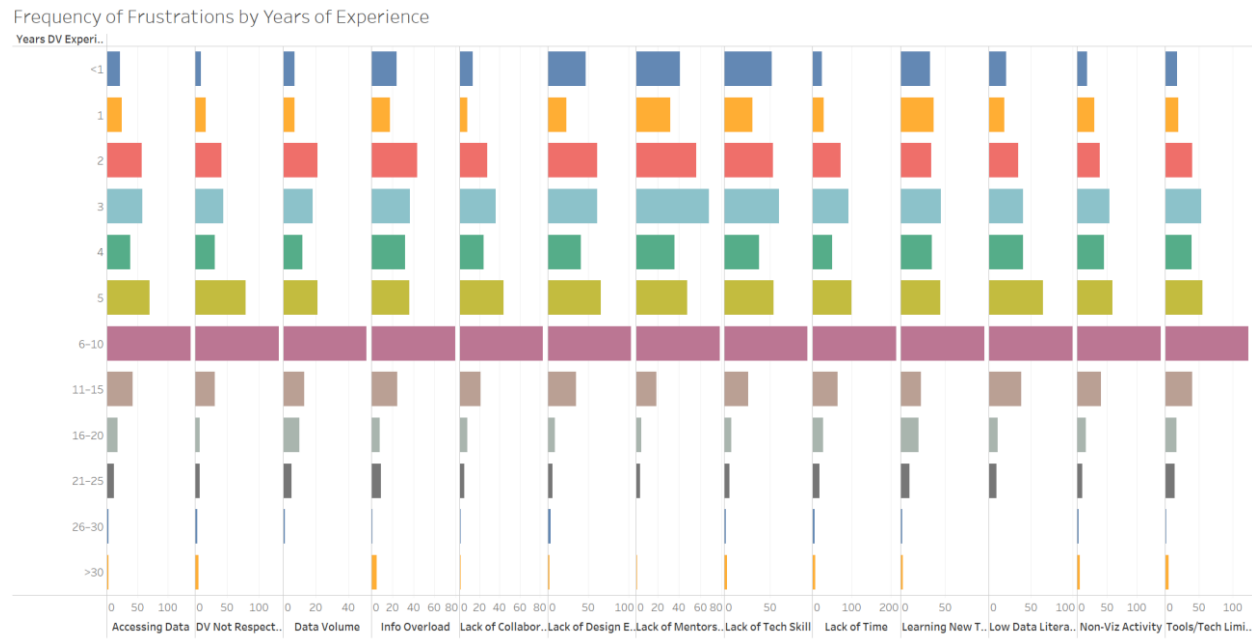


Career Problems in Data Visualization



Data Masons

Name (full name)	Purdue Email address
Caroline Dixon	dixon128@purdue.edu
Hao Liang	liang393@purdue.edu
Michael Maslowski	mmaslow@purdue.edu
Zaire Hayes	zmhayes@purdue.edu

Project URL: <https://datamasons.weebly.com>

Table of Contents

Introduction	3
Background	3
Questions	3
Problem Statement	3
Methodology	3
Results	3
Discussion and Conclusion	7
References	7
Appendix A – Resources Used	8
Datasets	8
Tools used	8
Appendix B – Project Web Page	8
Appendix C – Percent Contribution	9
Group Contributions	9
Individual Contributions	9
Appendix D – Individual Contributions	10
Team Member #1: Hao Liang	11
Team Member #2: Caroline Dixon	12
Team Member #3: Zaire Hayes	13
Team Member #4: Matthew Maslowski	14
Appendix E - Diversity Statement	15
Appendix F – Team Consensus	16
Team Consensus	16

Career Problems in Data Visualization

Introduction

Our group chose to explore the correlations between the top frustrations in data visualization, years of experience in that practice, and education level to research how gripes with data visualization vary in commonality across different people that use the practice.

Background

This year's Hackathon data set was a survey conducted on people who use data visualization as a part of their full-time career, and involved a lot of questions that measured different aspects of working in data visualization. We also cited an article for more context for our topic in the presentation - [Frustration at Work](#).

Questions

The questions our team has chosen to address are:

- Do the top frustrations change as years of experience of Data Visualization grows?
- Does education level affect the top frustrations?
- What are the most common top frustrations?

Our audience is people interested in pursuing a career that utilizes data visualization practices and we wish to demystify the education requirements and downsides of using data visualization as part of a full time career. A lot of different groups in the hackathon had chosen similar topics, and there have been a number of studies about workplace frustrations that we had looked through to find something related, so this topic has been researched in the past before.

