Hands-On Lab

Federal Revenue / Spending Visualization

Lab version: 1.0.0

Last updated: 6/27/2015

Contents

[Overview 3](#_Toc421994478)

[Exercise 1: Gather and prepare data 4](#_Toc421994479)

[Exercise 2 – Smooth & combine data 6](#_Toc421994480)

[Exercise 3 – Present the data 8](#_Toc421994481)

[Summary 9](#_Toc421994482)

Overview

For our first lab, we will be walking through this fairly common pattern for non-interactive visualizations:

* 1. Gather the data for the visualization
  2. Parse, sort, or otherwise organize that data, looking for patterns or insights
  3. Create a chart to give visual insight to the data
  4. Visually refine that chart

# Prerequisites

The following prerequisites are required to gain the most from this hands-on lab:

* + Microsoft Excel is required
  + Photoshop
    - Or Paint.NET (for PC) <http://www.dotpdn.com/downloads/pdn.html>
    - Or GIMP (for Mac or PC) <http://www.gimp.org/>

**Lab Structure**

This lab includes three exercises with the following tasks:

* + Finding and downloading a data set (link for monthly treasury report)
  + Refining that data set, creating a “rolling average” to smooth your data
  + Charting the data using Excel
  + Creating a chart snapshot and bringing it into an image manipulation program for visual refinement

# Estimated completion time

Completing this lab should take at least 60 minutes.

Exercise 1: Gather and prepare data

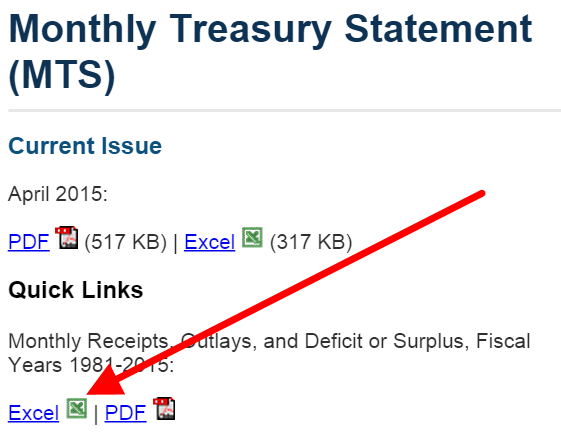
In this exercise we will download and view our data set and create the first chart to get a visual sense of the data.

**Task 1 - Download and view the data**

* 1. First let’s download our data set. Navigate to

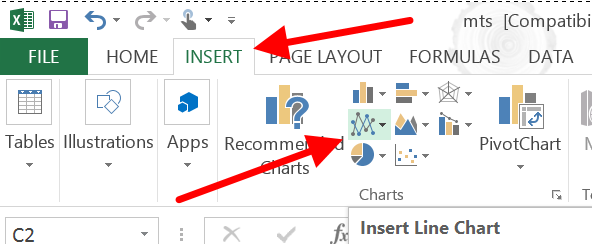
<https://www.fiscal.treasury.gov/fsreports/rpt/mthTreasStmt/current.htm>

And download the data for the “Monthly Receipts, Outlays, and Deficit or Surplus, Fiscal Years 1981-2015”

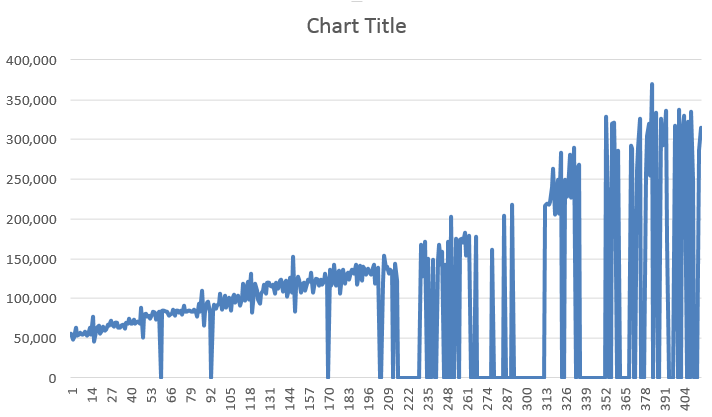


This will download as **mtr.xls**

* 1. Open it up. Delete columns E-G. This information is valuable in other contexts, but not right now.
  2. Let’s start by looking at federal spending (outlays). Highlight all the cells from C2 to C416. Then Click the “Insert” tab and click “Insert a Line Chart”

