Project Documentation

Lab 3 Data Visualization

1752095 范思琪

I. Data Analysis task

There are three datasets about college salaries. They are *salaries by college type*, *salaries by region* and *degrees that pay-back*, each of them focuses on different points of college salaries:

- *Degrees that pay-back* focuses on different undergraduate major has different percent change from starting to mid-career salary.

- *Salaries by college type* focuses on salaries changes from starting to mid-career differ in school type.

- *Salaries by region* focuses on salaries changes from starting to mid-career differ in region.

As a result, I choose one data analysis task for each dataset of college salaries and visualize correlated data to complete the task.

For *salaries by region*, two comparison tasks should be completed in the visualization:

- Salaries changing with regions in same period of career stage.

- Salaries changing with different periods of career stage in same region.

For *salaries by college type*, two comparison tasks should be completed in the visualization:

- Salaries changing with different college types in same period of career stage.

- Salaries changing with different periods of career stage in same college type.

For *degrees the pay-back*, only percent changing from starting to mid-career salary changes with different undergraduate major should be shown.

# II. Layout of designed dashboard

According to the five tasks above, I designed three different kinds of figures to show the three datasets of college salaries.

For *salaries by region*, a scatter plot is designed. In the plot, different regions' salary condition is showed in different color. X-axis and y-axis both represent salary in