Problem Statement

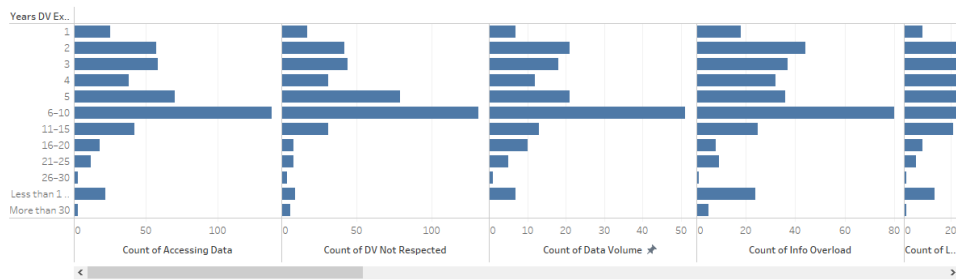
Our work is incredibly significant to specifically college kids who need to know the amount of experience and education needed to work in these fields, as well as the potential frustrations they could have within those careers. Demystifying data visualization fields help to foster potential interest in learning good data visualization practices, and establish how widespread the use of these practices are. Our team wants to specifically research how experience and education affect the commonality of those frustrations, honing in on these specific aspects of careers instead of focusing on salary, which a lot of other groups do for similar reasons.

Methodology

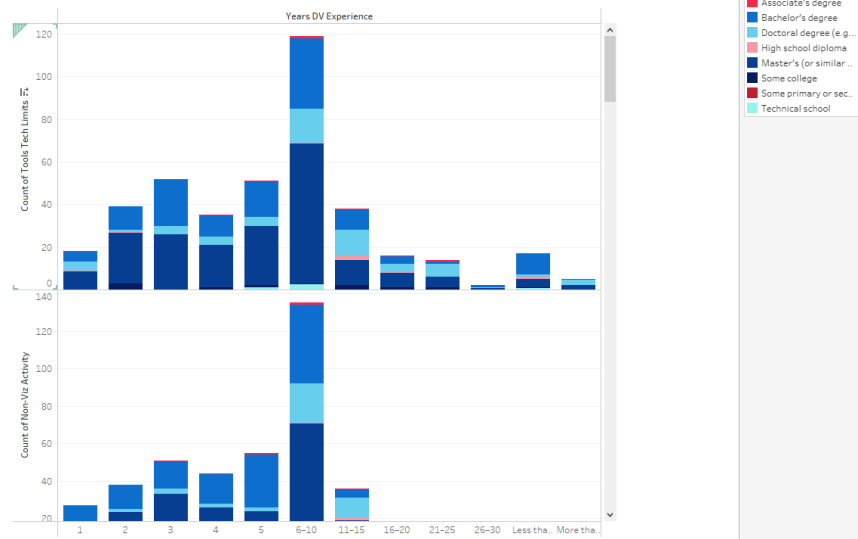
After obtaining the Hackathon data, we decided to focus on the relationships between Frustrations, Years of Experience in Data Visualization, and Education Level measures. We then isolated and sorted the data we wanted to use, parsing out the null values and renaming certain values, and creating three different copies of the parsed set of data in order to find correlations between our three variables. We decided to create multiple different visualizations testing out the correlations between the different variables, and then refined the ones we liked into their final versions.

