Data Analysis Report

Analysis on Gender Parity in Hiring, Wages and Promotion

Report prepared for Black Saber Software by Research and Developing Insight (RADI)

2021-04-21

Contents

Executive summary

Guidelines for the executive summary:

- *No more than two pages*
- *Language is appropriate for a non-technical audience*
- *Bullet points are used where appropriate*
- *A small number of key visualizations and/or tables are included*
- *All three research questions are addressed*

Technical report

Introduction

Employees at Black Saber have been raising concerns about potential bias in the company's hiring and remuneration processes. This report will cover the topic of gender parity in wages, promotions, and hiring at Black Saber. We will be working with the hiring data for Black Saber's new-grad program and the data about promotions and salary for Black Saber's current employees, which span from 2013 to 2020. We will cover all data manipulations we make to the provided datasets, exploratory analysis, models considered and fitted, and our results and findings.

Research questions

The research questions in this report are all based around gaining more insight about Black Saber's salary, promotion and hiring process and whether or not they are fair on the basis of gender parity. These research questions include:

- What factors, such as productivity, leadership for level and gender, etc. are related to an employee's salary?
- What factors, such as productivity, leadership for level and gender, etc. are related to the number of promotions an employee receives?
- What factors, such as gpa, work experience, extracurriculars, gender, etc. are related to new-grad candidates being hired?

Analysis of Gender Parity in Black Saber's Employee Salaries

The dataset we are working with contains 6906 observations of current employees at Black Saber. We began by filtering out all observations that had listed "Prefer not to say" in its gender column, which left us with 6789 observations. Next, we seperated the financial_q column into its respective columns for year and financial quarter. We also made a manipulation to convert the salaries in the data from a character type to an integer type, as well as factor releveling the role_seniority column, so that it was in the appropriate ascending order from "Entry-Level" to "Vice President".

A premise of this analysis is that we suspect that an employee at Black Saber's salary can be explained from their gender from the data between 2013 and 2020.

From Table 1, we can clearly see that, on average, men at Black Saber have higher salaries than women. Fitting a boxplot comparing salaries of men and women across the years, we also observe that after 2016, men tend to have higher salaries than females. Therefore, we will be considering

Table 1: Summary of salaries by gender at Black Saber from 2013 to 2020

gender	observations	minimum salary	maximum salary	mean salary	median salary
Man	4064	27100	154200	48490.92	42400
Woman	2725	25000	149400	45810.50	40000

gender as a fixed covariate in our model to test its significance in predicting an employee's salary at Black Saber.

Upon fitting a histogram to salaries from the current employees dataset, we noticed that there is a right skew to the distribution. Considering the seniority covariate, we are able to correct much of that skew.

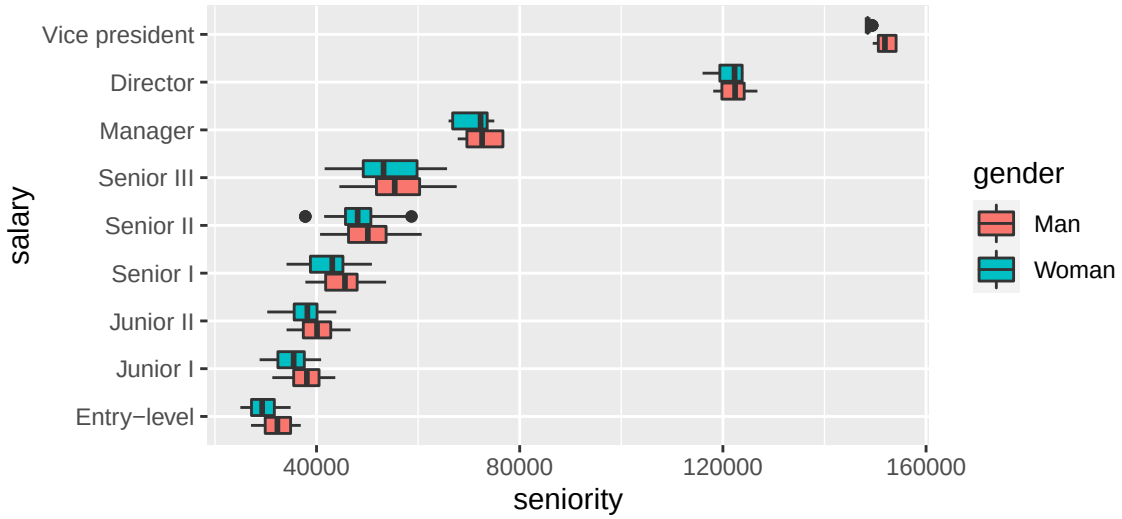

Figure 1: Boxplot of Salaries and Seniority for Men and Women

From Figure 1, we can see that after conditioning salary on seniority and roles, the boxplots indicate more of a normal distribution within each roles for both men and women. Thus, these visualizations and exploratory analysis helped us with selecting a linear mixed model to predict salaries at Black Saber.

