1. Topic

We plan to examine gender bias in online job screenings, plus bias in specific industries of job listings. We want to investigate the possible presence of gender bias in the online job screening process across different industries. We also want to explore how gender-coded language found in job listings and resumes may influence hiring decisions.

2. Motivation

As women in STEM fields, we are acutely aware of the barriers that gender bias, whether subtle or implicit, may impose in hiring, advancement, and representation. The STEM field is a rapidly growing employment sector and one where women remain underrepresented.

Furthermore, as college students, we are beginning to experience the benefits and harms of online job screenings. As individuals who hope to have jobs one day, we are interested in understanding the systems in place today that can both enhance and hinder our work experiences.

When firms use algorithmic systems for screening, such as automated resume parsers, coding tests, or AI-assisted tools, there is a risk that these systems amplify or encode social biases.

Our project, therefore, examines not only whether gender bias exists in online job screening but also how contextual features such as gender-coded words in job descriptions and industry differences influence candidate outcomes. By combining resumes (Kaggle, HireItPeople), screening systems (HackerRank, CodeSignal, ChatGPT), and analysis of job description text, we aim to provide evidence that speaks directly to questions of fairness in hiring algorithms.

3. Research questions

We propose the following research questions:
- Do male and female candidates with similar qualifications have different job screening outcomes?
- Does gender bias in job screening vary across industries (specifically gender-dominate fields)?
- Are gender-coded words (e.g., "collaborative" vs. "assertive") in job descriptions associated with differences in the gender distribution of job applicants?

4. Dataset

These are the two potential datasets of resumes that we will be using for our project. Both datasets include information about applicants' skills, past experiences, and the positions they are applying for. However, neither dataset includes details on applicants' gender or pronouns. Thus, to align with our research focus on gender bias, we would add a randomized "Gender" and "Pronouns" column to simulate applicant gender.

a. *Dataset 1*

The first dataset consists of resume records containing the following variables: an applicant ID, a plain-text version of the resume (Resume_str), an HTML-formatted

version (Resume_html), and the job category (Category) that the applicant is applying for. Each observation in this dataset represents a single resume. To prepare the data for analysis, we will extract structured features from the resume text, number of technical skills, and evidence of leadership experience. We will also add new columns, "Gender" and "Pronoun", which will be randomly assigned to simulate applicant gender, since the original dataset does not contain demographic information.

b. *Dataset 2*
The second dataset will be constructed through web scraping of resumes available on Hire IT People. Each resume entry contains text under sections such as Summary, Technical Skills, and Professional Experience. After scraping, we will preprocess the HTML to plain text and parse these sections to extract key variables. Specifically, we will identify the applicant's years of experience, skills, number of past jobs, and industries worked in. As with the first dataset, we will add a randomized "Gender" and "Pronoun" column to test our research question about bias.

5. Expected Findings
We hypothesize that there will not be a large bias based on gender alone, but algorithms' heavy reliance on other variables, such as gender-coded keywords in specific industries, may reflect implicit biases. We plan to use a regression model to evaluate whether hiring rates are different based on gender. A separate regression can evaluate industry-specific hiring decisions to see if industry-specific hiring preferences reflect gender bias.

6. Analysis of Data
Our data analysis will focus on testing whether outcomes differ by gender. The analysis will proceed in these stages:

a. *Exploratory Data Analysis (EDA)*
We will begin by summarizing the data with descriptive statistics and visualizations. For example, we will examine average years of experience, and skill frequencies across the dataset to ensure our extracted features are consistent. This stage will help us detect anomalies or cleaning issues before formal analysis.

b. *AI-Based Screening and Outcome Generation*
Next, we will use AI-assisted job screening tools to simulate automated hiring assessments. Each resume will be input into these systems to obtain an evaluation outcome such as predicted job fit, or shortlist likelihood. These AI-generated outcomes will serve as our dependent variable, representing how an automated system might rank or classify applicants in a real-world screening process.

c. *Regression/Predictive Modeling*
We will then use machine learning models such as linear regression, logistic regression, and classification algorithms to assess how gender and gender-coded language predict resume outcomes. These models will allow us to estimate the independent effect of

gender while controlling for other resume features such as years of experience, technical skills, and industry. If gender or gendered language remains a significant predictor after controlling for these factors, this would suggest that the screening system encodes or amplifies gender bias.

d. *Textual and Contextual Feature Analysis*

We will analyze how specific word choices and phrasing in job descriptions influence AI screening outcomes. Using natural language processing (NLP) techniques, we will identify gender-coded language and test whether resumes containing these terms systematically receive higher or lower AI evaluation scores based on their association to gender stereotypes, and whether job descriptions that use gendered wording are more likely to favor one gender label over another.

e. *Robustness and Fairness Checks*

Finally, we will perform robustness checks by matching resumes with similar professional attributes (experience, education, skill count) but different assigned genders. Comparing AI evaluation outcomes within these matched pairs will help confirm whether gender-driven differences persist when resumes are otherwise equivalent.

Group Contract and Work Agreement

After discussing our plans and hopes for this project, we (Maya, Anh, and Zoe) have come to the following conclusions:

- We will share most of the work equally (No one person will be responsible for any large section of work)
- We will meet weekly to discuss and work together (We will meet in person and collaborate on the work we have not done alone)
- We will focus on pair programming (We will monitor our distribution of time and energy of programming. Every 15-20 minutes we will rotate who is coding and who is "navigating")
- We will have open communication through in-person conversations, text, email (Communication about planning to meet, questions, and other information will be done clearly and efficiently.)
- We will meet before each deadline to agree on the work we submit (Before each deadline we will meet in person to finish and review the work that will be submitted, to ensure that everyone is satisfied with the results.)
- If conflict arises, we will talk it out first and if a resolution is not reached we will go to the professor (We will speak up about any miscommunications and concerns about how the project is going. If there is no clear resolution, we will get outside guidance.)

These agreements set our expectations of each other, the group as a whole, and the work that we will produce for this project. As stated in the last point, if these expectations are not met after trying to find compromise and solutions, we will seek outside supervision and guidance from Professor Stapleton. We are all excited about this topic and look forward to collaborating on this research project throughout the semester.