Results

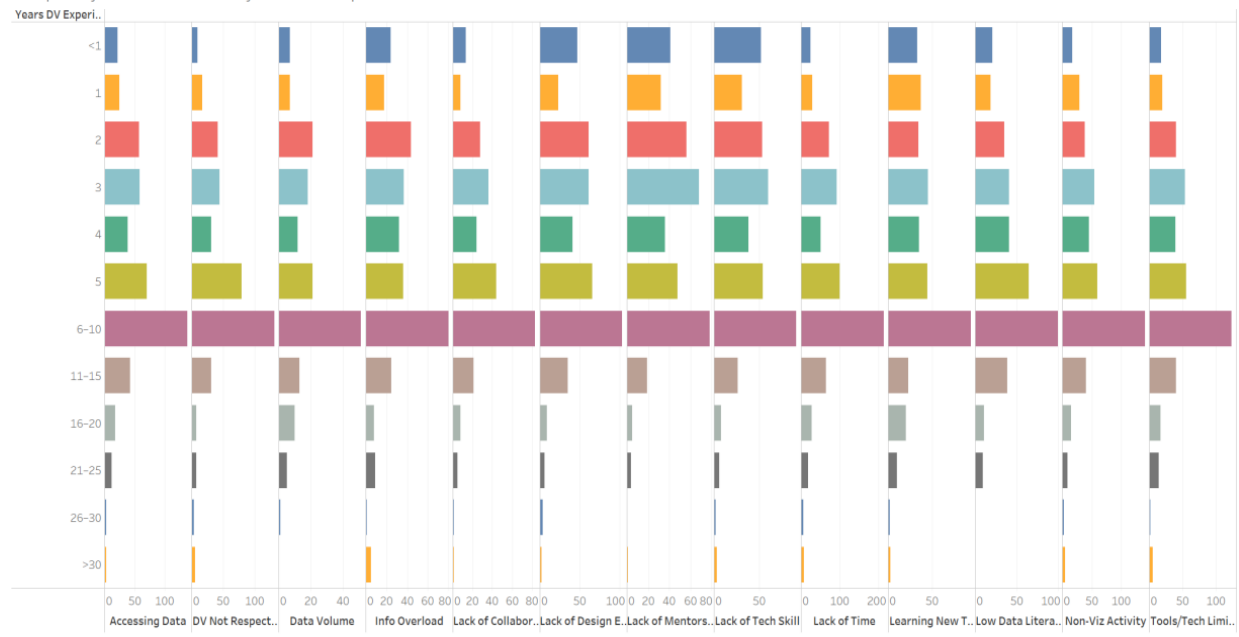
Count of Frustrations by Years of Experience



Top Frustrations Based On Years of DV Experience and Education level



Frequency of Frustrations by Years of Experience



Top Frustrations by Different Years of DV Experiences

Years DV Experience

Value

Measure Names

- Accessing Data
- DV Not Respected
- Data Volume
- Info Overload
- Lack Collaboration
- Lack Design Expertise
- Lack Mentorship
- Lack Tech Skill
- Lack Time
- Learning New Tools Etc
- Low Data Literacy
- Non Viz Activity
- Tools Tech Limits

Years DV Experience	Accessing Data	DV Not Respected	Data Volume	Info Overload	Lack Collaboration	Lack Design Expertise	Lack Mentorship	Lack Tech Skill	Lack Time	Learning New Tools Etc	Low Data Literacy	Non Viz Activity	Tools Tech Limits
3	58	45	18	38	35	62	68	60	92	45	42	55	53
6-10	138	132	52	80	83	103	78	92	215	93	100	140	123
16-20	18	8	8	7	7	8	3	6	28	20	10	15	17

