
Report title

Subtitle

Report prepared for Black Saber Software by Zeusolutions

2021-04-21

Contents

General comments (you can delete this section)	2
Executive summary	3
Technical report	4
Introduction	4
Informative title for section addressing a research question	4
(Hiring Question) Fairness of the AI Recruitment Pipeline Manager	6
(Promotion Question)	12
(Salary Question) Fairness of Salary assignment for current employees at Black Saber	15
Discussion	15
Consultant information	16
Consultant profiles	16
Code of ethical conduct	16

General comments (you can delete this section)

You can delete this section, and if you want to check what it said, just open a template from the package again. You don't have to use this particular template, but you DO need to write your report in RMarkdown and include a cover page.

The cover page must have:

- *A title and subtitle*
- *"Report prepared for Black Saber Software by" your company name*
- *Date (assessment submission date is fine)*

You can change the colour of this cover to any colour you would like by replacing 6C3082 in the YAML above (line 11) to another hex code. You could use this tool to help you: <https://htmlcolorcodes.com/color-picker/>

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

This part of the report is much more comprehensive than the executive summary. The audience is statistics/data-minded people, but you should NOT include code or unformatted R output here.

Introduction

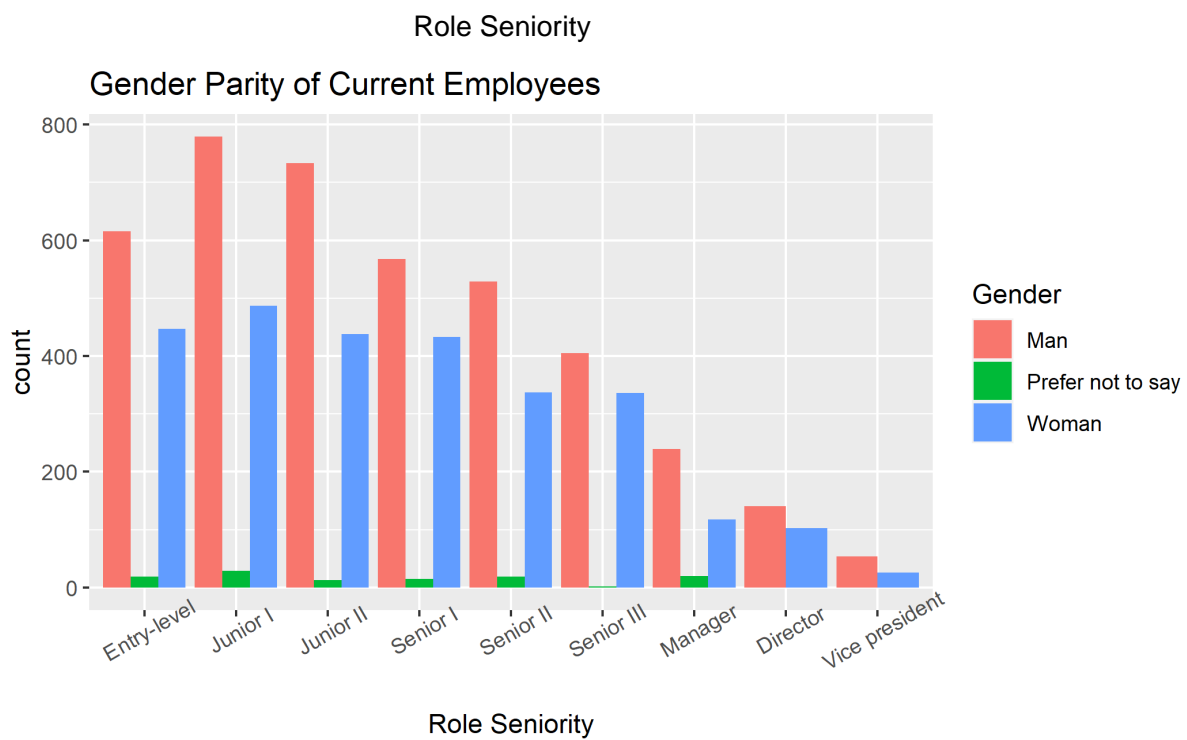
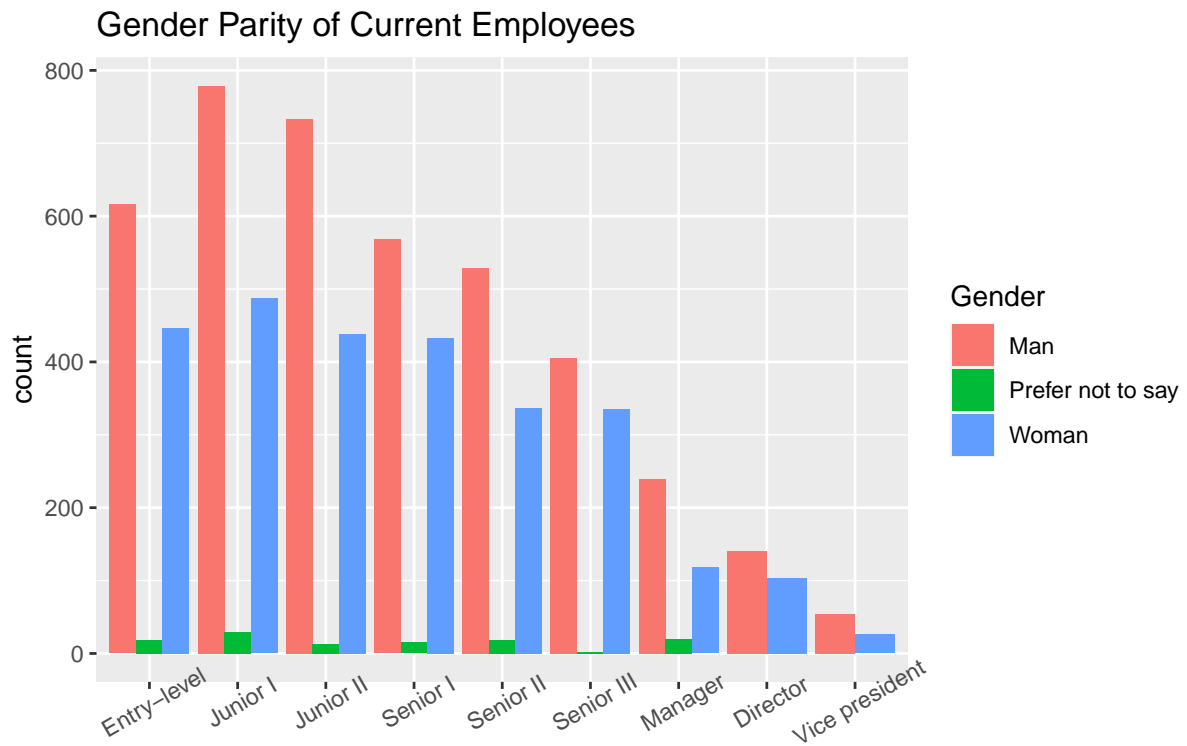
Provide a brief introduction to your report and outline what the report will cover. This section is valuable for setting scope and expectations.

Research questions

Use bullet points to describe the research questions you are going to address. Write in full sentences.

Informative title for section addressing a research question

For each research question, you will want to briefly describe any data manipulation, show some exploratory plots/summary tables, report on any methods you use (i.e. models you fit) and the conclusions you draw from these



(Hiring Question) Fairness of the AI Recruitment Pipeline Manager

The company Black Saber has a new AI recruitment pipeline manager that has been trialing for the Data and Software teams. The board has concerns about the process and questions whether or not it is fair and based on value and talent of each applicant. We have been given the data for each phase of the process, in which there are a total of three. Through each phase there is certain data collected through question and tasks in which the AI uses to narrow down the field of applicants from one phase to another. We want to determine what factors play a major role in being hired through Black Saber's recruitment pipeline manager.

