"Microsoft Excel (Essentials)" Handout

menu bar

tabs

add-ons (in tabs)

natural language help

title bar

ribbon

groups

formula bar (fx)

column "referents" / labels

work space

active cell

worksheet tabs

zoom function

status bar (during processing)

macro recording status

interrelated sheets (in sequence, in set)

normal layout view

page layout view (side-by-side pages)

page break preview (zoomed-out view)
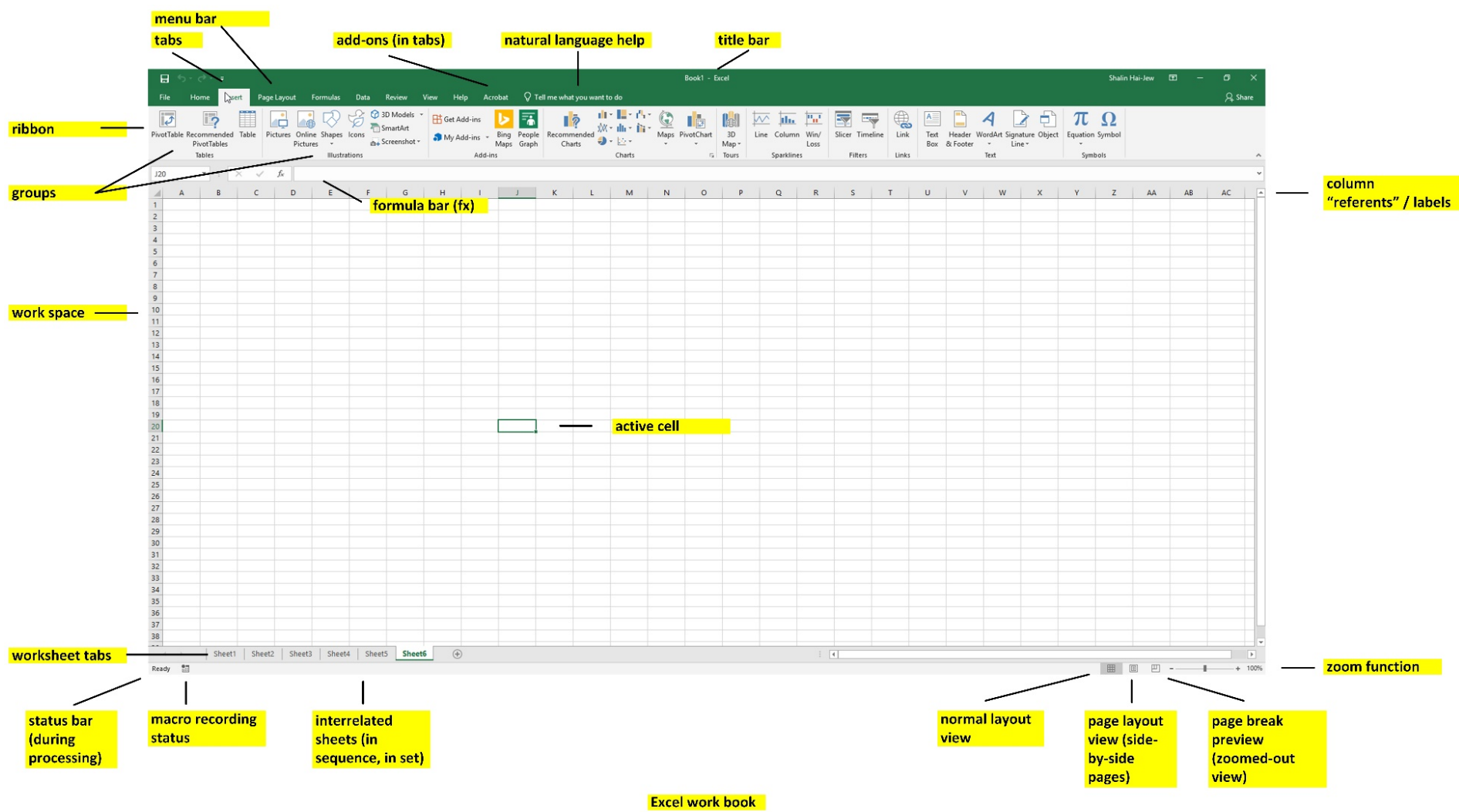
Excel work book

The Excel User Interface

Part 1. Session Contents

1. Self-Intros
2. What are your interests regarding Excel?
3. What have your work experiences been with Excel (if any)?
4. What questions would you like to have answered in this session (to be satisfied)?

