

HCI – Lab3: Data Visualization Report

1852448 Li Yuanfeng

1 Data Analysis Task Specification

1.1 Executive Summary

The graduation income of different college students has always been a hot topic in society. It is closely related to school choice, major choice, residence choice, etc. This report will provide a brief analysis of the graduation earnings of students from different universities in the United States in different dimensions.

1.2 Dataset

The dataset The dataset is divided into 3 sub-datasets according to the different classification methods. They are divided by “major”, “school type”, and “region” where the school is located.

Among them, the dataset classified by major is in terms of profession. One major corresponds to one data. While the dataset classified by school type/region is in terms of school. One major corresponds to one data.

For each piece of data, the following attributes are included:

1. Start Median Salary.
2. Mid-Career Median Salary.
3. Mid-Career percentage Salary (10th, 25th, 75th, 90th).

➤ What is the percentage salary?

A percentile salary tells how much a certain percentage of an overall population in a geographic area or withing a given industry or field makes. The most commonly-used percentile salaries are the 10th, 25th, 50th, 75th and 90th percentiles. As an example, if the 10th percentile wage for a job was \$10,000 annually, 10 percent of the people made \$10,000 per year, while the remaining 90 percent made more than \$10,000. Percentile salaries are usually annual figures, though they can also be reported as hourly wages.

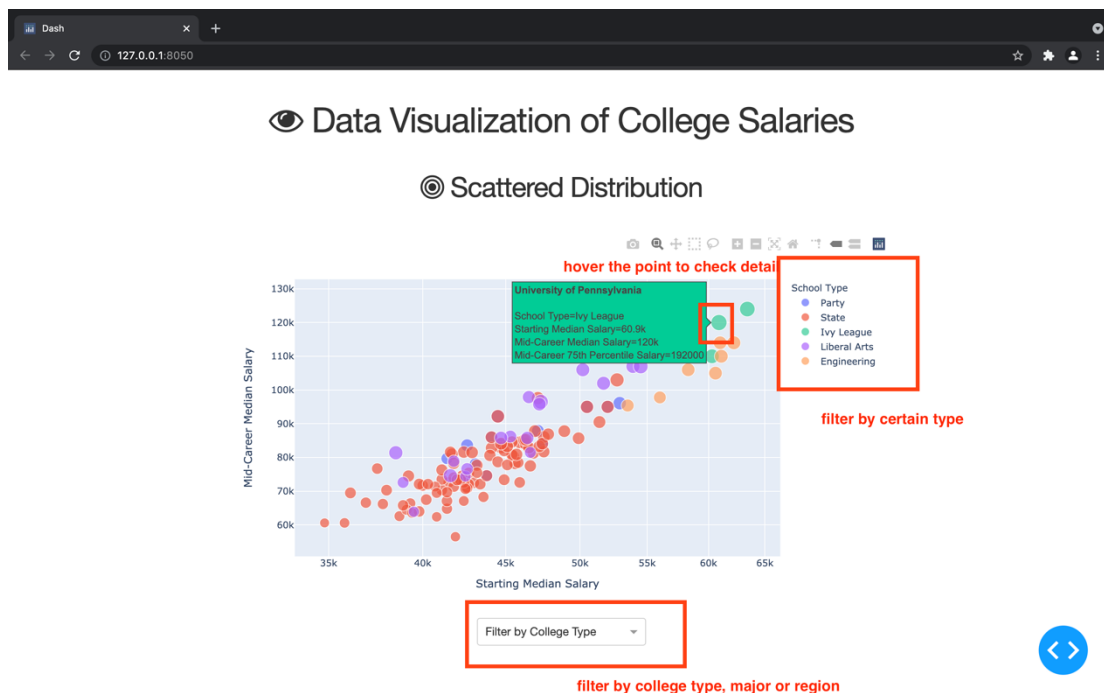
1.3 Dataset Characteristic

1. Multiple dimensions exist in the dataset to reflect students' wage characteristics from multiple perspectives.
2. There is no mutual constraint relationship among the data sets, but there is a comparative relationship. It may be necessary to use line or bar charts.
3. For the attribute "mid-career percentile salary", it may not be very meaningful to analyze it alone, but combining it with the process of change can clearly reveal the overall situation of the salary range.

4. Some specific conclusions can be drawn by combining some of the columns in the table:
 - Combine different Mid-Career Percentile Salary reflects differences in income of people in the same career period.
 - Combine starting median salary and mid-career median salary reflects the development potential occupied in certain major or graduated from certain university.
5. The dataset also implicitly shows the distribution of income groups.

