Course Project (50%) – Computer Science 6934

Project Proposal Due: 11:59pm, Oct. 25, 2021

Final Project Due: 11:59pm, Nov. 29, 2021

Introduction

You will act as a freelance data visualization specialist for your course project. You will have a business or some other organization (your "client") ask for your help to answer their key business needs or questions using a data set that they have provided. You will need to create visualizations from this data set to satisfy their needs.

Client

Your client is large software company. They have a few different teams that provide customer support to their customers. Each member of customer support spends their day responding to customer tickets that have been created through email, a web form on the company's website, or live chat on their website.

The company has provided a data set of tickets that have been worked by members of their customer support teams (`tickets.csv`). This data contains the customer support specialist that worked the ticket, the team to which they belong, when the ticket was first created, as well as when it was opened and closed by customer support. Customer feedback on each ticket is provided in terms of a 5-star rating and whether the customer felt the issue was resolved.

The broad goal your client wishes to answer is, "How efficient is our customer support?"

Instructions

Using the provided data set, you will need to:

- identify two sub-goals related to your client's main goal, and
- create at least 4 unique visualizations that address the main goal and your sub-goals.

The two sub-goals that you identify should be more specific than the main goal. One possibility is a more strict or specific definition of "efficient." For example, you might view efficiency through the lens of customer support as a whole, or through individual teams or customer support specialists.

Of your 4 visualizations, there must be:

- 1 explanatory visualization,
- 1 exploratory visualization,
- 1 interactive visualization, and

- 1 visualization that contains derived data types.

Each visualization may contain more than one of the above elements, but all 4 of the above elements must be represented at least once in your set of visualizations. For example, you could create 2 explanatory visualizations using derived data, 1 explanatory visualization with interactive elements, and 1 exploratory visualization.

Furthermore, of your 4 visualizations, there needs to be at least 3 different standard types of visualizations (e.g., bar plot, scatter plot, etc). That is, you will need to create a diversity of visualization types. For example, do not include 3 histograms in your set of visualizations, although it would be acceptable to have 2 histograms, a scatter plot, and a box plot.

You must work on the course project by yourself.

Project Proposal (10%)

Due: 11:59pm, Oct. 25, 2021

Submit a proposal to your client outlining the two sub-goals you have identified and the visualizations you plan to make. Describe the purpose of each visualization and how it addresses the main goal or your sub-goals.

Each planned visualization should include a rough sketch or protoype of what you envision the final visualization to look like. What is important is that the sketch conveys a sense of what you intend to make. The sketch can be made with whatever tools you like. A picture of a hand-drawn sketch would be acceptable, for example.

Your proposal should be submitted as a `.pdf` with standard 8.5×11 size (US Letter). A maximum of 2 pages is allowed.

Final Project 40%

Due: 11:59pm, Nov. 29, 2021

Your final project submission consists of 3 parts:

1. Your 4 visualizations. (**12%**)
2. The code used to create those visualizations from the original dataset. (**13%**)
3. A pre-recorded video presentation to your client walking through your visualizations. (**15%**)

Include a text file describing the files in your project submission.

Your visualizations will be graded on their:

- Representation of explanatory and exploratory visualizations. (**2%**)
- Interactive components of your interactive visualization(s). (**2%**)
- Diversity of visualization types. At least three visualization types are expected. (**2%**)

- Appropriateness of visualization type for the chosen data and alignment with stated goal or sub-goal. (**2%**)
- Originality of design. Identify one of your visualizations to be marked for originality. Explain why it is original and what is original about it. (**2%**).
- Elements of style, such as choices of colour, annotation, line or point style, and other stylistic components. (**2%**)

Your code will be graded as follows:

- Code to derive new data types and any pre-processing of data. At least one of your visualizations must include a derived data type. (**5%**)
- Code implementation of visualizations. (**8%**)

For your video presentation, you will walk your client through the visualizations that you have made. Introduce the problem in terms of the stated broad goal and any sub-goals you have identified. Describe each visualization to your client. Provide context for each visualization in terms of how it relates to the goal and sub-goals and why you have chosen this design.

Your video presentation should be at least 5 minutes, and absolutely no longer than 10 minutes. Your video presentation will be graded on your:

- Explanation of each visualization. How should your client understand your visualization? How can your client use any interactive elements? What is the purpose of the visualization? How does it help the client achieve what they need? What are the key takeaways that your client should read from each visualization? (**7%**)
- Alignment with original proposal. If one of your visualizations differs significantly from your original proposal, then provide justification as to why. It is not uncommon that final products differ from original plans, so this is acceptable, provided you explain why the original design was deficient and why your new design is better. (**3%**)
- Overall cohesiveness of the story that you tell through your presentation / visualizations. (**3%**)
- Length of video. Videos over 10 minutes will receive no marks on this aspect. Videos under 5 minutes will receive half marks. (**2%**)

Submission

Submit your project through Brightpsace.

Your 4 visualizations can be submitted in whichever file format makes sense. They could be `.pdf` or `.png` files, or submitted inside a Jupyter notebook (`.ipynb`; this may make the most sense for interactive elements).

Your pre-recorded video presentation can be in an uploaded video file or a YouTube link.

Late submissions will be subject to a 20% penalty for each day past the deadline (proposal + final submission deadlines).

Alternative Data Sets

If you have a data set that you wish to use instead of the provided data set, you *may* be allowed to use it provided that you have discussed this with me beforehand and I have given my approval. I will need to verify that your suggested data set is suitable for the course project.

Do not submit your project proposal using an unauthorized data set. You need approval to use any alternative data sets before you submit your proposal.

Attribution

Submissions should include an attribution section indicating any sources of material, ideas or contribution of others to the submission.

Submissions must represent your independent work.

You are encouraged to use any resources to help with your solution, but your solution must represent independent work. If your submitted work includes unacknowledged collaboration, code materials, ideas or other elements that are not your original work, it may be considered plagiarism or some other form of cheating under MUN general regulations 6.12.4.2 (4.12.4.2 for graduate students) and academic penalties will be applied accordingly.

Avoid academic penalties by properly attributing any contribution to your submission by others, including internet sources and classmates. This will also help distinguish what elements of the submission are original. You may not receive full credit if your original elements are insufficient, but you can avoid penalties for plagiarism or copying if you acknowledge your sources.

Github

I encourage you to store and version your work on github. It is good practice to do so as everyone uses git in the real world.

However, **it is a requirement that git repositories containing assignment material be private.** University regulations (undergraduate 6.12.4.2 and graduate 4.12.4.2) consider it cheating if you allow your work to be copied. There will be zero tolerance for this.