### From Bad to Good Graph: Transforming Data Visualization

### Enhancing Insights: A Journey Through Effective Data Presentation

Akhila Kudupudi  
*Data Analytic Engineering*   
*George Mason University*Fairfax, VA, USA  
akudupud@gmu.edu

*Abstract*— This paper explores the transformation of a data visualization focusing on the gender pay gap across U.S. industries. The original graph, characterized by its complexity and lack of clarity, undergoes a redesign in Tableau to enhance interpretability and accessibility. Key improvements include the introduction of horizontal bar charts, aligned baselines, enhanced color contrast, and interactive tooltips, achieving greater clarity and a more intuitive understanding of the data.

Keywords— Data Visualization, Gender Pay Gap, Tableau Redesign, Interactive Tooltips, US Census Bureau Data, Pay Disparity, Industry Comparison, Visualization Techniques, Accessible Information, Analytical Methods.

# Introduction

The gender pay gap persists across all sectors of the U.S. economy, with disparities varying by industry. This paper delves into the magnitude of wage differences between men and women, using data from the US Census Bureau to highlight median earnings in 2019. With finance and insurance showing the largest gap, where men earn significantly more than women, the paper aims to provide a detailed analysis of gender income inequality. This introduction sets the stage for a comprehensive examination of the factors contributing to the wage gap and explores potential solutions to bridge this divide.

# PROBLEM STATEMWENT

The problem this paper tackles is the ineffective communication of data due to poorly designed visualizations, focusing specifically on enhancing a graph that depicts the gender pay gap across various industries. The goal is to refine this visualization to ensure it clearly communicates the underlying data, making it easily interpretable and engaging for the audience. This transformation highlights the importance of good design principles in data visualization and aims to set a standard for how complex data, like the gender pay gap, can be presented more effectively.

# PURPOSE

## The purpose of this paper is to showcase the transformation of a poorly constructed graph into a more effective and insightful visualization, specifically focusing on the gender pay gap across different industries. By applying best practices in data visualization, this study aims to demonstrate the impact of clear and intuitive graph design on the understanding of complex issues like wage disparities. The ultimate goal is to highlight the significance of visual clarity in data presentation for both informing the public and guiding policy discussions.

# Methodology

The methodology section of your paper should detail the process used to transform the poorly designed graph into a more effective visualization. This involves several steps:

## Data Source Identification

Retrieved reliable gender pay gap data from the US Census Bureau.

## Data Preparation

Conducted thorough data cleaning, filtering, and transformation for visualization readiness.

## Data Visualization Design

Implemented essential design principles to enhance visual clarity, choosing suitable graph types, color schemes, and layouts.

## Tool Utilization

Employed Tableau for visualization, applying specific settings to improve the presentation.