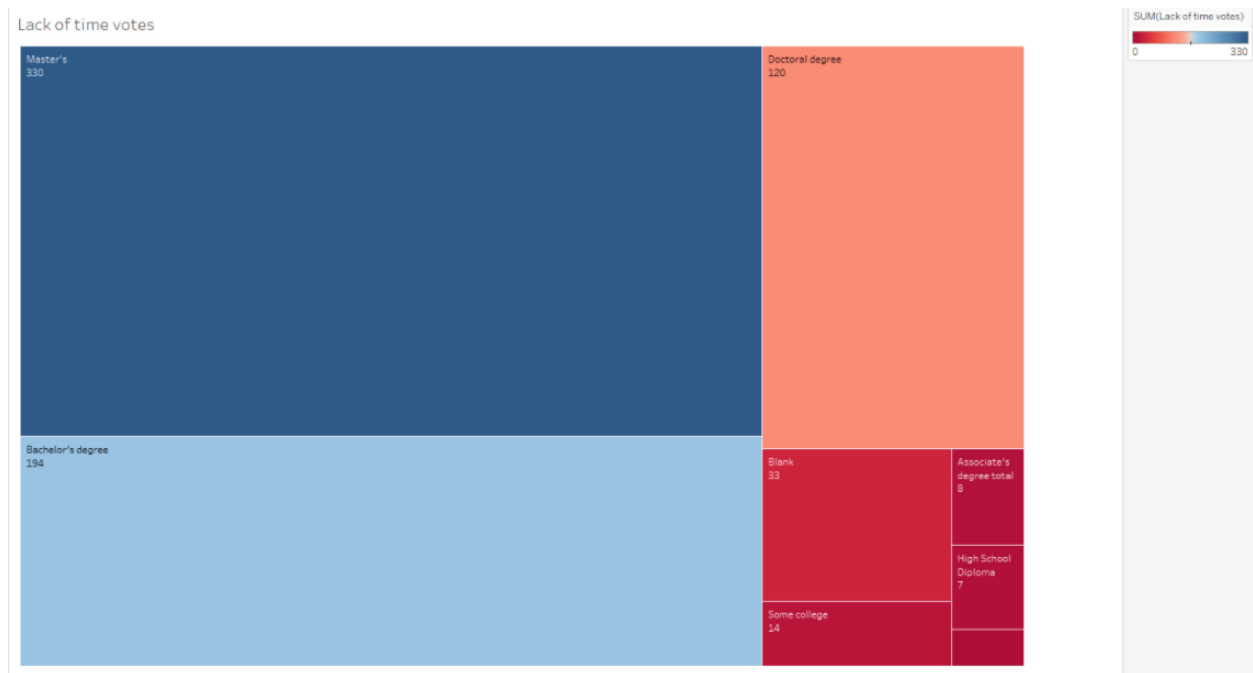
Education level

Education level	Count
Master's	231
Bachelor's degree	129
Doctoral degree	96
Blank	25
High School Diploma	7
Some college	6

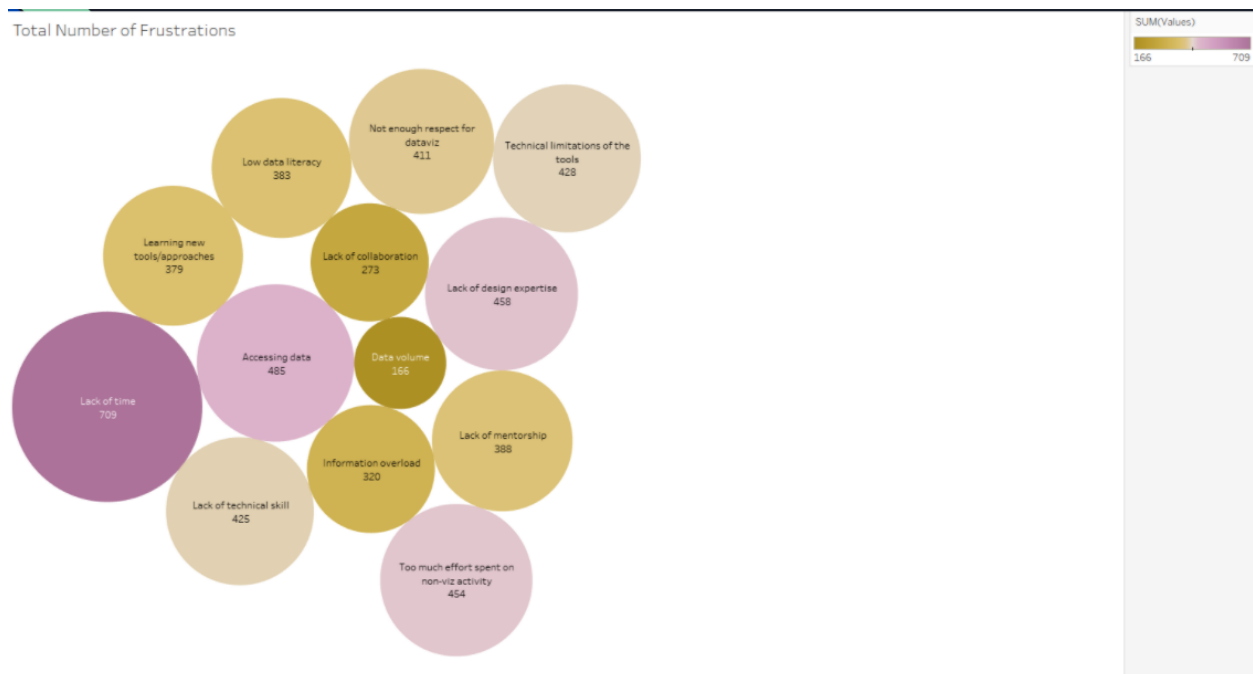
SUM(Lack of design ex...)

0 231

(Made by Zaire Hayes, area chart of instances of the “lack of design expertise” in frustrations in data visualization measured by level of education)



(Made by Zaire Hayes, area chart of instances of the “Lack of time” in frustrations in data visualization measured by level of education)



(Made by Zaire Hayes, bubble chart of the total instances of each frustration in the “frustrations in data vis” variable)

Discussion and Conclusion

We've found that the most common frustration across all of our visualizations and research is a lack of time, followed up by accessing data and a lack of design expertise, which seems to be in line with what students during our presentation had issues with as well. However, the most common frustrations did not vary between years of experience and levels of education as distinctly as would have made a better presentation mainly due to the drawbacks of the data source. There were several variables that seemed to have been collected via multiple choice answers and thus the years of experience category could not be pared down into different years in the 6-10 years of experience category. A recommendation could be that people working on data visualizations and charts be given more time to work on the design for these images.

