Data Analysis and Visualization

Assignment 02

E02.1. Download and copy the file *Sample - Superstore.xls* into your course assignment folder. You may find an archived version of the file in the Moodle Examples and Data resource folder, copied from https://community.tableau.com/s/question/0D54T00000CWeX8SAL/sample-superstore-sales-excelxls .

Note: The course book mentions that the file Sample - Superstore.xls would be available in your Tableau installation data folder. At the time of writing, however, the file seems no longer be included in the Tableau (Public) installation.

Open the file in Excel and check what in includes. (Tip: Using your TAMK OneDrive folders gives you direct access to the associated Office 365 tools, including Excel --- without local pc installations.)

In your own words, briefly describe the file: What kind of information does it include? How many tables or sheets it includes, with how many rows and columns? In principle, what kind of information is available per order (id)?

Note that the superstore example dataset is very widely used for learning Tableau (and various versions can be downloaded from multiple sources) --- but the business objective(s) and the relations of the dataset are sadly enough quite vaguely described. (So when compiling your own business critical data, pay attention to documenting your dataset!)

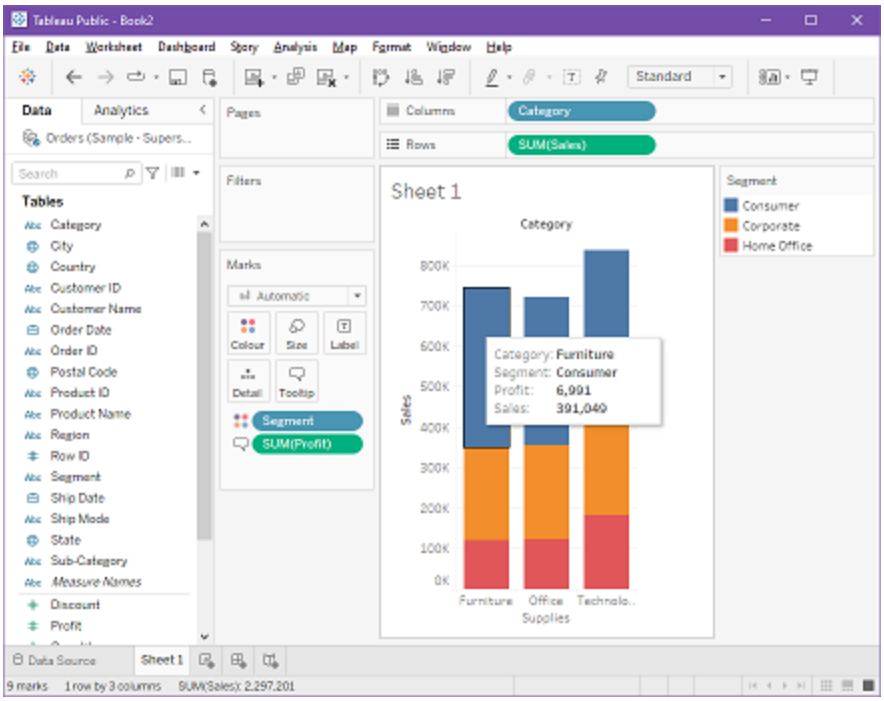
Open the (Excel) file in Tableau Public (drag Orders sheet to the middle pane) to and check the Data source pane. Did the data load correctly? Check and major fix issues, if any.

Tip: Check also our Tableau course book pp. 10-13 (starting from pdf page 42):

- Loth, Alexander; Vogel, Nate; Sparkes, Sophie, (2019). *Visual Analytics with Tableau.* Wiley. Available via Andor for TAMK students free from charge (link).

Note: When working with tableau, note that the Tableau (Public) interface (concepts) might have slightly changed since the publication of the book or other material. This is common with modern applications. In other words, the GUI images might not match 100% --- use common sense when working with archived instructions.

E02.2. With the superstore data, create a Tableau Public visualization like the following:



Tip: To review the essential steps, check the Tableau course book pp. 17-21 (starting from pdf page 49).

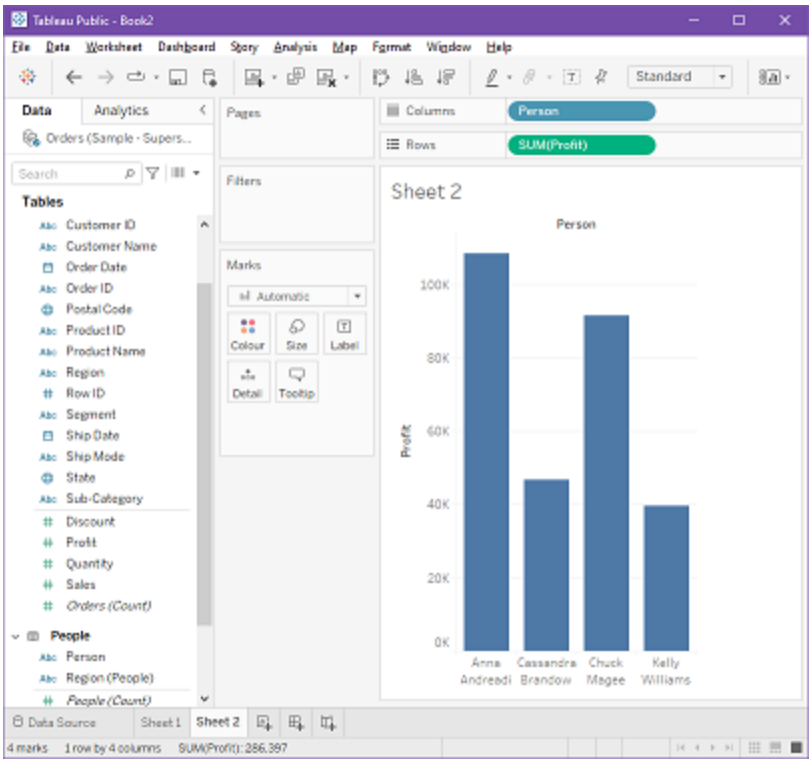
Save your viz into your public Tableau Public account, and copy the link to the viz to your solution (note that you may include the exercises below into the same viz and hence use the same link several times).

