The New Norm? Exploring Job Satisfaction & Retention in Remote Work Environments

BTMA 431 Fall 2024 (L02)

Professor Duy Dao

December 11, 2024

Bryce Corrin

Darren Keilty

Executive Summary

The shift to remote work, largely driven by the COVID-19 pandemic and the need for flexibility, has changed today's work environments. Our study examines the impact of remote work on employee satisfaction and retention, and we will be focusing on key factors that drive satisfaction, and exploring the relationship between remote work and employee retention. We have learned that pay, benefits, and workplace culture are the top drivers of satisfaction across all seniority levels and age groups. Senior employees prioritize financial security and comprehensive perks, such as healthcare and retirement plans, while junior employees emphasize financial growth. Mid-level employees value stability and workplace culture, highlighting their focus on career development and positive work environments. Younger employees (aged 20–30) prioritize financial stability, while older employees (40+) focus more on benefits like healthcare.

Remote work flexibility emerged as the most significant predictor of employee retention, surpassing other influential factors like job satisfaction and work-life balance. Employees with remote work options were significantly more likely to stay with their organizations, with retention rates reaching 75% for remote workers compared to 47% for on-site employees. Other key retention factors include work-life balance and job role. For organizations, the findings underscore the importance of designing tailored compensation packages that address the needs of different demographic groups, such as competitive pay for younger employees and healthcare benefits for older ones. Supporting work-life balance through flexible schedules and wellness initiatives further contributes to a positive work environment. Future research will be looking into 4 day work weeks potentially, as studies have shown that people feel less stressed and find

more time for family given that extra day. Overall, our research has shown that remote work plays a critical role in employee satisfaction and retention, offering organizations the value in potential channels of work environments going forward.

Table of Contents

Introduction

The rise of remote work has truly reshaped modern work environments, largely due to COVID-19 (Statistics Canada, 2024), but also to the flexibility and various benefits it provides employees. Over the last few years, we have seen the reduction of the traditional office norms, and a shift to hybrid/flexible work arrangements, especially in the tech industry. This shift has proven beneficial for both employers and employees, as studies have shown that it has enhanced productivity for employees, not to mention help with their commute and flow of traffic in busy cities where individuals work centrally in downtown locations. Additionally, remote work supports work-life balance and inclusivity by accommodating diverse needs and lifestyles.

Satisfied employees are more engaged, motivated, and committed to their roles, leading to higher performance levels when they are given the flexibility to work remotely. Organizations with satisfied employees tend to build stronger workplace cultures, which enhance their ability to attract top talent, which can be due to word of mouth as well for people interested in joining companies that support remote work. In competitive markets, prioritizing job satisfaction and retention not only boosts organizational efficiency, but it is also viewed as a job perk recently because many people value the ability to work remotely and are willing to take a pay cut to do so. In this project, we will be further exploring employee job satisfaction and retention in remote work environments.

Methodology

Question #1 of our study honed in on the most common factors that influence satisfaction among remote workers. Based on these identified factors, we looked into how these vary in importance among different seniority levels and age groups. To extract insights, we analyzed

employee reviews from Indeed, focusing on remote work experiences across prominent organizations (i.e. Amazon, Apple) and industries. The dataset included over 10,000 reviews based on 11 distinct sets of reviews, offering a broad perspective on employee satisfaction factors. To prepare the data, we organized and cleaned the text to ensure consistency, removed duplicates, and excluded irrelevant or incomplete reviews. Using specialized tools, we analyzed the frequency of key satisfaction themes and then grouped responses by seniority levels (Junior, Mid-Level, Senior) and age groups (20-30, 30-40, 40+). Key factors like pay, benefits, and workplace culture were identified by examining how often they appeared in the reviews. These insights helped us understand what matters most to employees and how priorities differ across demographics. While the findings are revealing, there are some limitations. The analysis relied on publicly available reviews, which may not fully represent the broader remote workforce, and the self-reported nature of the data could introduce bias. Still, the results offer valuable guidance for understanding job satisfaction in remote work settings.

The analysis for Question 2 and its follow-up integrated data from the Remote Working Survey, which collected responses from remote workers in 2020 and 2021. This survey included information about respondents' industries, job roles, and satisfaction levels with remote work policies, providing valuable insights into trends across different sectors. The combined dataset featured responses from thousands of participants, offering a robust foundation for cross-industry comparisons. To ensure the reliability of the analysis, data preparation involved standardizing column names, handling missing values, and filtering out incomplete or ambiguous responses. Satisfaction ratings, reported on a scale ranging from strongly negative to strongly positive, were converted into numeric values to facilitate comparison. Only industries with sufficient data and

average satisfaction scores above 4.0 were included to maintain analytical rigor and focus on high-satisfaction contexts.

