Productivity of Remote Work in the Digital Era

Saransh Puri

Thoai Phan

Govinda Sahoo

UC San Diego

June 2024

Problem of this Project:

In the contemporary world defined by digital technology, working from home has emerged in many professional organizations. As people adapt to this transition, understanding the aspects of remote work productivity becomes tremendously important. A popular debate regarding remote work during the COVID-19 pandemic is whether or not remote work has an impact on one's health, well-being, social interaction, and balance between their professional and personal life.

The goal of this project is to dive into the relationship between remote work environments and productivity, and how factors such as interactions, social isolation, and one's health are affected by this. We aim to do this by analyzing the various ways in which remote work conditions impact both productivity, but also by examining external factors such as job satisfaction, time spent working, commuting, and with family under both work from home and in person work. This alongside analysis on remote work sentiment can provide insights on the various social and emotional impacts of the work from home lifestyle, which can also be used to extrapolate effects of employee interactions and health.

By undergoing a comprehensive analysis of these factors, we want to produce a project that reveals the complex relationship between remote work environments, productivity, and other remote work dynamics as these factors are still present to drive business success in the contemporary digital age.

Relevance/Inspiration of Problem:

Working from home is a critical aspect of the modern world as it becomes a widespread adoption, especially during the COVID-19 pandemic. This project can provide insights into this problem by highlighting the impact that remote work has on productivity and possible factors that either promote or hinder work practices from home. One cannot deny the continuous widespread adoption of working with technology and from home as the world continues to become interconnected through technology.

By illustrating the benefits and challenges of working from home, this project can encourage people to promote better working practices for employees and employers, as we continue to live in the digital age where remote work is becoming a popular workplace. Similarly, this project could help investigate the changes in the work force since the covid pandemic, more specifically if it has had any impacts that are likely to be longstanding. This project also has relevance in terms of how employers can best utilize technology moving forward, as technology could both enhance and complicate remote work setups for employees. Thus, this can be used as a source of information with which to provide recommendations regarding remote work policies and practices, to find ways employers can increase productivity and overall employee fulfillment, given we are moving towards a more digital and remotely connected workplace.

Data Sources:

Our project used two data sources primarily, both provided different angles of analysis on remote work and its impacts. These sources were what enabled us to come to conclusions regarding the impact of remote work on employees, especially its impact on productivity and employee sentiments.

One of these data sources is a Kaggle data set, which included survey responses based on the experience of 1,500 remote workers in New South Wales, Australia. The data was collected at two different points in 2020 and 2021, which allows the data to provide insight into employee changes in sentiment and productivity resulting from the COVID-19 pandemic. This data set was developed to analyze productivity and the general impact of remote work on employees and to aid in creating policies for a more efficient work environment post-pandemic.

The second data set was from the UK office for national statistics regarding home working in the UK. It includes detailed data from the annual population survey administered in the UK that is broken down based on several demographic and job characteristic-related factors. This data set was created to use in internal UK government analysis on how working from home varies across sectors, demographics, and job types within the UK.

These data sources allowed us to analyze the evolution of employees' feelings on work from home, both during the pandemic and in the time that followed. Due to the detailed nature of the surveys administered we also had the option to more closely examine the various impacts of demographic information on remote work statistics. This could allow us to more closely examine the impacts of remote work on specific groups of people, to determine who is more positively and negatively affected.

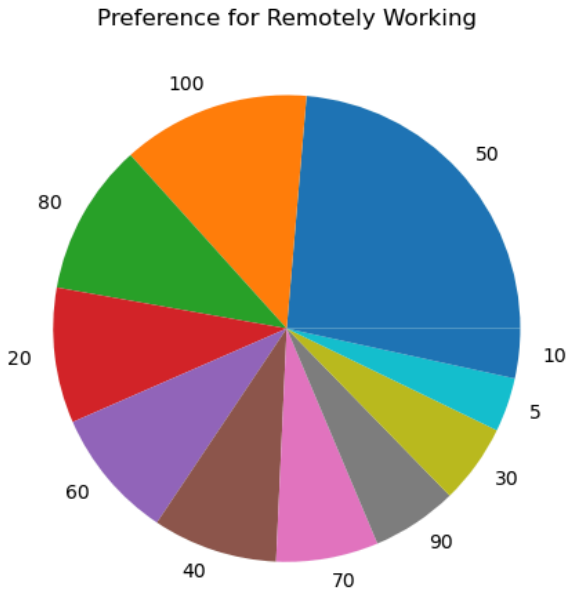
Data Description/Exploration:

