Post-grads Chasing the American Dream

Process Book

Audasia Ho (audasiah)
Elizabeth Ji (emji)
Jacob Zimmerman (jezimmer)

Overview & Motivation

As a child, I asked my father “What is the American Dream?” He was quick to respond: “It’s that everyone be able to grow up, own a house, start a family, and be happy.” It was only later that I learned of the canonical definition by James Truslow Adams: “life should be better and richer and fuller for everyone, with opportunity for each according to ability or achievement.” While maybe not the entire picture, home ownership is a large part of attaining the American Dream for many people.

As college students about to graduate, thinking about our own future potentials, we got to thinking about others like us. Earning potential post-graduation is intimately tied to area of study in college. Given the earning potentials for college majors around the country, where in the United States can they afford a home?

As a stepping stone to home ownership, most graduates choose to rent immediately after graduating, while saving for the future. We’ll investigate how earning potential interact with rental prices to understand where college graduates can afford to live.

Related Work

We found that there are many sites that have tables of average salaries per college major, however we were unable to find visualizations of this data. FiveThirtyEight, whose data we are using for this project, did an article about best paying college salaries, however they also just displayed the data in tables. We wanted to display the data in a way that would visualize the difference in college major salaries.

Trulia has a visualization of home prices for the entire country, creating a heatmap to show where in the country housing prices are the most and least expensive. The map is interactive, allowing users to input cities, counties, or states to see the heatmap for that area. We want to create a map similar to Trulia’s heat map, except instead we want to use housing rental prices. We will also color the heatmap based on the median salary for a college major, rather than use a static number scale.

We draw inspiration from Million Dollar Blocks, which visualizes the amount of money the state of Illinois spends to keep people incarcerated, and juxtaposes it with the city block that

inmates live on. This visualization brings together two datasets that are seemingly unrelated, and shows the correlation of the two datasets. We felt that this visualization brought up many interesting points about incarceration, that most people would not normally think about.

Questions

Our primary aim is to explore the relation between what people choose to study in college and where they may be able to live after graduation. This will be accomplished by exploring two relations; one is how college major affects median salary and the other is how expensive renting is in different parts of the US. By combining these datasets, we can show how a student's choices impacts their future. Some benefits include students being able to better understand where they afford to live, how much of their salary they will be spending on their rent, and possibly how other choices (how many roommates, how many bedrooms, etc.) can impact their cost of rent.

Data

We used FiveThirtyEight's dataset of earnings based on college major to understand the relationship between college majors or industries and their median starting salary. We are using Zillow's rental data set to understand the cost of renting in various locations. FiveThirtyEight and Zillow have produced this information in the anticipation that people will use it, and the data reflects this: as the data is well-structured and easy to understand.

The first thing that we did to process the data was to delete columns of data from both datasets that we did not need for our visualizations. We didn't want to import unnecessary data and clutter up the datasets.

We had gone through the FiveThirtyEight college major data in order to calculate the average median salary for industries as well. However, we also found that the number of majors per industry is unbalanced, as shown in the table below:

major_category	**count**
Engineering	29
Education	16
Humanities & Liberal Arts	15
Biology & Life Science	14
Business	13
Health	12

Computers & Mathematics	11
Agriculture & Natural Resources	10
Physical Sciences	10
Psychology & Social Work	9
Social Science	9
Arts	8
Industrial Arts & Consumer Services	7
Law & Public Policy	5
Communications & Journalism	4
Interdisciplinary	1

Because of the imbalance of majors per major category, we were unsure about the validity of an average median salary per major category, for now we have decided not to have a bar chart with the average median salary per category, as the statistics may not be accurate. Instead, we will have different bar charts for the different majors within a category.

For the Zillow dataset, we deleted all of the columns that were not pertinent to our project. We found that in order to plot the data for the cities, we had to find the coordinates of all of the cities. Once we found these points, we will be able to plot the colored circles for the map visualization.

Exploratory Data Analysis

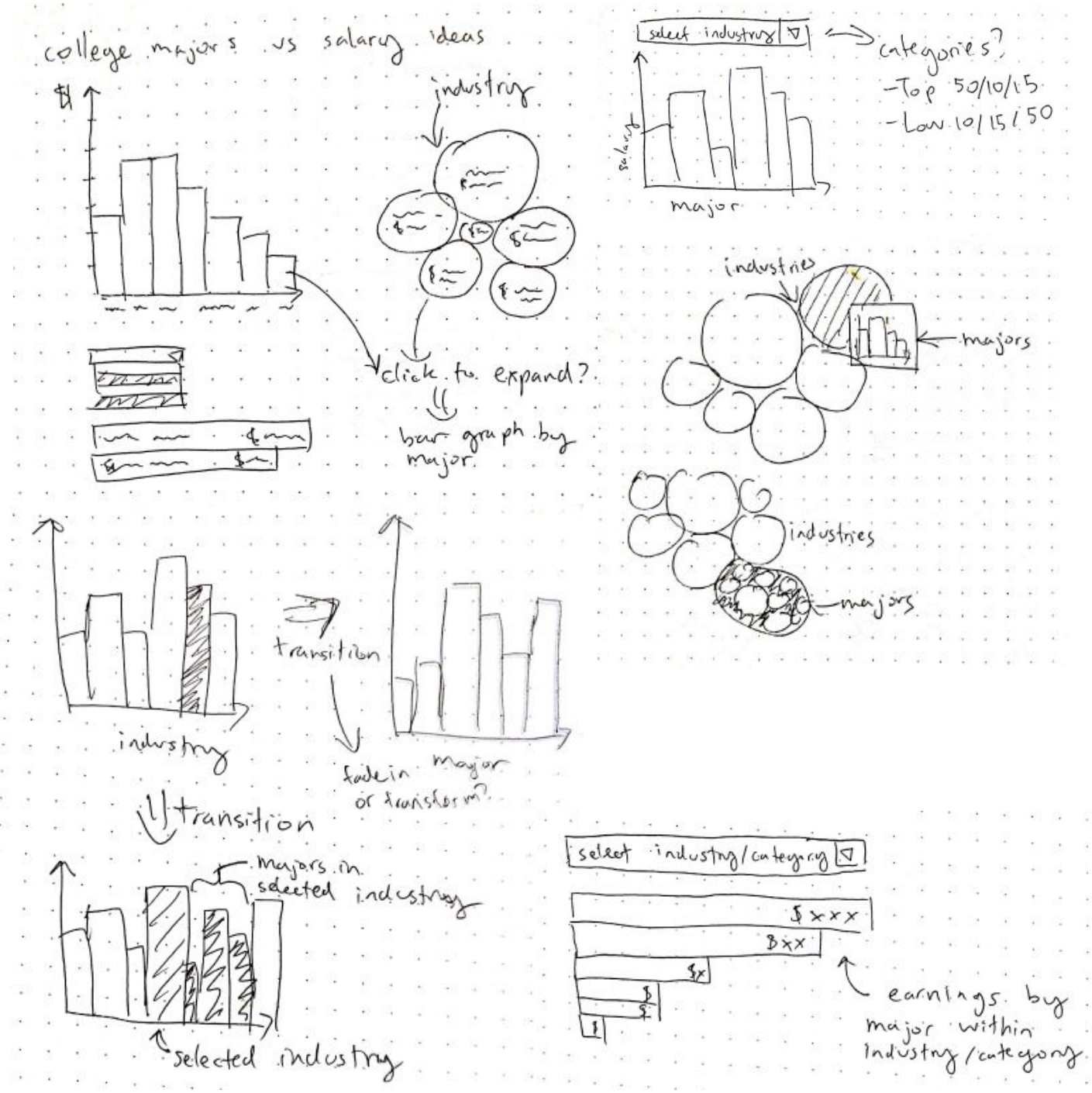
When we first started this project we initially wanted to create a heatmap of the entire country based on Zillow's dataset. However, after analyzing the data we found that the dataset does not contain all zip codes in the country, and thus we would be unable to create a heatmap that had the granularity of zip codes. Instead, we decided that we would focus on metropolitan areas in the United States, as we assume that most college graduates would prefer to move to a metropolitan area after graduation. Our map view will now have data points that are color coded like a heat map, but it will only be for metropolitan areas instead of the entire country.

