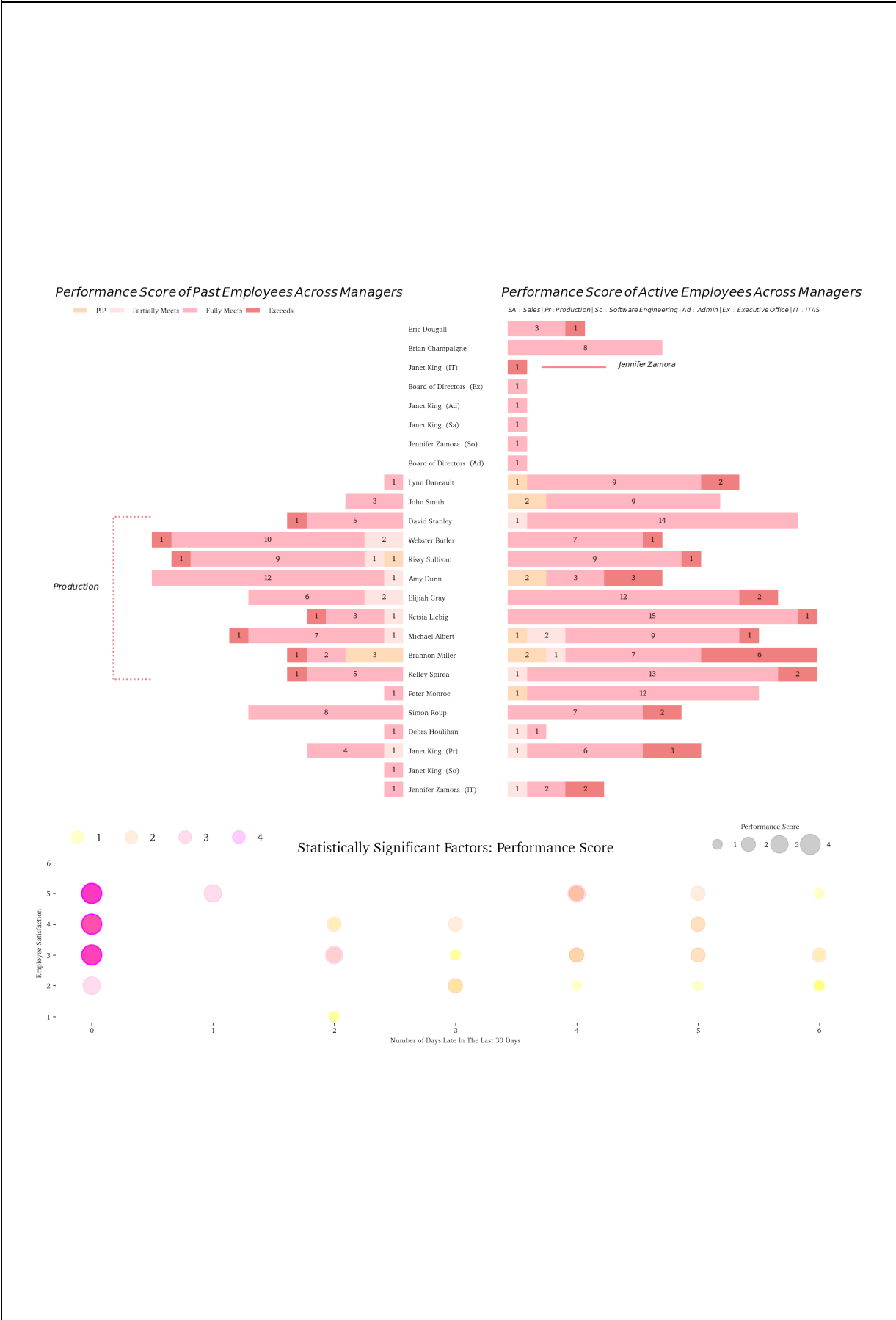


SMM635 — Mid-Term Project Submission Template

!! Do not remove the blank page between the sections!!
— That's the space for the marker's feedback —

Visualization #1

How do employees' performance scores change within and across managers?



