

Final Exam

CDA 500

Building your own Data-Driven Story

Overview: For your final examination, you will be required to use the skills you've learned to create a story about your selected dataset using visuals and supplementary details (e.g., a written summary) to tell your story. You will select one of the four provided datasets (detailed below), which are available on UBBox (link on UBLearns) to complete this assignment. You only need to select ONE dataset for this assignment.

You must choose one of the four audience personas we discussed in class to center your visuals around. You will document this choice in the write up portion of your deliverable (detailed below).

You are encouraged, but **not required**, to do additional research or pull in additional data to support your visuals.

Deliverables:

1. A minimum of **five** visuals of your choosing, built with Tableau. **For each visual**, you must provide a screenshot of the visual and a 6-8 sentence description of the visual. 1-2 sentences must tell the story of the visual as though you are communicating it to your target audience. The remainder of the write up must be devoted to describing how your visual utilizes the design methodologies and audience considerations we discussed in class (such as the Gestalt principles).
 - a. Note: While more than one visual can be of the same type (e.g., line chart), you must have at least **three** different chart types across the five required visuals. You are welcome to build more visuals if you find benefit in doing so.
2. At least one dashboard, which combines visuals together to communicate your insights to your audience. You must provide a screenshot of your dashboard, along with a 6-8 sentence description of the dashboard: 1-2 sentences to tell the story of why you chose those visuals and what they are communicating about the data and the remainder discussing how your dashboard utilizes the dashboard design concepts discussed in class.
3. An additional 1-page double spaced write up discussing your process in developing visuals, issues you encountered, insights you were able to derive, and your chosen audience persona, and other outcomes of your activities on the assignment.

Your final submission should contain the above, along with a copy of the tableau workbook file you built your visuals in.

Dataset details (pick 1):

1. AirBnB Boston Dataset
 - a. This dataset contains scraped information from Inside Airbnb (<http://insideairbnb.com/about/>).

- b. A data dictionary, describing the fields present in the data, can be found here:
<https://docs.google.com/spreadsheets/d/1iWCNJcSutYqpULSQHINyGInUvHg2BoUGoNRIGa6Szc4/edit?pli=1#gid=1322284596>
- 2. COVID-19 Case Surveillance Public Use Data
 - a. This dataset was collected and released by the CDC (<https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data/vbim-akqf>).
 - b. **NOTE:** The original dataset for this is massive, over 98 million entries. I have truncated this data to only reflect 500,000 sampled datapoints to avoid you having to download and use such a massive dataset.
 - c. The data dictionary can be found on the linked website.
- 3. Bank Marketing Dataset
 - a. This dataset is a modified version of the following Kaggle dataset (<https://www.kaggle.com/datasets/muhammedsal98/bank-marketing>) and reflects the efforts and results of bank marketing campaigns based on customer profiles.
 - b. **Note:** This dataset has been modified and you should refer to the UBBox details **not** the Kaggle details should you choose this.
- 4. FIFA 18 Soccer Data
 - a. This contains a number of pieces of information regarding player statistics from the 2018 FIFA video game: <https://www.kaggle.com/datasets/thec03u5/fifa-18-demo-player-dataset>
 - b. Details are available in Box to describe the different types of statistics contained within the file.