Design Evolution

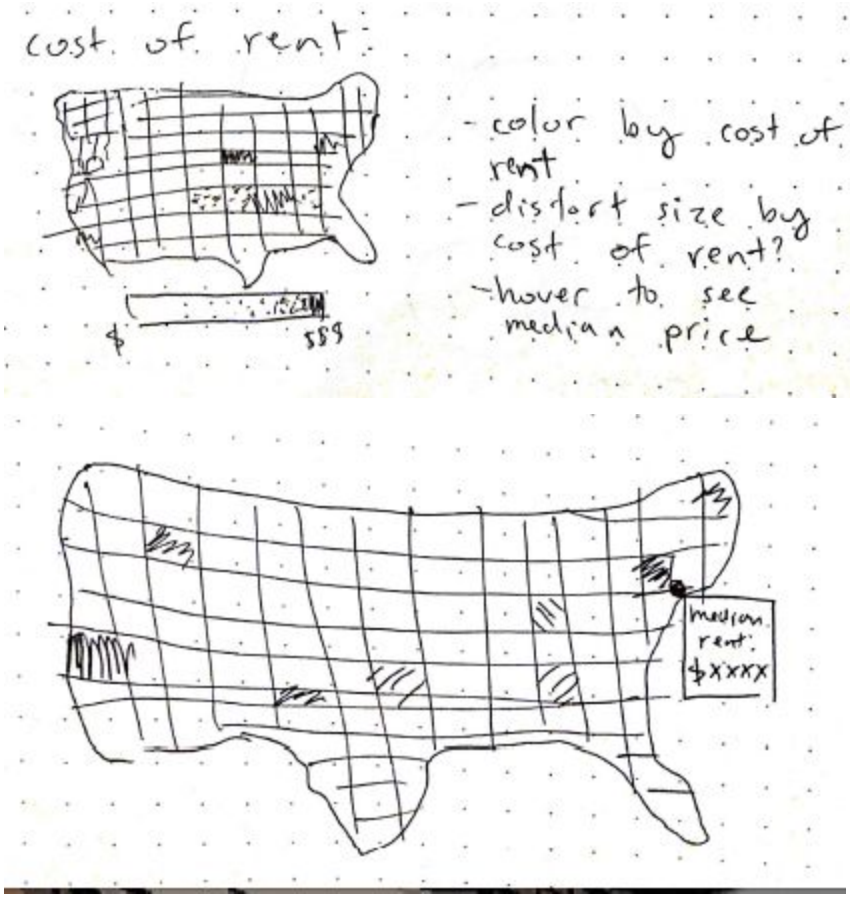
Initial Sketches

When we first started working on this project we decided that we wanted three visualizations of the data, one of each dataset alone, and then the final one would connect the two datasets. Below are our initial sketches for the three different views:

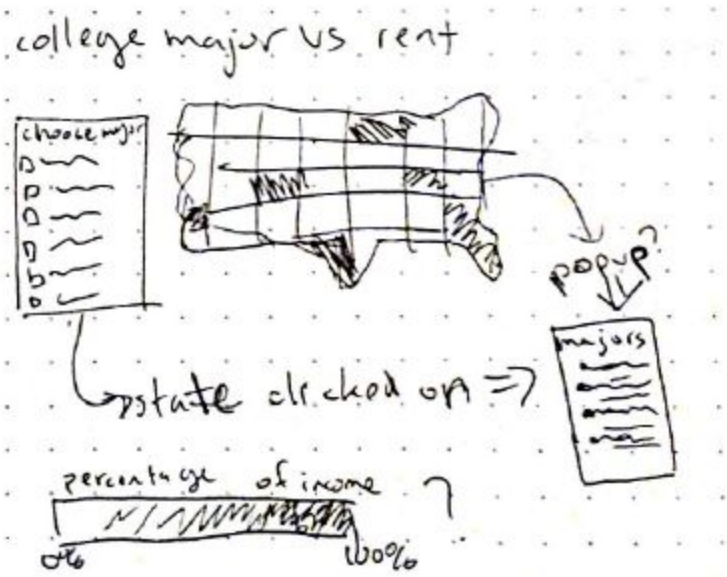
View 1: Bar Chart of College Salaries

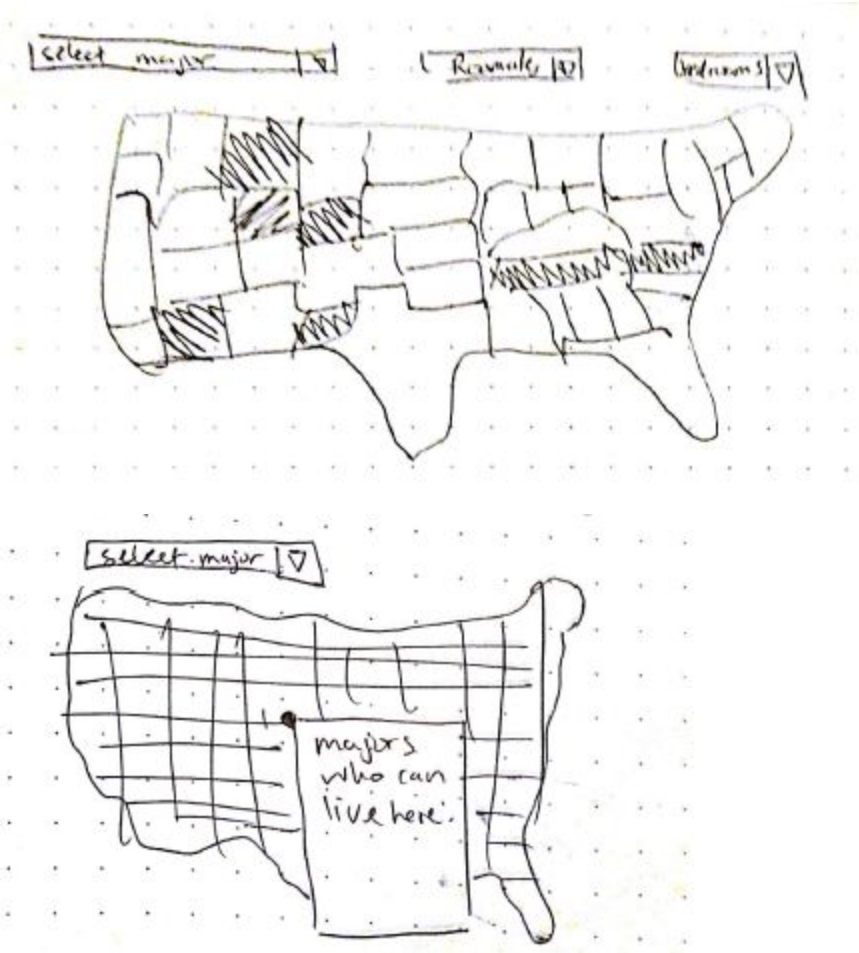


View 2: Heat Map of Rental Prices in US



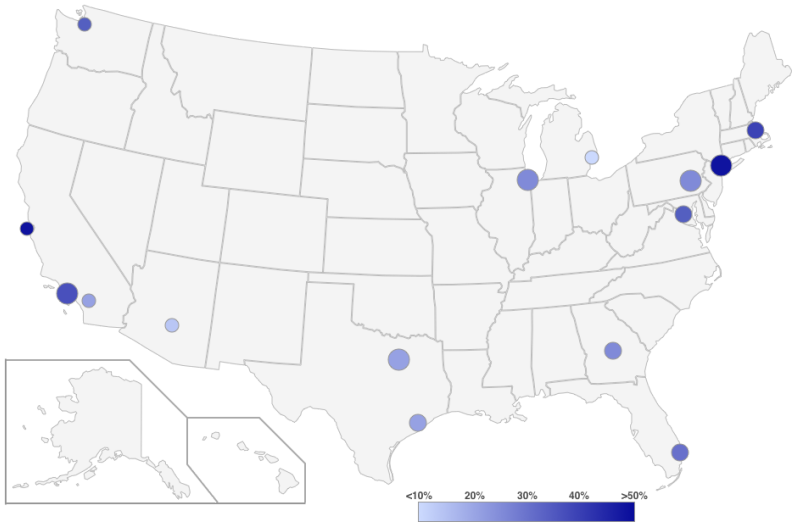
View 3: Interactive Map of Rental Prices and College Salaries





Digital Sketches

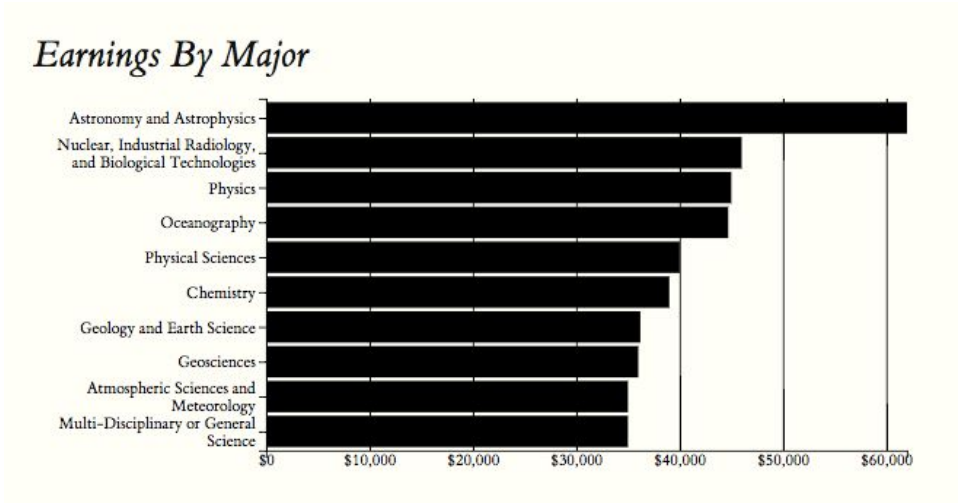
After we had our pencil sketches, we decided to create digital sketches in order to better visualize what our views would look like. We created the below digital prototypes of our two static views:

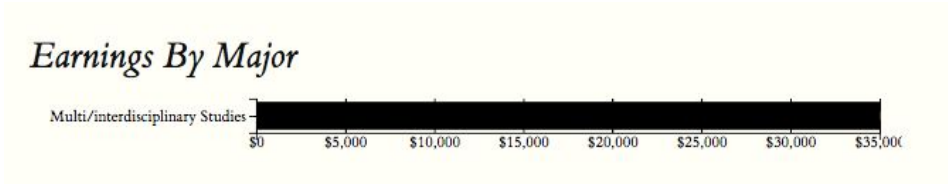


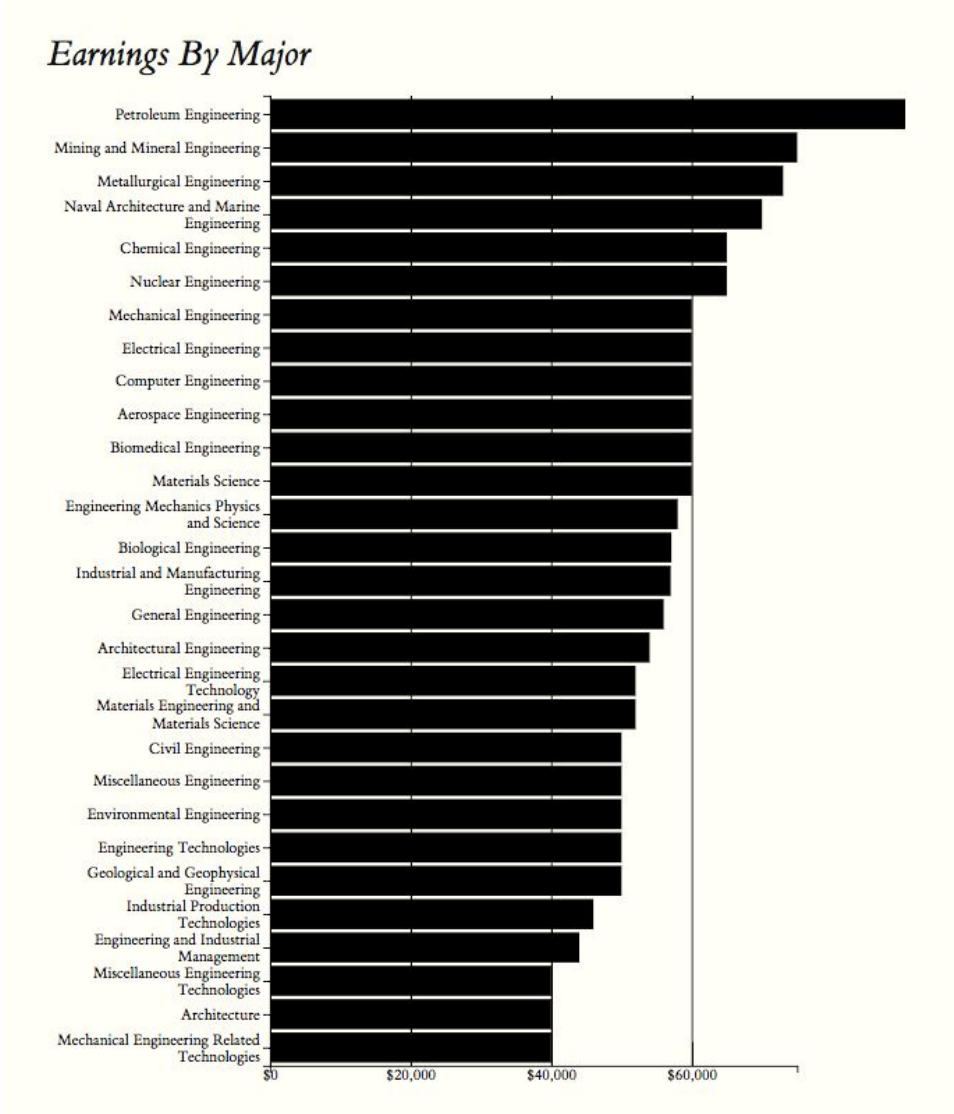
After reviewing our views and sketches, we decided that having all three views was redundant and unnecessary. We decided to get rid of View 2 and merge it with View 3. Thus, our idea slightly changed from our proposal, and we instead will have two views. The first view will be as we initially said, a bar chart of median salaries for different college majors across different major categories. Our second view will be an interactive map view, with the default view being average rent prices of housing in the top metropolitan areas, scaled to the average salary for all college graduates. We connect our two views, by allowing users to select majors or categories from the bar chart, and the map will update to show the affordability of housing for the particular selection.

Bar Chart of Median Salary for College Majors

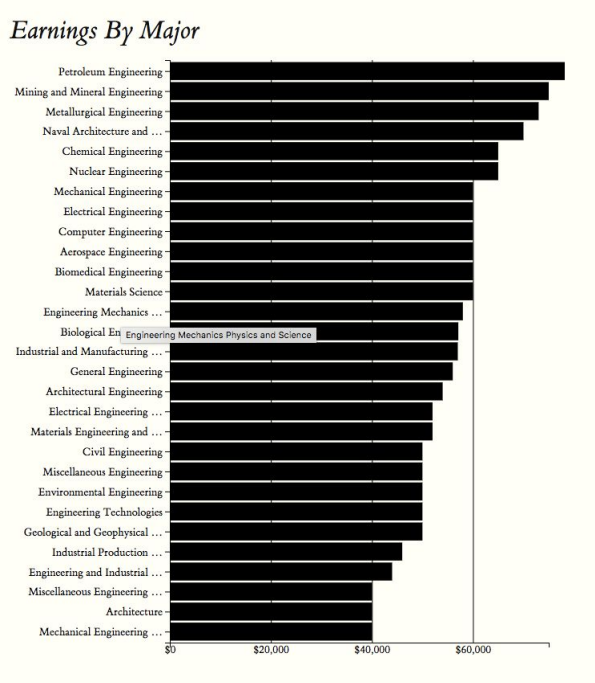
We initially started with a basic bar chart for our median college major salaries. We found that the visualization of the bar chart for different major categories was unbalanced, as some major categories have more majors than others. We are unsure about how to solve this issue, but we have attached our preliminary prototype views below:

