Project Proposal: Wage Gap Analysis

1. List of data sources

Our plan is to utilize the Bureau of Labor Statistics website to gather our data. The data that we plan on gathering is Salaries broken down in the following categories: locations at the state level, occupation, and gender.

https://www.bls.gov

2. Inspiration!

For decades we have always heard that women are often paid lower than a man for the same job. Most companies say they are working on building an inclusive organization and focused on promoting women and minorities. We wanted to see if we could analyze the wage gap trends. Our inspiration project we are using is from Kaggle, Exploring the Wage Gap. We are also using some Tableau visualizations as an inspiration.

https://www.kaggle.com/drgilermo/exploring-the-wage-gap/notebook
https://public.tableau.com/app/profile/the.elearning.guild/viz/2018GenderPayGap552px/Gender552
https://public.tableau.com/app/profile/poojagandhi/viz/GenderPayGap-Australia/GenderPayGap-Australia

We do not want to do the stereotypical blue and pink color scheme for our data set. We found the following webpage shows that most media companies also shy away from the blue/pink and that there are no set standard colors used in the industry. We plan on using teal for men and purple for women.

https://blog.datawrapper.de/gendercolor/

3. Project Structure Breakdown

With this project we would like to answer a few questions such as:

*   Has the gender wage gap reduced over the years?
*   Are there certain age groups the gender wage gap is higher than others?
*   Does occupation affect the gender wage gap?
*   Can we predict what someone's salary might be based on occupation and gender?
*   We predict that our model will be 80% accurate, is that achievable?

Our project will be hosted on Heroku utilizing HTML, Java and Flask. We plan on having a total of 7 pages which will consist of the following:

*   Overview – Project description
*   Visualization Dashboard – Will be filterable

- Salary Predictor – Machine learning with visualization of output
- Data – Data table that is filterable
- Analysis – Overall project results
- About Us – Picture and short write up of team members
- Work Cited – List of any resources we utilized for the project.

4. Roles

Proposal – Ericka
Gather & Clean Data – All
Machine Learning Model – Ericka
Website Full Stack – Kelli
Visualizations:
- Macro Visualizations – Ali
- Micro Visualizations - Kelli
Presentation – Ali
Final Write-up – All