Zoom Screen Sharing for a F2F Session: Zoom setup for the F2F session, so those sitting at the back can see my screen:
https://ksu.zoom.us/j/6752623021   (I will not be sharing sound via Zoom.)
Fully Online Zoom: For a fully online Zoom presentation, sound will be used.

My One Goal Today: I want you all to give Excel a try. (Please don't try to memorize how to use Excel. I want you to just remember capabilities and maybe a referent name to that capability…and those can all be found online. While there are some references to some more sophisticated capabilities, my plan is not to do walk-throughs of all of them…because the pacing of actually working these out is slower than the pace of a presentation.)

If you already use Excel for work or pleasure, please just try some new function or feature of the tool.

Here is the general plan.

HOUR ONE: The Walk-throughs and Demoing

- Creating a workbook (with spreadsheets)
- Ingesting data (.csv / comma separated values; .tsv / tab separated values; database data tables, and others?)
- Creating spreadsheets
- Cleaning data
- Filtering data
- Applying simple formulas to data
- Running a fill handle / using autofill (properly)
- Navigating in the data
- Searching and finding in the data
- Creating data visualizations (based on the data, based on data visualization conventions)
- Using built-in visuals [automated data visualizations (and starting to line up visuals with the data)]

- Troubleshooting together

HOUR TWO: Some More Advanced Walk-throughs

- Intensity matrices / intensity tables by frequencies (conditional formatting - > color scales)

- Custom cell formatting (positive numbers; negative numbers; handling of zero) ;
- Conditional formatting

- Excel templates (Business, Personal, Planners and Trackers, Lists, Budgets, Charts, and Calendars)

- Other topics (based on your own preferences)

Part 2. Some Basic Uses of Excel 2019

Excel is often used to achieve the following:

- conduct calculations,
- clean data,
- graph or plot data (to identify data patterns),
- create data visualizations (both static and interactive),
- apply visual styles to data tables and data visualizations,
- create interactive data visualizations (interactive pivot tables, interactive dashboards with slicers, and others),
- create simple macros (sets of directions, scripts) for continuing data handling (with Visual Basic programming language),
- process data for analytics in other software programs,
- enable the share-ability of information,
- access online survey data (in analyze-able format), **and more**

Starting a Data Workbook:
- An Excel workbook may be created from scratch manually.
- It may be created from pasted data.
- It may be created from available data in an existing file (which is loaded into an Excel workbook).
- It may be downloaded from a database.
- It may be downloaded from an online survey system (like Qualtrics).
- Or it may be created from some combination.

Avoiding Data "Lossiness" (Data Loss). Worksheets are limited to 1,048,576 rows by 16,384 columns. (When using the computer "clipboard" to copy-paste data, those limits are lower, so be careful when moving data to make sure that you don't permanently lose anything.)

Data Extractions from Application Programming Interfaces (APIs): With add-ons and other packages, it is possible to conduct some data extractions from social media platforms (NodeXL/Network Overview, Discovery and Exploration for Excel), create data streamgraphs, and achieve other aims. These are sometimes free. (Make sure to go to trusted sites only to download these to avoid downloading malware lurking in some programs.)

There is a free Web version with a Microsoft account. This is part of the Free Office Online apps (https://products.office.com/en/office-online/), which require a desktop browser and a Microsoft log-in.

Figure 1. The Quick Analysis Tool in Excel for an Initial Glance at the Data
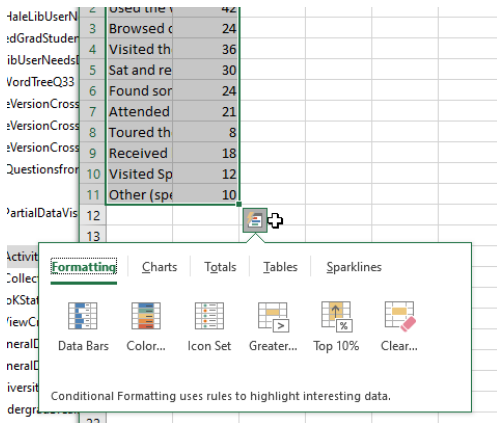
Engaging Quick Analysis:

There are a lot of built-ins in Excel that enable fast exploration of data. To access one, just highlight your data…and click the button at the bottom right.

