Link to site: https://ds4200-finalproject.netlify.app/
Abigail Swanson, Emma Penn, Chelsa Batara, Grace Koo

Visualization 1:

Marks:
This graph uses line marks to track the trends over time. There are two separate lines, one for “Entry-Level” and one for “Students” both represented by a different color.

Channels:
One channel is position. The X and Y axis have different values which are important in understanding the purpose of the visualization. The X axis is representative of the month out of the year in which the job posting is made available. The Y axis is representative of the count of postings. The colors of the marks also act as a channel for this graph because it helps differentiate the two attributes for the audience.

Why you chose this vis:
We chose a line graph for this visualization because it best suits the need of tracking trends of numerical data overtime, especially when we are tracking categorical data alongside numerical data. For example, in this visualization we are keeping track of the month in which job postings are made available and just how many of those are posted.

Interaction Explanation:
The interaction for this plot allows users to see a visual representation of the availability of certain jobs within the NYC public sector. This graph has two different career levels, entry-level and student. The user can use the color key on the right to differentiate between the two distinct marks. This graph is static so there is no immediate interaction between the user and the graph. Overall, the interaction will provide more insight into trends and patterns of job postings within the entry-level and student roles.

Visualization 2:

Marks:
The marks of this visualization are the rectangular bars itself on the graph but also the stacking of the different career levels within the bar. The bar as a whole is used to represent the count of records for each job role across agencies. Each segment of the bar corresponds to a specific career level.

Channels:

The visualization uses position (horizontal and vertical) as a channel to differentiate points. The X-Axis represents the total count, while the Y-Axis corresponds to each agency. Color also is a channel in this graph which is used to distinguishes the job levels, and the stacking visually shows the proportion of each level within the total. These channels work together to compare workforce distributions both across and within agencies effectively.

Why you chose this vis:
This visualization was chosen because it allows for a clear comparison of job-level distributions across agencies using a horizontal layout. The use of different colors for job levels makes it easy to distinguish between categories, and the stacked format effectively summarizes the total workforce while showing the breakdown by job levels.

Color Explanation:
The colors in this visualization are used to represent different job levels from 'Executive, Manager, Experienced (non-manager), and Entry-Level'. Each job level is assigned a distinct color, as indicated in the legend, to make it easy to visually differentiate between them within the stacked bars. The choice of colors helps viewers quickly identify patterns or trends of job levels across agencies, and ultimately enhances the readability for the audience.

Visualization 3:

Marks:
The marks in this visualization are rectangular heat map cells. Each cell represents the average salary range for a specific career level and month. The size of the marks is uniform to emphasize the differences in salary levels through color intensity.

Channels:
The primary channel used is the color scale. Dark purple indicates higher average salaries, while bright yellow represents lower average salaries. This color gradient allows for immediate recognition of salary hierarchy across career levels. The horizontal position (x-axis) represents the posting month, while the vertical position (y-axis) differentiates career levels such as Executive, Manager, Experienced, Entry-Level, and Student.

Why you chose this vis:
This heat map was chosen because it effectively communicates trends in salary ranges and their variations across career levels and months. The color intensity provides a clear visual cue for identifying patterns, such as higher salaries for executives and lower salaries for students. The simplicity of the layout makes it easy to interpret, even for audiences unfamiliar with the data, highlighting seasonal salary trends and job posting dynamics.

Color Explanation:
The colors in the heat map range from dark purple to bright yellow, providing a clear distinction between high and low salaries. Dark purple cells represent positions with the highest salaries, such as executive roles, while yellow cells indicate lower salaries, typically associated with student and entry-level roles. This colorblind-friendly Viridis palette ensures accessibility and inclusivity, allowing all viewers to interpret the visualization accurately.

Visualization 4 (the interactive scatter plot):

Marks:
The marks include points to represent each data point.

Channels:
The channels include position (x and y axis) as well as color to represent the different career levels.

Why you chose this vis:
We chose this visualization to look at how the number of positions compares to the salary posted as well as to represent how this would change based on the career level. Since both the salary and the Number of Positions for the posting are numeric variables we chose to use a scatter plot to represent these variables because scatter plots are used to plot 2 numeric variables.

Interaction Explanation:
The interaction will allow users to select a specific career level by filtering out any other points that are not a part of that career level allowing for a clearer view on if there is any relationship between the number of postings and average salary between each of the career levels. Additionally, it helps the user to better identify if there are clusters among different career levels as the plot that includes all career levels has overlapping points which can make it harder to identify clusters/trends among certain groups.

Visualization 5 (The Boxplot):

Marks:

The marks in the boxplot are area for the quartiles, lines that represent Q1 and Q4 as well as points to represent the outliers.

Channels:

The channels for the plot include position (x axis and y axis) to represent the different categories as well as the area in which bigger quartiles/larger distributions will have more area for the area mark. Another channel is length, the length of the lines in Q1 and Q4 differentiate among different categories to represent bigger/shorter distribution.

Why you chose this visualization: We chose this visualization to see the distribution of a numeric variable, the average salary, amongst different categories of a categorical variable. The options to view the distribution are histograms as well as a side-by-side boxplot however, since we wanted to be able to easily compare these distributions, the side-by-side boxplot is best for looking at more than 2 categories without the graph feeling too cluttered. For these reasons, the side-by-side boxplot was best to achieve these goals.

Visualization 6 (The Interactive Line Plot):

Marks:
The marks to represent that data points include lines.

Channels:
The channels which different the data points include the position (x and y axis) additionally the color to represent the different career levels.

Why you chose this vis:
We chose to use a line pilot to represent the interactions between the date and the average salary because line plots are best used to look at the change of a variable over time. The line plots help to show the trends of the mean average salary over time which we wanted to see how it would change, even within the year.

Interaction Explanation:
The interaction for this plot allows users to select a certain year to see how the average salary changes over the course of the months of the selected year. This will allow users to have a selected “zoom” into each year as we previously plotted the distribution of the average salary per year. This graph includes the differences in the career levels so the user can look at if there is a specific career level that might have contributed more to the increasing distributions per year. Overall, the interaction will provide more insights to a selected year and its average salary distribution.