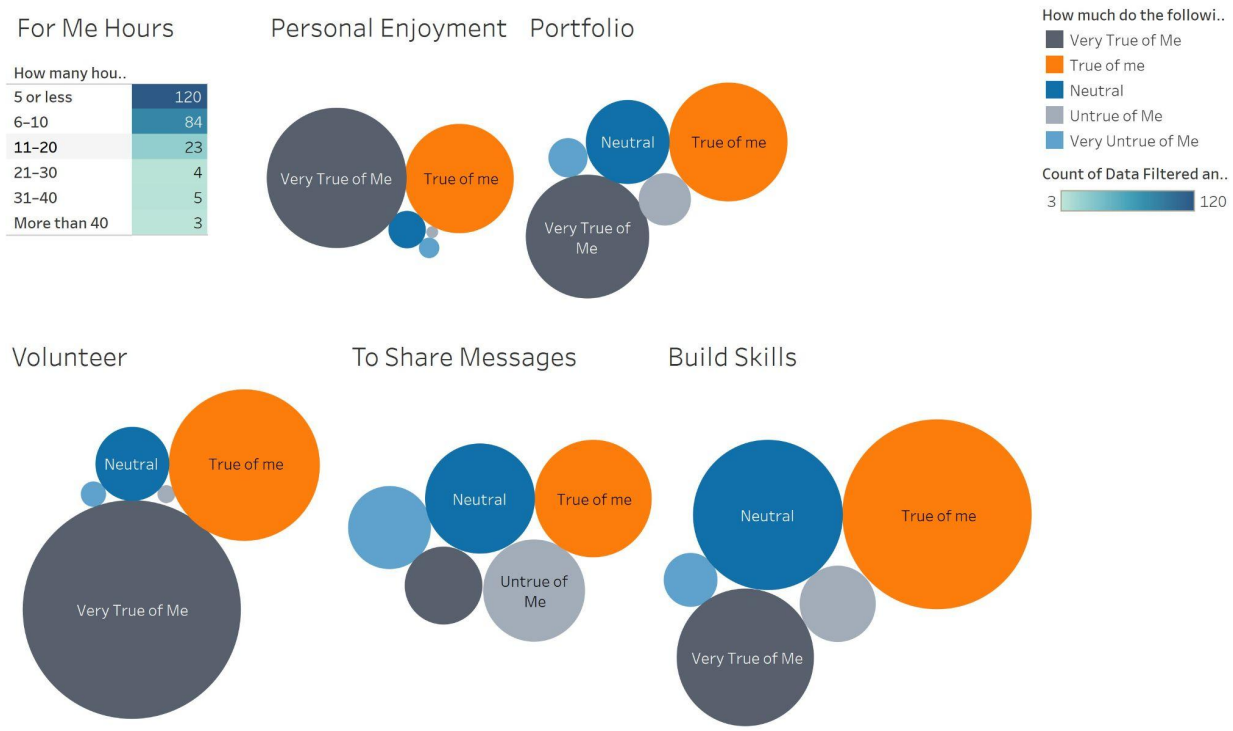
Data Visualization 2021 Demographics



Team name

Name (full name)	Purdue Email address
Robinson J. Witt	witt31@purdue.edu
Noah P. Egierski	negiers@purdue.edu
Keegan C. Palonis	kpalonis@purdue.edu

Project URL: Team Octopodes' Website

Table of Contents

Data Visualization 2021 Demographics

Introduction

In our project, we wanted to answer a few questions about the field of Data Visualization in general. We wanted to understand what field Data Visualizers have or are receiving their degrees, as well as what role data visualization played in receiving their degree. Then, we wanted to know whether data visualizers worked more as freelancers, or part of an organization. We also wanted to compare the two different types of jobs by looking at their job title, and who they make their visualizations for. Finally we looked at what data visualizers did outside of their careers in the field of data visualization.

Background

The data we used was from the Data Viz. Society's annual survey. This survey polls people within the data visualization industry about a variety of issues, such as why they make visualizations to tools to effects of COVID. We chose a page from the compiled set of results about the 2021 data.

Questions

What were the degrees that data visualizers pursued in school? What role did data visualization play in you getting your degree? Do data visualizers work more as freelancers or as part of an organization? Who do freelancers and employees of an organization make visualizations for? What sector are those organizations a part of? What do data visualization hobbyists spend their data visualization time on?