(Phase 1) Defining Features of the Application Form

The first phase of the hiring pipeline is the submission of an application form, cover letter and CV. Each line of `phase1-new-grad-applicants-2020.csv` contains the information of each applicant in phase 1:

- `applicant_id` = A unique ID given to each applicant in Phase 1
- `team_applied_for` = Software or Data
- `cover_letter` = 1 if present, 0 if not.
- `cv` = 1 if present, 0 if not
- `gpa` = 0.0 to 4.0
- `gender` = "Man", "Woman", or "Prefer not to say" as the only options
- `extracurriculars` = 0, 1 or 2. Description in which is accessed automatically by keywords where 2 indicates high, 1 indicates some and 0 indicates no relevant skill building extracurriculars.
- `work_experience` = 0, 1 or 2. Description in which is accessed automatically by keywords where 2 indicates high, 1 indicates some and 0 indicates no relevant work experiences.

Similarly, we are also given the list of applicants who have made it to phase 2 along with the factors assessed in phase 2. However, in this section we are only interested in which applicants made it to phase 2. For this reason we will create a variable called `passed_1` and set it to 1, if the applicant made it from phase 1 to phase 2 and 0, if not. This will be the response variable we are interested in for developing our model. The other variables will be fixed effects besides `team_applied_for`. We believe that `team_applied_for` should be added as a random effect because each team has a different set of required skills and/or experiences. Other than that, since each row corresponds to a unique applicant, this suggests that the observations are independent, hence our assumptions are not violated. We want to determine which is the most impactful component of the application form that allows the applicants to proceed to phase 2 of the hiring process.

Table 1: The first five applicants from phase 1 of the hiring pipeline

applicant_id	team_applied_for	cover_letter	cv	gpa	gender	extracurriculars	work_experience
1010	Software	0	1	1.3	Man	1	1
1020	Software	0	1	3.4	Woman	2	1
1030	Data	1	1	2.4	Woman	2	0
1040	Software	0	1	2.7	Man	1	1
1050	Data	1	0	2.1	Prefer not to say	0	1

Given the binomial response variable and random effect, we will be constructing a generalized linear mixed model. Before constructing the model, we want to first consider the gender distribution for those who passed phase 1 and those who did not. In addition, consider also the marginal probabilities and odd ratios.

Table 2: The number of applicants that passed phase 1 based on gender

	Did Not Pass Phase 1	Passed Phase 1
Man	146	145
Prefer not to say	8	3
Woman	159	152

Table 3: The probability of an applicant's gender given that they pass phase 1

	Did Not Pass Phase 1	Passed Phase 1
Man	0.4664537	0.4833333
Prefer not to say	0.0255591	0.0100000
Woman	0.5079872	0.5066667

Hence from the table above, given that the applicant passes phase 1, the probability that their gender is a “Man” is 48%, whereas a “Woman” is 51%. Also, the odds ratio of Woman vs Man who pass phase 1 is approximately 0.96. That is, the odds of a “Woman” passing phase 1 is less than that of a “Man”. This odds ratio is relatively close to 1, hence there is no reason to suspect that there are any gender biases in those who passed phase 1.

Then, based on the response variable and random effect we will be considering a generalized linear mixed model.

Table 4: Generalized linear mixed model of applicants who passed Phase 1

Characteristic	log(OR)	95% CI	p-value
cover_letter	63	-24,344, 24,471	>0.9
cv	52	-41,425, 41,528	>0.9
gpa	13	6.1, 19	<0.001
gender			
Man	—	—	
Prefer not to say	0.63	-25, 26	>0.9
Woman	0.91	-1.0, 2.8	0.4
extracurriculars	10.0	5.1, 15	<0.001
work_experience	12	6.1, 17	<0.001

OR = Odds Ratio, CI = Confidence Interval

It is evident that **gpa**, **extracurriculars** and **work_experience** have the most significant effect on passing Phase 1, based on their log odds. We cannot interpret the odds ratio of **gpa** because there is no one with a 0.0 GPA, however we can interpret the odds ratio of **extracurriculars** and **work_experience** to an extent. Since **extracurriculars** and **work_experience** are given scores of either 0, 1 or 2, the odds ratio is an average of scores 1 and 2. Nonetheless, we find that the odds of passing Phase 1 with some to high relevant skills from extracurriculars is 21267.4 times that of those without and the odds of passing Phase 1 with some to high relevant work experience is 107591 times that of those without. Overall, we can conclude that the most impactful component of the application form that allows the applicants to proceed to phase 2 of the hiring process in order of is **gpa**, followed by **extracurriculars** and **work_experience**.

