

CoOp'd

Pranathi Alluri*

Rachael Cheung[†]

Eduardo Massud[‡]

Aritra Saharay[§]

Northeastern University

ABSTRACT

This visualization will provide Northeastern University's students with insightful statistics on the co-op program. Co-ops are the heart of every Husky, but Northeastern does not provide any tools to facilitate co-op exploration and transparency. With this tool, students will be able to better understand compensation distributions and see transparent ratings for co-ops. This allows students to not only explore co-ops, but also assist them in making an informed decision in this big career milestone.

<https://github.com/DS4200-S22/final-project-co-op-d>

1 INTRODUCTION

Co-ops are the number one reason many students choose Northeastern University. If not the main reason, at least a crucial decision factor. Northeastern currently offers students resources for co-op through NUWorks. NUWorks is a great platform to find co-ops but provides insufficient information for the applicant. A job listing gives you the job description, location, and the currency the salary will be paid in. The student has no insight into the actual experience, however. The aim of this project is to change that. An app has already been created for students to share their experiences anonymously. This visualization tool will use the app's data to generate intuitive and insightful visualizations on the whole co-op program. Students will be able to see company ratings, comments, compensation trends, locations stats and more. The combination of this tool and the app will be a powerful way for students to explore co-ops, as well as make more informed decisions.

2 RELATED WORK

For our references we'll focus on review visualization and opinion extraction, as our data is a collection of reviews about co-op and internship experiences from previous Northeastern students.

2.1 Summary Visualization

Before getting into the specific details, it is important for a visualization tool to clearly represent the big picture. Especially in our case of Co-op experiences, a summary of all the data should be readily available, in order for our users to initially understand what they can expect during the Co-op process. Once hired, more specific details regarding companies would be helpful, but someone opening our tool with no background of where they should apply to for co-op, expected pay scale, opportunities based on their year, and/or general experiences with employers, should be able to easily have these questions answered. Sarikaya et al [1] take this further, and emphasize the need for a good summary visualization along with the key factors that make up a summary visualization. More specifically, various ways to summarize data based on data type,

possible challenges one can run into, the impact of different summarization methods on overall analysis, and patterns in real-world uses of summary visualizations are mentioned in detail in the article, which are going to be helpful when creating our tool, as our first step will be coming up with the best method to group all the reviews in a clear way.

2.2 Visualizing Review of Local Businesses

Although, there aren't many visualization tools on summarizing internship experiences, there are many visualizations that include opinion extraction and presenting reviews from various customers. Specifically, Wang et al [6] created a tool that compares local businesses using online reviews. Similar to our app data, Wang and his team were left with multiple reviews that varied in style, length, accuracy, information, etc, and needed to find a way to compile them all together to create a holistic review of businesses. They mention in detail the entire process, including various type of visualizations they found to be effective, working from a broad picture to specifics. Furthermore, they also mention user feedback to their created visualizations. The extensive details on their process is helpful when creating our own tool, as we can take into account the shortcomings based on the user feedback and challenges the team ran into, and try to better our methods. Also, the general purpose of the review tool on business is the same as our tool, which hoped to compare companies, so we most likely will also be using similar visualizations to answer our questions,

2.3 Internship Search Stress

The need for our visualization tool is evident when looking at various studies and articles on the toll searching for internships and jobs can have on a students. In particular, a study done by Parent et al [4], illustrates how the entire internship application process was viewed negatively. Although the study was specific to psychologists, the experience can be considered similar for other majors as well. The general consensus was that the idea of not matching with a company, due to an imbalance of applicants to positions, was very intimidating and resulted in lots of stress. This is something we are hoping to overcome with our visualization tool. By being able to hear about the application process from previous students, and see the percentage of students who were actually able to get employed, along with a list of companies one might not have previously considered, hopefully the stress of finding a co-op is alleviated. Through our visualization tool we would like to make the process of finding a job enjoyable and exciting. Also, the study done by Mark and his team is also helpful in creating our tool itself, as the data analysis they completed is similar to what we will be doing. For example, in order to understand the application experience they sent out a survey with question on whether or not they found a job, demographics, and an open-ended prompt on how the application experience was. Then, in order to easily interpret the open-ended question, they broke the reviews into smaller chunks and grouped individual attributes, measuring the frequency of the most common terms. We will most likely be taking a on a similar approach, as the comments in our data are also open-ended, but share similar attributes.

*e-mail: alluri.p@northeastern.edu

[†]e-mail:cheung.ra@northeastern.edu

[‡]e-mail:massud.e@northeastern.edu

[§]e-mail:saharay.a@northeastern.edu

2.4 Visualizing Employee Expectations

As previously mentioned, the co-op process plays a large role in the college experience of a Northeastern student. After all, the basis of participating in a co-op is to apply the skills and knowledge gained in classes to one's desired career path. However, a major concern for many students is whether they do in fact have enough skills to meet the employers expectations. In a study done by Alwin et al [2], readiness and work experience of students was compared to employer expectations, to try getting a better answer to this question. The results of this study are helpful, when creating our visualization tool, as we would like to compare reviews on job experience to initial job descriptions posted by employers, and help co-op applicants understand what technical and transferable skills are needed to be successful. The article would act as a guide in this process, as there is a breakdown on data analysis and example visualizations which compare student's experience and employer expectation on technical and transferable skills, which is an attribute we would like to consider in our tool.

2.5 Visualizing Employee Reviews for Companies

Despite there not being many visualization tools that compare various job position and companies, the one created by Andres [5] is a great starting off point. The created tool, analyzes reviews left by employees for large tech companies. This is relevant to our project as it contains examples of possible visualizations one could make with employee reviews to compare different companies. Also there are many tips on how to clean data and analyze data, such as taking into account who is leaving the comments and what companies have the most comments. These approaches can be directly applies to our data, as we are also working with employee reviews and multiple companies. Overall, the purpose of the visualization tool created by Andres was to give users an idea of what major tech company aligns with their needs, which is what we're also trying to achieve, just with a larger list of employers.

3 USE CASE

A Northeastern University student just received an offer letter to work at MadeUpCompany, Inc for the upcoming fall semester. That student has no idea if that is a good opportunity or not and cannot find any useful Northeastern resources to help them out. The student is smart enough to look through LinkedIn. Every post on LinkedIn is an amazing review about an awesome experience. But then again, who posts bad reviews of their co-ops on LinkedIn. The student is getting desperate, but he finds out about this tool! The student signs up and reads honest reviews of the co-op. Now they are more well informed and prepared to make such a difficult decision. Another scenario would be to analyze compensation across colleges and companies. This could help the student decide whether the offered pay is fair or not.

4 DATA

Our dataset [3] contains co-op reviews exclusive to Northeastern University students. The data is straight from the mobile application CoOp'd, which is designed to bring more transparency to the Northeastern co-op program. Although exclusive to NEU, the program is not affiliated with the university, and therefore there was no real incentive (such as extra credit) for users to send in their data other than the desire to boost transparency within the community. The app has been advertised through the NEU subreddit, which could bring in some biases to the user base. We assume that due to this focus on the Reddit community, the co-op reviews might be leaning more towards Computer Science and Engineering, but not so much to some of the other areas of study. Besides that, the app uses a single attribute for rating. Although this makes it simpler for users, it is not holistic. For example, a user might take off a couple of stars from a

co-op simply because the commute was long. A broken-down rating system could have eliminated these issues.

5 TASK ANALYSIS

See task analysis in "Tasks" Table 1.

- The primary consumer of the visualization will be Northeastern students who are currently applying for co-ops
- The visualization will be developed mainly for discovery purposes. This is due to the fact that most students applying for co-ops don't have much information about average salaries, ratings, past experiences, etc, even if they do have access to the job listings offered at Northeastern. Thus, our tool will be used to help students find this new knowledge not previously known.

6 IMPLEMENTATION PLAN AND PRELIMINARY WORK

We plan to make the co-op ratings the main metric for visualization. From there we would dig deeper into how that's correlated with compensation and colleges. We would also like to relate compensation with location and nTh co-op. Another interesting visualization we'd like to implement is a road map showing where students go after their first co-op. For example, a roadmap leaving Wayfair, showing the companies students have gone to. For the comments we would likely make a word cloud or similar visualization. Since we also have information on when the co-op was completed, we can make a tool to see how the pay has been increasing throughout the semesters. Students always wonder if Fall co-ops are superior or inferior to Spring semester, so by combining some attributes we can see how the ratings relate to the term.