Our audience is anyone that is interested in pursuing data visualization whether it is through school, as a freelancer, through an organization, or as a hobby.

Problem Statement

The significance of this data is to show people who are working in data visualization, or students interested in a career in data visualization what kind of education, career, and other work are common in the field. With this information, people can both better select their field of study, but also have a better understanding of what kind of jobs they might be doing. This will allow them to make more educated choices in order to get to a career that they would enjoy.

Methodology

Throughout our project, we followed the data visualization process in order to make the most effective visualizations and presentations that we could. Very quickly in our process, we determined that we were interested in what Data Visualizers did in their careers, as well as what fields they worked in. From there, it was simple to select the questions we felt were relevant to our question, and put them in a separate excel file in order to more easily understand them. It was at this point that we found interest in both the education and hobbyist portions of the data, and so we included them in our spreadsheet to utilize. Then, we started to understand what each question of the data was asking, and how we could use them. We brainstormed what chart types would be most useful for each question, and assigned different portions of the data to each of us in order to make it easier to explore. Each of us then mined our portion of the data, education by Noah, freelancer and parts of organizational by Keegan, and the rest of organization and hobbyist by Rob. We then reconvened and showed each other our prototype visualizations, and critiqued each other and further developed what things we wanted to spend our time on. We further filtered the

data, as well as made our second assumption as mentioned above, and made a second round of visualizations. Then we started to prepare for the video, compare conclusions, and make a story with our data.

Results

Choose one of your team's "BEST" visualizations and insert it here. This visualization should be the best representation of the team's effort. Provide a figure caption. If the team would like to include additional visualizations, add them to this section. Make sure each visualization has a figure caption AND includes the name of the person who created the visualization. Also make sure the appropriate page layout (portrait or landscape) is used as well as the appropriate chart type and layout (see best practices for visualization data and data visualization checklist).

