An Analysis of the Fairness in Hiring, Promoting, Salary Distributing Process

A Special Focus on Potential Gender Bias

Report prepared for Black Saber Software by KoCad

2021-04-21

Contents

Executive summary

Background & Aim

Current research was conducted by KoCad to investigate potential biases as well as the gender disparity issues that may exist in the company's recruitment, wage, and promotion schemes. To study the potential gender bias in the company's salary and promoting systems, we evaluated whether employees at the company were being paid and promoted fairly and that those processes were not influenced by discriminatory factors such as gender. We also investigated whether or not any kind of bias was included in each hiring phase that uses either the AI service or the score ratings of interviewers who evaluated applicants.

Key findings

- Male employees' average salary was approximately $1785.6 dollars more than that of female.
- Among those who have never been promoted, the proportion of female employees was slightly greater than that of male employees; However, unlike the proportion of male employees that increased with the number of promotions, the proportion of female employees decreased as the number of promotions increased.
- The promotion rate per financial quarter for the male employees was about 1.38 times greater than female employees.
- In the first hiring phase, the proportion of applicants who got approved to next phase was approximately evenly split between male and female.
- In the second hiring phase, only 7% of the applicants were invited to an interview and among them, the number of male applicants were slightly greater than female applicants.
- In the third hiring phase, the ranges of interview scores given by interviewer 1 and 2 were similar to each other. Also, AI did not give advantages/disadvantages with regard to gender of an applicants

Limitations

The lack of information on salary determination methods, promotion policies, and evaluation criteria in the hiring process has limited our models' ability to produce more accurate results.

Conclusion

There was a gender bias in both wage and promotion systems in Black Saber Software but not in the hiring process.

Figures

The key findings of the analysis are summarized in the following figures:

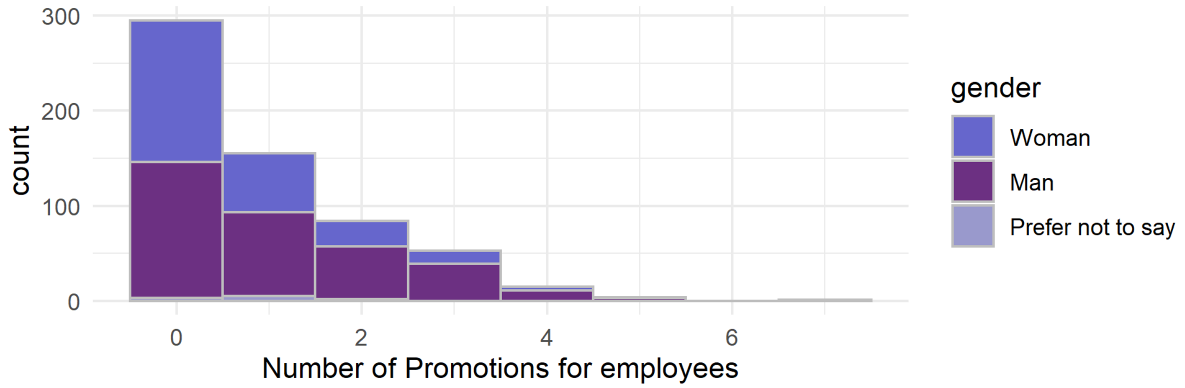

Figure1: Distribution of employees' number of promotion, showing genders as different colour.
Each bar's height represents the count for number of promotions

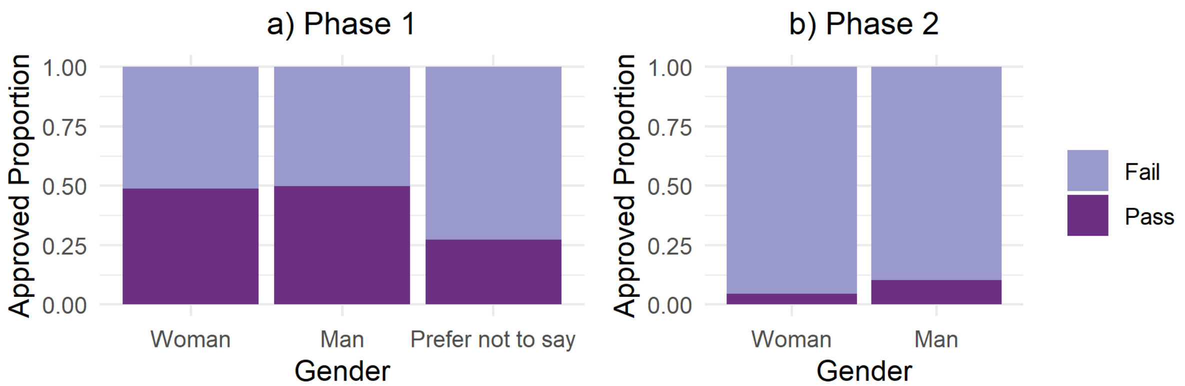

Figure 2: Proportion of Applicants who Passed/Failed in (a) Phase 1 and (b) Phase 2 by Gender