The analytical approach centered on summarizing satisfaction trends across industries and identifying the top job roles in sectors with high average satisfaction. Average satisfaction scores were calculated for both industries and job roles, and visualizations were used to display these trends effectively. Satisfaction was measured using survey questions related to employees' feelings about their employers' remote work policies and the overall impact of remote work on their experience. While the survey provided valuable insights, there were limitations. The data reflects self-reported responses, which may be subject to individual biases and may not fully represent the diversity of remote workers across all sectors. Additionally, missing data in some variables, such as collaboration and preferred remote work time, limited their inclusion in the analysis. Finally, the timing of the survey during the COVID-19 pandemic may have influenced responses, reflecting temporary conditions rather than long-term trends.

In our third question, we focused on employee retention in remote work environments, utilizing a synthetic employee attrition dataset sourced using the Kaggle API. The dataset, designed for HR analytics, included detailed records of over 70,000 employees, ensuring a robust basis for our analysis, with the inclusion of factors such as remote work, job satisfaction, work-life balance, performance rating, and attrition. To ensure accuracy in our analysis, data preparation methods included the exclusion of rows with missing values, removing duplicate rows to prevent redundancy, and utilizing testing and training sets for model development. To perform analysis on the categorical features of our dataset, these were encoded as factors, and due to their ordinal nature, reference levels were set to facilitate meaningful interpretation of regression results (e.g., "Medium" for Job Satisfaction, "Mid" for Job Level).

In our analysis, we conducted various tests to investigate the relationship between remote work and retention, and then further explored how other factors compare to remote work in predicting retention. First, we utilized a chi-squared test to assess a potential relationship between remote work and attrition, where a significant result would indicate an association between these two variables. To investigate and quantify this relationship further, a logistic regression was fit with remote work as the sole predictor. This model served as a baseline to isolate the effect of remote work options on employee retention. A proportion table of attrition outcomes and remote work status was then constructed and graphed to visualize this relationship. Subsequently, a multivariate logistic regression model was fit to incorporate additional predictors including job satisfaction, work-life balance, job level, and monthly income, among others. This allowed for a comprehensive examination of the marginal impact of other factors in predicting attrition compared to remote work. After checking our model for multicollinearity fitting issues, we evaluated the significance and impact of each predictor. Significant estimates were then visualized through a bar plot, where the relative importance of each predictor could be seen by the magnitude and direction of their coefficient estimates. While the analysis provided comprehensive insight, we must be aware of the synthetic nature of the data we utilized. Although designed to simulate realistic HR scenarios, the data may not fully capture specific nuances in employee behavior or organizational dynamics. Despite this, our analysis provides a strong framework for understanding the critical elements that impact employee retention, providing insights that organizations can leverage to reduce attrition rates.

Results

For our first question, the primary analysis of remote work reviews identifies pay, benefits, and support as the most prominent factors influencing job satisfaction, with pay

standing out as the leading driver across reviews. Other factors such as culture, flexibility, and work-life balance also play a role, indicating that employees value both tangible rewards and qualitative aspects of their work environment. Less frequently mentioned factors, like collaboration and career advancement, suggest these elements, while relevant, may not universally drive satisfaction compared to more dominant factors.

As mentioned in the Methodology section, our follow-up question investigates how the aforementioned satisfaction factors differ in significance among different levels of seniority and age groups. Regarding seniority levels, senior employees prioritized pay as the most critical factor, consistent with their expectations for financial compensation in high-responsibility roles. Benefits ranked second, underscoring the importance of comprehensive perks like retirement plans and healthcare. Culture, support, and flexibility held relatively equal importance, reflecting a holistic view of satisfaction that extends beyond financial incentives. However, flexibility and collaboration received fewer responses, suggesting that senior employees may already expect flexibility or focus less on teamwork due to leadership responsibilities. Mid-level employees had the highest overall number of responses, emphasizing their substantial representation in the workforce. Pay overwhelmingly emerged as the top factor, even more so than in other seniority levels. Benefits and culture followed, highlighting this group's focus on career stability and a positive workplace environment. Similar to senior employees, flexibility and collaboration were less frequently cited, indicating these factors may be secondary to financial and professional growth. For junior employees, pay was the most cited satisfaction factor, reflecting their early-career focus on financial growth. Benefits and culture were tied as secondary factors, demonstrating their need for a supportive environment and tangible employment perks. Flexibility and collaboration were again less emphasized, potentially due to the structured nature

of entry-level roles, where these aspects are predefined. Across all seniority levels, flexibility, collaboration, and work-life balance had notably low responses. This may reflect that these factors are either taken for granted in remote work or overlap in employee perceptions, particularly between work-life balance and flexibility.

