INST627: Suarez Final Project Report

Over the course of the semester, I have focused my efforts on analyzing gender pay disparities in the United States since roughly 1980. There were several factors contributing to my motivation behind this topic. First is the idea of fundamental equality and fairness, especially the notion that everyone should be compensated equally for equal work. In 2021, the U.S. Census Bureau reported that the female-to-male earnings ratio was 0.837. Another way to interpret this finding is that women earn roughly 84 cents for every dollar men earn (Guzman & Kolar, 2023). There are also economic factors at play, including potential missed opportunities for economic growth by stifling the financial well-being of women. Implications associated with closing the gender pay gap include promoting workforce diversity, reducing poverty, ensuring women receive fair compensation for educational achievement, and, overall, empowering women to pursue careers without fear of facing discrimination in their compensation.

My project aims to provide a historical perspective on the issue, identify potential explanations for this longstanding discrepancy, and visualizations that help others understand the severity of the wage gap. This report will walk the reader through each question step by step, explaining where the data was collected, how I formatted and prepared it for analysis, and my results and insights from these efforts.

The first of my three questions was, how does the gender pay disparity vary by variables such as race, age, and education? For the race variable of this question, I used gender wage gap data from “The State of Working America Data Library,” created by the Economic Policy Institute. This dataset, spanning from 1973 to 2022, contains average and median wages for wage and salary workers ages 16 and older. It also sorted the data by race, explicitly focusing on white, Black, and Hispanic workers. For age, I utilized data from the U.S. Bureau of Labor Statistics (BLS), providing the inflation-adjusted median of usual weekly earnings by age groups for full-time wage and salary workers between 1979 and 2020. To ascertain the impact of educational attainment on the gender wage gap, I worked with more BLS data. This dataset encompasses inflation-adjusted median usual weekly earnings, by academic achievement, for full-time wage and salary workers 25 years and older between 1979 and 2020.

Before working with the race data, I established a historical baseline, which included calculating the difference between the 1973 and 2022 median and average hourly wages by men and women. Based on this data, since 1973, men have seen their average salary increase by $8.86 and women by $10.55. Men's median hourly salary difference over the years is less dramatic. Male median hourly wages increased by $0.91, while women increased by $5.64. The figures demonstrate that women saw a more significant increase in average and median hourly wages than men, but it did little to alleviate the disparity. For example, men in 1973 had an average hourly salary of $26.96, while women had an

average of $17.31, for a difference of $9.65. In 2022, men had an hourly average wage of $35.82, while women earned $27.86, for a difference of $7.96. These findings further establish that while hourly wages have increased over time, we have observed inconsequential changes to curb the wage disparity.

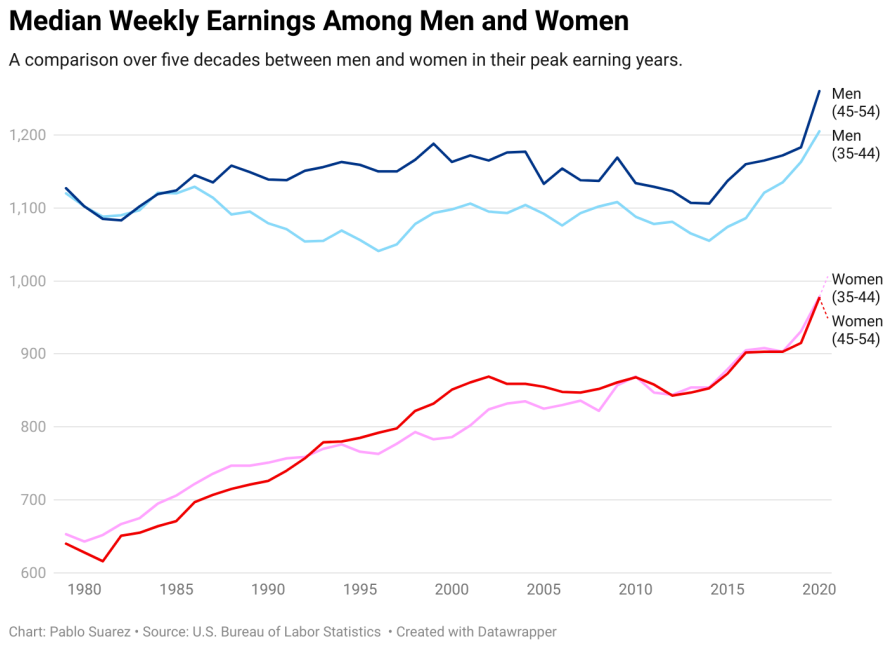
I then applied a similar workflow to the race data. Looking at wages among white men and women, I found that between 1973 and 2022, the average hourly wages of white men increased by $11.17 and $11.93 for white women. The hourly wage gap between white men and women in 1973 was $10.36, now $9.60, a difference of $0.76. The difference between white women and men in 2022, or the amount white women earn less than men, is -$6.32. An interesting finding from this section is that the difference between white women and women overall was $1.64, signifying pay disparities even among women based on race.

