Investigation on Gender Bias

Focus on Hiring, Promotion and Salary

Report prepared for Black Saber Software by BDSTA.CO

2021-04-21

Contents

Executive summary

The promotion of gender equality can be seen in every dimension of life, from private to public, from work to home. However, many still think the male generally has more opportunities than the female when getting high-salary occupations and being leaders (Pew Research Center, 2020). This study analyzes the hiring data for the new grad program and the data about promotion and salary for all the staff in Black Saber Software to access the gender parity in hiring, wages, and promotion. The results of the key findings are summarized below.

Table 1: Number of Female and Male in Each Hiring Phase Promotions

Characteristic	**Female**, N = 339[1]	**Male**, N = 341[1]
round		
phase1	178 (53%)	173 (51%)
phase2	152 (45%)	145 (43%)
phase3	7 (2.1%)	15 (4.4%)
phase4	2 (0.6%)	8 (2.3%)

[1] n (%)

- According to **Table 1**, the proportions of male and female enter round 1 and round 2 are similar. Only 2.1 percent of females enter round 3, whereas 4.4% of males enter round 3. The effects of gender do not play a role in all the phases of the hiring process.

Table 2: Number of Female and Male in Each Number of Promotions

Characteristic	**Man**, N = 340[1]	**Woman**, N = 257[1]
promotion		
0	143 (42%)	149 (58%)
1	88 (26%)	62 (24%)
2	55 (16%)	27 (11%)
3	39 (11%)	14 (5.4%)
4	11 (3.2%)	4 (1.6%)
5	3 (0.9%)	1 (0.4%)

7	1 (0.3%)	0 (0%)

[1] n (%)

- According to **Table 2**, 51% of the never promoted employees are female; the percentages of men exceed that of women under all numbers of promotions.
- When the employee is male, the odds (probability of promotion/probability of not being promoted) of promotion are about two times that of when the employee is a female, when their time of employment is the same.
- Over the period from 2013 to 2020, the median salary for men is $42,400, whereas the median salary for women is $40,000 across all roles.

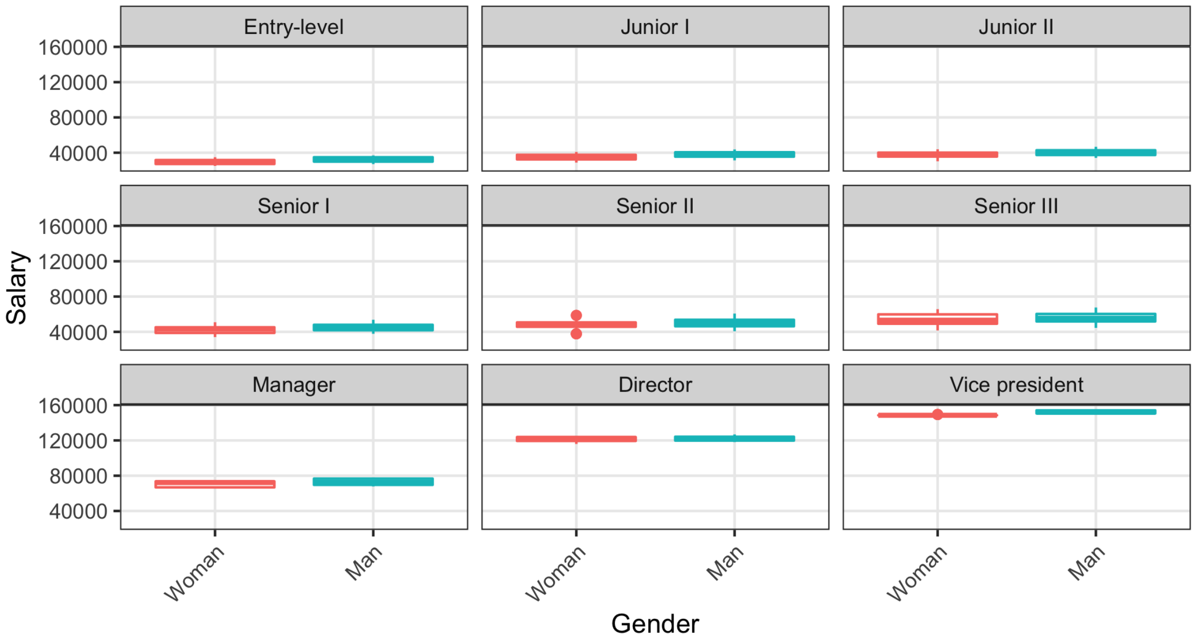


Figure 1: Boxplot of Salary vs. Gender across Role Seniority

- **Figure 1** shows that women have a lower salary than men under each role.
- We estimate female yields $2746.51 lower than males on average across various seniority.

The limitations of this study are summarized below.

- The study focuses on whether the employees have been promoted or not during their time of recruitment, instead of the frequency of promotions they had. Therefore, we do not consider the differences between the employees who were promoted once during five years and those who were promoted fifth during five years.
- A few outliers have high salaries, which may affect our results in the gender distribution.

Technical report

Introduction

There have been improvements in gender equality over the past decades, with more and more women's voices being heard by the public. Countries worldwide have published legislation for gender equality to guarantee female rights as it is a foundation for the well-being of society (LegislationOnline).