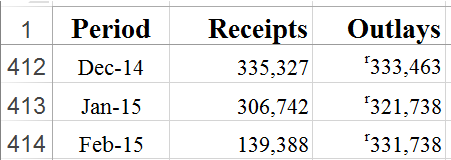


You should get something that looks like this:



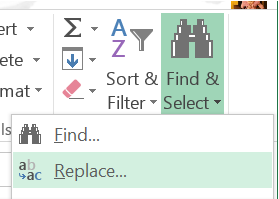
Does this look terrible? Yes it does.

This is because our data has been annotated in a way that makes it impossible for Excel to interpret. If we look at the Outlay cells for January, February, and March 2015, we see that they have all been marked with an “r” to indicate that the numbers have been “revised”.

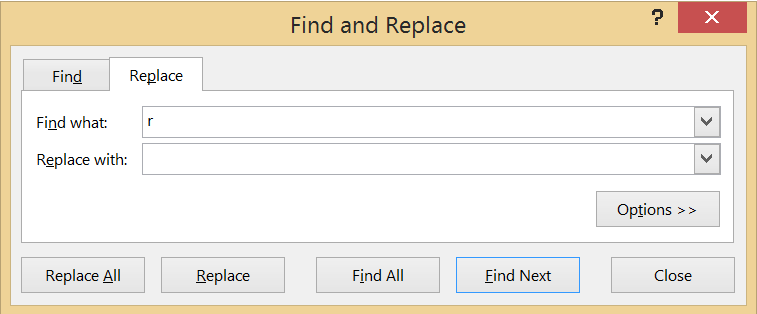


But Excel can’t chart numbers with an “r” in front of them. So let’s fix that.

* 1. From the “Home” tab, select “Find and Replace” or use the keyboard shortcut “ctrl-F” or “apple-F”



Click on the “Replace” tab and replace the letter “r” with nothing. Click “Replace All”



This will do some things we don’t want, like changing “Surplus” to “Suplus” but we don’t care about that. We only care that our data charts appropriately. Create a new chart using the previous method and it now looks like this.

Create another chart for the Revenue column (column B) by highlighting the B2-B416 and using the same method.

Exercise 2 – Smooth & combine data

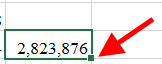
You may notice that our revenue chart is a little “dirty” with lots of dips and spikes. That is because the federal government collects revenue unevenly through the year with the biggest spike being in (you guessed it) April. Let’s smooth this data so we can get something a little more visually appealing.

* 1. Select cell E417 and type:

=SUM(B417:B406)

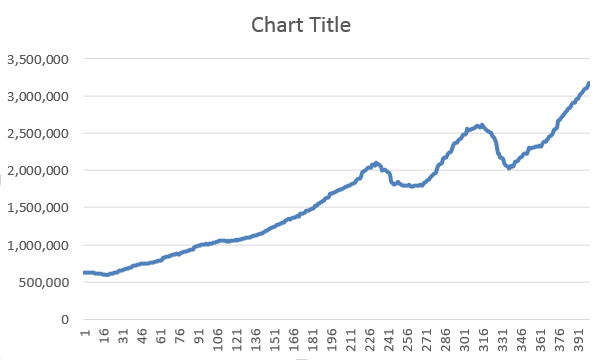
This adds together a full year of receipts.

With cell E417 selected, click on the cell “handle” (the colored square in the bottom right corner of the selected cell) and drag this formula up to the top of the Excel spreadsheet. Note: Your data may not match the number seen below since this data set is updated frequently.



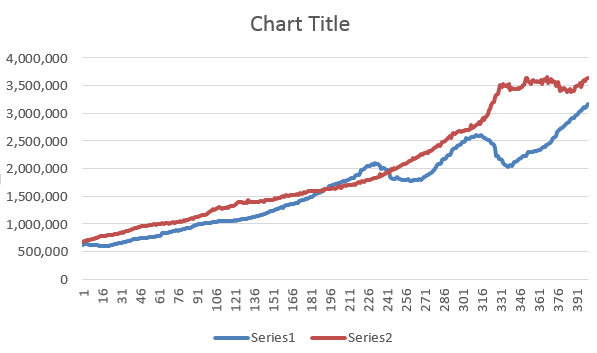
* 1. This will shift the references in those cells in all of the months back to 1982. We’ve just created a “rolling sum” calculation in which each cell contains a sum of the previous 12 months of receipts.