What are the key design features of your visualization? (MAX 200 WORDS)

To visualize the changing performance for the employees across their managers, we decided to investigate both the current employees and those who have been terminated. We added the employees in groups based on their manager's names, Department, and their Performance Scores. Therefore, we chose a horizontally stacked bar chart to demonstrate the change, illustrated as "Performance Scores Across Managers" in the graph. Furthermore, the left side of the chart captures the data of employees who are not currently a part of the organization, whilst the one on the right presents the data for current employees.

Further, we created a second graph to explain the relationship between 'Performance Scores' with factors such as 'Days Late Last 30' and 'Employee Satisfaction'. This bubble graph is labelled as 'Statistically Significant Factors: Performance Score'.

Why did you choose the above-mentioned design features? (MAX 200 WORDS)

The reason behind the chosen graph is to gain an idea of the performance of employees in relation to their respective Managers. We wanted to visualise the performance score for all employees within the identified Managers without choosing a central measure, in such a way that we avoid losing out on leverage points or extreme values. The horizontal stacked chart gave us the option of visualizing the data at scale whilst giving us a fair summary of; overall team size of different units, managers with the highest attrition/team size, managers with the lowest attrition/team size, and the highest and lowest performing teams.

The second chart gives us a good direction to explore when investigating Performance Scores and factors which can improve the same. The sizes and colours codes help us bring clarity as we try to identify a correlation among the 3 factors - 'Employee Satisfaction', 'Days Late Last 30' and 'PerfScoreID'.

For these reasons, we can argue that this is a good direction to further examine the variation in Performance Scores.

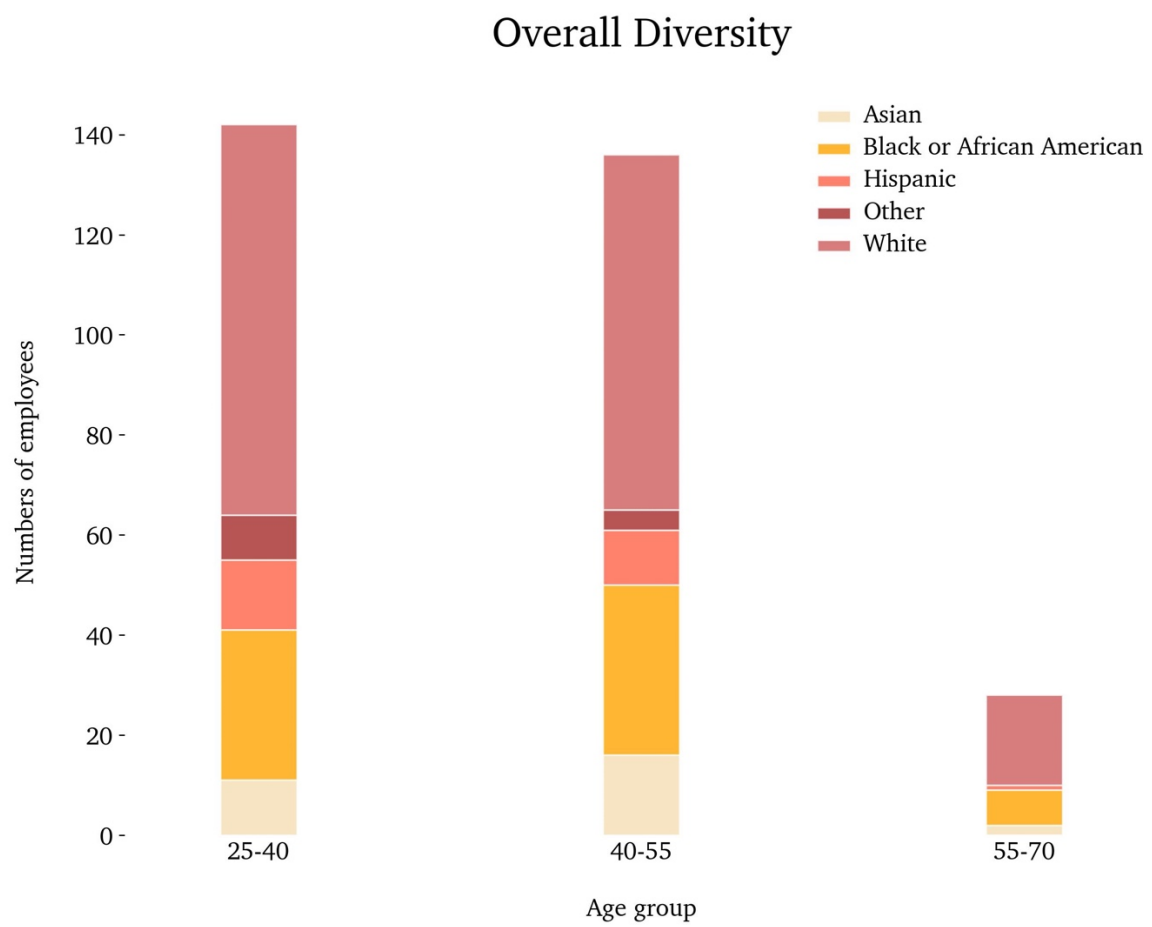
What is the main insight of the chart? (MAX 100 WORDS)