The objective of this report is to assess the gender parity in Black Saber Software from the aspects of hiring, wages, and promotion. With four datasets containing information about the applicants for each recruiting phase, we investigate gender discrimination in the recruitment process with AI. Using the dataset about current employees, we evaluate whether gender plays a role in promoting and determining the remuneration amount.

Research questions

- Does the gender discrimination exist in the recruitment process by using AI?
- Do males have more opportunities to get promoted than females?
- Is the salary level is affected by the gender and how does it influence?

Gender bias in Recruitment

Data Wrangling

We are given four raw data sets from Black Saber Software to analyze the hiring process, each containing information about the applicants for each recruiting phase. A total of 613 participants applied to Black Saber Software are included in the first phase, 300 in phase 2, 22 in phase 3, and 10 in phase 4. For every two adjacent phases, we combine the data sets as a new one to mark the passing applicants in that round. For instance, applicants who pass the first phase and enter the second phase will be assigned "1" in the new data set for round 1. Hence, we create three data sets in total, where each stands for the result of one hiring round and contains all the information available for each participant. In this way, we can conduct logistic models to estimate the odds of passing each round and explore the factors that may affect the odds.

Since we aim to analyze the gender effect, we remove all the observations with unknown gender. In the Round1 dataset, we have information on applicants' gender, cv, and cover letter submission condition, GPA, and conditions on extracurriculars and work experience. We remove observations with a cv and cover letter submission and no experience in extracurriculars since people without these qualifications are automatically dropped according to observations in phase 2. There are four test scores in the Round 2 dataset for each applicant, technical skills, writing skills, leadership presence, and speaking skills. The first two are conducted on a 100 scale through writing tasks, and the last two in 10 through a video, with greater scores representing better performance. Hence, we reclassify the tests into two types (write skill, video skill) based on their scales and assessment approach to simplify the model. In the Round 3 dataset, we sum the two interviewer scores as a final score.

Data Visualization

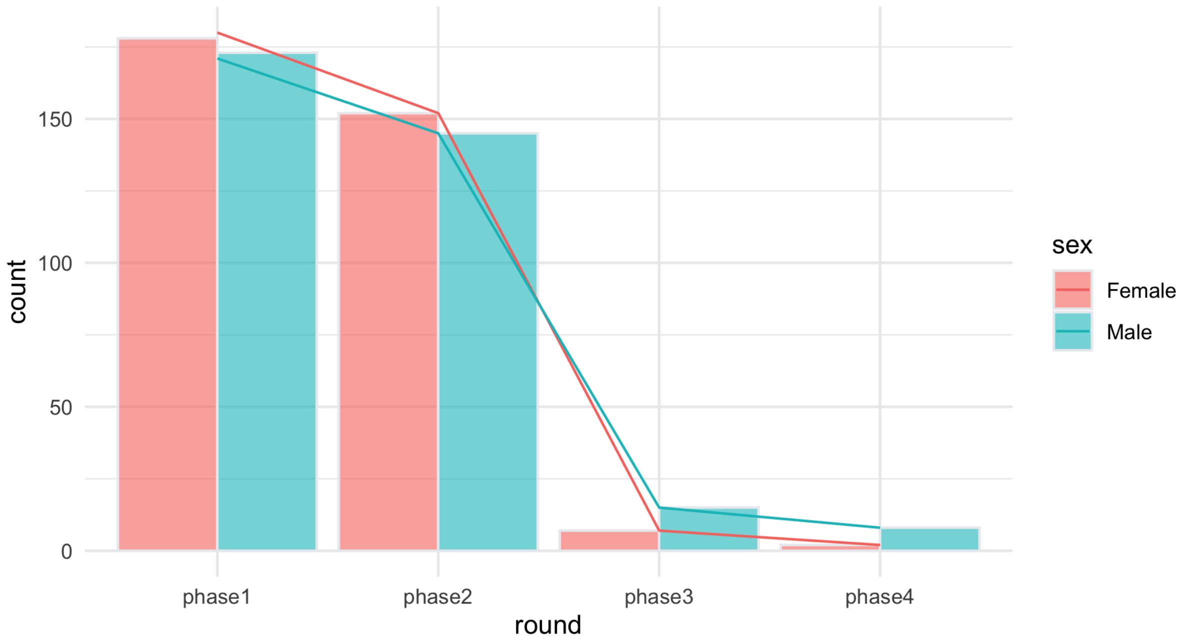


Figure 2: Bar Plot for Applicants in each Hiring Phase

The recruitment process consists of three rounds of assessment, with the first two graded and chosen by an AI algorithm and the last round held by employers. To generate an overview of the hiring condition, we plot the number of applicants within each gender in each phase with grouped bar plots. As shown in **Figure 2**, the proportional decreases in round 1 (phase 1 - phase2) and round 3 (phase3 - final phase) are similar between the two genders. However, it seems like females could be subject to gender discrimination in the second round (phase 2 - phase 3). Therefore, we put our emphasis on investigating second-round recruitment in this report. To produce a thorough investigation, we also model and analyze the other two rounds. Specifics are included in **Appendix 1: Logistic Regression for 1st Round Hiring** and **Appendix 2: Linear Regression for 3rd Round Hiring**.