Concerning Black men and women, between 1973 and 2022, the average hourly wages of Black men increased by $6.34 and $8.16 for women. The hourly wage gap between Black men and women in 1973 was $5.26 and is now $3.44, a difference of $1.82 and greater than what I observed between white men and women. The difference between Black women and men in 2022 is -$11.83. When I tallied the difference between Black women and the rest of their genders, I found that they earned $3.87 less overall.

For Hispanic men and women, the average hourly wages of the former have increased by $5.96 and by $7.21 for the latter. In 1973, the hourly wage gap between Hispanic men and women was $5.28. That figure is now $4.03, a difference of \$1.25, and the lowest difference among the races observed in the data. The difference between Hispanic women and men overall in 2022 was the most prominent difference I saw in this data, -$13.22. Also, the difference between Hispanic women and women overall equaled -$5.26.

Part two of this first question is where I examined age as a variable. I began work on this section by reformatting the BLS data frame of weekly earnings mentioned above. I also calculated total earnings each year by gender, just in case I needed it for later analysis, including finding the percentage of an age group's median weekly earnings compared to the total earnings by gender for a specific year. Given the different data points across age groups and years, the data in this section was difficult to visualize. However, according to the BLS, workers earn the most between the ages of 35 and 54. Therefore, I focused my visualization efforts on data points within those age groups. Using Datawrapper, I visualized this data with a line chart for peak earning ages between men and women over 50 years. As indicated in Figure 1, men have seen a slight salary increase as they transition into the older earning age group. While their overall salaries have increased, women do not see that same jump between age groups. Their lines coincide frequently and have almost wholly overlapped in more recent years.

(Figure 1)



The final step of this first question examined how the level of completed education impacts wage earnings. Unfortunately, the BLS dataset of median weekly earnings by educational attainment required reformatting as the original version was not readable in Rstudio due to how the two tables were stacked on top of each other. Instead, I split the tables into separate CSV files and imported them individually. After completing that step and confirming they loaded correctly, I merged them by year. I also renamed the columns because the previous names contained special characters and awkward spacing that were making it difficult to read the data.

Once again, turning to Datawrapper, I took the cleaned dataset and visualized the results. Upon seeing the initial results, I created two distinct line charts. One compared various levels of education between genders, and the other solely compared bachelor's degrees and more. For the first chart (Figure 2), I found that the weekly median pay disparity since the early 1980s has remained between $100 and $300 across various levels of educational attainment. The gap seemingly increases as levels of educational attainment increase. The second chart (Figure 3) helped visualize this trend and showed that since the 1980s, the median weekly earnings between men and women with bachelor's degrees or more have remained close to nearly $400. It is especially strange to see this gap occur at the highest represented level of educational attainment. However, my understanding is that salaries at these levels are much higher, hence the greater disparity as opposed to jobs for people without college degrees, which often have lower wages and a lower earning ceiling.

(Figure 2)

Median Weekly Income by Gender and Level of Completed Education

Pay disparity between genders based on educational attainment since the 1980s has remained between $100-$300 for median weekly income.

