click Open and choose the second “open” shape
       3. Click Retractable and choose the last “retractable” shape
7. Click OK.
8. Click the Size shelf and slide the bar to the right to make your shapes big enough so you can see each of them.
9. Check the results on page 210. The background has been changed by selecting a map layer to be dark (Map Menu, Map Layers, Style: Dark)

**Pages 211 - 215**

**Custom Maps and Paths**

1. Create a new tab and name it Hurricane Map. Use the Sheet1 (CDWT\_ch11\_2012Hurricanes data source.
2. Drag Longitude measure to Columns (NOT generated field, use the actual data pill)
3. Drag Latitude measure to Rows (NOT generated field, use the actual data pill)
4. Drag Name dimension to the Label shelf in the Marks card. The blue circles represent the location of hurricanes in 2012.
5. Change the Chart type on the Marks card to Line. One connected line appears across the map, and also there is a Path shelf in the Marks card.
6. Drag the ISO time dimension to the Path shelf. Circles appear again.
7. Right click the ISO time pill and choose the SECOND More on the list. Click Hour.
8. CTRL drag the Name pill that is currently on the Label shelf and drag a second instance to the Color shelf. The chart should now resemble page 214.
9. Click the down arrow on the ISO Time pill and try different choices (Year, month, week, day, etc.) in order to see the differences in time of the hurricane.
10. CTRL drag the ISO Time pill to the Pages shelf to automate the map. This option works best when the ISO Time is set to the Continuous Day option.

**Pages 218 - 223**

**Custom Maps and Parameters**

1. Create a new tab and name it US Voting. Use the Sheet1 (CDWT\_ch11\_USVotingDemographics data source.
2. Review pages 218 and 219. The instructions state to create a parameter and a calculation, with the intention of making the chart interactive. This is A LOT of typing. To make this easier, perform the following:
   1. Make this instance of Tableau smaller and put it on one half of your screen.
   2. Open the workbooks folder. Open the Chapter 11 folder. Double click the US Voting Demographics solution workbook.
   3. Right click the parameter called Compare With and select Copy.
   4. Right click in the white/empty space in the data pane on Chapter 11 – Advanced Maps and select Paste.
   5. Repeat this process to copy/paste the Chosen Demographics calculation in the solution.
   6. Right click both of these pills after they are moved and choose Edit. Review the code.
3. Drag Chosen Demographic measure to Columns.
4. Drag Voted for Obama measure to Rows.
5. Drag State dimension to the Detail shelf in the Marks card.
6. Drag State ID dimension to the Label shelf in the Marks card.
7. Drag Voted for Obama measure to the Color card.
8. Change the drop down chart type in the Marks card to Map.
9. Change the color legend to Red-Green Diverging.
10. Add a trend line as you have done in previous labs (Go to analytics tab. Drag a linear Trend Line to the view canvas.
11. Right click the Compare With parameter and click Show Parameter Control.
12. Use the parameter on the right side of your screen to work through all the different demographics. Notice that your trend line follows the result set.
13. Check the results on page 221.

NOTE: On Page 222, there is another “dual axis” map that you can create by adding a second Chosen Demographic pill to the Columns. This is an optional visualization.

**Chapter 12 – No Lab**

**Chapter 13 – Building Dashboards**

**Lab Time: Approximately 60+ Minutes (Pages 243-275)**

**Chapter 14– Advanced Dashboards**

**Lab Time: Approximately 60+ Minutes (Pages 278-301)**