Model

In this model, we aim to explore the relationship between gender and the chances of passing the second round of recruitment. With a binary response on whether applicants have passed the hiring round, we employ a logistic model to estimate the odds of passing. Since we do not know how the algorithm works specifically, we assume that the AI will only look at the demographics and the round 2 test scores as evaluation criteria and include those variables as predictors. With different combinations of predictors, we choose the model with the smallest AIC. The independent variables include gender, write skill, and video skill. The null hypothesis of the model is that there is no gender bias with the AI assessment. The model is shown below:

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_{gender} X_{gender} + \beta_{skillwrite} X_{skillwrite} + \beta_{skillvideo} X_{skillvideo}$$

where p_i is the probability of the ith applicant passing the second round (1=pass; 0=fail), β_0 denotes the intercept value, represent the log odds of a male with 0 score on all tests, β_{gender} denotes the slope value for gender groups, represent the log odds difference between females and males; X_{gender}: $=1$ if the respondent is identified as a female; $\beta_{skillwrite}$: the slope values for candidates' writing skill scores; represent the log odds difference among writing scores; $X_{skillwrite}$: the candidate's writing scores; $\beta_{skillvideo}$: the slope values for candidates' video skill scores; represent the log odds difference among writing scores; $X_{skillvideo}$: the candidate's video skill scores

We conduct model diagnostic checks on the above final model, through multicollinearity and influential points test. First, one assumption is that there is no multicollinearity between the predictors so that the inference will not be affected. We use the Variance Inflation Factor (VIF) to check if there are any correlated independent variables. Based on **Appendix 3: Multicollinearity (2nd Hiring Model)**, the VIFs for each predictor is less than 2, which is smaller than the convention cutoff of 5. Thus we do not need to remove the variables in the logistic model. Second, the existence of influential points may affect the fitted model, which can be examined by the cook's distance. According to **Appendix 4: Influential Point (2nd Hiring Model)**, there are no points with a value greater than convention cutoff of 0.5 in the cook's distance plots; thus, there are no influential plots.

Results

Table 3: Summary of Logit Regression (2nd Hiring)

	Estimate	Std. Error	z value	Pr(>|z|)	2.5 %	97.5 %
(Intercept)	-20.344	3.651	-5.572	0.000	-28.667	-14.171
genderWoman	-0.554	0.705	-0.786	0.432	-2.003	0.801
skill_write	0.085	0.019	4.579	0.000	0.053	0.127
skill_video	0.781	0.152	5.124	0.000	0.517	1.122

According to the **Table 3**, the gender effect on the odds of passing the second-round recruitment is non-significantly negative. The p-value for the gender being female is 0.432, larger than the convention cutoff of 5%. It implies no evidence against the null hypothesis that the AI algorithm treats males and females equally in the second round of recruitment. Similarly, the 95% confidence interval ranges from -2.00 to 0.80, including 0. Thus, both the p-value and confidence interval suggest that there is no evidence that gender is significantly associated with the pass of second-round recruiting. The interpretation of the gender effect can be omitted under the context.

Conclusion

To conclude, the second-round hiring process is free of gender discrimination under the AI algorithm. Although the data collected may propose that females are treated unequally in the second round since there is a larger decrease in the number of passes in females than males, the model results do not support the argument. Therefore, the hiring pipeline with the trailing of AI services assesses the candidates fairly and equally. Black Saber Software is not exposed to the risk of being sued on potential bias in its hiring process.

Gender Bias in Promotion

Data Wrangling

The performance and the basic information about the employee are separated by season. In order to conduct the experiment, the initial data set needs to be cleaned. Firstly, our group believes the number of years that the employee has worked for this company would affect the promotion opportunity. Therefore, based on the number of seasons presented on the data set, a new variable is created called year, which shows the duration of employment. Secondly, for each season, the original data set provides the role seniority. A new variable is created called promotion to record the number of promotion opportunities for each employee since she/he worked in the company. To fit the logistic regression, a new associated binary variable is created. If the number of promotions is greater than 0, then the binary variable is recorded as 0. Otherwise, the value is recorded as 1. Also, those employees whose gender is recorded as prefer not to say are deleted. Finally, after using the distinct function, the targeted independent and dependent variables are selected, which are gender, team, year, and promotion. The final data set is tidy because every column is a variable, every row is an observation, and every cell is a single value.