## Feedback Integration

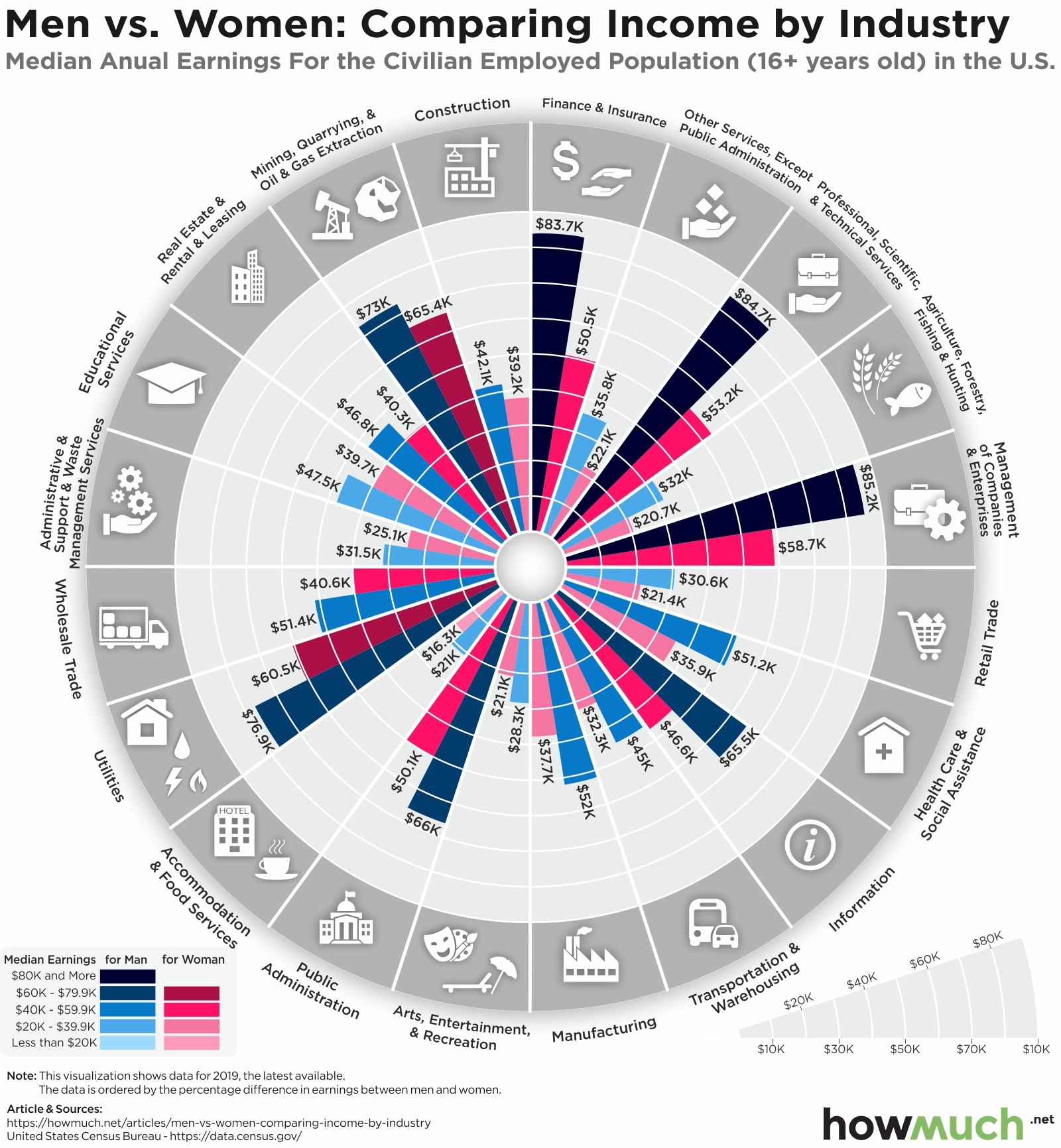
Incorporated peer feedback into the visualization process, refining it for greater accuracy and impact.

# Analysis

The analysis section of the paper examines the transformation process from a poorly designed graph to a more effective visualization of the gender pay gap across U.S. industries. The original graph's weaknesses, such as design complexity and unclear scale, were identified. Utilizing Tableau, significant improvements were made, including adopting a horizontal bar chart format, aligning baselines for easier comparison, enhancing color contrast for better accessibility, and incorporating interactive tooltips for detailed information. These changes resulted in a visualization that not only presented the data more clearly but also facilitated a more intuitive understanding of the gender pay gap's severity across different sectors, highlighting the most and least affected industries.

## The Bad Graph

Figure 1 presents a radial chart aiming to depict the median annual earnings for men and women across various industries. While the intention is to provide a comparative view, the complexity of the graph’s design detracts from its clarity and interpretability. The overlapping circular segments representing men and women’s earnings are not immediately intuitive, requiring readers to spend additional effort to understand the disparities. The similarity in color choices for the segments corresponding to men and women further complicates the distinction between the data sets. Furthermore, the graph's scale is not consistent, and the figures for median earnings are small and challenging to read, hindering quick comparisons. The visualization also falls short in effectively communicating the extent of the income disparities, as the radial layout makes it difficult to compare lengths and areas accurately. While comprehensive in data, the graph’s structure requires a redesign to aid in the straightforward conveyance of the significant information it contains.