Finally, in your own words, briefly describe what Tableau means by dimensions, measures, and marks, and how they are in principle used in visualizations.

E02.3. Let's then examine relations in data and use them to create visualizations that aggregate several (sub) data sources (in this case, simply within a single Excel file).

Switch back to Data source pane and add the People sheet to your Tableau data model by dragging it onto the right.

With the new improved data model, create a viz like the following:



Intuitively, the viz tells how much profit is credited to which person.

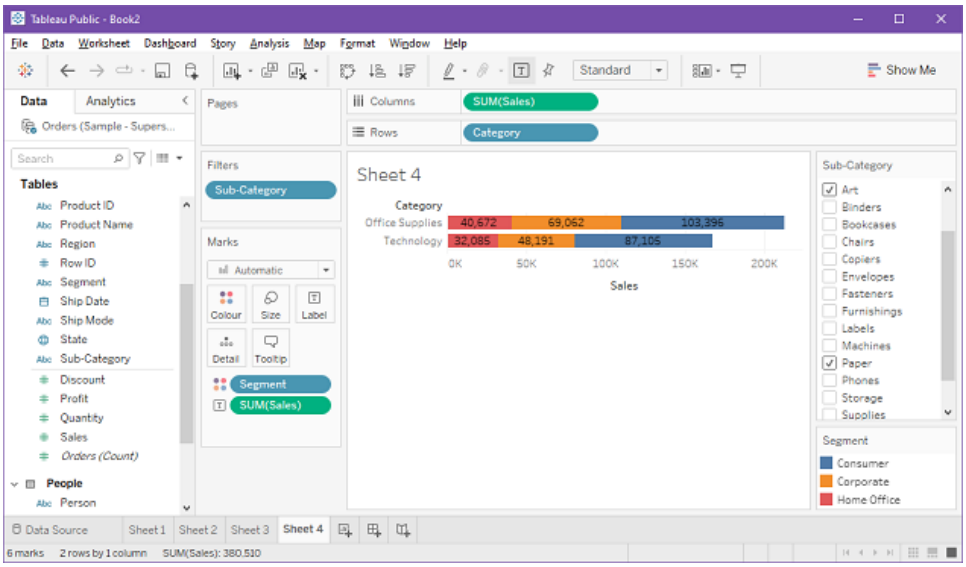
Save your viz into you public Tableau Public account, and copy the link to the viz to your solution.

Tip: To review the essential concepts of joins and unions, check the Tableau course book pp. 31-34 (starting from pdf page 63).

In your own words, briefly describe what was the use of relation or a join in this case.

In your own words, briefly describe what is the difference between a join and an union. (Check the Tableau course book pp. 34.)

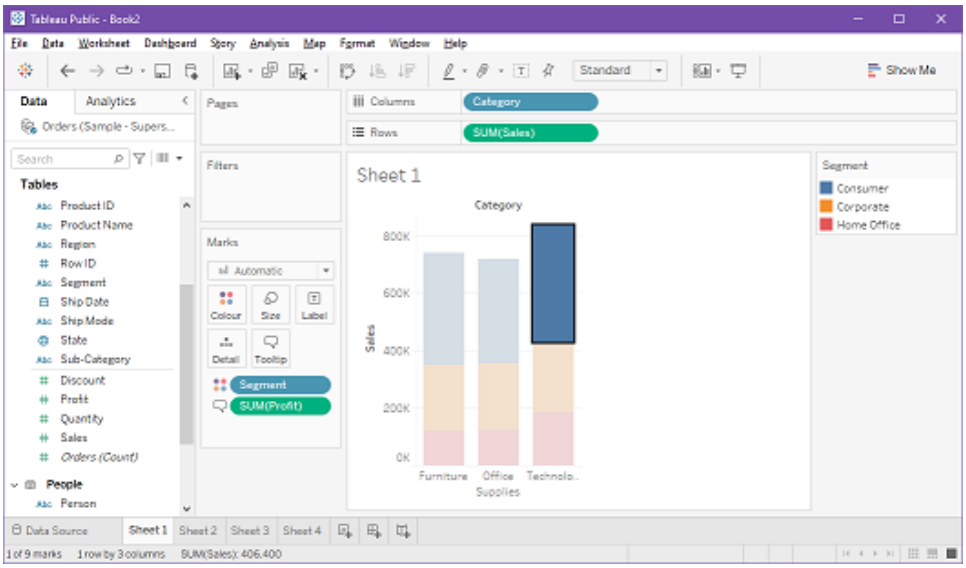
E02.4. Examine the use of filters. Create the following viz:



Tip: Check the Tableau course book pp. 55-57 (starting from pdf page 87).