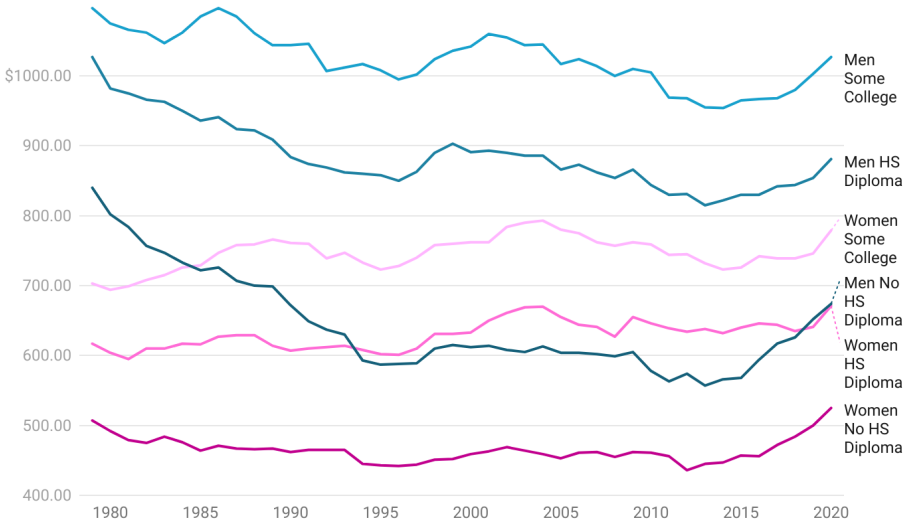


Chart: Pablo Suarez • Source: U.S. Bureau of Labor Statistics • Created with Datawrapper

(Figure 3)

Median Weekly Earnings of Higher Education Degree Earners by Gender

Since the 1980s, the pay disparity between men and women has remained close to nearly $400.

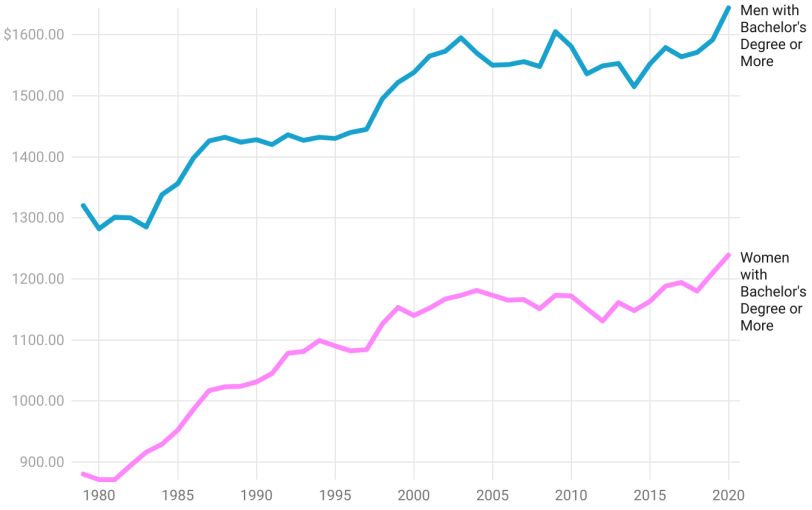


Chart: Pablo Suarez • Source: U.S. Bureau of Labor Statistics • Created with Datawrapper

My second question was, which industries have the most and least significant wage discrepancies? I acquired two data frames from the U.S. Department of Labor to answer this question. One featured sectors and occupations with the most significant wage gaps, while the second featured industries and professions with the most minor wage disparities

between genders. Each data frame measured the wage disparity with women's earnings as a percentage of men's earnings. I then merged the two data frames by the occupation column and arranged the results by earnings in descending order.