Those who prefer to make their own formulas may do so using pre-written formulas… Or they can write their own formulas based on basic operands and directions (operations). A relative cell reference is assumed in most cases; for absolute cell references, use the F4 or the $ in front of each part of the formula ($columnreferent,$rowreferent) to lock down the referent. A locked-down referent uses the exact indicated cell again and again, and the referred-to cell does not move. (The default setting is to have relative cell references, unless otherwise indicated.)

Data Visualizations from Excel:
1. The data visualizations may be kept in the Excel file (.xlsx, not .csv file formats), and the Excel file may be shared.
2. The data visualizations may be shared online on Office 365's various Groups and SharePoint and others…for interactive experiential access and downloads.
3. The visuals may be shared in a document as a linked visual to the underlying dataset. Any changes to the underlying dataset will affect the data visualization. (This is done with a simple copy and paste from Excel directly. The underlying data files should not be moved from the respective locations, or that will result in an empty visualization. The files are linked locally, so folder structures and names and such matter.)
4. The visuals may be shared as a .pdf file (with text-readable text using screen readers).
5. The visuals may be shared as a data graphic in either vector or raster / bitmap formats. (In this case, the graph may be captured with a screenshot and rendered in a digital image editing tool.)

Import / Export File Types and Features

Import		Export

Work Sequences: Knowing what may be imported (into) and exported (out of) Excel is important because Excel may be part of a basic work sequence involving other technologies. For Excel to be able to "talk" with these other software tools, it is important to know what the imports / exports file types are.

File Versioning: Also, these file types enable versioning…or creating different versions of a basic file…for uses in different contexts. (The first "pristine" version of the dataset or workbook should be preserved in its original state, without changes, in order to protect the data against mishandling or unintended changing. The individual can always return to the original data and start fresh.)

Avoiding File Lossiness: Some file types are "lossy" and will lose information, so even if one starts out with a particular file type, it may make sense to save out the file in a different format. (One example is opening a .csv / "comma separated values" file of data but saving out as a .xlsx (XML spreadsheet) if you create data graphs and plots…because the .csv version will lose your visuals and result in repeated work.)

Part 3. A Bit about Structured (Labeled) Data

Basic Table Structure of Structured Data

	Column Headers (usually variable data) …					
Row Headers (usually individual records) …	Data cells					

Many datasets forgo actual labels to the columns if the contents are generally understood, or the documentation for the data is documented in a data dictionary or README documentation. Excel accepts a number of structured data as imports.

Common "Data Types" in Cells
- **Labels** (usually as alphanumeric or text-only characters)
- **Values** (usually raw numbers, dates, and others)
- **Formulas** (directions to Excel for how to perform calculations)

"**Structured**" data is **labeled data,** usually in data arrays or data tables. They are "structured" because they are formatted in ways that enable their processing in different software programs and of recognizable types. [By contrast, **semi-structured** or **un-structured data** are textual documents, imagery, videos, audio files, mixed-modal files, and others, because these have more subtle types of structuring…not the explicit data table structures of classic quantitative data.] Excel can handle both structured and unstructured data.

   Layout-wise, Excel has some challenges in handling images, though. It takes some work to center a series of thumbnails in a data cell… Those who want the thumbnail to link to a full-size image need to make the image a link to the actual image (hosted on an available site).

Data types take particular formats, and changing the data types in the column data can affect how the spreadsheet is ingested into databases (or not). Saving a workbook as a particular file type may also result in lost data if the format is not the right type to maintain the level of information. [For example, a file opened as a .csv but then edited to include data visualizations should be saved out as a .xlsx file to keep the visuals.] Data visualizations are usually summary data, to capture higher level data patterns. Some may be made interactive in Excel.

Data formats. The data formats may be set for each column of data, assuming that the data is structured data and all of a type in the column. The settings will affect how the cells are displayed and may change (numerical) input data…so it helps to know what the formatting is applied to each cell and how it affects the data. (This is especially true with inherited data from databases, data tables sent by colleagues, downloaded datasets, and so on. With inherited data or other types of data, always save a "pristine" master set before making any changes to the data…so you can go back to an anchor point before changes were made.)