We chose a linear mixed model to predict salaries at Black Saber because our simple linear model assumptions are violated since our observations are not independent due to employees being measured more than once within and across the years. Also, we can see that the response type, salary, follows a normal distribution when considering the grouping of seniority, and that random effects can be considered for each employee, year, and team.

We considered two linear mixed models, with fixed effects for gender, productivity and leadership for level, and random intercepts for the team, year, and employee ID. One model was fitted with

Table 2: Likelihood Ratio Test for The Linear Mixed Models on Salaries

#Df	LogLik	Df	Chisq	Pr(>Chisq)
10	-57747.22	NA	NA	NA
15	-57733.25	5	27.92321	3.77e-05

a random intercept for the role/seniority, and another model was fitted with a random slope for role/seniority for different leadership for level. We performed a Likelihood ratio test (Table 2), which is appropriate given the two models have the same fixed effects and are only comparing nested random effects, and observed a chi-square test statistic of 27.923 at a p-value of 3.768e-5. Thus, random slopes for role/seniority explains the data better than a random intercept for role/seniority and we have very strong evidence against the hypothesis that the simpler model fits the data just as well.

Table 3: Summary of Coefficients for The Linear Mixed Model on Salaries

	Estimate	Std. Error	t value
(Intercept)	67863.73900	14034.628191	4.835450
genderWoman	-2251.21165	276.815074	-8.132547
productivity	-1.08661	1.223819	-0.887885
leadership_for_levelExceeds expectations	-328.35494	216.836898	-1.514295
leadership_for_levelNeeds improvement	328.70762	245.603619	1.338366

In conclusion, from Table 3, our linear mixed model shows that women at Black Saber have salaries that are, on average, $2251.21 less than men at Black Saber. At a t-value of -8.132547, and given that our 95% confidence interval for the gender covariate is (-2799.25, -1712.30), we observe that the gender covariate is significant in predicting salaries of current employees at Black Saber, which could potential lead to some gender parity issues.

Analysis of Gender Parity in Black Saber's Employee Promotions

We began by taking the dataset used in the analysis of gender parity and employee salary section, which had 6789 observations. Next, we created a role/seniority index variable that assigned a value from 0 to 8 to the roles of "Entry-level" to "Vice President", respectively. We also grouped our dataset my employee IDs and created a new variable called promotions, which took the difference between an employees most recent role and their first role at the company to calculate how many promotions they have had at Black Saber.

A premise of this analysis is that we suspect that gender explains the number of promotions an employee has at Black Saber in the data between 2013 and 2020. Upon fitting a histogram to promotions from the current employee dataset, we noticed that there is a right skew to the

distribution. Since our response variable, promotion, is a count, we will be considering a Poisson regression model.