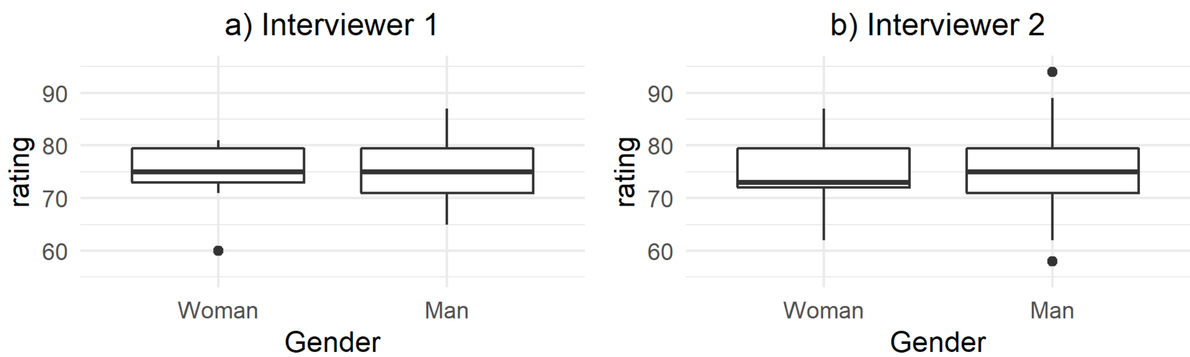

Figure 3: Distribution of Interview Scores evaluated by (a) Interviewer 1 and (b) Interviewer 2, by Gender

The horizontal line in the middle of each box represents the point separating the highest 50% interview scores from the lowest 50%. The horizontal lines at the top and bottom of each box represent the lowest 75% and 25% of the scores, respectively. The vertical lines and dots in the plot give the full range of values, with some extreme values further away from the middle of the box.

Technical report

Introduction

Society has been discriminating against women. They received unequal opportunities to get jobs and earn money based on their assets. With increasing concern about EDI, companies are moving toward implementing EDI initiatives. Considering this social atmosphere, our client, Black Saber Software also requested to launch an investigation on potential bias in their hiring, promoting and salary distributing processes. Using the hiring data for their new grad program and data about promotion and salary for their entire staffs, the present report investigated whether employees of Black Saber Software were hired, promoted, and paid based on their value to the company, not on discriminatory factors such as gender. First, we explored whether the employees were being paid fairly and that salary was not affected by one's gender, using a linear mixed effects model. Then, we studied the relationship between promotion frequency and some potential factors, particularly focusing on identifying the presence of gender bias in the promotion process. Here, we used a generalized linear model with the number of promotions as the response. In addition, with a simple linear regression and generalized mixed effects models, we explored the fairness of the recruitment process by investigating whether the implementation of AI algorithms and the interviewers who evaluated job applicants had gender bias. Finally, we discussed the limitations and strengths of our methods and suggested a future direction for improvement.

Research questions

Throughout the analysis, we focused on assessing the fairness in the hiring, promoting, and salary distributing processes at Black Saber Software. Below are our questions of interest:

Salary

- What factors, such as gender, team, role, leadership level, and productivity, are related to increased rate of salary?
- Does gender have significant influence on the salary?

Promotion

- What are some potential factors affecting the number of promotions of employees?
- Is the common assumption that male employees are more likely to be promoted than female employees also true at the Black Saber Software?

Hiring process

- Are there gender biases throughout the recruitment phase, especially phase 1 and phase 2 in which AI service was implemented?

A Potential Gender Bias in Salary

Data Wrangling and Visualizations

For data wrangling, we first changed the salary variable in the current employee data from characters to doubles so that we can fit a regression model with the salary variable as a response. We removed the dollar sign in the front and a comma sign that separates the thousands value. Then, we reordered the factor level of the employee's seniority of role such that it matches the actual hierarchy in the company rather than being in an alphabetical order. We have also set woman as the baseline for gender variable so that we can make comparisons of each of salary and promotion by gender easily.

Based on the current employee data, leadership level and productivity of an employee seemed to be the variables that shows the employee's value to the company. However, we also suspected that team and seniority of role may also have fixed effect on salary since based on those variables, the employee will have differential value to the company. For instance, some teams like software and data team, might be more valuable to the company because Black Saber is a software company. Also, for seniority of role, the higher position an employee takes, the more challenging work they have to do and thus more valuable to the company.

To figure out whether such speculation is true, we fitted a boxplot.

