

Contents

Executive summary	3
Technical report	5
Introduction	5
Investigation on the factors and potential gender bias of being hired	6
The potential gender bias towards females on the promotion	23
The difference in employees' salaries between males and females	31
Discussion	40
Consultant information	43
Consultant profiles	43
Code of ethical conduct	43

Executive summary

Background and aims

Influenced by the traditional male-dominated job market, gender bias and gender stereotypes may still exist in the workplace. Recently, negative news on the gender discrimination in the recruitment and remuneration processes of Black Saber Software Company is reported. In order to dismiss the public concerns on the potential gender bias, we plan to explore whether the gender bias actually exists in the Black Saber Software. We are provided the data about the new hiring program and the information of current employees by the vice president Valin Hess from the Data Team, Black Saber Software. This report provides an analysis of factors and bias that will influence the hiring process and an examination of the potential gender bias on promotion and salary distribution. The results demonstrate gender is not the determining factor of being hired by the company and other factors like their academic achievement, professional competence in their skills and leadership have become more important considerations in the hiring process. However, there exists potential gender bias on female applicants in the promotion and salary distribution since females are less likely to get promoted to a higher place and females receive fewer salary.

Key Findings

- During the hiring process, female applicants face higher rejection rates of 50.8% and 52.8% respectively in Phase 1 and Phase 2 and males have relatively lower rejection rates. This indicates women have slightly higher probability of being rejected in Phase 1 and Phase 2 than men (Figure 1.1).
- For the 10 applicants who were hired, they all received an average interviewing score that is higher than 76.5. This indicates average interviewing score is the determining factor of being hired by the company in the last phase, while gender does not influence the final hiring phase.
- Female employees with higher productivity do not tend to obtain more opportunities on promotion and this is an unfair treatment towards women. Conversely, some of the male employees with low productivity got promoted for more than 4 times.
- There is less likely for female employees to get promoted more than twice than males. This illustrates that female employees are difficult to get a higher position in the company.
- The median salary of female employees is 2400 dollars lower than male employees in the annual quarter. Employees with a high salary are mainly dominated by males and thus there exists wage differences among males and females.
- Except for the team of “people and talent”, the average salaries for female employees in all the teams are lower than male employees and thus it demonstrates there is potential gender bias towards females (Table 3.1).

Limitations

- Since only a single dataset is provided for the hiring process, there may exist randomness in data and thus the final result may not reflect the actual outcome.
- For the current employees, the ratings on their leadership are given by people instead of machines. There may exist potential bias that is caused by the judges' personal perspectives.

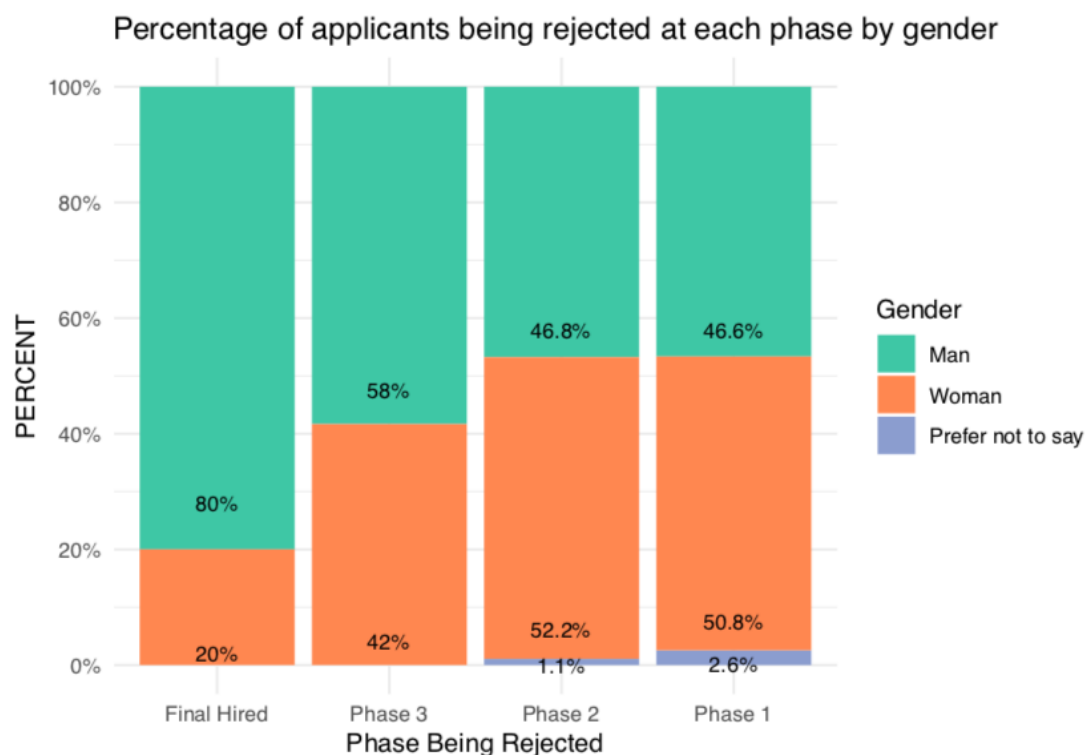


Figure 1.1. The bar chart illustrates the percentage of male and female applicants that were rejected at each phase.

Table 3.1. The table shows the average salary in each team by gender.

	Salary for Male	Salary for Female
Team Client services	43956.36	35414.49
Team Data	51954.29	43482.05
Team Design	64083.33	40350.00
Team Legal and financial	54871.43	53707.69
Team Marketing and sales	46490.16	38783.72
Team Operations	50166.67	43702.86
Team People and talent	46800.00	53094.12
Team Software	46796.12	44829.63

Technical report

Introduction

Recently, the Black Saber Software company receives concerns about potential bias in their hiring and remuneration processes. We aim to study whether the hiring process, promotion and salary distribution are fair in Black Saber Software company. The first dataset is focused on a three-phase hiring process with 613 applicants, including 291 males, 311 females and 11 unknown. For the second dataset, it consists the information of 607 current employees. The study includes three research questions, analyzing potential bias in the hiring, promotion and salary, respectively. Three different models are used for studying the research questions. The study uses binary logistic models to analyze the hiring process, generalized linear mixed models to analyze the promotion and linear mixed models to analyze the salary. In the research problem of hiring, some important factors like gender, gpa, working experience and relevant skills are investigated to reveal whether the gender bias exists in the hiring process and what will affect the final hire. In the research question of promotion, variables like gender, average productivity and team are chosen to study whether the company will promote the current employees based on their genders, the contribution to the company, and the team they belong to. In research question for analyzing salary, some variables like gender, team and role seniority to see what factors will affect the employees' salary.

Research questions

- During the recruitment process, what are the determining factors of being successfully hired by Black Saber Software?
- How effective does gender parity act on the promotion of the current employees in Black Saber Software?
- In Black Saber Software, what are the difference in salary distribution between female and male employees?

Investigation on the factors and potential gender bias of being hired

Methods

In research question 1, we mainly focus on the hiring process and we aim to investigate what are the determining factors of being hired by the Black Saber Software Company. Since we also want to explore the gender bias in each phase of the recruitment process, we need to find out whether gender is also a determining factor of being hired. The recruitment process of the company consists of three phases: (1) In phase 1, it was the initial hiring process with 613 applicants' basic information like their gender, GPA, working experience and extracurriculars, (2) In phase 2, only 300 applicants entered and it also contains the assessment on their technical skills, writing skills, speaking skills and leadership presence, (3) In phase 3, only 22 applicants entered and they would be given interviewing scores by two different interviewers. After these three recruitment phases, only 10 applicants entered the final hire.

The original datasets provided the information of the applicants in phase 1, phase 2, phase 3 and final hire. We create a categorical variable "Result" to indicate the result of entry. If the applicant successfully enters the next phase, then "Result" will be "Proceed". If the applicant fails to enter the next phase, then "Result" will be "No". Thus, we have our dependent categorical variable as the result of entry ("Proceed" or "No") and we have several independent variables. The binary logistic model will be the best model to explain the relationship of the categorical dependent variables and several independent variables. We have a categorical independent variable like gender and other independent variables are numeric. The binary logistic model will predict the logit transformation of the probability of entering the next phase. Thus, we build three binary logistic models for each phase where model 1 is for phase 1, model 2 is for phase 2 and model 3 is for phase 3. The models will show us how these factors(gender, working experience, leadership, and etc.) will influence the result of entry and it will reflect the determining factors of being successfully hired by the company. If gender is not the determining factor and will not influence the hiring, then we could conclude the company has no gender bias during the recruitment process.

Model 1

The given dataset of Phase 1 provided the basic information of applicants like their GPA, gender, extracurricular involvement and work experience. Unfortunately, 313 of the applicants were phased out at Phase 1 and only 300 of them left. Since we want to investigate what factors will influence their entry from phase 1 to phase 2, we need to have a binary dependent variable that could indicate the result of entry. Therefore, we created a new variable (Result) to show whether the applicants entered phase 2 where "Proceed" indicates they entered the next phase and "No" indicates they failed to enter. The purpose of this model is to analyse how GPA, gender, working

experience and extracurricular would influence the applicants' entry of Phase 2. Additionally, we aim to find out whether gender is influential to the applicants' entry of Phase 2. Therefore, we have our null hypothesis that GPA, working experience, gender or extracurriculars do not influence probability of entering Phase 2. The binary logistic model will be the most appropriate model to analyze the relationship between a binary dependent variable and a set of independent variables.

By building a binary logistic model (model 1),

$$\text{logit}(p) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$$

p is the odds ratio where

$$p = \frac{p}{1-p} = \frac{\text{probability of entering Phase2}}{\text{probability of failing to enter Phase2}}$$

x_1 is the numeric independent variable that represents the GPA of each applicant that ranged from 0.0 to 4.0. x_2 and x_3 are the categorical independent variable which represents gender as female and as preferred not to say respectively. x_4 is the score of extracurricular involvement of applicants where 0 is the no extracurricular, 1 is some relevant extracurriculars and 2 represents high relevance extracurriculars. x_5 is the score of work experience of applicants where 0 is the no working experience, 1 is some relevant working experience and 2 represents high relevance working experience.

Model 2

Similarly, we use the given dataset of Phase 2 which provided applicants' basic information, applicants' skill assessment scores and rating on their leadership presence. There are 22 applicants entered Phase 3 and 278 applicants are eliminated at Phase 2. Since we want to investigate the factors that will affect their entry from phase 2 to phase 3, we also need to have a binary dependent variable that could indicate the result of entry. Thus, we created a new variable (Result) like the method used in the first model to show whether the applicants entered phase 3 where "Proceed" indicates they entered phase 3 and "No" indicates they failed to enter. The purpose of this model is to analyse how GPA, gender, working experience, extracurriculars, technical skills, writing skills, speaking skills and leadership presence would influence the applicants' entry of Phase 3. Additionally, we aim to find out whether gender is influential to the applicants' entry of Phase 3. Thus, we have our null hypothesis that GPA, working experience, gender, extracurriculars, technical skills, writing skills, speaking skills and leadership presence do not influence the entry of Phase 3. The binary logistic model will be the most appropriate model to analyse the relationship between a binary dependent variable and a set of independent variables.