Data Visualization

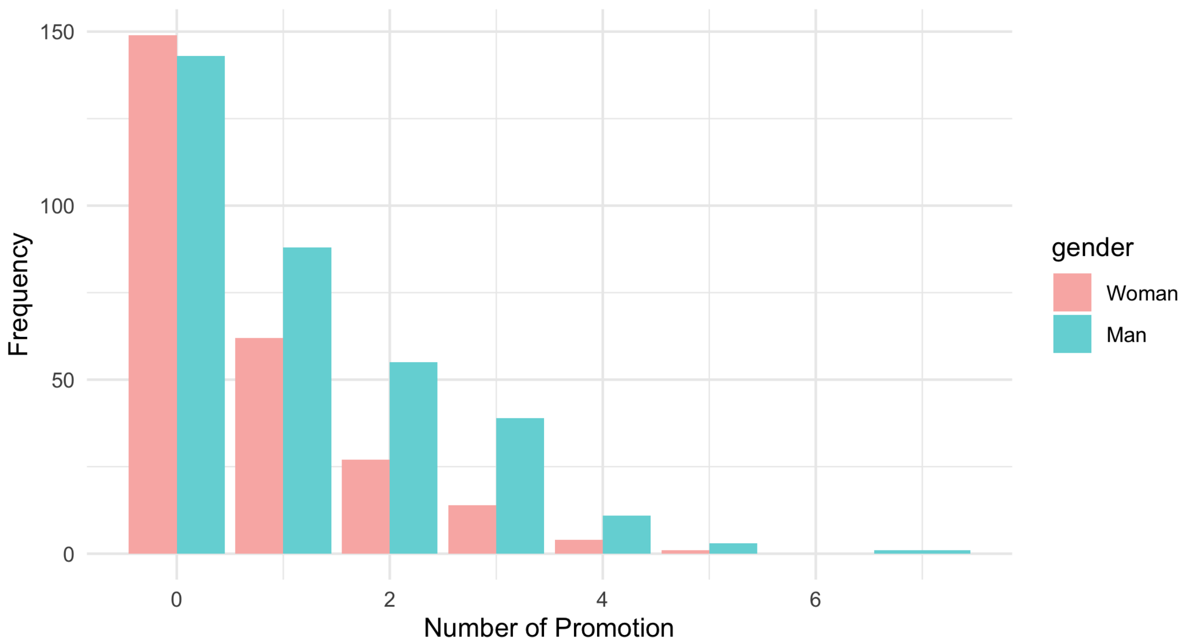


Figure 3: Bar Plot for Promotion Opportunities

Our group chooses to plot the frequency for a different number of promotions and separate the frequencies based on gender, male and female. According to the **Figure 3**, for those employees who have promotion opportunities, the number of male employees is consistently higher than the number of employees, regardless of the number of promotions. Thus, based on the visualized data, male employees are more likely to get promotions than those female employees.

Model

The logistic regression is chosen because the dependent variable is binary, which represents whether the employee has been promoted since he/she joined this company. The null hypothesis is that the promotion opportunities are the same among female employees and male employees. When conducting the logistic regression model, three independent variables are included, which represents the employee's gender, their working team and the number of years that they have worked for this company. The model is shown below:

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_{gender}X_{gender,i} + \beta_{team}X_{team,i} + \beta_{year}X_{year}$$

where p_i: the probability of the employee get promotion opportunities, β_0: the intercept value; represents the log odds for an male employees who are at client services team, β_{gender}: the slope values for gender; represent the difference of promotion opportunities between female and male, $X_{gender,i}$: $=1$ if the employees are female, β_{team}: the slope values for different team groups; represent the difference of promotion opportunities between eight working teams, $X_{team,i}$: $=1$ if the employees are in the working teams other than client service, β_{year}: the slope values for working years; represent the difference of promotion opportunities between people who worked for different years, X_{year}: the number of years that the employees have worked in the company

To ensure the feasibility of the model, we will check the multicollinearity and the influential points in the data set. Firstly, one of the assumptions that are required to be checked is multicollinearity between the predictors, which may influence the inference. Using the Variance Inflation Factor (VIF), we need to remove one of the independent variables if two or more predictors are correlated. Based on **Appendix 5: Multicollinearity (Promotion Model)**, the VIFs for each predictor is about 1, smaller than the conventional cutoff of 5. Thus we do not need to remove the variables from the logistic model. Secondly, the influential points will considerably affect the fitted model, which can be checked by the cook's distance. Based on the criteria, points with a cook's distance that is greater than 0.5 should be noticeable. Referring to **Appendix 6: Influential Point (Promotion Model)**, there are no points with a value greater than 0.5 in the cook's distance plots; thus, there are no influential plots.

Results

Table 4: Summary of Logit Regression (Promotion)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.015	0.360	-8.370	0.000
genderWoman	-0.684	0.256	-2.675	0.007
teamData	-0.098	0.419	-0.233	0.815
teamDesign	0.364	0.793	0.459	0.646
teamLegal and financial	0.147	0.614	0.239	0.811
teamMarketing and sales	-0.740	0.380	-1.948	0.051
teamOperations	-0.263	0.428	-0.613	0.540
teamPeople and talent	0.154	0.689	0.224	0.823
teamSoftware	-0.027	0.363	-0.074	0.941
year	1.339	0.107	12.573	0.000

Based on the regression results in **Table 4**, the P-value for gender is statistically significant because its P-value is less than 0.05. We have strong evidence to reject the null hypothesis that female employees have the same opportunity to get promotions as males. Instead, holding other variable constant, the estimate of the odds ratio (the odds of promotion when the employee is female compared to the odds of promotion when the employee is male) is 0.5047. It demonstrates that the odds that a female gets the preferment since she has entered the company is about half of that when a male employee. Additionally, based on **Appendix 7: Confidence Interval of Logit Regression (Promotion Model)**, the 95% confidence interval for gender does not include 0. It means that gender is statistically significant, which corresponds to the results from the P-value.

Conclusion

Based on the statistical analysis, the promotion is affected by the employee's gender. Male employees are more likely to promote than female employees. Therefore, holding other variables constant, we can conclude that male employees are more likely to upgrade than female employees. According to the provided data set, the company is considered to have gender discrimination regarding the promotion. The firm may need to make adaptations to the system.