In terms of age groups, the 20-30 age group ranked pay as the most important satisfaction factor, followed by benefits and culture, which were closely tied. This reflects younger employees' focus on financial stability and supportive workplace environments that foster growth and belonging. The lower emphasis on flexibility and collaboration suggests that these employees are more focused on foundational financial and professional needs than lifestyle-driven factors. Similarly, the 30-40 group also ranked pay as the highest factor, with benefits and culture following closely. This group's responses highlight the importance of financial security and a positive work environment as they advance in their careers while balancing personal and professional responsibilities. The 40+ group had fewer responses overall, limiting the diversity of observed patterns. However, pay and benefits were the most frequently cited factors, with a slightly higher emphasis on benefits, likely reflecting the older demographic's focus on comprehensive perks like healthcare and retirement plans. This group's lower representation may indicate that remote work satisfaction is less frequently expressed or varies less for employees nearing the latter stages of their careers. Overall, the analysis for question 1 consistently highlights pay as the top satisfaction driver across all seniority levels and age groups, with benefits and culture also playing important roles. While flexibility, collaboration, and work-life balance had lower representations, their potential overlap (e.g., between flexibility and work-life balance) may explain their limited presence. For organizations,

this suggests prioritizing financial and tangible benefits while recognizing that cultural factors may hold varying levels of importance depending on an employee's career stage or age.

Question 2 analyzed which industries respond the most favorably to remote work environments. Industries such as Professional, Scientific, and Technical Services, Information Media and Telecommunications, Arts and Recreation Services, Accommodation and Food Services, Electricity, Gas, Water and Waste Services, and Administrative and Support Services emerged as high-satisfaction industries (see Figure 4 in Appendix B), with average satisfaction scores exceeding the threshold of 4 on a 5-point scale calculated using two different satisfaction metrics. This indicates that these industries have effectively fostered positive remote work experiences through policies and practices that align with employee preferences.

Among these high-satisfaction industries, specific job roles stood out for their particularly high satisfaction levels (see Figure 5 in Appendix B). For instance, in the Accommodation and Food Services sector, Specialist Managers exhibited the highest satisfaction levels, likely due to the autonomy and leadership opportunities afforded by their roles. Similarly, in both the Arts and Recreation Services and Professional, Scientific, and Technical Services industries, Clerical and Office Support Workers stood out as the most satisfied, reflecting the adaptability of support roles in creative industries to remote work environments. Within Administrative and Support Services, Health and Welfare Support Workers reported the highest satisfaction, highlighting effective remote work policies that cater to service-oriented roles. In Electricity, Gas, Water and Waste Services, Inquiry Clerks and Receptionists were most content with working from home, benefiting from technological tools that facilitate seamless remote workflows. Lastly, in Information Media and Telecommunications, Carers and Aides emerged as the top role by satisfaction, showcasing the adaptability of support-based positions to remote settings. These

findings underline the importance of tailoring remote work policies to specific roles within industries, addressing unique job requirements to enhance satisfaction. While satisfaction was generally high in these industries, the results also underscored variations within and across roles. For example, Administrative and Support Services had a broader range of job roles with varying satisfaction, highlighting potential disparities in how remote work policies are perceived by employees in different capacities. Overall, the findings emphasize that high-satisfaction industries often provide clear policies, flexibility, and support that cater to the unique needs of specific job roles, making remote work a positive experience for their workforce. The results of this question reveal the importance of tailoring remote work policies to individual roles to maximize satisfaction.

In our third question, we explored the relationship between remote work offerings and retention rates. A chi-squared test of independence showed a statistically significant relationship between remote work and attrition ($p\text{-value} < 0.05$), indicating that remote work flexibility is associated with retention rates. Further verifying the existence of this relationship, a logistic regression was fit to predict the likelihood of attrition based on remote work offerings. The model coefficients revealed a significant positive relationship between remote work and retention. This means that if employees are able to work remotely, they are more likely to stay with their organization. This described relationship can be observed in Figure 6 (see Appendix B) illustrating the proportion of attrition outcomes by remote work status. As seen in the figure, of those within our sample who did not work remotely, 52% left their company, while 47% stayed. And as predicted, we see that of those able to work remotely, 25% left, while 75% stayed. These findings suggest that remote work flexibility may serve as a critical factor in improving

employee retention. Employers and managers could consider implementing remote work options into their existing organizational structures to improve retention rates.