By building a binary logistic model (model 2),

$$\text{logit}(p) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9$$

p is the odds ratio where

$$p = \frac{p}{1-p} = \frac{\text{probability of entering Phase3}}{\text{probability of failing to enter Phase3}}$$

x_1 - x_5 are the same variables as shown in model 1. x_1 is the numeric independent variable that represents the GPA of each applicant that ranged from 0.0 to 4.0. x_2 and x_3 are the categorical independent variable which represents gender as female and as preferred not to say respectively. x_4 is the score of extracurricular involvement of applicants where 0 is the no extracurricular, 1 is some relevant extracurriculars and 2 represents high relevance extracurriculars. x_5 is the score of work experience of applicants where 0 is the no working experience, 1 is some relevant working experience and 2 represents high relevance working experience.

x_6 and x_7 are numeric independent variables that represent the score on technical skills and writing skills of each applicant respectively and the score is ranged from 0 to 100. x_8 is the rating of leadership presence from a 1-to-10 scale where 10 represents strongest leadership presence and 0 represents absence of leadership presence. x_9 is the rating of speaking skills from a 1-to-10 scale where 10 represents the strongest speaking ability and 0 represents the weakest speaking ability.

Model 3

Though 22 applicants entered Phase 3, only 10 of these applicants successfully entered the final hiring stage. Since the given dataset of Phase 3 only contains the information of the interviewing scores, we join the dataset of Phase 3 and Phase 2 together by the applicant id. This provides us with a complete dataset including their basic information, skill assessment score, leadership presence and interviewing scores of applicants at Phase 3. Similarly, we want to investigate the factors that will affect their entry from phase 3 to final hiring, we also need to have a binary dependent variable that could indicate the result of entry. We also created a new variable (Result) to show whether the applicants entered phase 3 where “Proceed” indicates they entered the final hiring phase and “No” indicates they failed to enter. The purpose of this model is to analyse how GPA, gender, working experience, extracurriculars, technical skills, writing skills, speaking skills, leadership presence and interviewing scores would influence the applicants’ entry of Final Hiring Phase. Additionally, we aim to find out whether gender is influential to the applicants’ entry of the final hiring phase. Thus, the binary logistic model will be the most appropriate model to analyze the relationship between a binary dependent variable and a set of independent variables.

By building a binary logistic model (model 3),

$$\text{logit}(p) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9 + b_{10}x_{10}$$

p is the odds ratio where

$$p = \frac{p}{1-p} = \frac{\text{probability of entering Final Hire Phase}}{\text{probability of failing to enter Final Hire Phase}}$$

Since it does not contain the variable for gender as prefer not say. x_1 is the same variables as shown in model 2. x_2 is the categorical independent variable which represents gender as female. x_3 is the score of extracurricular involvement of applicants where 0 is the no extracurricular, 1 is some relevant extracurriculars and 2 represents high relevance extracurriculars. x_4 is the score of work experience of applicants where 0 is the no working experience, 1 is some relevant working experience and 2 represents high relevance working experience. x_5 and x_6 are numeric independent variables that represent the score on technical skills and writing skills of each applicant respectively and the score is ranged from 0 to 100. x_7 is the rating of leadership presence from a 1-to-10 scale where 10 represents strongest leadership presence and 0 represents absence of leadership presence. x_8 is the rating of speaking skills from a 1-to-10 scale where 10 represents the strongest speaking ability and 0 represents the weakest speaking ability. x_9 represents the interviewing score given by the first interviewer on a scale of 0 to 100 and x_{10} represents the interviewing score given by the second interviewer on a scale of 0 to 100.

Result

Interpretation of model 1

Table 1.1: The table shows the estimated coefficient and the 95% confidence interval of variables in model 1.

	Estimated coefficient	95% confidence interval
baseline odds(b_0)	-6.40	(-7.4597669, -5.3477224)
GPA(x_1)	1.98	(1.5618904, 2.3897057)
genderPrefer not to say(x_2)	-0.58	(-2.1396362, 0.9818577)
genderWoman(x_3)	-0.06	(-0.4644362, 0.3352827)
extracurriculars(x_4)	0.26	(-0.1208189, 0.6496537)
work experience(x_5)	0.76	(0.2526958, 1.2618562)

From Table 1.1, b_0 is the intercept which indicates the log of odds of all the applicants entering Phase 2 with no predictor variables in the model and it has a value of -6.4037. We can also convert the log of odds into simple odds by taking $\exp(-6.40374) = 0.00165$. The number indicates that the odds of an individual applicant entering phase 2 decreases by $(1-0.00165) = 99.8\%$ if we have no predictor variables.

The coefficient for GPA (x_1) is 1.97580 in the logit of odds terms and we can convert it into odds by taking $\exp(1.9758) = 7.212$. This indicates as the GPA increases 0.1 unit, the odds of an individual applicant entering phase 2 will increase by $(7.211-1)*0.1 = 62.1\%$.

Out of three levels in gender, male is the referenced variable. For the variable gender as female(x_3), the coefficient for this variable is -0.06458 and the odds is $\exp(-0.06458) = 0.9375$. This illustrates the odds of an individual female applicant entering phase 2 decreasing by $(1-0.9375) = 6.25\%$ than the applicants with gender as male. Similarly, the coefficient for variable gender as preferred not to say(x_2) is -0.57889 and the odds for this variable is $\exp(-0.57889) = 0.56$. This indicates odds of an individual applicant with unknown gender entering phase 2 decreasing by $(1-0.57889) = 43.94\%$ than the male applicants.

The coefficient for the work experience (x_4) is 0.75728 and odds is $\exp(0.75728) = 2.13$. This indicates as the score of work experience increases one unit, the odds of an individual applicant entering phase 2 will increase by $(2.13-1) = 113.24\%$. For the variable of extracurricular involvement(x_5), it has the coefficient of 0.2644 and odds of $\exp(0.2644) = 1.302$. This demonstrates as the score of extracurricular involvement increases one unit, the odds of an individual applicant entering phase 2 will increase by $(1.302-1) = 30.2\%$.

For model 1, we have our null hypothesis that GPA, working experience, gender or extracurriculars do not influence probability of entering Phase 2. Since the p-value of GPA and working experience is smaller than a significant level of 0.05, this illustrates GPA and working experience is highly related to the entry of Phase 2. For 0.1 unit increasement in GPA, the odds of an individual applicant entering phase 2 will increase 62.1%. Moreover, as the score of working experience increases one unit, the probability of an individual applicant entering phase 2 will increase by 113.24%. This means if the applicants have abundant working experience and high GPA, it will largely increase the probability of entering Phase 2. Conversely, extracurricular involvement and gender have relatively larger p-value and thus we have no evidence to reject the null hypothesis that gender and extracurriculars would not affect the entry of Phase 2. Therefore, in the process of entering Phase 2, gender and extracurriculars are not the determining factor for the entry of Phase 2. This provides us the information that GPA and working experience are the determining factors for both male and female applicants to successfully enter Phase 2.

Furthermore, we can also calculate the 95% confidence interval for these variables. Since the 95% confidence interval does not contain zero for GPA and working experience, this indicates GPA

and working experience will affect the entry of Phase 2 and thus it may influence the probability of being successfully hired. However, the 95% confidence interval contains zero for gender and extracurricular and this means gender and extracurricular will not contribute to the successful entry of Phase 2. Thus, GPA and working experience could increase the probability of entering Phase2 while gender and extracurriculars could not.

Interpretation of model 2

Table 1.2: The table shows the estimated coefficient and the 95% confidence interval of variables in model 2.

	Estimated coefficient	95% confidence interval
baseline odds(b0)	-19.90	(-2.761374e+01, -12.1880805)
GPA(x1)	-0.77	(-2.270682e+00, 0.7300574)
genderPrefer not to say(x2)	-16.62	(-4.015097e+03, 3981.8671827)
genderWoman(x3)	-0.52	(-2.016804e+00, 0.9715275)
extracurriculars(x4)	-0.36	(-1.755065e+00, 1.0397073)
work experience(x5)	0.01	(-1.507837e+00, 1.5313010)
technical skills(x6)	0.09	(4.541985e-02, 0.1426114)
writing skills(x7)	0.10	(5.154750e-02, 0.1554407)
leadership presence(x8)	0.97	(5.438201e-01, 1.4049595)
speaking skills(x9)	0.76	(4.092769e-01, 1.1169932)

In model 2 , we have our null hypothesis that GPA, working experience, gender, extracurriculars, technical skills, writing skills, speaking skills and leadership presence do not influence the entry of Phase 3. Since the p-value of technical skills, writing skills, speaking skills and leadership presence are smaller than a significant level of 0.05, this illustrates these skills and leadership presence are highly related to the entry of Phase 3.

From Table 1.2, the coefficient for the technical skills (x6) is 0.09402 and odds is $\exp(0.08620) = 1.098582$. This indicates as the score of technical skills increases one unit, the odds of an individual applicant entering phase 3 will increase by $(1.098582-1) = 9.86\%$. For the variable of writing skills (x7), it has the coefficient of 0.10349 and odds of $\exp(0.10349) = 1.109035$. This demonstrates as the score of writing skills increases one unit ,the odds of an individual applicant entering phase 3 will increase by $(1.109035-1) = 10.9\%$.

The coefficient for leadership presence (x8) 0.97439 and odds is $\exp(0.92488) = 2.64955$. This

shows as the rating of leadership presence increases one unit, the odds of an individual applicant entering phase 3 will increase by $(2.64955-1) = 165\%$. For the variable of speaking skills (x9), it has the coefficient of 0.76314 and odds of $\exp(0.76314) = 2.145001$. This demonstrates as the score of speaking skills increases one unit, the odds of an individual applicant entering phase 3 will increase by $(2.145001 - 1) = 114.5\%$.

Conversely, GPA, working experience, gender and extracurriculars have relatively larger p-value and thus we have no evidence to show GPA, working experience, gender and extracurriculars have influence on entering Phase 3. Though GPA and working experience are highly important in entering Phase 2, they do not have any impact on the entry of Phase 3. Additionally, gender is not the factor that will influence the result of entering phase 3. This indicates the result of whether entering phase 3 or not is not determined by applicants' gender, but highly related to their score or rating on technical skills, writing skills, speaking skills and leadership presence.

Interpretion of Model 3

Table 1.3: The table shows the estimated coefficient and the p-values of variables in model 3.

	Estimated coefficient	p value
baseline odds(b0)	-530.52	1
GPA(x1)	-3.41	1
genderWoman(x2)	-58.17	1
extracurriculars(x3)	8.32	1
work experience(x4)	-9.56	1
technical skills(x5)	-16.95	1
writing skills(x6)	-17.45	1
leadership presence(x7)	-163.65	1
speaking skills(x8)	-165.79	1
first interviewing score(x9)	60.58	1
second interviewing score(x10)	3.16	1

We have our null hypothesis that GPA, working experience, gender, extracurriculars, technical skills, writing skills, speaking skills, leadership presence and interviewing scores do not influence the entry of Phase 3. Since all the variables in the model 3 have p-values that are larger than significant levels of 0.05, this indicates all the variables are non-significant. Therefore, GPA, gender, working experience, extracurriculars, technical skills, writing skills, speaking

skills, leadership presence and interviewing scores will not influence the entry of the final hiring phase. The p values are approached 1 and this may be caused by the small sample size and we still need to examine the factors that will influence the final hiring phase by plotting the figures.

