Potential Gender Biases in the Workplace

Investigating the Hiring Process, Promotions, and Salaries in Black Saber Software

Report prepared for Black Saber Software by The Hive

April 21, 2021

Contents

Executive summary

Background & Aim

A critical area of concern in today's workplace is gender bias, and Black Saber Software's culture is no exception. It is critical that a company shows that their workplace practices are not only unbiased, but also that they embrace diversity. As a result, the purpose of this report is to investigate any potential gender bias in Black Saber Software's hiring, promotion, and salary processes.

We began by analysing the new graduate hiring process, in which 2 of the 3 stages are automatically rated by an artificial intelligence algorithm to determine who moves onto the next round. The third and final phase involves interviews by two people.

We see that as the company grew, the number of men in higher level roles (ie. above Entry-level) is noticeably greater than the number of women (Figure A (a)).

Moreover, we see that in their most recent financial quarters (2019-2020), men on average made close to $\$50,000$ while women (and those in the "Prefer not to say" group) made on average under $\$45,000$ (Figure A (b)). We investigate whether or not this is due to gender alone, or if other factors affect these initial observations.

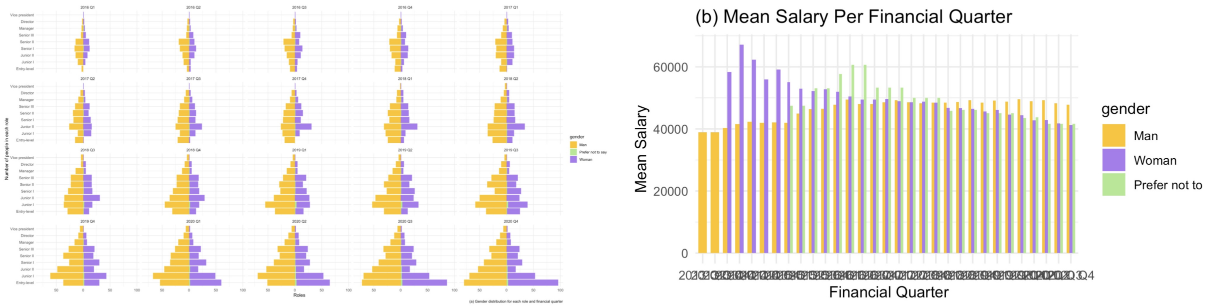

Figure A. (a) Gender Distribution for Each Role in Financial Quarter (b) Salary Per Financial Quarter

Key Findings

Hiring

- The AI algorithm did not prefer any specific gender. That is, it chose who should move forward in the recruitment process fairly.
- More women than men initially applied to Black Saber, but only 2 of the 10 final hires were women. Black Saber did not hire anyone who identified as "prefer not to say" in 2020.
- The interviewers favoured men slightly more than women.

Promotions

While accounting for productivity and role seniority, we found that:

- Women were about 40% less likely to be promoted than men.
- There is significant evidence that gender affects whether an employee is expected to be promoted

- There are no significant conclusions to be drawn about the group "Prefer not to say" most likely due to the small sample size of this group.

Salary

While accounting for the variables gender, team, productivity, leadership and role seniority, we found that:

- Women received a salary of approximately $2248.2 less than men.
- Those in the "Prefer not to say" group received a salary of approximately $1104.1 less than men.

- There is significant evidence that gender affects an employee's salary

Limitations

- We did not have enough data to determine whether the biases we found were necessarily due to a biased culture at Black Saber, or if it is due to the systemic barriers women (and those who prefer not to specify) face in the professional world.
- Because only 10 people were hired, it is difficult to determine whether or
- The number of applicants and employees in the "Prefer not to say" group was too small to make meaningful conclusions.
- The data does not account for those who may reject promotions for personal reasons which could inflate the amount of non-promotions we see. This in turn indirectly impacts salary expectations as well.

Technical report

Introduction

Background

A critical area of concern in today's workplace is gender bias, and Black Saber Software's culture is no exception. It is critical that a company shows that their workplace practices are not only unbiased, but also that they embrace diversity. As a result, we have been hired as an external, third-party consultancy to review Black Saber Software's hiring and promotion processes, as well as their employee salaries in order to determine whether or not the company is biased in their practices.

Research questions

- Does the hiring algorithm favour a certain gender? Do the humans conducting the interviews have bias towards a certain gender?
- Does Black Saber as a company favour a certain gender when promoting employees?
- Does Black Saber pay a certain gender more?

Methods

Hiring Process

Black Saber's current new graduate hiring process proceeds in 3 stages, the first two of which are assessed by an artificial intelligence algorithm. It is not until the third and final round that a human becomes involved in the recruitment process. At the beginning of the process, each applicant is assigned a unique ID number that follows them throughout the process to help anonymize the data, as well as keep track of their progress. The applicants specify their gender (male, female, prefer not to say), and the team they wish to apply to (data or software). They then have the option of uploading a cover letter, resume, their GPA (scale from 0.0 to 4.0), extracurricular activities, and work experience. In phase 1, the algorithm rates each applicant's level of extracurriculars and work experience (0, 1, or 2; 2 being the best). These in conjunction with their GPA and the presence of a cover letter and resume are used to decide which applicant moves onto phase 2. Phase 2 consists of a technical task, writing sample, and re-recorded video.

