Data Exploration with Pivot Tables and Charts 2016

Excel files referred to in this five-part, rough introduction to pivot tables and pivot charts are available for download at datatools.tamscenter.com under the Data category, with filenames such as 1-VideoDemoDataAnalysis\_XXX.xls, 2-VideoDemo\_\_.xls, etc.

The demonstrations begin with a file downloaded from a commonly-used datalogger that collects ozone, PM2.5, and met data in hourly increments: 1-VideoDemoDataAnalysis\_RawData.xlsx.

**Data tips:**

* Use Cntrl-down arrow and Cntrl-up arrow, and right and left arrows to zoom around the sheet. (See **1-control-arrow** video.)
* When you have a lot of columns, highlight the last column of your data in yellow to better identify the end of your data. See **2-highlight last column** video.)
* Use the zoom slider at the bottom right to see more of your data, especially when you add a chart, so that you can more quickly see if you have selected all the data you intend to. (See **3-zoom slider** video.)
* Because blanks slow down excel and with this much data may cause it to crash with a lot of blanks, replace the blanks with a text string such as no. You can use find and replace, but be sure not to select the whole sheet, only the block of data or it will take a long time. (See the **4-get rid of blanks** video.)

 **Excel Table:**

1. Now, continuing with 2-VideoDemoDataAnalysisNoBlanks.xlsx:
	1. Review each (important) column using Autofilter, and check for error values, etc. (Negative values are fine as long as they are greater than the MDL for your method—see the AQS MDL file to see what the MDL is for your method; for TEOMPM10 it is -50 microg/m3 so do not invalidate negative values if they are above the MDL.) Note that because this file is so huge, the Autofilter doesn’t work well on the dates. There are error codes that need to be removed from this table, and noted in logbook/rows saved in an error code sheet in this file. Keep a copy of the rows that were edited as a sheet in the file. (See the **5-Autofilter** video.)
2. Now picking up at your file which should look like 3\_VideoDemoDataWoErrors.xlsx, continue:
	1. With the cursor anywhere in the table (that now has no blanks) click Insert, Table. Verify that excel selected only the block of contiguous data and no extra columns to the right. (See the **6-make excel** **sheet into excel Table** video.) This table will automatically expand as you add data to it, which will also automatically add to the pivot graph you will make, so that you will have a dynamic chart that will expand as you gradually add to your data table.
3. PivotTable:
	1. Now to insert a different kind of table, the Pivot Table, by Insert, Pivot Table: And ask excel to put the pivot table into a new sheet. (See the **7-create pivot table** video.) At this point, your data should look something like 4\_VideoDemoDataWPivotTable.xlsx.
	2. In this new sheet, there will be a PivotTable Field List, and drag the fields into the boxes so that they look like this:  (See the **8-dragging fields in pivot table field list** video.)
	3. If this panel (pivot table field list) disappears, click in its table on the left, or after clicking in the pivot table, click PivotTable Tools, Options, Field List. (See the **9-field list panel disappearing solution** video.) 
	4. In the pivot table field list panel, click on the Count of O3 and select Value Field Settings to change the values to Average, and format the values so that they show no digits after the decimal. (See the **10-pivot table O3 avg and formatting** video.) 
	5. Suppress the annoying error to show just blanks where there is no data for that day (PivotTable Tools, Options, Options, click the down arrow) then check the box that says For Error Values Show…and leave the box blank. (See the **11-pivot table suppressing errors** video.) 



* 1. At the top of the pivot table rename Row Labels to Date and Average of O3 to O3 with a space after it. (See the **12-change headers in pivot table** video.)
	2. Now for excitement! With the cursor in any date, right click, select Group, and select day, month, quarter, and year. (See the **13-pivot table date grouping** video.)
	3. To clean it up a bit, click on the Design Tab, Report Layout, Tabular. (See the **14-tabular format pivot** **table** video.) 
1. Now you have a file that looks something like 5-VideoDemoGroupedPivotTable.xlsx.
	1. Click the minus sign to the left of each quarter to collapse each quarter, to see a table of the avg O3 concentration in each quarter. (See the **15-expanding or collapsing pivot table groupings** video.)
2. With the cursor in the pivot table, click Options, PivotChart.
	1. Select Column chart (usually select XY scatter chart, but those are not available in PivotCharts) See the **16-pivot chart creation** video.
	2. To make it pretty before we do the magic with it, change the title, legend, column format and chart background to make it look how you want it. Save a template so you do not have to do this again (Design, Save template). See the **17-customizing pivot chart** video.
3. Then working with a file like 6-VideoDemoPivotChartPrettyO3.xlsx for 6-VideoDemoPivotChartLineO3.xlsx, make magic:
	1. Left-click the minus sign to the left of any quarter in the pivot table, to expand both the table and the graph; see the **18-pivot chart equal grouping by month** video.



* 1. Then add PM2.5 to the table by dragging it down from the field chooser in the top panel to the Values panel. Change the Total to Average and format to show no decimals. (See the **19-adding 2nd pollutant to pivot chart** video.)
	2. At this point your excel file should look something like 7-VideoDemoPivotChart\_O3\_PM25\_Column.xlsx. You can see two things: first, that column charts do not always work well for showing two parameters, so evaluate whether a different chart type would present your story better by clicking Change Chart Type:  (See the **20-chart types incl pivot** video.)
	3. Second, because pm is in different units than the ozone, it needs to be charted on a secondary axis (see the **21-secondary axis pivot chart** video): 

Then 

* 1. Change the chart as you prefer; I think it makes most sense without markers, as we are going to be showing averages by day, month, etc. and markers make the chart messy looking. Do these by formatting the series (either select using the menu, as above, or right click the series and select Format). You may consider keeping certain colors associated with certain pollutants, for example I use brown for PM and orange for ozone.
1. At this point the chart looks something like 8\_VideoDemoPivotChartLineO3\_PM25\_2ndaryAxis.xlsx if you prefer a line chart, or 8-VideoDemoPivotChart\_O3\_PM25\_Column.xlsx if you prefer columns.
2. Tools that the pivot table gives you:
	1. Double-click any value in the pivot table, and excel will automatically create a new sheet with all the data that makes up that value that you double clicked. This is very useful to generate subsets of data. (See the **22-pivot table extract** function.)
	2. Any pivot table can be easily copied as a table into a Word file (after you have changed the Report Layout to Tabular) just by selecting the pivot table cells, cntrl-C to copy, the click in a Word doc where you want the table, and cntrl-V to paste (or use the menu).
	3. The Table (in our sheet AllData) will expand to include any rows that you add, and then you can refresh the pivot table and pivot chart and they will all expand to include the new data. (See **23-automatic inclusion of new data when appended to table** video.)
	4. Any formulas you create reference the header, not the cell, so entering formulas is easier.
	5. Column headers are always visible.
	6. You can print just the table, rather than the whole sheet.
3. Tools that the Pivot Chart gives you:
	1. Can select subsets of data to display using the built-in buttons on the chart. (See the video **24-Two-ParameterPivotTableFeatures**.)