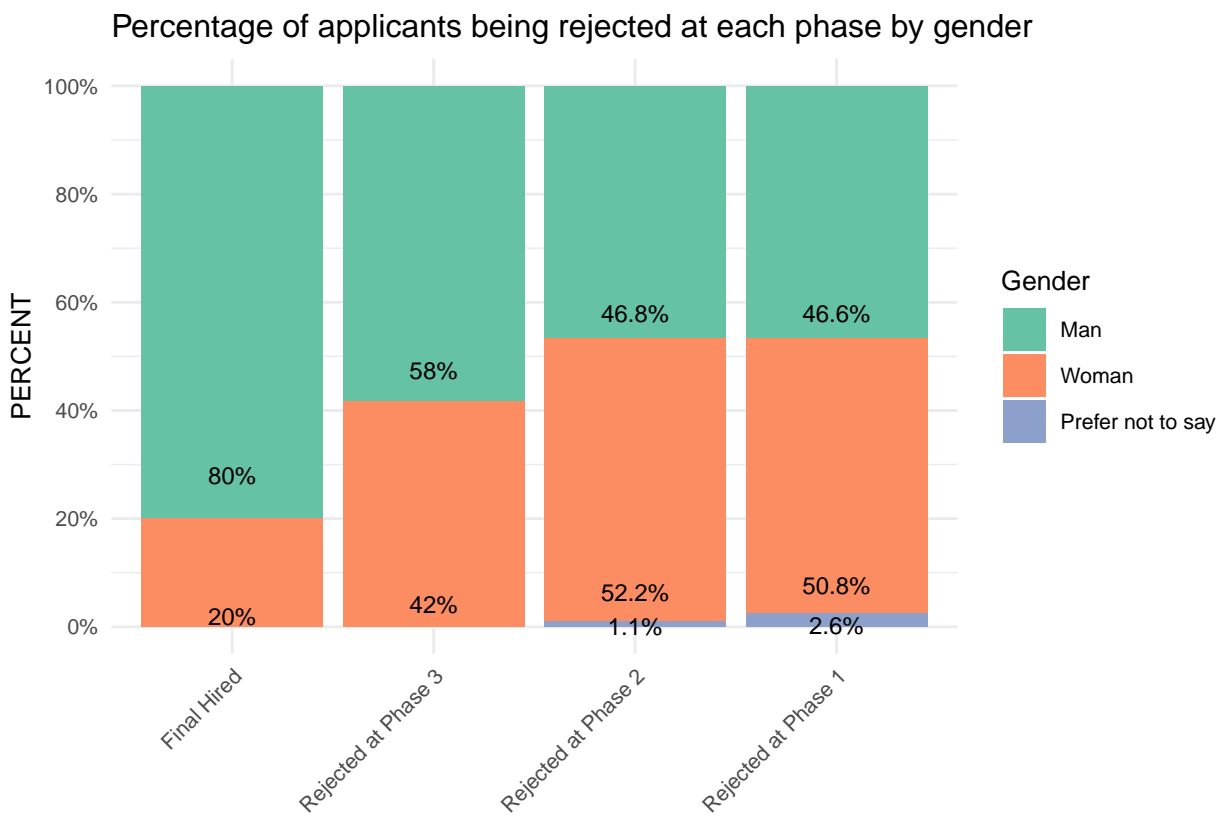
Figure and Tables

Figure 1.1: figure 1.1 illustrates the percentage of male and female applicants that were rejected at each phase.

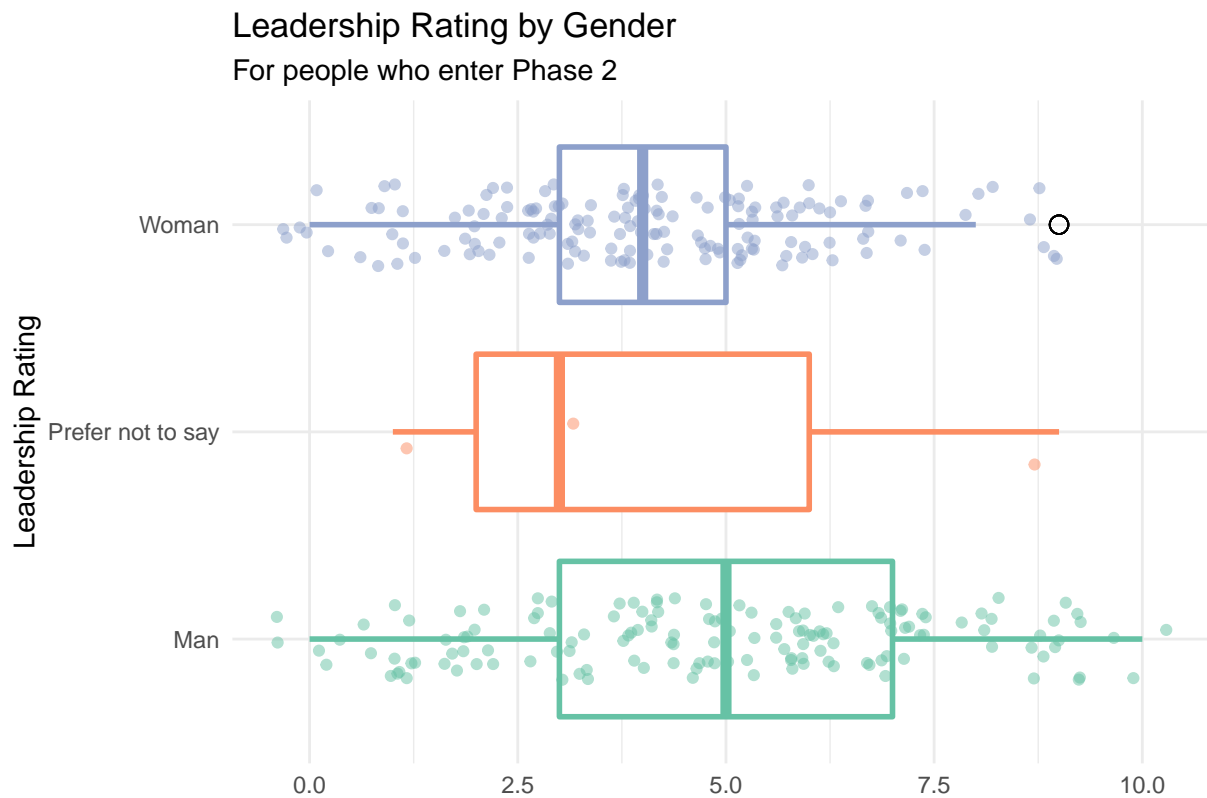


Figure 1.2: Figure 1.2 demonstrates the rating of their leadership presence based on their gender.

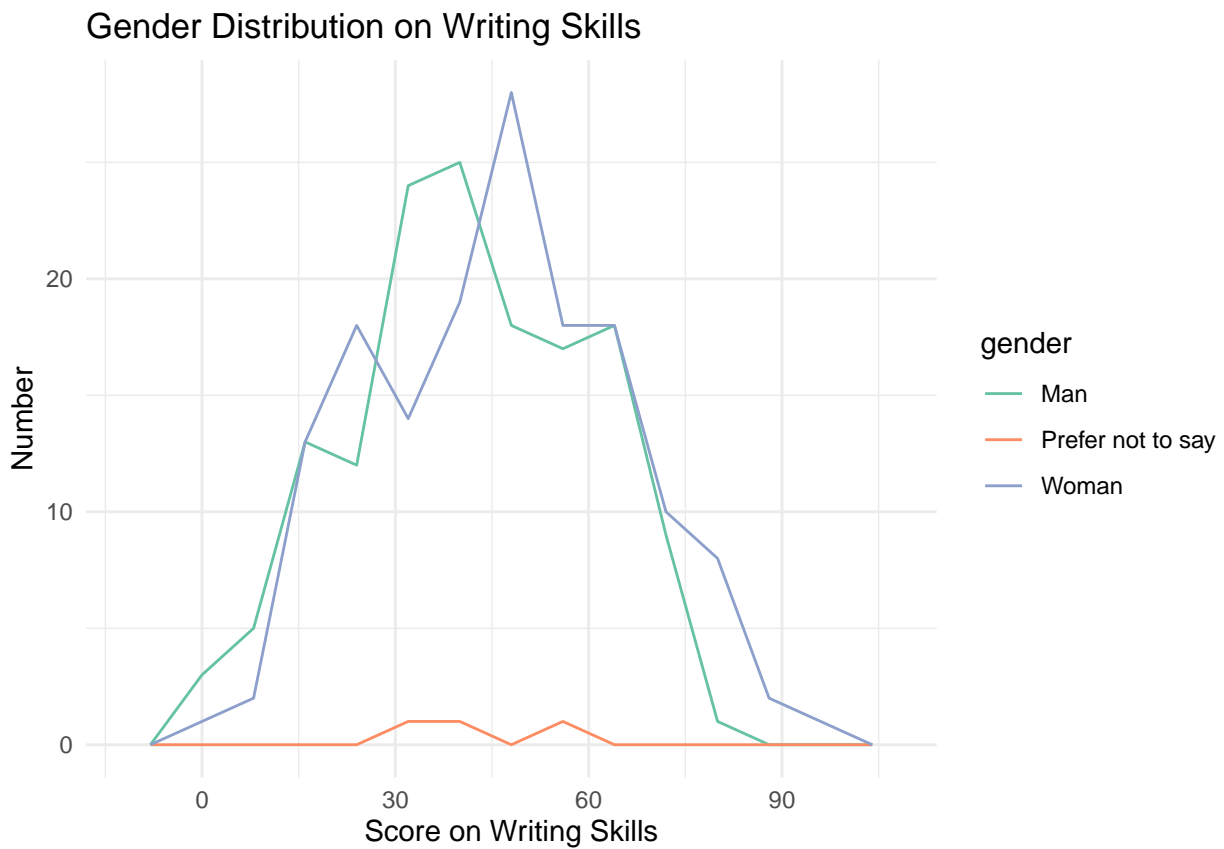


Figure 1.3: Figure 1.3 demonstrates the score of their writing skills based on their gender

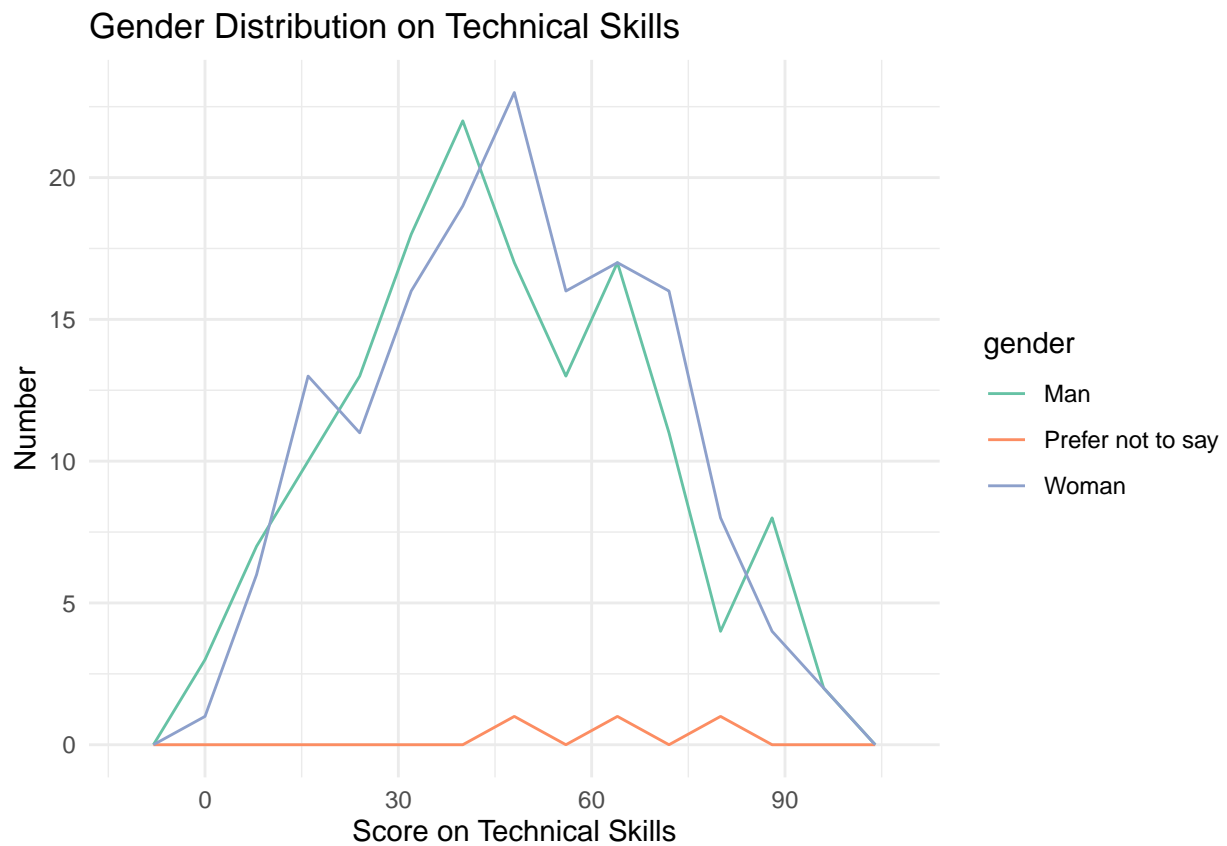


Figure 1.4: Figure 1.4 demonstrates the score of their technical skills based on their gender