Gender Bias in Salary

Data Wrangling

Given the raw data from Black Saber Software, the key data used contains information about each employee's gender, their affiliated team, role seniority from least senior to most senior, leadership for level, distinct productivity rated on a 0-100 scale, and the salary level at each financial quarter in dollar. The primary purpose is to analyze whether there is a potential bias on the salary between males and females. To get the tidy data for the later modeling, the dollar signs are removed as well as the comma symbol between letters, so that we possess the numeric salary levels. Among a total of 6906 current employees, 59% are males, whereas 39% are females. The rest, 1.7%, prefer not to say, which is the minority. Those 117 cases are not considered as well.

Besides, under the variable role seniority, the order is inappropriate when fitting the model because the Director is at the top of the list. As a result, they are reordered to be listed from the least senior (Entry-level) to the most senior (Vice President).

Data Visualization

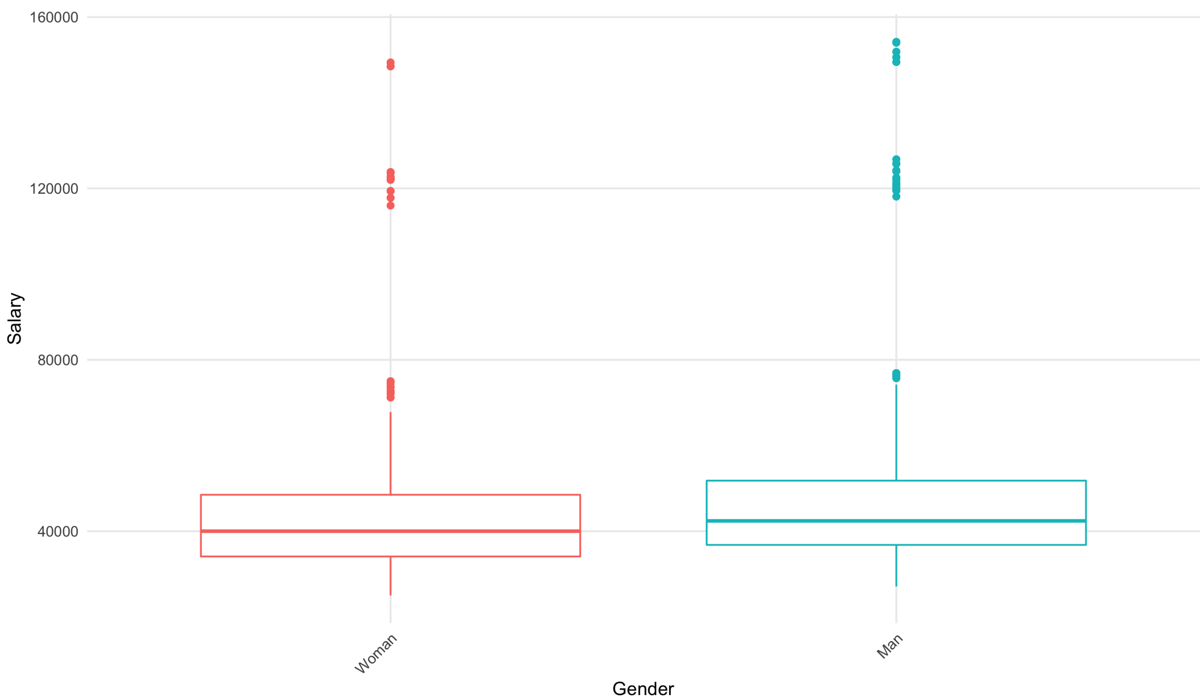


Figure 4: Boxplot of Salary vs. Gender

After all the essential data manipulations, a simple boxplot can be generated between gender and salary. According to the **Figure 4**, several outliers yield way more wages for both male and female employees. That is probably due to higher role seniority. Clearly, the median salary for males is higher than the one for females, yet by how much remains unknown.

Model

Three linear mixed effect models are fitted further to discuss a more precise relationship between salary and gender. This type of model allows both fixed- and random-effect terms; thus, it is considered the most suitable one.

In the very first basic model, gender, productivity, and a random effect intercept for the employee ID are included as variables. The second one is a nested model with an extra random effect intercept for the role seniority. To consider impacts from other variables, the third model contains the random intercept-slope that permits the explanatory variable gender to have a different effect for each level of role seniority.

While constructing a likelihood ratio test between model 1 and 2, it suggests that including an additional random effect for role seniority explains the data better. We have a shred of strong evidence against the hypothesis that the simplest model fits the data just as well. Similarly, when a likelihood ratio test is constructed between model 2 and 3, we obtain a piece of strong evidence that the third model fits the best since the additional random slope term explains the data better than the extra random effect for role seniority. Therefore, model 3 is the final model used for the research analysis. The model is shown below:

$$Y_{salary} = \beta_0 + \beta_{gender} X_{gender} + \beta_{productivity} X_{productivity} + \alpha_{employeeid} + \gamma_{roleseniority} X_{gender}$$

where β_0 is the overall intercept; β_{gender} is the slope for gender; $\beta_{productivity}$ is the slope for productivity; $\alpha_{employeeid}$ is the random intercept of Employee ID; $\gamma_{roleseniority} X_{gender}$ is the random slope term of gender

It is also essential to check whether the model achieves the following assumptions because it involves the model's validity. Under this specific case, salary is considered continuous, and variables chosen to be included in the model have certain connections with the response, thus being considered appropriate. Undeniably, the cases are independent. Since the estimates of random effect components are robust to violation of normal distribution, it may cause the model to be less precise, but the results remain unbiased.