a certain period of career stage. These two periods of career stage can be chosen in two labels. As a result, the figure is shown as below.
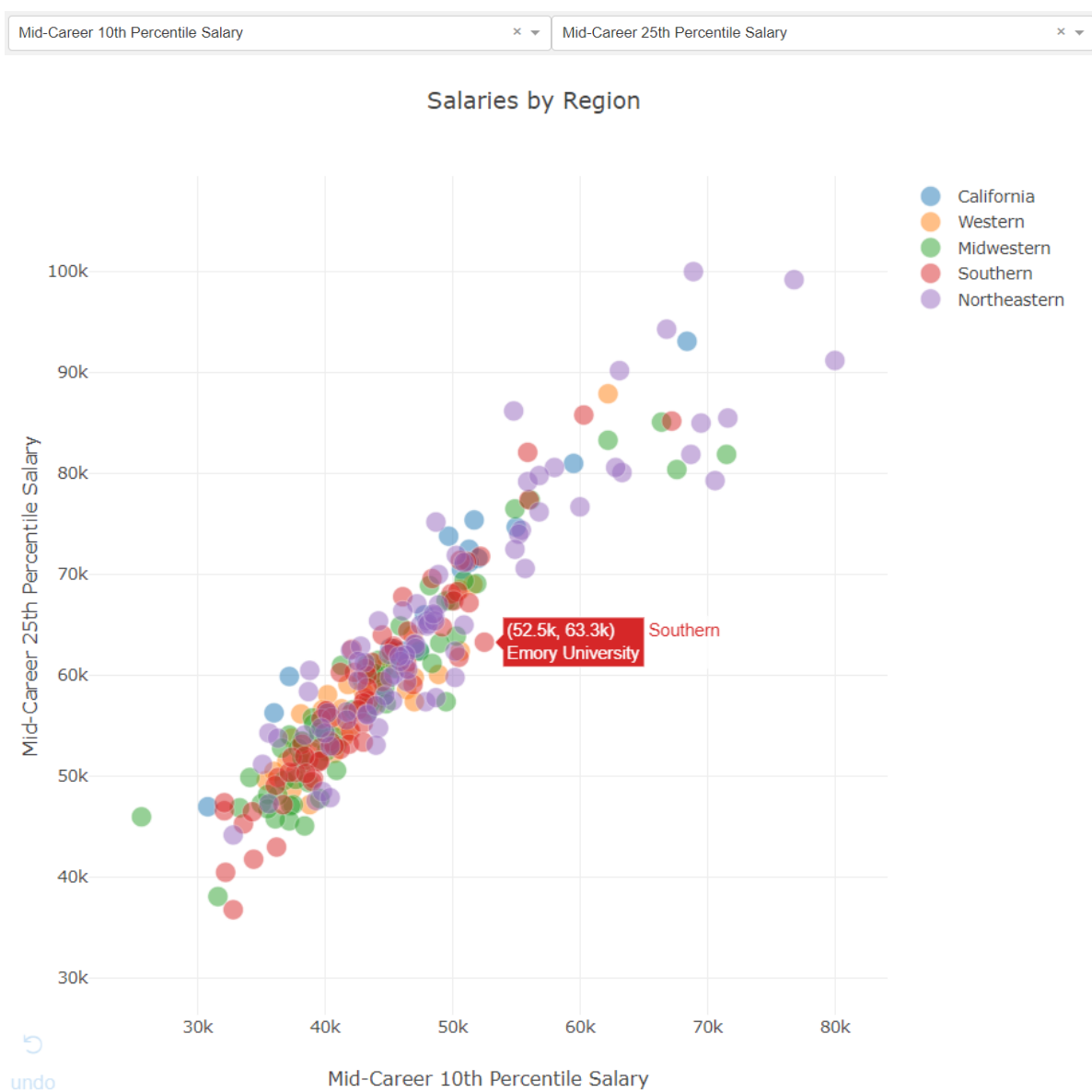
The text of the point will show the university where the data from. You can change the period of career stage and whether to display the data of every region. You can see the effect as
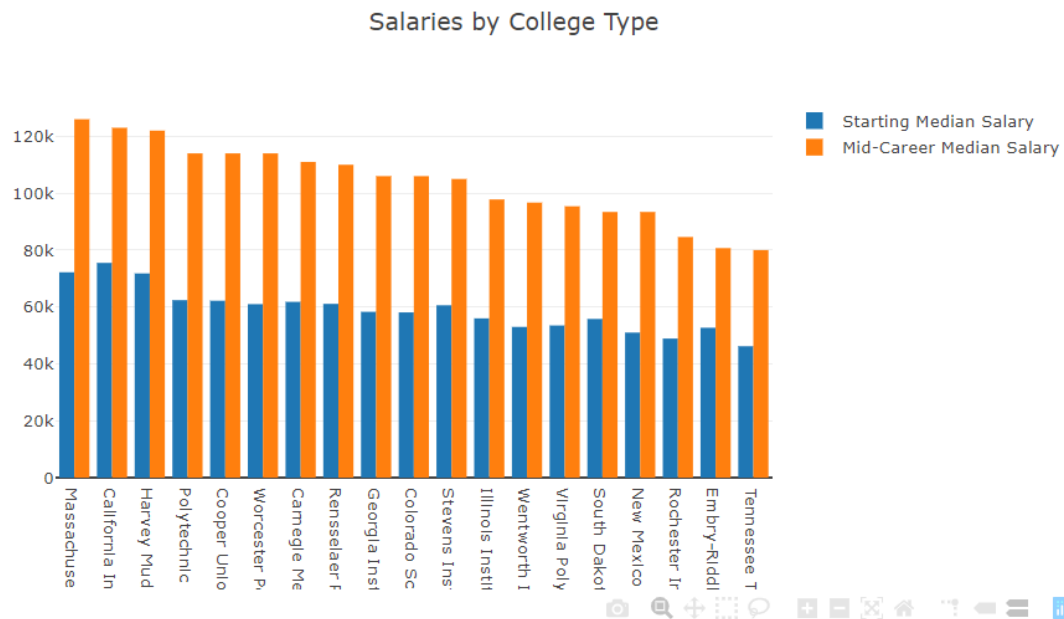
below.
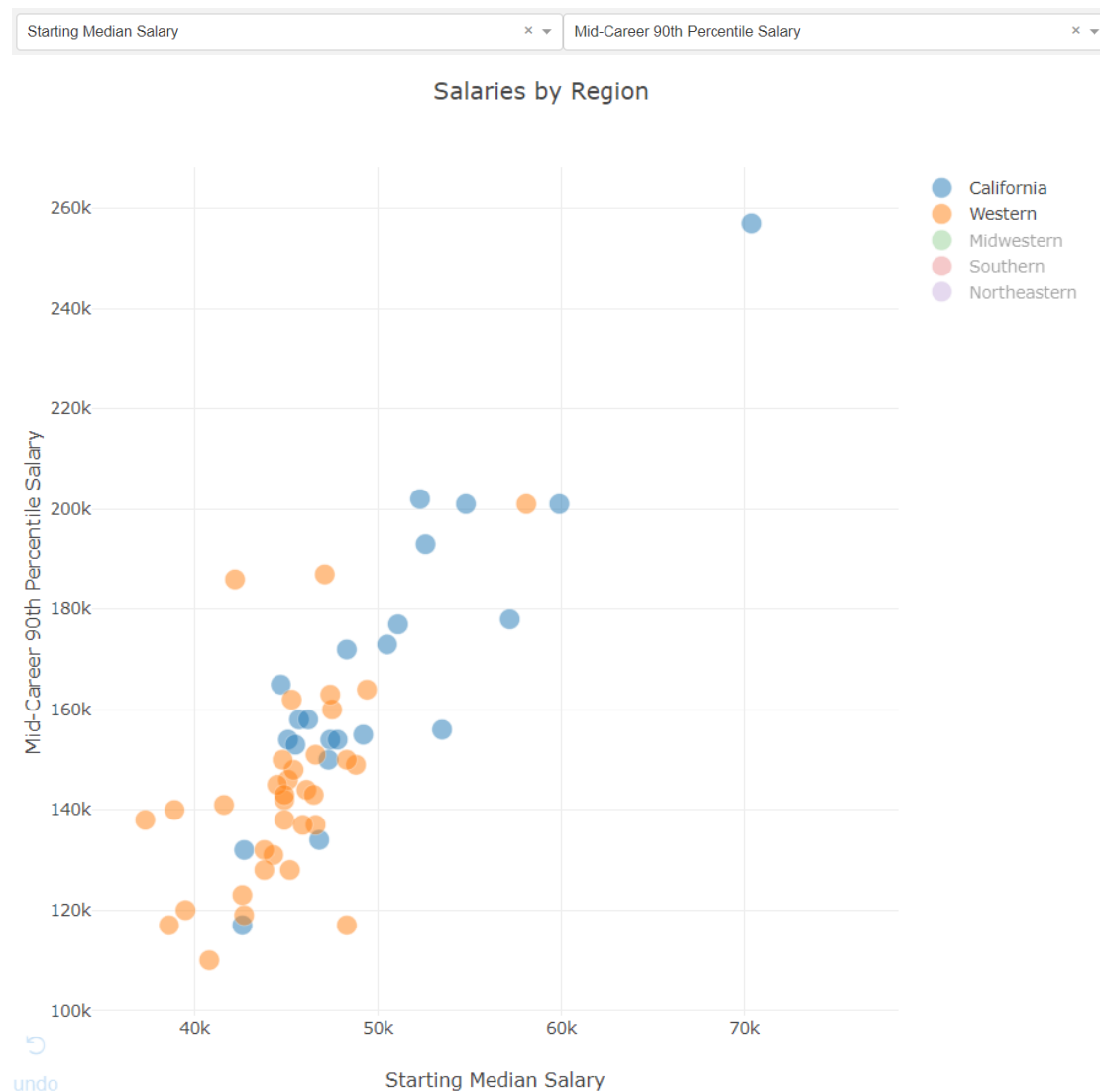
For *salaries by college type*, a bar plot is designed. In the plot, x-axis represents different college of a certain type while y-axis represents the salary data of the correspondent college. There are two bars to show different periods of career stage. Like the scatter, these two periods of career stage and the college type can also be chosen.
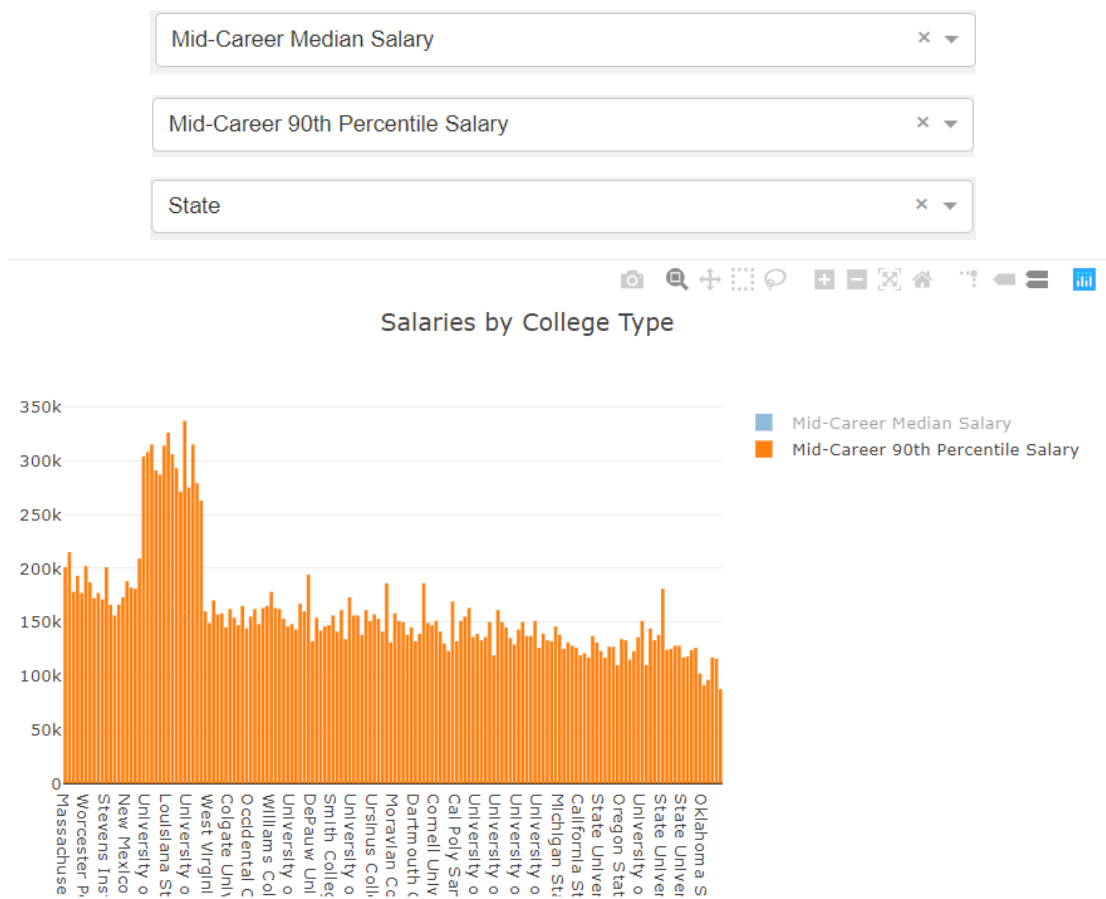
Salaries by College Type