As presented in the graph, out of the 19 Managers, few managers have more than 1/3 of their team exceeding expectations. The Production Department has the highest attrition rate. Further, employees with low 'DaysLateLast30' and high 'Emp Satisfaction' surveys have reported to have better performance score.

Our recommendations to the client would be to identify and implement organisational practices, that improves employee satisfaction and promotes punctuality. Brian Champaign is a manager who has avoided attrition in his team since his arrival in 2016. Further examining Brian's managing style, and employees can help gain insights that improve attrition.

Visualization #2

What is the overall diversity profile of the organization?



What are the key design features of your visualization? (MAX 200 WORDS)

To visualise the diversity profile of the organisation, we created a stacked bar chart focusing on the variables of age and race.

The visualisation represents the 3 biggest age groups in the organisation, only disregarding a tiny minority of people above 70 for the purposes of statistical accuracy. It further defines the proportion of people from different race groups among these 3 age groups in an attempt to illuminate the overall diversity profile of the organisation as we have defined it.

Why did you choose the above-mentioned design features? (MAX 200 WORDS)

We have defined diversity as the employment of people from different racial backgrounds and sexes, as well as the existence of a younger demographic of employees, among the older, established ones.

More than half of the employees are female; hence we have concluded that the organisation is quite diverse in terms of gender. Therefore, on our visualisation, we focus on age and race.

This visualisation is effective in illustrating upon first glance that the organisation is quite diverse when it comes to age, as most employees are between 25 and 50 years old, and the majority of employees are White. They represent almost, or more than half of each age group. The second most represented race group in the organisation is Black or African American.

What is the main insight of the chart? (MAX 100 WORDS)

The main insight of the graph is that each age group, employees from all 5 identified race groups are quite similarly represented proportionally.

Although the company employs more of the white race than other race groups, we cannot conclude that it lacks racial diversity as a whole, as we would have to make assumptions about the culture of the company, and whether it fosters diversity other than quantifying it.

Our recommendations would include reshaping the company's policies to accommodate for a more diverse work culture, in addition to recruiting on specific demographics to increase diversity.

Visualization #3



What are the key design features of your visualization? (MAX 200 WORDS)

We used a scatter plot to investigate how employees from different race groups are recruited through the nine types of recruitment sources used by the organisation. Larger sized dots represent larger groups of employees recruited. The visualisation intends to illustrate the count of what we have defined as employees 'from a diverse background' and the recruitment sources which they have used, in addition to White employees and their most efficient recruitment sources. This was created in order to get an understanding of where the company can focus on with its diversity recruiting strategy.

