Work Life Balance Capstone

Collin Hearn, Michael McCarthy, Zach Corlazzoli

Department of Business and Technology, Seton Hill University

SDT 100 01 Introductory Data Science and Business Analytics

Dr. Mirwais Parsa, Ph.D.

1. **Executive Summary – one paragraph summarizing findings and implications.**

This analysis will give an in-depth insight into how several factors influence the modern U.S workforce in terms of work attitude and job satisfaction. Some of these factors include age, compensation, work type, and job experience. After cleaning and wrangling the dataset of several thousand respondents, we were able to conclude that younger workers showed the strongest relative preference for in-person work. Older workers demonstrated a clear and increasing preference to work remote or hybrid environment. To our surprise, salary and job satisfaction were not significantly influenced by work preference; however, descriptive statistics showed developers across all work types tended to be the most satisfied at higher salary ranges. Inferential tests were conducted that confirmed meaningful differences between age groups and work preferences as stated above. These findings suggest that organizers can improve retention by aligning remote work and benefits with employee age and preference, with little impact on job satisfaction and work attitude for all employees.

2. Introduction – background, motivation, and research question.

The American dream has been changing in recent years. Purchasing power, job security, and average ages for large life events have all changed over the decades. We hypothesize that the younger generation has a focus shifted to fulfillment outside of work over compensation. Inversely, older generations value compensation for their work and are more willing to work. In this analysis, we will compare Age, Compensation, Work type, and Years of Experience. We will make useful recommendations to best fit different workforces and predict future trends. The Goal of this research is to wrangle data, interpret data, and provide actionable recommendations.

3. Data and Methods – description of the dataset, cleaning process, and tools used.

In order to properly work with this dataset, it was necessary to first clean it and trim it to a more manageable size. The initial survey results csv file was enormous, having thousands of rows and spanning across 170 columns, so it could not easily be loaded into an excel file. The dataframe was loaded into python, and was filtered to only include U.S. respondents that were employed, who had 30 or less years of experience and had salaries from $30,000 to $500,000. This cleaning constituted enough to save the file as a separate cleaned dataset. The data was then cleaned even further, as many of the rows would be unrelated to our specific analysis. Another dataset was simply made which only included the response-ID, age, job satisfaction, work preference, and salary columns. This dataset was then exported to excel for further cleaning because the age column could not easily be grouped in python (due to being represented by a range rather than one number in each cell). In excel, new columns were made next to the age column, and Flash-Fill was used to extract the minimum and maximum numbers in the range, from which an average was then calculated. A very small number of values in the age column said “prefer not to say,” and these rows were simply deleted because our sample was quite large (having over 3,900 rows). Next, we noticed that roughly 200 cells in the Job Satisfaction column were blank, so we computed the average for the values that were present (which was 7) and imputed this average to all of the empty cells to retain the size of our dataset. The Work Preference Column was then trimmed using Flash-Fill so that only the first term would be displayed rather than a whole sentence. 14 blank rows in the Work Preference column were found, which were then simply removed because of the small size of these blanks relative to our large dataset.

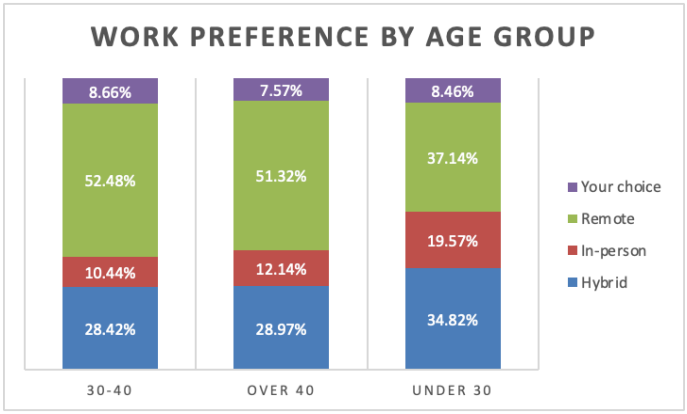
4. Results and Discussion – descriptive and inferential findings supported by visuals.

After cleaning the data thoroughly, the descriptive statistics could be computed to better understand the dataset.

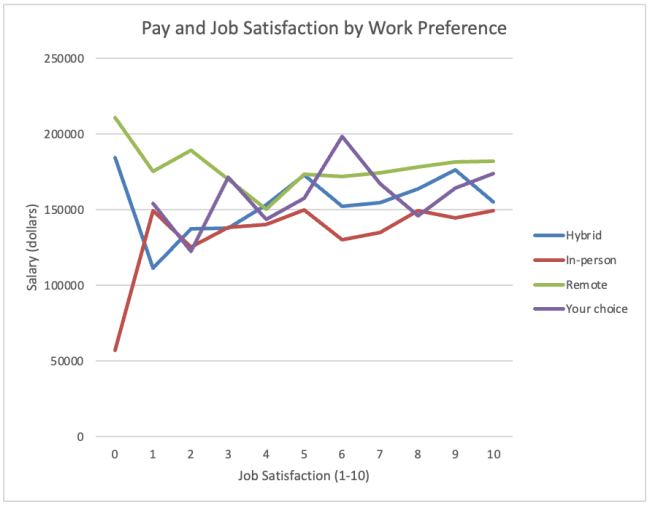
Statistic	***Age***	***Job Satisfaction***	***Work Exp***	***Salary***
Median	39.5	7	12	150000
Mode	39.5	7	10	200000
Standard Deviation	9.072	1.827	8.034	78139.049
Skewness	0.348	-1.058	0.361	1.330
Range	44	10	29	470000
Minimum	21	0	1	30000
Maximum	65	10	30	500000
Count	3893	3893	3893	3893

Firstly, this table displays that the age statistic is actually centered mostly around individuals who fall between 30 and 40 years old, rather than being younger and having a rather large age range, which was rather unexpected. The Job Satisfaction column seemed to center around 7, and its skewness statistic suggests that most of the values are in the upper portion, meaning that most developers are relatively satisfied with their jobs. The Work Experience column appeared to center around 12 years with little skewness, which was higher than we had expected. The salary column had quite a high center, with a large range. This was expected, especially given the median years of experience being so high for this dataset.