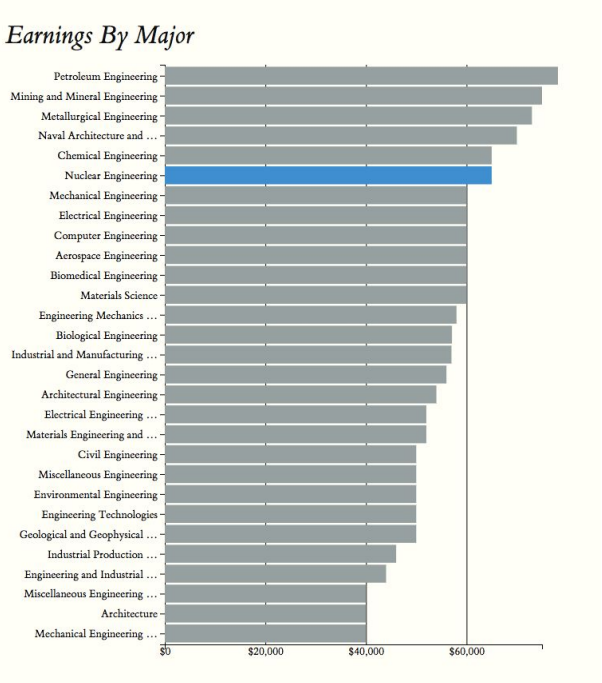




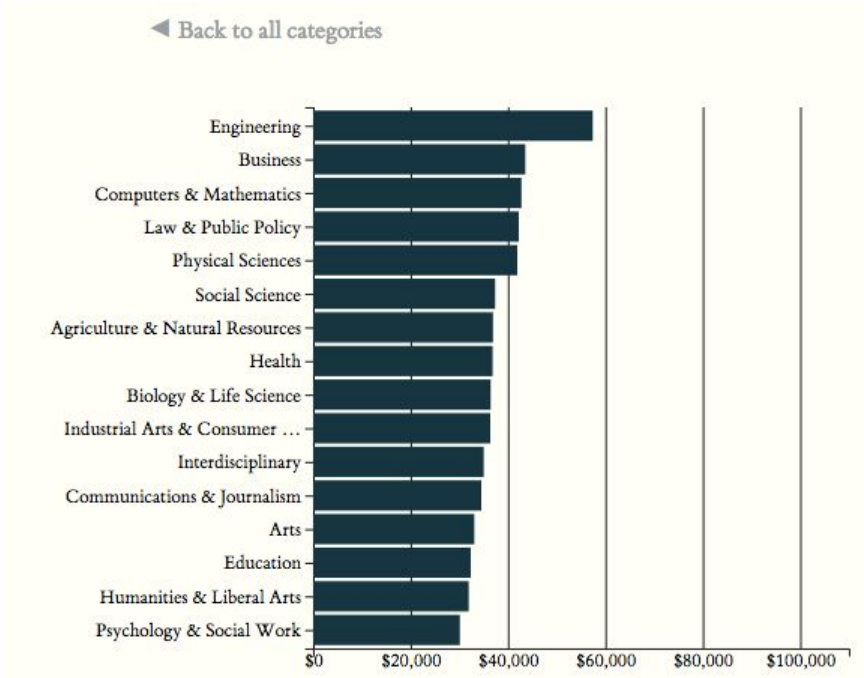


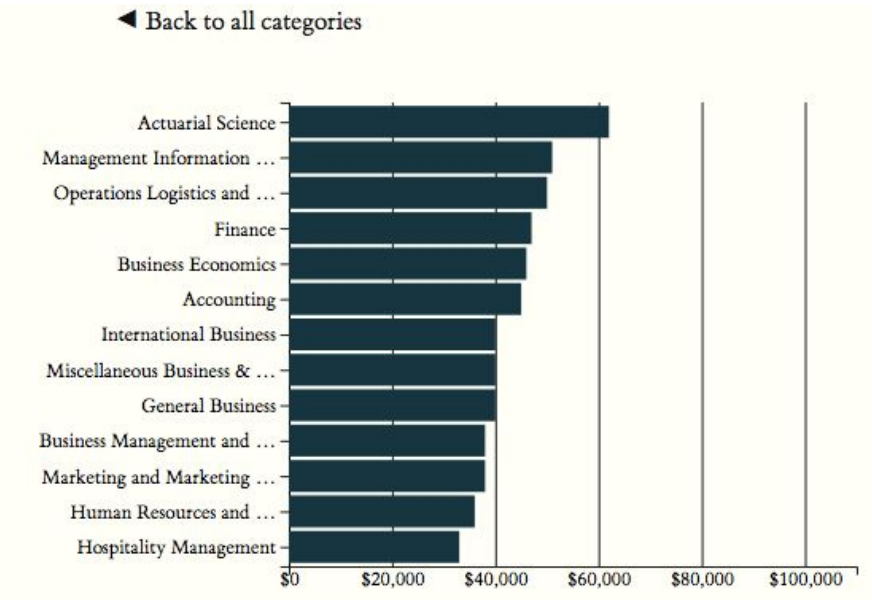
We then changed the colors of the bar chart, as well as truncating the majors and having them be displayed with a hover instead. This is because we didn't want the text labels to get too long and close together, which would be harder to read for the user. In the future we may change the coloring of the graph into a color scale, we are still not completely decided.





After we had a preliminary bar chart, we decided to have a top level view of majors by category, so we have the initial bar chart be all of the categories, with the average salary for the entire category. The user can then click on a category bar, which will transition the bar chart into showing all the majors within the category. We add a button at the top of the chart in order to go back to the category view.

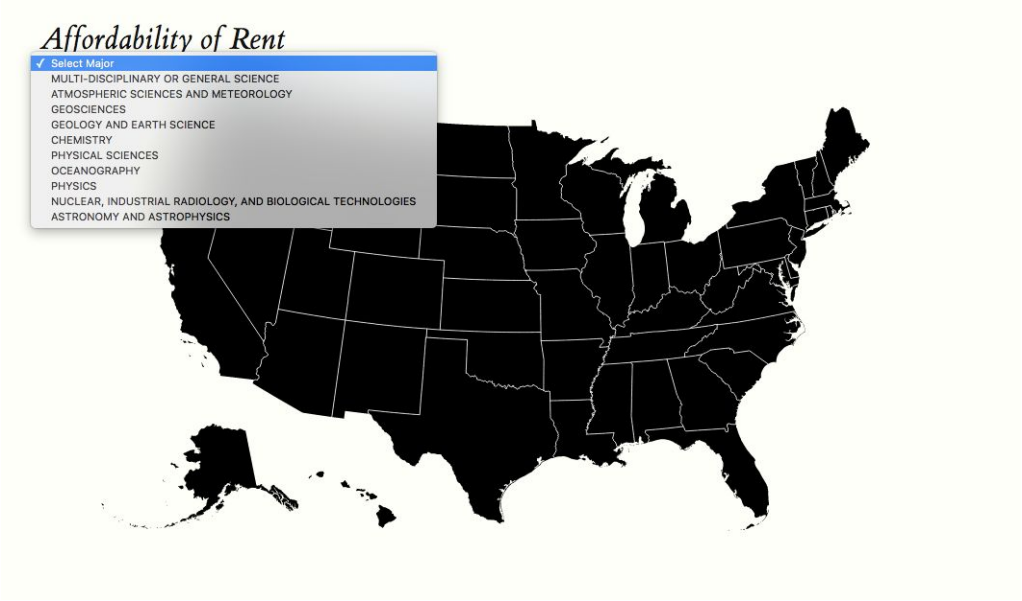




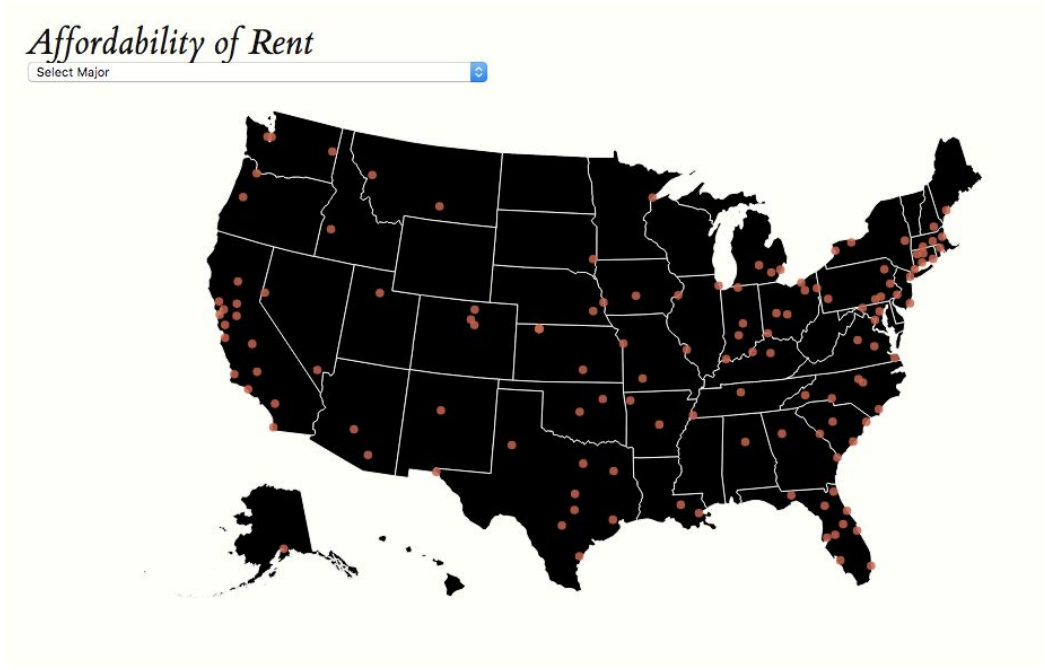
Heat Map of Average Rental Prices by Metropolitan Area

We have created an interactive map view that displays rental prices for the top metropolitan cities in the country, and allow users to select the college major category that they would like to see data on.