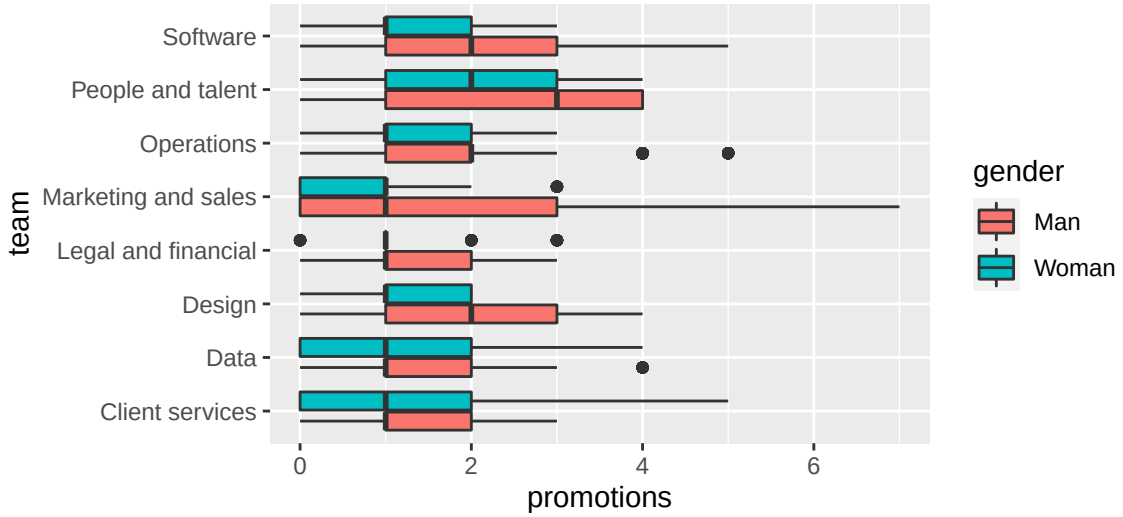

Figure 2: Boxplot of Promotions and Teams for Men and Women

From Figure 2, we can see that by grouping employee promotions into the different teams at Black Saber, the boxplots indicate that for half the teams, client services, data, legal and finance, and marketing and sales, men usually have a higher median number of promotions during their time at Black Saber than women do. Hence, these visualizations and exploratory analysis helped us with selecting a generalized linear mixed model with a Poisson distribution to predict the number of promotions for an employee at Black Saber.

We chose a generalized linear mixed model to predict salaries at Black Saber because our observations are not independent due to the groupings of employees into different teams. Also, we can see that the response type, promotion, follows a Poisson distribution, and that potential random effects can be considered for each employee, year, and team.

We began by fitting a Poisson model with fixed effects for gender, leadership for level, and productivity. Looking at the summary, we found that both gender (z-value: -13.197 and p-value: $<$ 2e-16) and the exceeds expectation (z-value: 6.037 and p-value: 1.57e-9) covariates were statistically significant. However, a goodness-of-fit test reveals that there remains significant lack-of-fit (residual deviance: 8758.7 with only 6784 df; p-value $<$ .001 based on chi square test with 6784 df).

Then, we fitted a Generalized linear mixed model with Poisson family, fixed effects for gender, leadership for level, productivity, and random intercepts for the team and year. Looking at the summary, we found that gender (z-value: -13.244 and p-value: $<$ 2e-16) the exceeds expectation

(z-value: 3.268 and p-value: 0.00108), and the productivity (z-value: 2.836 and p-value: 0.00456) covariates were statistically significant.

Table 4: Drop-in-Deviance Test/Analysis of Deviance Table for GLM and GLMM on Promotions

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	6788	9022.107	NA
gender	1	228.766018	6787	8793.341	0.0000000
leadership_for_level	2	31.703849	6785	8761.638	0.0000001
productivity	1	2.961048	6784	8758.677	0.0852926

We performed a Drop-in-deviance test (Table 4) using an Analysis of Deviance table to compare the two models and test for the significance of our model coefficients. The drop-in-deviance test statistic for gender was 228.766 on 1 df, producing a p-value of $< 2.2e\text{-}16$ and the drop-in-deviance test statistic for leadership for level was 31.704 on 2 df, producing a p-value of 1.305e-7. Thus, there is strong evidence against the null hypothesis that a simpler model without gender or leadership for level covariates would better explain our data on promotions.

In conclusion, our Poisson mixed model shows that women at Black Saber, on average, receive 0.742 times the promotions that men do during their time at Black Saber. Given that our 95% confidence interval for the gender covariate is (0.718, 0.782), we observe that the gender covariate is, again, significant in predicting the number of promotions for current employees at Black Saber, which could also potentially lead to some gender parity issues.

Analysis of Gender Parity in Black Saber's Hiring Process

The dataset we are working with contains 613 observations of Black Saber phase 1 applicants. We began by filtering out all observations that had listed "Prefer not to say" in its gender column, which left us with 602 observations. Next, we performed a left join with the phase 2 applicants dataset and another left join with the phase 3 applicants dataset in order to combine it into one comprehensive dataset for analyzing the AI automated component of the hiring process. We also created a "move forward from phase 1" and a "moved forward from phase 2" variable, and removed the interviewer rating variables from this dataset. Next, we took the phase 3 applicants dataset and performed a left join with the phase 1 applicant dataset to get the gender and "team applied for" into this dataset for analysis of the human component of the hiring process. Lastly, we left joined the dataset with the final hire dataset in order to get the applicants that passed phase 3 of the hiring process and created a combined rating variable for the combined score on both interviews in phase 3.

A premise of this analysis is that we suspect that gender explains the candidates that are hired

in Black Saber's hiring process in the data. Upon fitting a scatterplot of technical skills against those that move forward from phase 2 in the AI automated hiring dataset, we noticed that men had increasingly higher probability of moving forward than women at a scores greater than 25. Since our response variables, moving forward in phase 1 and 2 and status of hire in phase 3, are all probabilities, we will be considering a Logistic (Binomial) regression model.