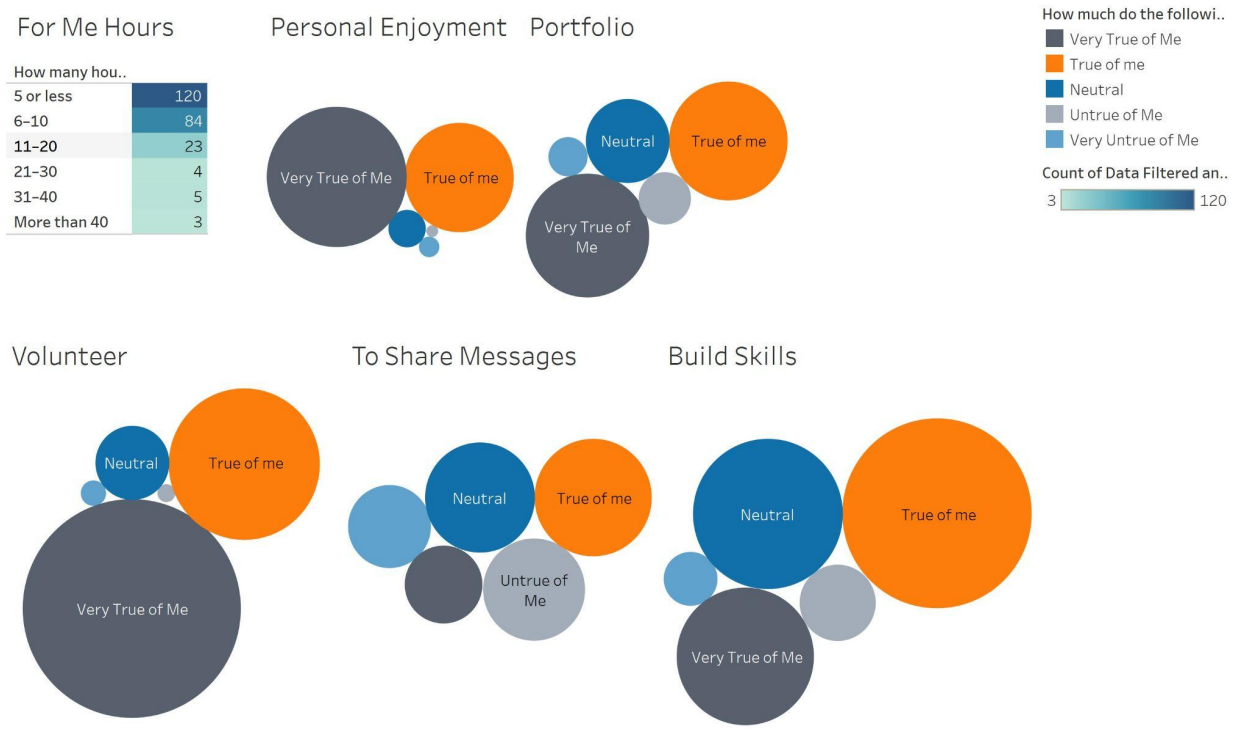


Fig.1 Responses of Visualizers who Create in their Spare Time

Made by Rob Witt

Discussion (What's the story?) and Conclusion

Discuss your results (the figures in the Results section). Do your visualizations address the problem stated in the Problem Statement Section? Explain. What insights did the team uncover? What recommendations can be made based on these insights?

We believe that our visualizations address the problem in the problem statement. We explore what education data visualizers receive, from field to level of education. We also analyze the different careers in both freelance work and organizational work, as well as the level of focus on data visualization in those careers. Finally, we explore other ways that people do data visualization, as well as the reason behind it.

Through our project, we can conclude a few things. First, we determined that the majority of Data Visualizers received their degrees in STEM fields by a large majority. We also determined that the majority of responders received a graduate degree. What was interesting was that only 40% of responders said that data visualization was a required course within their degree. This was intriguing since all of them experienced data visualizations in their careers, but only a minority was introduced to it during their education. This also shows that data visualization is still not a core part of college education, even though a large number of people do data visualization related work.

Next, when we looked at freelancers and people who worked at an organization there were a few interesting points. That there was such a small portion of responders that had worked as a freelancer in the past year (~28%) was interesting, as it means that a large majority all work as part of an organization. We had expected a slightly larger portion to be freelancers, as data visualization lends itself to contract or individual project work. Next, when we looked at what careers people had in both freelance and organization, there were some surprising differences. We did expect that the most popular career field overall would be analysts, as that position involves a lot of understanding and presenting data. What was unexpected was that the most popular career as a freelancer was a designer, and second in an organization was a leadership position. Therefore, we can determine that the most common field was analyst, but depending on whether you work as a freelancer or as part of an organization, your responsibilities might be widely different. Another interesting observation is that there is a difference in level of focus on data visualization between freelancers and organizational workers. Half of freelancers stated that data visualization was their primary focus, only 30% of organization employees did so. This is an understandable difference, as freelancers are oftentimes more concentrated into a specific area, but it was interesting to see a minority of organization employees consider it their primary focus.

Then, we compared who freelancers and organization workers made their visualizations for. We were less surprised by this data, as the largest category of freelancers made visualizations for the general public, and employees in an organization made the most for analysts. What was interesting was the number of visualizers who said they made visualizations for themselves. A final interesting observation was the top seven categories for both freelancer and organization were the same, just in different orders. This means that no matter if you work as a freelancer or as part of an organization, you will find yourself doing work for similar groups.

After working on Freelancers, we shifted our focus onto Data Visualizers who create for their organizations. When comparing job roles to the significance of data visualization, we found that a majority of jobs have visualization as a secondary focus, we found this to be most surprising for analysts. Another surprise was who we found visualizing to be a major focus for: Designers, Developers, Journalists, and Teachers. When looking into what sectors Data Visualizers work in, we found, unsurprisingly, that Information Technology, the Private & Public sectors were the most popular. However, we also found that Non-Profits were also one of the most popular sectors, which we thought was surprising. Another surprise was that marketing was the least popular sector for visualizers to work for. We found that, unsurprisingly, Analysts and Project Managers were the highest responses for who visualizers made for. We also found that visualizers also frequently made for themselves, but the data did not expand on what this exactly meant.

We also focused some of our efforts on people who made visualizations for personal reasons. We found that, weekly, just over half of data visualizers who create for themselves work for five hours or less. The next largest group, at ~35%, work six to ten hours. When looking into the reasons they create, we found all responses to be positive, with the most popular being volunteer work, building a portfolio, and developing their skills. The only reason with a sizable "untrue of me" response was to share messages. We

collected all of the data onto a single dashboard because the responses were measured on a Likert Scale, so we felt it would make for the easiest comparison to consolidate our findings.

References

If references are listed, make sure they are cited in the body of the document. See Purdue Online Writing Lab for how to cite and list full citations. Improperly cited work will be treated as plagiarism and handled accordingly.