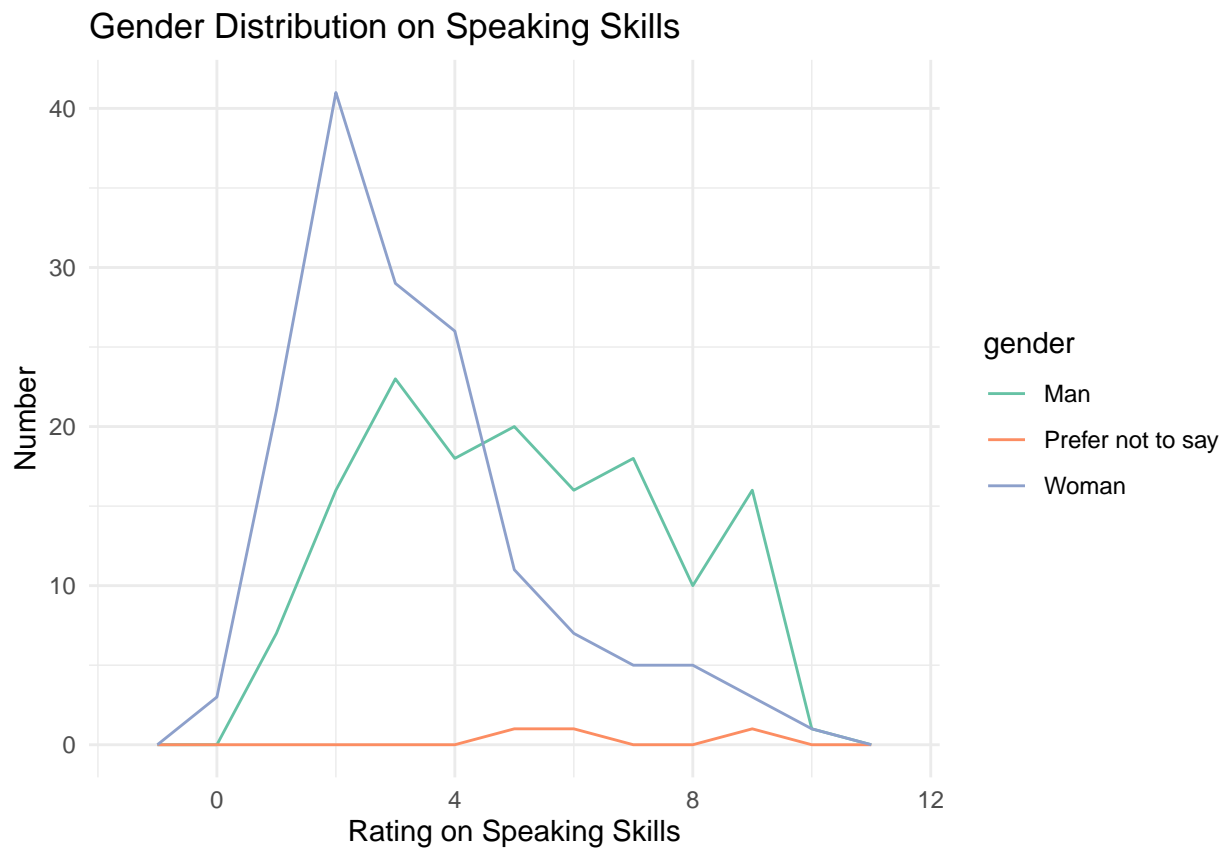


Figure 1.5: Figure 1.5 demonstrates the score of their speaking skills based on their gender

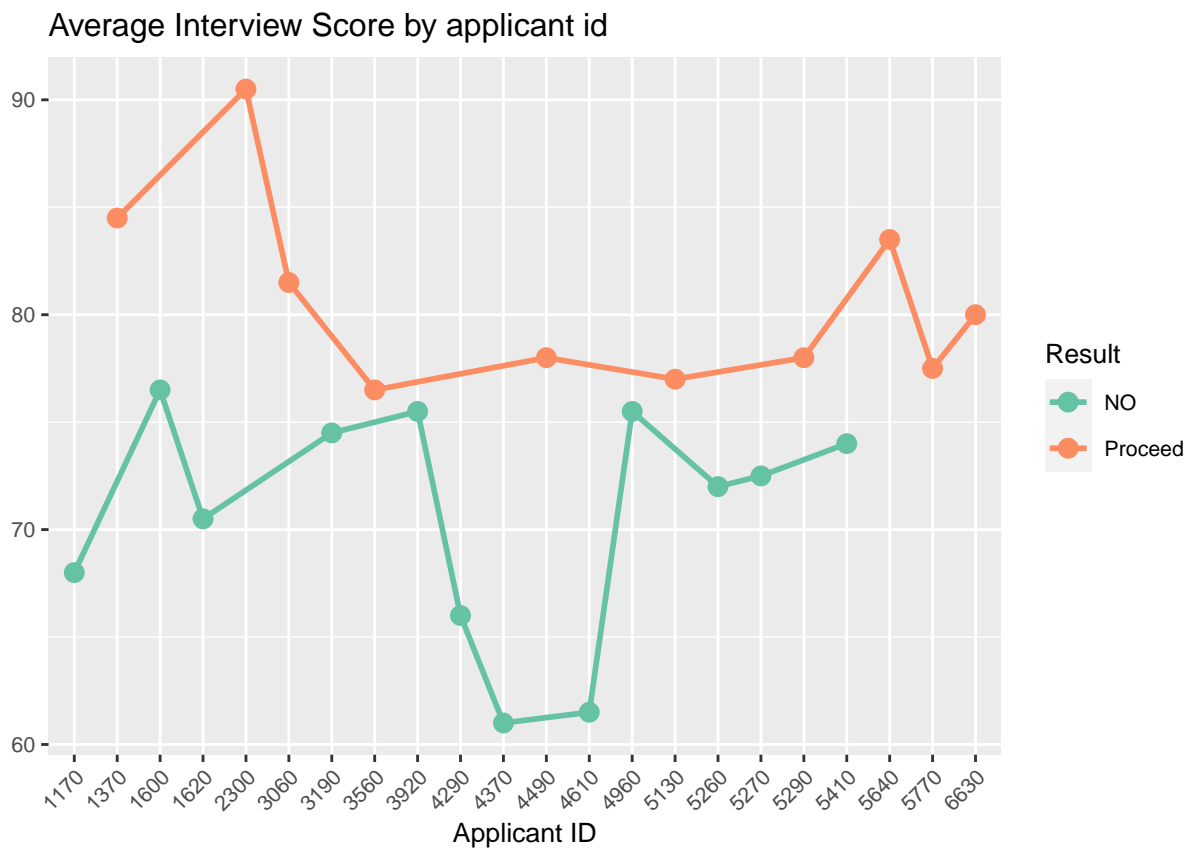


Figure 1.6: Figure 1.6 demonstrates the applicants' average interviewing scores based on whether they were hired or not.

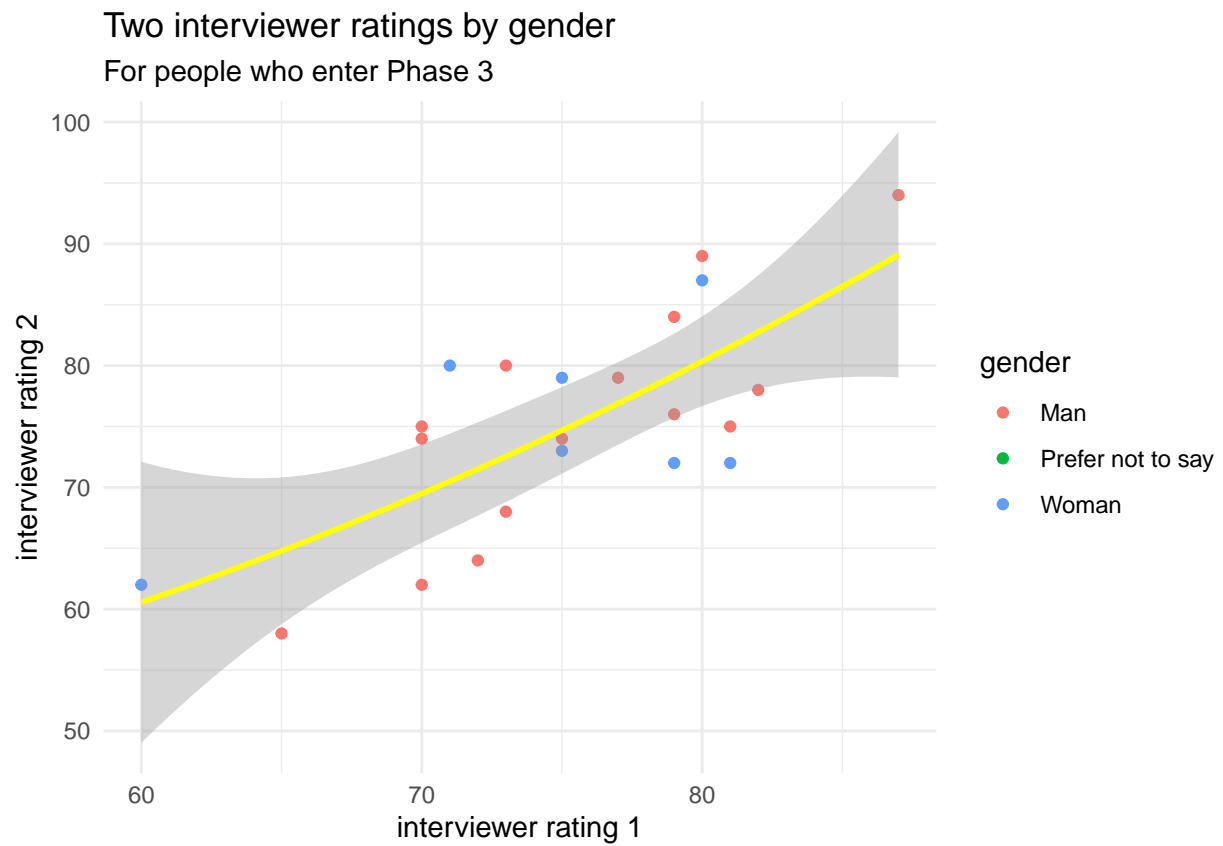


Figure 1.7: Figure 1.7 demonstrates the interviewing scores giving by two interviewers based on their gender.

