Avery Schiff

Once you've completed the coding portion of this assignment, use the data you produced to answer the following questions. Data scientists think critically about how to turn data into actionable information – the programming and quantitative pieces are only part of the job. Turn in your answers to these questions as a PDF file in your Github repo along with your code and output. A few sentences for each question is fine.

1. **What are the potential advantages and disadvantages of seniority-based layoff policies, from the company owner's perspective, the senior employee's perspective, and from the perspectives of the employees involved in the DeGraffenreid case? How might these stakeholders weigh these pros and cons differently?**

Seniority-based layoff policies have a variety of advantages and disadvantages that differ based on one's perspective. For company owners, the benefits are that it is a simple process and results in the retention of the most experienced employees. The disadvantage is that the company risks losing young innovative talent and the layoff will produce less cost savings by continuing to employ the senior employees who tend to be the highest-paid employees overall. The company will also face a setback in any progress they have made towards more fair hiring practices and may risk a loss in public opinion and perhaps customer approval. For senior workers, the policies are advantageous because they ensure job security and recognition for their loyalty, but they may then be forced to deal with intense workloads in the absence of younger employees. Newer employees, which in the case of DeGraffenreid v. General Motors were more likely to be Black and women, are at a distinct disadvantage with these policies because discriminatory practices prevented them from being hired early enough to avoid the layoff. Therefore, each group of stakeholders will likely weigh the pros and cons of the layoff policies quite differently, as company owners and senior workers mostly benefit from this layoff technique, while newer and minority employees rarely benefit.

2. **This homework project is based on one of the law cases Dr. Kimberlé Crenshaw analyzed to describe how Black women's lived experiences differ from White women's lived experiences and simultaneously differ from those of Black men. Black women exist in a space where the realities of race and gender overlap. Within the American social structure, it is at times a toxic place where racism and sexism exist simultaneously. Professor Crenshaw named the place "intersectionality" [cite]. Her exploration of intersectionality brought to light complexities that would have remained hidden without listening to and incorporating the perspectives of Black women. In the context of data science, how can we apply this lesson to ensure that we aren't missing deeper narratives within our datasets? In other words, what are some limitations of just looking at data?**

Solely looking at data without acknowledging the context or situation associated with the data and its collection can lead one to overlook the deeper, meaningful narratives within the dataset. Dr. Kimberlé Crenshaw's concept of intersectionality highlights how race and gender create unique experiences when combined that cannot be understood by analyzing

each identity in isolation. In data science, this means that it is important to recognize how the separate analysis of race or gender alone might limit our understanding of the intersection of these identities and how they are affected differently in combination. One can attempt to avoid this by looking at data that combines these two identities. However, even taking that approach may lead to the oversight of certain situational biases that cannot be tracked in a dataset. Overall, there are a multitude of deeper narratives within data that cannot be understood by solely observing raw or even slightly manipulated data.

3. **Data science is a powerful tool for uncovering information about the world, but it often grapples with imperfect data. In the context of this project, what limitations did you encounter regarding the data or analytical methods? Were there noticeable gaps in the data's representation of individuals' identities and experiences? How might these limitations impact the conclusions and insights derived from the data analysis?**

The data taken from this project reveals the issues with failing to represent intersectionality for minority groups like Black women. As shown in the data, Black women and women from other minority groups were eliminated after the layoffs. If one had only looked at gender or race separately, it would be apparent that black people and women were laid off at slightly higher rates as compared to white people and men, but not that Black women were laid off completely. The separated data on gender and race does not confirm that minority women were laid off in higher numbers, suggesting that discrimination is not captured when looking at race or gender alone. This shows a gap in capturing intersecting identities, like being a Black woman, which could lead to biased conclusions. Therefore, this data analysis risks missing the true impact of decisions on marginalized groups and may hide underlying patterns of inequality.

EXTRA CREDIT QUESTIONS - OPTION 1

1. **What were the primary factors you considered when designing your layoff/salary reduction algorithm? Justify why you chose these factors over others.**

While designing the reduce_company_costs algorithm the primary factors I considered were employee salary levels, employee roles, and the target reduction amount. I prioritized salary levels because they have a direct impact on overall payroll expenses. However, certain roles in the company will automatically make more money than other roles, so I ensured that those with the lowest salary in each role were laid off first. This way, the company prioritizes keeping enough people in each role as well as reducing costs by laying off those with lower salaries and less contribution. I also prioritized total payroll costs before layoffs and the target reduction amount to focus on achieving a specific target reduction rather than laying off too few or too many people. This factor was crucial as it set a clear goal for how much the company needed to save, guiding the decision-making process. Therefore, my factors were related solely to salary levels, roles, and target reduction goals to ensure that decisions are based on objective financial criteria rather than subjective opinions or biases.

2. **Reflecting on this exercise, what are your thoughts on using algorithms for impactful decisions? Consider the benefits and risks of algorithmic**

decision-making, and discuss whether you believe it's appropriate to use algorithms in this context.

Using algorithms like the reduce_company_costs function for impactful decisions has benefits including efficiency and objectivity but also has risks like the potential for discrimination and the overlooking of qualitative factors in certain situations. Objectivity allows decisions to be made entirely on quantitative data reducing the possibility of bias and allowing for efficient processing of large data sets. Therefore, algorithms that prioritize this efficiency and objectivity allow for easy application across a wide range of situations or organizations. However, there are risks associated with this type of algorithm, including the potential for overlooking important qualitative factors like performance. Also, these algorithms have the potential to create biases, potentially resulting in discrimination against marginalized groups. They may also overlook hidden bias, such as how bias may affect what appears to be "objective" data. An algorithm based solely on salary may seem objective, but salary can be based on the bias of the bosses and how they assign employees' salaries. Therefore, while algorithms can provide a useful framework for making financial decisions, they should be used cautiously and with human judgment to ensure fairness and consideration of the broader implications within the data.