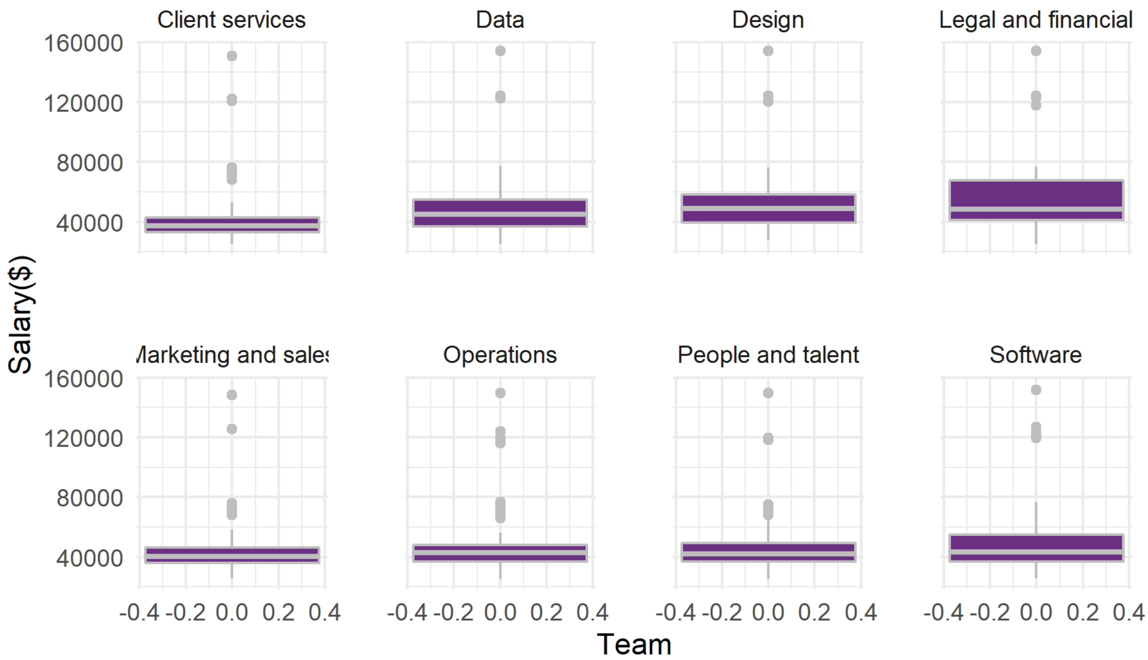

Figure1: Boxplots for Salary by Team

Indeed, the software and data teams generally receive higher salary than other teams as shown in Figure 1. Thus, we also included team and seniority of role as predictors to a regression model.

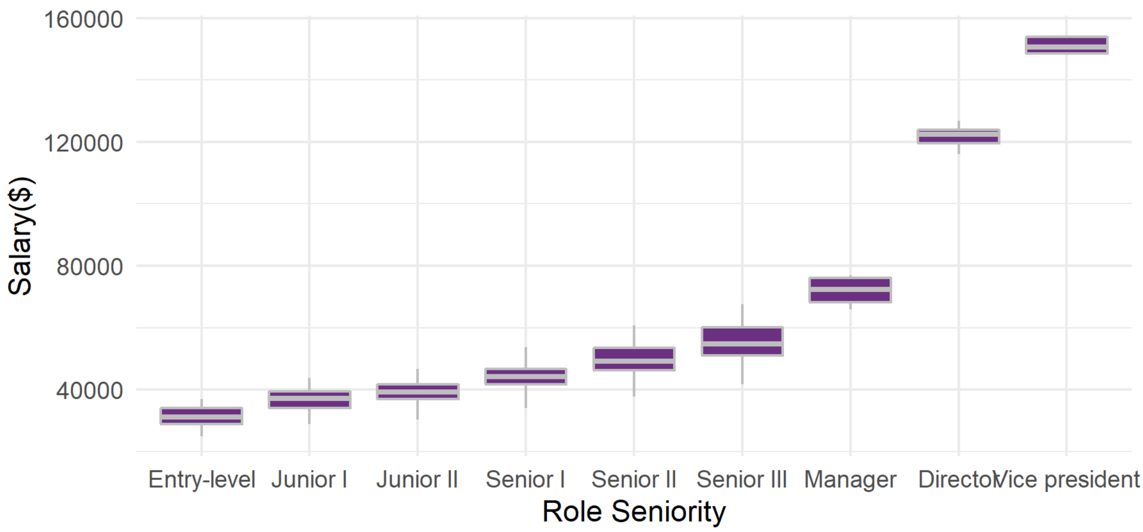

Figure2: Boxplots for Salary by Seniority Role

Also, we see that salary significantly depends on the person's role seniority level. In particular, manager, director, vice president receives salaries significantly different from other positions (see Figure 2).

We also looked at a boxplot for salary by gender to see whether there is a gender difference in salary (see Appendix Figure 1). Although gender does not seem to create sysemtatic difference, the one-way anova of gender as a predictor and salary as a response suggests a significant difference. However, we need to see this along with other variables related to an employee's value to the company. Thus, we will see whether a regression model with and without gender as a predictor significantly differs after we fit each model and make comparison.

