Roles and Responsibilities in the Data Visualization Industry

Chart, bar chart

Description automatically generated

**Team name**

|  |  |
| --- | --- |
| **Name (full name)** | **Purdue Email address** |
| Connor Colbert | [colberj@purdue.edu](mailto:colberj@purdue.edu) |
| Tyler Nielson | <tnielson@purdue.edu> |
| Ema Westerfeld | <ewesterf@purdue.edu> |

**Project URL:** <https://teamairfryer.weebly.com/>

[Introduction 3](#_Toc102092060)

[Background 3](#_Toc102092061)

[Questions 3](#_Toc102092062)

[Problem Statement 3](#_Toc102092063)

[Methodology 3](#_Toc102092064)

[Results 6](#_Toc102092065)

[Discussion (What’s the story?) and Conclusion 9](#_Toc102092066)

[References 9](#_Toc102092067)

[Appendix A – Resources Used 10](#_Toc102092068)

[Datasets 10](#_Toc102092069)

[Tools used 10](#_Toc102092070)

[Appendix B – Project Web Page 10](#_Toc102092071)

[Appendix C – Percent Contribution 11](#_Toc102092072)

[Group Contributions 11](#_Toc102092073)

[Individual Contributions 11](#_Toc102092074)

[Appendix D – Individual Contributions 12](#_Toc102092075)

[Team Member #1: Tyler Nielson 13](#_Toc102092076)

[Team Member #2: Connor Colbert 14](#_Toc102092077)

[Team Member #3: Ema Westerfeld 15](#_Toc102092078)

[Appendix E - Diversity Statement 16](#_Toc102092079)

[Appendix F – Team Consensus 18](#_Toc102092080)

[Team Consensus 18](#_Toc102092081)

Roles and Responsibilities in the Data Visualization Industry

# Introduction

Our project focuses on the role and pay disparity in the data visualization industry. We focus on the characteristics and gender that lead to one’s role assignment and salary, and how their role impacts their pay. We decided to tackle these from the outset of the project as we felt they were most prevalent to those looking to enter the data visualization industry or those who are not a part of it.

# Background

We chose the [*2021 State of the Industry*](https://docs.google.com/spreadsheets/d/1lDkxioTmT5--JufJuYryiV5fKsdQQopvlSJO4Gh0ors/edit#gid=1462100456) survey as our dataset for this project as we wanted to focus on the most recent data to accurately represent the climate of today’s data visualization industry. This data is from the survey conducted by the Data Visualization Society to gain insight into the current standing of the industry. Our team did not gather additional data to the original dataset; all visualizations and findings are based on the responses found in the survey data.

# Questions

Our group asked the questions “What characteristics result in the highest annual salary?”, “What role in an organization produces the highest paying jobs?”, and “How does one’s gender affect their role in their organization and their annual salary?”

Our audience consists of those who are interested in entering the data visualization industry and our findings are meant to guide our audience towards their ideal position and salary in the industry.

The problem we are addressing is the lack of information available to provide insight into how one may be placed in the data visualization industry regarding their position and salary.

# Problem Statement

Going into this project, we wanted to focus on the contributing factors toward pay and role disparity in the data visualization industry. The findings our team have come to in this project are important to those who are looking to enter the data visualization industry as it provides insight into what kinds of roles and pay one may receive based on their characteristics, gender, and work experience.

# Methodology

Once we acquired the data set, we got to work finding what sections would be most beneficial to our problems and placed them in a separate Excel sheet. The next step was to get rid of any null values and to put the data into a more usable format. This way, we were able to find the total number of relevant responses. When it came time to create the visualizations, we first sketched out a few ideas and reformatted our data for use inside of Tableau, then used tableau to create visualizations we could work from. Using that process we were able to arrive at several finalized visualizations.

Our Sketches:

A picture containing text, receipt

Description automatically generatedA picture containing text, whiteboard

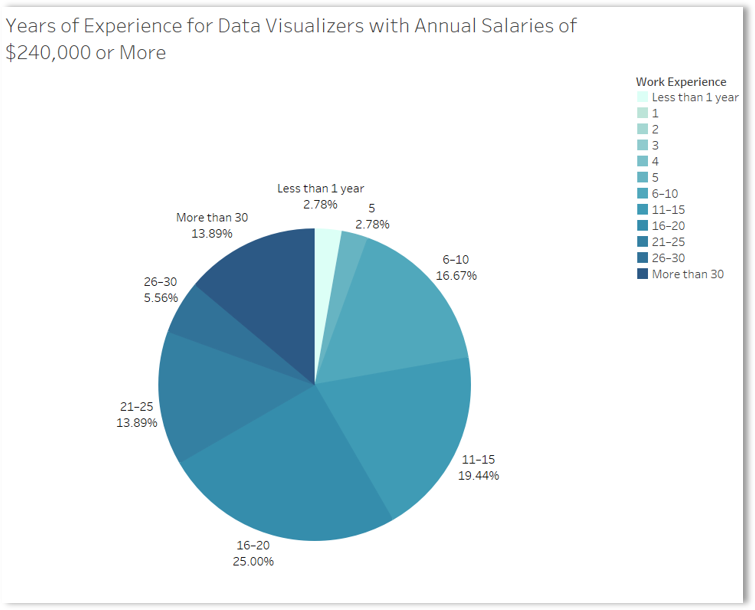
Description automatically generatedDiagram

Description automatically generatedDiagram

Description automatically generatedText

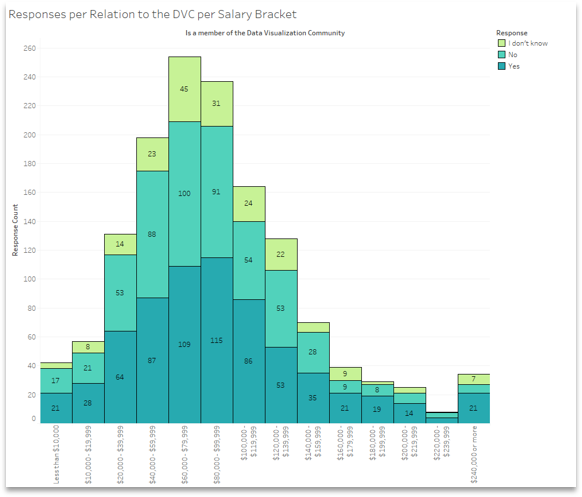
Description automatically generated with medium confidence

# Results



Visualization 1 – Distribution of Years of Experience for those making $240,000 or more (Ema Westerfeld)

This visualization focuses on the years of experience for the highest paid respondents in the survey. The chart shows the percentages of respondents in order to show the proportion of each bracket of experience more accurately. As shown, only 5.56% of respondents have anything below six years of experience, with the vast majority of respondents having anywhere from 6 to over 30. The largest chunk, at 25%, has 16-20 years of experience, showing that this is not an industry that pays its workers highly right off the bat.



Visualization 2 – Data Vis Society Membership by Salary (Tyler Nielson)

This visualization displays the proportions and values of the survey answers in relation to people’s membership in the Data Visualization Community and grouped by their salary bracket. It’s clear that, by looking at the ratio of respondents between the Yes and No categories, as salary increases past $160,000, more people tend to be a member of the DVC. The ratio of respondents who answered “I don’t know” does not appear to fluctuate with salary.

Chart, bar chart

Description automatically generated

Visualization 3 – Organization Role Distribution Across Genders (Connor Colbert)

This visualization shows the organization role/job one receives based on their gender. Percentages are of respondents who fall into a specific gender response rather than of individual roles (e.g., 15.7% of all female respondents occupy leadership positions). Male respondents are more likely to take leadership, developer, and engineering positions by significant margins. Meanwhile, women are nearly 1.3 times as likely to be an analyst than men (9.9% disparity).

# Discussion (What’s the story?) and Conclusion

The first visualization, made by Ema Westerfeld, details the years of experience held by the highest paid members of the data visualization industry. This visualization seeks to answer the question of how a person’s role in an organization can affect their salary. The data is filtered to only show respondents with a salary of $240,000 or more per year in order to be more focused. Only around 5% of respondents had 5 or less years of experience, expressing that it is uncommon to be paid top dollar at an entry level. Based on the chart, the largest number of visualizers in this salary bracket have 16 to 20 years of experience, at 25%. This data supports the conclusion that in order to earn the most money in the data visualization industry, one should expect to earn around 20 years of experience.

The second visualization, created by Tyler Nielson, displays the ratio and count of survey answers regarding people’s membership in the Data Visualization Community, grouped by salary brackets. This visualization aims to explore the relationship between DVC membership and salary. This chart tells the viewer that as salary increases past $160,000, a higher proportion of respondents claim to be a member of the DVC. Below $160,000, the proportion is roughly 50/50 between being a member of the DVC and not being one. This information gave us the insight that being a member of the DVC slightly trends with a higher salary.