1. The original radial graph: Radial Comparison of Median Annual Earnings by Industry and Gender.

## Data Visualizations

The redesigned graph in the image is a dual-axis bar chart, created to display the median annual earnings of men and women by industry, along with a representation of women's earnings as a percentage of men's earnings.

The visualizations were created using Tableau:

A graph of women's earnings

Description automatically generated

1. The redesigned graph: Bar Chart Analysis of Gender Earnings Gap by Industry.

Here are some observations and potential areas of improvement:

## Clearer Comparison

This design allows for a direct comparison between men’s and women’s earnings within each industry. It is easier to compare side-by-side bars than the segments of a circle.

## Dual-Axis Representation

The dual-axis approach, with absolute values on one side and percentages on the other, is helpful for understanding both the actual earnings gap and the relative earnings ratio. However, this can sometimes be confusing if the viewer doesn't notice the two different scales on the left and right.

## Color Coding

The use of blue for men and red for women is a common convention and helps in differentiating the data points quickly.

## Readable Labels

The industry labels and values are more readable in this design compared to the circular chart, making the data more accessible.

## Percentage Calculation

Having the percentage shown directly on the graph provides immediate insight into the relative differences in earnings.

## Scale Issue

It seems there might be an issue with the scale or alignment of the bars and percentages. The bars for men's earnings and the corresponding percentages for women's earnings don’t seem to align properly. For example, the longest bar, which should represent the highest earnings for men, does not align with the highest percentage for women.

## Overall Design

While the design is an improvement over the radial chart for data comparison, ensuring that scales and axes are correctly aligned and that the percentages correctly represent the data they correspond to is crucial for accurate interpretation.

## Legibility

Some labels and numbers are cut off or overlap, which can make the chart hard to read. It would be beneficial to ensure all text is fully visible and clearly separated from other elements.

# Interpretations

The redesigned bar chart provides a visual comparison of the median annual earnings between men and women across various industries for the civilian employed population aged 16 and above in the United States. Each pair of horizontal bars represents one industry, with the Right bar indicating men's median earnings (in blue) and the Left bar showing women's median earnings (in brown). A secondary axis on the right displays women's earnings as a percentage of men's earnings, which is depicted by the length of the red bars with sequential coloring.

From the chart, we can observe that in every industry, men's median earnings surpass those of women, reflecting a consistent gender pay gap. The industries where this gap is most pronounced are 'Management of companies and enterprises' and 'Professional, scientific, and technical services', where women's earnings are significantly lower than men's when considering both absolute dollar amounts and percentages. Conversely, industries such as 'Accommodation and food services' and 'Arts, entertainment, and recreation' show a comparatively smaller gap, indicating a closer parity in earnings between genders.

The percentage values provide additional insight into the extent of the earnings disparity. For example, in 'Health care and social assistance', women earn approximately 93.17% of what men earn, suggesting a narrower gap. In contrast, the 'Construction' industry shows a wider gap, with women earning only about 60.31% of men's earnings.

It is important to note that while the chart effectively highlights disparities in earnings between men and women, it does not explain these differences. Factors such as hours worked, occupation within industries, education levels, work experience, and discrimination could contribute to the observed discrepancies and are worth examining in further research.

Overall, the redesigned bar chart effectively presents the median earnings data by industry and gender, revealing clear patterns of income inequality across different sectors of the economy.

# Discussion

While the redesigned graph significantly improves the understanding of the gender pay gap, it doesn't address all issues. Data collection methods and potential biases might limit the complete representation of wage disparities. Future research could benefit from a more nuanced approach that considers intersectionality, including how race and age intersect with gender to affect earnings. This broader perspective could unveil deeper insights into the complexities of pay gaps, fostering more targeted and effective solutions for achieving wage equality.