(Phase 2) AI Scoring Influences

The second phase of the hiring pipeline invites applicants to do a timed technical coding task, writing sample and submit a pre-recorded video. Each line of `phase2-new-grad-applicants-2020.csv` contains the AI graded scores for skills relevant to each task:

- **applicant_id** = A unique ID given to each applicant in Phase 1
- **technical_skills** = AI graded score from 0 to 100 for the timed technical task
- **writing_skills** = AI graded score from 0 to 100 for the timed writing task
- **speaking_skills** = AI graded rating from 0 to 10 for the speaking ability in the pre-recorded video

- **leadership_presence** = AI graded rating from 0 to 10 for “leadership presence” in the pre-recorded video

In this phase, we are interested in what factors from the application form affect the AI scoring/rating of technical, writing, and speaking skills and leadership presence. Thus we will be using the data from `phase1-new-grad-applicants-2020.csv`, which has been explained in the previous section. We will have 4 models with the response variables being **technical_skills**, **writing_skills**, **speaking_skills** and **leadership_presence**. For the fixed effects and random effects, we will be using the exact same found in the final model presented for phase 1 except for **cover_letter** and **cv**. The reason for not including **cover_letter** and **cv** is because only those who had submitted both a cover letter and CV made it to phase 2. Based on the response variables, we will be using a linear mixed model for each.

Table 5: Linear mixed model of the
Technical Skills Score

Characteristic	Beta	95% CI
gpa	8.8	3.9, 14
gender		
Man	—	—
Prefer not to say	23	-1.2, 47
Woman	1.0	-3.8, 5.8
extracurriculars	4.6	-0.39, 9.6
work_experience	-0.62	-6.6, 5.4
CI = Confidence Interval		

From the linear mixed model of the Technical Skill Score, it follows that the most influential fixed effect is **gpa** followed by **extracurriculars**. Similar to before, we cannot interpret **gpa** because there is no one with a 0.0 GPA. As for **extracurriculars**, those with relevant skills from extracurriculars have a higher score than those who do not by an average of 4.6. We also notice that those who put “Prefer not to say” for their gender have higher score than those who put “Man” by an average of 23. This is not concerning because the sample of those who put “Prefer not to say” is very small and thus can be seen as outliers to the overall average. Besides that, those who put “Woman” for their gender only have a higher score than those who put “Man” by an average of 1.

Table 6: Linear mixed model of the Technical Skills Score

Characteristic	Beta	95% CI
gpa	9.9	5.8, 14
gender		
Man	—	—
Prefer not to say	5.2	-15, 26
Woman	3.8	-0.23, 7.8
extracurriculars	-1.5	-5.6, 2.7
work_experience	4.1	-0.88, 9.2

CI = Confidence Interval

The linear mixed model of the Writing Skill Score shows similar results to the one for the Technical Skill Score with the most influential fixed effect being **gpa**. Unlike the Technical Skill Score, the second most influential fixed effect for the of the Writing Skill Score is **work_experience**. That is, those with relevant work experience have a higher score than those who do not by an average of 4.1. We also find that those who put “Woman” for their gender only have a higher score than those who put “Man” by an average of 3.8.