The third visualization, created by Connor Colbert, focuses on the distribution of job positions in an organization based on gender. This visualization highlights the disparity of certain positions between men and women in the data visualization industry. Although data for those who responded with “Prefer not to say” and “Self-described” are also included, due to the small sample size the data is not as conclusive. The third question in our problem statement section is partially answered by this visualization as it focuses on the job distribution, leaving a second visualization ([which can be found on the Team *Air Fryer* website](https://teamairfryer.weebly.com/results.html)), to handle the rest. We found, based on the visualization, that men are more likely to occupy leadership positions by a significant margin while women are more likely to be analysts. Based on other visualizations that indicate pay ranges for jobs in an organization, this indicates that men make a significantly higher salary than women on average.

# References

# Appendix A – Resources Used

## Datasets

2021 Data Visualization Society State of the Industry Survey

## Tools used

|  |  |
| --- | --- |
| **Tool/Application** | **Description** |
| Excel | Data cleaning & formatting |
| 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 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 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, brainstormed 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 and grabbed extra data.* | *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%* |
|  | 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** |
| *Tyler Nielson* | *Developed the project web page, created, and uploaded the video, created two visualizations.* | *33%* |
| *Connor Colbert* | *Helped format and aggregate data in the mining stage, created four visualizations, created third question of focus in presentation.* | *33%* |
| *Ema Westerfeld* | *Parsed the data set, helped to mine the data, created three visualizations.* | *34%* |
|  | 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: Tyler Nielson

Group Topic: *Pay and role distribution in the data visualization industry*

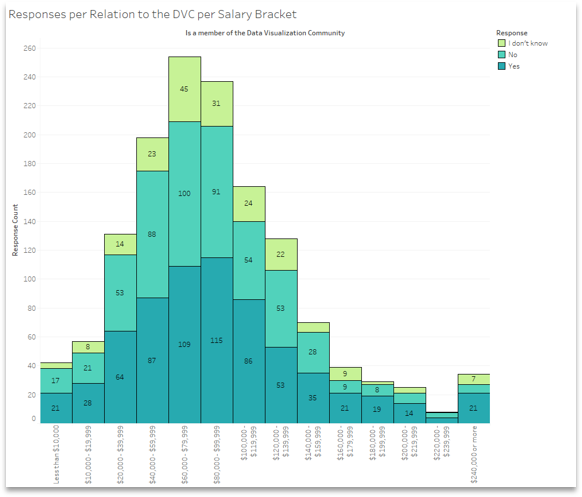
Your Topic/Question: *What characteristics outside of one’s job result in the highest salary in the data visualization industry?*

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

I am a freshman who is double majoring in Animation & Visual Effects. I am originally from Northern Michigan and went to a Catholic school system. I bring the experiences from my upbringing and hobby as an artist in order to view things in a new light and provide insights in a way that nobody else has done before.

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

One of the questions our group decided to tackle was “What characteristics outside of one’s job result in the highest annual salary?” I produced this visualization by first asking myself whether there might be a correlation between a person’s membership in the Data Visualization Community and their salary. Then, I went into Tableau with my imported data and assigned Salary to the X-Axis and Response Count to the Y-Axis for the 3 categories of Yes/No/I don’t know. This placed the data into a format in which both count and ratio of responses can be easily seen by the viewer. I added numbers inside of each of the boxes to prevent any confusion between sizes. I also changed the default colors to a more visually appealing palette while maintaining the distinction between categories. The resulting insight, which was that people with higher salaries in Data Vis positions were more likely to be a part of the Data Visualization community, supported the rest of the data found by our team which suggested that a higher involvement within aspects of the Data Visualization industry contributed to higher pay. (Such as years of experience, DVC involvement, and education level.)



## Team Member #2: Connor Colbert

Group Topic: *Pay and role distribution in the data visualization industry*

Your Topic/Question: *How does one’s gender affect their salary and role in an organization?*

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

I am in the Game Development & Design major, which gives me a unique perspective on what data means and how it impacts people. I work a lot with numbers and code as I specialize in game programming, especially systems programming in games. I also have a unique perspective in that I took engineering courses all throughout high school, so I used data we gathered to improve and tweak our robot AI (Artificial Intelligence) and performance of other devices we made (e.g., bridge trusses). I am also from central Indiana, so I share a domestic perspective with midwestern culture which carries over to Purdue.