References

Murphy, Mark. "Frustration at Work." *Leadership IQ*, Leadership IQ, 10 Mar. 2022, <https://www.leadershipiq.com/blogs/leadershipiq/frustration-at-work>.

Appendix A – Resources Used

Datasets

List the name of the data set provided and a description of the additional data set acquired.

2021SOTI_SurveyDataForPublic: <https://www.datavisualizationsociety.org/soti-challenge-2021>

Datasets must be available on the project web page.

All data must be accessible to the public. No paid data sources allowed.

Tools used

Tool/Application	Description
Excel	Data cleaning
Tableau	Data visualization
Weebly	Web development

Appendix B – Project Web Page

The project web page will be an extension of the final report. You will be allowed to add content to the project web page up to last day of classes. The project web page should contain (at a minimum) the following sections:

About The team

List each team member, provide a short bio (150 words or less) for each team member, Provide photo (headshot only) dress appropriately.

The Hackathon Challenge

Describe the team's focus/goal related to the challenge, Who's the audience? What assumptions are made?

Methodology

Describe the team's data visualization workflow and process.

Deliverables

5-minute video (1 pt deduction for each minute over if over 5:00:00 minutes), Hackathon Report, Team agreement (signed by all team members)

Results

This the team's time to shine! Visualizations created by the team that support the team's solution to the challenge, Visualizations must be relevant to the question(s) the team is answering in regards to the visualization challenge.

Conclusions

What insights are presented? What recommendations did the team make?

- Project Web Page: <https://datamasons.weebly.com/>

Appendix C – Percent Contribution

Group Contributions

In this section list the tasks that were completed by all team members for example: contributed to the data visualization process, brain stormed topic ideas, served as rotating team leader, contributed content to the short story (summary), contributed content to the 5-minute video, reading the final deliverable before submission.

Individual Contributions

Table 1 shows an example of what a team contributions table might look like.

Table 1 Example Team Contribution Table.

Team Member	Contribution	Contribution
<i>Example Team Member 1</i>	<i>Developed the project web page, acquired additional data for the project</i>	<i>25%</i>
<i>Example Team Member 2</i>	<i>Responsible for gathering written contributions from the team and combining them into a cohesive story, data wrangling (parsing, filtering) .</i>	<i>20%</i>
<i>Example Team Member 3</i>	<i>Videographer for the 5-minute video (recording and editing)</i>	<i>15%</i>
<i>Example Team Member 4</i>	<i>Creating visualizations of the data, revising and refining</i>	<i>40%</i>
Total		100%

In the table below list each team member's full name, their contribution (body of work) and their % of the work completed. The total must add up to 100%.

Team Member	Contribution	Contribution
Caroline Dixon	<i>Developed the web page, wrote a lot of the report, refined some of the visualizations, contributed to slides + voices</i>	<i>30%</i>
Hao Liang	<i>Parse and Mine, One visualization, Outline of Slides, record 5 slides.</i>	<i>20%</i>
Zaire Hayes	<i>Helped format the slides and made three visualizations</i>	<i>20%</i>
Michael Maslowski	<i>Mined, parsed, and formatted the dataset, created a visualization, edited the presentation video, contributed to slides and voices</i>	<i>30%</i>

Total contributions must equal 100% 100%

Appendix D – Individual Contributions

In this appendix each team member must contribute a one-page document relating the team's topic/data. The one-page document must contain: (1) a description of the problem, (2) a comparison to the team's findings with insights related to the hackathon data (3) a visualization to support items (1) and (2).

Each person should create their individual page (**1-page only**) and make it available to the designated team member who will upload the final document.

This will be viewed and assessed as part of each person's individual contribution.

Leave this page as is.

Start adding individual page content on the next page.

REMOVE any blank pages before submitting.