The algorithm uses these materials to rate each applicant's technical skills (0-100), writing skills (0-100), speaking skills (1-10), and leadership presence (1-10). These scores determine who moves onto the final phase, the only one that has human involvement on the company's side. Phase

Table 1: Gender Count for Phase 1

Gender	Quantity
Man	291
Woman	311
Prefer not to say	11

3 is an interview with 2 interviewers, who each score the applicant on how fit they are for the job on a scale from 0 to 100. We will use this information to investigate whether or not there is gender bias in the rating system both in the algorithm, but also in the interviewers.

Phase 1 To wrangle the data, we added a column that specified whether or not an applicant moved onto the next round (denoted 0 and 1 for no and yes, respectively). This was done by first checking if there were any missing values in the dataset, which there were not. Then, we (fully) joined the phase 1 and phase 2 datasets, and noting which applicants had a value for "technical skills" in a new "next round" column. This is because technical skills were rated in phase 2; thus, if an applicant did not have a technical skill rating, then they did not make it to the next round. This could be done because it was confirmed prior that there were no other missing values in the dataset. We marked a 0 for applicants who did not have a technical skill rating and 1 for those who did. Then, we kept only the columns that were rated in phase 1, along with the new "next round" column.

We then compare the number of applicants that identify as either male, female, or preferred not to specify (Table 1). We see that the applicant pool seems to be fairly even between 311 women and 291 men; with a smaller portion of 11 applicants who preferred not to specify. The even distribution between women and men is a good basis to investigate whether or not there is gender bias in the AI algorithm that determines who moves onto the next round. We investigate this effect with generalized linear models and generalized linear mixed models, with the response variable being a binary response of whether or not the applicant moved onto the next round. Since the algorithm considers each factor (ie. existence of cover letter, resume, level of GPA, work experience, and extracurriculars), these will be the fixed effects in the base generalized linear model.

Next, we created models with a fixed effect for gender. This second model is the same as the base linear model, but with an additional fixed effect of gender. The third model is a generalized linear mixed model, and adds an additional random effect for the team that the applicant applied for. We want to see if this potential bias exists in one, or both of the teams.

Table 2: Log Likelihood Test for Random Effect of Team in Phase 1

# Df	Log Likelihod	Df	Chi-squared	P-value
8	-17.45	NA	NA	NA
9	-17.45	1	1e-07	0.9997

Table 3: Log Likelihood Test for Fixed Effect of Gender in Phase 1

# Df	Log Likelihod	Df	Chi-squared	P-value
6	-17.91	NA	NA	NA
8	-17.45	2	0.9144	0.6331

Since the second model is nested in the third, we can first compare these last two models with a log likelihood test to see if there is a significant difference between the model that includes the teams, and the one that does not (Table 2). Since the p-value value is large and close to 1 (0.9997), we conclude that there is not a statistically significant difference that the algorithm is biased towards a certain team. Thus, we can use the simpler model to compare to our base linear model that does not include gender as a fixed effect.

Since the base model is nested within the second linear model with a fixed effect for gender, the log likelihood test can be used again to compare them (Table 3). We find a large p-value again (0.6331) that shows that there is not a significant difference between the two models. This shows that the algorithm is not significantly biased towards gender in the first round, otherwise the random intercepts between these models would be different and the p-value value would be very small (<0.05).

Phase 2 We wrangled the data in phase 2 similarly to how we did in phase 1. We fully joined the phase 2 and phase 3 datasets, and added a column denoting which applicants had an interviewer rating, which is how we determined who moved onto the third round. We denoted "moved forward to the next round" a 1, and "did not move forward" as 0. Again, if the applicant did not have an interviewer rating, it meant that they did not move forward in the process since we confirmed in the beginning that there were no missing values prior to the wrangling process.

Table 4: Gender Count for Phase 2

Gender	Quantity
Man	145
Woman	152
Prefer not to say	3

Table 5: Log Likelihood Test for Random Effect of Team in Phase 2

# Df	Log Likelihod	Df	Chi-squared	P-value
7	-34.62	NA	NA	NA
8	-34.43	1	0.3808	0.5372

Table 6: Log Likelihood Test for Fixed Effect of Gender in Phase 2

# Df	Log Likelihod	Df	Chi-squared	P-value
5	-35.52	NA	NA	NA
7	-34.62	2	1.784	0.4099

We can see that in Table 4, the split between men and women is still fairly even, and there is a similar ratio of people in each gender category removed by the algorithm. This makes sense from our last statement that the algorithm as not biased in phase 1.
Then, we want to investigate whether or not there exists gender bias in this phase. Since the algorithm necessarily considers technical skills, writing skills, leadership presence, and speaking skills, these will be the factors in our base generalized linear model.
Next, we created a second and third model with a fixed effect for gender. The second model is the baseline model, with an additional fixed effect for gender. The third model is the same as the second model, but with a random effect for the team each applicant applied for. First, let's compare the second and third model to see if the algorithm is more biased for one team than the other (Table 5). We test this with a log-likelihood test, since the second model is nested within the third. We see that the p-value is 0.5372, which is insignificant at the 5% level, signifying that there is not a significant difference between the model with the random effect for team and the one without it. Therefore, we can move forward with our comparison using the similar model without the random effect.
Next, we compare this simpler model with a random effect for gender with its nested generalized linear model without this effect using a log-likelihood test (Table 6). We see that the p-value is 0.4099, meaning it is insignificant at the 5% level. Thus, we can see that there is not a statistically significant difference between the model accounting for gender and the one that doesn't. Thus, we can say that there is insufficient evidence to suggest that the algorithm is biased in Phase 2 of the hiring process.