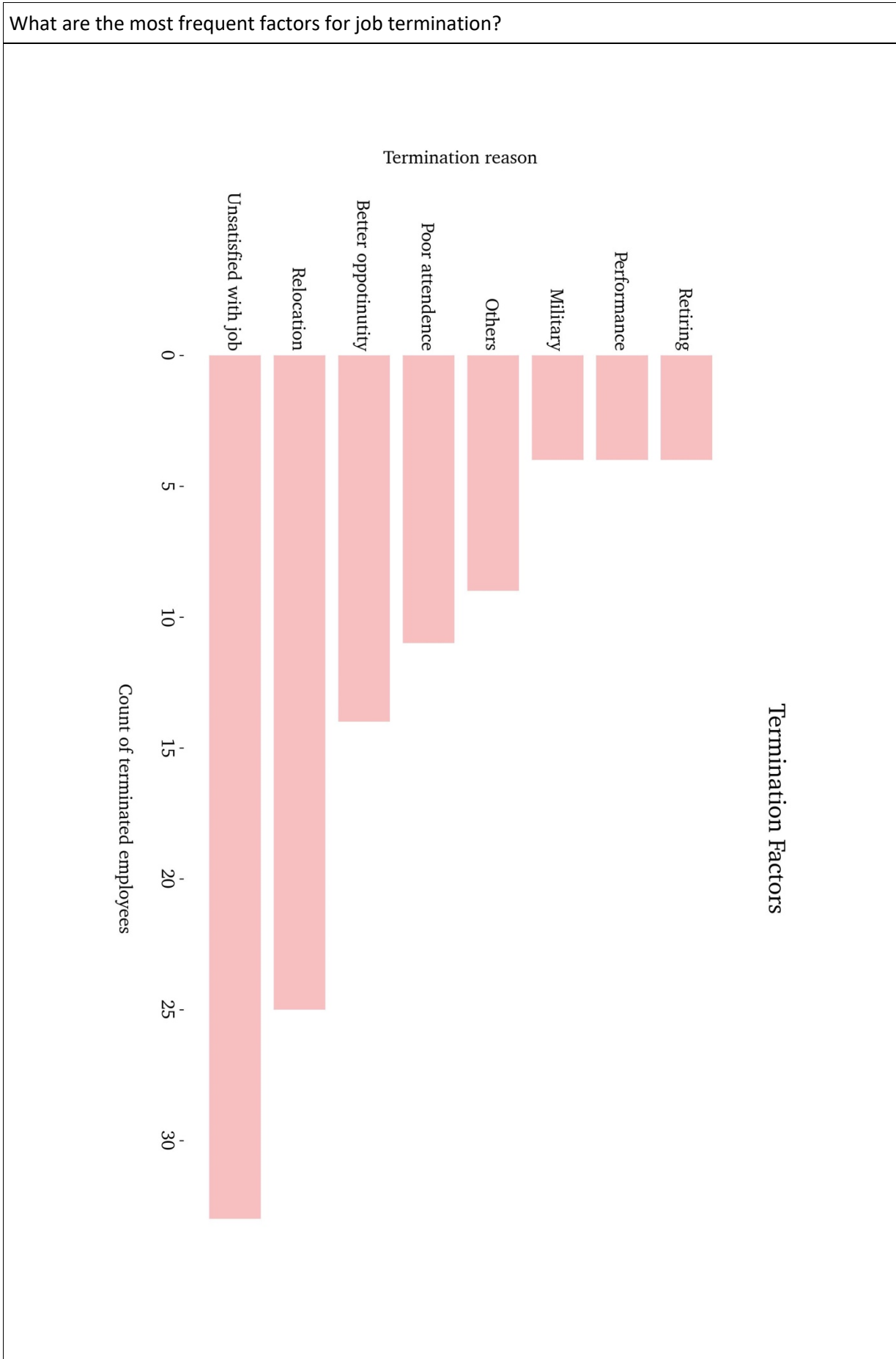
Why did you choose the above-mentioned design features? (MAX 200 WORDS)

We have chosen this design feature to get an insight of the recruiting strategy and whether diversity candidates have as much presence on recruiting sources as White candidates. In other words, the plot illustrates inclusivity in the different recruiting sources regarding the 5 race groups we have identified. Overall, "indeed and LinkedIn" is shown to be the greatest recruitment source. However, it is important to take into consideration that the "white" race has an advantage when applying for jobs online in comparison to all these other races reported in the company.