Methods

We chose a linear mixed effects model for the regression model because financial quarter will also produce sysematic difference, since for instance, in financial quarters where Black Saber earns a lot of money, employees will receive bonuses. However, we are not interested in this variable because we only care about the employee's value. Further, statistically speaking, employees' salaries were sampled across various financial quarter and so this would violate the independence assumption for observations. Thus, it was important to add financial quarter as a random intercept.

Results of the maximum likelihood test comparing models with and without gender as predictor while including seniority of role, team, leadership, and productivity as other predictors was significant (see Appendix). The significant p-value shows that we have no evidence against the claim that simpler model is better. In other words, we should go with the complex model, which is the model that includes gender as predictor, to have better predictive ability for salary. This result implicates that gender cause notable difference in salary and so there is a gender bias in salary.

Note that we also tested whether we should include team as predictor for salary becuase it might be ambiguous of whether team has fixed effect on salary based on the above plot as compared to seniority of role. The result showed that we should also include team (see appendix).

Results

Table 1: Partial Table of Coefficient Estimates for Salary Model Including Gender

	Estimate
(Intercept)	27851.5437
genderMan	1785.6056
genderPrefer not to say	499.8409

Table1 shows that for our final model predicting salary, man receives salary approximately 1785.6

dollars more than woman, on average, while controlling for variables related to employee's talent to the company. Similarly, on average, people who did not prefer to say their gender also received salary higher than woman by approximately 500 dollars. For full information about coefficient estimates for the final model for salary, please see Table1 in the Appendix.

In conclusion, there is a gender bias in salary since gender improves salary prediction above and beyond inclusion of other variables, which are the factors that show the employee's value to Black Saber Software. Moreover, such factors not only includes employee's productivity and leadership, but also the team they are in and the role they take in the company.

A potential Gender Bias in Promoting Process

Data Wrangling and Visualizations

For this particular topic, employee data that was cleaned in the previous section was used to study several potential factors that affected the company's promoting process. Each line of this data contained information of the 607 employees at the Black Saber Software.

For this analysis, the response variable was the number of promotions per employee after joining the company. Since there may be some employees who have been promoted more than once, it was thought that having a Bernoulli variable (which indicates whether or not an employee has been promoted at least once) as our response would not take into account the variability in different number of promotions of the employees; therefore, we chose our response to be the number of promotions, instead.

Since the main purpose of investigating this research topic was to see the potential bias present in the promotion process, we first created a new variable that stored each employee's roles numerically rather than categorically to track changes in their roles more easily. This variable helped us determine the initial and current roles of each employee, and the difference between the initial and current roles gave us the number of times each employee was promoted since he or she first joined the company.