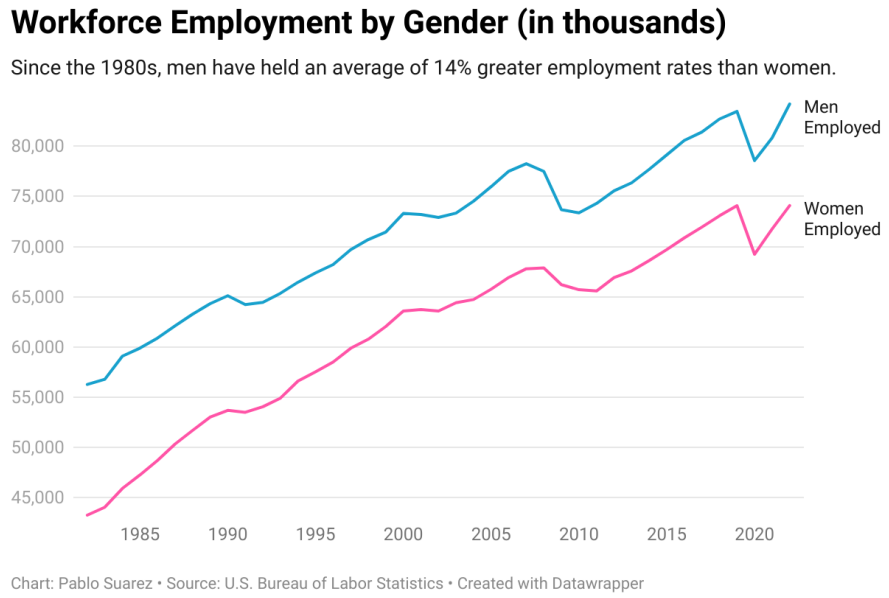
Figure 4 (see references page) visualizes this merged dataset. The top five professions with the most considerable salary disparity were judicial workers – judges, magistrates, administrative positions – financial clerks, title examiners, property appraisers, and sales positions. Conversely, childcare employees, health care social workers, barbers, and administrative positions like executive secretaries and payroll clerks were positions where women earned more than men. Women in childcare occupations earned almost 11% more than their male counterparts in those positions.

Regarding judicial workers, women in these positions earned roughly 56% of what their male counterparts in similar positions took home. That figure is an especially stark contrast with the Pew Research Center's finding that in 2022, "women earned an average of 82% of what men earned" based on median hourly earnings of both full- and part-time workers (Aragão, 2023). The earnings discrepancy becomes even more egregious when considering that women earned 80% as much as men in 2002, marking an almost insignificant 2% improvement over two decades. These findings suggest that women working in judicial positions are 26% off of current average earnings from women, which still face an 18% difference from men's earnings.

Finally, my third question was, at what rate do men and women participate in the workforce? To answer this, I called in data from the BLS' Labor Force Statistics from the Current Population Survey, which provides the employment status of the civilian non-institutional population aged 16 years and over by sex, starting in 1982. This BLS data frame was formatted similarly to the education data frame used in question one. Therefore, I had to organize this data into two separate data frames and load each into Rstudio. After confirming that each dataset was loaded properly and accounted for the correct data, I renamed the columns to avoid confusion between the men's and women's values. I merged them by year to create a new data frame named employment status.

From there, I used the mutate function to create an additional column called employed percentage difference, which showed the disparity in employment between genders each year. I found that historically, men have held an average of 14% greater employment rates than women, but the percentage difference in weekly median wages over that time favored men by 21%. This finding suggests that workforce participation is not necessarily to blame for the disparity, although further testing of these variables is needed to reach a more concrete conclusion. This section's visualization used the columns providing the raw count of employed individuals in the U.S. by gender from 1979 to 2020 and measured in thousands. Figure 5 shows the output of these columns.

(Figure 5)



Upon completing the three primary research questions, I created a new section in the markdown file showing the setup for conducting a deeper statistical analysis of the variables included in the project, including t-tests and linear regression. The ggplot2 and corr (t-test) packages in Rstudio, used for data visualization and analysis, require the data to be stored in one data frame. Therefore, I merged three primary data frames from the project – gender wage gap, education, and employment status – into one titled merged data.

I employed the reduce and intersect functions to merge the data frames, which joined them by the same starting and ending years, 1982 to 2020. While this approach was effective, it ultimately required omitting the data from the post-COVID years of 2021 and 2022. There are methods I could have used to add some of that data back into this new data frame. However, this would result in “NA” values in columns from data frames that did not extend to 2022. Following merging the data frames, I arranged the newly created version by year.