Now if we create our chart of the receipts it will look a little smoother.



And we can make some inferences and see that out chart is starting to tell a little bit of a story about the economy. We can see revenue dips during the most recent recessions and spikes in revenue during recovery periods.

* 1. Highlight the E416 cell and click on the handle and drag it to the right. This shifts the references in the cell so that they now refer to the “outlays” column. Drag that column up to 1982.
  2. Now highlight all the cells from E17 to F417 and create another chart.



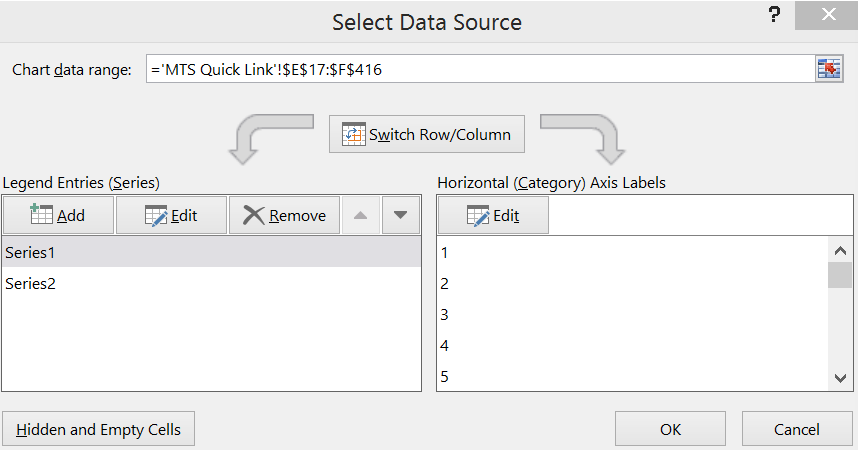
We can see federal spending and revenue together. The distance between the red line (spending) and the blue line (revenue) is the deficit. Or, for a short time in the late 90’s, the surplus.

But the chart by itself doesn’t tell this story. There is no good labeling, no months or years, just a set of abstract numbers that don’t mean anything unless we know what we’re looking at.

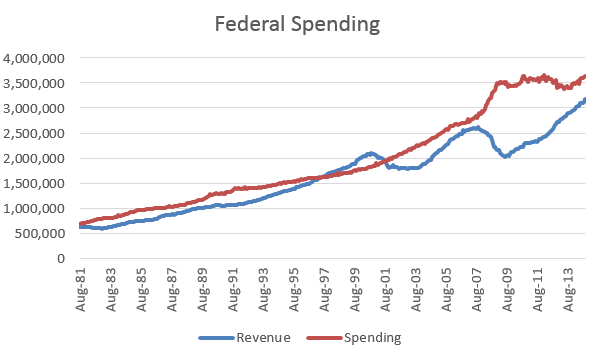
Exercise 3 – Present the data

Now that we have data that makes sense to us, we need to make sure it makes sense to other people as well. We’ll do this the “easy way “ and then the “better way”.

1. Click into the “Chart 1” text. Change it to “Federal Spending”
2. Right-click on the chart and choose “Select Data”. This will bring up the “Select Data Source” dialog.



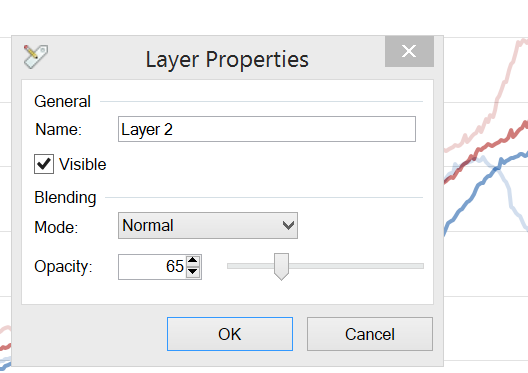
1. Select “Series1” from the “Legend Entries (Series)” section and change the name to “Revenue”
2. Select “Series2” and, using the same steps, change the name to “Spending”
3. Now let’s change the horizontal axis. Click the “Edit” button in the “Horizontal (Category) Axis Label” region. Click and drag to select the cells A13-A416.
4. Click “OK”. Your chart should look something like this:



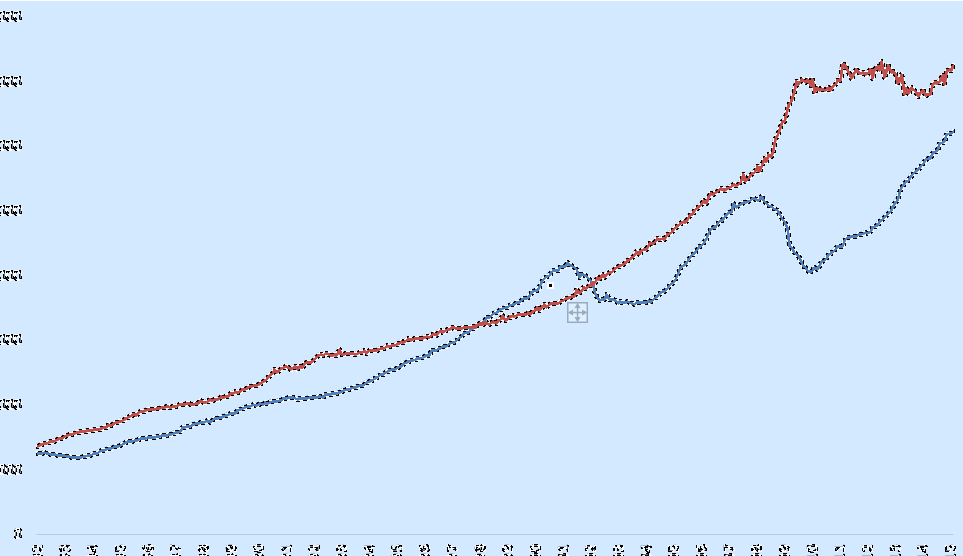
1. Now we have something that is starting to make sense to an outside observer. There are still a few difficult spots, such as the lack dollar signs on the spending scale and the fact that the scale is still in thousands (4,000,000 thousand is equal to $4 billion ($4,000,000,000)). We could release this as it if… or we could continue to improve upon it and get it ready for a more polished and formal release.
2. Make a copy of this chart by using the “snipping” tool. On a PC, press your Windows button and type “snipping tool” to open the Snipping Tool application. On a Mac, you can jump right into a snip by pressing Command-Shift-4.
3. With your Excel chart visible and sized to fill most of your screen, use your respective snipping tool to capture the graph as an image. Save it to the lab folder as “Fed Revenue And Receipts.png”.
4. Now (and this is going to sound weird so stick with me here) go back to Excel and select the horizontal lines on your graph by clicking on them. Hit the “delete” key to get rid of them and then snip that chart and save it as “Fed Revenue And Receipts Clean.png”.

The reason for doing this is so that we can have a very clean version of our spending trend line. That will be incredibly valuable as we design our final visual.

1. Open Photoshop (or Paint.NET for PC or GIMP for Mac) and load the both images into your application. (The following images will be from Paint.NET, so tool icons may not be exactly the same.)
2. Using the selection tool , select the “clean” version of the graph, copy it, and place it on a layer over the image of the graph with grid lines.
3. Right-click (or double click on the clean layer and set the opacity to 50%). Move it so that it is laying directly on top of the version of the graph with grid lines.



1. Using the “magic wand” tool, select the white portions of the “clean graph” and delete them. This will give us a layer where we only have the graph lines. We can draw our own more elegant graphing structure on top of this.



1. Draw vertical lines every two years (2014, 2012, etc) to indicate the timeline of congressional budgets. Now we can see how each congress has influenced spending.
2. There are many, many things we can do in terms of design with this graph. If we were redesigning this entirely, we might actually use the excel chart as a template and trace our own custom thickness graph on top of it. We could shade areas to show which president was in office during which periods of spending or revenue growth.
3. Things to try on your own:
   1. Change the Y axis labels to be 4 trillion, 3.5 trillion, etc
   2. Change the X axis labels to indicate years, ignoring the month data.
   3. Move the legend into the empty space on the graph in order to save space and maintain a clean visual.
   4. Alternate the coloring of the fiscal years in order to create visual differentiation
   5. See the example below (created a few years ago) as a point of comparison