To dig deeper into the above question, we sought to understand the marginal impact of remote work flexibility compared to other factors that have a known influence on retention rates. Variables such as work-life balance, job level, and job satisfaction were added to a logistic regression model, where significant predictors were identified based on p-values ($p \leq 0.05$). Figure 7 (see Appendix B) graphs the coefficient estimates of significant predictors, illustrating the overall influence each of these factors has on employee retention. Positive predictor coefficients (in blue) indicate factors that increase the likelihood of an employee staying at their organization, while negative coefficients (in red) represent factors that may contribute to an employee's decision to leave. Predictors with large magnitudes are considered to have a greater impact on retention, and we observe that the most influential factors in predicting attrition are remote work, work-life balance, and job role. Employees with remote work options and senior-level job roles are more likely to stay at their company, while employees with poor work-life balance and entry-level roles are more likely to leave. Since remote work options emerged as the largest predictor of retention, this reinforces our view that managers should consider implementing remote work programs if looking to retain employees.

Discussion

Through our analyses, it is evident that remote work has a substantial impact on employee satisfaction and retention. Across all demographics, pay consistently emerged as the most critical factor, followed by benefits and workplace culture. Flexibility and work-life balance, while important, were less emphasized, possibly because they are considered inherent to

remote work. High-satisfaction industries, including Professional, Scientific, and Technical Services and Information Media and Telecommunications, demonstrated that clear remote work policies and adaptable roles significantly enhance employee satisfaction. Retention analysis revealed that remote work is the most significant predictor of employee retention, surpassing other influential factors like work-life balance and job satisfaction.

These findings highlight several actionable steps for decision-makers. Benefits should be tailored to specific demographics: younger employees prioritize financial stability through competitive pay, mid-career professionals value comprehensive benefits and supportive workplace cultures, and senior employees focus on retirement and healthcare benefits. Organizations should implement clear, role-specific remote work policies that address unique job requirements while promoting autonomy and leadership development in managerial roles. Work-life balance should be supported through flexible schedules and wellness initiatives. Finally, integrating remote work options into retention strategies can reduce attrition and strengthen employer branding to attract top talent.

This study has some limitations that should be acknowledged. The reliance on self-reported data introduces potential bias, and the focus on specific industries and age groups may limit the findings' generalizability. Additionally, the synthetic data used for retention analysis, while useful, may not fully capture real-world organizational dynamics. Future research should include longitudinal studies to track trends over time, expand to global samples, and explore additional satisfaction factors, such as mental health support and career advancement opportunities. By addressing these areas, future studies can deepen our understanding of remote work's transformative role in employee satisfaction and retention.

References

Statistics Canada. (2024, January 18). *Working from home in Canada*.

https://www150.statcan.gc.ca/n1/daily-quotidien/240118/dq240118c-eng.htm

Appendix A: Technical Details

Q1: What are the most common factors affecting job satisfaction among remote workers? / Follow-Up: How do specific satisfaction factors vary in importance among different age groups or seniority levels?

The data for this analysis was collected from employee reviews on Indeed, focusing on remote work experiences across various organizations. HTML files containing reviews were manually downloaded and parsed using the rvest package in R to extract review text. The raw data underwent a preparation process to ensure consistency and usability. This included converting text to lowercase, removing duplicates, and filtering out irrelevant or incomplete reviews. The tidytext package was used to tokenize the text into individual words for further analysis.

Key satisfaction factors such as pay, benefits, culture, and flexibility were identified through a predefined list of keywords and phrases. The frequency of these keywords was calculated using the str_count function in R. Reviews were grouped into demographic categories based on seniority levels and age groups. Seniority was determined by pattern matching for terms like "manager," "intern," and "senior," while age group categorizations relied on context phrases like "early career" or "seasoned professional." For reviews with ambiguous classifications, random sampling was applied to assign them to appropriate groups. Data manipulation and summarization were conducted using the dplyr package.

Visualizations were created using the ggplot2 package for both the primary question and follow-up question. The primary question produced a line chart that highlighted the most common factors that affect remote worker satisfaction, in ascending order (see Figure 1 in

Appendix B). For the follow-up question, stack bar charts were generated to illustrate satisfaction factor frequencies by seniority level and age group (see Figures 2 and 3 respectively in Appendix B). These visualizations included customizations for readability, such as sorted factors, clear labeling, and flipped axes. The analysis relied on a range of R packages, including rvest for web scraping, tidytext for text mining, dplyr for data manipulation, stringr for pattern matching, and ggplot2 for visualization.

Despite the robustness of the methodology, there were limitations. The data was limited to publicly available reviews, which may not fully represent the broader remote workforce. Reviews are inherently subjective, introducing potential biases. Additionally, some reviews lacked clear indicators for seniority or age, requiring random assignment for classification. Nonetheless, these methods provide a systematic framework for analyzing employee satisfaction data, balancing rigor with the constraints of publicly available information.

