Gender Wage/Unemployment Gap Project Proposal

Basic Information:

- **Project Title:** A Deeper Dive into the United States' Gender Wage/Unemployment Gap
- **Name:** Cassidy Newberry
- **Email:** crnewberry@dons.usfca.edu
- **Github Link**: https://github.com/CassidyNewberry/Gender-Wage-Unemployment-Gap

Background and Motivation:

My future career goals include using data science to influence public policy and promote social justice. I am particularly interested in gender equality and racial equality and this project focuses on both. I have been an active member of USF's women in tech club for the last four years, and I think this project is a great opportunity for me to not only shed light on the general wage gap in the United States, but also specifically discuss how this impacts women in tech. I have spent my entire undergraduate career dedicating my projects to promoting social justice and I think this project would be a way for me to tie everything together that I have been focusing on over the years. Within this project I will look at the general wage gap, wage gap for specifically women in tech, how COVID has impacted the gender unemployment gap, and the wage gap by race in the United States (also relating it back to gender). I am also motivated to do this project during this time because COVID is causing the wage gap to increase, so the progress feminists have made over the years is currently being depleted. I think this is a great time to shed light on this issue, and I am very excited and motivated to do so.

Project Objectives:

- Does a wage gap exist in the United States, and if so how large is the gap?
- How are women in tech impacted by this wage gap?
- How are women of color impacted by this wage gap?
- How has COVID impacted all of this?

Data:

- General wage gap
    - Dataset collected by Glassdoor
    - **https://www.kaggle.com/nilimajauhari/glassdoor-analyze-gender-pay-gap/activity**
- Tech wage gap
    - Dataset collected by Glassdoor

    - https://www.kaggle.com/nilimajauhari/glassdoor-analyze-gender-pay-gap/activity
- Race and gender wage gap
    - https://www.bls.gov/news.release/pdf/wkyeng.pdf
    - See table 7
    - Will have to find a way to make this table a csv file
- COVID impact on unemployment gap
    - Male: https://beta.bls.gov/dataViewer/view/timeseries/LNS14000025
    - Female: https://beta.bls.gov/dataViewer/view/timeseries/LNS14000026

Data Processing:

I do not expect to have to do much cleaning of the data I am using. The datasets are all very clean, and do not have any NA values. I will have to find the best way to make the COVID dataset a csv file because it is currently only in a table format. From the Glassdoor dataset I am planning on calculating the average base salary for males vs. females. I then plan on filtering the data to only include tech jobs and doing the same calculation again.

For the rage and gender wage gap dataset I plan on looking at the median weekly earnings for each race and gender and comparing them for each category. I also am planning on running a significance test for this data to test if the median weekly earnings is significantly different for any categories.

For the COVID impact on unemployment rates dataset I plan on comparing the average monthly unemployment rate of females and males.

For the data processing I will use Pandas and R in order to filter the data when needed, as well as calculate averages of the data. I will save these calculations into a csv file to visualize using D3.

Visualization Design:

I included my sketches below of what the visualizations will look like. I will make many of them interactive because the values are very close in number. I chose many bar charts in order to include the categorical variable gender. I also did several scatterplots to show how age and seniority impact salary per gender. I chose the color scheme because I think it is what people are familiar with.