I also created a for loop to convert hourly wages from the data frame used in question one into weekly wages. Creating this loop was crucial because the wage data must share the same parameters to test it properly. The loop takes the names of the columns I wanted to convert. In this case, it was median hourly wages to weekly wages, as the latter was abundant across the other data frames. Then, it renamed the columns to “weekly” versions and multiplied their values by 40 because we can assume employees worked a 40-hour work week. Using the merged data frame, I added a few examples of how linear regression models and t-tests would work with this data at the bottom of the markdown file. One example tests variables such as men's weekly median earnings and women's earnings after obtaining a bachelor's degree or greater.

During the final project video presentation, I shared how I would proceed to work on this project moving forward. One potential future output highlights significant policies or economic events that have contributed to or shaped these wage trends. Visualizing this would include a timeline and contrasting those milestones with wage values. Another option is creating an interactive map of the U.S. showing the difference in annual earnings between genders by state.

In conclusion, this project examines gender pay disparities in the United States from 1980 to the present. It reveals long-standing wage gaps and the impact of variables, including race, age, education, and industries. Analysis conducted during the project revealed distinct patterns for white, Black, and Hispanic individuals, underscoring the intersectionality of race within the broader issue. It also shares critical insights about the issue's presence across industries, with professions like judicial workers and financial clerks exhibiting substantial pay gaps. Conversely, positions such as childcare employees challenge norms, showcasing instances where women out-earn men.

The workforce participation analysis indicates a historical employment rate disparity favoring men, but the magnitude of the gender pay gap exceeds this differential, suggesting broader factors contribute to earnings disparities. The practical implications of this research can hopefully inform policymakers and lead to organizational practices aimed at closing the gender pay gap. Future research on this topic may support these goals by exploring the impact of specific policies or economic events on wage trends. It may also benefit from a greater number of interactive visualizations to enhance its depth and impact.

Above all, this research demonstrates the importance of advocating for women and remaining cognizant of systemic barriers that prevent them from fully participating in the workforce. The research also reiterates that women deserve to be compensated equally to men when working in the same industries and providing the same output level. Addressing this issue involves aligning with fairness and contributing to a diverse, inclusive, and economically robust society.

References

Aragão, C. (2023, March 1). Gender pay gap in U.S. hasn't changed much in two decades. Pew Research Center. https://www.pewresearch.org/short-reads/2023/03/01/gender-pay-gap-facts/

Economic Policy Institute. (n.d.). Data Library. The State of Working America Data Library: Gender Wage Gap. https://www.epi.org/data/#/cpsorg?subject=wagegap-mf

"Guzman, G., & Kolar, M., Income in the United States: 2022 10 (2023)."

U.S. Bureau of Labor Statistics. (2021, September 1). Earnings of full-time workers. U.S. Bureau of Labor Statistics, Table 17. https://www.bls.gov/opub/reports/womens-earnings/2020/

U.S. Bureau of Labor Statistics. (2021a, September 1). Earnings of full-time workers. U.S. Bureau of Labor Statistics, Table 19. https://www.bls.gov/opub/reports/womens-earnings/2020/#table-19

U.S. Bureau of Labor Statistics. (n.d.). Employment status of the civilian noninstitutional population 16 years and over by Sex, last 40 years. U.S. Bureau of Labor Statistics. https://www.bls.gov/cps/cpsaat02.htm