Q2: How does remote work satisfaction vary across industries? / Follow-Up: Within high-satisfaction industries, what specific job roles report the highest satisfaction levels?

The technical details for the analysis of Question 2 and its follow-up relied on the Remote Working Survey datasets from 2020 and 2021. The datasets were retrieved using the Kaggle API, ensuring direct and consistent access to the files. Once downloaded, the datasets were unzipped and imported into R for analysis. The column names were cleaned using the janitor package to remove inconsistencies, and missing or ambiguous responses in key variables were filtered out to maintain the integrity of the analysis. Relevant columns included "Satisfaction with Employer's Remote Working Policy," "Sentiment Toward Remote Working (Positive/Negative)," "Collaboration While Working Remotely," and "Preferred Remote Work

Time." Satisfaction-related ratings, originally provided in categorical terms (e.g., "Strongly Agree," "Somewhat Positive"), were transformed into numeric scales ranging from 1 to 5 using the dplyr::mutate() function, enabling quantitative aggregation and comparison.

The analytical process involved grouping data by industry to calculate average satisfaction scores for each category. This was achieved using the dplyr::group_by() and summarize() functions, allowing for precise identification of industries with satisfaction scores exceeding a threshold of 4.0. For the follow-up question, data for job roles within high-satisfaction industries was further analyzed. The top job role in each high-satisfaction industry was identified by selecting the job role with the highest average satisfaction score using the dplyr::slice_max() function. To address ties in satisfaction scores, the first occurrence of the maximum score was selected, ensuring only one role was highlighted per industry.

Visualizations were created to present average satisfaction scores by industry and the top job roles within high-satisfaction industries. Data was reordered to emphasize trends, and bar charts were used to ensure clarity in communicating the results. Limitations in the analysis included missing data for certain variables, such as collaboration and preferred remote work time, which resulted in NA values in summary statistics. These gaps restricted the interpretability of certain trends. Additionally, the reliance on numeric conversions for satisfaction ratings may have simplified complex employee sentiments. The analysis used R with key packages including dplyr, tidyverse, and janitor to handle data transformation, grouping, and cleaning, ensuring a robust and reproducible workflow.

Q3: Is there a measurable correlation between remote work flexibility and employee retention rates? / Follow-Up: What other factors could influence employee attrition, and how do they compare to the impact of remote work flexibility?

The dataset was again retrieved using the Kaggle API, this way, file access is reproducible and consistent. The downloaded file was then unzipped into testing and training sets, for regression analysis. The removal of rows with missing values was done using the na.omit() function while checking for duplicate rows used the duplicated() function. The as.factor() function was used to convert all categorical columns in our dataset to factors, and base levels were changed using the relevel() function. The base levels of these factors were changed to the more moderate responses in ordinal categories to facilitate the analysis of more extreme factors in our follow-up.

Each of our regression models was run with both testing and training data, where we saw similar coefficient estimates and significance levels between regressions. This ensured that when running multivariate regressions, we could detect overfitting within our model. Multicollinearity was assessed using variance inflation factor (VIF) scores calculated using the vif() function. The results confirmed no multicollinearity issues with all VIF scores below 3 within our logistic models.

Our proportional bar chart was created using ggplot2 to display attrition outcomes by remote work status. The stacked bar chart shows the proportions of remote workers that either stayed with or left their company. Proportions were calculated with the prop.table() function using the remote work and attrition columns as parameters. Our bar chart of significant factors in predicting attrition was created through the analysis of our multivariate regression output.

Predictors that had a significant p-value ($p < 0.05$) were included in the visualization, and their coefficient estimates were graphed in ascending order. Including all significant features in the resulting visual gives a better representation of the relative impact of specific factors.

Appendix B: Visualizations

Figure 1: Key Factors Affecting Remote Work Satisfaction (Q1 primary)