6.1 Revised Implementation Plan

- Visual Encodings
 - Bar chart (Stacked, Group Bar Chart)
 - Scatter plot - with company logos
 - Nice to have
 - * Heatmap / map
 - * Word Cloud
- Interactions and interactive components
 - brush/link
 - filtering (using drop down)
 - highlighting
 - selection
 - Nice to have
 - * tooltip
 - * have bar graphs change to remove all irrelevant information based on filtering and highlight relevant bars
 - * have new bar graphs (for rating and salary) and reviews appear on the side of the screen when a company is clicked, and not be there when there is no company clicked.

We will be implementing this visualization by using html, CSS, and d3.js connected to a local python server in order to work with the data easily. We will be using the library, react.js, to assist in building user interfaces and allowing for login capabilities. In terms of server-side components the visualization tool and data will only be loaded and displayed if correct username and passwords are provided to the login.

Task ID #	Domain Task	Analyze Task (high-level)	Search Task (mid-level)	Analysis Task (low-level, "query")
1	As an undecided major, I want to know what kind of co-ops different majors or schools at Northeastern get.	Consume Discover	Explore	Summarize
2	As first-time co-op searcher, I want to know the average salary per major and location.	Consume Discover	Explore	Compare
3	As a 2nd or 3rd time searcher, I want to know the average rating of a group of companies in a location.	Consume Discover	Explore	Compare
4	As a first-year student, I want to become familiar with the companies that NEU students tend to go to, and get a general feeling of the quality of experience (ratings) to start building a list of target companies	Consumer Discover	Explore	Summarize
5	As a student who just received an offer, I want to evaluate it based on peers' experiences and stats. This would include reading past reviews of that company and seeing the ratings as well as comparing compensation based on college and location	Consume Present	Lookup	Identify

Table 1: Tasks

7 VISUALIZATION DESIGN

Our final visualization tool is composed of two main sections—the first section is company-centric with three different visualizations. The main visualization is a zoomable scatterplot that plots companies on axes for their average pay and average rating. Two smaller distribution bar graphs show the distribution for the pay and rating values for either all the companies, or the selected companies on the scatterplot.

The second section of the visualization contains two other graphs. The data shown on these graphs are dependent on a drop-down menu within the container. The first graph is a stacked bar graph. Each bar represents a location, and each bar is sectioned into areas that represent the nTh-Coop. The width of these areas is dependent on the feature selected in the drop-down menu. The second graph is a grouped bar graph, each section of bars representing the nTh-Coop, and each bar representing a college. The height of the bars is dependent on the feature selected in the drop-down menu.

Interactions that the visualization can facilitate including selecting companies on the scatterplot. Doing so will change the pay and rating distribution graphs accordingly to display data for the selected company. It will also change the data shown in the graphs second section of the visualization, such that only the company data will be shown. Another interaction that is facilitated by the visualization is the selection of a bar. Doing so will highlight the corresponding companies that contributed to that specific section of data.

The placement of the graphs within the visualization allow for intuitive use by users with some visualization experience. The grouping of the graphs and the discoverability of the filtration options all contribute to the easy understanding and usage of the visualization overall.

Overall, the visualization allows for filtration and exploration of co-op data, which can help our target users. The visualization will allow users to compare different average ratings and compensation by college, nTh-Coop and by location. It also helps users discover different companies that may suit their goals and explore different co-ops that their peers have done in the past to evaluate their options.

REFERENCES

- [1] D. A. S. A. Sarikaya, M. Gleicher. Design factors for summary visualization in visual analytics. *Computer Graphics Forum*, 37(3):145–156, July 2018. doi: 10.1111/cgf.13408
- [2] T. M. Ajisha Alwin, Yostina Geleta. Toward conceptualizing student outcomes in undergraduate field programs and employer expectations for field positions. *The Bulletin of the Ecological Society of America*, 102(2):18–20, Jan. 2021. doi: 10.1002/bes2.1820
- [3] E. Massud. Co-op. CoOp'd, 2020.
- [4] M. W. E. A. J. L. C. Mike C. Parent, Tyler C. Bradstreet. “The worst experience of my life”: The internship crisis and its impact on students. *Journal of Clinical Psychology*, 72(7):163–169, Mar. 2016. doi: 10.1002/jclp.22290
- [5] A. Vourakis. Analyzing employee reviews: Google vs amazon vs apple vs microsoft. Towards Data Science, April 2019.
- [6] C. S. Y. W. X. Z. S. F. H. Q. Yong Wang, Hammad Haleem. Towards easy comparison of local businesses using online reviews. *Computer Graphics Forum*, 37(3):63–74, July 2018. doi: 10.1111/cgf.13401

A GROUP CHARTER

A.1 Group Purpose

Our group was formed to collaborate on building a visualization tool for the final project of the course DS 4200, Information Visualization and Presentation at Northeastern University. Our purpose is to build a visualization tool that can be used by Northeastern students to explore co-op data and to create a demonstration and paper documenting our research and process.

A.2 Group Goals

- Project Goals: to create a visualization tool and create a demonstration and paper documenting our research.
- Process Goals: to be able to distribute work evenly amongst members and maintain effective communication throughout the process, and avoid last-minute work
- Quality Goals: produce our best quality work possible, adhering to the project guidelines
- Our group members are willing to commit 7-8 hours a week of work for this project and are aiming for an A grade for this course.

A.3 Group Member Roles/Responsibilities

- Pranathi Alluri: Meeting coordinator, Documentation Coordinator; python
- Rachael Cheung: Time facilitator; HTML/CSS/JS
- Eduardo Massud: Information manager, Main Point of Contact
- Aritra Saharay: Web Development; JS/TS/HTML/CSS

A.4 Ground Rules

- Our group will primarily meet online through Zoom. We will have weekly meetings on Sunday evenings. However, any updates, information, and questions will be put in the group chat.
- Ground Rules:
 - Complete your share of work
 - Communicate in the chat about updates on work
 - Give advance notice if unable to go attend the meeting
- Discussions will be facilitated through questions and decisions will be made after considering the opinions and views of the members
- Dissenting views amongst members will be taken care of through discussion and careful consideration of compromises before making a final decision
- We will hold each other accountable for living by these rules through check-ins in the group chat and check-in meetings or work session meetings
- We expect to commit up to 8 hours of work for the project each week. Some other expectations include active participation and communication, and completion of responsibilities.

A.5 Potential Barriers and Coping Strategies

- Conflicting schedules and maintaining communication may be potential barriers.
- If such barriers materialize, then our group will hold a meeting to discuss compromises that can be made, or ways that we can meet briefly in order to divide up work and stay updated on each other's progress.
- Group dynamic issues from our previous experiences include miscommunication or lack of communication, not fulfilling their responsibilities, and lack of active participation in group discussions. In the event that these issues come up again, we will try to communicate the presence of such issues and try to resolve it amongst ourselves first before escalating the situation to teaching staff.

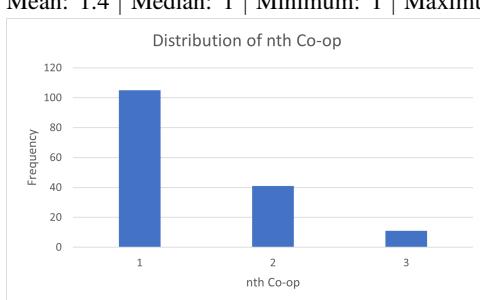
A.6 Revisiting Group Charter

- We have you all been abiding by your agreed-upon guidelines.
- We feel comfortable with the group roles.
- We seem to have trouble finding meeting times each week, so we plan to create at least one fixed time a week to meet with each other.
- One positive thing they have seen other group members contribute to the project:
 - Ari: Eduardo has done a good job with preparing the dataset for the project. Pranathi has been effective in organizing meetings and making sure all the tasks are finished appropriately on time. Rachael has been a big help with her knowledge of latex and for preparing the digital sketches of the final visualization tool.
 - Eduardo: the team works like a big united family. We're all communicative and do our work in time. We spend most of the time discussing our ideas and sharing concerns. It really feels like every member wants to achieve the best possible result.
 - Rachael: Pranathi actively contributes ideas during group meetings and also actively responds to others' ideas. Eduardo facilitates discussions and helps include everyone in group discussions. Ari participates in group meetings and fulfills responsibilities.
 - Pranathi: Rachael is always hands-on and willing to step up to the plate when things need to get done or if someone is busy that week and needs a helping hand. She checks-in on us and makes sure we're all doing okay, and is the first to answer any questions that come up in our messages. Eduardo is super passionate about this project and very knowledgeable regarding the data set. He is always open to listening to ideas, and contributes to productive discussions about what we would like to create. Ari shows up to all the group meetings, and makes sure to do the work he is told.