Results

Table 5: Summary of Linear Mixed Model (Salary)

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	65497.187	14120.612	7.989	4.638	0.002
genderMan	2746.519	385.249	34.740	7.129	0.000
productivity	-2.360	1.124	6305.743	-2.100	0.036

The third model considers the gender male as the reference level. Based on the **Table 5**, the p-value for gender female is a significantly small number 2.7x10-8 which can be rounded to zero, indicating that the sample results are not consistent with the null hypothesis that gender has no effect on the salary level. Moreover, according to the **Appendix 8: Confidence Invertal of LMER (Salary Model)**, the 95% confidence interval ranges from -3504.96 to -1957.95, which does not include 0. Both p-value and confidence interval suggests that gender has a certain extent of an effect on salary. If we explore the coefficients, the estimate for the female gender is negative. Holding all else constant, it signifies that if the employee is a female, then the mean of her salary is $2746.513 lower relative to the mean of male's salary.

Conclusion

In general, the boxplot offers a rough idea that there is a potential bias between males' and females' salaries. A further linear mixed effect model statistically proves that gender has an effect on salary level. Specifically, the results powerfully show that the female yields $2746.513 lower than males on average across different role seniorities. Hence, it is undeniable that a gender bias in the remuneration process exists, and further actions are recommended to fix it.

Discussion

With the appropriate usage of statistical methods and models to analyze three main research questions, we BDSTA.CO discussed in detail whether there is a potential gender bias regarding the hiring, promotion, and remuneration process in Black Saber Software. Unfortunately, only the hiring process is fair between men and women, whereas the rest are not. In the hiring process, we mainly emphasized the result from the second round (phrase2 – phrase3) because rounds 1 and 3 obtain similar proportional decreases. By using a generalized linear model, not only the p-value is not statistically significant, but also the 95% confidence interval includes zero. Both measures imply that the second-round hiring process is free of gender discrimination under the AI algorithm. However, for both the promotion and salary process, the p-values are statistically significant. The 95% confidence intervals do not contain zero using the generalized linear model and the linear mixed effect model. Those measures suggest that the gender woman has a certain impact on the odds of whether she can get a promotion and to what extent her salary is. In particular, when a man and a woman have precisely the same time of tenancy, the odds of that man's promotion are about two times when the employee is a woman. Similarly, the female yields $2746.513 lower than males on average, holding all else fixed.

In conclusion, with all the hiring data for the new grad program and the data about promotion and salary for all the staff in Black Saber Software, we are able to offer an answer regarding the research questions of whether there is gender discrimination in the recruitment, promotion, and remuneration processes. Essentially, the hiring process is free of gender bias under the AI algorithm. Nevertheless, the promotion and salary levels are not fair between men and women. Specifically speaking, holding other variables constant, male employees are more likely to be upgraded than female employees, and they earn a higher salary on average than females.

Limitations

Yet, limitations exist when analyzing the hiring process. When we conduct comparisons between the models, we only focus on the AIC value for selecting the final one. Since we make assumptions on how the AI algorithm evaluates the candidate, there could be differences from reality. For instance, the AI may also consider the team each applicant applied for when assessing their scores, which could act as a random effect in the model.

Similarly, limitations also exist when analyzing the promotion process. Firstly, this study only considers whether the employee has been promoted before, rather than the promotion opportunities for different years or separate quarters. Though two employees may get promoted, the situation that people who enter the firm only one year is different from those who enter the firm several years. Based on the model, we have concluded that males are likely to get promoted.

However, we cannot investigate how easily male employees will get promoted compared with those female employees. Secondly, centered on the results, employees who stay longer in the company are more likely to be upgraded. This model cannot study whether female employees need to wait longer than male employees for the next opportunity after the first promotion. Thirdly, to analyze the previously mentioned hypothesis, using the binary variable is not comprehensive, introducing the total amount of preferment for further investigation. For the next step, we will investigate whether the change of leadership in two consecutive years will correspond to the promotion status for different genders. This will help to demonstrate whether the working status is related to the promotion opportunities.

Next Steps

- We will perform more models with different variable combinations, with or without random effects, to check whether there will be any difference for the gender effect if a random effect is added.
- We will investigate whether the change of leadership in two consecutive years will correspond to the promotion status for different genders. This will help to demonstrate whether the working status is related to the promotion opportunities.
- We will further explore the difference in salary for a specific role. It is possible that in entry level, the difference is not big, whereas the gap gets larger as the seniority raises

Consultant information

Consultant profiles

Huiyi Lu. Huiyi is a senior consultant with BDSTA.CO. She specializes in data analysis and consultation. She achieved her Bachelor of Arts, Specialist in Finance & Economics and Applied Statistics, from the University of Toronto in 2022.

Xinyi Qin. Xinyi is a senior consultant with BDSTA.CO. She specializes in data manipulation and data visualization. She achieved her Bachelor of Arts, Specialist in Financial Economics and Major in Statistics, from the University of Toronto in 2022.

