# CS 416 Data Visualization Final Project: Narrative Visualization

By Aareana Reza

**Messaging**

Data Science is a field that is currently booming and has a growing number of job opportunities. This career field also encompasses a variety of different job titles. All of these jobs cover a diverse range of qualities, for example various levels of experience, skill sets, and salary expectations. There are a lot of possibilities for people who wish to pursue a career in this field, but since there are so many options it can be intimidating to find a role that is best suited to you. The graph in this narrative visualization is a tool to help people find something that fits their preferred salary expectations, in addition to other preferences like experience level, employment type, remote ratio, and location. By looking at patterns in the graph we can see that jobs that have a lot of data points are a lot more popular than those with less. We can additionally see that some jobs have a wide range of salaries offered while others are either skewed in the higher or lower direction. It is also easy to spot that there are a few very high salaries that are outliers, these positions are sparse and tend to be reserved for people with more experience. So, the graph also allows users to easily sort through and compare particular data science positions based on things like popularity and availability.

**Narrative Structure**

My narrative visualization is designed to follow the martini glass structure. The martini glass structure is when the author delivers the message by initially guiding the user through the page and then allowing the user to explore on their own at the end. My web page starts off with a general summary of the message that the display is trying to portray. It also gives a brief description of the graph, so that the reader can have a general understanding of what kind of information they are about to look at. In this intro paragraph, the user is also told how to navigate the graph. Then, through buttons the user is guided through a few scenes that have different annotations that highlight data points and trends in the graph. All of these things are examples of author-driven content since the user is outright being told to look at specific observations about the graph. After the user has clicked through all the annotations, they are allowed to explore the graph and its data points on their own. This part is user-driven content since the user is now allowed to analyze and “drill-down” on the graph at their own pace.

**Visual Structure**

The visual structure that I choose for each scene is a column scatter plot. The x-axis is the job title which is an ordinal discrete variable. The y-axis is the

salary of a specific job in USD which is a quantitative continuous variable. Each job title is treated like its own scatter plot, with each column containing many data points that represent individual jobs and their correlated salaries. Having all the columns together then creates an array of multiple scatter plots that the user can evaluate by itself or compared with one another.

All of the scenes contain the same visual structure in order to have a sense of consistency which ensures that the user understands the data when they are moving from page to page. Highlighting is utilized in the web page through annotations, tooltips, and an info box. In the first few scenes, annotations are displayed on the graph. They point out trends and anomalies present in the graph which gets the viewer to focus on the important parts of the data in each scene. When the user is allowed to explore the graph, there is also a tooltip available on each of the data points. When the user hovers over a data point, they get some details on demand, like what type of experience level is required for that job. The user also has a help box available to them that gives them some pointers on how they can explore the graph and a breakdown of what some of the abbreviations in the tooltip mean.

There are also clear and easy ways for the user to transition between scenes using buttons. In the introduction, the user is told how to navigate through the narrative using the NEXT and BACK buttons located near the graph. Since the chart remains the same in each of the scenes, the user can focus on the changing annotations and direct their attention to the message that is being illustrated.

**Scenes and Visual Ordering**

There are 5 scenes in total in this narrative visualization. The first scene is the introduction page. The user first sees the introduction paragraph and then scrolls to see the graph of the data without any annotations. Only the graph is shown because then the viewer can get an understanding of the axis and structure of the graph before delving into it. The second scene then presents the first set of annotations. The third and fourth scene are formatted similarly, both having their own set of annotations. Each set of annotations are grouped together in a scene because they cover a particular trend in the graph. This way, when the user is moving from scene to scene, they are looking at a group of annotations that are all related and that further expand upon a certain topic. This helps with the cohesiveness of the story telling. The fifth and final scene is when the viewer gets to examine the data of their own. This scene includes an information box to let the user know what things they can explore and a tooltip functionality. The ordering of the scenes in this way then allows the user to seamlessly transition from the author-driven content, where the user just reads

the annotations, to the user-driven content, where the user then has full control and can research on their own.

**Annotations**

In my narrative visualization there are three scenes that contain annotations. The first collection of annotations are in the second scene. They reveal that for some columns there are a lot of data points, which leads to the conclusion that these job titles are currently more popular and available. For the jobs that are more popular, the annotations point out that the range of salaries are larger. The two jobs these annotations focus on primarily in this scene is Data Scientists and Data Analysts. The second collection of annotations is in the third scene and indicates that there are also roles with less data points and are less common when compared to other data science roles. In this scene, the positions of Product Data Analyst and Lead Data Scientist are highlighted. There is one more annotation on the fourth scene which identifies that there are some high salaries in the graph that are outliers. The annotations follow a template where each set of annotations are grouped together in a scene because they cover a particular trend in the graph. All the annotations in the narrative are placed in an order that allows the reader to clearly understand the main takeaways of the graph.

**Parameters**

Parameters and states are utilized on the final scene through the inclusion of a drop down selection bar. When the user gets to explore the data on their own, they are given the option of selecting a filter to put on the graph using a drop down bar. The drop down menu, which is what the parameter is bound to in this case, contains 5 different options which are different types of experience level. The states are then each of the experience levels available, which are entry level, mid level, senior, executive level, and all of them combined. The viewer can select a particular filter and the state of the graph would change accordingly.

**Triggers**

In the intro paragraph, the reader is told to use the NEXT and BACK buttons to be guided through the narrative. The user is able to use these buttons as triggers to go through each of the scenes and move from the author driven content with annotations to the user driven content where they get to explore the data on their own. Another trigger that is available is the drop down menu where the viewer can select different experience levels and filter out certain data in the graph which changes the state. Lastly, the tooltip that is available when a reader hovers over a data point can also be considered a

trigger since there is a popup with more information everytime the details on demand feature is triggered.

There are two main affordances that are available in this web page. The first one is the initial intro paragraph that gives the user a brief overview of the narrative visualization and introduces the next and back buttons in order for the user to be able to get started. The second one is the information boxes that are presented on the last scene of the narrative when the user driven content is introduced. One of the boxes gives a small breakdown of what types of features are available to the reader, for example the drop down menu or the tooltip. The info box also notes that the graph has a large x axis that can be further explored by scrolling to the right. It also points to another box that gives a quick overview of the abbreviations that are present in the graph in order to clear any confusions that the user may have.

**Reference Links**

- Narrative Visualization: https://aareanareza.github.io/
- Dataset: https://www.kaggle.com/datasets/ruchi798/data-science-job-salaries
- Source Code: https://github.com/AareanaReza/AareanaReza.github.io