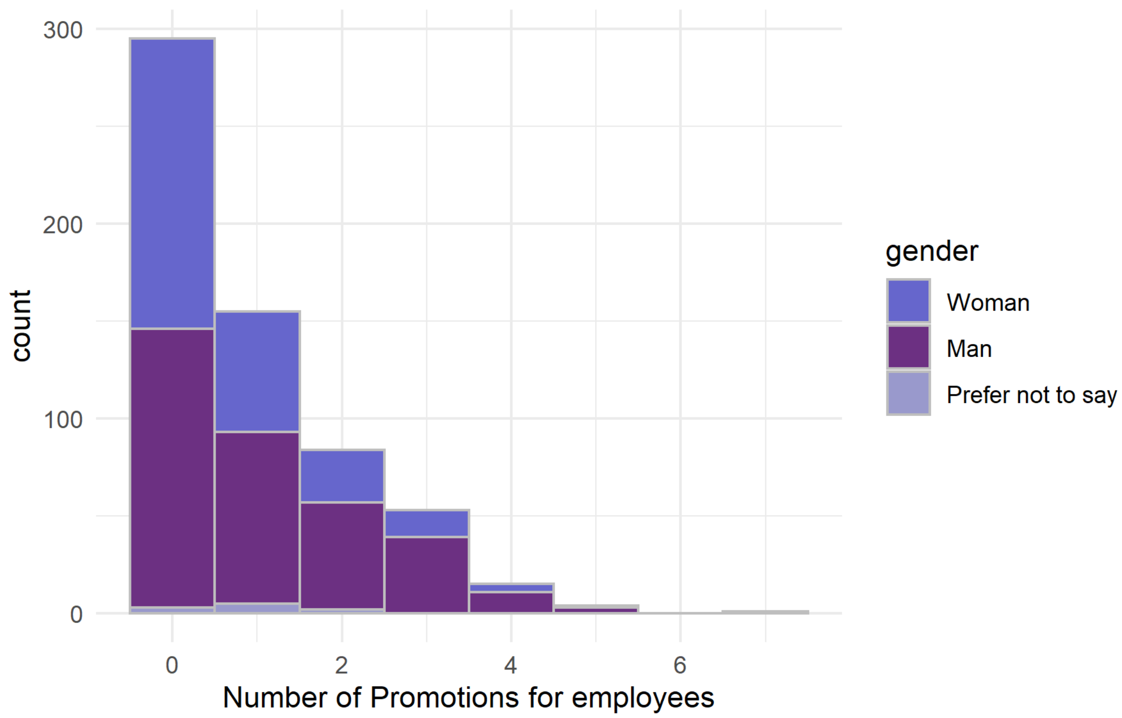

Figure3: Distribution of Employees' Number of Promotions by Gender

Starting with our response, the number of promotions, it was found that around 48.6 % of the employees at the Black Saber Software have never been promoted since they first joined the company. (Figure3 illustrates the observed number of promotions for the employees by gender). The mean number of promotions recorded in the employee data was around 0.94 with a standard deviation of 1.17 and the number of promotions was highly right-skewed. However, this is not very surprising because since our response is a count, it is expected to follow a Poisson model with lambda equal to the observed mean number of promotions, and a Poisson distribution with a low mean is usually highly skewed.

Meanwhile, we were also interested in investigating whether gender bias was present in the promotion process at the Black Saber Software. According to Figure 3, there are more female employees who have never been promoted than male employees. As the number of promotions increased, this trend was becoming more pronounced, and among those who were promoted at least once, male employees accounted for more than women. Thus, this trend further highlighted the importance of identifying the potential gender bias present in the promotion process.

Methods

Since the response was a count, a Poisson regression was used to model our data. However, there were several assumptions that we had to check before fitting a Poisson model:

1. **Poisson Response:** Since the response was a count it followed a Poisson model.

2. **Independence:** This assumption was also met because each observation in the data contained information of different employees.

3. **Mean = Variance:** Accounting for the fact that the number of promotions varied from 0 to 7, the discrepancies between the mean and the variance, ranging from 0.3 to 0.5, were considered small; therefore, we assumed that the assumption of variability equal to the mean was also satisfied.

4. **Linearity:** Since gender was not a continuous predictor, it was difficult to identify the linearity of $\log(\lambda)$.

As we were interested in studying the potential gender bias in the promoting process of the Black Saber Software, we created two different models to determine whether including the gender improves our model. In both models, we included an offset, which was the log of the work period, so that the number of promotions could be adjusted to be comparable across employees with different working periods[1].

The initial model contained covariates such as leadership, productivity, salary and role seniority and it was compared with another model with an additional gender predictor. With a small p-value from the drop-in-deviance test, we concluded that adding gender actually improved the model.

Results

This table shows the estimates and 95% confidence interval for Gender variable, calculated from the final model:

Table 2: Partial Table of Coefficient Estimates and 95% Confidence Interval for Gender in the Final Model

	Estimate	2.5 %	97.5 %
(Intercept)	-20.6975	-3004.9658	-3004.9658
genderMan	0.3255	0.1367	0.5180
genderPrefer not to say	0.2164	-0.5394	0.8405

From the summary of the model, only the p-value for male employees was found to be significant[2].

[1] Employees who have worked longer than others are more likely to be promoted at least once.
[2] The full table of the estimates and 95 % confidence interval for the model is included in the Appendix

Also, considering that the 95% confidence interval for male employees does not contain 0, there is evidence against the hypothesis that there is no difference in promotion frequency between men and women (which was our baseline for gender). Specifically, Table 2 suggests that the promotion rate per financial quarter for male employees was nearly 1.38 times greater than that of females, after controlling for other predictors. In other words, a gender bias was actually present in the promoting process at the Black Saber Software.

Fairness in Hiring Process

Data Wrangling and Visualizations

We first made variable called phaseNaccepted telling us whether an applicant pass that phase or not. Since our main purpose is testing whether genders of applicants gives affects to the acceptance of the phase or not, this ordinal variable will help us to make response variable to have bernoulli distribution. Some phases do not have information about genders of applicants, so make all phases have gender information of all applicants. Also, in phase 2, we can see that no one get accepted. We will omit the Prefer Not to Say gender from phase 2.

Nearly half (50.7%) of our 613 applicants are females.