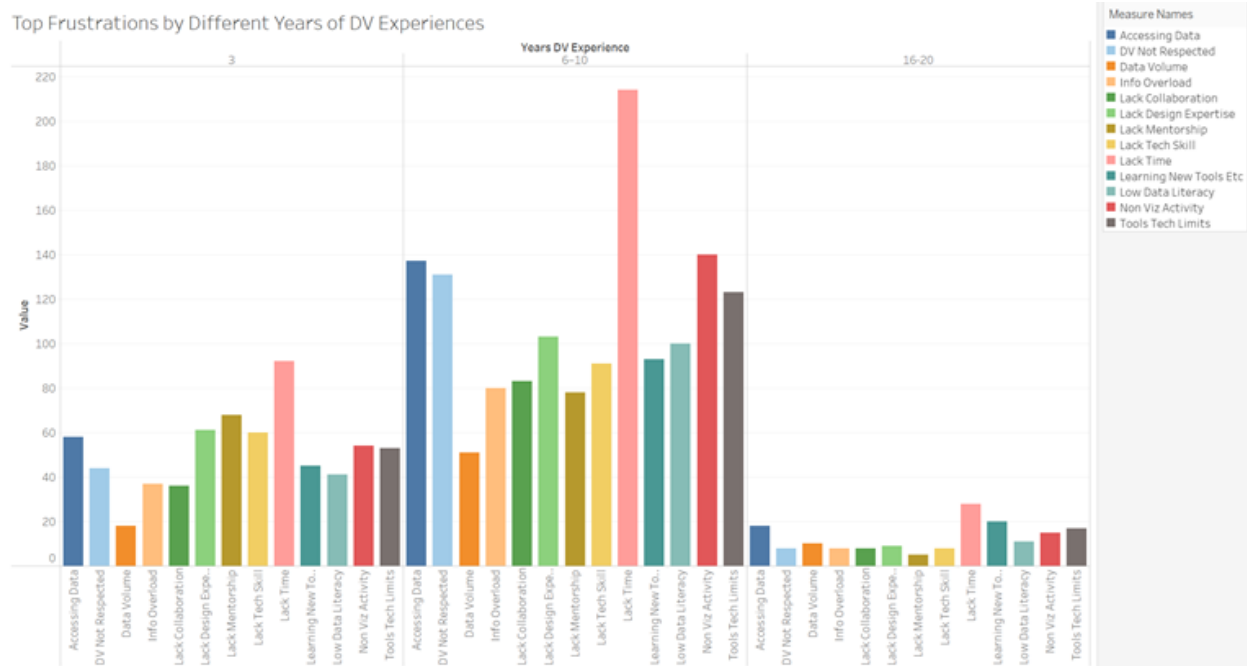
Team Member #1: Hao Liang

Group Topic: The relationship between Top Frustrations, Years of DV Experience, and Education Level
 Your Topic/Question: The relationship and trends between Top Frustrations and Years of DV Experience.
 Describe the diversity YOU bring to the group (150 words or less):

The diversity I bring to the group is being Asian and Chinese International students. I'm also majoring in Game Design. My family is middle class so I will also bring my values and worldview that are from my parents.

Include your story and visualization below (**do not go over one page**). Single spaced, 11-pt font, Times New Roman.

After researching the first version of visualizations created by Michael, I got inspiration from one of his graphs, which is the one you see in the first page of this document. The inspiration I got is that I want to make other visualizations with only three different Years of Data Visualization experiences and compare their differences and similarity. To make such a visualization, I need to find out three different Years of Data Visualization Experiences that are representatives of all periods of all the data we have. Then, I created my first graph, which is a graph that shows the numbers of people in different Years of Data Visualizations experiences and try my best to separate them into three periods. Finally, I choose three years, six to ten years, and 16-20 years of DV experience as the representative periods that I will use to create Visualizations and see what insight I can find. Gladly, I find out that some Frustrations are relatively remain high for all three periods, such as Lack of Time, Accessing Data, and Non Viz Activity. Since even 16-20 Years of DV experience reflects high on these three Frustrations, I believe they are a good goal for company and office to solve. Besides Visualizations, I also created half of the 5-minutes video Slides and set up the outline of it.



Team Member #2: Caroline Dixon

Group Topic: Correlations between Years of Experience, Frustrations in Data Visualization, and Education Level

Your Topic/Question: Correlation/trends regarding education level

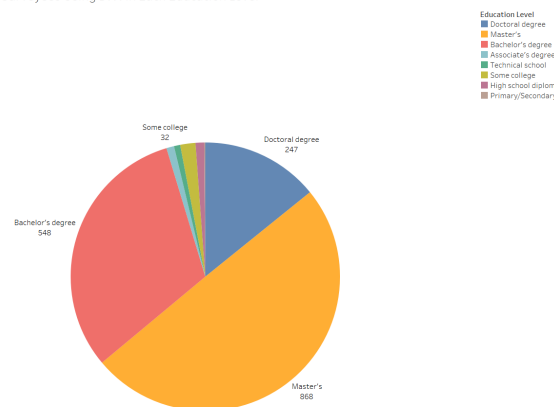
Describe the diversity YOU bring to the group (150 words or less):

My diversity is being from California, as well as being Irish-Scandinavian American and a member of the LGBT community. I am also part of a small number of women pursuing careers in computer science, and part of an even smaller number of women pursuing careers in game development.

Include your story and visualization below (**do not go over one page**). Single spaced, 11-pt font, Times New Roman.

