Visualizing the Gender Gap in Employment and Education

Sarah Pursley[1], Erin Szeto[1], Madeleine Thompson[1], and Erica Xia[1]

University of Southern California, Los Angeles CA

Abstract. Gender discrimination is an issue that has plagued modern economies, and although female participation in the workforce has been climbing, there are still many issues left to be addressed, especially outside of the US and in specific disciplines such as STEM. This project aims to create an interactive infographic that allows users to explore data documenting the gender gap in employment and in education. The resulting website and infographic can be used by everyone curious about learning more about how this gender gap affects women across the world.

Keywords: Gender Disparity · Pay · Education.

1 Introduction

Women have been fighting for gender equality for centuries. Although progress has been made in this area in recent decades, gender disparities continue to exist in employment and education. For instance, women earn considerably less than men, indicated by the gender wage gap, a measure that reports the difference between earnings of men and earning of women relative to earnings of men [11].

Achieving gender equality is fundamental to human rights and prosperity around the world. Visualizing gender differences will allow viewers to understand the depth of this issue and drill down to specific details. The website and infographic created in this project are designed to be used by the general public as well as men and women in STEM to help call attention to gender discrepancies in the workforce and education and bring awareness to this continuing issue.

In this report, we will discuss the visualizations of related work in section 2, describe the datasets used in the infographic in section 3, explain the design process in section 4, display the main features of the application in section 5, and present the conclusions in section 6.

2 Related Work

There have been many efforts to visualize the gender gap in terms of economic opportunities, employment, and education. Because of the amount and variety of data available, visualization styles used in these efforts can vary widely and can often be quite unique. In our work, we attempted to emulate the work that

has been already been done and expand upon the insights that have already been gathered from previous work.

One such effort, a paper by Christina Boll and Andreas Lagemann [9] focuses on visualizing the gender pay gap in the European Union. This paper looks at the gender differences in wages across the EU and creates an adjusted and unadjusted pay gap metric with which to compare between countries. Boll and Lagemann use bar charts and scatterplots in their paper to visualize this gap.

Helen Shen's feature in Nature [12] discusses the gender gap and how it affects women in science, from applying to and completing graduate school to choosing a career to getting funding for research. This work uses a static infographic with various decorative elements to visualize the gender gap.

The No Ceilings project [8] by the initiative of the Clinton and Gates Foundations assesses the data-based evidence on both improvements and remaining gaps for women and girls' education and economic indicators over twenty years. This project gathered and analyzed data from multiple sources to visualize key indicators and trends on an interactive website and in report form.

Our work builds off of this previous work and integrates the ideas of these authors into an interactive infographic. This approach is similar to the interactive infographics presented in the No Ceilings project, allowing users to explore data from aspects such as womens' education and legal rights, but also integrates other data sources and visualization formats.

3   Data

Four datasets were used in the creation of the infographic: Organisation for Economic Co-operation and Development (OECD) data, No Ceilings dataset, Institute for Women's Policy Research (IWPR) Status of Women in the States data, and the FiveThirtyEight college majors data. Descriptions of each dataset are shown in Table 1.

The OECD data [4] needed to be preprocessed in Python using Pandas. Two OECD datasets were combined by year and country and countries that had 5 or more years of gender wage gap data were selected for the resulting preprocessed OECD dataset.

The No Ceilings dataset [7] was also preprocessed using Python and Pandas. The data sourced from the No Ceilings Project came from the International Labour Organization (ILO). The wage data had 2,910 total observations for 2 genders and 85 countries. For the circle packing chart, 22 separate employment data files for employment by sector and by gender were combined into one dataset. Since the circle packing chart required the data to be in hierarchical nested JSON format, the JSON package was used to reformat the data. The employment data had a total of 686 observations for 35 countries.

The IWPR Status of Women in the States data [5] was preprocessed into a geojson format for use in Mapbox. The data was aggregated over each state in the United States and separated by gender. Out of 17 unique indicators, the indicator for median annual income for women employed full-time, year round by

state was used for the Mapbox chloropleth map. This indicator had an average value of $37,388 and a standard deviation of $5,680.

The FiveThirtyEight College Major data [6] documents the representation of women in different college majors and the economic prospects of students in those respective majors. This data was collected from the American Community Survey 2010 - 2012 Public Use Microdata Series by FiveThirtyEight and compiled into several csv files. For the radial bar chart, the share of women in each major indicator was used, which had an average value of 0.522 and a standard deviation of 0.231.