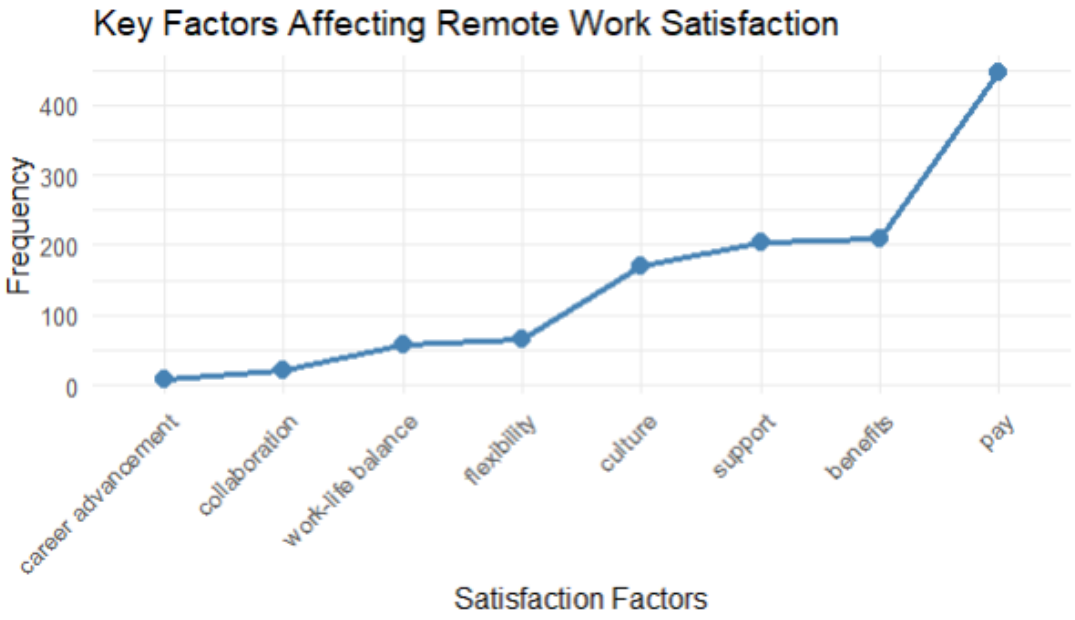

Figure 2: Remote Work Satisfaction Factors by Seniority Level (Q1 follow-up)

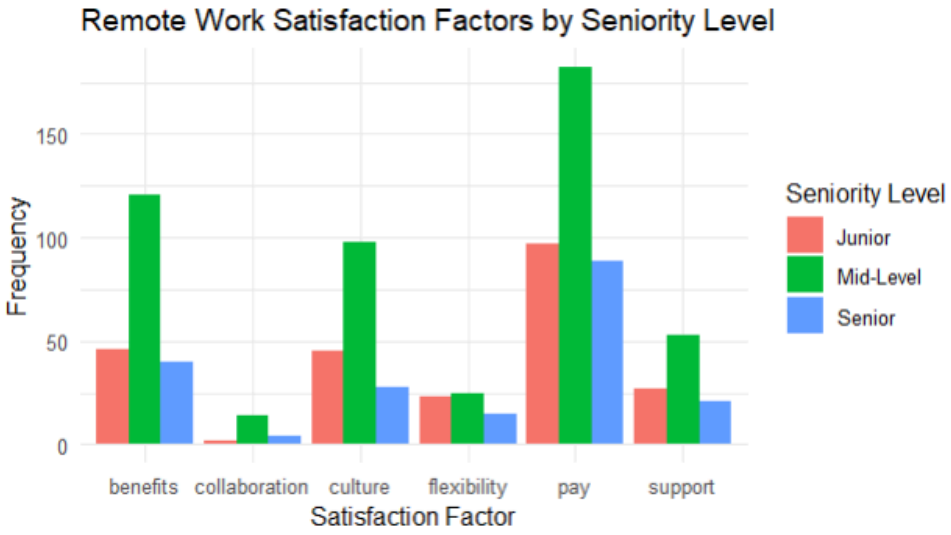

Figure 3: Remote Work Satisfaction Factors by Age Group (Q1 follow-up)