Interpretion of Figures

By plotting the bar chart of being rejected in each phase, Figure 1.1 illustrates the percentage of male and female applicants that were rejected at each phase. In the process of entering phase 2 and phase 3, women had a slightly higher rejection rate than male applicants. By the conclusion drawn from Model 1, gender is not the determining factor of entering Phase 2 and thus this slight higher rejection rate in female applicants is acceptable. Therefore, gender is not considered as a potential bias that will influence the probability of being eliminated in Phase 1 or entering Phase 2.

Though the rejection rate for female applicants at Phase 2 was also higher than male, this could not directly indicate there exists gender bias in the entry of Phase 3. We still need to examine whether there are potential biases in gender by comparing their score on technical skills, writing skills, speaking skills and leadership presence.

For applicants who had entered Phase 2, the boxplot (Figure 1.2) demonstrates the rating of their leadership presence based on their gender. Since there are only 3 applicants with unknown gender, we mainly focused on the male and female applicants. Majority of the ratings for female applicants are in the range between 2.6 to 5 and only around 25% of them exceed a rating of 5. Conversely, the median rating of male applicants already reached 5 which is significantly higher than the average median score of female. This indicates male applicants had better overall performances in leadership presence and female applicants still needed some improvements in their leadership presence.

Conversely, female applicants performed better on the rating of speaking skills by observing the line graph (Figure 1.5). Only male applicants received the lowest rating of 0 and high ratings that above 8 were dominated by female applicants. This leads to the opposite result as the rating on leadership presence and this illustrates female applicants achieved better performance than male in their speaking skills. Additionally, according to Figure 1.4 and 1.3, the score on technical skills and writing skills for both male and female applicants are normally distributed with a mean score between 45 and 55. There is no distinct difference in score on technical skills and writing skills between male and female. Since ratings on speaking skills and leadership presence are the main influences on the entry of Phase 3, the better performance of male applicants in leadership presence and higher rating of female applicants in speaking skills balanced with each other. Therefore, in the process of entering Phase 3, male and female applicants show a similar level of performance and there is no obvious gender bias since the scores were given by AI.

By measuring their average interviewing scores from two interviewers, we plot a line chart (Figure 1.6) to demonstrate the applicants' average interviewing scores based on whether they were hired or not. The orange line represents the applicants that were hired by the company and the green line represents the applicants that were eliminated at phase 3. For applicants who were

hired (orange line), all of them have an average interviewing score that is higher than 76.5. This indicates average interviewing score is the determining factor of being hired by the company since applicants who have an average interviewing scores lower than 76.5 were eliminated at Phase 3. Only applicants with average interviewing scores ranked top ten could be successfully employed by the company.

Only 7 women entered Phase 3 and 15 male applicants entered Phase 3. In Figure 1.7, the high interviewing scores that were higher than 75 were dominated by male applicants. Since the interview scores were given by the interviewers, there may be some potential gender bias. However, the graph indicates the interviewing score of both male and female applicants are uniformly distributed. Both male and female applicants have the probability of receiving low interview scores.

Summary of the results

The models and figures aim to examine the factors that will influence the probability of being hired by Black Saber Software Company. We also want to find out whether there is any potential gender bias during the recruitment process. In order to enter Phase 2, the applicants must remain competitive in GPA and their working experience. For applicants who satisfy the basic requirements in their GPA and working experience, they will participate in a systematic assessment on their skills and leadership presence. The scores on writing, technical skills, speaking skills and leadership presence are especially important and will strongly affect the entry of Phase 3. Female applicants could have further improvement of their leadership presence and male applicants could try to improve their speaking skills. During the final hiring stage, the average interviewing score from two interviewers directly determines whether the applicants will be hired. Therefore, for applicants who desire to enter the company, they need to possess a strong academic achievement in their GPA and relevant working experience. Additionally, competitive skills in technology, speaking and writing, strong leadership and high interviewing scores are the determining factors to be successfully hired by the company. Though females have slightly higher rejection rates at Phase 1 and Phase 2, there is no obvious gender bias in each recruitment phase since gender does not influence the entry of each phase. Both male and female have equal chances to be hired by the company during the recruitment process. Therefore, GPA, working experience, skills (include speaking skills, writing skills, technical skills), leadership and interviewing scores are the determining factors of entering the company, but gender is not the determining factor of being hired.

The potential gender bias towards females on the promotion

Methods

In our research question 2, our purpose is to consider the factors that affect the number of promotions. The number of promotions for the current employees in Black Saber Software is counted to check whether the promotion is realized for all workers in the Software after the baseline of gender is introduced.

In order to analyze the factors that influence the promotions, we need to have a new dataset that contains the promotion-related variables. As we only have the original dataset from the information of the current employees for Black Saber Software, we should tidy up the original dataset by changing the data types of variables (gender, team, leadership for level and etc.) and removing the specific symbols or units for the variable names. In order to check the status of promotion for each employee, we group by the employee ids and count the numbers of role seniority for each id. Then, count each employee's missing values of roles and calculate each employee's number of promotions by using the number of all roles to subtract the number of missing values in roles. The subset of the number of promotions in each employee id is created and rejoined with the original dataset into a new dataset. Moreover, we mutate another new variable called average productivity by calculating the mean of current employees' productivity, summarizing the average productivity by employee id, and rejoining the variable into the new dataset. After tidying up the data in the new dataset by changing variables' data types, this is the new dataset we prepare to fit in the model that contains the variables of employee ids, number of promotions, average productivity, gender, team and leadership for level.

According to research problem 2, we want to explore whether gender affects the number of promotions of current employees in Black Saber Software. So, we introduce the null hypothesis that gender does not have an effect on the promotion of current employees. As the gender has a group of males, females and people who prefer not to say, we can set our baseline is male. Then, there are two null hypotheses for females and people who do not tell their genders. The first null hypothesis is that the gender of being a female does not have an effect on the promotion of current employees, while the second hypothesis is that the hidden gender of current employees does not affect on promotion.

For our model, we choose the number of promotions as the response variable, and gender(0 = male, 1 = female), team and average productivity as the explanatory variable with a random effect of employee id in the dataset for the information of the current employees in Black Saber Software. The generalized linear mixed model is chosen because we have a random effect (employee id) and it can measure the response variable (number of promotion) that repeats for more than one time on the same person. As the generalized linear mixed model is flexible to determine

different types of response, this model is quite similar to the generalized linear models and the linear mixed models. For the reason that our response variable is discrete within a range and the random variability in the outcome due to the random effect, we will fit a generalized linear mixed model with the family of poisson for the effects of gender, team and average productivity for each employee id by using the generalized linear mixed function.

For the generalized linear mixed model with poisson family, we assume our response as a poisson response that is a number of promotions per employee. The assumption of independence should also be satisfied in the model as each variable in our model is independent of one another. As the mean of random effect (employee id) is equal to its variance, we satisfy the assumption of the equality between the mean and variance. For the general equation of generalized linear mixed model with poisson family, we have

$$\log(\lambda_i) = \beta_0 + \beta_1 X_i \text{ where } Y_i \sim \text{Poisson}(\lambda_i) \text{ with } \lambda = \lambda_i$$

$$E(Y_i) = e^{\beta_0 + \beta_1 X_i}$$

Y_i represents the number of promotions for the i^{th} employee id, β_0 has an intercept of the number of promotions, $X_i\beta$ has fixed effects for gender, team and average productivity.

We fit an overall model with all predictors of gender, team and average productivity and a random effect of employee id as our first model. As we want to study the gender parity for promotion in the Black Saber Software, we prepare to find the best model by decreasing one of the predictors in the generalized linear mixed model with poisson regression, and comparing each model, respectively. For the second model, we remove a predictor of average productivity from the overall model. Then, we use the only predictor of the team and fit the third model. At last, we compare all the models by using the two-by-two likelihood ratio tests.

Results

Final model interpretation and table

For the comparison of models, we use the likelihood ratio test since this test is used to compare the goodness of fit for two statistical models. At first, we compare the overall model and the second model that lacks a predictor of average productivity. We can see that the p-value is 0.58 shown in the likelihood ratio test, so the second model with only two predictors of gender and team fits significantly better than the overall model with three predictors as the p-value is less than 0.05. Then, as the second model is a better model, we compare it with the third model by using the likelihood ratio test again. After calculating the p-value in the test, we find that p-value is 0.0031, which is less than 0.05. So, we conclude that the complex model with two predictors (gender and team) shows a better fitness than the third model with only

one predictors. Therefore, by comparison of all the models, the model with gender and team predictors is the best fitted model among all the fitted models and we can choose this second model as our final model.

Table 2.1. The table shows the estimated intercepts and p-values for the model with two predictors gender and team.

	Estimate Coefficient	P-value
Baseline	0.62	0.00
Gender Prefer not to say	-0.09	0.71
Gender Women	-0.21	0.00
Team Data	0.13	0.23
Team Design	0.30	0.08
Team Legal and financial	0.11	0.43
Team Marketing and sales	0.08	0.40
Team Operations	0.18	0.08
Team People and talent	0.40	0.00
Team Software	0.14	0.14

We summarize the best fitted model with the predictors of gender and team in the table 2.1. From the table, the intercept for baseline shows the log of risks for the number of promotion for a male employee from the client services, which is 0.62. By taking exponential of the value for baseline intercept, we have $\exp(0.62) = 1.87$, which indicates that the risks of the number of promotion for a male employee from the client services increase by $(1.87-1) = 87\%$.

The slope intercept for the female is -0.21, the slope for “prefer not to say” is -0.09 and other slope intercepts are both positive. For the female employee, the difference in the logs of the expected number of promotion is expected to decrease by 0.21, given the other predictors are constant. Similarly, for the people do not tell the gender, the difference in the logs of the expected number of promotion is expected to decrease by 0.09, with other constant predictors. This means that as a group of females and someone does not say their genders, gender is a factor that affects their chances of promotion. For other slope intercepts of different teams, the difference of staying in different teams for the current employee do not affect the promotion except for the team of People and talent as all p-values are greater than 0.05. For employee in the team of People and talent, we have strong evidence that they have more chance of getting promoted as p-value is

less than 0.05.

To be more specific for the gender parity of employees for the Black Saber Software, the risks of getting promoted for female employees is calculated by taking exponential of $\exp(-0.21) = 0.81$, which means that for being a female employee, the risks of the number of promotion decrease by $(1 - 0.81) = 19\%$ than that of male employee. By checking whether the result is significant, we should interpret the p-value of females' slope intercept. Since p-value is about 0 and is less than 0.05, we consider the result is significant and we have strong evidence against the first null hypothesis that the gender of being a female does not have an effect on the promotion of current employees. This means that the factor of being a female employee affects the promotion in the Black Saber Software. In addition, the people prefer not to say their genders has the risks of getting promoted in a value of $\exp(-0.09) = 0.92$, which means that the employees with unknown information of gender has less promoting chance by $(1 - 0.92) = 8\%$ than male employees. However, the p-value of people who do not want to say their genders is 0.71, which is larger than 0.05. This shows that the result for people who do not tell their genders is not significant and we have no evidence against the second null hypothesis that the hidden gender of current employees does not affect on promotion. By concluding this model, we find that compared with the baseline of males, females face the gender biases on promotion in the Black Saber Software, which is not a fair competition for promotion.

The model illustrates the potential bias on females towards the promotion, so we try to visualize the data.

