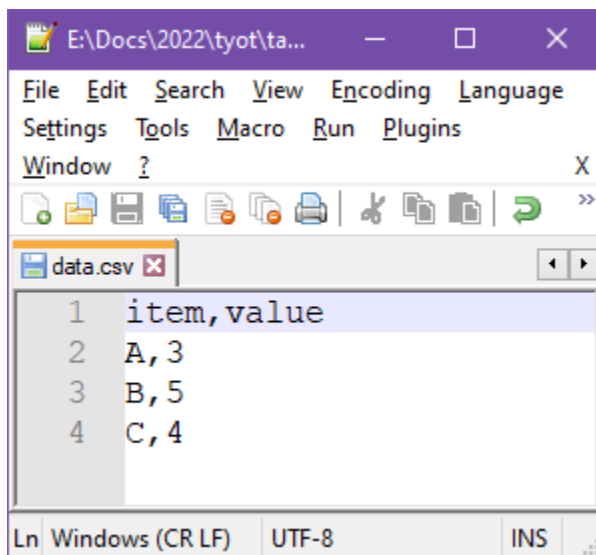## Data Analysis and Visualization

Topic 02 (week 3): Tableau and viz basics
Today's agenda

- Brief introductions (skipped during the first day)
- What makes a good visualization?
  - Note: We come later back to the underlying question "What makes good data (and statistics)"
    - *...but leave the even more fundamental question "What is a good business question (...and the vizs to support it)" largely unanswered in detail*
- Learning Tableau
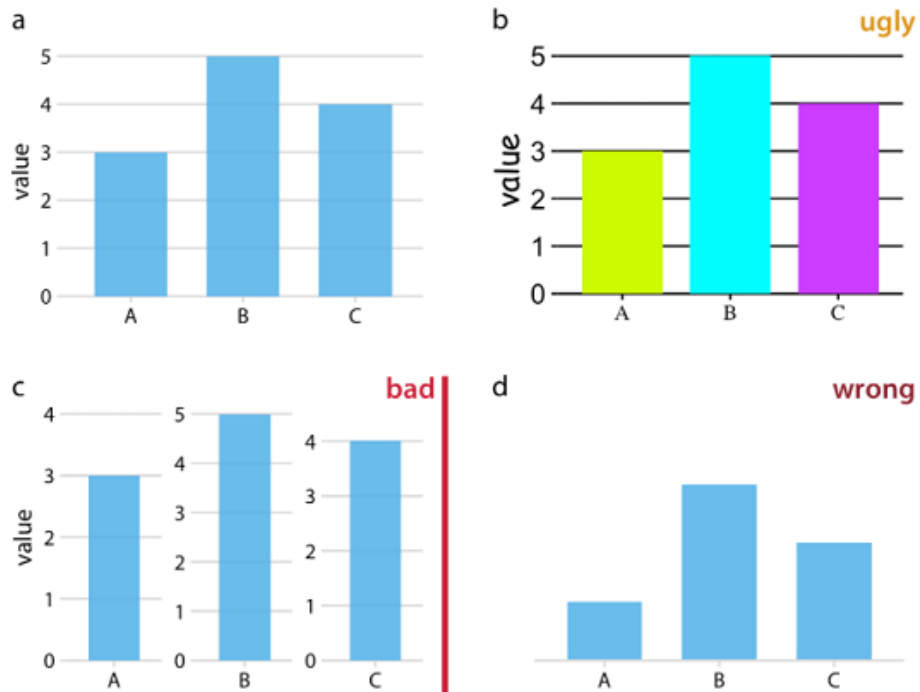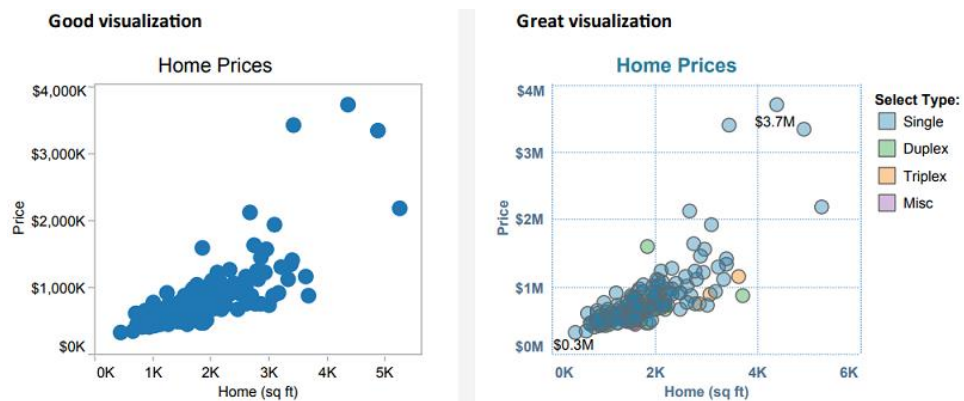- Assignment 02