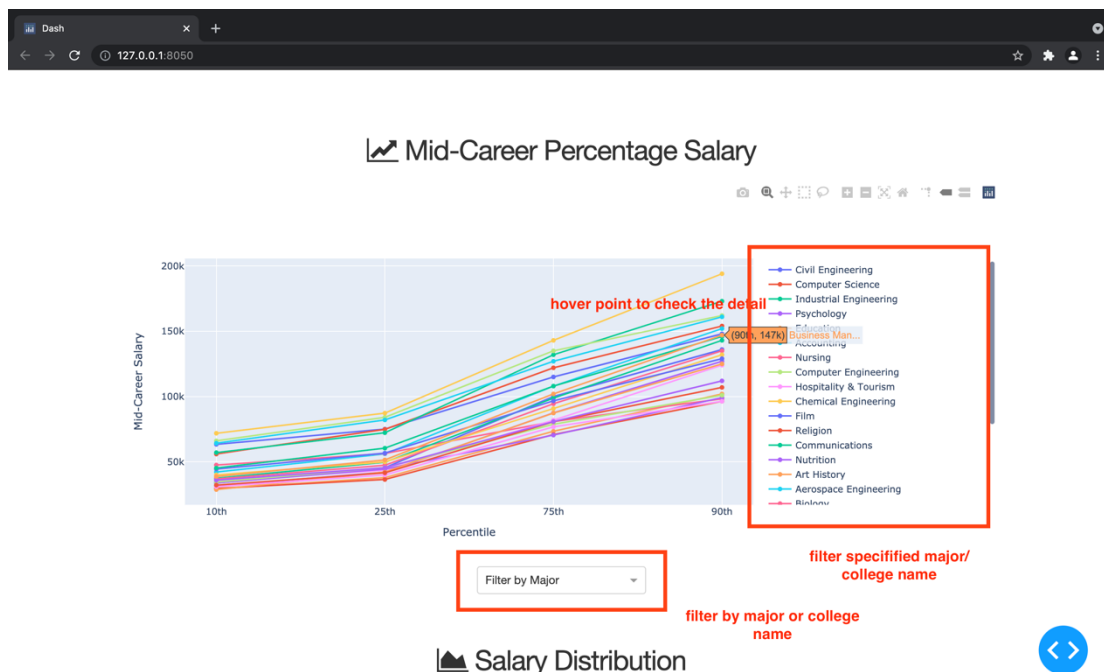
2 Dashboard design

2.1 Bubble Chart: Scattered Distribution



- Function: Users can intuitively access a particular data through this graph. Also, the density of the points gives a good indication of the distribution of salaries.

2.2 Line Chart1: Mid - Career Percentage Salary



- Function: Users can visually derive the wage differences between similar populations based on the altered graph. If the curve is flatter, it means that the difference between the actual income of the high-income group and the low-income group is smaller in this type of population. If the curve is steeper, the difference between the actual income of the high- and low-income groups is larger.

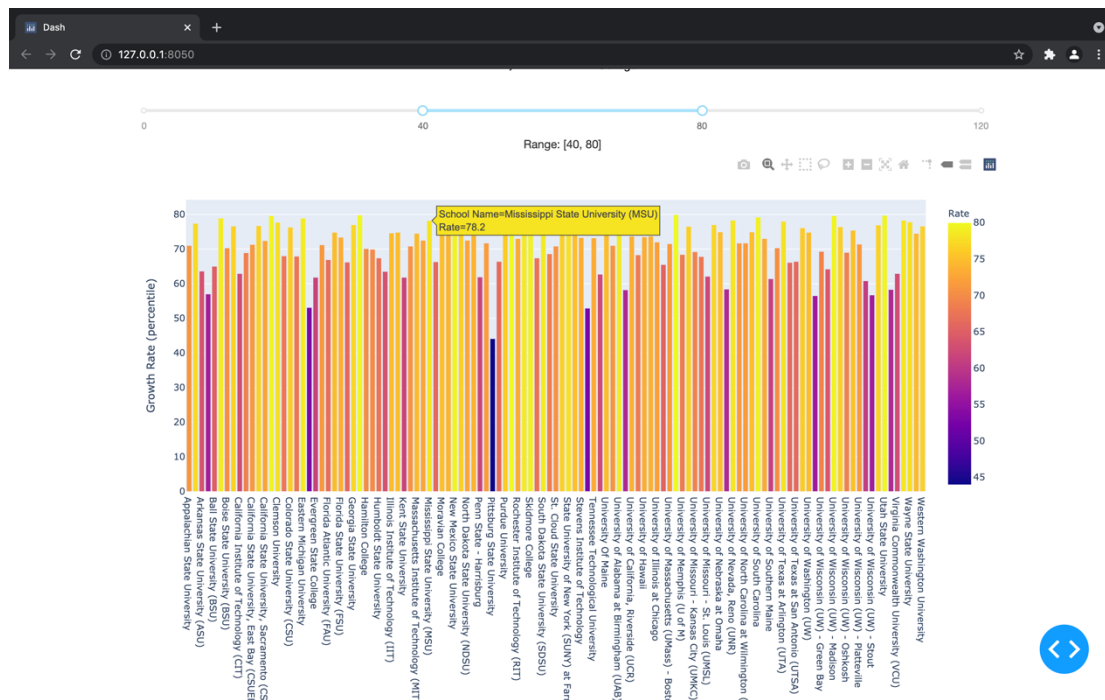
2.3 Line Chart2: Salary Growth Rate of different major