B DATA EXPLORATION

B.1 Data Review

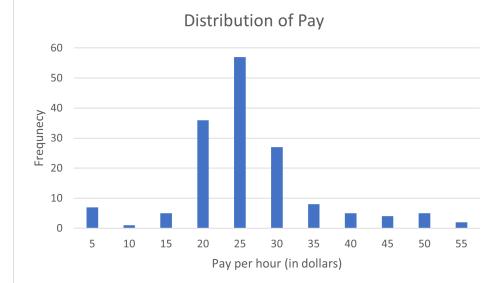
- College
 - Categorical
 - Levels: Khoury College of Computer Science; College of Engineering; College of Social Sciences and Humanities; College of Arts, Media, and Design; D'Amore-McKim School of Business; Bouvé College of Health Sciences; School of Law, College of Professional Studies; College of Science
- Comments
 - Qualitative
 - Free form text
- Company
 - Categorical
 - Levels: Many - can be any company name that students work at
- Country
 - Categorical
 - Levels: Many - can be any country in which a student worked in
- Created By
 - Levels: Many - auto-generated ids
- International Student
 - Categorical
 - Levels: True; False
- Anonymous
 - Categorical
 - Levels: True; False
- Is Co-op?
 - Categorical
 - Levels: True (is co-op); False (is an internship)
- Drug Test
 - Categorical
 - Levels: True (was drug tested); False (not drug tested)
- nth Co-op
 - Ordinal
 - Levels: {1, 2, 3}
 - Mean: 1.4 | Median: 1 | Minimum: 1 | Maximum: 3



- The distribution is skewed right, with the highest frequency being for 1. Indicating that most reviews are from student's first co-op, and the least are from a student's third co-op.

- Pay

- Quantitative, Continuous
- Range: (\$) 0-52 (per hour) | Mean: 24.45 | Median: 24



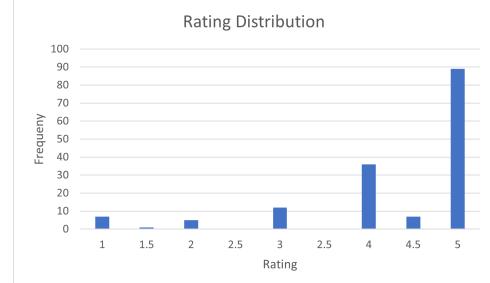
- The distribution is relatively normal, being slightly skewed to the right. This indicates that most students had an average pay between \$20 to \$30.

- Position

- Categorical
- Levels: Many - can be any position a student takes in a company

- Rating

- Ordinal
- Range: 1-5, steps of 0.5
- Average: 4.3 | Median : 5



- The distribution is skewed to the left, indicating that most people had good experiences and gave high ratings for their co-op experiences. The highest frequency was a rating of 5.

- State

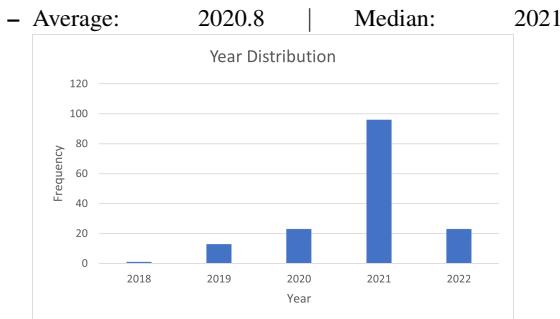
- Categorical
- Levels: Many

- Term

- Categorical
- Levels: {Spring; Fall; Summer}

- Year

- Ordinal
- Levels: {2018, 2019, 2020, 2021, 2022}



- The distribution is skewed to the left, with the highest frequency of co-ops being in 2021.
- Identify issues with your data.

- Bias - Majority Khoury College: This dataset is a small sample of the student population and not a census, so it only represents a small proportion of students who have gone on co-op and may not be wholly representative of the entire student body. Specifically, this app was posted on Reddit, which caters to a majority of Khoury College of Computer Science students, resulting in much of the data to be from computer science students (31%), while there is no co-op data for Bouve Health Science, School of law, and College of Professional Studies.

* Solution: We are sending the link of the app out to more people, in order to collect more data and have a more diverse sample. More efforts are being made to spread it through LinkedIn, Instagram, and other platforms.

- Bias: Students with higher compensation are more likely to publicly self-report their hourly wage. This skews the mean wage of the dataset and makes it much higher than the mean wages reported by the school, who have access to a much wider gamut of data.

* Solution: We make sure students know they can report anonymously.

- Quality issues - double majors: The college doesn't account for double and combined majors.

* Solution: An issue we will take into account when creating the graphs, however this is a data collection issue, that needs to be changed through the app, which is beyond the scope of this project.

- Quality issues - Covid Co'ops: Only 8.9% of co-op reviews were before Covid. Covid made it harder to find co-ops, and it also made the experiences very different, such as people working remotely. This impacts the reviews and experiences that are left by the users.

* Solution: With time people will start adding more post-Covid co-ops, as the job environment shifts back to "normal", and this could turn into an opportunity for us to visualize changes in data during Covid and after Covid.

B.2 Insights

Some trends we found were that a majority of the students that reviewed had left generally good reviews of their experiences. Also, we found that the co-op pay increases with experience, such that subsequent co-ops tend to pay higher. In the same realm, we realized that students who had received lower wages also tended to want to keep their wages hidden in their review. Furthermore, one surprising observation we made during the exploration was that international

students had a lower average pay than non-international students. One possible reason for this is that typically international students have a smaller pool of employers to choose from as some companies will only hire citizens, causing less competition which would normally drive up wages.

In general, there did not appear to be any errors or issues with the data being messy or confusing. The issues were mainly caused by the potential problem of biases that come from data collection, as mentioned earlier. This should be taken into account when discussing another perplexing insight we garnered, which is that students enrolled in the College of Engineering were far more likely to be drug tested than students enrolled in any other college. One possible cause for this could be that engineering co-ops may involve a larger share of employers that are involved in government contracts, which would require drug tests for employees. Another reason may be because we have more responses from College of Engineering than some of the others schools.

B.3 Screenshots

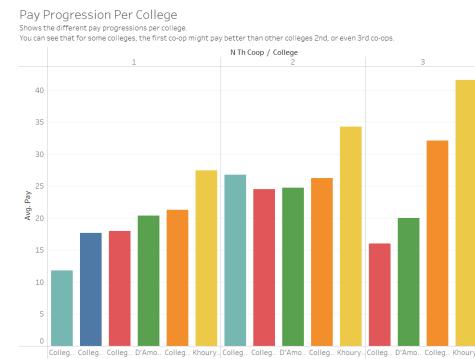


Figure 1: This bar graph shows the different pay progressions per college within Northeastern. It displays items, and uses area as marks, and color and area as channels. The area channels are used on each bar to show the quantity of the average pay for each college, and the color is used to distinguish between each college to show which category the bar belongs to. The visualization shows the trend that generally, the pay increases with each subsequent co-op.

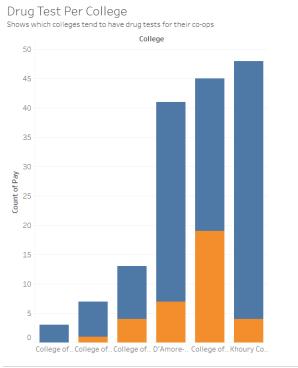


Figure 2: This bar graph shows us the different rates at which students in Northeastern's different colleges are likely to get drug tested in their co-ops. It displays items, and uses area as marks, and color and area as channels. The area channels are used on each bar to show the quantity of coops with drug test data for each college, and the color is used to distinguish the proportions of co-op experiences in which drug tests did or did not occur. The main trend we can observe in this bar graph is that students enrolled in the College of Engineering are roughly three times more likely to be drug tested than students enrolled in any other college.