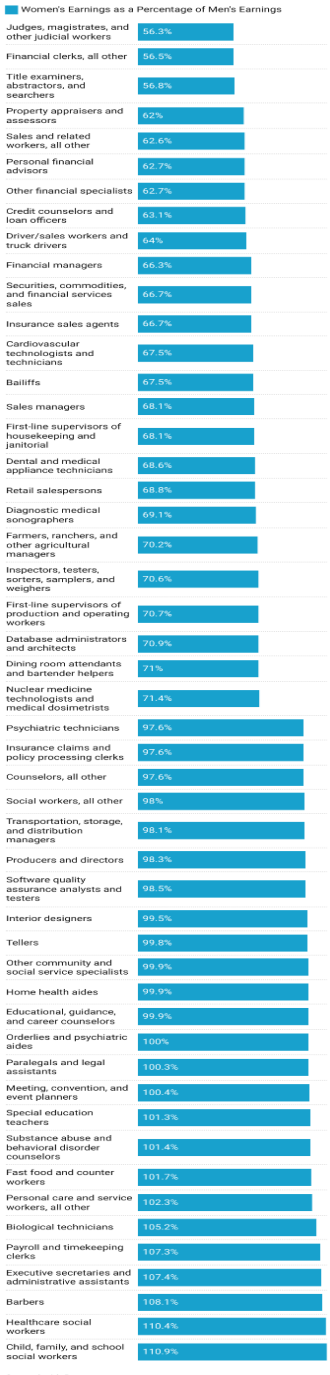
U.S. Bureau of Labor Statistics. (n.d.-b). Median weekly earnings increased 5.7 percent over the year ending in the second quarter of 2023. U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/ted/2023/median-weekly-earnings-increased-5-7-percent-over-the-year-ending-in-the-second-quarter-of-2023.htm#:~:text=%E2%80%8B%20Source%3A%20U.S.%20Bureau%20of,End%20of%20interactive%20chart.&text=Usual%20weekly%20earnings%20were%20highest,workers%20ages%2055%20to%2064.

U.S. Department of Labor. (n.d.-a). Occupations with the largest gender earnings gap. DOL. https://www.dol.gov/agencies/wb/data/earnings/occupations-largest-gender-gap

U.S. Department of Labor. (n.d.). Occupations with the smallest gender earnings gap. DOL. https://www.dol.gov/agencies/wb/data/earnings/smallest-gender-earnings-gap

Link to Final Presentation: https://www.canva.com/design/DAF2s1xrnhw/8iRwMnvEen7ZBTupZPdOBw/edit?utm_content=DAF2s1xrnhw&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton

(Figure 4)


Women's Earnings as a Percentage of Men's Earnings
Judges, magistrates, and other judicial workers
56.3%
Financial clerks, all other
56.5%
Title examiners, abstractors, and searchers
56.8%
Property appraisers and assessors
62%
Sales and related workers, all other
62.6%
Personal financial advisors
62.7%
Other financial specialists
62.7%
Credit counselors and loan officers
63.1%
Driver/sales workers and truck drivers
64%
Financial managers
66.3%
Securities, commodities, and financial services sales
66.7%
Insurance sales agents
66.7%
Cardiovascular technologists and technicians
67.5%
Bailiffs
67.5%
Sales managers
68.1%
First-line supervisors of housekeeping and janitorial
68.1%
Dental and medical appliance technicians
68.6%
Retail salespersons
68.8%
Diagnostic medical sonographers
69.1%
Farmers, ranchers, and other agricultural managers
70.2%
Inspectors, testers, sorters, samplers, and weighers
70.6%
First-line supervisors of production and operating workers
70.7%
Database administrators and architects
70.9%
Dining room attendants and bartender helpers
71%
Nuclear medicine technologists and medical dosimetrists
71.4%
Psychiatric technicians
97.6%
Insurance claims and policy processing clerks
97.6%
Counselors, all other
97.6%
Social workers, all other
98%
Transportation, storage, and distribution managers
98.1%
Producers and directors
98.3%
Software quality assurance analysts and testers
98.5%
Interior designers
99.5%
Tellers
99.8%
Other community and social service specialists
99.9%
Home health aides
99.9%
Educational, guidance, and career counselors
99.9%
Orderlies and psychiatric aides
100%
Paralegals and legal assistants
100.3%
Meeting, convention, and event planners
100.4%
Special education teachers
101.3%
Substance abuse and behavioral disorder counselors
101.4%
Fast food and counter workers
101.7%
Personal care and service workers, all other
102.3%
Biological technicians
105.2%
Payroll and timekeeping clerks
107.3%
Executive secretaries and administrative assistants
107.4%
Barbers
108.1%
Healthcare social workers
110.4%
Child, family, and school social workers
110.9%
Created with Datawrapper