Moreover, if we look at the 95% confidence interval for the coefficients in the model with a fixed effect for gender (Table 7), we can see that 0 lies within the interval for women, and that the estimate for those that prefer not to specify is negative, and the upper bound for the estimate is positive, so we know that 0 also lies within this interval. Therefore, we can say that there is insufficient evidence that suggests gender is associated with moving onto the interview round from phase 2.

Table 7: 95% Confidence Interval for GLM with Gender Effect

	Estimate	2.5%	97.5%
Baseline	-20.773	-29.329	-14.434
Technical Skills	0.081	0.046	0.127
Writing Skills	0.092	0.050	0.145
Leadership Presence	0.896	0.545	1.360
Speaking Skills	0.716	0.416	1.083
Woman	-0.567	-2.055	0.823
Prefer not to Say	-16.192	NA	243.426

Phase 3 and Final Hires In phase 3, we wrangled it similarly to Phases 1 and 2 and created a new binary column, but instead of a new column indicating if the applicant moved onto the next round, it indicated whether or not the person was hired. First, we had to create two intermediate datasets. Let's call them intermediate dataset 1 and 2 respectively. By right joining the phase 2 and phase 3 datasets respectively, we were able to see the variables critical to our analysis (such as gender and team applied for) from phase 2 on only the applicants that made it to phase 3. Then, we remove the columns we don't need to analyse phase 3: the scores the algorithm gave each applicant in phase 2. Thus, we get a dataset with each applicant's ID, gender, team, and both interviewer ratings - intermediate dataset 1.
Then, by fully joining intermediate dataset 1 with the IDs of the final hires, we were able to get a dataset with all the variables from intermediate dataset 1 but for only the applicants that got hired. Then, by adding a column of all 1's, indicating that these applicants were hired, we get intermediate dataset 2. By fully joining the two intermediate datasets, and changing the missing values in the hired column (for those that were not hired) to 0's, we get a dataset that can be used for analysis. For the purpose of this analysis specifically, however, we know that both the interview scores are taken into consideration when hiring an applicant. We noticed that the 10 applicants with the highest average score were hired, so we consolidated the two scores into one column for the average of the two scores to run our analysis.

We see that there was only 1 woman hired for each team (Table 8). From this we can say two things: (1) The number of women hired is fairly even, so we don't need to create and test a model with a random effect for team; especially since the sample size is so small, it will be hard to make good statistical inferences even with the model. (2) Even though more women applied and made it through the unbiased algorithm twice, only 2 out of the 20 final hired applicants were women. We should investigate this and see if the actual people conducting interviewers are significantly biased or not.

For our analysis, we used two generalized linear models - one with a fixed effect for gender, and one without. We use a log-likelihood test to test whether there is a difference in models when

Table 8: Gender Breakdown of Final Hires

	Man	Woman
Not Hired	7	5
Hired	8	2

Table 9: Log Likelihood Test for Fixed Effect of Gender in Final Hires

# Df	Log Likelihod	Df	Chi-squared	P-value
2	-1.386	NA	NA	NA
3	0.000	1	2.773	0.09589

we account for gender (Table 9). We see that the p-value is 0.096, which is significant at only the 10% level. Since we are considering significance only at a 5% level, we can say that this test did not show significant evidence for gender bias in the interviewers; however, this result is still important to note. The significance at a 10% level should not be ignored, especially since the phases in the hiring process done by the AI algorithm did not show any significant results. This may suggest that the people involved in the interview process may have a slight bias in gender, even though we technically did not find significant results at the 5% level.

Promotions

Our next task was to look into Black Saber's promotions. We were given data of all the current employees by financial quarter which included each employee's gender (man, woman or prefer not to say), the team they worked on (Client Services, Data, Design, Legal and Financial, Marketing and Sales, Operations, People and Talent and Software), role seniority (Entry-level, Junior I, Junior II, Senior I, Senior II, Senior III, Manager, Director, Vice president), leadership for their role (needs improvement, appropriate for level and exceeds expectations), productivity on a scale of 0 - 100 (with 100 being very productive) and lastly, the salary in the given financial quarter. To wrangle the data and prepare it for analysis, we first factored and numerated role seniority as this is the only indication we have of someone being promoted. Entry-level, Junior I, Junior II, Senior I, Senior II, Senior III, Manager, Director and Vice president was coded as 1,2,3,4,5,6,7,8,9 respectively. This also allowed us to visualize the gender distribution by role for each financial quarter as a starting point.
Disclaimer: Upon analysis we decided not to use data from years 2013 to 2015. Totaling the number of employees from each quarter resulted in 100 or less employees. With such a small pool of data per quarter, it would be difficult to apply any meaningful analysis.