My Story & Visualization:

Our goal in this project was to reveal the traits of those who had attained certain roles and salaries to better understand how those new to the industry may be placed in their organization. I worked on mining the data from the survey and formatting it to be easily understood in Excel so deciding which visualizations to focus on became much simpler. With some of this data, I created four visualizations: *Data Vis Hobby Time by Salary*, *Salary by Organization Role*, *Organization Role by Gender*, and *Salary by Gender*. These visualizations gave me insight into how to answer our problem statement and three main questions. I would say my most insightful visualization would be *Organization Role by Gender*.

Chart, bar chart

Description automatically generatedWhen approaching this visualization, the problem statement/question that was tied to it did not yet exist. I was curious how organization roles were distributed across genders to see if that gave any insight into pre-existing stereotypes and possible discrimination in the data visualization industry. What I found led me to create an entire topic question about it and explore it deeper with a second visualization. The distribution of organization roles across genders is significant as it illustrates how societal expectations of men and women have an influence on the role they receive. Due to the smaller sample size for those who responded with “Prefer not to say” and “Self-described,” the data for those categories do not provide reliable conclusions; however, I still opted to include the data to show that not all respondents fell into two categories. In this visualization, men have a significantly larger hold on leadership positions than women. Additionally, men are more likely to be an engineer than a woman is. I found these points to be significant as it appears to enforce stereotypes in the workplace. Upon speaking to some of my friends who are women, they often feel that they cannot hold leadership positions due to their instructions or requests often come off as “bossy” or rude. Meanwhile, men often are revered for their dominant leadership strategies.

If I were to change anything about my Role by Gender visualization, it would be to create comparable charts for previous years to view the trends and see how it has changed. I believe this would provide greater insight into what needs to be changed in the data visualization industry for greater equity. I would also like to see more diversity-centric questions featured in the survey to further analyze the distribution of roles and salary across race, ethnicity, age, etc. to find trends in the industry’s inclusivity.

## Team Member #3: Ema Westerfeld

Group Topic: *Pay and role distribution in the data visualization industry*

Your Topic/Question: *What role in an organization produces the highest paying jobs?*

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

As an animation major, I often pay attention to aesthetics and visual clarity, both of which are vital to creating functional visualizations. But even outside of my general knack for visuals, I have plenty of experience in graphic design, a discipline which has a lot in common with the represent stage of data visualization. These skills allow me to create straightforward visualizations.

Chart, pie chart

Description automatically generatedChart, bubble chart

Description automatically generated

Chart, treemap chart

Description automatically generatedMy first visualization, *Years of Experience for Data Visualizers with Annual Salaries of $240,000 or More*, was made so that we could compare the different years of experience for the highest paid respondents to the survey. Since my question dealt with what allows a person to make the most money in the Data Visualization industry, this was a helpful insight, which indicated that the majority of the highest paid respondents had 16-20 years of experience. My second visualization, *Organization Size of Highest Paid Data Visualizers ($240,000 or More Annual Salary)*, took a similar approach, focusing yet again on the highest salary bracket. This visualization produced a similar conclusion, with most belonging to larger organizations. My third visualization, *Education Level of Data Visualizers by Salary*, focused on the overall education level of the industry, with annotations on my slide to indicate the most important aspects of the chart. This chart indicated two conclusions. One, that, generally speaking, one needs a bachelor's degree or higher to succeed in the industry, with only a small portion of the chart being dedicated to anything below. Two, that the majority of highest paid visualizers have a master's degree, indicating that is in line with the rest of the industry.

# 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](https://www.merriam-webster.com/dictionary/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 boot 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 *Air Fryer* is made up of three members majoring in Animation and VFX (Ema Westerfeld, Tyler Neilson) and Game Development and Design (Connor Colbert). Connor is from central Indiana, while Tyler is from Northern Michigan. Ema is also from Indiana; however, she has lived in Wuhan, China for three and a half years. Connor has a background in programming, which gave him experience in Excel, letting him format the data easily. He also took part in a robotics team and took engineering classes throughout high school. Tyler has experience in graphic design and made charts for his statistics course in high school. He also did some volunteer work during high school. Ema also has a background in graphic design and participated in theatre in high school. She also spent all four years of high school studying Chinese and is conversational in the language. Each member’s background and experiences contributed to our project and allowed us to work well together.

# 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.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Connor Colbert |  |  | Connor Colbert |  | 4/25/2022 |
| Print Team Member Full Name |  |  | Signature |  | Date |
| Ema Westerfeld |  |  | Ema Westerfeld |  | 4/25/2022 |
| Print Team Member Full Name |  |  | Signature |  | Date |
| Tyler Nielson |  |  | Tyler Nielson |  | 4/25/2022 |
| Print Team Member Full Name |  |  | Signature |  | Date |