Ying Xiong. Ying is a senior consultant with BDSTA.CO. She specializes in reproducible analysis and data visualization. Ying earned her Bachelor of Commerce, Specialist in Finance & Economics and Major in Statistics, from the University of Toronto in 2022.

Yixin Liang. Yixin is a senior consultant with BDSTA.CO. She specializes in data wrangling and data visualization. Yixin will receive her Bachelor of Commerce, Specialist in Finance & Economics and Applied Statistic Specialist, from the University of Toronto in 2022.

Code of ethical conduct

- We promise to contribute to and employ best practices throughout the consultation service.
- We safeguard the privacy of all employees, balancing the privacy rights with our needs to collect and analyze the information.
- We promise that all the conclusions and findings are given with fairness, impartiality and objectivity.

Appendix

1. Logistic Regression for 1st Round Hiring

Summary of Logit Regression (1st Hiring)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-15.777	2.630	-5.998	0.000
genderWoman	0.182	0.542	0.335	0.737
gpa	5.948	1.013	5.872	0.000
as.factor(work_experience)1	4.099	0.726	5.644	0.000
as.factor(work_experience)2	15.645	1262.795	0.012	0.990

2. Linear Regression for 3rd Round Hiring

Summary of Logit Regression (3rd Hiring)

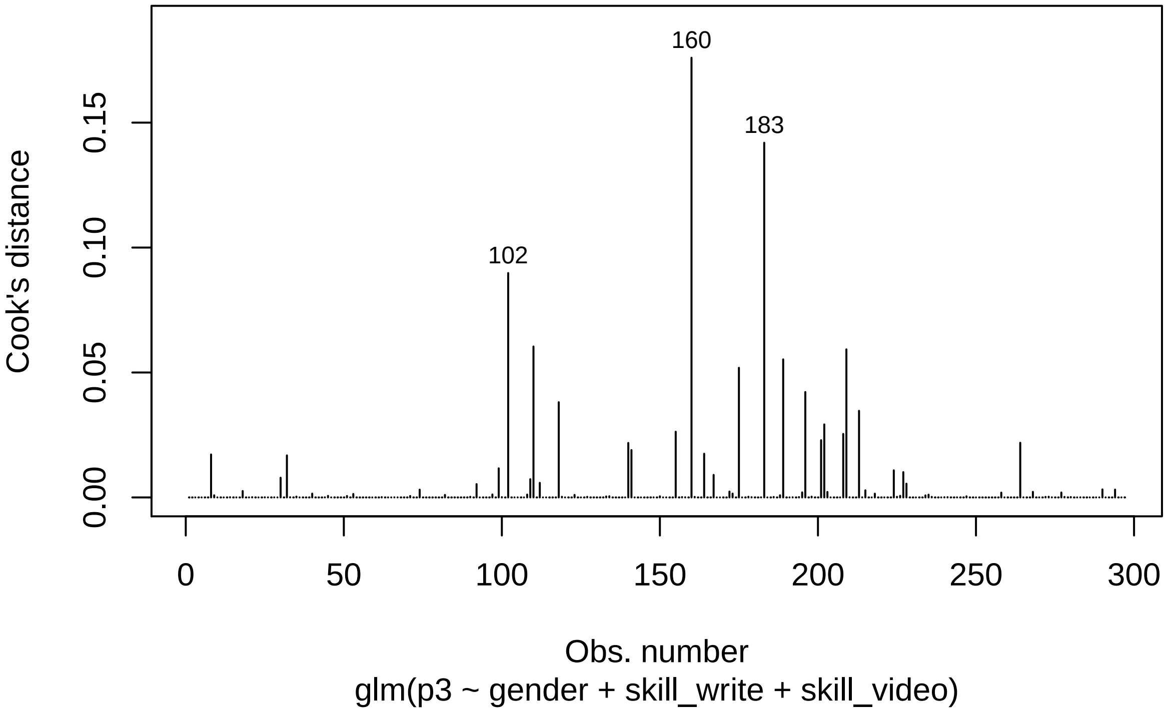
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.098	17.212	0.703	0.493
genderWoman	2.117	2.604	0.813	0.429
gpa	0.655	2.532	0.259	0.799
skill_write	0.765	0.086	8.872	0.000
skill_video	4.282	0.723	5.919	0.000
as.factor(work_experience)1	-13.476	5.842	-2.307	0.036
as.factor(work_experience)2	-16.531	6.448	-2.564	0.022

3. Multicollinearity (2nd Hiring Model)

VIFs for 3 Predictors (2nd Hiring Model)