Table 7: Linear mixed model of the Speaking Skills Rating

Characteristic	Beta	95% CI
gpa	-0.30	-0.81, 0.21
gender		
Man	—	—
Prefer not to say	1.6	-0.90, 4.1
Woman	-1.8	-2.3, -1.3
extracurriculars	0.47	-0.05, 1.0
work_experience	0.93	0.31, 1.6

CI = Confidence Interval

The linear mixed model of the Speaking Skills Rating is significantly different from the previous models. The model shows that the most influential factor of the rating is **work_experience**. More specifically, those with relevant work experience have a higher rating than those who do not by an average of 0.93. The second most influential factor of the rating is **extracurriculars**, showing a higher rating on average of 0.47 for those who have relevant skills from extracurriculars compared to those who do not. Surprisingly, there is a negative coefficient for **gpa**, showing that those with higher GPAs have a lower rating on average. We also see a negative coefficient for those who put “Woman” for their gender as they have a lower rating compared to those who

put “Man” by an average of 1.8.

Table 8: Linear mixed model of the
Leadership Presence Rating

Characteristic	Beta	95% CI
gpa	0.22	-0.32, 0.75
gender		
Man	—	—
Prefer not to say	-0.44	-3.1, 2.2
Woman	-0.91	-1.4, -0.39
extracurriculars	0.76	0.22, 1.3
work_experience	0.16	-0.48, 0.81

CI = Confidence Interval

The linear mixed model of the Leadership Presence Rating shows that the most influential factor is no other than **extracurriculars**. This is followed by **gpa** and then **work_wxperience**. By interpreting the coefficients of **extracurriculars** and **work_wxperience**, we find that those who have the relevant skills for each, compared to those who do not, have a higher rating on average of 0.76 and 0.16, respectively. Similar to the Speaking Skills Rating, we also see a negative coefficient for those who put “Woman” for their gender as they have a lower rating compared to those who put “Man” by an average of 0.91.

After analyzing all of the possible factors that can affect the scores and rating given by the AI, we conclude that the top most influential factors are no other than **gpa**, **extracurriculars** and **work_experience**. We also find that those who put “Woman” as their gender, score higher than those who put “Man” on average for the Technical and Writing Task. On the other hand, those who put “Man” as their gender, were rated higher than those who put “Woman” on average for the Speaking Skills and Leadership Presence. Ultimately the differences between the genders are small relative to the other factors, hence there is no concern about bias in gender parity for the AI scores and ratings.

(Promotion Question)

Black Saber aims to promote an equitable and inclusive work environment based on merit. The board wants to assure that its promotion processes are not only fair, but strictly based on talent to the company. We have been given the data for the company's current employees and their key performance indicators throughout their time at the company marked by financial quarters. We wish to determine which factors contribute to employees moving up the ranks of Black Saber seniority.

Features Recorded of Black Saber's Current Employees

While employed at Black Saber Software, employee performance indicators are recorded every financial quarter. Each line of `black-saber-current-employees.csv` contains the information of each employee at the financial quarters they were employed:

- `employee_id` = A unique ID given to each employee of Black Saber
- `gender` = "Man", "Woman", or "Prefer not to say" as the only options
- `team` = Divisions of Black Saber staff categorized into: "Software", "People and talent", "Operations", "Marketing and sales", "Legal and financial", "Design", "Data", and "Client services".
- `financial_q` = Financial quarter where following features were recorded.
- `role_seniority` = Hierarchy of employees within each team. Rank order from least senior to most senior: "Entry-level", "Junior I", "Junior II", "Senior I", "Senior II", "Senior III", "Manager", "Director", "Vice president".
- `leadership_for_level` = Quality of demonstrated leadership, taking into account role level. (i.e. "Appropriate for level" requires much less for entry-level employees than for a manager). Categorized into "Needs Improvement", "Appropriate for Level", and "Exceeds Expectations".
- `productivity` = Work output in relation to job description, rated on a 0-100 scale with 50 being satisfactory and above 50 indicating better than expected productivity.
- `salary` = Salary at the given financial quarter (note: these are effective yearly values for the current wage, but don't take in to account previous salary steps in the same year, etc.)