Try out how the filter behaves. Briefly explain how it could be used to examine the data and how the viz changes according to filter selections.

Go back to the viz created in exercise E02.2. Try clicking some of the viz elements, thus highlighting it:



In your own words, briefly describe what is the difference between a highlight and a filter.

Right-click and element (also highlighting it). Then, from the context menu, choose View data...

In your own words, briefly describe what this means and note which rows of the data source contribute to that viz element.

Tip. Check the Tableau course book pp. 79 (starting from pdf page 111).

E02.5. Let's then take a step back and reconsider the bigger picture of visualizations: In reality, they are always designed for a reason:

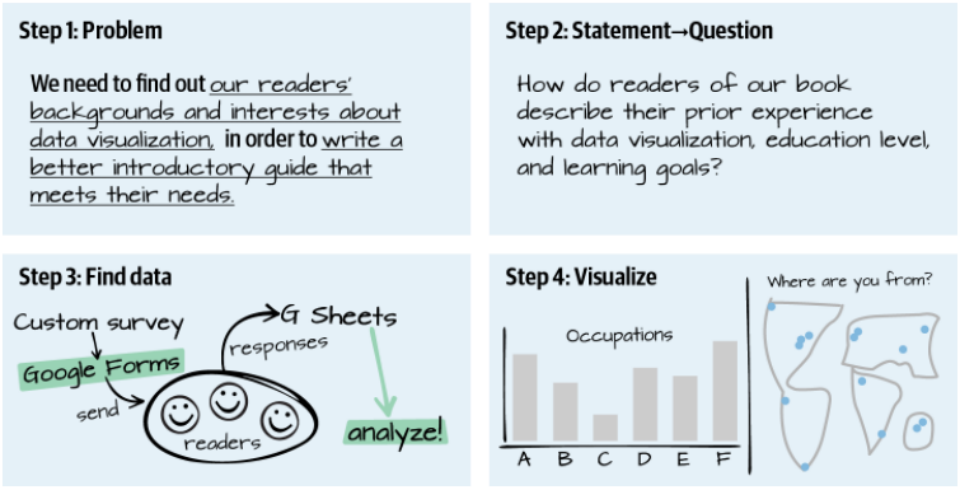


Image source: https://andor.tuni.fi/permalink/358FIN_TAMPO/1j3mh4m/alma9911194976405973

Let's imagine that you are running a superstore and would be interested in learning which products make the most of your sales or something similar. (If you want a challenge, come up with some nearby problem of your own!)

Part 1/3 of the solution: Using the four-step example above, come up with a similar, VERY SIMPLE four-step plan for yourself, how to get insight into the topic. For this exercise, it is sufficient to write a few lines of text, including textual description of the viz sketches. (If you wish to do this more "for real", design a simple Powerpoint presentation that includes the four steps, and add it as a snapshot picture to your solution.)

Part 2/3 of the solution: Using the example superstore dataset, try implementing the related viz using Tableau Public. (I.e. with the textbook data, not with proper data actually collected by yourself.) If you can't implement the experimental visualization "fully", briefly explain what kind of data would be needed.

Note that in the real world, coming up with a "final" visualization might take several quite laborious iterations --- which we skip for now.

Part 3/3 of the solution: Briefly consider your findings (with the textbook data). What kind of actions might be taken based on the (final) results? How to follow-up if the actions lead into a desired direction?

Note that the actionability of visualizations is very important. Unless simply for exploring some application (data), visualizations should be actionable --- otherwise they might simply be efforts wasted (from a business perspective).

Note also that in practice, e.g., optimizing sales is a very tricky question with many tacit dependencies and things to consider. For instance, from the perspective of a local grocery store, it might superficially seem that selling milk, bread, and toilet paper might not (directly) be that profitable while selling cheese might be --- but if the shop stops selling the very common stuff, costumers might not bother visiting the (general) grocery store in the first place. (Such things might be analyzed in terms of, e.g., market basket analysis, which seeks understanding customer purchasing patterns.)

Tip: By "VERY SIMPLE" we mean that completing this exercise should not take several hours.

Once you have completed the exercises of this assignment:

- **Submit** (brief notes, code(s) or link(s) about) your solutions to the correct Moodle discussion forum (start a new discussion; check the instructions for more details and how to update already submitted solutions, etc.); and
- **Write** peer feedback to two other students' solutions in the correct Moodle discussion forum (check their solutions and add a brief comment as a reply, written in a constructive tone).