General: default numbers, no understood format
Number: positive numbers, negative ones, up to 30 decimal places, may include the comma separator at each of three digits
Currency: any of hundreds of symbols for currency indicators, three-letter indicators for currency indicators, and others
Accounting: number representations for accounting, ( ) to indicate something of the amount, and others
Date: variety of representations of month, day, year; month; month-day, and other representations of

Figure 2. The Representation of Large Numbers

Time: hour, minute, second and AM/PM; and other variations… and can filter by locale (for sense-making)

4.3552E+04
4.3552E+04
4.3567E+04
4.3572E+04
4.3573E+04

Percentage: percent values with a certain amount of decimal places (%)
Fraction: numbers of digits for fractions, fractional representations
Scientific: represents large numbers in exponential notation

Text: "Text format cells are treated as text even when a number is in the cell. The cell is displayed exactly as entered."
Special: customized based on user needs (and enables text description of the "Special" type
Custom: custom data type with scripted handling of certain types of numeric data for more effective visual communications of numbers (directions, thresholds, frequencies, and others); applying rules or classes of rules to whole worksheet for every instance of the number: [];[];[];[]; can use symbols to represent numbers

[**Resource Note:** A fun resource on using "custom" settings is available here: https://youtu.be/tGY70sdpaLc. There is a video on Excel shortcut key combinations: https://youtu.be/Xe4U_-o_EWw.]

Maintaining a pristine master set of workbooks. The data types settings applied to the columns may change the data in the workbook, so save an original as a pristine master first before engaging with the data. User error in filtering can also break row data integrity.

3D Mapping with General Place / Location Data

- 3D Mapping Tour [Open Excel with the data. Open the 3D Maps. Select the data you want to include. In the 3D Map dropdown, "Add Selected Data to 3D Maps," and select desired data features to visualize.]
- The map is interactive and three-dimensional.

Figure 3. A Built-in 3D Maps Feature in Excel (requiring very light informational inputs)
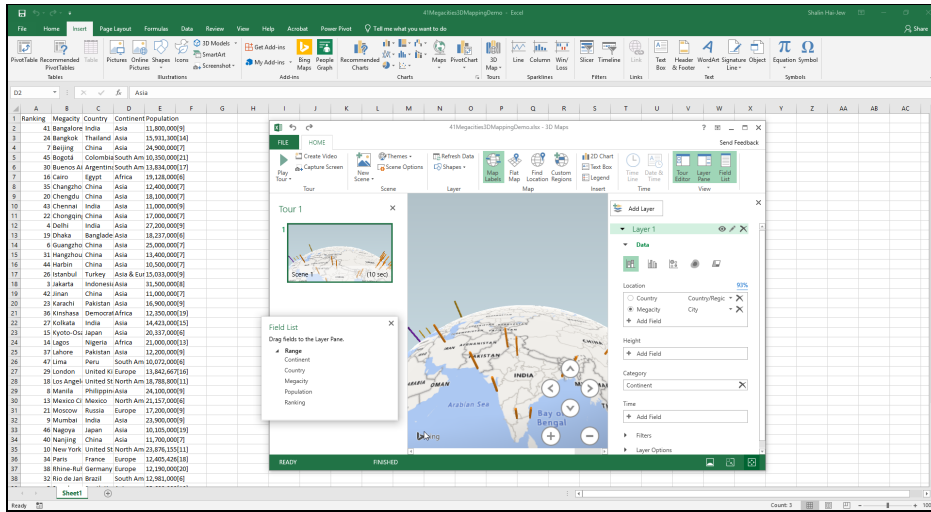
Pivot Tables (and Pivot Charts) for Summary Data Exploration

- Pivot tables for data summarization (includes recommended ones based on the data)
- Pivot charts for data summarization

Figure 4. A Pivot Table View
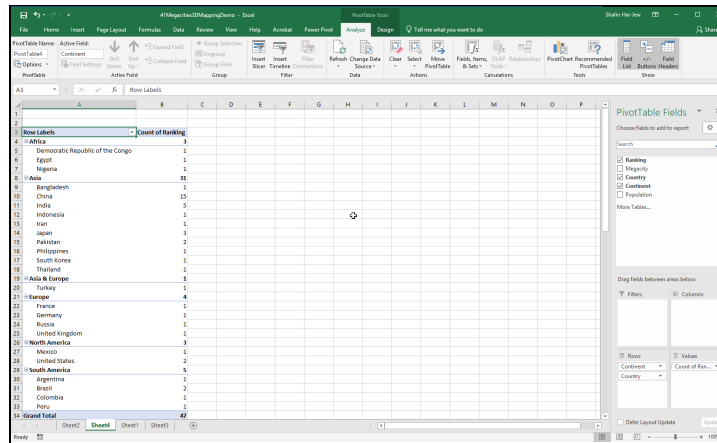
Figure 5. (Vertical and Horizontal) Lookups for Targeted Queries

Lookups for Individual Record Data Exploration / Queries:

The lookups are by record…based on particular variables of interest.