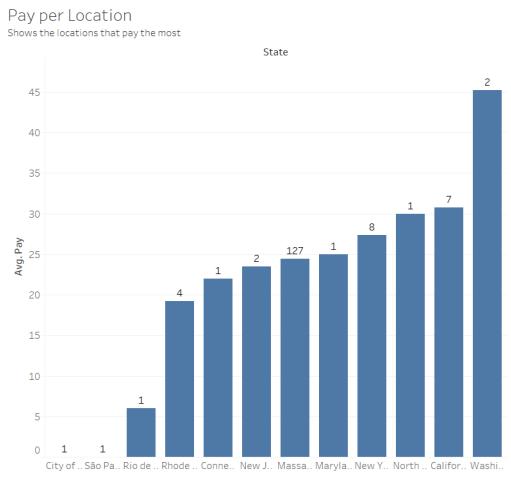


Figure 3: This bar graph shows the different average pay rates based on the location of where Northeastern students worked . It displays items, and uses lines as marks, and length and position as channels. The vertical position is used on each bar to show the quantity of the average pay, while the horizontal position illustrates the location of where co-ops have taken in place. The visualization shows the trend that generally, the average pay is above the national minimum wage. Furthermore, larger more developed states and cities, such as New York, California, and Boston had higher pay, which is understandable as large companies are moving to these areas. Furthermore, the visualization shows that international locations had the lowest average pay.

B.4 Data Snippet

Figure 4: First few data items from cleaned and revised data.

C INTERVIEW

C.1 End User Personas

- User 1: A first-year undecided major student at Northeastern that is not looking for co-ops, and has had minimal experience with using visualizations; The student is trying to look into fields that they would be interested in, and is looking to find out more about the co-op experience and the nature of work in certain fields; The visualization tool could provide the student with insights on a wide range of fields; This task is currently performed through word-of-mouth interactions and internet search, so the tool could help perform the task better by gathering all such data in one place and offering a comprehensive understanding of the information. Especially data on co-ops grouped by major and specific colleges at Northeastern, to help in making a decision about majors the student might wish to pursue.
- User 2: A Northeastern student looking for their first co-op that has basic understanding of visualizations and has had exposure to using visualizations; The visualization tool can be used to help the student by providing summarized insight on roles that the student may be interested in, or act as a foundation on where to start when applying for jobs. Without the visualization tool, the task is currently performed through word-of-mouth interactions, advice from a Co-op advisor, NU Work listings, and social media posts; The visualization tool will help perform the task more easily by providing a comprehensive overview that provides general statistics on average salary, locations of past Co-ops, past companies that have hired Northeastern students, etc. and allows for more in-depth interactions with co-op reviews. This information goes very far when applying for internships and co-ops for the first time.
- User 3: A Northeastern student who has already finished one or two co-ops and is looking for their final one, and is also familiar with visualizations and has used them before. The visualization tool would assist this user's co-op experience by providing relevant insight about the company they were hired for, whether it be previous employee's experience, drug testing, salary, or other positions offered. As of right now, this task is mainly achieved through word-of-mouth, google search, company website, Reddit threads, and suggestions from a co-op advisor. Our visualization tool will increase the ease with which this task can be completed, by providing a comprehensive overview of company-specific co-op statistics and allowing for more in-depth interactions with co-op reviews.

C.2 Interview Script

Thank you for coming to this interview. For a class project, we are working on creating a visualization tool for Northeastern students that can be used to gain insight on other students' co-op experiences with different companies. We are interested in finding out how you would use it and what you think of it! For this purpose, I have a couple of questions to ask you. Please feel free to ask for further clarification at any time.

- Are you currently searching for co-ops?
 - If not, why are you interested in looking into co-op reviews?
 - If so, how do you currently go about finding reviews and insights on companies that you are considering for co-op?
 - What are some obstacles that you have faced while looking for information?

– Is this your first time searching for a co-op or have you been on Co-op before?

- * If this is your first time, where and how did you start your co-op search?
- * If not, what is something you wish you would've known during your first search?
- * What are you going to differently searching for co-op in the future?

• What kind of information regarding co-ops are you most interested in seeking? What kind of information do you find the most helpful?

• What are some features in a visualization tool that would make this information search easier for you?

- Have you used a similar tool in the past?
- What features did you find easy to follow?
- What features could you have done without?
- What are some core tasks this tool needs to support to be successful?

• Is there anything else you would like to share?

That brings us to the end of our interview today. This has been a very fruitful interview. Thank you very much.

C.3 Interview Notes

- User 1: Northeastern student undecided major
 - No I am not currently searching for a co-op
 - The co-op offered at Northeastern is one of the main reasons I decided to come here, which is why I would love to learn more about it. I think more than co-op reviews I'm interested in seeing the overall summaries of past Northeastern student's experiences.
 - Followup: When you say overall summaries, do you mean visualizations summarizing companies students have worked at, salaries, majors, or do you have something specific in mind?
 - Yeah, something along those lines. I am an undecided major, so I think being able to see the co-ops grouped by major or specific colleges would be good to have. Just so I can get a general idea of what students in each major end up doing, and use that to help declare my major. I also want to study abroad or do a co-op abroad, so maybe seeing if students have done that in the past, and their experiences would be helpful.
 - Follow up: Of course, those are definitely features we will have. You said co-ops grouped on specific colleges or majors. Did you want to see average salaries of colleges, average ratings, being able to read reviews filtered on colleges or like companies that offer positions for specific colleges?
 - I think all of those features would be helpful, so I can see what my future experiences would entail if I picked that major
 - That makes sense. You kind of touched on this already, but is there any other information regarding co-ops you are interested in seeking? What kind of information do you find the most helpful? I know you said co-ops grouped by specific colleges, but is there anything else?

- Definitely everything I have mentioned already. Also, salaries, reviews, and majors based on where people have had co-ops would be nice to see, so I can get an idea of where I could head to outside of Boston in the future.
 - In terms of features, just like overall graphs would be good to see, and like you said filtering based on location or major would be nice as well. Just so I get all the information on one page, and don't have to click through multiple things.
 - Followup: Yeah, so for example a graph that has all the average salary for co-ops based on colleges and then you can filter for number of jobs, specific companies, and reviews. Do you think something like that would be helpful? If so, what would you want the first graphic to be about salaries or number of jobs?
 - I think salaries are good, as a college student that is something that is important.
 - Followup: Have you had experience with similar visualization tools in the past? Or are there specific graphs and features you find to stand out.
 - I haven't had much experience using visualization tools, but I think simple is better. Just like a bar graph or scatter plot rather than something too complex that is hard to read.
- User 2: Northeastern student just starting co-op search
 - I am currently searching for a co-op
 - I go on NUWorks to find listings to apply, but there aren't any reviews on there, so I'm not sure about that. What I hear about companies is usually from peers, my co-op advisor, and when I attend events put on by my college.
 - An obstacle I have faced is figuring out where to start my search. Like what companies offer the most people from my major, and also NUWorks doesn't list salaries for all companies so that would be helpful.
 - This is my first time searching for a co-op.
 - I started by meeting with my academic advisor and just filtering based on locations and positions through NU-Works, but that was definitely very overwhelming.
 - Followup: Yeah of course, the first time job search can be intimidating. What are some resources or tools that you think would've made it easier for you to start your search?
 - I think just knowing what companies are known to hire Northeastern students, where past Northeastern students have worked, like in-state or out-of-state, and previous user's experience who've had my major. I think this is important information and we don't have access to this information, unless we ask them during co-op events
 - Follow up: Yeah definitely, having previous experiences and stats can be helpful. Are there features you would like to see, or past visualization tools that you've enjoyed using in the past.
 - I think the first things that comes to mind, is when we were searching for colleges there was a tool I used where I can put in my major, school I wanted to go to, GPA, and like test scores and it would tell me where I fell according to people in the past. I think something like that would be helpful. Obviously, GPA isn't available for everyone but something similar, with our major and the company we're applying for.
 - User 3: Northeastern student who has been offered a position
 - Some features I need is definitely showing the average salaries per major, and also location, so I know what to expect when interviewing and asking for salary. Also, I think specific company salaries would also be good, if someone wants, but can be overwhelming if it's the first thing that is seen. Same thing for reviews.