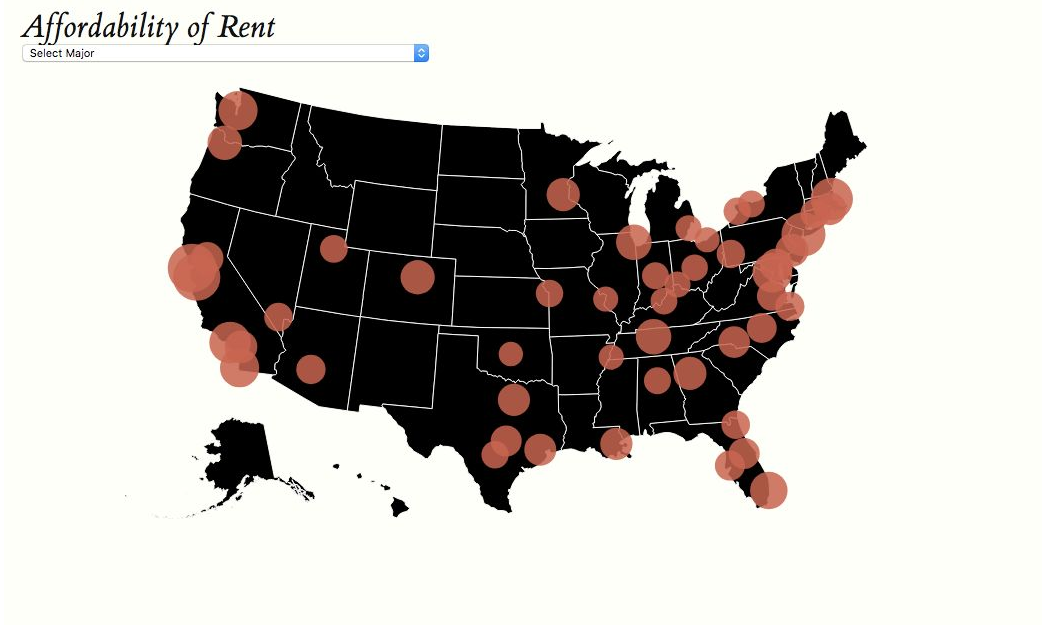
We first started with displaying a map of the United States and creating a drop down menu to allow users to select a specific major category.



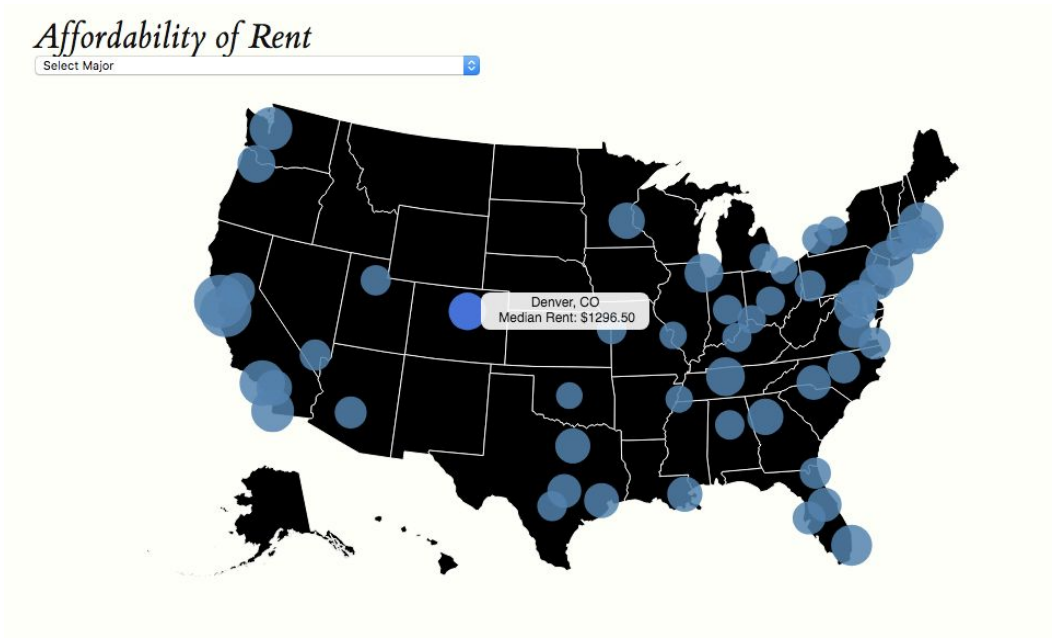
We then created data points for the different metropolitan areas on the map, in order to show a user the exact location of a metropolitan area for context.



We decided to scale the size of the data points based on the average price of rent in that metropolitan area. This will allow users to tell at a glance which metropolitan areas have the highest rent.



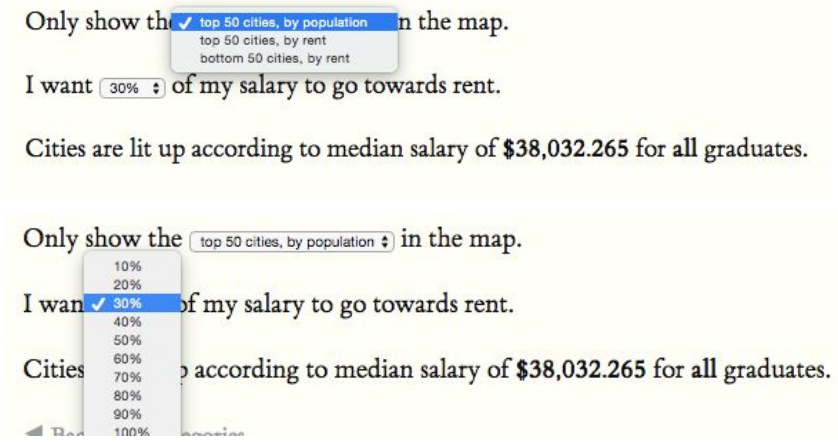
We decided to change the color of the data points to blue, as we want the entire project to feel cohesive. We also have a tooltip that appears when the user hovers on a metropolitan area, which displays the name of the metropolitan area as well as the median rent price.