- If the column headers are used , then the VLOOKUP function is used.

- If the row headers are used (as the main labels), then the HLOOKUP function is used.

- **VLOOKUP** (Vertical Lookup) across rows (finding a value related to the leftmost column of data and a related corresponding value from another column, but in the same row, based on a defined relationship / formula)
- **HLOOKUP** (Horizontal Lookup) across columns (using a value related to the topmost row and finding a related corresponding value across columns…in rows)
   - **(V/H)Lookup Formulas:** (lookup_value, table_array, col_index_num, [range_lookup])

- **lookup_value** = first value of interest (usually identifier) in the leftmost column or the topmost row
- **table_array** = entire data table (extraneous data does not matter)
- **col_index_num** = column of interest by basic count (1, 2, 3…)
- **[range_lookup] =** true (if fuzziness or approximation or proxemic okay), false (if exact result desired) … The square brackets mean that an input is not needed. The default is "true" for approximated values.

- **For both, "named ranges"** may be created to enable referencing multi-columns simultaneously using a named variable.
- The arguments have to be set correctly the first time and can be re-used from there-on out…and shared with others. The point is to test with simple data first…to ensure that everything is working as desired.

Figure 6. An Index-Match Application on Open-Source USDA Data about Entrees and Accompanying Condiments

	Index and Match
Select Dis	Roast beef (eat lean & fat)
Condimer	Horseradish sauce
Condimer	Ketchup (catsup)
Condimer	None
Condiment	4
Condiment	5

● [**Resource Note:** One helpful video is at https://youtu.be/E7gQ-PgYkMc. For more on the new XLOOKUP, please check https://youtu.be/xnLvEhXWSas.]

Index- Match: A more sophisticated approach involves using Index - Match. A helpful video is at https://youtu.be/F264FpBDX28. This involves indexing or mapping data (a column or a whole table, location-to-location in the data array) and then using a matching function (from one column to particular data)…to create an interactive dropdown that enables light data exploration.

Figure 7. Creating a Dropdown List in Excel
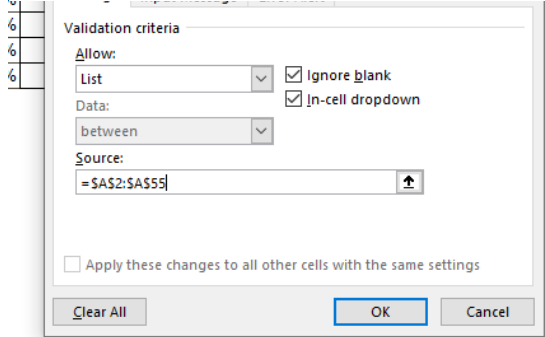
Making a Dropdown List:
1. Click on a cell to activate.
2. Go to Data tab. Go to Data Tools area and Data Validation.
3. Select Data Validation in the Dropdown.
4. In the Dropdown, choose List.
5. Identify the column source for the list (of identifiers).
6. Select all. (Keyboard shortcut: CTRL + SHIFT + downwards arrow key (↓ ▼)
7. Click OK.

Creating Interactive Data Dashboards:
- Set up the data in Excel.
- Set up pivot tables in a new sheet (with the selected features). Name the new sheet.
- Analyze -> Pivot Chart… Clean up the pivot chart data visualization. Swap in the desired text for proper labeling. Style it in the way that you want visually.
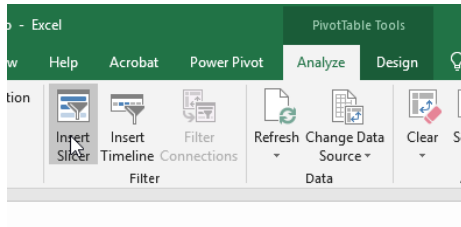- Make a new sheet for the dashboard. Emplace the pivot table through copy-and-paste.
- Continue until there are sufficient pivot charts. Place them in a logical order from top-to-bottom (and left-to-right).