- User 3: Northeastern student who has been offered a position
 - I am currently searching for a co-op, and have actually been offered a position from Company X
 - I go on NUWorks, LinkedIn, and just my own networks to find positions to apply. I think just talking to my friends is how I get information on each company.
 - An obstacle I've faced is finding reviews from previous Northeastern students, drug testing info, and salary info, so I know what offers to accept.
 - I think I would've liked to know average salaries based on college, so I can see where I fell on the scale. And just like what companies are hiring would've been nice to see.
 - In the future, I have moved away from NUWorks and have used my personal networks. I know this is not accessible for everyone, so hopefully a tool like this can help people in that process.
 - Followup: You seem very experienced in the search. Would you still find it helpful to use the tool as someone who has already been on a co-op?
 - Of course! I think information on companies and reviews from previous students is important no matter what stage of the application process you are in.
 - I am most interested in seeking information on the companies that have offered me a position. Like the average ratings, reviews, and drug testing.]
 - I have not used a similar tool in the past, but I think having summary graphs to begin with showing average salaries and where people are doing their third co-op would be nice. Like filtering based on co-op number, is helpful for someone like me. And then being able to search a company name and seeing their drug testing policies, reviews, and ratings is ideal. If possible, maybe even having a feature that compare two company ratings and reviews side-by-side, to make it easier for which offers I should accept s reject.

C.4 Interview Results

Q: Are you currently searching for co-ops? If not, why are you interested in looking into co-op reviews? If so, how do you currently go about finding reviews and insights on companies that you are considering for co-op? and what are some obstacles that you have faced while looking for information?

Answers:-

User 1: I am not yet looking for a co-op.

User 2: Yes, I am looking for my first co-op.

User 3: I have already done 1 (or 2) co-ops so I am searching for my final one.

General answer: All users are Northeastern students interested in finding a meaningful co-op, but are in different stages of the process. Currently, students go about finding reviews and insights on companies based on postings on NUworks, co-op advisors, and networking. A common theme is that besides word-of-mouth and internet searches, which are not always reliable, there is currently no solid place for students to start their search and find generalizations specific to Northeastern, which is difficult.

Q: What kind of information regarding co-ops are you most interested in seeking? What kind of information do you find the most helpful?

Answers:-

User 1: As an undecided major, I want to know what kind of co-ops different majors or schools at Northeastern get.

User 2: I want to know what companies are known to hire Northeastern students, where past Northeastern students have worked, like in-state or out-of-state, and previous user's experience from my major.

User 3: As this is my second (or third) time searching for a co-op, I want to be more selective in my search. I want to be able focus my attention on companies that have been rated highly by the students they have hired in the past.

General Answer: Information based on average salaries per major/college is something that is very common. Also, where students have had co-ops in the past, and the salary for various locations, is good to know. Furthermore, being able to filter and find more details for specific companies and colleges would be a nice addition to have, for those who are a little further along in their search.

Q: What are some features that would make this information search easier for you?

Answers:-

User 1: I would like to see graphs summarizing information that I can also filter based on location and/or major.

User 2: I want to be able to roughly estimate my chances of securing a co-op at any specific company I have in mind, by seeing how many students any company has hired from Northeastern that are also in my major.

User 3: I would like to see graphs summarizing information such as compensation and location that I can also filter based on whether the student was hired for their second or third co-op, and more specific details such as likelihood of drug tests and reviews/ratings.

General Answer: Summary tools that are able to show average salaries per location, college, co-op number, etc. seem to be very common. Make these easy to read, and have the values easily found. Beyond that, have features where users can click to find more detailed info about specific companies, can filter and compare based on specifics they are personally looking for, and maybe even compare multiple companies and positions.

D DESIGN SKETCHES

D.1 Individual Sketching

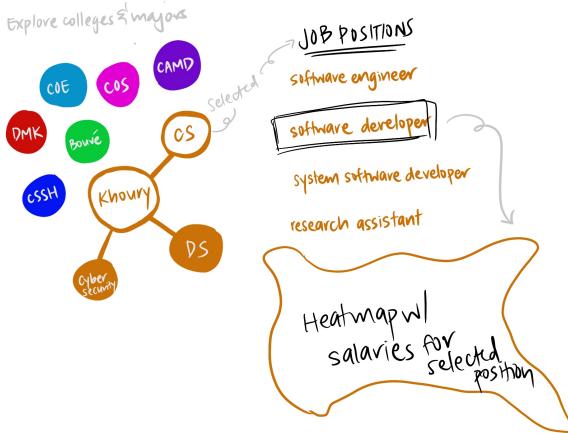


Figure 5: [FAVORITE] Rachael: The network of items/nodes and uses points and lines as marks, and color as a channel. The channel is used to differentiate the college, and the points and lines show connections between groups. The heat map uses areas as marks and position (both vertical and horizontal) and color as channels. The colors are used to show the scale of salary and position is used to show the location of a specific item. This visualization covers tasks 1 and 2. It displays the types of positions a certain major may go into, and also the average salary depending on location and position.

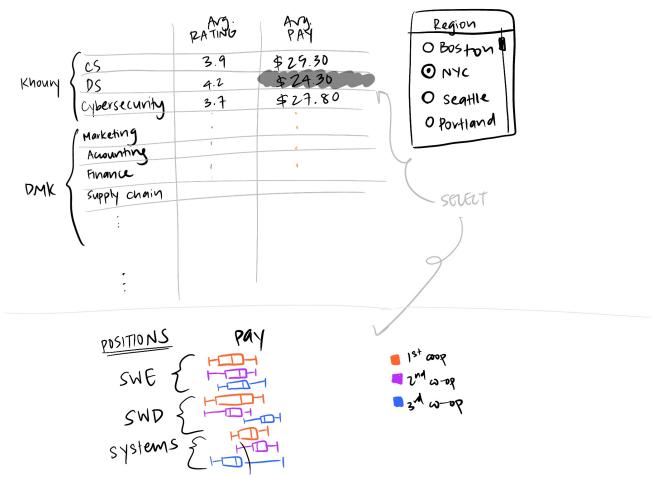


Figure 6: Rachael: Lines and areas are used as marks, and position and color are used as channels. The marks use the channels to communicate the distribution of pay and average rating across different job positions for each job position. This visualization is able to help with tasks 1, 2, and 5. It shows the positions that certain positions that specific majors get, and allows for comparison between positions.