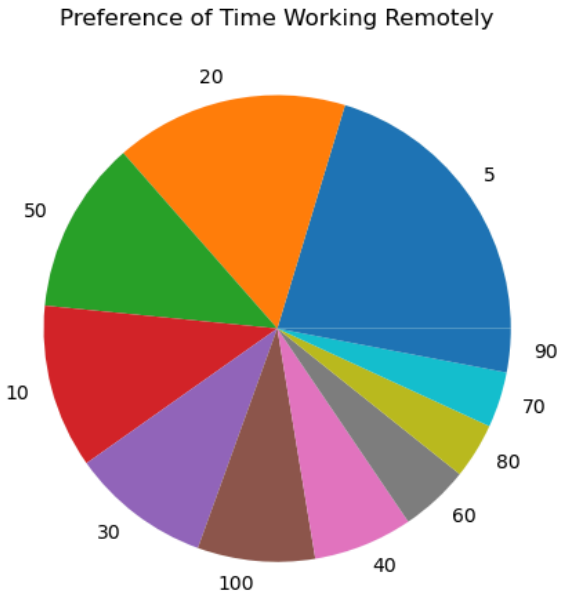
The Kaggle dataset contains detailed information on the responses of each survey recipient, this data includes employee age, gender, industry, specific title, and occupation, and asks each individual several questions regarding their personal sentiment on remote work. These questions were asked both during and shortly after the pandemic, and could thus be used to track changing employee sentiment on remote work. The survey asked numerous questions regarding factors, such as difficulty in getting approved to work remotely, self-reflections on productivity, specifics on time spent during the workday, and their desire or lack of desire to return to in-person work after the pandemic ended, amongst many others.

The UK national statistics data set covered many of the same attributes in its data. It included a wide range of variables related to remote work. These included age, gender, industry, occupation, employment status, and daily employee metrics including hours worked, their incomes, and satisfaction with their employer. This dataset covered these factors as well as many others regarding specific employee compensation but placed special emphasis on breaking down each of these factors for many individual groups, such as different demographic groups, professions, and genders. Similarly, this dataset contained categorical and continuous variables covering a wide variety of remote work logistics and sentiment in the UK.

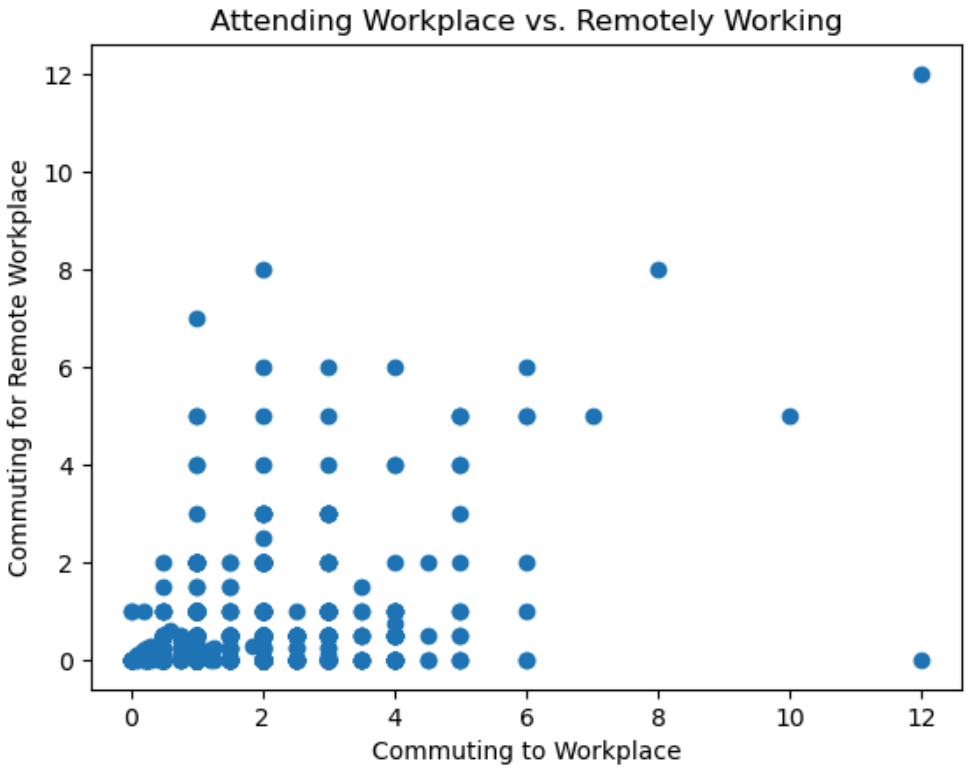
These data sources will allow us to look more closely at the relationship between remote work and related performance measures, including productivity by comparing the amount of time spent by employees working compared to other time consuming tasks such as commuting. Similarly, these data sources enable us to view similar questions from the same group of respondents, both during the height of the pandemic and as it was coming to a close, allowing us to track changes in employee views of remote work. Additionally, due to the information available regarding various demographic variables, we can view the impacts of remote work on different groups, allowing for the possibility of creating recommendations tailored to specific groups.