Table 1. Description of datasets used in infographic.

| Data Source	Description	Statistics
OECD [4]	Gender wage gap at median, 1st decile, and 9th decile earnings, labor force participation rate, GDP ($USD/capita), and female board seats data	1,225 observations, 25 unique countries, timeframe is from years 2000, 2005, 2010, and 2015-2019
No Ceilings [7]	Monthly wages for female and male employees in local currency units, Employment participation rates per sector for female and male employees by country	2910 observations for time series chart, 686 observations for circle packing chart, 35 unique countries, timeframe for the time series chart is 1995-2011, year for the circle packing chart is 2012
IWPR [5]	Median annual earnings for women employed full time, earnings ratio between men and women, and labor participation rates, all specific to the Unites States and aggregated to the state level	17 unique indicators for each of the 50 states and the District of Columbia, only one indicator (median annual income) was used ($\bar{x} = \$37,388$, $\sigma = \$5,860$), timeframe is 2013
College Majors [6]	List of the percent of women in college majors and employment statistics such as unemployment rate and median income for graduates of those majors	18 indicators for 173 observations (majors). Used the percent women statistic ($\bar{x} = 0.522$, $\sigma = 0.231$), timeframe is 2010-2012

4 Approach

The website was designed so that it follows a cohesive story of the gender gap in employment and the economy and the gap in education. Thus, two separate sections of the website were created - one page focusing on employment-related visualizations, and one page for education-related visualizations. A sidebar menu allows users to easily navigate between different visualizations. The employment

page gives users an overview of the general trends that characterize the gender gap story from aspects of both salary and employment by sector. In order to choose the most effective visualization types, we considered the type of data being displayed. For example, we used a line plot to display trends over time. We also added additional links to certain visualizations allowing the user to explore the data in greater depth. The interactive elements allow users to explore the data in detail and use different visual queries to understand the data being presented.

5   System

The base application for the infographic was built using Vue.js [3]. To format the application, Bootstrap-Vue [2] and Bootstrap [1] templates were utilized. In order to create and implement the visualizations for the website, we used D3.js [10].

One such visualization (see Figure 1) is an interactive bubble chart included in the education page of the website. This visualization highlights the share of women in different STEM fields and median salary of each of those fields. Another visualization (see Figure 2) is an interactive chloropleth map included in the employment page of the website. This visualization allows the user to see how the gender wage gap for certain countries changes over time and allows users to drill down by country for additional data about the gender gap in that country.
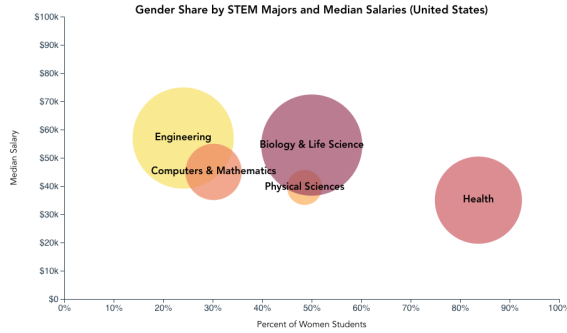
Fig. 1. An interactive bubble chart

6   Conclusion

In conclusion, the website that was created fulfills its purpose as a tool for users to explore data documenting the gender gap. If time permitted, further work
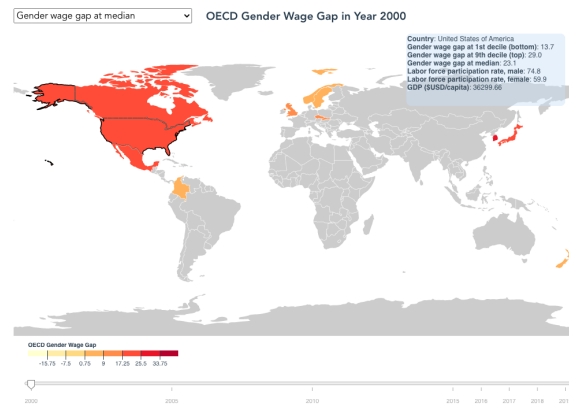
Fig. 2. An interactive chloropleth map

on this website would consist of the addition of more visualizations utilizing the data that has already been collected and used in this website, as well as potentially integrating more data sources into the infographic.

Each member of the team - Sarah, Erin, Madeleine, and Erica - contributed to making at least one visualization for the website. In addition to these shared contributions, Erin and Erica worked on designing and implementing the structure of the website and creating a narrative for the website. Sarah worked on creating additional visualizations for the education page. Madeleine created and gave the final presentation and finalized the paper.

References

1. Bootstrap, https://getbootstrap.com/, last accessed 2 Dec 2021
2. Bootstrap-vue, https://bootstrap-vue.org/, last accessed 2 Dec 2021
3. Vue.js, https://vuejs.org/, last accessed 2 Dec 2021
4. Gender wage gap (2011), https://data.oecd.org/earnwage/gender-wage-gap.htm, last accessed 29 Nov 2021
5. Status of women in the states (2013), https://statusofwomendata.org/explore-the-data/state-data/united-states/, last accessed 29 Nov 2021
6. College majors (2014), https://github.com/fivethirtyeight/data/tree/master/college-majors, last accessed 29 Nov 2021
7. Data for no ceilings: The full participation project (2015), http://www.noceilings.org/data, last accessed 29 Nov 2021
8. No ceilings: The full participation report (2015)
9. Boll, C., Lagemann, A.: The gender pay gap in eu countries–new evidence based on eu-ses 2014 data. Intereconomics **54**(2), 101–105 (2019)
10. Bostock, M.: Data-driven documents - d3.js (2014)
11. Nikolka, T.: The gender wage gap in oecd countries. CESifo DICE Report **11**(1), 69–71 (2013)
12. Shen, H.: Inequality quantified: Mind the gender gap. Nature News **495**(7439), 22 (2013)