Appendix A – Resources Used

Datasets

Data Set Provided: 2021SOTI_SurveyDataForPublic

Data Set Used: data_2021_main filtered into Data Filtered and Adjusted (made by us)

Tools used

List all tools used in the project and a brief description (see the *examples* below); update accordingly.

Tool/Application	**Description**
Excel	Data cleaning/Filtering
Tableau	Data visualization
Weebly	Website 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 the 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 is 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?

-

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

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

Table 1 Example Team Contribution Table.

Team Member	**Contribution**	**Contribution**
Example Team Member 1	*Developed the project web page, acquired additional data for the project*	*25%*
Example Team Member 2	*Responsible for gathering written contributions from the team and combining them into a cohesive story, data wrangling (parsing, filtering) ,*	*20%*
Example Team Member 3	*Videographer for the 5-minute video (recording and editing)*	*15%*
Example Team Member 4	*Creating visualizations of the data, revising and refining*	*40%*
	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**
Robinson J. Witt	*Created Visualizations and mined/filtered data relating to organizational and personal data of survey takers, recorded organizational and personal portion of the video, refined/critiqued visualizations*	*30%*
Noah P. Egierski	*Created visualizations and mined/filtered data relating to academic history of survey takers, recorded academic related portion of video, critiqued/refined visualizations.*	*30%*
Keegan C. Palonis	*Created Visualizations, and helped mine and filter data relating to freelance and organization. Recorded the freelance and parts of the organizational portion of the video, edited the video, made the website, and refined and critiqued visualizations.*	*40%*
	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.

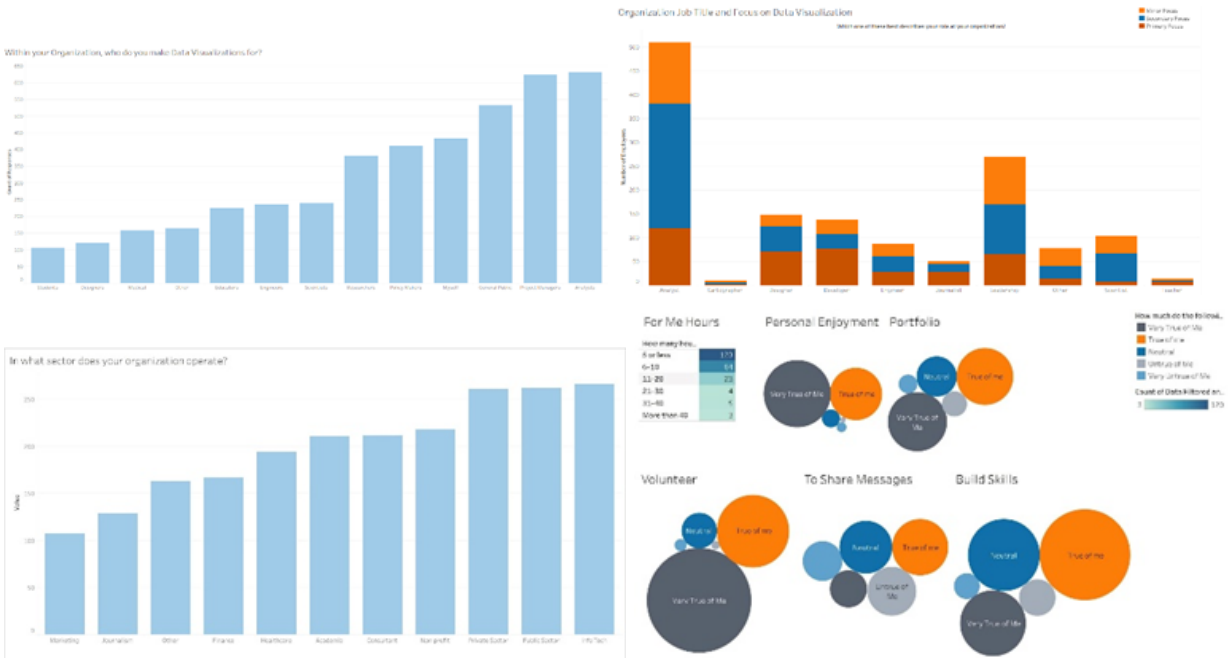
REMOVE any blank pages before submitting.

Team Member #1: Robinson Witt

Group Topic: 2021 Data Viz Demographics
Your Topic/Question: Organizational & Personal Data
Describe the diversity YOU bring to the group (150 words or less):

I'm a sophomore in UX Design with a minor in Psychology. I'm also an Eagle Scout, demonstrating that I have a background in project management. I also have experience in excel both from personal life and from that I used to want to be an engineering major. I'm also a member of the LGBTQ+ community.