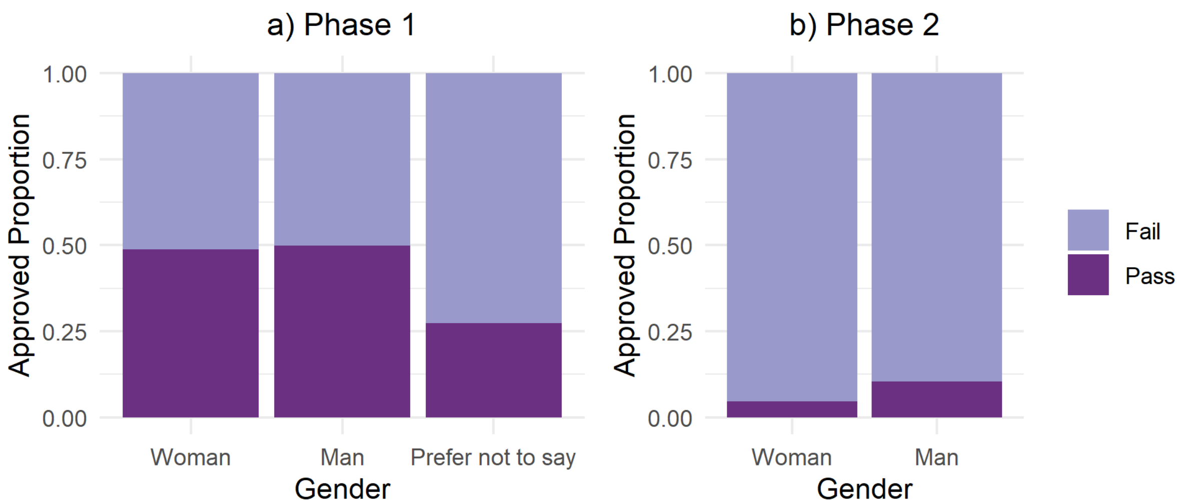

Figure 4: Proportion of Applicants who Passed/Failed in (a) Phase 1 and (b) Phase 2 by Gender

The proportions of applicants who are accepted in phase 1 has similar proportion between male and female. We can see 25% of difference in the the accepted rate in phase 1 with the applicants who prefer not to say their gender, so we will see more detaily by models. Remind that this is proportion so the proportion depends on the number of the applicants. We have only 11 Prefer not to say applicants, which is 2% of whole applicants.

In phase2, we can see male has slightly higher percentage than female. However, the number of people who accepted in phase 2 has small proportion(7%), so we need to be careful for the difference.

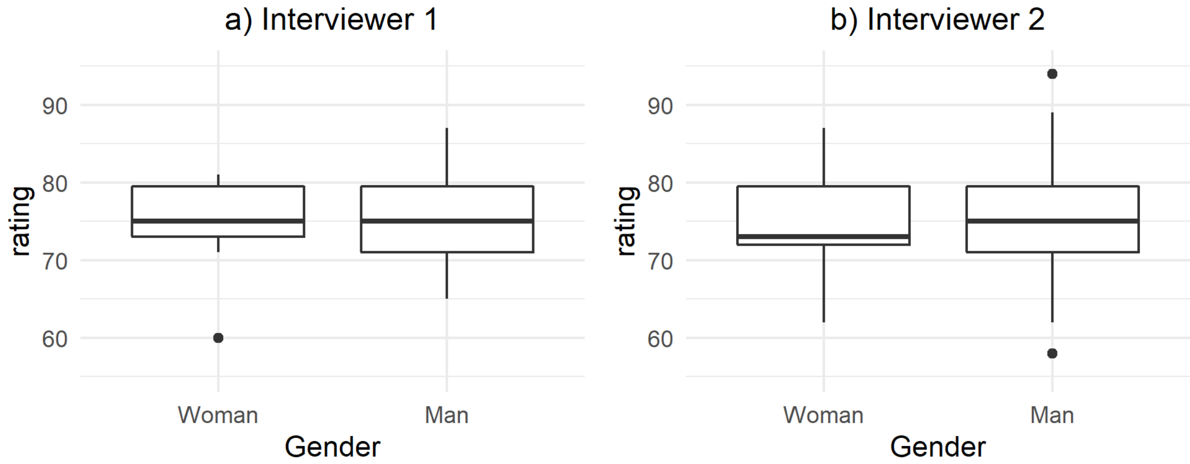

Figure 5: Distribution of Interview Scores evaluated by (a) Interviewer 1 and (b) Interviewer 2 by Gender

In phase 3, interviewer 1 and 2 gives similar range of grades to an interview. Even though there are slight mean difference between male and female for interviewer 2, it is just 2.5 points difference so we can assume that gender does not give effect to the interview score.

Methods

We do not know how AI make grade for each applicants but AI will consider all the variables, so we put all variables as fixed effects. Response variable follows bernoulli, which is passed each phase or not. Therefore as use logit link function, we can assume that there is a linear relationship between the transformed response and explanatory variables. We can assume that information of each applicants are independent. Therefore we used generalized mixed model.

In initial model, where gender is our only predictor, confidence interval of all genders includes 0. Therefore at the 95% confidence level, it is believable that female applicants and male applicants with applicants who prefer not to say gender have the same log odds of accepted by AI in phase 1. We will see more carefully that AI is biased with gender with other variables in phase 1.

In the models with all other variables, the deviance test tells us that it is not preferable that adding gender as an indicator variable to the model. The deviance difference between models with gender and without gender is 1.21, which is pretty small. Also, the large p-value also supports that adding new indicator variable, which is gender, to the model with all the other variables is not helpful.

Similar as phase 1, we can see that there are small deviance(0.62) and large p-value(0.43) in deviance test. Therefore adding gender to the model does not explain acceptance of applicants in phase 2.