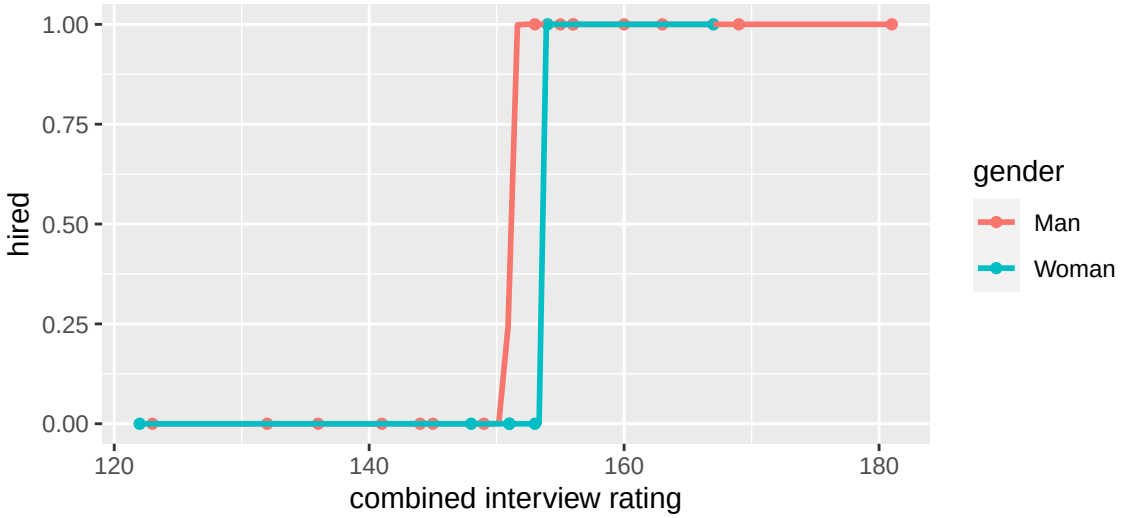

Figure 3: Scatterplot of Combined Interview ratings and Hires for Men and Women

From Figure 3, we can see that in the phase 3 hiring dataset, a scatterplot of the combined interviewer ratings against the applicants that were hired follows a Binomial/Logistic distribution. Thus, these visualizations and exploratory analysis helped us with selecting a generalized linear mixed model with a binomial distribution to predict the probability of being hired in the Black Saber phase 3 hiring process. We also considered fitting a generalized linear mixed model with a binomial distribution to predict the probability of moving forward in phase 1 and phase2 in the AI automated dataset.

We chose a generalized linear mixed model to predict salaries at Black Saber because our observations are not independent due to the groupings of applicants into different teams that they applied for. Also, we can see that the response type, moving forward in phase 1, moving forward in phase 2, and being hired, all follows a Binomial distribution, and that potential random effects can be considered for each applicant and team applied for.

We began by fitting a Generalized linear model with "moving forward from phase 1" as the response, fixed effects for cover letter, cv, gpa, gender, extracurriculars, and work experience, and a Binomial family. Then we fitted a Generalized linear mixed model with the same response, fixed effects and family, and the addition of random intercepts for each applicant and the team applied for.

Table 5: Drop-in-Deviance Test/Analysis of Deviance Table for GLM and GLMM on Phase 1

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	601	834.44289	NA
cover_letter	1	414.6648073	600	419.77808	0.0000000
cv	1	118.3933204	599	301.38476	0.0000000
gpa	1	150.7314046	598	150.65336	0.0000000
gender	1	0.8572304	597	149.79613	0.3545149
extracurriculars	1	28.2076517	596	121.58848	0.0000001
work_experience	1	86.6990088	595	34.88947	0.0000000

We performed a Drop-in-deviance test (Table 5) to compare the models and test for the significance of model coefficients and found that gender was the only covariate that yielded a p-value > 0.1 (p-value $= 0.3545$). Thus there is no evidence against the null hypothesis that gender explains applicants moving forward from phase 1 in the AI automated hiring dataset.

Next, we fitted a Generalized linear model with "moving forward from phase 2" as the response, fixed effects for gender, technical skills, writing skills, leadership presence and speaking skills, and a Binomial family. Then we fitted a Generalized linear mixed model with the same response, fixed effects and family, and the addition of random intercepts for each applicant and the team applied for.