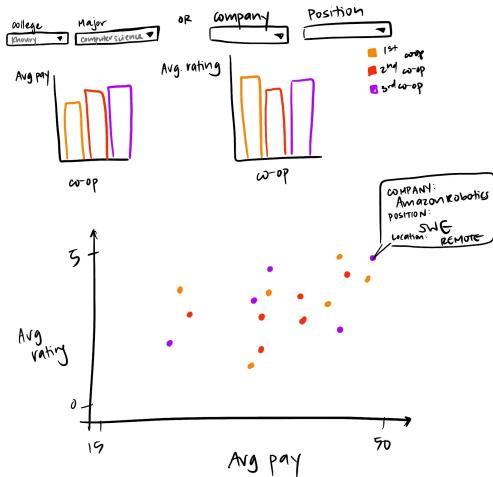
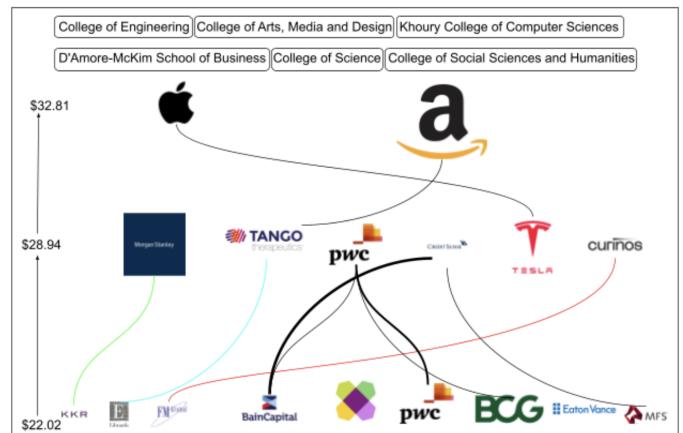
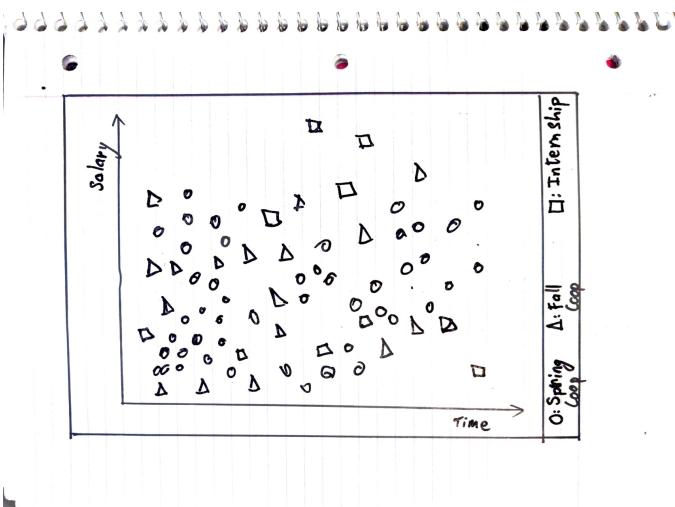
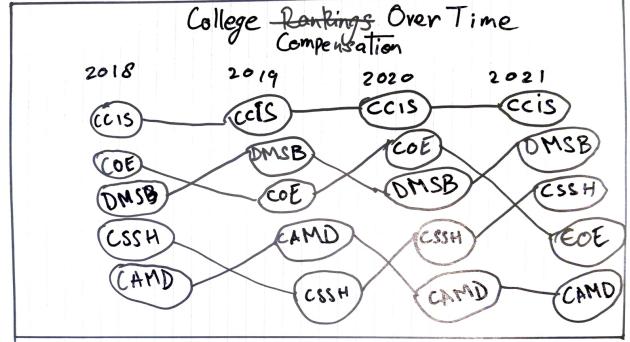
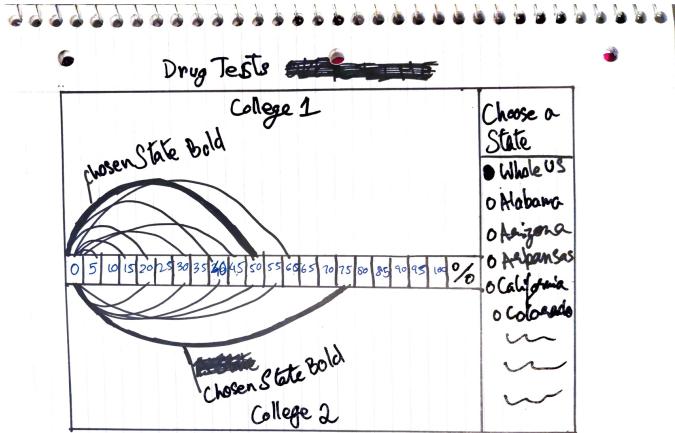


Figure 7: [FAVORITE] Rachael: Areas and points are used as marks, and position, channel and area are used as channels. The marks and channels are able to be used to compare and communicate data efficiently and concisely, and presents the information so that it can be understood intuitively. This visualization covers tasks 4 and 5. This is because it allows the user to get an idea of the ratings and salaries of companies and an overview of how it would compare to other companies.



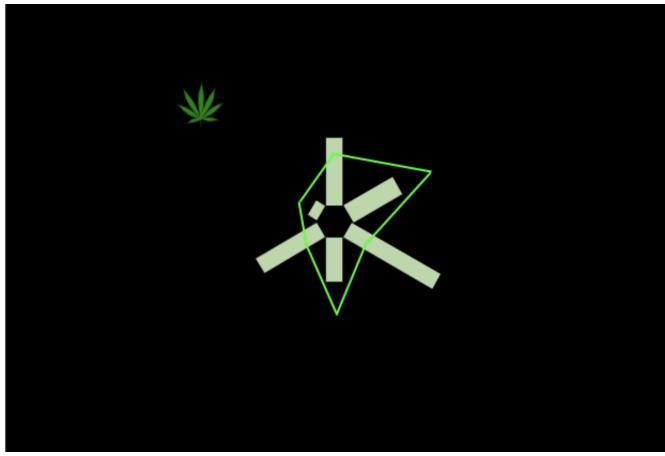


Figure 12: Eduardo: This radial graph would have a bar for each college with the respective drug test percentage. The green line would show the total drug tests. When students click on a bar, the visualization could bring up the companies that have been known for drug testing within that college.

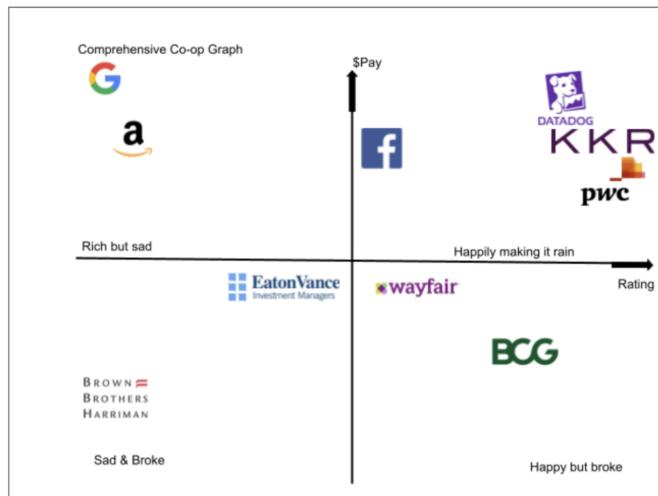


Figure 13: [FAVORITE] Eduardo: This visualization is a scatter of the aggregated company stats. With this students can easily explore the companies and see if they are above or below average when it comes to both rating and compensation.

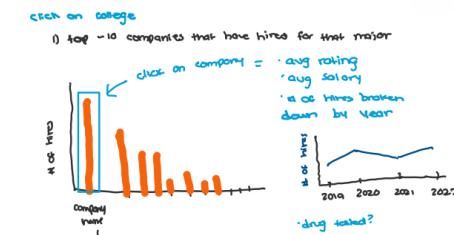


Figure 14: [FAVORITE] Pranathi : A majority of the graphs in this view are bar graph and scatter plots, which use lines as their marks, and x and y position for their channel, along with color to differentiate between various groups. This is because they are easily understandable by everyone. The main graph allows for the user to clearly see the average pay for specific colleges, broken down by co-op number. This covers part of task 2, where a co-op searcher can know average salary based on location. In this graph the co-op number and college are clickable, taking users to visualizations that show location people have co-op for their n-th co-op, along with average rating and salary for the company, which further illustrates task 2 and also task 3, showing average salaries and ratings per location. Clicking on a college, shows top companies that hire and the number of hires, which addresses task 1 and 4, of exploring co-ops different majors get and companies that hire.



Figure 15: Pranathi : This visualization focuses on showing companies based on location, and uses x and y position, size, and area to do so. A user can click on places on the map and zoom in, to see the places most co-ops have been, which addresses task abstraction 2 and 3. They can also click on a location to see companies that are known to hire there (top 5), and click on the company to get a word cloud to show common reviews, which can be further filtered by rating, which addresses task 4 of becoming familiar with companies that hire and their ratings.