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



For the Hackathon, I chose to take on the sections involving data visualizers who make visualizations for employers/organizations and who make them for personal reasons. I started by using the PDF version of the survey to isolate which questions dealt with these two parts. Beyond that, I removed questions involving tools, issues, and salary. This was not because these questions were not important, but that they were not the most relevant to the questions we were answering.

With the personal data visualizers, I decided to create a bubble chart to display the results of a Likert scale. To consolidate the responses, I altered the color scheme to be consistent with each other and put them in a dashboard. In that dashboard, I also included the weekly hours worked as a colored table. This was to further consolidate the relevant, included data into a single visualization.

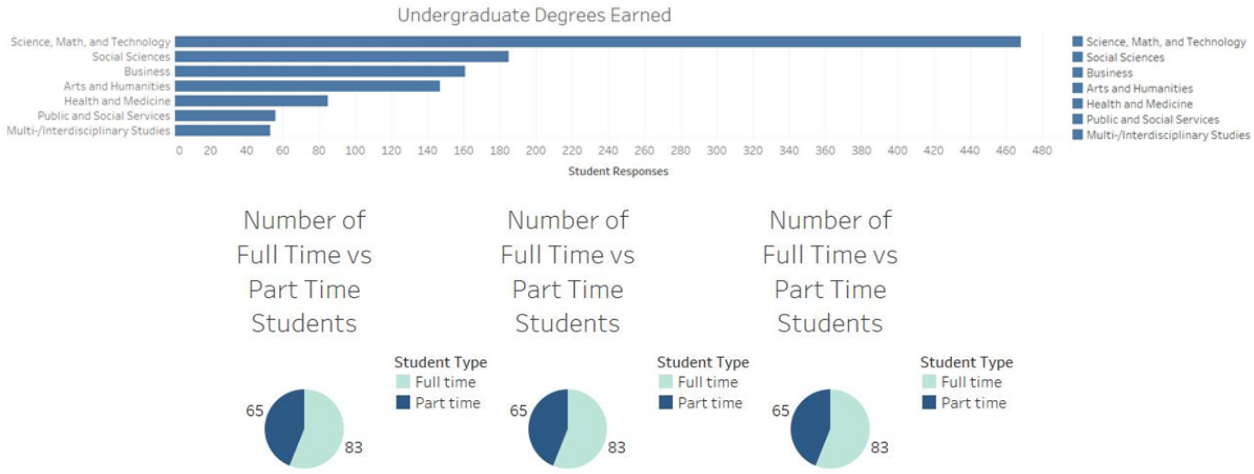
Working on the organizational data, I kept to bar charts, as they were an easy, comprehensive way to compare data from the same question. The exception to this was the stacked bar chart, which I made mimicking the style of another visualization made by a teammate. This was because, as opposed to the other bar charts, this visualization measures a third variable for each answer.

Team Member #2: Noah P. Egierski

Group Topic: 2021 Data Visualization Demographics
Your Topic/Question: Academic Related Data
Describe the diversity YOU bring to the group (150 words or less):

I am a freshman in Game Development and Design. Relating to skills, I have used Excel for the past 8 years of my schooling career to work with data, and I have experience presenting said data. I also have had plenty of prior experience working with teams.

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



For the Hackathon, I decided to work on the section relating to the academic history of people within the data visualization field. Looking through the pdf for the survey, my group mates and I all found questions that best suited what we wished to find and isolated them. Looking through the isolated data, I assumed that all "null" values were questions where people decided not to answer, and where it was acceptable, I turned the "nulls" into "no's".

First, I decided to look into whether students were full or part time, and what degrees they were pursuing. For these visualizations I decided to use pie charts because I really wanted to show the comparison between the variables, and because of the chart types I tested, they turned out to be the easiest to understand. For the color scheme, I went with an easy on the eyes blue scheme.

Looking further into the education, I looked into academic areas of specialty, degrees earned, and whether data visualization was taught within their field. Again, I used pie charts to show simple comparisons, but I used bar charts to not just show the values for all the selectable survey options, but to show off the extremes in the data. For the pie charts I again used the blue color schemes, but for the bar charts I stuck with a solid blue color to not make anything too complicated for viewers.