# Conclusion

The transformation of the graph in this study illuminates the gender pay gap across various industries with newfound clarity, providing a more accessible and insightful view into a persistent societal issue. While the redesigned visualization marks a significant improvement in data presentation, it also opens the door to exploring deeper aspects of wage disparities, emphasizing the need for a comprehensive approach that considers intersectional factors. This work underscores the critical role of effective data visualization in fostering understanding and dialogue around complex issues, advocating for continuous improvement in both data representation and equality initiatives.

##### Acknowledgment

I am grateful to my instructor, Richard Sigman for his invaluable guidance and to my peers at the Department of Data Analytics Engineering for their insights. Special thanks to the US Census Bureau for providing the crucial data that formed the foundation of this study. Their resources were indispensable in conducting a comprehensive analysis. Additionally, the constructive criticism received from peer reviews significantly enhanced the quality of the work, making this research more robust and well-rounded.

##### References

1. HowMuch. (n.d.). Men vs. women: Comparing income by industry. HowMuch, from <https://howmuch.net/articles/men-vs-women-comparing-income-by-industry>
2. Data Visualization Techniques for Effective Data Analysis. Medium. <https://medium.com/@analyticsemergingindia/data-visualization-techniques-for-effective-data-analysis-4e3d6f635ff9>
3. Tableau. (n.d.). Viz in Tooltip. Retrieved, from <https://help.tableau.com/current/pro/desktop/en-us/viz_in_tooltip.htm>

Appendices

**Appendix A: Tableau Specifications/Screenshots**

To create a plot like this in Tableau, I followed these steps:

* **Connect to Data in Tableau:** Start Tableau and connect to the data source.
* **Drag Fields to Columns and Rows:** Drag 'Industry' field to the Rows shelf and the 'Median Earnings' fields for men and women to the Columns shelf. (opt for Dual axis graph)
* **Choose the Right Chart Type:** For a bar chart like the one shown, use the 'Bar' mark type.
* **Color Code the Bars:** Drag the 'Median Earnings' fields for men and women to their respective Marks card to differentiate men's (blue) and women's(brown) earnings by color.
* **Drag Fields to Columns and Rows:** Drag the percentage of women’s earnings relative to men's fields to the Columns shelf.
* **Color Code the Bars:** Drag the percentage of women’s earnings relative to men's fields to Marks card.(Red)
* **Add the Percentage Calculation:** Drag the calculated field for the percentage of women’s earnings relative to men's to the Label shelf on the Marks card to show the values on the bars.

To add annotations to the plot in Tableau:

* **Select the Mark:** Click on the specific data point or bar you want to annotate.
* **Annotate Option:** Right-click (control-click on Mac) to open the context menu and select "Annotate."
* **Edit Annotation:** In the annotation dialog box, type your text. You can include dynamic fields that update with changes in the underlying data.
* **Format Annotation:** Adjust the font, size, color, and alignment as needed for clarity and emphasis.
* **Finalize:** Click 'OK' to place the annotation on your chart.
* **Finalize Visualization:** Add titles, adjust fonts and colors, format axes, and make sure all labels are clear and readable.
* **Create a Dashboard:** Combine the elements into a dashboard and arrange them for the final presentation.

Bad Graph:

A circular chart with many colored circles

Description automatically generated with medium confidence

Redesigned Graph:

A graph of women's earnings

Description automatically generated

**Appendix B: Data Sets**

* Data Source: Source: howmuch
* Authors: Irena
* Publication years: 4 March 2021
* Data Source: US Census Bureau
* URL: <https://howmuch.net/articles/men-vs-women-comparing-income-by-industry>

Data Structure:

The data set includes several columns: Serial Number (SNO), Industry, Median Earnings for Men and Women, and Women's Earnings as a Percentage of Men's Earnings. Each row represents an industry and contains data points for both genders' earnings and the calculated percentage that highlights the pay gap. This data can be analyzed to understand the gender pay disparity across different industries.

Actual Data:

A screenshot of a chart

Description automatically generated

Cleaned data:

A table with numbers and text

Description automatically generated