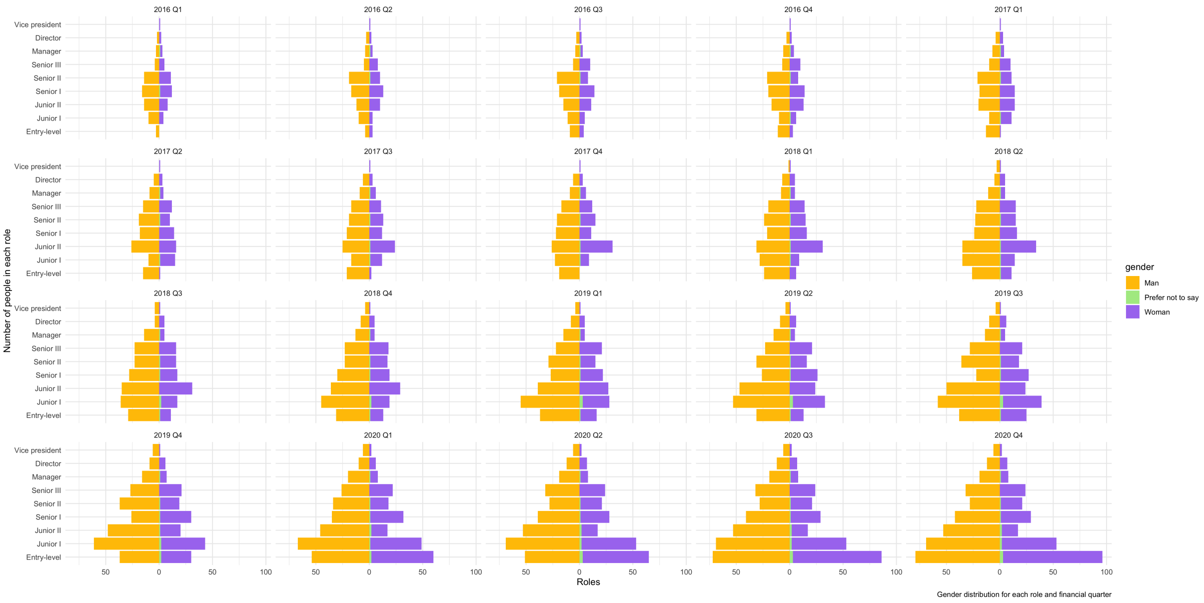

Figure 1: Gender Distribution for Each Role in Financial Quarter

Visually, we can see that as Black Saber expanded as a company and hired more people, there seems to be slightly more women in entry level roles than men. On the other hand, while we move into the more senior roles, there seems to be more men. This is something we want to consider and look out for in our analysis. Is this simply due to chance? Are men in these roles because they are on average more suited for them by looking into factors such as productivity and leadership? Or does Black Saber have a bias towards promoting men. Note: This plot was made referencing Robert Lanfear's blog post (https://www.robertlanfear.com/blog/files/visualising_gender_balance_R.html).

Quantifying and Modeling Promotions To be able to model promotions, we created a "promoted" column in the data set to indicate whether a person has been promoted from their previous role or not. An employee is considered to be promoted if their role seniority is greater than the previous quarter (eg. In 2015 Q1, role seniority = 2 < In 2015 Q2, role seniority = 3). This change was indicated as a 1 if promoted, and 0 otherwise.

Now, we can visualize the proportion of men, women, and those who have preferred not to disclose their gender. The image below shows us that of the promotion of those promoted, men have be promoted the most (Figure 2).