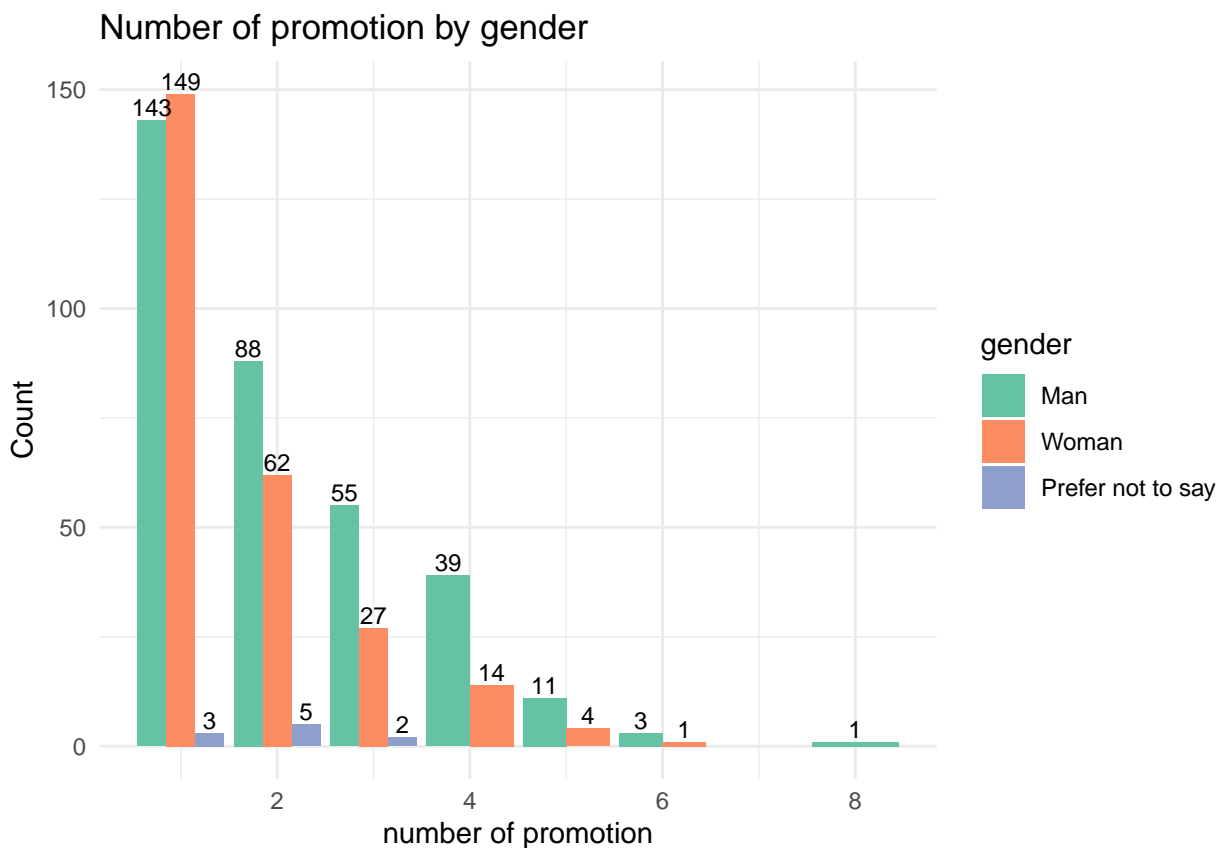
Interpretations of figures

Figure 2.1. The histogram shows the number of promotion by gender.

Potential bias can be created during the promotion process. In Figure 2.1, the histogram shows the number of promotions by gender. On the graph, the number of males and females promoted for one time is roughly equal. However, the gender ratio changes in the population promoted more than one time. The number of men having job promotion for five times is twice that of women. In other words, the number of males promoted for more than one time is much greater than females. Therefore, the graph shows that males are more likely to get promoted than females. So, gender will be a bias during the promotion process.

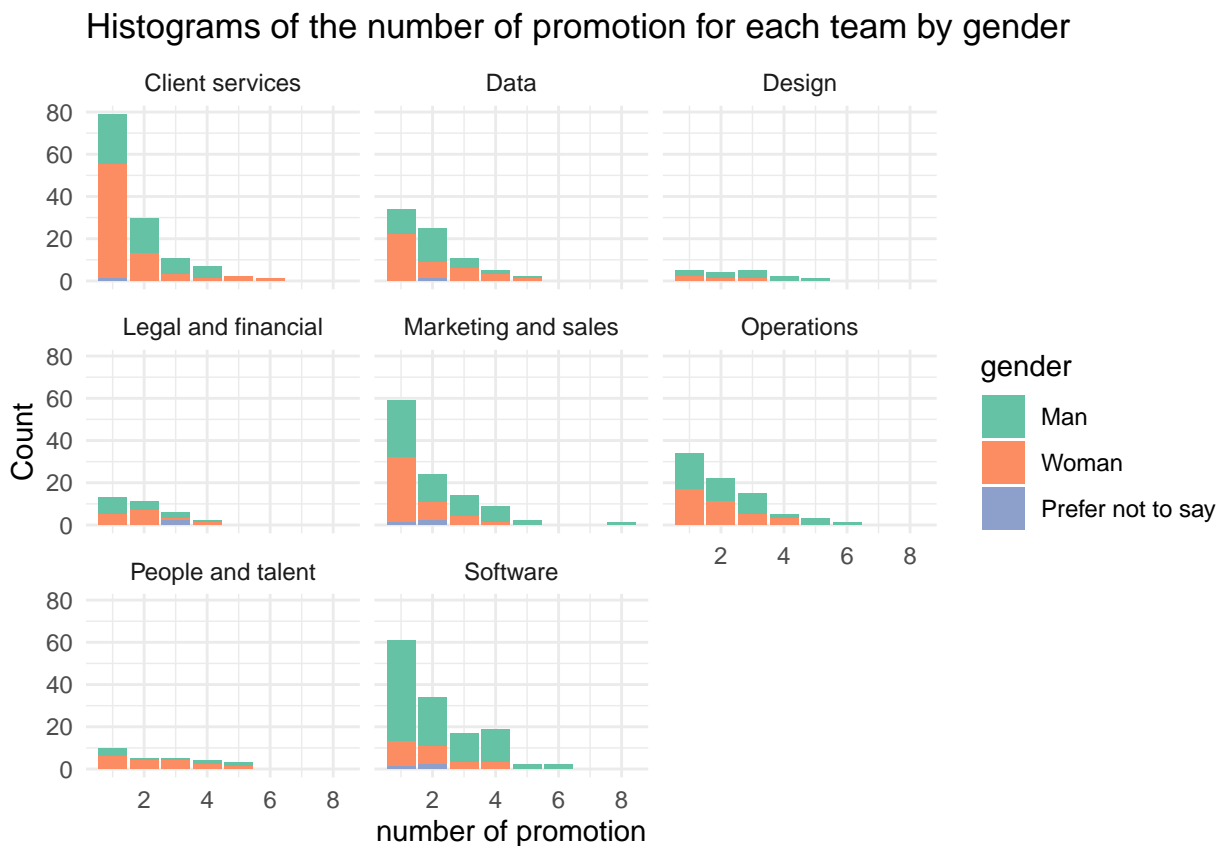


Figure 2.2. The histograms shows the number of promotion for each team by gender.

In Figure 2.2, the histograms show the number of promotions for each team. More females work in the Client Services Team than males. When comparing the promotions for both genders in that team, we can find that most females get promoted once and only a few females get promoted more than one. While, more males are promoted more than once compared with those promoted once. That is to say, it is easy for female employees to get one promotion. However, they are less likely to get promoted for more times than male employees. Similar results can get from other teams. The Software team has more employees promoted more than two times than other teams. While only a few employees working there are female. So, this figure also shows gender bias during the promotion process in this company.

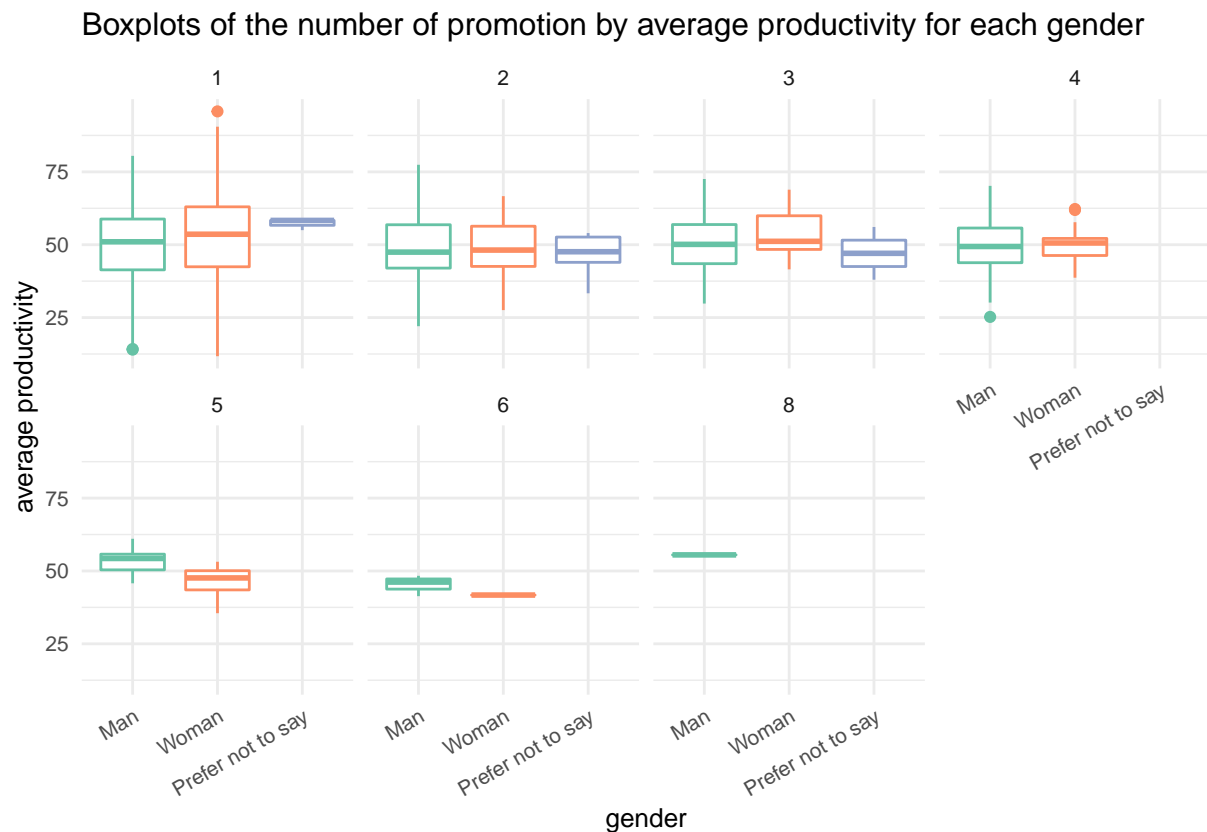


Figure 2.3. The boxplots of average productivity by gender for each number of promotion.

Figure 2.3 shows the boxplots of the number of promotions by average productivity for each gender. Some outliers are found on the boxplot. A few female employees promoted once have higher productivity than rest of the employees. While, a few male employees who have low productivity get promoted four times. The employees with more contributions to the company should have more chances of getting promoted. However, in this company, a few male employees with low productivity still get promoted many times. The figure shows gender bias during the promotion process in this company.