For this question, we are interested in how many promotions an employee received while working at Black Saber Software. For this reason, we will create a variable called `number_of_promotions` counting the change in `role_seniority` of each employee. This will be the response variable we are interested in for developing our model. The other variables will be fixed effects besides `team_applied_for`. We believe that `team_applied_for` should be added as a random effect

Table 9: Five Anonymous Employees at Black Saber Software

<i>employee_id</i>	<i>gender</i>	<i>team</i>	<i>financial_q</i>	<i>role_seniority</i>	<i>leadership_for_level</i>	<i>productivity</i>	<i>sal</i>
24149	Man	People and talent	2020 Q4	Entry-level	Appropriate for level	15	\$3
24061	Woman	Software	2020 Q4	Junior I	Appropriate for level	61	\$3
24219	Man	Operations	2020 Q4	Junior I	Appropriate for level	56	\$3
24484	Man	Client services	2020 Q4	Senior III	Appropriate for level	41	\$5
24368	Man	Legal and financial	2020 Q4	Entry-level	Appropriate for level	59	\$3

because each team has a different set of required skills and/or experiences. Other than that, since each row corresponds to a unique applicant, this suggests that the observations are independent, hence our assumptions are not violated. We want to determine which is the most impactful component of the application form that allows the applicants to proceed to phase 2 of the hiring process.

Given the ... response variable and random effect, we will be constructing a ... model. Before constructing the model, we want to first consider the gender distribution in number of promotions of current employees. In addition, consider also the joint probabilities.