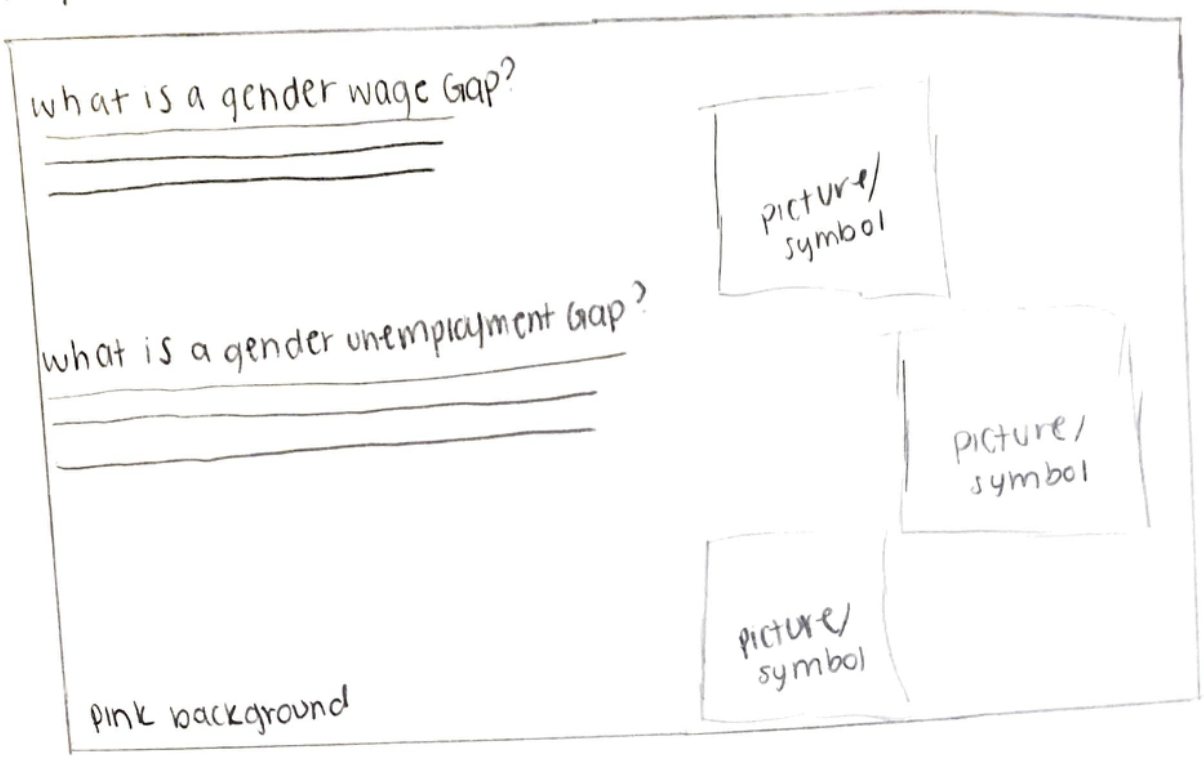
Here is how my visualizations map to my objectives:

- Page 2 includes 4 graphs that all address objective 1

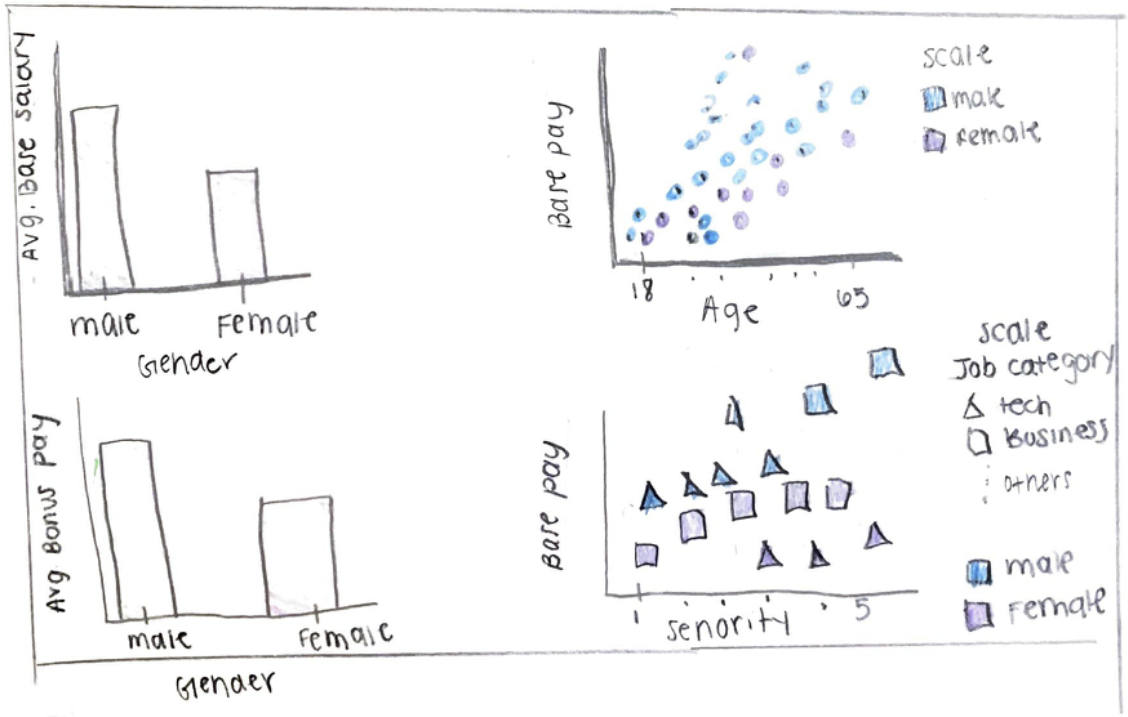
    - The two bar graphs show the general wage gap between males and females and how big the gap is
    - The scatter plot shows how age impacts the general wage gap
    - The four dimensional scatter plot shows how seniority and job category impact the general wage gap
- Page 3 includes 3 graphs that all address objective 2
    - The small multiples address the wage gap based on different tech positions
    - The bar chart shows the overall wage gap within the tech field
    - The four dimensional scatter plot shows how seniority and education impact the general wage gap
- Page 4 includes 2 graphs that both address objective 3
    - The bar chart shows how women on different races are impacted by the wage gap
    - The line chart shows if there were any changes in this wage gap from 2019 to 2020
- Page 5 includes 2 graphs that both address objective 4
    - The line chart shows how COVID has impacted the gender unemployment gap in the United States.
    - The small multiples show how the unemployment gap has changed from 2019 to 2020