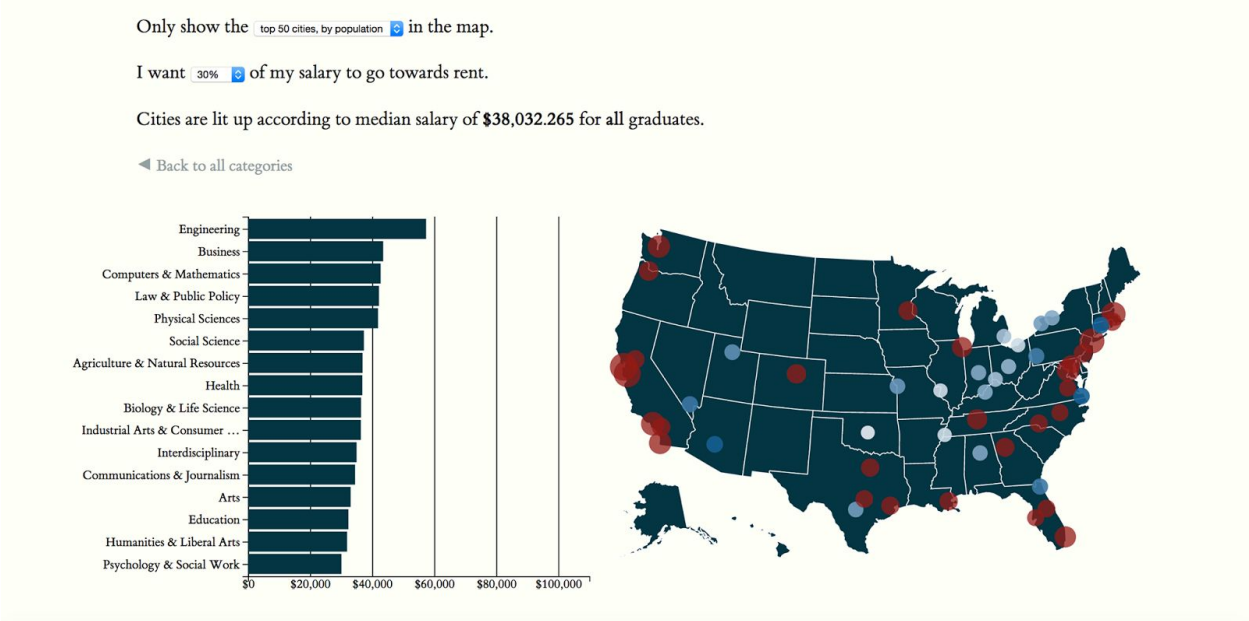
Connecting the Views

After this, we decided to start connecting our two views so that the user could interact with the map by clicking bars in the bar chart and changing parameters for our map.

We started by allowing the user to click on bars in the chart, and displaying the major or category and the salary. Then, we changed the drop down menu on the map, so that users could filter by top 50 cities by population, top 50 cities by rent, and bottom 50 cities by rent. This way, users could see the affordability of rent for different cities in the country. We also added a second drop down menu to allow the user to specify what percentage of a graduate's salary they would like to go to rent.

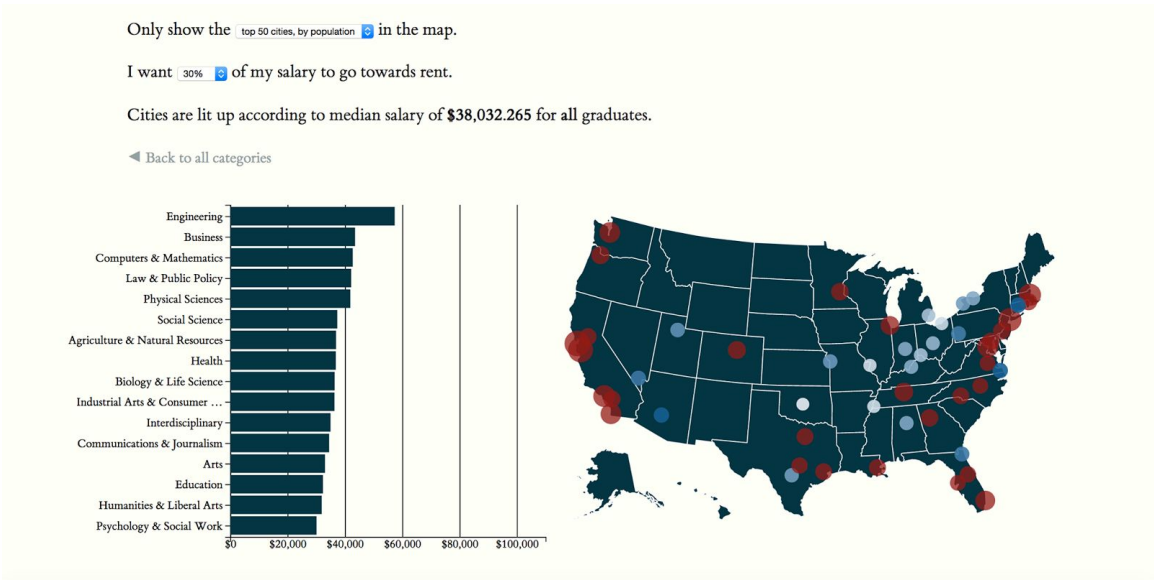


After this, we were able to change the colors and locations of the circles depending on what the user selected. When a user hovers above a city on the map, they are shown the name, median rent, and percentage of salary rent would be for that city. We took some time to find the right colors, and finally decided to split the data into two buckets. The first would be affordable cities, with a scale from most affordable to less affordable. The second would be unaffordable cities, where median rent was above the selected salary percentage. In the end our visualization is an interactive bar chart that allows users to see what cities are affordable for college graduates.