Personally, I mainly created drafts of some visualizations in order to test out certain associations between our chosen research variables. Most of my work that made it to the final version was done in the website and final report, as well as obviously the presentation video. Regardless, I did go through the data myself and was the one who discovered the drawback of how the “years of experience in data vis” variable was measured. That variable had a fixed number of answers that suggested it was collected via a dropdown menu or multiple choice answer, and the largest number of surveyees who answered that question answered “6-10 years,” and there was no way to break down the specific years of experience because it was a general multiple choice category. It definitely tested our ability to work around the issue but there was no way we could change our topic or fully fix the issue in good faith to what the data suggested (people during our presentation suggested we calculate the “6-10 years” category’s specific years of experience but that would not have been faithful or accurate to what the data measured). Overall, I think I found a lot of good correlations between variables that later went on to be refined by other members into full visualizations, and sent the rest of our group down on a good path to investigating the relationships between these three variables. Below is one of the visualizations that didn’t make the final cut that I did, all the other ones are revisions for other people’s visualizations.

Number of Surveyees Using D.V. in Each Education Level



Team Member #3: Michael Maslowski

Group Topic: Correlations between Years of Experience, Frustrations in Data Visualization, and Education Level

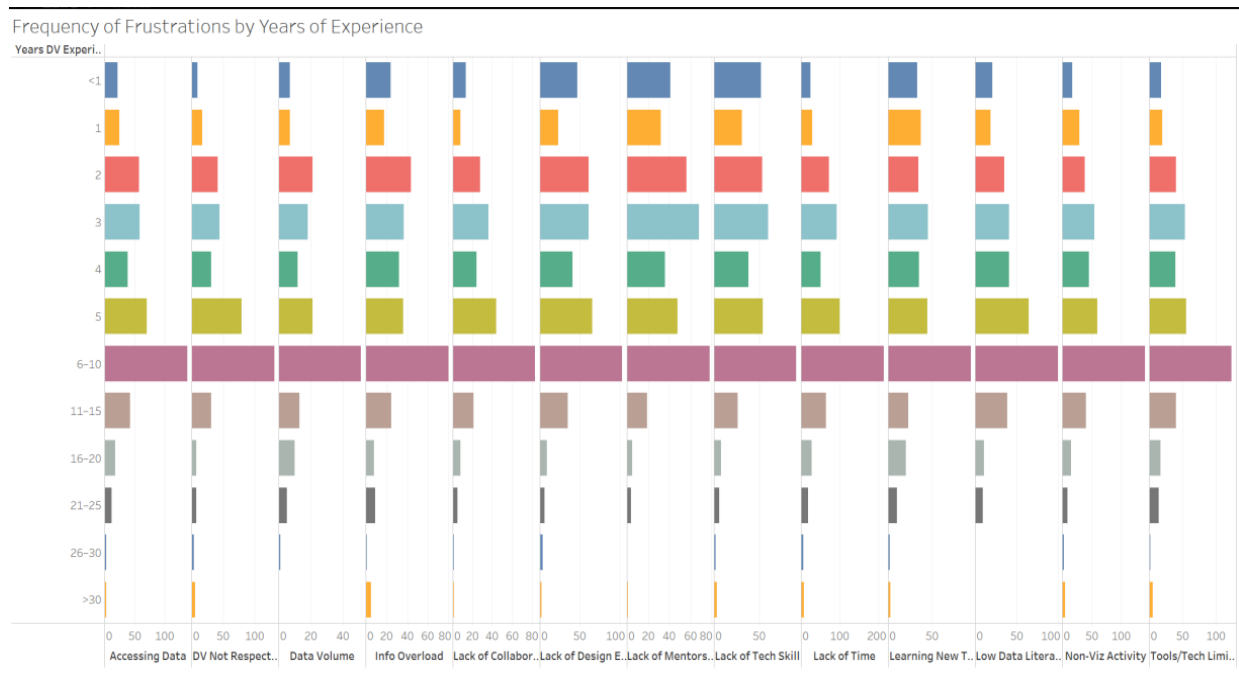
Your Topic/Question: Analysis of dataset

Describe the diversity YOU bring to the group (150 words or less):

The diversity I bring to the group is being the first generation child of two Polish immigrants. With that, I bring my hard work and dedication that I have been raised with by my parents. I am a junior that recently switched into the Game Design major. I bring a unique perspective by having experience at Purdue from my previous major.

Include your story and visualization below (**do not go over one page**). Single spaced, 11-pt font, Times New Roman.