To Make a Slicer in the Dashboard, …

Figure 8. Inserting a Slicer after a Pivot Table or Chart has Been Made for Easier Interactivity with the Data by Users

- Highlight a **pivot table or pivot chart** (interactive data visualization). Go to the activated Analyze tab above. Insert Slicer (an interactive filtering tool based on a selected dimension). Select the data column variable used for the filtering. The Slicer appears in the spreadsheet. Copy and paste it into the dashboard. The Slicer is a visual interactive filtering feature on the dashboard.
- Then, connect all the charts with the slicers. Right click the slicer and select the Report Connections Region, and connect the slicer to all pivot tables. Do so for each slicer. When mouse over-ing the slicers, the data visuals will change. [**Resource Note:** One video resource about creating a dashboard with slicers is available at https://youtu.be/RM8T1eYBjQY.]

New: Dynamic Arrays in Excel… A fairly new feature in Excel involves the built-in formulas that return arrays of data (collections of entities). The main built-in array functions are FILTER, UNIQUE, SORT, SORT BY, SEQUENCE, RANDARRAY. A detailed video about this functionality is available at https://youtu.be/BsUTlhib4Tw. A simpler and shorter video is here: https://youtu.be/2USJsIyIzvo. (This works on Excel 2019…with array formula features released in late 2018 and the update to Microsoft Excel's calculation engine.) [Ctrl+ Shift + Enter was used in the past for array formulas.]

About **keyboard shortcuts**, many of the common ones in the Microsoft Suite (esp. in Word) work the same in Excel. More are available here: https://support.office.com/en-us/article/keyboard-shortcuts-in-excel-1798d9d5-842a-42b8-9c99-9b7213f0040f.

In Excel, you can work with the graphical user interface / visuals, built-in pre-set macros, and / or line formulas (for those who like command line-interactions). Functions are represented multiple ways simultaneously for meaning-making.

Original Release Date of Excel 1.0: Sept. 30, 1985 (for our trivia fix)

Presenter: Dr. Shalin Hai-Jew
     ITS, Kansas State University
     785-532-5262
     shalin@k-state.edu

Copies of this handout (evolving) may be accessed at the following URLs:
   PDF version: https://idme-test.ome.ksu.edu/MSExcel2019.pdf
   MS Word version (as a download): http://www.k-state.edu/ID/MSExcel2019.docx

[Updated: June 2021]

Part 4. Typical Work Sequence (and sub-sequences) in Excel (extra)

1. Open Excel.
2. Input data manually (using copy and paste or other means). (Or capture by downloading data from a website or exporting data from a database.)
3. Check the data for correctness.
4. Save a pristine master of the data before any further processing. Use informative naming protocols for the Excel workbook.
5. Merge / combine / aggregate (union) selected data as needed. (This assumes that there is an identifier column which enables the ability to match data between tables.)
6. Clean and process the data (as needed).
7. Select particular data ranges for study. (Usually, data arrays are depicted visually in sections. In some cases, full datasets may be studied simultaneously such as for factor analyses.)
8. Explore and analyze the (selected) data through various queries, statistical analyses, and data visualizations. Sanity-check the findings and results.
9. Create a number of finalized data visualizations, and exported (or screenshot).
10. Export the data in various formats for analysis in outside software tools.
11. Archive.
12. Create a sharable Excel workbook of data that is sufficiently de-identified / anonymized and with a README section. (A README file usually addresses data sources, data representations, data table setup, contact information, and other relevant information.)

A related workflow follows on the next page.

```
1 Open Excel

2 Input Data          2 Open            2 Download        2 Export the
Manually             Existing           the Data          Data from a
                     Dataset or         from a            Database
                     Spreadsheet        Website

          Validate /        Save a Pristine
          Verify the        Master Set (and Set
          Data              Aside)

                    3 Merge Data
                    As Needed

                    4 Clean Data
                    (reduce noise, address empty
                    cells, disambiguate)

                    5 Select Data for
                    Study

6 Explore the Data   6 Run          6 Create Data          6 "Sanity Check" the
                     Statistical    Visualizations          Findings and Results
                     Analyses       (for discovery, exploration
                                    and analysis)

                    7 Create Finalized Data
                    Visualizations (for
                    Presentation and
                    Publication)

                    8 Export Data in Proper
                    Formats for Analysis in
                    Other Tools

                    9 Archive Workbook and
                    All Related Data, Visuals,
                    and Documentation

                    Share or                                    no
                    Not

                    yes

                    10 Create a Sharable
                    Excel Workbook that is
                    De-identified /
                    Anonymized and with a
                    README Section                        Close Excel
```