(Reasonably) Good, Ugly, Bad, vs. Wrong Figures

Figure 1-1. Examples of ugly, bad, and wrong figures. (a) A bar plot showing three values (A = 3, B = 5, and C = 4). This is a reasonable visualization with no major flaws. (b) An ugly version of part (a). While the plot is technically correct, it is not aesthetically pleasing. The colors are too bright and not useful. The background grid is too prominent. The text is displayed using three different fonts in three different sizes. (c) A bad version of part (a). Each bar is shown with its own y axis scale. Because the scales don't align, this makes the figure misleading. One can easily get the impression that the three values are closer together than they actually are. (d) A wrong version of part (a). Without an explicit y axis scale, the numbers represented by the bars cannot be ascertained. The bars appear to be of lengths 1, 3, and 2, even though the values displayed are meant to be 3, 5, and 4.

Source: https://andor.tuni.fi/permalink/358FIN_TAMPO/1j3mh4m/alma9910690477705973

Going beyond "reasonably good" visualizations?

Source: https://www.tableau.com/learn/whitepapers/tableau-visual-guidebook

From great to effective visualizations

Emphasize the most important data

- Utilize readers' natural perception attention to highlight the most important aspect(s)
- A rule of thumb is to put the most important data on the X- or Y- axis and less important data on color, size, or shape

Orient your views for legibility

- In the Western countries, people prefer reading from left to right (rotate viz, not reader's head)
- If you find yourself with a view that has long labels that only fit vertically, try rotating the view

Organize your views

- Support making comparisons and hypothesis-making --- but don't force the reader to struggle on "obvious" conclusions (write them down)

- E.g., a bullet chart combines a bar chart with reference lines to create a great visual comparison between actual and target numbers

Avoid overloading your views

- Focus: The reader does not have to see "everything" and Break down: Use subcharts when appropriate
- E.g., instead of stacking many measures and dimensions into one condensed view, break them down to small multiples

Limit the number of colors and shapes in a single view

- Human cognition can absorb only so much; having too much on a screen is perceived as a decoration (unless one knowingly aims for very data intensive visualizations, such as heatmaps, where most pixels might convey some specific meaning)
- Limit the number of colors and shapes in one view to 7-10 so that you can distinguish them and see important patterns

Check examples from the source: https://www.tableau.com/learn/whitepapers/tableau-visual-guidebook (pp. 15-20).

Tableau Demonstration

Connecting to Excel data

Managing and joining multiple data tables

Tableau GUI basics:

A. Workbook name. A workbook contains sheets. A sheet can be a worksheet, a dashboard, or a story. For more information, see Workbooks and Sheets.

B. Cards and shelves – Drag fields to the cards and shelves in the workspace to add data to your view.

C. Toolbar – Use the toolbar to access commands and analysis and navigation tools.

D. View – This is the canvas in the workspace where you create a visualization (also referred to as a "viz").

E. Click this icon to go to the Start page, where you can connect to data. For more information, see Start Page.

F. Side Bar – In a worksheet, the side bar area contains the Data pane and the Analytics pane.

G. Click this tab to go to the Data Source page and view your data. For more information, see Data Source Page.

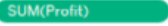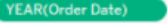**H.** Status bar – Displays information about the current view.

I. Sheet tabs – Tabs represent each sheet in your workbook. This can include worksheets, dashboards, and stories. For more information, see Workbooks and Sheets.

Source: https://help.tableau.com/current/pro/desktop/en-us/environment_workspace.htm

Visualization worksheet basics, field (or attribute) types: Dimensions and Measures, Blue and Green; basic view area concepts (columns and rows)

Blue versus green fields

Tableau represents data differently in the view depending on whether the field is discrete (blue), or continuous (green). *Continuous* and *discrete* are mathematical terms. Continuous means "forming an unbroken whole, without interruption"; discrete means "individually separate and distinct."

- Green measures `SUM(Profit)` and dimensions `YEAR(Order Date)` are continuous. Continuous field values are treated as an infinite range. Generally, continuous fields add axes to the view.

- Blue measures `SUM(Profit)` and dimensions `Product Name` are discrete. Discrete values are treated as finite. Generally, discrete fields add headers to the view.
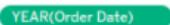
Possible combinations of fields in Tableau

This table shows examples of what the different fields look like in the view. People sometimes call these fields "pills", but we refer to them as "fields" in Tableau help documentation.

Discrete Dimensions	`Product Name`
Continuous Dimensions (dimensions with a data type of String or Boolean cannot be continuous)	`YEAR(Order Date)`
Discrete Measures	`SUM(Profit)`
Continuous Measures	`SUM(Profit)`

Source: https://help.tableau.com/current/pro/desktop/en-us/datafields_typesandroles.htm

Filtering view data vs. highlights and filters

Assignment 02