Making Sense of Data with Excel

Temple Analytics Challenge 2018

Mary Garcia Charumilind, Director of Performance Analytics

Three parts of analytics

- 1. Back end: data and databases
- 2. Analysis: statistics
- 3. Communication: visualization

Today's Topics

- 1. Back end: data structure, format, and preparation
 - Examining and navigating data
 - Tables
 - Merging data
 - VLOOKUP
 - INDEX and MATCH
 - Data types and formats
- 2. Analysis: statistics
 - Pivot tables
 - Slicers
- 3. PowerQuery: Data >1M rows

Excel

- Do you have Excel 2016?
 - Figure out which version you have: https://support.office.com/en-us/article/About-Office-What-version-of-Office-am-I-using-932788B8-A3CE-44BF-BB09-E334518B8B19?ui=en-US&rs=en-US&ad=US
 - Excel 2016 can be downloaded from TUPortal: Look for the "Microsoft Download" link in the list on the left-hand side
 - Really Big data? Choose the 64-bit version and use a PC
 - Microsoft Excel 2016 for Mac doesn't have "Get & Transform" feature. This is a new feature which only applies to Excel 2016. "New Query" and "Query Editor" are included in "Get & Transform", so they are not available either
 - The TU TECH Center has PCs with Excel 2016 https://computerservices.temple.edu/tech-center-software
 - Fox MIS lab in Alter 602 has powerful machines, but 32-bit Excel

Meeting your data

- Read the data dictionary
 - Unique Identifiers
 - How many rows, columns
- Getting around
 - Ctrl or Cmd + End
 - Ctrl or Cmd + Home
 - Click A1, Ctrl or Cmd + down arrow (same with up, down, right, Page Up, Page Down)
- Selecting data
 - All: Ctrl or Cmd + a to select all, or Ctrl + Shift + 8 to select the dataset I am in right now
 - A bit: Click and drag the puffy plus, or Shift + arrow key
 - A lot: Ctrl or Cmd + Shift + down arrow (same with up, down, right)

Preparing your data

Benefits of making a table:

- Defines range once and for all
- Automatically turns on filters/sort drop downs
- See column headings as you scroll
- Autofills formulas
- Can use slicers
- Names column ranges for you
- Format as Table: Home->Styles->FormatAsTable (Table tab on Mac), or type Ctrl+T (CMD+T on a Mac)
 - Go to the Design tab to name it

Merging datasets: Vlookup

- =VLOOKUP(ID cell in destination sheet, click top of columns of the data to search, number of column the answer is in, FALSE)
 - If using TRUE option to find near fits rather than exact fits, be sure to sort the ID ascending in your sheets!
 - Filling Formulas (usually not needed when working in a table): Using the skinny plus, double click the bottom right-hand corner of the cell to be copied all the way down through consistent cells (check that it reaches the bottom)
- Paste values of your formulas to prevent Excel working too hard
 - Copy and paste: Ctrl or Cmd + c, Ctrl or Cmd + v
 - Values only: Right click or Cmd + click to copy, paste, or, paste special (values only, formats only, formulas only)
- (Repeat as needed)

Common Problems with Vlookups

- The range reference was not anchored, did not use whole column, or was missing an important column
- The ID wasn't to the left of the target value
- The data wasn't in the column we thought
- The ID had a space at the end in one sheet but not the other
- The ID formats did not match (number stored as text in one of the sheets, etc.)
- The formula was left in (values weren't pasted), and then things moved

Preparing data

- Numbers stored as text
 - Left vs right alignment, green triangles
 - Use the VALUE function to return just the numeric value of the text.
 - Or, the Text to Columns button is typically used for splitting a column, but it can also be used to convert a single column of text to numbers. On the Data tab, click Text to Columns
 - Or Click through the warning sign to convert (uses more memory)

Merging data: Index and Match

- Benefits:
 - Doesn't matter what order the columns are in
 - Faster processing time
 - If seeking inexact matches, don't need to sort the lookup value
- =INDEX(select columns where the ID and target value are, MATCH(the ID where you are, the column that contains the ID in the target sheet),0), column with the data you want)

Preparing Data

- Date formats:
 - If your dates look like 41913, they need to be formatted as short dates.

Gaining insights: Pivot Tables

- Click a cell in the table; Insert -> Pivot Table
 - Can add to data model if on PC
 - Click Ok
- "To build a report" field list; drag them into the boxes
 - Rows: Use multiple fields and drag to reorder to show hierarchy
 - Expand and collapse rows
 - Columns
 - "Σ Values": Click the drop down arrow inside the box and select Value Field Settings Average to change from the default
 - Click the Show Values As tab to Show as %, index, etc.

Gaining insights: Pivot Tables

- When you click out of the pivot table, the field list disappears.
- Sort by values: Right click (cmd+click) on a value in the table
- Conditional formatting
- Filter
- Copy and paste tables
- The Analyze tab
 - Refresh (data)
 - Resurrect field list

Learning about your data: Descriptive stats

- Pivot tables
 - Count, Mix, Average, StDev, Max

- Formulas: Count min, max, median, mean, stdev...
 - SKEW only works on up to 255 values

- Data analysis toolpak function uses a lot of memory
 - Range, min, max, mean, standard deviation, median, mode, skewness, kurtosis...

Exploring Data

- **Insert slicer** to filter
 - A table, when columns are hard to keep track of
 - One or multiple pivot tables in a sheet
 - Insert -> Filters -> Slicer

Working with Datasets >1M rows*

- Data -> Get data -> Query options
 - Default query load settings (under Global -> Data Load)
 - Specify custom, Load to data model only, Click ok
- Data -> Get data -> Launch Query Editor
 - New source -> file -> Excel [repeat for each file if multiple]
- Close & Load -> Close and load to -> only create connection, check "add to data model"
- Insert -> Pivot Table -> Select "Use this workbook's Data Model"

^{*}Microsoft Excel 2016 for Mac doesn't have "Get & Transform" feature (Query Editor). This is a new feature which only applies to Excel 2016.

Questions?

mgcharumilind@temple.edu

215-204-3208

Office hours: Mon & Tues 12pm-1pm through October, Alter 731