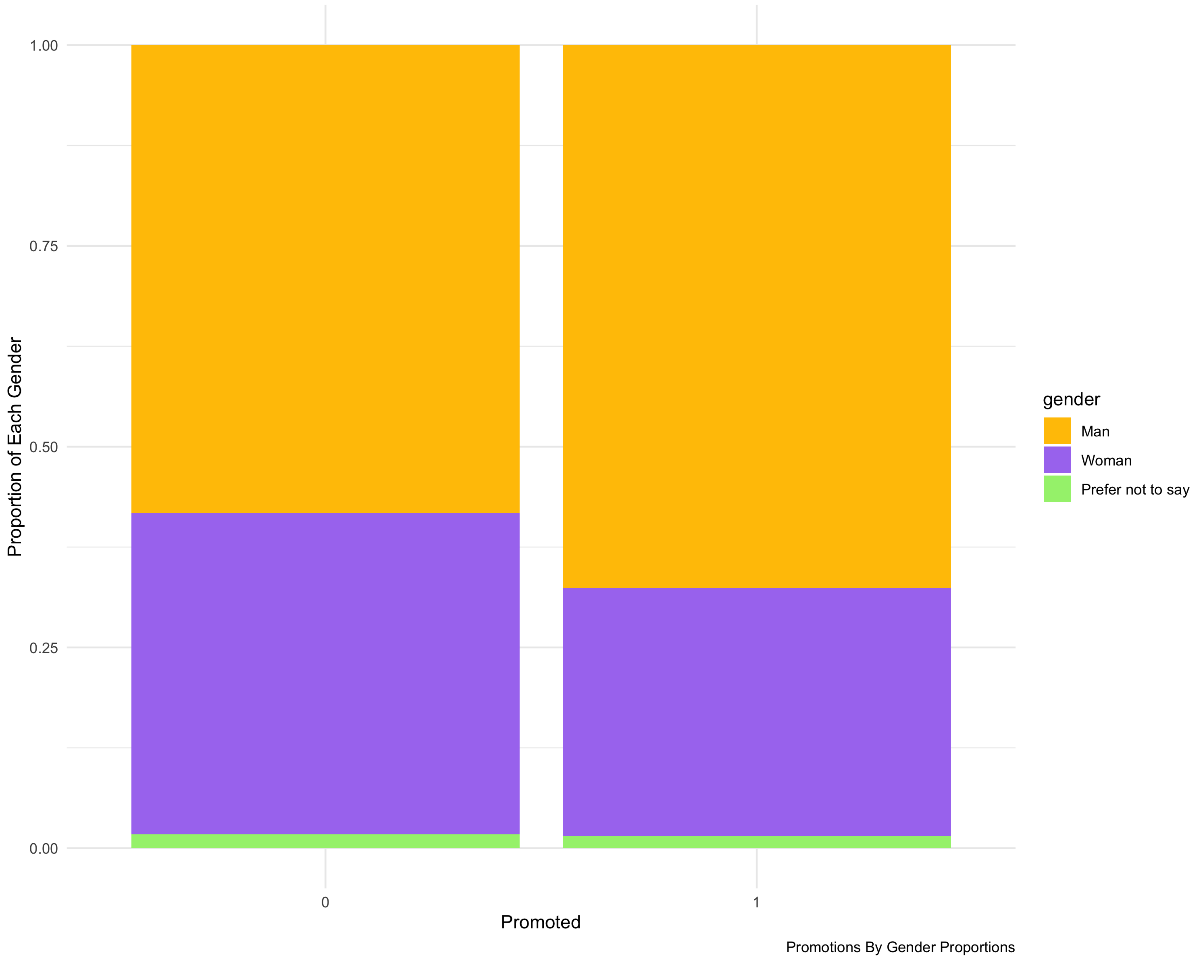

Figure 2: Promotions By Gender Proportions

Do these proportions change as we separate the different teams? As seen below, the only team where women seem to have more promotions is in People and Talent. While in the data team it seems to be pretty even. However, it is important to consider if these proportions are significant when considering all the other variables, such as leadership, productivity, as well (Figure 3).

Table 10: Model 1: Promotion and Gender Values

	Estimate	Std. Error	z value	Pr(>|z|)
Baseline	-2.299	0.056	-41.122	0.00
Woman	-0.406	0.099	-4.111	0.00
Prefer not to Say	-0.256	0.371	-0.690	0.49