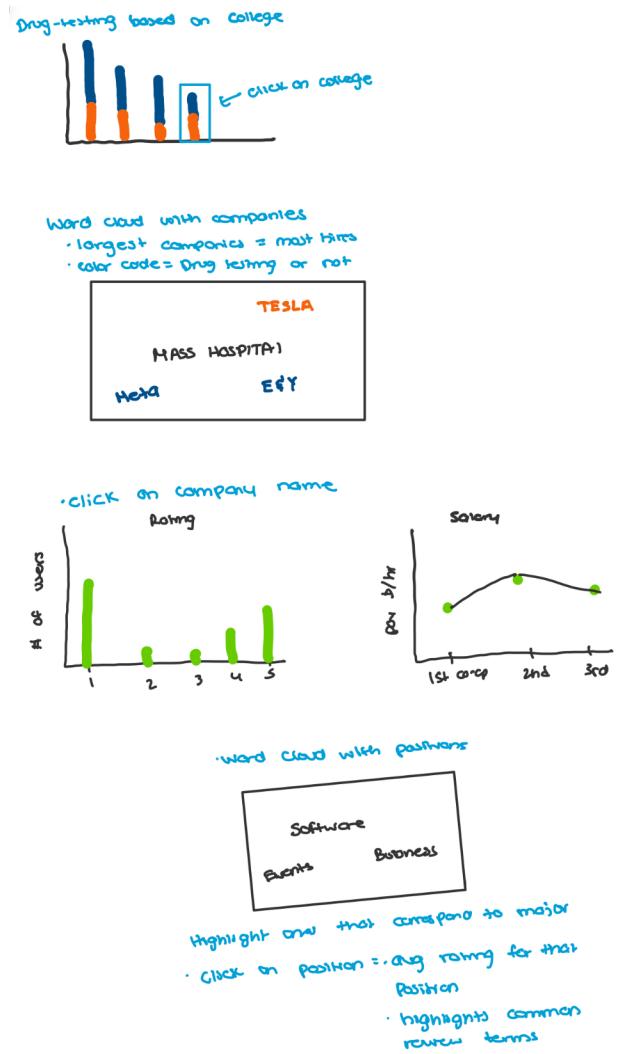


Figure 16: Pranathi : This visualization focuses on showing company specific information, addressing task abstraction 4 and 5, of exploring companies and comparing companies ratings, drug testing, and salary with each other. This is done through the use of size, color, and x and y position. Specifically, the company that have the most northeastern hires, are larger, and companies that drug test are shaded in orange, while those who don't are blue. Clicking on a company name, will show a distribution of the rating for that company, so people understand how the average rating for the company is being calculated. Furthermore, there is the reviews and various positions in that company listed, for information purposes and additional details on request.

D.2 Favorites

We picked our favorites mainly by observing which ones could be connected easily to each other to create a cohesive coordinated view, and views that could communicate meaningful information that would be useful to the user.

The first favorite, in Figure 7, is able to be used to let the user explore data to gain insight on ratings and salaries of companies. It can also be used to compare companies with other companies based on two key features. This visualization supports quantitative and categorical data. The depicted visual encodings are effective because it is able to use the channels to effectively communicate data.

The second favorite, in Figure 13, is able to be used for the user to explore companies to see if they are above or below average when it comes to rating and compensation. This visualization supports quantitative and categorical data, and the depicted visual encodings are effective because it uses position and lines effectively to establish intuitive boundaries for easier interpretation.

The third favorite, in Figure 14, is able to be used to discover average salary based on location, and to show average salaries and ratings per location, and explores different colleges and companies. This visualization supports quantitative and categorical data, and the depicted visual encodings are effective because there is a clear point of connection between each encoding, which allows for intuitive understanding.

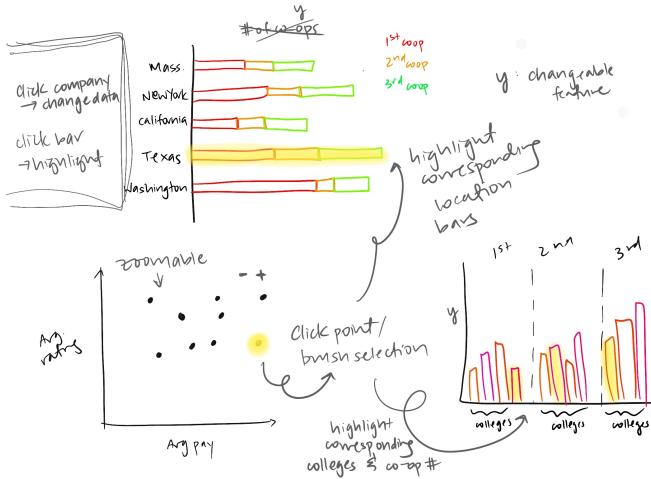


Figure 17: Visualization Tool Design — The final sketch uses points and areas as marks, and color, position and area as channels to convey information that is arranged in separate, ordered, aligned regions and axis orientation (through stacked bar graphs and a scatterplot.) The colors are used to classify and group different areas as the same group, and areas are used to convey the volume of certain features. Position is used to compare different companies in terms of two different measures. The coordinated visualization tool supports the data and tasks that were previously identified by allowing users to interact with it to explore more about insights to the data which include: co-op positions for a certain college, comparisons of pay by location and college, exploration of pay for certain colleges, and exploration of company ratings.

E DIGITAL SKETCH

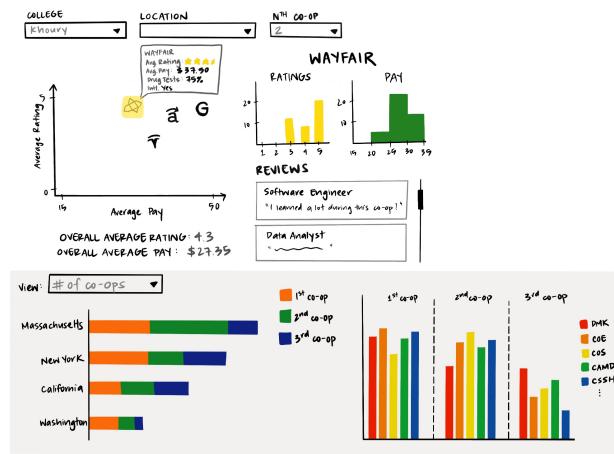


Figure 18: Visualization sketch



Figure 20: Visualization brushing and linking 1

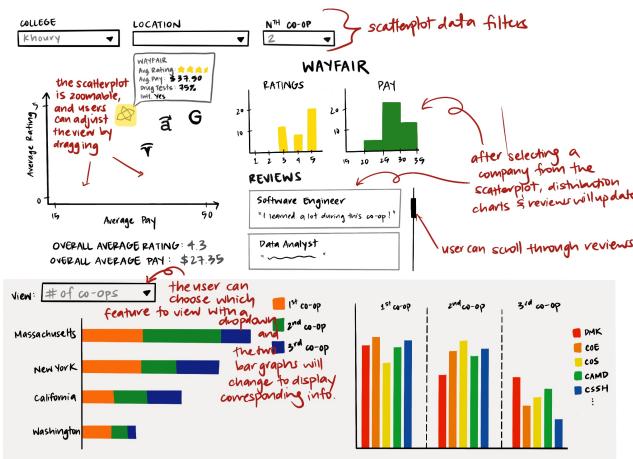


Figure 19: Visualization interaction details

Most of our end users want to compare the different average ratings and compensation for the college they are in and the location(s) they are interested in. This can be achieved by clicking on the filter drop-down menus and selecting the appropriate items. The average pay and rating at the bottom will be updated every time a filter is changed and the graph too would only show the companies that have had co-ops from the chosen college in the specific selected location.

Anyone wanting to compare more specific details about multiple companies would also be able to do so by clicking the logos of the companies they want to inquire about. Every time the cursor hovers over a company logo on the scatter plot, a tool-tip will popup informing the user about the name, average rating, average pay, percentage of drug tests, and whether they have hired any international students based on the currently selected filters. Whenever a company logo is clicked, two bar plots will be displayed on screen, showcasing the distribution of ratings and pay, so the user can understand how the data was aggregated to show the average pays and filter. Color would be used to separate the different attributes. Below the bar charts, all of the reviews that were left by students (which pass

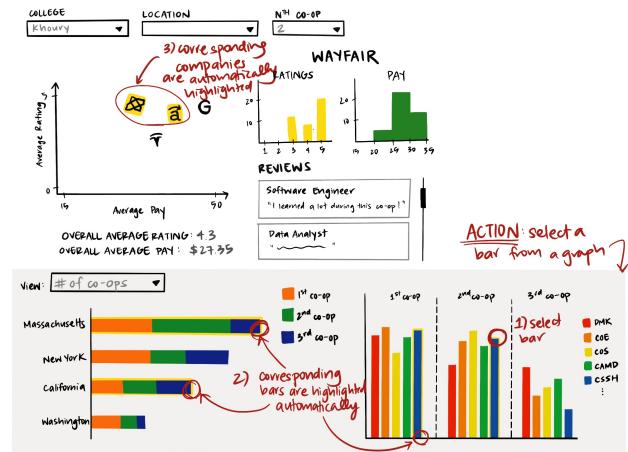


Figure 21: Visualization brushing and linking 2

through the filter) who previously co-op'd there would also be listed, along with their positions. There will be a filter that when a user clicks on, filters the review based on rating, as most of our users want thorough information on the company experience, ratings, and reviews.