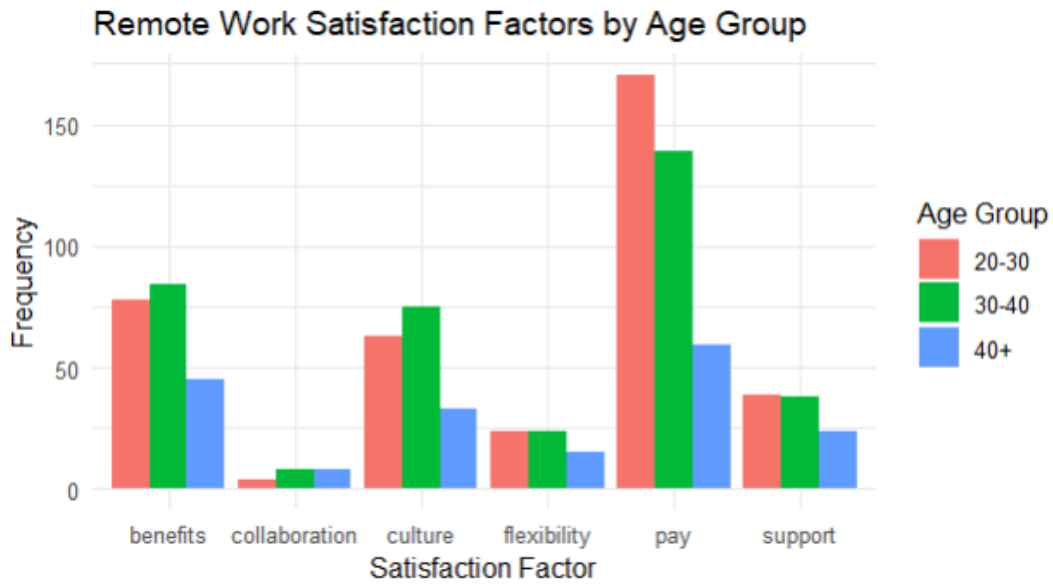

Figure 4: High Remote Work Satisfaction by Industry (Q2 primary)

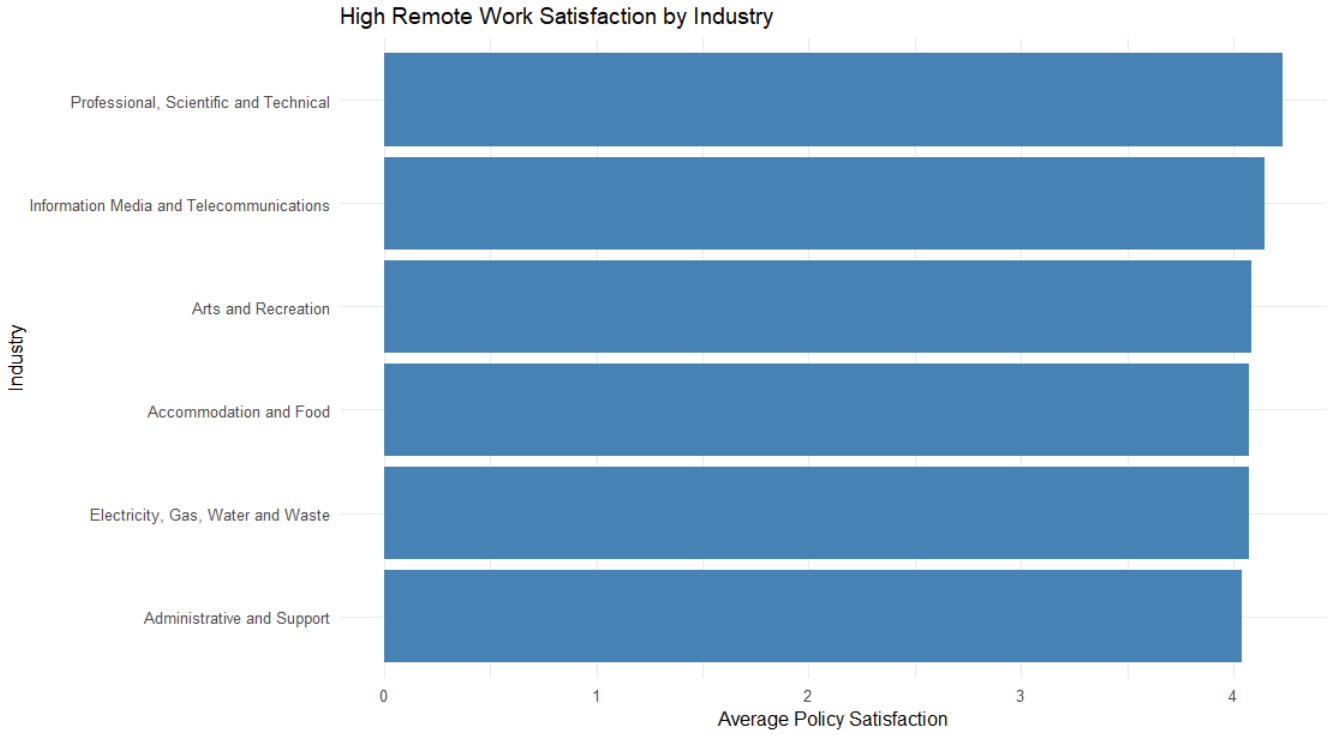

Figure 5: Top Job Role by Satisfaction in Each High-Satisfaction Industry (Q2 follow-up)