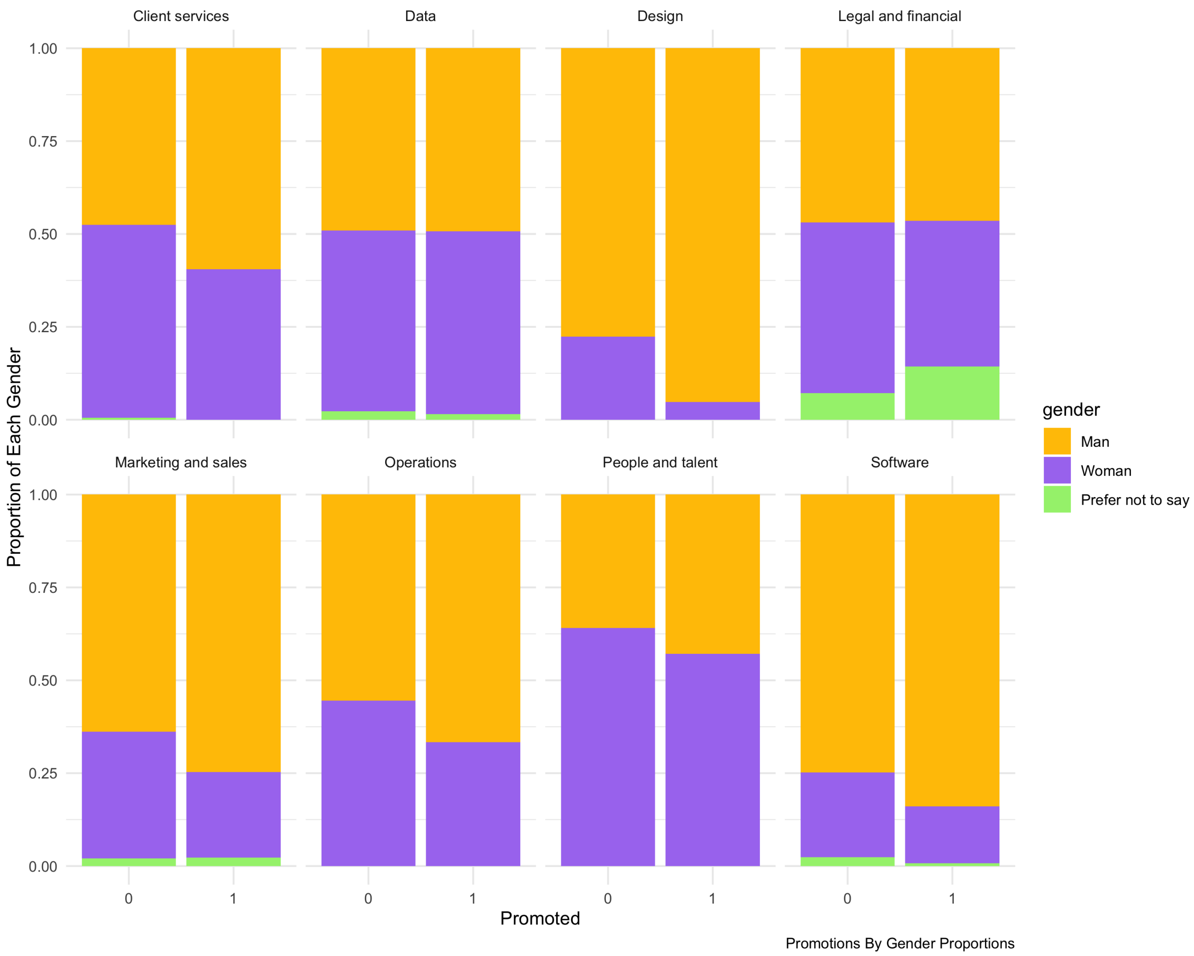

Figure 3: Promotions by Gender Proportion within Teams

Promotion Model For our model, we used Generalized Linear Models with the variable promoted as the binary response variable. First, we started off with a model that predicted being promoted with gender as the only variable. This gave us a significant coefficient at the 0.001 level of -0.41 for woman. This can be interpreted as the odds a woman is promoted is 0.66 times less likely than men. For the subgroup, prefer not to say, the odds are also lower, however this value is not significant.

Table 11: Model 2: Promotion with Gender and Team Values

	Estimate	Std. Error	z value	Pr(>|z|)
Baseline	-2.364	0.125	-18.943	0.000
Woman	-0.410	0.102	-4.013	0.000
Prefer not to say	-0.241	0.375	-0.644	0.520
Data	0.001	0.174	0.004	0.996
Design	0.180	0.259	0.693	0.488
Legal and Financial	0.077	0.231	0.333	0.739
Marketing and Sales	-0.049	0.162	-0.299	0.765
Operations	0.051	0.168	0.304	0.761
People and Talent	0.420	0.214	1.966	0.049
Software	0.132	0.151	0.877	0.380

Table 12: Model 3: Promotion with Gender and Leadership Values

	Estimate	Std. Error	z value	Pr(>|z|)
Baseline	-2.276	0.359	-6.349	0.000
Woman	-0.431	0.101	-4.260	0.000
Prefer not to say	-0.281	0.372	-0.757	0.449
Leadership Appropriate for Level	0.002	0.354	0.006	0.995
Leadership Exceeds Expectations	-0.471	0.447	-1.055	0.291

But is something else effecting the promotions? We decided to investigate further and figure out if any of the other variables could explain our results better. First, we looked at adding the team they are in. Is there bias in only some teams? Table 10 shows the results from modeling promotions with gender AND teams. Again, we see that for women, the odds of being promoted remains at around 0.66 times less likely than male counter parts.

Then, we looked at whether a person's leadership affected their odds of being promoted after all (Table 11); we would expect to see those with higher leadership skills be promoted. Even with leadership being added as a variable, we do not see a change in the gender promotions, nor does it seem significant to include in our model.

Does productivity explain this relationship we see?

Table 13: Model 4: Promotion with Gender and Productivity Values

	Estimate	Std. Error	z value	Pr(>|z|)
Baseline	-1.792	0.153	-11.743	0.000
Woman	-0.390	0.099	-3.944	0.000
Prefer not to say	-0.300	0.372	-0.808	0.419
Productivity	-0.011	0.003	-3.492	0.000

Table 14: Model 5: Promotion with Gender, Productivity and Role Seniority

	Estimate	Std. Error	z value	Pr(>|z|)
Baseline	-2.370	0.181	-13.081	0.000
Woman	-0.380	0.099	-3.829	0.000
Prefer not to say	-0.295	0.373	-0.791	0.429
Productivity	-0.010	0.003	-3.189	0.001
Role Seniority	0.137	0.022	6.352	0.000

Table 15: Log Likelihood Test for Gender in Promotions

# Df	Log Likelihod	Df	Chi-squared	P-value
5	-1783	NA	NA	NA
3	-1791	-2	15.39	0.0004549

We found that productivity, specifically, whether someone was more productive than expected, lowered their odds of being promoted, which was an interesting find. However, it still did not change the relationship with women and promotions by very much. That is we still find that women are less likely to be promoted.
Finally, we explored the relationship of role seniority with promotions since on the surface, role seniority has an ambiguous relationship with being promoted (Table 13). On one side, role seniority could be an indication of productivity and leadership levels and affect promotions positively. But on the other hand, the expectations and skills required to be promoted in the higher roles, may make it more difficult to be promoted (there are also less spots, eg. entry level roles vs vice-president).