Here are my sketches:

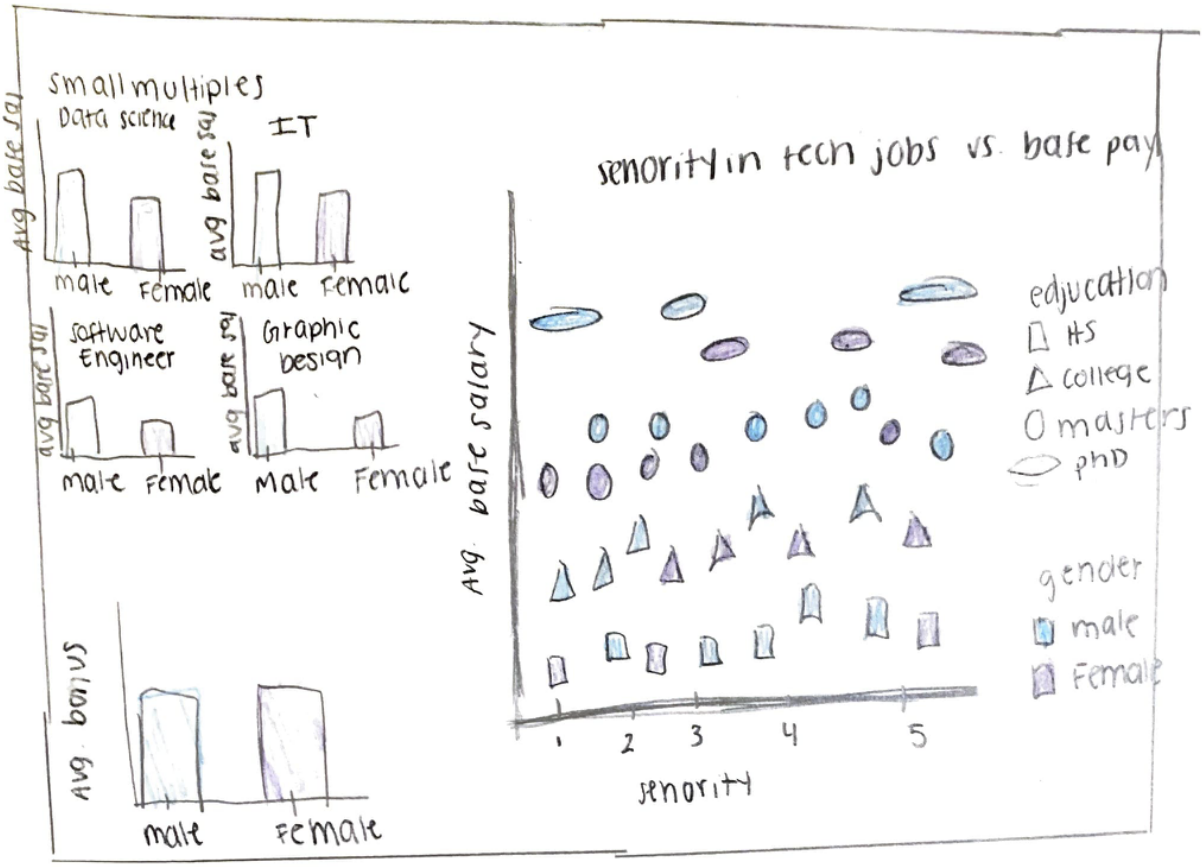
Page 1: Cover / Information



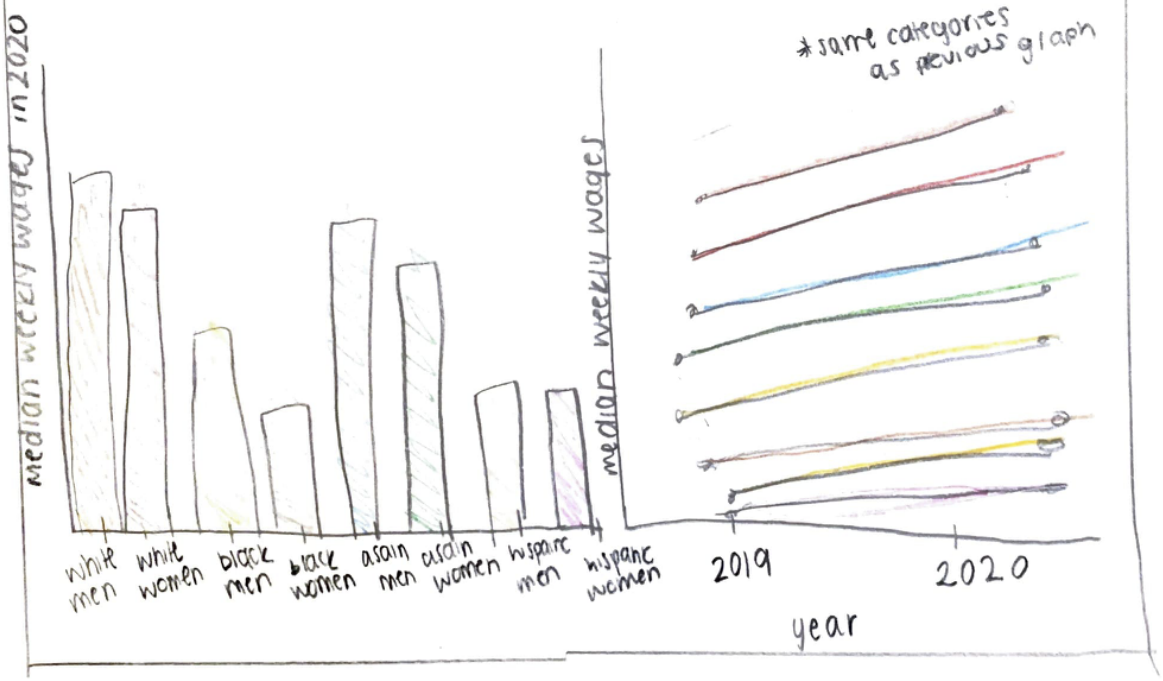
Page 2: General wage Gap



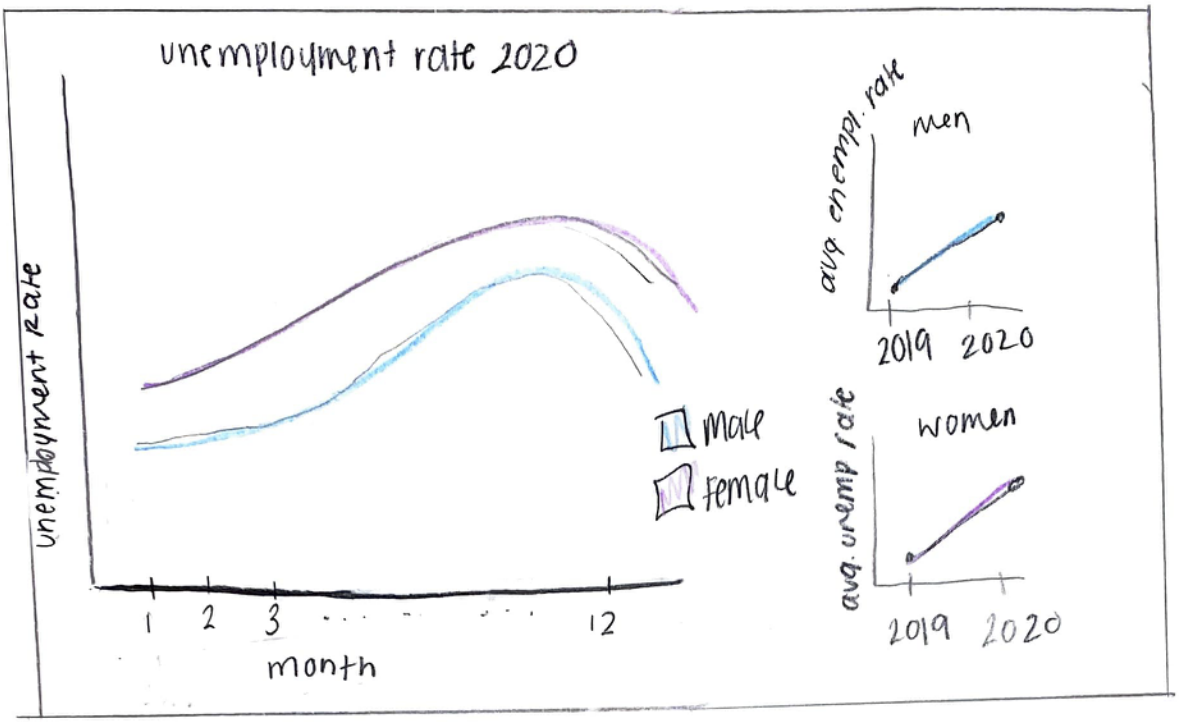
Page 3: Gender wage Gap in Tech



page 4: Race and Gender wage Gap



page 5: COVID impact on ^gender unemployment Gap



page 6: Resources page

- how to get involved in change?
- unemployment resources
- others

Must Have Features:

This project must have clear visualizations for each category listed above. The project must have the two 4 dimensional graphs I included in my sketch. I think those two graphs answer objective one and two clearly. I also think my project must have labelling for nearly all graphs because the numbers are very close, but in this case a small difference is a big deal.

- Page 2 must include both bar charts because they directly answer objective 1
- Page 3 must include the small multiples because they directly answer objective 2
- Page 4 must include the bar chart because it directly answers objective 3