- Function: This chart visually reflects the development potential of different major. If the

growth rate is high, it means that the salary of the major is growing fast. This chart can provide help in choosing a major.

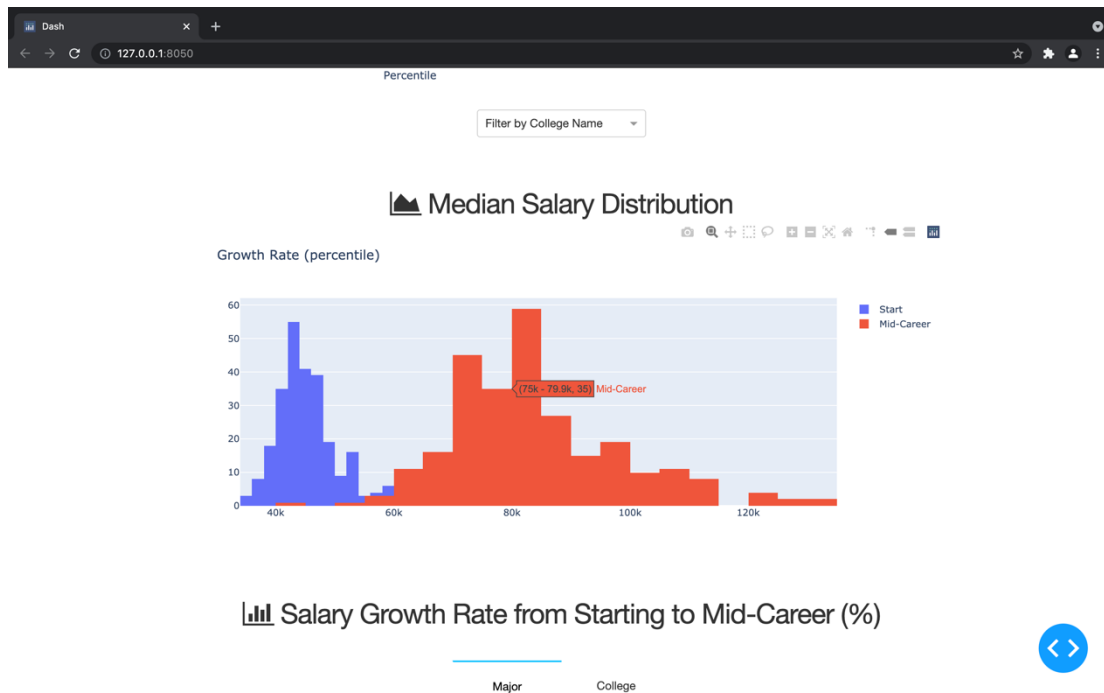
2.4 Bar Chart: Salary Growth Rate of different college



- Function: This graph visually reflects the potential for talent development in colleges and universities. If the value of the bar is higher, it means that the talents graduated from that college can grow rapidly in salary after they are employed. This chart can and does reflect to some extent the quality of education in colleges and universities.

(Since there are so many colleges, the line chart design makes the folds very dense and difficult to see. So here the bar chart design, and make the color changes with the value, to provide a more intuitive visual effect.)

2.5 : Histogram: Salary Growth Rate of different college



- **Function:** The vertical axis of this graph is the frequency of occurrence of each type of income. This chart reflects the distribution of graduate earnings by college and university in the United States. By looking at the peaks of the chart users can quickly find the average earnings of graduates and provide a measure of graduate earnings. It also allows for a comparison of the difference in earnings distribution between the beginning of a career and the middle of a career.