

Project Paper

For the final project, you will collect your own data, (most preferably cybersecurity related data) conduct exploratory data analysis, develop hypotheses, and then create data visualizations to answer the hypotheses. The project tasks are listed as follows:

1. Collect a set of data, which you will create data visualizations for. Based on your needs, you may process the data such as cleaning or removing out some data items before you use it. Transform the data into a CSV or xlsx file. Describe the data file, including its content and attributes so that others can understand the data without referring to any other materials.

Here are a few links to data portals:

Google dataset search: <https://datasetsearch.research.google.com> (Links to an external site.)

Kaggle: <https://www.kaggle.com/datasets> (Links to an external site.)

Open government data portals such as: <https://www.data.gov>, (Links to an external site.) <https://data.texas.gov> (Links to an external site.)

2. Conduct an exploratory data analysis (EDA) of the data and develop at least three hypotheses that you would like to answer. Based on EDA results, explain why you think the data set is appropriate to answer the hypotheses.

3. Create a group of data visualizations (at least three graphs) to answer the hypotheses. The data visualizations must include at least 2 interaction features. Explain how data visualizations answer the hypotheses. Explain the development of data visualizations: you need explain what elementary perceptual tasks you chose, what visualization formats, such as scatter plots or bar charts, you created. Explain the rationales for choosing the 2-dimensional space design, the spatial layout, the colors, etc.

4. Write a report that includes the description and explanation of your project development and all data visualizations. You may create the data visualizations using Tableau, and then copy the graphs into the report.

Submit your paper, in docx or pdf format, and submit the data file in csv or xlsx format. Keep in mind that the submitted paper should contain **at least 1000 words excluding references**.