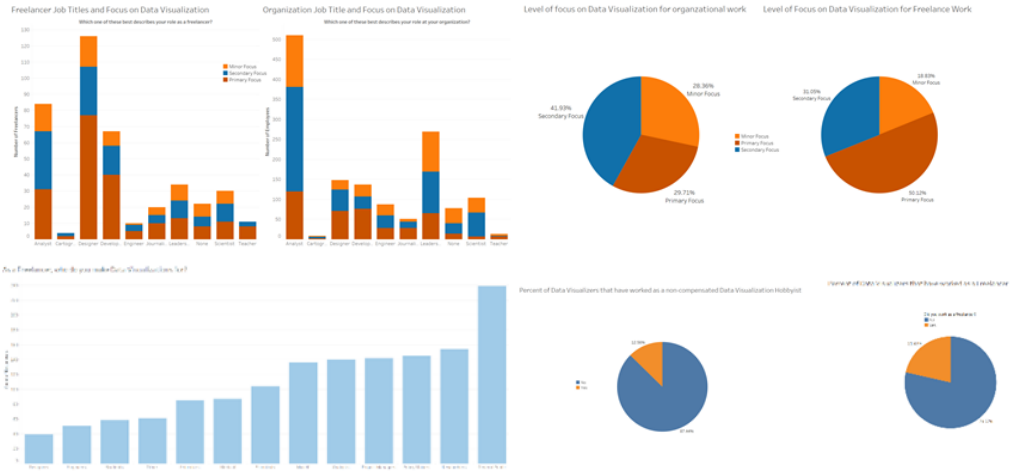
Team Member #3: Keegan Palonis

Group Topic: 2021 Data Viz Demographics

Your Topic/Question: Freelancer and Organizational Data

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

I'm a Junior in Data Visualization with a double minor in Communications and Economics. I was previously a Physics major. I've held several leadership positions, such as the Boy Scouts of America, where I was both an Eagle Scout and the Senior Patrol Leader, which is the head of the entire troop. I have also been a part of the marching band in both high school and here at Purdue. I have also done work in the data visualization field for my local school district.



In the Hackathon, my responsibility was to look at the portion of respondents that said that they had worked as a freelancer in the past year, as well as the level of focus that data visualization plays for both freelancers and organizations. I also looked at the percentages of people that stated they both worked as a freelancer, and as a hobbyist in order to get an idea of the size of the population in both sections. Overall, I contributed by building the website, participating and editing the five-minute video, as well as helping to hammer out a lot of the details in what we wanted to talk about both before and after the video, and making the visualizations above.

In my personal visualizations I mainly kept to bar, stacked bar, and pie charts. In the first visualization it shows the number of respondents per top title in both Freelance positions and Organizational positions, and the colors represent whether they responded that data visualization has a primary, secondary, or minor focus within their career. Next to that is a pie chart that shows the percentages of freelance and organizational workers that considered data visualization a primary, secondary or minor focus. With these, we tried to see if there was a major difference between the two in both job titles and focus on data visualization, which you can see that there were some. The next bar chart shows who freelancers made visualizations, to compare to Rob's Organizational bar chart of similar nature. The last two pie charts show the portion of respondents that said they had worked as either a freelancer or data visualization hobbyist in the last year, so we could further understand the portion of data visualizers that do either kind of work.

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 makeup 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.

The Octopodes, while not being culturally diverse per say, is relatively diverse when relating to skills and experience. Our team is made up of individuals ranging from freshman to junior, and we all found ourselves in different areas of study. Rob Witt is studying UX Design with a minor in Psychology, Keegan Palonis is studying Data Visualization with a double minor in Communication and Economics, and Noah Egierski is studying Game Development and Design. Having members from such varied areas of study and differences in college experience really opened this group up to ideas and solutions. Also, even though we are all similar in many ways, we each have our own ideas and experiences that lead us to a variety of ideas and solutions to problems we discover. We all also have had experience with tools such as Excel and Tableau so that helped us out a lot. It really helped when one of us had an issue with one of our visualizations because someone could help us get the kinks out.

Appendix F – Team Consensus

Team Consensus

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

Print Team Member Full Name	Signature	Date
Robinson J. Witt	Robinson Witt	4/29/2022
Noah P. Egierski	Noah Egierski	4/29/2022
Keegan C. Palonis	Keegan Palonis	4/29/2022

Save this document as:

HackathonTeamName_CGT270Spring2022_FinalReport.pdf