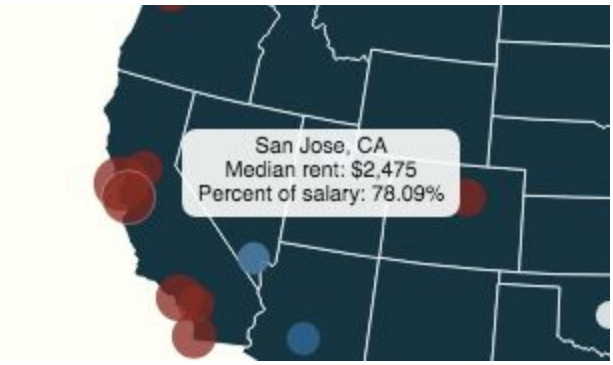


Implementation

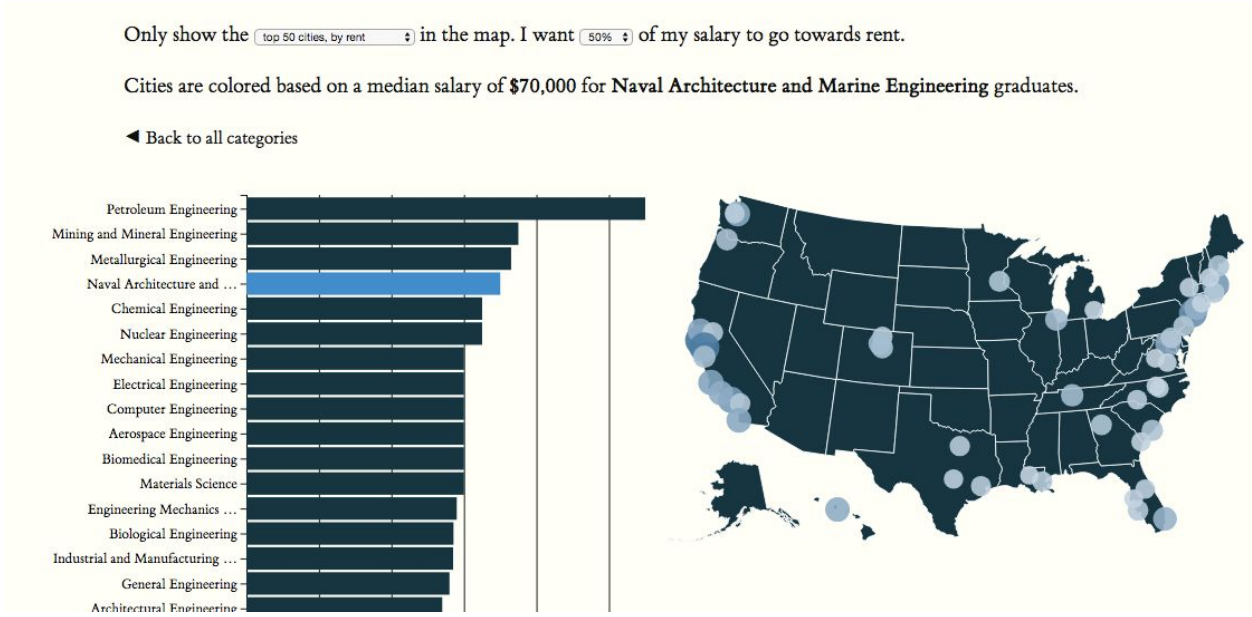
The intent of our visualization is to allow users to visually see what cities in the United States are affordable based on salary. Users are first shown the average salary for all graduates, with 30% of their salary going to rent, for the top 50 cities by population in the country.

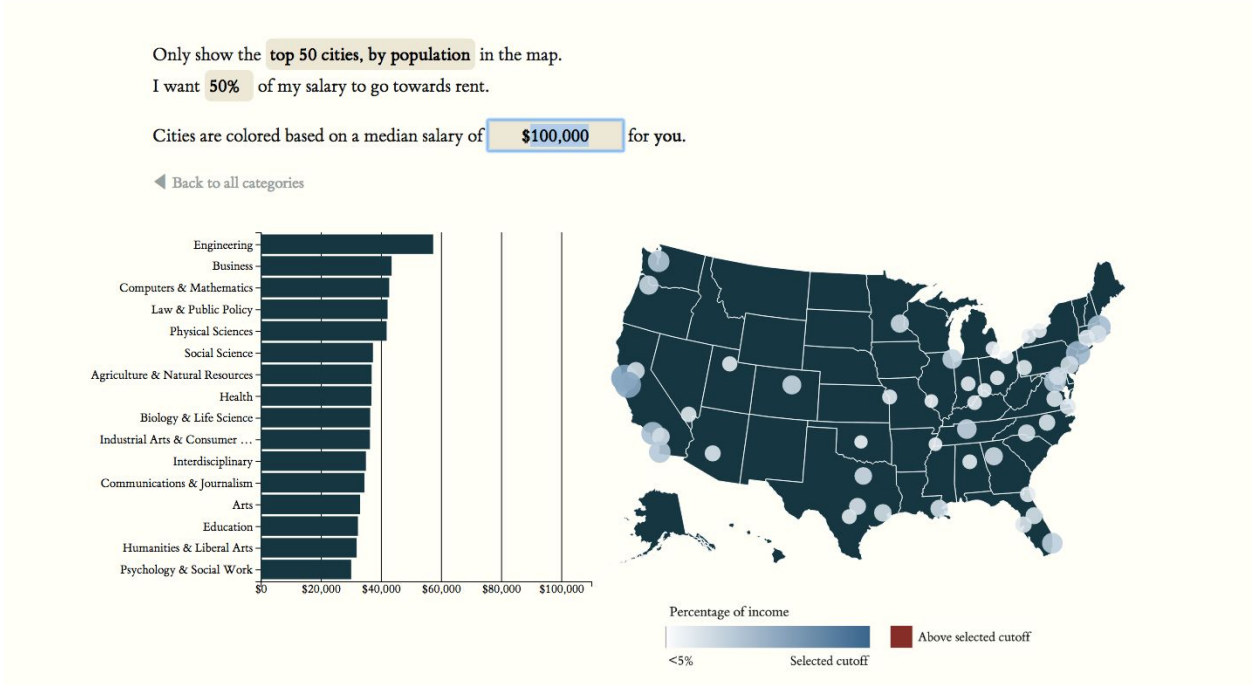


Users are able to hover over cities in the map, seeing the data on a per city basis. The tooltip allows the user to see the exact numbers of what is being visualized.



Users can then change the category or major that is being visualized by clicking on the bar chart, as well as change which cities are being viewed and what percentage of their salary they would like to contribute to rent. The user can also input their own salary in order to see what cities they would be able to afford on their salary.





We believe that our visualization allows the user to easily explore and examine the affordability of housing for college graduates.

Evaluation

By using our visualization, we found that for many college majors, in order for them to be able to afford living in a city, they would have to spend around 50% of their income on rent. Even then, they were often not able to afford living in the top most expensive cities, like San

Francisco, New York, and Boston. Only a very select few of majors were able to afford living in most of the most populated majors, most of them being in the Engineering category. We found that only Petroleum Engineering majors are able to afford living in all cities with 30% of their income going to rent. It was also interesting to note that for Computer Science majors, if they made the median salary they would have to spend 60% of their salary to afford to live in the San Francisco Bay Area. This is interesting, as Silicon Valley is considered a high paying area, where many Computer Science students eventually go to work. Conversely, it was interesting to see how the majors with the lowest median salary ($22k) would be able to afford living in some cities and spending 40% of their salary. We were able to see how major selection affected where a college graduate could live, and feel that rental prices and median salary for most majors do not scale together. Many areas are unaffordable for recent college graduates, which makes it very difficult for them to find work.

Our visualization works well to show users the average median salary for college majors and how affordable major metropolitan cities in the country would be for them. Users are able to see category averages, in order to see if they would like to pursue a degree in that category. Some improvements that our visualization could have would be to allow users to examine rental data for apartments that are not 1 bedroom 1 bath, and be able to factor in housemates. This would be a more realistic view of how college graduates live, as many are unable to afford to live by themselves.

Overall, we feel that our visualization answers the questions that we were thinking about, allowing users to see the relationship between a college graduate salary and rental prices.