What is the main insight of the chart? (MAX 100 WORDS)

There is a clear difference in job recruitment between the "white" race compared to the other races. The "white" race is shown to have less of a difficulty finding employment through online recourses such as LinkedIn, Indeed and Google Searches. In contrast, "Black or African Americans" and "Hispanics" managed to find more employment through "Diversity Job Fair" in comparison to the online recourses. For these reasons, we can conclude that there is clearly discrimination between the races in regard to "non-white" races finding employment compared to the "white" race.

Visualization #4



What are the key design features of your visualization? (MAX 200 WORDS)

There were multiple termination reasons stated in the dataset, however, we decided to focus on the most common ones, which were eight in total. The chart clearly states the different reasons for termination, in addition to how many employees that fits this description. The design shows a clear comparison between all the termination factors. Noticeably, the main reason for job termination at this company is unsatisfied employees.

Why did you choose the above-mentioned design features? (MAX 200 WORDS)

The chart attempts to give us a clear overview of the main reasons as to why employees who are no longer active in the company decided to terminate. It is important to understand the causes of termination, especially when these have to do with employees being unsatisfied.

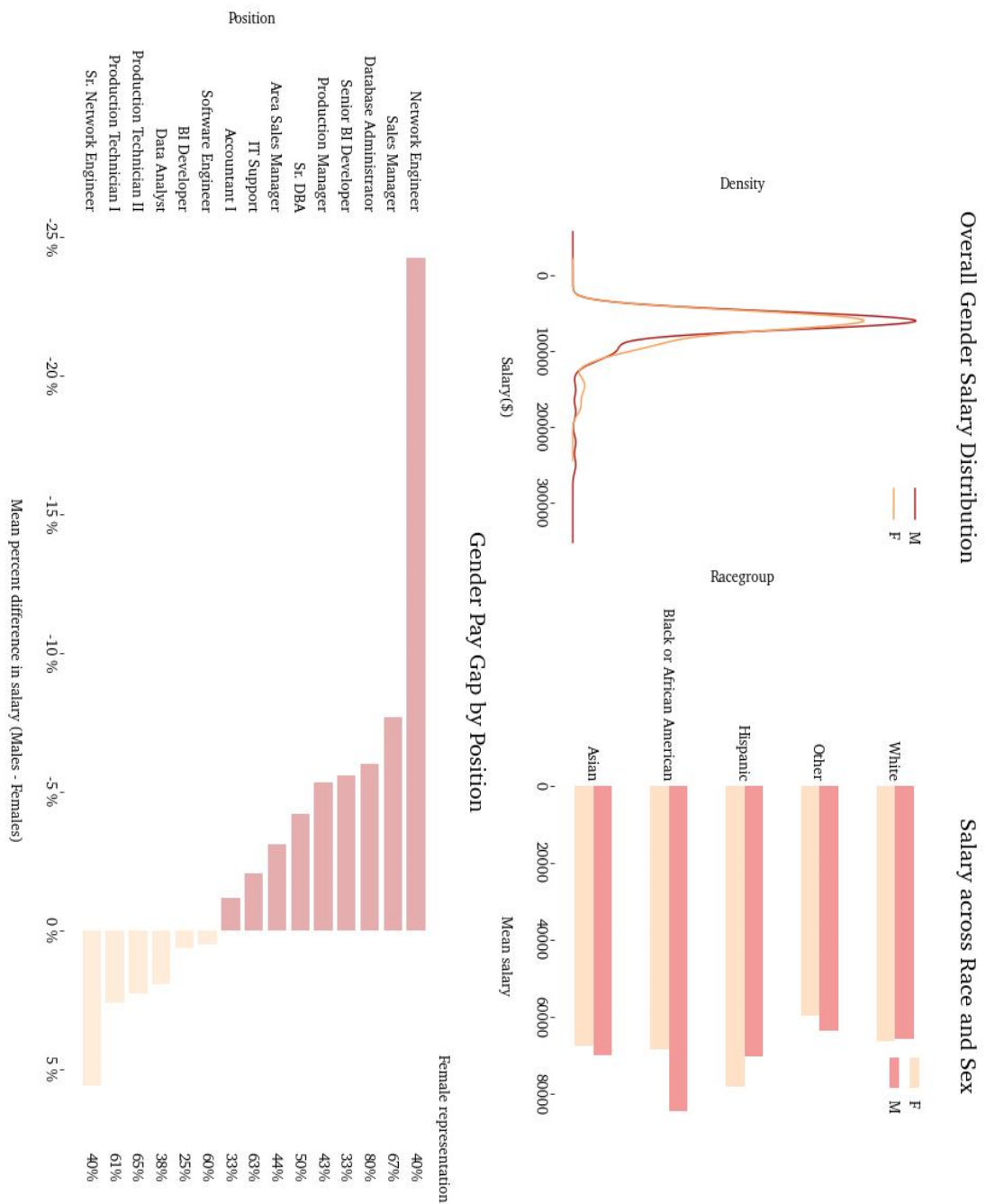
The chart is designed to illustrate the most common termination reason as the bottom bar. The ascending order with which the plot has been designed is intended to give a sense of priority for when it comes to the company's turnover policies.