My major involvement with the project was the mining, parsing, and filtering of the dataset. I broke up all the separate elements that the group would utilize to create the different visualizations. These elements included the years of DV experience, education level, and top frustrations. I created a tableau workbook using this information. The tableau workbook had some trouble initially with the top frustrations. Since the question asked for a list of multiple answers, each response would include their top frustrations and then a list of nulls for the missing responses. To counteract this, I flipped the way tableau read the dataset. I then filtered out any left-over null values to create a basis of our analysis. I then took years of DV experience, education level, and top frustrations to create several rough visualizations to see if there were any correlations between these elements. These visualizations were then refined to more optimal and visually appealing visualizations that we used for the presentation. I took one of these initial visualizations and refined it more by adding color.



Team Member #4: Zaire Hayes

Group Topic: Correlations between Years of Experience, Frustrations in Data Visualization, and Education Level

Your Topic/Question: Total frustrations

Describe the diversity YOU bring to the group (150 words or less):

My name is Zaire Hayes and I bring diversity to my group and class in general. I am an African American male making up one of the lowest groups of people in CGT (and Purdue overall). I bring in a way of looking at data that is looked at as rare due to the number of people that are in the category I fit in. I also major in game design which also gives another perspective to data rather than majoring in game design like other people in the class.

Include your story and visualization below (**do not go over one page**). Single spaced, 11-pt font, Times New Roman.



I made three visualizations, this being one of them. When we made our original questions I already knew which parts I felt were going to be key and can also be built upon in our hackathon. Total frustrations is something I wanted to focus on, which was the first one I made. I also made the two tree charts that we had in our presentation. I used these in order to compare the amount of frustrations from the different educational groups seeing if there was a trend to be spotted within the data. I also helped format the slides in the presentation to make sure it lined up and made sense.

Appendix E - Diversity Statement

Some of the most enlightening outcomes are generated by diverse teams working together to solve complex problems. What does diversity mean and why is it important? Merriam-Webster defines [diversity](#) as: 1) the quality or state of having many different forms, types, ideas, etc., 2) the state of having people who are different races or who have different cultures in a group or organization. When solving complex problems having adequate representation is important. In the context of the hackathon, diversity could mean (but is not limited to): varied perspectives, varied points of view, different academic majors represented, different academic levels (Freshmen, Sophomore, Junior, Seniors) on the team, different ethnicities (state this professionally). Having a diverse team from different backgrounds can boost engagement and productivity and make us smarter (read short article: “How diversity actually makes us smarter”).

In the space below, provide a statement describing the group’s diverse make up and how the diversity of the group contributed to the outcomes of the team’s deliverables for the hackathon. Every team member must contribute to the development of the diversity statement.

Team Data Masons has a makeup of entirely Game Development and Design Majors, one of the majors included in the Computer Graphics Technology department. The team is also mostly freshmen, save for Michael Maslowski who is a junior. Data Masons is also diverse in ethnicity, background, nationality, and hometown; Hao Liang is from Dongguan, China, and while the rest of the team is from the United States they all hail from different parts of the country. Caroline Dixon is from San Mateo, California, Zaire Hayes is from Fort Wayne, Indiana, and Michael Maslowski is from Lafayette, New Jersey, all wildly different regions of the country. Hao Liang is an international student from China and who believes his Chinese value will contribute to the team during the project. Zaire Hayes is an African American male making up one of the lowest groups of people in CGT (and Purdue overall); he brings in a way of looking at data that is looked at as rare due to the number of people that are in the category he fits in. Michael Maslowski is a first generation child of two Polish immigrants, and has only recently switched into a Game Development and Design major, bringing with him the skill and dedication from not only his background but his previous major. Caroline Dixon is a member of the LGBT community as well as an aspiring illustrator, bringing her skill in visual design with her into the project. While all of us were radically different from one another, each of us brought a different cultural mindset as well as unique skill to the project based on our backgrounds.

Appendix F – Team Consensus

Team Consensus

I have read and approved of the content as a representation of the team's work and my contribution.

<u>Hao Liang</u> Print Team Member Full Name	<u>Hao Liang</u> Signature	<u>4/25/2022</u> Date
<u>Caroline Dixon</u> Print Team Member Full Name	<u>Caroline Dixon</u> Signature	<u>4/25/2022</u> Date
<u>Zaire Hayes</u> Print Team Member Full Name	<u>Zaire Hayes</u> Signature	<u>4/25/2022</u> Date
<u>Michael Maslowski</u> Print Team Member Full Name	<u>Michael Maslowski</u> Signature	<u>4/27/2022</u> Date

Save this document as:

HackathonTeamName_CGT270Spring2022_FinalReport.pdf