After grouping the data by age group, the age groups could be cross-referenced with work preferences.



Clearly, most respondents across all age groups enjoyed working remotely the most, and hybrid work the second most, which was what we expected. What we did not anticipate was that the respondents under 30 actually showed the most preference for in-person work when compared to the other respondents. We expected the younger developers to prefer working from home more



than others, but it actually seems that they dislike in-person work less than all other age groups.

By comparing the salary and job satisfaction by work type, we can see that there is a general trend towards a similar salary range for all types of work. It does seem that those who work in person are generally happier than those who work remotely, even though they are generally paid less. The hybrid preferring workers seem to fall relatively in the middle of remote and in-person workers with regard to their satisfaction and salaries. Overall, this graph suggests that the most satisfied workers across all work types generally have relatively high salaries, but those who prefer in-person work (who may put less emphasis on work-life balance) are easier to satisfy with slightly smaller salaries.

In order to see if age groups differ significantly in their work preferences, a chi-square test was performed on a table that contained age groups and work preferences. The test had a very small p-value, meaning that age groups do differ significantly in their work preferences. This is to be expected, as our first graph shows that younger respondents seem to prefer in-person work much more than older respondents, while older workers seem to prefer more flexible work arrangements.

In order to see if job satisfaction was significantly associated with work type, a one-factor ANOVA test was performed on a table showing job satisfaction by work preference. This test resulted in a very large p-value, meaning that work preferences did not actually have a significant effect on the respondent's job satisfaction.

5. **Limitations, conclusion, and Implications – key takeaways and managerial Recommendations**.

With any data collection, there are some limitations in the process that affect the accuracy of the analysis. I would say a major limitation for our data is the nature of the job satisfaction question. The survey form of collection has inherent issues with the misunderstanding of

questions and the meanings of answers. There are ways to further clarify this particular variable to increase accuracy. We feel a question regarding life satisfaction would give further insight into factors that may influence job satisfaction, as well as preference for the type of work. It would also be beneficial to collect information from an employee's management or supervisor regarding their perceived level of enthusiasm and satisfaction with work. This could provide an objective measure of the variable to further its usefulness. According to Visier, a company that specializes in human resources analytics, employees find it "most difficult to be honest about" questions relating to "overall job satisfaction (Gonzales, 2024)." The company cites the driving force for dishonesty in questionnaire answers as "a lack of trust in how survey responses will be handled (Gonzales, 2024)" and distrust related to anonymity. Part of the data collection process must be educating employees on the systems used to keep responses anonymous, to ensure accuracy and honesty in responses.

One major implication of the data is the increasing preference toward remote work as employees age. This is a very insightful variable that can be used by management in a multitude of ways. As a retention tool, opportunities for hybrid and remote work should be offered at increasing rates to workers, in relation to their time at the company. This aligns with employee preference and encourages employees to remain with the company. This also gives insight into employee needs in terms of benefits. Older employees likely prefer remote work as a result of increasing levels of responsibility outside of the workplace. Childcare benefits are increasingly important as employees age. Furthermore, the quality of healthcare benefits should be thoroughly examined for workers in the highest age bracket, as the difficulty of commute and travel is likely influential on the preference toward remote work.

The data presented shows higher rates of satisfaction for workers classified as in-person. According to Forbes, “good relationships with colleagues”, “good relationships with superiors”, and “appreciation for work”, are all top 10 factors of “employee happiness (Morgan, 2015).” In-person employees likely gain a sense of community that hybrid and online workers miss. From a management standpoint, this information should be used to inspire events such as employee outings and company parties. It is imperative to foster a sense of community amongst workers, and even more important to ensure remote employees are involved in these events. It is also necessary for management to personally reach out to remote employees consistently and maintain a relationship.

In conclusion, the data analyzed shows a preference toward in-person work for younger employees, and an increasing preference for remote work as employees age. This suggests the use of increasing remote work opportunities to employees as they progress in the company to reward loyalty and increase retention. These statistics also suggested that as employees age, benefits such as childcare and healthcare become increasingly important. Higher rates of job satisfaction for in-person employees can likely be attributed to a sense of community, which can be fostered amongst the entire workforce through company events and outreach from management.

6. **References – properly cite the dataset and any external sources**.

Gonzales, S. (2024). *Employee Honesty & Productivity: The Truth Revealed | Visier*. Visier.com. https://www.visier.com/blog/new-survey-employee-engagement-productivity-impact/

Morgan, J. (2015, January 23). The Top 10 Factors For On-The-Job Employee Happiness. *Forbes*. https://www.forbes.com/sites/jacobmorgan/2014/12/15/the-top-10-factors-for-on-the-job-employee-happiness/