You can also choose whether to display every period of career stage as below.

Salaries by College Type

For *degrees the pay-back*, a line graph is designed. In the graph, you can see percent

change from starting to mid-career salary which changes with undergraduate major.

Percent change from Starting to Mid-Career Salary
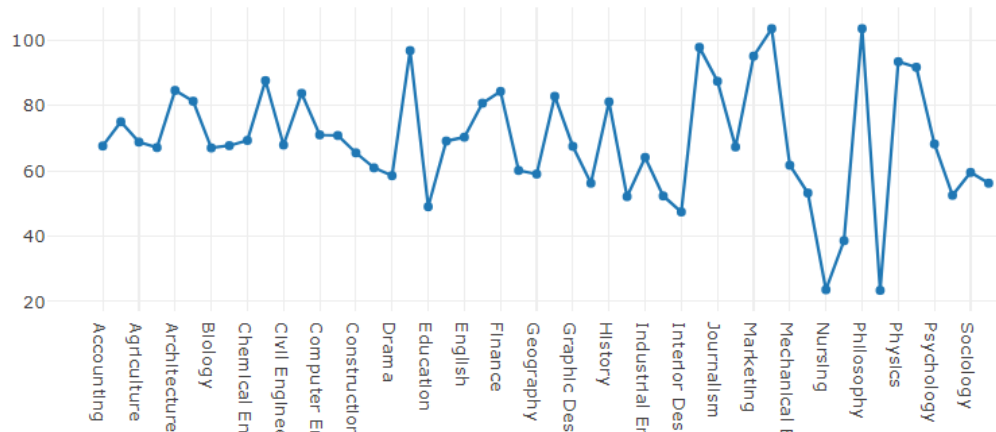
After all, you can see the three figures as below.