- Page 5 must include the line chart because it directly answers objective 4

Optional Features:

I may possibly add a heat map for the COVID impact section to show the unemployment rate change over time. I cannot decide if the line graph for the heat map would be more useful. I am thinking about possibly including both.

I am also thinking about adding a section that discusses non-binary people and how they are also included in the wage gap but I had a really hard time finding data on this. I will keep searching and possibly include this in my final project.

Project Schedule:

Deadline	Tasks
March 31st	Proposal Due
April 7th	• Make any advised changes to proposal • **Submit final proposal**
April 9th	• Create the general wage gap section • Create cover page
April 11th	• Create tech wage gap page • Create race and gender wage gap page
April 12th	• Create COVID impact page • **ALPHA release due**
April 12th-April 26th	• Make all changes necessary • Edit

April 26th	• **Beta release due**
April 26th- May 10th	• Edit
May 10th	• **Final Presentation**

Related Work:

"Gender Pay Gap in Tech - Storytelling with Data Viz." *Jency Francis*, www.jencyfrancis.com/storytelling-with-data-viz.

Iversen, Katja. "7 Data Visualizations That Opened the World's Eyes to Gender Inequality." *Medium*, Medium, 24 Aug. 2017, medium.com/@Katja_Iversen/7-data-visualizations-that-opened-the-worlds-eyes-to-gender-inequality-75ee03b60589.

Meara, Katie, et al. "The Gender Pay Gap in the USA: a Matching Study." *Journal of Population Economics*, Springer Berlin Heidelberg, 5 Sept. 2019, link.springer.com/article/10.1007/s00148-019-00743-8.

"Exploring Social Issues Through Public Data: The Gender Wage Gap." *Enigma*, enigma.com/blog/post/exploring-social-issues-through-public-data-the-gender-wage-gap.

Blau, Francine. *THE GENDER WAGE GAP: EXTENT, TRENDS, AND EXPLANATIONS*, www.nber.org/system/files/working_papers/w21913/w21913.pdf.

Link to Project Website:

- https://cassidynewberry10.github.io/