We will use simple linear regression model in phase 3. We have one fixed effects, which is interview rating and gender. we can assume interview rate is related to gender linearly. By residuals and fitted value plot we can see the mean of the erros is 0. Also, standardized residuals and fitted value graph shows us the errors are homoscedastic. And obviously, the errors are uncorrelated.

From confidence interval, we can see the coefficients of gender contains 0, so we can say interviewers are not biased by gender. In other words, gender does not give effect to interview scores.

Discussion

In this section you will summarize your findings across all the research questions and discuss the strengths and limitations of your work. It doesn't have to be long, but keep in mind that often people will just skim the intro and the discussion of a document like this, so make sure it is useful as a semi-standalone section (doesn't have to be completely standalone like the executive summary).

Overall, the results of statistical analyses suggest that there is a gender bias in salary distribution process and promotion process of current employees in Black Saber Software, but not for the hiring process using trialling AI and the human interviewer; For biased processes, men received higher salary and were more likely to get promoted than women, despite controlling for their talents and values to the company.

In particular, for salary, we found that employee will be affected by gender when other explanatory variables such as productivity, leadership, and team are controlled. Furthermore, for promotion, number of promotion also get affected by gender when other factors such as productivity, leadership, salary and role seniority are fixed. In hiring process, we can find that both AI and human are not biased with gender. Specifically, we found that applicants will not be affected by gender from AI when AI decide the applicant is pass to phase or not. Also, the grade of interview will not be affected by gender from interviewer either.

Strengths and limitations

One limitation of our analysis is that the provided data did not contain information on how the company would select employees for promotion. Understanding the company's core values or knowing more about employee promotion policies will help us indentify the right factors that affect employee promotion.

Another limitation with making models in hiring process is that we do not know the evaluation criteria for each phase. If we know how score is being calculated by the implemented AI, it would be easier to identify what effects are correlated, and make better model with interaction terms. Same for interview score, if we are given with a list of interview questions and an evaluation criteria template that contains detailed components of the grading schemes, we would be able to explore potential biases in the interview phase more in depth.

For future consideration of our study, we could look into salary and promotion distributions by team, role and financial quarter and identify segments with pronounced biases. This way, we will not only be able to figure out whether or not the company's wage and promotion systems are biased but we can also find out specific teams, roles or financial quarters in which the biases are present.

Consultant information

Consultant profiles

Yoon Young Lee. Yoon Young is a junior data analyst with KoCad. She specializes in data visualization and data analysis. Yoon Young earned her Bachelor of Science, Specializing in Psychology and Majoring in Statistics from the University of Toronto in 2022.

Guemin Kim. Guemin is a junior data engineer with KoCad. She specializes in data analysis and machine learning. Guemin earned her Bachelor of Science, Majoring in Actuarial Science and Statistics from the University of Toronto in 2022.

Woolim Kim. Woolim is a junior data analyst with KoCad. He specializes in data visualization and statistical communication. Woolim earned his Bachelor of Science, Majoring in Statistics and Economics from the University of Toronto in 2022.

Hojung Kim. Hojung is a junior data analyst with KoCad. He specializes in mathematical statistics and statistical communication. Hojung earned his Bachelor of Science, Majoring in Statistics and Mathematics from the University of Toronto in 2022.

Code of ethical conduct

This section should be fairly short, no more than half a page. Assume a general audience, much like your executive summary.

- *Make at least three relevant statements about your company's approach to ethical statistical consulting. These should be appropriately in line with professional conduct advice like the (Statistical Society of Canada Code of Conduct)[https://ssc.ca/sites/default/files/data/Members/public/Accreditation/ethics_e.pdf] or the (Ethical Guidelines for Statistical Practice from the American Statistical Society)[https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx]. For example, "the customer is always right" ISN'T the type of thing an ethical statistical consultant would include.*
- *Be very careful not to just copy and paste from these other documents! Put things in your own words.*

Responsibility of Statistician - KoCad acknowledges the statisticans' responsibility of manipulating data, analyzing data, and interpreting the result of analysis in a transparent way. KoCad promises not to manipulate data in a way that will introduce bias, in order to create a significant result. KoCad approaches the analysis with the appropriate model for the given data and not by the model that we want. KoCad communicates the results in a manner that considers indiviudal differences in beliefs, opinions, and background.

Responsibility to Clients - KoCad cannot guarantee that the results of the analysis will be exactly aligned with the expectations of the client. KoCad is responsible for retaining full knowledge and understanding of statistical methods, deducing valid conclusions from the data provided, and identifying and explaining any limitations to the conclusions that can be drawn.

Responsibility to other fellow statisticians - KoCad ensures a supportive working environment to fellow statisticians in their professional development. To motivate and inspire fellow professionals, questions and debate on projects are recommended. Avoiding direct criticism of the person, conflicts should be directed and resolved according to procedures. All fellow KoCad consultants are responsible to act with integrity toward others.