Our analysis tells us that role seniority is an important variable in our model. Specifically, as role seniority increases, we expected the odds of being promoted to increase as well (by about 1.14 times more likely). However, although it does reduce the effect we see on women being promoted, it does not change it by much. The odds of women being promoted is still around 30% less than men.
Using Model 5 as our final model, we see that this relationship between gender, women specifically, is significant to the 1%. Even when accounting for significant variables such as productivity and role seniority, we found that the odds of women at Black Saber being promoted was about exp(-0.38) = 0.68 or about 40% less likely to be promoted.\ To further solidify these findings, we compared Model 5 without the variable gender and found that Model 5 (with gender) described the data better. Furthermore this was significant to the 1% level, giving us evidence that gender does affect whether an employee gets promoted (Table 15).

Salary

Our last area of interest was the salary. Again, we utilized the current employee data set. First we visualized the mean salary for each gender in each quarter. In 2013 we see women on average earning more than men at Black Saber. However, its difficult to obtain any conclusions from this as the sample is very small (about 4 people in 2013). However, Figure 4 does tells us that as the company was starting, women and prefer not to say, on average were making more, however, as we move closer to 2020, we see men beginning to make more. Is this difference significant?

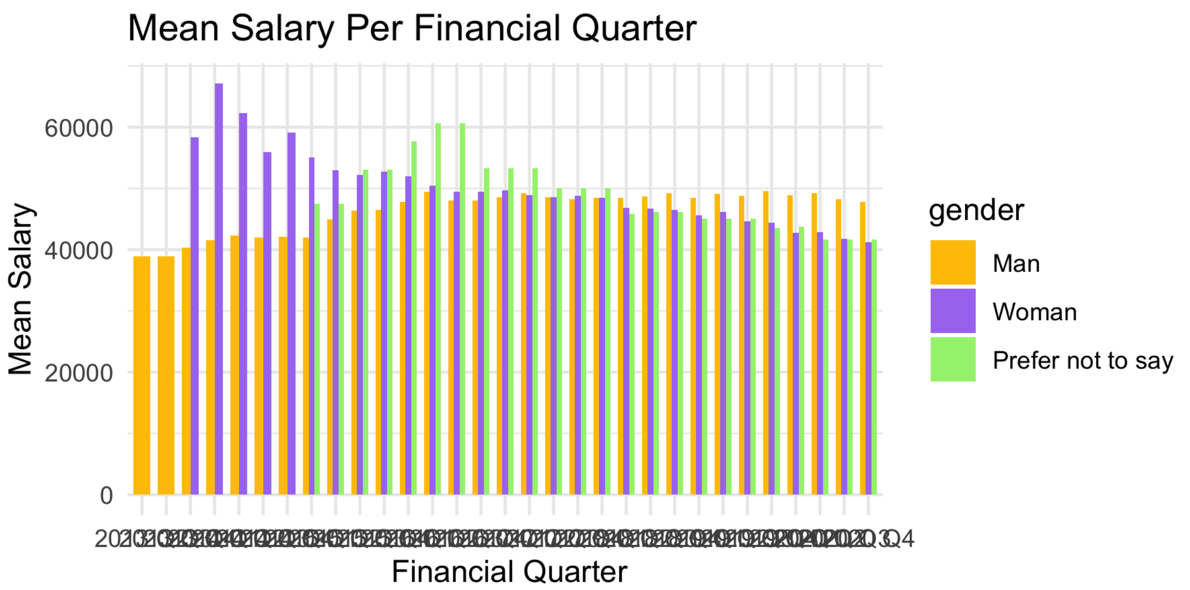

Figure 4: Mean Salary Per Financial Quarter

We can further see this relationship in the Figure 5 boxplots below. At first glance the salaries per seniority level per gender is generally even, though women seem to be making less than men on average.These visualizations give us an inclination of what our models should look like.

Table 16: Log Likelihood Test for Fixed Effect of Gender in Salary

# Df	Log Likelihod	Df	Chi-squared	P-value
13	-58720	NA	NA	NA
11	-58760	-2	92.29	0

Figure 5: Average Salary in Each Role

Using Linear Mixed Models, we tested different random and fixed effects on our data. The model we found that best described our data was a fixed effect on gender, a random effect of employee and team to account for the different employees and teams they are in throughout the financial quarters (treating them as independent) and a random slope for leadership and role seniority as these factors seem to be related with each other as well.
We then performed a log likelihood test to compare the model mentioned above to a model with gender removed (Table 15). If there was no bias, we would expect to see insignificant differences between the two models. However that is not the case.
In fact, our model with gender was significant at the 5% level with $p < 0.05$. Thus, we have evidence to reject the null hypothesis that gender does not impact an employee's salary.

More specifically, our model tells us that we can expect to see women and those who have indicated "prefer not to say" under gender making approximately \$2248.2 and \$1104.1 less respectively than their male counterparts (Table 16).