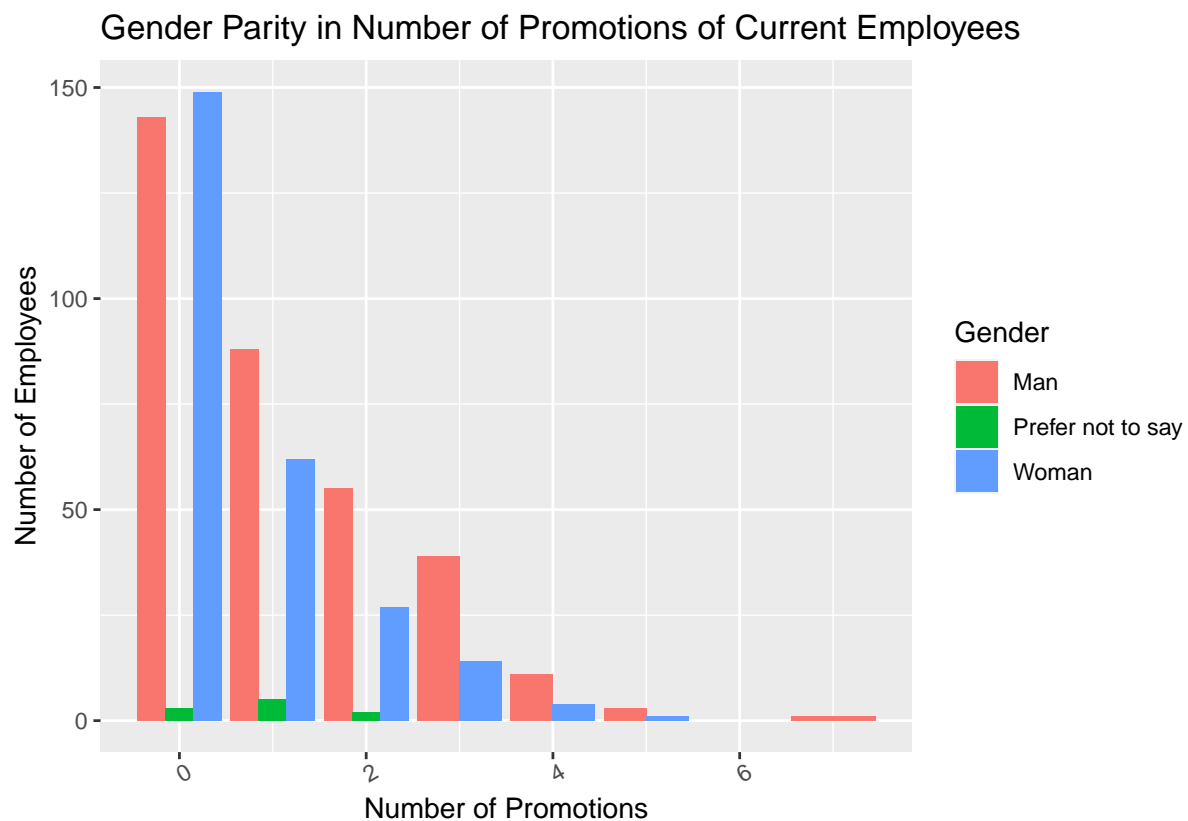


Table 10: The number of employee promotions based on gender

	0	1	2	3	4	5	7
Man	143	88	55	39	11	3	1
Prefer not to say	3	5	2	0	0	0	0
Woman	149	62	27	14	4	1	0

Table 11: The probability of an employee's gender given number of promotions

	0	1	2	3	4	5	7
Man	0.4847458	0.5677419	0.6547619	0.7358491	0.7333333	0.75	1
Prefer not to say	0.0101695	0.0322581	0.0238095	0.0000000	0.0000000	0.00	0
Woman	0.5050847	0.4000000	0.3214286	0.2641509	0.2666667	0.25	0

From the tables above, the only time the probability of an employee being a woman is higher is given the employee has not had any promotions. Conversely, seeing that given a higher number of promotions, the probability of being a man becomes proportionally larger, a gender bias may be suspected however we will withhold any conclusions until we consider our model.

(Salary Question) Fairness of Salary assignment for current employees at Black Saber

Information of each employee at the financial quarters they were employed will be utilized again from `black-saber-current-employees.csv`.

In this section, we are interested in the fairness of salary assignment for current employees at Black Saber. We will be looking how `role_seniority`, `leadership_for_level`, `productivity`, and `gender`.

Discussion

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

Strengths and limitations

Consultant information

Consultant profiles

Complete this section with a brief bio for each member of your group. If you are completing the project individually, you only need to complete one for yourself. In that case, change the title of this section to “Consultant profile” instead. Examples below. This section is only marked for completeness, clarity and professionalism, not “truth” so you can write it as if we’re a few years in the future. Put your current degree in as completed and/or add your first choice grad school program, whatever you like. What skills related skills would you most like to highlight? What job title do you want?

Andy Vu. Andy is a junior consultant with Zeusolutions. He specializes in statistical modeling. Andy earned his Bachelor of Science, Specializing in Mathematics & Its Applications (Probability/Statistics) and Majoring in Statistics from the University of Toronto in 2021.

Ethelia Choi.

James F. Kanu. James is a junior consultant with Zeusolutions. He specializes in data visualization. James earned his Bachelor of Science, Majoring in Mathematics and Minor in Computer Science and Statistics from the University of Toronto in 2021.

Justin Lee. Justin is a junior consultant with Zeusolutions. He specializes in data analytics. Justin earned his Bachelor of Science, Majoring in Statistics and Mathematics from the University of Toronto in 2021.

Code of ethical conduct

Zeusolutions strives to be a forefront in professional statistical work while maintaining the integrity, security and privacy of our customers. We believe in executing our work with proper due diligence and informative research. This is why we believe in the following.

1. Maintaining the trust of our customers as a vital part of our company’s asset.
2. Privacy of our customers is essential in all our statistical work. We do not disclose or seek personal gain from any information given to us by our customers. Any possible interest that can be regarded as influencing the results of our statistical work will be clearly communicated.
3. Appropriate security measures and practices take place amongst our employees to ensure sensitive information is kept safe.

4. Support for all employees. All statistical work performed with Zeusolutions is appropriately credited to all employees involved.
5. Ensure the disclosure of all limitations and assumptions made in relevance to data of any particular work.

We recognize these importance's and responsibilities we ought to uphold to our customers.

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

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

Final advice: KNIT EARLY AND OFTEN!