What is the main insight of the chart? (MAX 100 WORDS)

The graph illustrates that overall, the most common termination reason amongst employees is being unsatisfied with the job. Some termination reasons such as joining the military and relocating, are outside of the organisation's control, as employees make those decisions considering factors that are external to the business. However, being unsatisfied with the job should illuminate a possible issue that the company may have to address as it is the most common reason for terminating.

Visualization #5

Are there areas of the company where pay is not equitable?



What are the key design features of your visualization? (MAX 158/200 WORDS)

In order to investigate whether there exist any areas where pay is inequitable, we decided to create a dashboard consisting of three visualisations, focusing on areas of gender and race. Empirically, gender and race are the most common discriminatory factors when it comes to inequitable pay among employees.

Firstly, we plotted the density distributions of female vs male salaries across the whole organisation, a graph labelled "Overall Salary Distribution".

The second graph is a positive/negative horizontal bar plot, labelled "Gender Pay Gap by Position". It is intended to dissect further into possible gendered salary discrepancies across positions in the organisation, visualising both female representation, as well as the mean percentage difference between female and male salaries for each position.

Finally, our dashboard is completed with a horizontal bar chart, visualising race and gender against salary, comparing the average salary of different race groups and their gender, labelled "Salary Across race and Sex".

Why did you choose the above-mentioned design features? (MAX 183/200 WORDS)

The density plot facilitates the comparison between female and male employees and their salaries. Looking at the overall distribution would help us identify whether a female employee is as likely to earn a particular salary as a male employee. Pay discrepancies in the middle of the distribution would cause a more significant concern, rather than discrepancies at the tails.

The "Gender Pay Gap by Position" plot considers the variable of position in the organisation in the investigation of the gender pay gap, in an attempt to compare salaries between female and male employees in the same positions, e.g., network engineers. The chart is intended to showcase whether there exist positions where although female and male employees are held to the same expectations and responsibilities, pay is inequitable between them.

The "Salary Across Race and Sex" plot is designed to consider the variable of race against salary and along sex, in an attempt to explore whether for example a black woman for example, would be discriminated against a white woman, who might also be in turn discriminated against a male colleague. It provides an insight as to whether both race and sex are discriminatory factors.

What is the main insight of the chart? (MAX 100 WORDS)

The "Overall Gender Salary Distribution" plot indicates that pay discrepancies in positions salaried around \$75,000 - \$115,000 and \$125,000-\$180,000 in favor of males.

The "Gender Pay Gap by Position" plot shows that the position of Network Engineer holds the biggest mean percentage difference in salaries between females and males. Although 80% of the employees at the position of "Database Administrator" are female, they earn on average 7.5% less than their male counterparts.

Lastly, the "Salary Across Race and Sex" plot illuminates that except for "Hispanic" and "White", for the rest of the Race groups, males earn more, on average.

Average salary for White people is almost identical for females and males, while the biggest discrepancy can be identified for Black or African American employees.

Overall, our recommendations to the client would be to re-examine salaries in the aforementioned areas where there seem to be discriminatory pay discrepancies. These would either have to be confirmed, and firm pay policies revised, or disconfirmed if management can attribute them to other professional factors.