Table 17: Coefficents of the Fixed Effect of Gender

	Estimate	Std. Error	t value
Baseline	68454.900	11226.823	6.097
Woman	-2248.184	275.248	-8.168
Prefer not to say	-1104.095	1030.859	-1.071

Discussion

In Black Saber's hiring process, we investigate whether the algorithm that grades each applicant is biased towards a certain gender. By running a log-likelihood comparison of two nested generalized linear models (one with a fixed effect for gender, one without) on each of the two AI-rated phases. We found that in both phases, there was insufficient evidence to show that the algorithm was biased towards any gender. Moreover, we considered whether or not the algorithm was more biased towards a certain gender within each team; however, we did not find a significant difference when we added a random effect for team in either phase.

The final, human-involved interview round of Black Saber's hiring process may be something of concern. By running a log-likelihood comparison of two nested generalized linear models (one with a fixed effect for gender, one without) on the average interview rating given by Black Saber employees, we find a slight bias towards males. Despite a larger number of women than men that initially applied to Black Saber's new grad program, and more women moving forward through the phases ranked by the algorithm, only 1 woman on each team was hired. We found bias at the 10% significance level, which we recommend further investigating - we will discuss the data required for this in our limitations section.

In Black Saber's promotion process, we looked to find important variables that the company considers when promoting their employees by comparing different GLM models that look to predict the odds of being promoted. What we found was that while role seniority and productivity are important considerations in the process, gender was the most significant across all models. More specifically we found that women were about 30% less likely to be promoted than their male counterparts which was significant in the 1% level, providing evidence that there is a bias in favour of men being promoted. It was difficult to find sufficient evidence of any bias against those who preferred not to disclose their gender, most likely due to the small sample size of that group.

In Black Saber's salaries, we investigated important variables that contributed to an employee's salary. We found that the best model included a fixed effect on gender which provided evidence that there is gender bias. We controlled for factors we found to have an impact on salary such as role seniority, productivity and the team they were in for that quarter and still found a significant difference in the salaries between the genders. More specifically, we found that women and those preferred not to indicate their gender reported salaries of $2248.2 and $1104.1 less respectively than their male counterparts.

Limitations

Due to the systemic barriers women and non-binary people face in the professional world, it is difficult to determine whether our results are directly due to a gender bias in Black Saber Software's culture, or if the system itself is flawed.
We did not find significant bias in the hiring algorithm, but found slight bias towards men in the interview phase. Thus, it is possible that the interviewers themselves should be wary of their internalized biases when interviewing potential employees. Moreover, because the sample size in the final round and final hires were so small, it is difficult to necessarily determine if the small bias we found was actually significant enough, especially because it was only at the 10% level. In order to properly investigate this result, we would require Black Saber to collect and provide us with data on information on the interviewers (eg. gender), and what information is gathered during the interview.We recommend Black Saber to continue with their hiring algorithm, and possibly implement some training over bias through Human Resources.
Given the conclusions about the data on current employees, we found that gender was consistently an important variable when predicting promotions. An important consideration to further investigate this relationship would be to see what the exact promotion process is, are managers and those in higher roles overseeing promotions? Do promotions have to be approved by vice-presidents? These insights would give us a better direction of the root of the bias we found.
Because sample size for the "prefer not to say" gender group was so small, it was difficult to make many meaningful or significant statistical analyses pertaining to their contribution to the company culture. However, the fact that the group is and has been such a small portion of Black Saber Software's community, we believe that this speaks to the potential that the company has towards hiring more diverse communities in the future.
Furthermore, this data and analysis does not consider those who may not want to take on more responsibilities in a new position for reasons such as work-life balance and thus turned down promotions. This analysis was run under the assumption that everyone wants to be promoted. This can also indirectly impact salary.
We recognize we may not have all the right answers - there may be additional insights that are also helpful. We are not declaring that our ideas are the best and should be necessarily followed - these are merely our honest suggestions from the analyses we have run.

Consultant information

Consultant profiles

Yian Wang. Yian is a junior data analyst at The Hive. She specializes in reproducible visualization and making actionable insights. Yian earned her Bachelor of Science, double majoring in Statistics and Economics, and minoring in Mathematics from the University of Toronto in 2021.

Claire Hsiung. Claire is a junior financial analyst at The Hive. She specializes in interpretable visualizations and Big Data. Claire earned her Bachelor of Science, double majoring in Statistics and Economics from the University of Toronto in 2021.

Code of ethical conduct

Not only is The Hive passionate about making actionable insights, but we also practice ethical statistics. Our main values lie in, but are not limited to:

- Confidentiality of client information and respecting their rights
- Using appropriate methods and interpreting them correctly and completely
- Reporting results impartially even if they may pose harm to the parties involved, so as to encourage action against our insights so as to not fall into trouble in the future
- Only using methods we have sufficient knowledge in to use in order to prevent any misinterpretation and summary of the data
- Only using data that is provided to us directly by Black Saber Software, and not to scrape other data that we do not have permission to access
- Declaring our relationship with the clients (ie. financial or other interests) to maintain transparency regarding the influence it may have on the outcomes