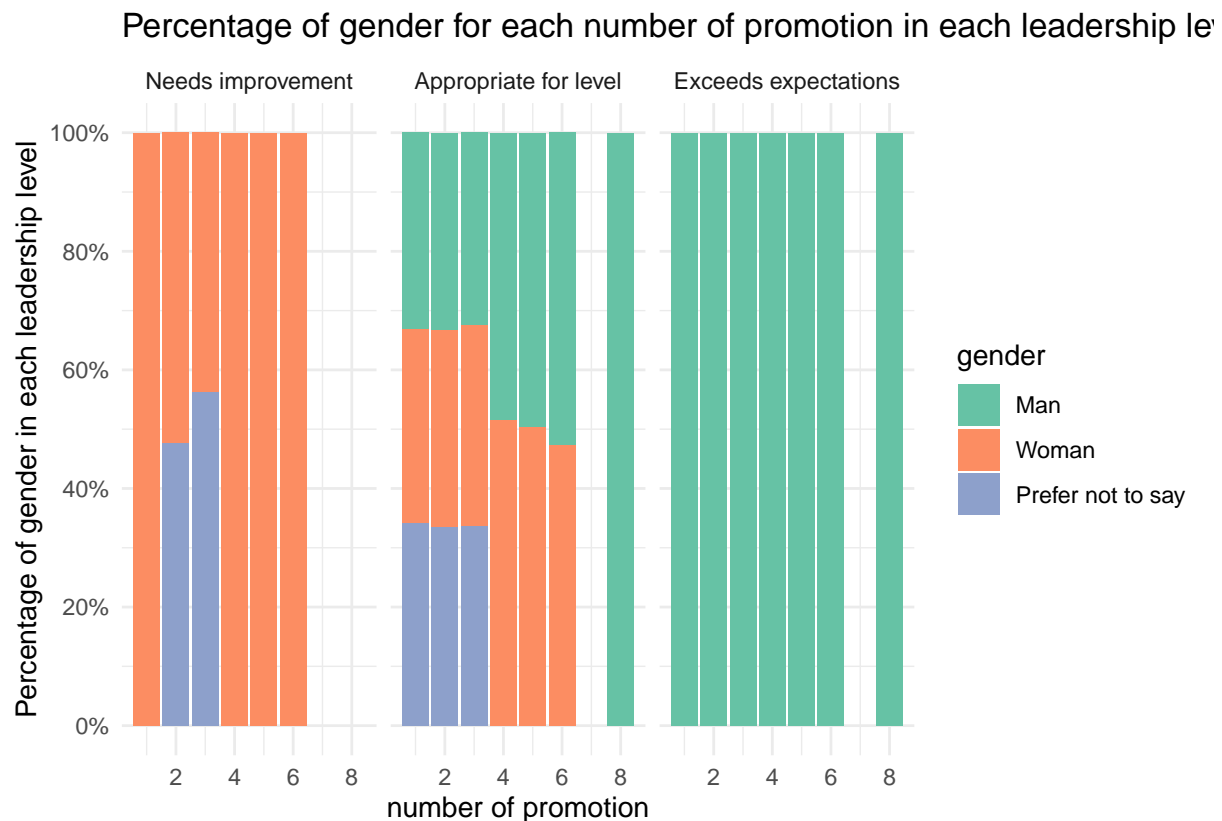


Figure 2.4. The bar chart shows the percentage of gender for each number of promotions in each group of leadership level.

In Figure 2.4, the bar chart shows the percentage of a number of promotions by gender at each leadership level. On the graph, the proportions of males and females received “appropriate for level” are roughly equal. However, no females receive “exceeds expectation” for their leadership level and no males receive “need improvement” for their leadership level. In other words, male employees are more likely to receive positive comments in this company.

Through the data analysis of the final model and plots, we can conclude that the female employee faces a serious gender bias on promotion, which means that the gender parity of current employees for the Black Saber Software is not effective.

The difference in employees' salaries between males and females

Methods

For the research question 3, we want to know about the factors that affect the employee's salary. From the data we get for the Black Saber Software Company, we can find salary term and we can set up the model that contains the variable we think might influence the salary. From those models we need to compare and choose which one is the most appropriate and also can show whether gender is the factor affects the salary.

In order to analyze the factors that influence the salary, we need to have a new dataset that contains the salary-related variables. As we only have the original dataset from the information of the current employees for the Black Saber Software. In the original dataset, we have Man, Woman and Not to say. Since we want to know that whether the gender difference influence the salary, we can filter term of Man and Woman that can directly show its gender. And then we should tidy up the original dataset by changing the data types of variables and removing the specific symbols or units for the variable.

We use the null hypothesis that gender does not have an effect on the salary of current employees. As the gender male and female, we can set our baseline is male and the null hypothesis is that the gender of being a female does not have an effect on the salary of current employees.

As we want to know whether gender affects the salary of current employees in the Black Saber Software. We introduce the linear mixed-effects models to help us set up and compare. The term "mixed" or, more fully, "mixed effects", denotes a model that incorporates both fixed- and random-effects terms in a linear predictor expression from which the conditional mean of the response can be evaluated. Also, we know that maximum likelihood or restricted maximum likelihood estimates of the parameters in linear mixed-effects models can be determined.

For the general equation of linear mixed model, we have

$$Y_i = X_i\beta + Z_ib + \epsilon_i \text{ where } b \sim \mathcal{N}(0, \psi_\theta), \epsilon_i \sim \mathcal{N}(0, \Lambda_\theta)$$

Y_i represents the salary for the i^{th} employee id, $X_i\beta$ has an intercept and fixed effects, Z_ib represents employee id i's baseline of salary, ϵ_i is the potential unmeasured confounders.

For the effect variables. In broad terms, fixed effects are variables that we expect will have an effect on the dependent or response variable. They are what we call explanatory variables in a standard linear regression. In this case, we are interested in making conclusions about how gender impacts the employee's salary. So gender is a fixed effect and salary is the dependent variable. On the other hand, random effects are usually grouping factors for which we are trying to control. They are always categorical variables. We will fit the random effect using the syntax

(1|variable's name). Although lots of the time that we are not specifically interested in their impact on the response variable, we know that they might be influencing the patterns we see. Note that adding or changing the random effects will not change our points estimates for our fixed effects.

As we get some models with some different predictors like gender, team, level, productivity and random effect. We should find out which one is the most appropriate one for our purpose. The likelihood ratio test is helpful that it can tell us the importance of additional variables. The likelihood ratio test performs a likelihood-ratio test of the null hypothesis that the parameter vector of a statistical model satisfies some smooth constraint. By comparing the two of those models, we can get one better and then use that model to do the next comparison. Eventually, we can get the best model from all of them.

Results

Model interpretation

Firstly, for the basic model of the relationship between gender and salary, we get the p-value of 0.000447 which is smaller than 0.05, it is significant that we can reject the null hypothesis which means being a female does have an effect on the salary of current employees.

For the comparison of models, we use the likelihood ratio test since this test is used to compare the wellness of the fit for two statistical models. At first, we compare the model1 and the model2 that lacks the predictor productivity. From the p-value, it is 0.00104 shown in the likelihood ratio test, as it is smaller than 0.05, it shows that the full model which is model2 is better. So the second model is better and we should add the productivity predictor. Based on our model2, we add the predictor role as the model3. Comparing these two models, we get the p-value by the likelihood ratio test is $2.2e^{-16}$, which is smaller than 0.05 so model3 is better, we should add the role seniority as the consideration. Since that adding or changing the random effects will not change our points estimates for our fixed effects. Based on model3, we change the random effects as the model4. We compare it with the model3 by using the likelihood ratio test again. After calculating the p-value in the test, we find that p-value is 0.99, which is much greater than 0.05. So, we conclude that the simple model model3 shows a better fitness. For the model5, we keep the only one predictor gender and compare the model5 and model3, we get the p-value is $2.2e^{-16}$, since smaller than 0.05, the complex model which is model3 is better. Therefore, by comparison so far, the model3 is the best. It is worth noting that our purpose is to find whether gender affect the employee's salary. So for the model6, we just delete the predictor gender based on the so far best model3. By comparing the model6 and model3, we get the p-value is $2.2e^{-16}$, since the number is very small, so model3 is better that we cannot delete the predictor gender, it shows that woman get lower income when other conditions are the same.

Overall, the model illustrates the potential bias on females towards the salary, so we do the visualization of the data.

Figures

Figure 3.1. The points plot shows the salary by gender.

Potential bias can be created for the salary by gender. In Figure 3.1, the points plot shows the salary by gender. It shows the number of salary males and females get is roughly not equal. The red points are roughly located higher of men having salary that higher than women. Therefore, the graph shows that males are more likely to get more salary than females. So, gender will be a bias for the salary.



Figure 3.2 The boxplots of salary by gender.

Figure 3.2 shows the boxplots of the salary for each gender. It shows the median salary of man is higher than the median salary of woman in the annual quarter. Also, there are some outliers are found on the boxplot that few employees get very high salary, even in this wage level, man get more than woman. Overall, this figure shows gender bias for the salary in this company.

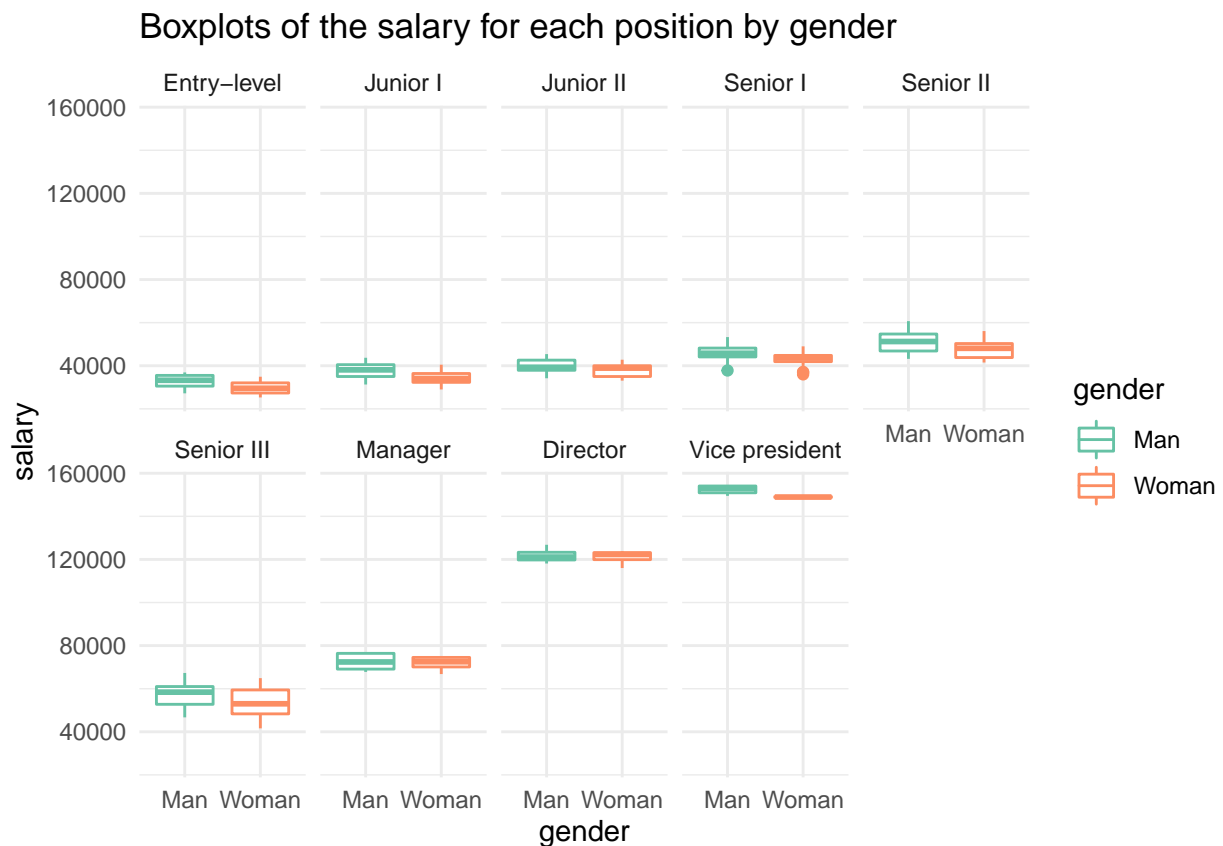


Figure 3.3 The boxplots of salary in each position by gender.