Table 6: Drop-in-Deviance Test/Analysis of Deviance Table for GLM and GLMM on Phase 2

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	296	156.84692	NA
gender	1	3.630446	295	153.21647	0.0567316
technical_skills	1	16.714182	294	136.50229	0.0000435
writing_skills	1	11.068057	293	125.43423	0.0008783
leadership_presence	1	29.674302	292	95.75993	0.0000001
speaking_skills	1	26.512490	291	69.24744	0.0000003

Again, we performed a Drop-in-deviance test (Table 6) to compare the models and test for the significance of model coefficients and found that gender was the only covariate that yielded a p-value $= 0.0567316$. Thus there is weak evidence against the null hypothesis that gender explains applicants moving forward from phase 2 in the AI automated hiring dataset.

Finally, we took a look at the phase 3 hiring dataset and fitted a Generalized linear model with "hired" as the response, fixed effects for gender, interviewer rating 1 and interviewer rating 2, and a Binomial family. Then we fitted a Generalized linear mixed model with the same response, fixed effects and family, and the addition of random intercepts for each applicant and the team applied for.

Table 7: Drop-in-Deviance Test/Analysis of Deviance Table for GLM and GLMM on Phase 3

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL	NA	NA	21	30.31641	NA
interviewer_rating_1	1	11.45155	20	18.86486	0.0007143
interviewer_rating_2	1	18.86486	19	0.00000	0.0000140
gender	1	0.00000	18	0.00000	0.9999591

Again, we performed a Drop-in-deviance test (Table 7) to compare the models and test for the significance of model coefficients and found that gender was the only covariate that yielded a p-value > 0.1 (p-value $= 0.9999591$). Thus there is no evidence against the null hypothesis that gender explains applicants being hired in the phase 3 hiring dataset.

In conclusion, our Logistic mixed models showed that in all 3 phases of the hiring process, gender had little to no effect on whether or not an applicant moved forward from the first 2 phases and whether or not an applicant was hired in the phase 3. This is potentially good news on the topic of gender parity in Black Saber's hiring process as it suggests that in each step, the covariates other than gender are all statistically significant (p-value < 0.001) and explain most of whether or not an applicant moves forward or is hired.

Discussion

We found that with regards to salaries and promotions at Black Saber, women have salaries that are $2251.21 less than men and have 0.742 times the promotions that men have while working at Black Saber, on average. In addition, we observed that with regards to Black Saber's hiring process, applicants' gender showed weak to no evidence (p-value > 0.05) in explaining whether or not they moved forward in phase 1 and 2 and whether or not they were hired. As for the topic of gender parity in hiring, wages and promotion, we believe there is a need to be out and ahead of the potential issues in the wages and promotions department.

Limitations

A limitations of this analysis may result from the definition of certain covariates used in addressing the research questions. Things such as productivity scores, leadership for level, leadership presence, etc, may be subjective to those acquiring the dataset at Black Saber. This may suggest some ineffective covariates for the response variables.

Another limitation is that size of the dataset for the hiring process may not be sufficient to make a significant conclusion on any factors and their effects on being hired. We might suggest acquiring a larger sample in the future and performing another statistical analysis on the covariates.

Finally, a limitation may be related to confidentiality. Depending on who is acquiring the datasets for the current employees and hiring process, there may be some bias for certain categorical covariates and/or in the interviewer ratings from those administering the performance review or interview.

Consultant information

Consultant profile

Ryan Jo Wong. Ryan is a junior consultant with Research and Developing Insight (RADI). He specializes in data wrangling and visualizations, and statistical analysis. Ryan earned an Honours Bachelor of Science, Majoring in Statistics and Minoring in Computer Science and Mathematics from the University of Toronto in 2022.

Code of ethical conduct

Here at Research and Developing Insight (RADI), we believe that it is essential to uphold high standards in our statistical practices and be responsible when delivering statistical insight to our clients and stakeholders. Thus, we adhere strongly to the Statistical Society of Canada's (SSC) Code of Ethical Statistical Practice. With regards to society, our mission is to provide objective, unbiased information with the goal of improving public knowledge and avoiding misleading information in our reports. We hold fast to the privacy laws and standards set forth by the SSC as it pertains to the collection, storage, and publication of data and findings. With our employers and clients, it is important that we remain constantly aware of any potential conflicts between our ethical practices and the interests of the client, and to refrain from any misleading summaries or findings from our data in reports. It is our utmost priority to avoid disclosing any confidential and/or sensitive information acquired without the explicit consent from our employers and clients. Our duty is to remain professional in keeping good statistical practices and standards by only taking on projects that are within the skill set and capacity of our team. In addition, we refuse to partake in projects where the outcomes or findings in our statistical reports dictate the financial gains and rewards for our services.