	x
gender	1.157
skill_write	2.022
skill_video	1.854

4. Influential Point (2nd Hiring Model)

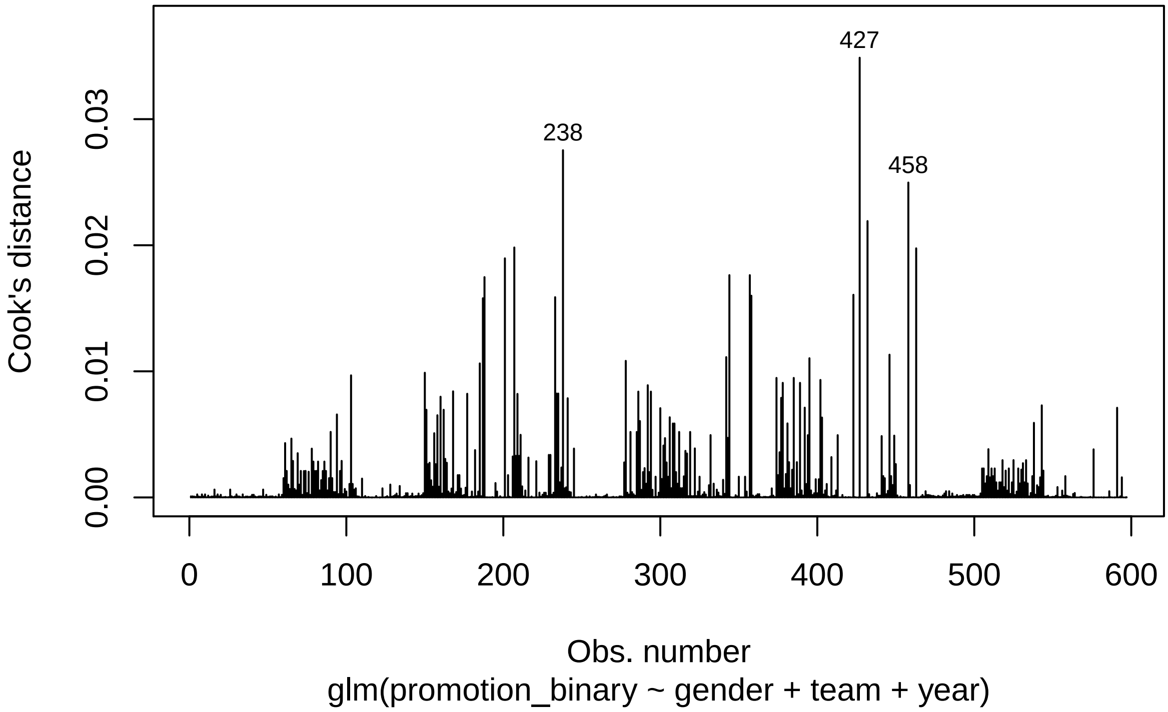


5. Multicollinearity (Promotion Model)

VIFs for 3 Predictors (Promotion Model)

	GVIF	Df	GVIF^(1/(2*Df))
gender	1.095	1	1.046
team	1.148	7	1.010
year	1.058	1	1.029

6. Influential Point (Promotion Model)



7. Confidence Interval of Logit Regression (Promotion Model)

Confidence Invertal of Logit Regression (Promotion)

	2.5 %	97.5 %
(Intercept)	-3.749	-2.333
genderWoman	-1.190	-0.186
teamData	-0.925	0.722
teamDesign	-1.170	1.956
teamLegal and financial	-1.066	1.348
teamMarketing and sales	-1.494	-0.002
teamOperations	-1.110	0.574
teamPeople and talent	-1.225	1.484
teamSoftware	-0.741	0.686
year	1.141	1.559

8. Confidence Interval of LMER (Salary Model)

Confidence Invertal of LMER (Salary)

	2.5 %	97.5 %
.sig01	3235.603	3629.734
.sig02	26752.859	68905.527
.sig03	-0.893	0.200
.sig04	433.739	1301.492
.sigma	948.472	982.519
(Intercept)	36356.696	94641.003
genderMan	1957.919	3504.961
productivity	-4.562	-0.155

References

1. Ben Bolker and David Robinson (2020). broom.mixed: Tidying Methods for Mixed Models. R package version 0.2.6. https://CRAN.R-project.org/package=broom.mixed
2. Benjamin Nutter (2021). pixiedust: Tables so Beautifully Fine-Tuned You Will Believe It's Magic. R package version 0.9.1. https://CRAN.R-project.org/package=pixiedust
3. Bob Rudis (2020). hrbrthemes: Additional Themes, Theme Components and Utilities for "ggplot2". R package version 0.8.0. https://CRAN.R-project.org/package=hrbrthemes
4. Daniel D. Sjoberg, Michael Curry, Margie Hannum, Joseph Larmarange, Karissa Whiting and Emily C. Zabor (2021). gtsummary: Presentation-Ready Data Summary and Analytic Result Tables. R package version 1.4.0. https://CRAN.R-project.org/package=gtsummary
5. Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.
6. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
7. Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2021). dplyr: A Grammar of Data Manipulation. R package version 1.0.5. https://CRAN.R-project.org/package=dplyr
8. John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/
9. Kazuki Yoshida and Alexander Bartel (2020). tableone: Create "Table 1" to Describe Baseline Characteristics with or without Propensity Score Weights. R package version 0.12.0. https://CRAN.R-project.org/package=tableone
10. Kuznetsova A, Brockhoff PB, Christensen RHB (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, *82*(13), 1-26. doi: 10.18637/jss.v082.i13 (URL: https://doi.org/10.18637/jss.v082.i13).
11. Legislationline Organization. (2021). Gender equality. Retrieved April 21, 2021, from https://www.legislationline.org/topics/topic/7
12. Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595
13. Richard Iannone, Joe Cheng and Barret Schloerke (2020). gt: Easily Create Presentation-Ready Display Tables. R package version 0.2.2. https://CRAN.R-project.org/package=gt

14. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

15. Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich

16. Yihui Xie (2020). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.30.

17. Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963