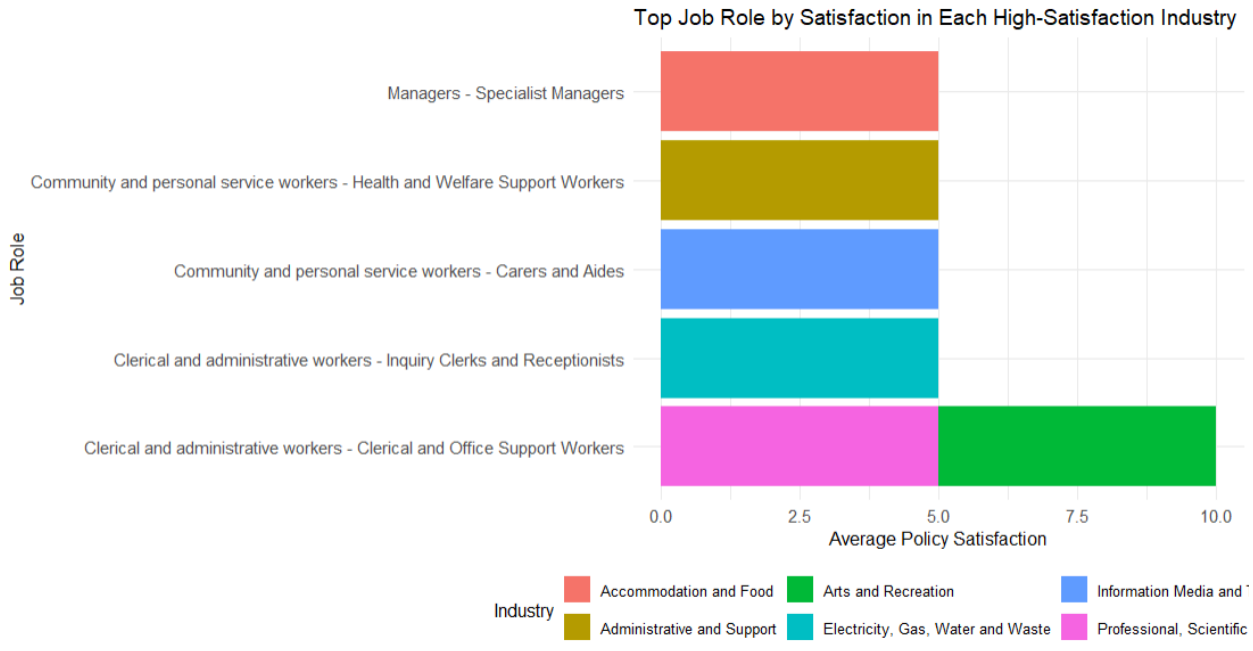

Figure 6: Attrition by Remote Work (Q3 primary)

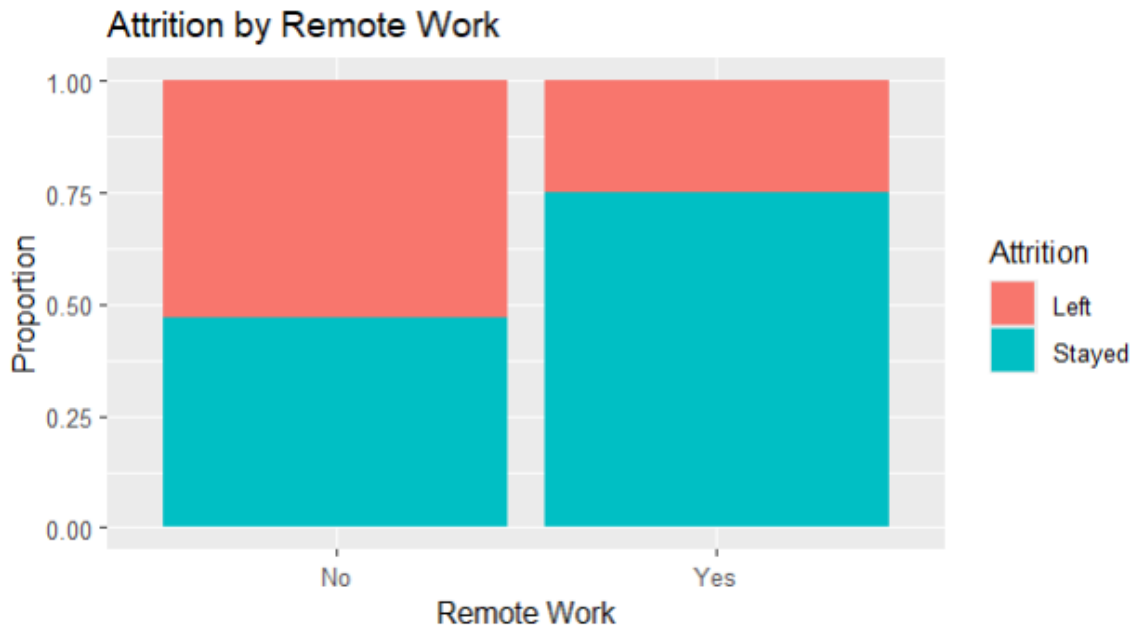

Figure 7: Significant Factors in Predicting Attrition (Q3 follow-up); interactive plotly graph