Appendix

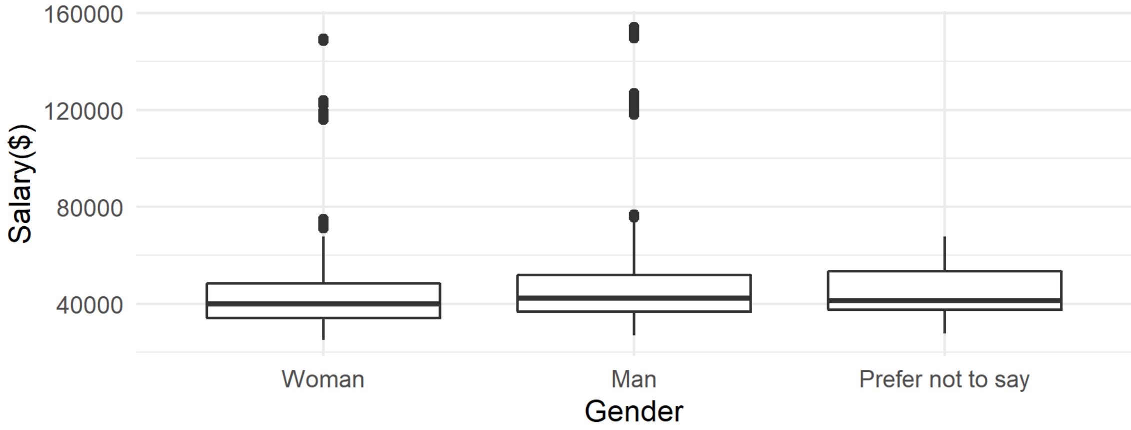

Figure1: Boxplots for Salary by Seniority Role

Table 3: Full Table of Coefficient Estimates for Salary Model

	Estimate
(Intercept)	27851.5437
genderMan	1785.6056
genderPrefer not to say	499.8409
teamData	3796.7121
teamDesign	1180.7819
teamLegal and financial	4545.6853
teamMarketing and sales	1796.9405
teamOperations	1273.5526
teamPeople and talent	-1220.5586
teamSoftware	4839.2788
role_seniorityJunior I	5372.1281
role_seniorityJunior II	7853.7190
role_senioritySenior I	13138.4922
role_senioritySenior II	18614.8519
role_senioritySenior III	24167.9741
role_seniorityManager	40309.9786

	Estimate
role_seniorityDirector	91071.2294
role_seniorityVice president	119589.8496
leadership_for_levelExceeds expectations	-288.2453
leadership_for_levelNeeds improvement	-182.8906
productivity	1.8654

Table 4: Full Table of Coefficient Estimates with 95% Confidence Intervals for Promotion Model

	Estimate	2.5 %	97.5 %
(Intercept)	-20.6975	-3004.9658	-3004.9658
genderMan	0.3255	0.1367	0.5180
genderPrefer not to say	0.2164	-0.5394	0.8405
leadership_for_levelExceeds expectations	-0.1277	-0.7814	0.4304
leadership_for_levelNeeds improvement	0.1843	-0.6863	0.8855
productivity	0.0017	-0.0045	0.0080
salary	0.0000	0.0000	0.0000
role_seniorityJunior I	17.6121	3000.2624	3000.2624
role_seniorityJunior II	17.7547	3002.0225	3002.0225
role_senioritySenior I	18.1421	3002.4100	3002.4100
role_senioritySenior II	18.0128	2173.3174	2173.3174
role_senioritySenior III	18.1659	3002.4339	3002.4339
role_seniorityManager	18.0216	3002.2900	3002.2900
role_seniorityDirector	18.0293	1344.3718	1344.3718
role_seniorityVice president	17.8445	3002.1172	3002.1172

Table 5: Full Table of Coefficient Estimates with 95% Confidence Intervals for Hiring Phase 1 Model

	Estimate	2.5 %	97.5 %
(Intercept)	-151.1283	-2449.2960	-3653.7517
team_applied_forSoftware	-0.9071	-2.7961	0.8373
cover_letter	59.0835	31.6094	1184.7225
cv	48.7949	-104.2644	1776.5233
gpa	11.9693	6.9696	19.5409
extracurriculars	9.5009	5.8364	15.2032
work_experience	10.8500	6.7927	17.1193

Table 6: Full Table of Coefficient Estimates with 95% Confidence Intervals for Hiring Phase 2 Model

	Estimate	2.5 %	97.5 %
(Intercept)	-20.6704	-29.0472	-14.4615
technical_skills	0.0789	0.0444	0.1228
writing_skills	0.0872	0.0473	0.1361
speaking_skills	0.7159	0.4206	1.0787
leadership_presence	0.9111	0.5586	1.3780

Table 7: Full Table of 95% Confidence Intervals for Hiring Phase 3 Model

	2.5 %	97.5 %
(Intercept_1)	69.4626	79.3945
genderMan_1	-4.9093	7.1188
(Intercept_2)	67.7454	82.2546
genderMan_2	-8.4524	9.1191

Final advice: KNIT EARLY AND OFTEN!