Figure 3.3 shows the boxplots of the salary in each position for each gender. It is obvious that as the position goes high, employee's salary also increase gradually. While it shows that from each position, the median of man is higher than the median of woman except the Manager and Director seems like they are on similar height. Overall, this figure shows gender bias for the salary in this company.

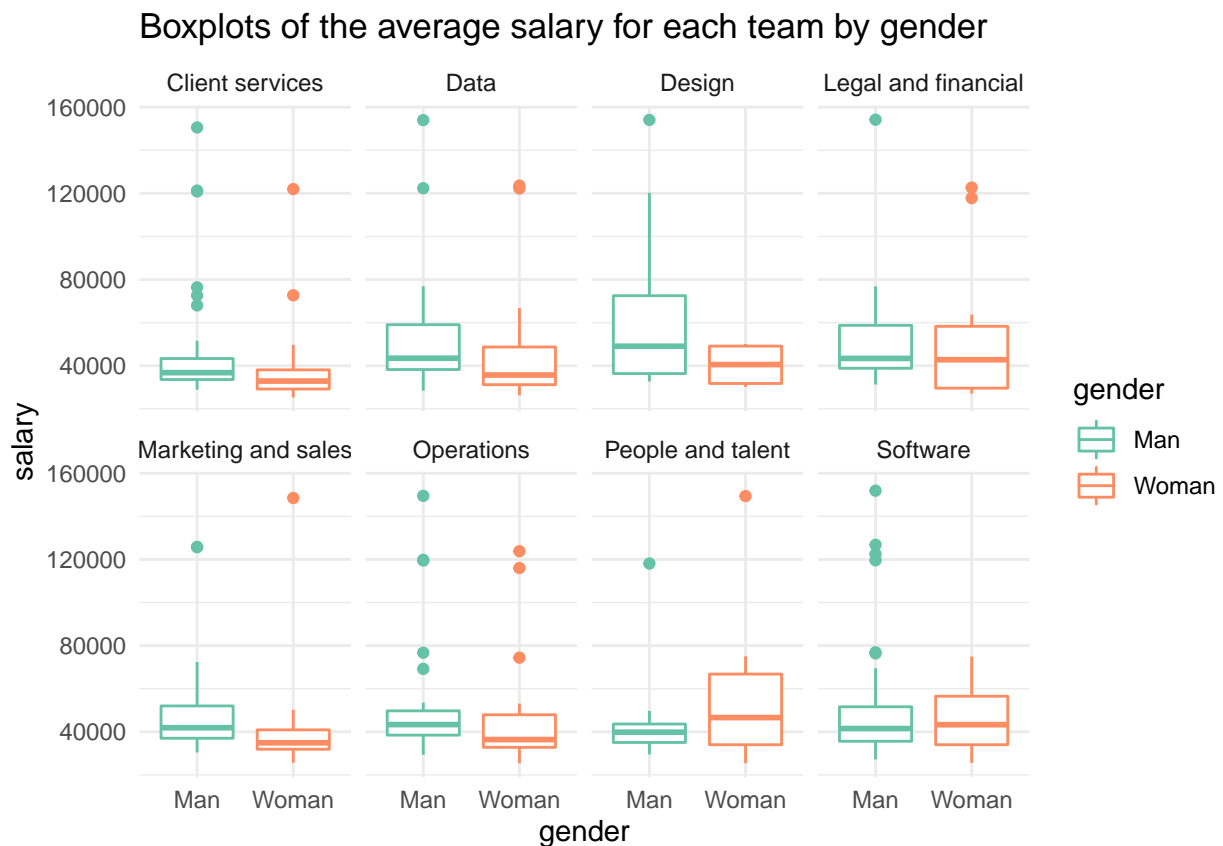


Figure 3.4 The boxplots of salary in each team by gender.

Figure 3.4 shows the boxplots of the salary in each team for each gender. It shows that from each team, the median of man is higher than the median of woman except the People and talent, Software these two teams seems like women's median salary is a little bit higher than men. And also in Legal and financial, the median salary of employee seems like they are on similar height. But in general, most of the team shows apparently that men get higher salary than women. Overall, this figure shows gender bias for the salary in this company.

Tables

Table 3.1. The table shows the average salary in each team by gender.

	Salary for Male	Salary for Female
Team Client services	43956.36	35414.49
Team Data	51954.29	43482.05
Team Design	64083.33	40350.00
Team Legal and financial	54871.43	53707.69
Team Marketing and sales	46490.16	38783.72
Team Operations	50166.67	43702.86
Team People and talent	46800.00	53094.12
Team Software	46796.12	44829.63

We summarize the average of each team's salary by gender in the table 3.1. From the table, by comparing the man and woman's salary in the same team, except People and talent team shows woman get higher average salary, other seven teams all shows male get higher salary.

Table 3.2. The table shows the average salary in each position by gender.

	Male	Female
Director	121891.67	121171.43
Entry-level	32948.61	29832.56
Junior I	37834.78	34394.34
Junior II	39837.74	38135.29
Manager	72436.84	71762.50
Senior I	46041.46	43075.86
Senior II	51117.86	48071.43
Senior III	57421.88	53900.00
Vice president	152383.33	148950.00

We summarize the average of each position's salary by gender in the table 3.2. From the table, by comparing the man and woman's salary in the same position, all the role seniority shows male get higher salary.

Summary

We firstly use some descriptive statistical methods to investigate the distributional characteristics of the related gender wage data. The dataset comes from the database of Black Saber Software Company. Then, we employ approaches of modeling and decomposition respectively to do analysis. The research aims on finding the key influential factors of gender wage and the most important purpose is whether the gender affect the employee's salary. The models, figures and tables all shows the gender bias on salary by Black Saber Company. By analyzing the data and models, the result illustrates there is a distinct difference in salary between male and female employees. Thus, the issue of gender bias still exists in distributing the salary (in annual quarter) towards females and the company should take actions to protect the rights of female employees.

Discussion

During the hiring process, applicants with higher GPA, abundant working experience, sufficient skill and relatively strong leadership are more likely to enter the final phase. In the final recruitment process, achieving a high interviewing score will be the determining factor of being hired. Gender equality is guaranteed in the hiring process since both male and female have relatively equal probability of entering each hiring phase. Though female applicants tend to have higher rejection rates at phase 1 and phase 2, it balances with the higher rejection rate on male at phase 3. Thus, gender is not the determining factor of entering Black Saber Software Company which indicates women's rights and interests are protected in the hiring process. The female applicants are less competitive in leadership than male and male applicants could have some improvement on their speaking skills. Though only 2 out of 322 female applicants were hired and 8 out of 291 male applicants were hired in the end, this is not representative since it may be caused by randomness and the limited number of job positions. Overall, applicants who want to enter the company should not worry about the potential gender bias in the hiring process. Instead, they should focus on enhancing their working experience, boosting up their GPA and relevant skills, increasing leadership presence and performing better during the interviews. These are the main factors that determine whether applicants will be successfully hired by Black Saber Software Company.

The potential gender bias that exists in the company significantly influences the result of getting promoted. In the Black Software Saber Company, the opportunities for promotion is not equal for both genders. It is easy for female employees to obtain one promotion. Nevertheless, female employees are less likely to get promoted for mutiple times. Though female employees do the same job as males, it is difficult for women to get the same opportunities of promoting. In a fair workplace, the employees with more contributions to the company should have more chances of getting promoted, but a few female employees with extremely high productivity only got promoted once. Conversely, male employees who have low productivity got promoted for more times. The abilities of female employees tend to be underestimated by the company, and thus gender equality is not achieved in this company. Moreover, compared with male employees, female employees are less likely to receive positive comments in this company. Overall, potential gender bias exist in promotion process since female employees may receive discrimination when working in this company. Therefore, the Black Saber Software Company should provide an environment with equal opportunities of promotions for both genders based on their contribution to the company.

There is a considerable difference in salary that male employees are more likely to receive higher salaries than female employees. Women have lower median salary than men though they work in the same position. Specifically, male employee who holds a similar position as female employees

earns a bigger paycheck. Additionally, female employees still earn less even when they work for different teams. This wage gap between male and female employees shows the existence of gender inequality in the Black Saber Software Company. The salary distribution discriminates unfairly against the female employees when they work for this company. Hence, the Black Saber Software Company should balance the average salary between females and males in order to minimize the wage gap.

Limitations

During the hiring process for the Black Saber Software, phase 3 may exist as potential biases since the ratings for phase 3 are graded by interviewers. Since the hiring process involves judges, we are not sure whether these judges are fair towards everyone. However, since phase 2 is automatically marked by the AI system, it is relatively fair and female applicants are less likely to face bias. Since we are only provided by a single set of data and it has a limited number of job positions, there exists randomness in analyzing the data and the final result may not reflect the actual outcome. We could not simply conclude there is gender bias by only examining this set of data. Therefore, if we are provided by more datasets of the hiring process, the result may be more accurate and we will find out whether gender bias is an issue in the company.

For the leadership level of current employees, there are three levels including “Needs improvement”, “Appropriate level” and “Exceeds expectation”. However, the data of the leadership for level counts in the annual quarters, which means the employees in the Black Saber Software are rated in their leadership in every annual quarters. Thus, the provided variable for the leadership is repeated for each employee id. Since we do not know who rates the quality of leadership for each employee, we cannot be sure whether the variable of leadership has potential bias that is caused by the judges’ personal perspectives. To improve this collected data and reduce the biases, we can divide the leadership variable for three levels and consider each level as a variable.

The initial proportion of current male employees and female employees is not the same. If we want to compare the opportunities of promotion for both genders, it is better to choose a dataset with equal number of men and women. We are supposed to find that more male employees get promoted since there are more male employees in the dataset. Thus, this may cause inaccurate results towards our findings. For the future considerations, we should compare similar number of male employees and female employees for the next time.

There are also some limitations and concerns on the model we use. Since we apply the likelihood ratio tests to compare models and get appropriate models in the data analysis on the promotion and salary, the tested models should be nested into each other in some ways. As we cannot use the likelihood ratio test to compare three models at the same time, it makes the model

comparison process more complex and we can only make the comparison with two models to find a better model.

Consultant information

Consultant profiles

Jinwen Tao. Jinwen is a Data Scientist with Queen Back Technology Company. She specializes in analyzing complex raw information and finding patterns. Jinwen earned her Bachelor of Science, Majoring in Statistics and Mathematics from the university of Toronto in 2022.

Zhifei Cao. Zhifei is a Machine Learning Scientist with Queen Back Technology Company. She specializes in researching new data approaches and algorithms. Zhifei earned her Bachelor of Science, Majoring in Statistics and Mathematics from the university of Toronto in 2022.

Yuanxi Gong. Yuanxi is a IT Consultant with Queen Back Technology Company. She specializes in implementing technological solutions. Yuanxi earned her Bachelor of Science, Majoring in Statistics and Economics from the university of Toronto in 2022.

Jingyi Zhang. Jingyi is a Business Intelligence Developer with Queen Back Technology Company. She specializes in business intelligence developers design. Jingyi earned her Bachelor of Science, Majoring in Statistics and Economics from the university of Toronto in 2022.

Code of ethical conduct

- As an employee, the responsibility is to carry out and record the work carefully, satisfy the diligence requirements of the client, and apply the statistical analysis scientifically, without forecasting outcome.
- Respect and acknowledge the contributions and intellectual property rights of others.
- In publications and reports, communicates findings in an honest and fair manner that makes sense to the readers. This includes figures, models and tables that convey the ideas more directly through visualizations.
- Avoid any actions that will adversely affect the statistic and the good reputation of the statistician.