Another task that is important to our end-users is exploring and getting information about co-ops, such as salaries, company names, ratings, based on location, and just generally seeing where past Northeastern students have gone. This is done through a bar graph, showing the various locations on an axis. We used this approach, because there isn't enough data to use an actual map, and found this way to be more effective and easily understandable.

Furthermore, most users are still exploring co-ops and don't know where to start, and would like to be able to do this easily. Hence, we have included a graph that shows co-ops broken down by college and nth co-op, so users can easily differentiate and understand where they would fit in.

For both of these graphs, showing location and colleges, it was important for the end user to be able to see ratings, salary, along with just number of hires/people who have co-op'd there in the past. Hence, there is a filter drop-down menu, which the user can click

and change the y-axis accordingly to what they would like to make their decision based on.

Also, to make this exploration even easier, the user will be able to use brushing and linking among the three graphs. To be specific, if the user want to see all location where a company hires, they can click a company logo on the scatter plot, and it will highlight all corresponding locations. Similarly, clicking on locations in the bar graph will highlight the company logos that are in that location. This functionality can also be used for the third bar graph based on co-op number and college. For example, clicking on a college for a co-op number will highlight the companies that have hired NEU students in that college in the past. On the other hand, clicking on a company name in the scatter plot will change the college bar graph to clearly show which colleges and co-op number were hired by that company, and how many of each.

F USABILITY TESTING

F.1 Preparation

Today we will be taking a look at a new visualization application called CoOp'd. CoOp'd can be used to give insight on other students' co-op experiences across different locations, colleges, and companies. It displays distributions that can be used to explore and compare average pay and average rating of co-ops, the distribution of compensation and ratings for individual companies, features relative to the co-op location, and features relative to the n'th co-op and college.

For this evaluation, you will be asked to complete 3 specific tasks with the tool. Feel free to ask any questions you may have about the tasks. You're encouraged to think out loud as you work through the tasks, so that we can better understand your thought processes as you use the app. Do you have any questions?

For the first task, let's say that you are a first-time co-op searcher and want to explore the average salary for a different companies. Show me how you would go about finding this information.

- We would like to test this task because this is one of the main methods of discovery and comparison for our visualization tool that can provide valuable insight to our end users.
- The ideal outcome is for the user to click on different points on the scatterplot pay graph and look at the corresponding information displayed.

For the second task, let's say that you are a second or third time co-op searcher, looking for the average rating of companies in the state of New York. Show how you would go about finding this average rating.

- We would like to test this task because typically, past their first co-op, students start looking into co-ops outside of Massachusetts, which would be the motivation behind this task. This would allow them to explore different ratings of companies that operate outside of Massachusetts, specifically in New York.
- The ideal outcome is for the user to click on the "New York" bar in the location bar graph, and then look at the corresponding information displayed underneath the scatter plot.

For the third task, let's say that you are a first-year student and want to get a general feeling of the quality of experience and building a list of target companies. Please show how you would go about doing this.

- Some companies are bound to provide more valuable co-op experience than others, and this can be reflected in the rating that students give their co-op. By exploring past reviews and aggregated values, students can get a general idea of which companies may provide a higher quality of experience through this task.
- The ideal outcome for this task is for the user to look at the average rating vs average pay scatter plot and observe and interact with the points that are on the higher end of the "average rating" axis to see what companies are being represented in that area.

Well done! We appreciate your time, and your feedback on CoOp'd will be very important as we continue to improve upon our tool design. We will reach out to you if any questions come up regarding your responses. Thanks again for your time!

F.2 Results

The usability test went smoothly and demonstrated no significant misalignment between the developers' intentions and the user's behavior. However, minor issues came to light, mainly relating to interaction with marks in the visualizations. For example, when users were asked to find specific statistics based on a state, it wasn't super clear that they could click on the bar related to the state to change the other graphs' data to reflect the selection. Overall, the test users liked the visualization, bringing attention to how aesthetically pleasing it is as well as intuitive and easy to understand.

Task 1: Explore average salary for different companies

- Positive aspects of our tool that were illuminated during this task included the immediate delivery of information through the scatter plot's tooltip function.
- The metrics and outcomes that we were looking for told us that the usage of our scatterplot and the interactions that it fosters is intuitive for the user. Users during our usability test immediately hovered over each scatterplot point and tried to interact with it by clicking.
- The test results indicate that a change in our design is not necessary for the purpose of this task.

Task 2: Explore average rating of companies in New York

- At the time of the usability test, the functionality for this task had not yet been implemented. However, upon description to the users, they decided it was not intuitive to have clicking a bar on the location graph to manipulate the data for the desired output.
- The specific outcome we were looking for told us that what we were expecting from the user was not intuitive and therefore not discoverable and not ideal for the task.
- The test results indicated that a change was necessary, which we plan to implement through the form of a filter instead.

Task 3: Explore companies with higher ratings

- A positive aspect of the tool illuminated was that our scatterplot was quite intuitive, and it was clear to our user, that companies that are higher up on the graph, have better ratings. However, because of our data, there are a lot of companies that are highly rated, so it's not the easiest to differentiate between each of the companies.
- Since, we were looking for the user to use the scatter plot, and mouse over the companies with higher ratings, which is what occurred, we realized that our scatter plot does it's intended job, of providing holistic information of co-ops. However, as this task was meant to be first year co-ops, still exploring their options and majors, we were hoping that the college bar graph would be used. Although, this was not implemented by the time our testing, even with a detailed explanation, users didn't tend to draw their attention to the specific bar of first year majors and locations. This indicated to us, that we need to allow for that exploration, in another measure, Furthermore, simply highlighting associated companies makes it difficult to go back, and view more information about that company, as once you remove your mouse, it's no longer highlighted.
- Similar to the previous task, this test results indicated that although a change isn't necessary, it would be extremely beneficial in order to provide the various exploration and tasks that can be done through our tool.

The visualization must make its capabilities more apparent to the users. For that reason, we decided there must be a more obvious way to filter data rather than just clicking on the bars. We decided to create some redundancy so that users have more than one way to filter data. For example, to see the information specifically for California, the user can click on the California bar on the locations bar chart but also click on the drop down list and select California from there. This decision will not affect the aesthetics, but it will make it easier for the user to navigate the data. Furthermore, this allows for filters based on only one criteria, rather than the two used for the bar graphs. Also, having the filters allows for the screen to be modified and stay there, while selecting the bar graphs just highlights the companies. For example, if a first year business major is trying to find companies with high ratings, rather than simply trying to pin point all the highlighted companies in the scatter plot when they hover over a major and co-op number, they now have the option to filter based only on co-op, and have those remain on the screen. This makes the process much clearer and less overwhelming.

Other modifications that we can make is including a zoom feature for the the scatter plot. Since, there are so many companies overlapping each other, zooming might help clear up the gaps, and allow for each company to stand on it's own. This makes it easier to click and find the rating for that company. In a similar manner, the location bar graph also has lots of small bars, making it difficult to select and view highlighted bars. However, zooming in results in important information loss. For instance, all the locations a company has offices in, should be seen in one panel rather than having to scroll up or down. Thus, we decided to use filtering to show locations, and not rely on brushing and linking between the two bar graph, as that wasn't clear in getting the information across. Now, one can filter for majors and co-op at the top, and get a concise graph that only shows the locations that previous co-ops have been to. Also, there is a clear count present, which is more difficult to calculate through highlighting a section of the stacked bar graph, which is the method we used when brushing and linking.