Exploratory Data Analysis:

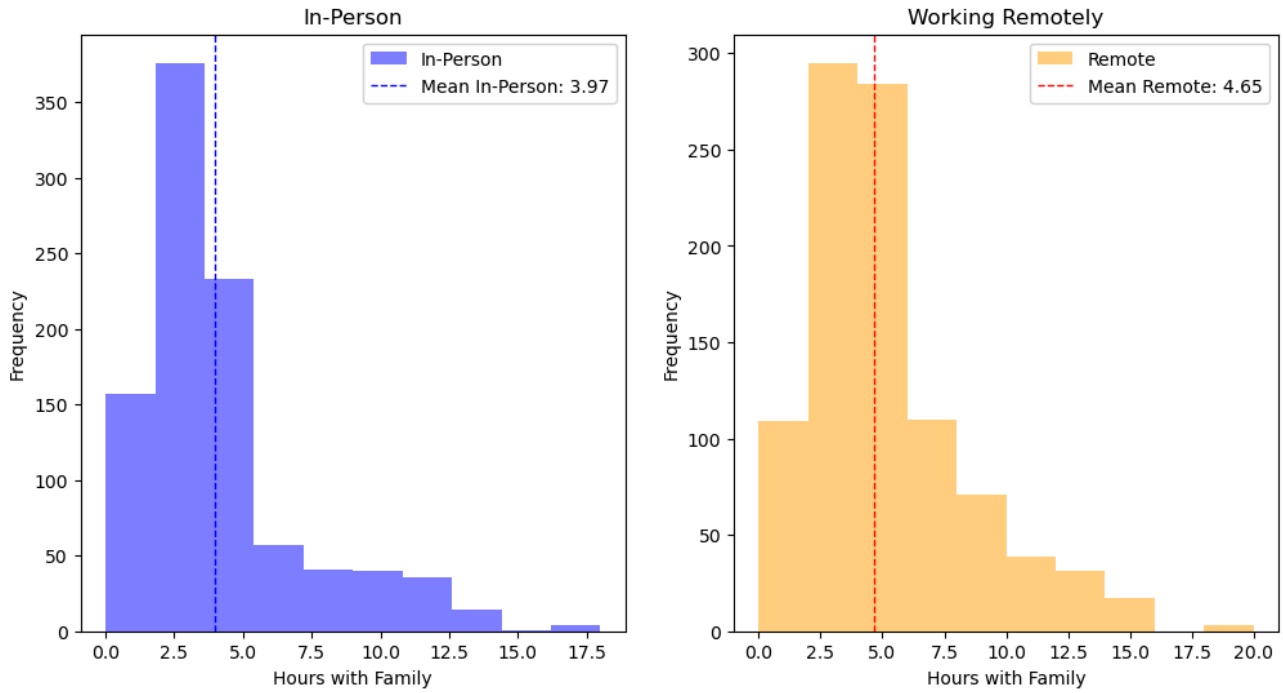




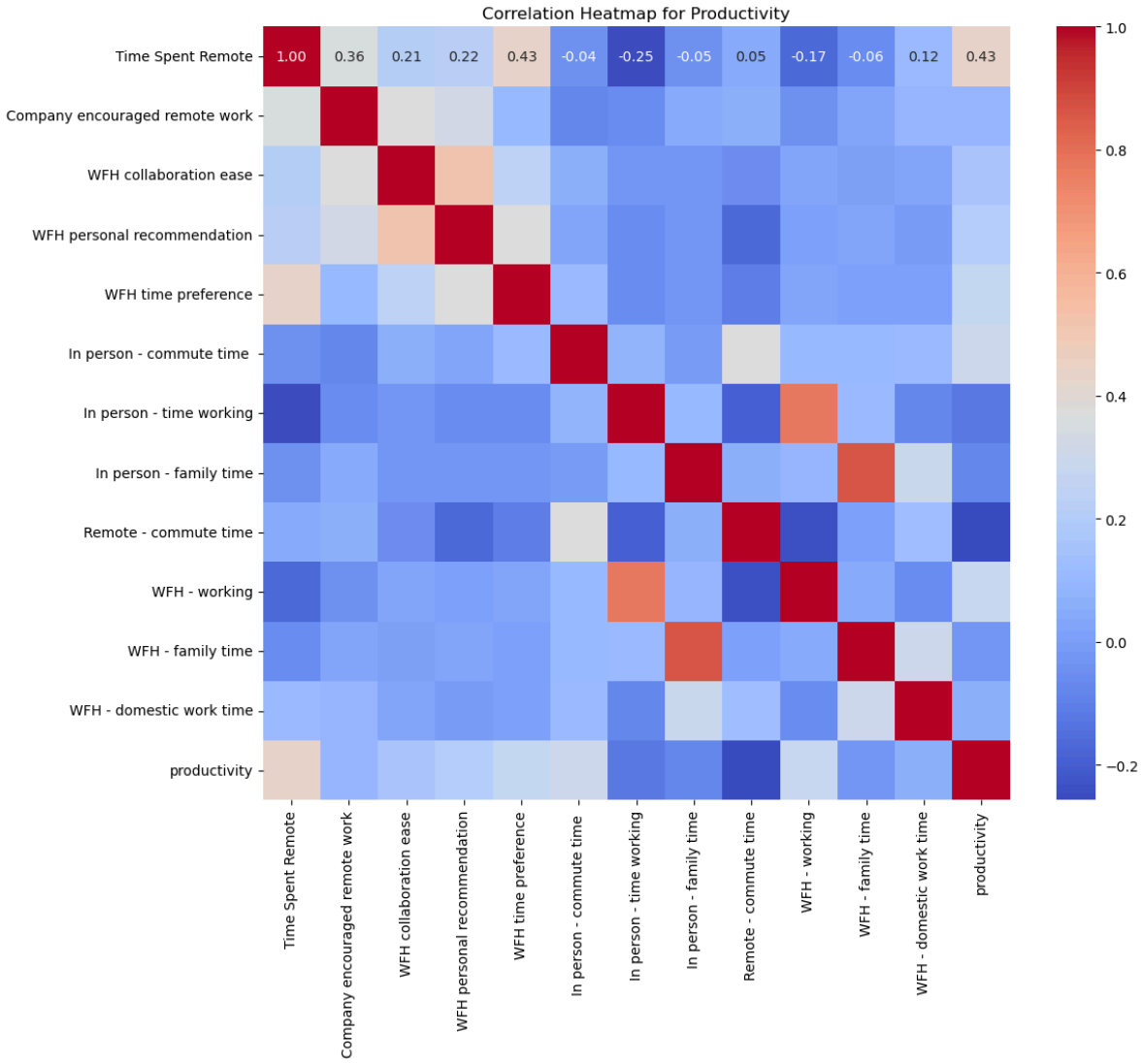
To start off, one cannot deny the tremendous changes that people experienced in the beginning of the COVID-19 pandemic as they were forced to transition to a completely remote environment for work, school, etc. We see that in the first pie chart, people spent a lot of hours working remotely in 2020 due to the pandemic. However, an interesting fact that we found is that according to the second pie chart that asked the surveyees if they would still prefer working remotely after the COVID-19 pandemic has been eradicated, we see that a lot of people would choose working remotely instead of fully transitioning back to an in-person workplace. We see that more than 25% of the surveyees said that they want to spend 50-100% of their work time remotely, while a very small portion of the surveyees said that they would prefer to spend little time remote (5-10%). This suggests that some people might believe that a flexible work environment improves their work productivity because they need optimizing comfort from their homes to work in order to enhance their focus and concentration on work.



This scatterplot illustrates how for the most part, we see that when people work remotely, they spend less time commuting. Although one can argue that there is a positive correlation, if we look at the cluster in the bottom left of the graph, we see that a lot of people who used to commute to their workplace stopped when they are allowed to work remotely and we see a small negative correlation. This is an interesting observation because it can improve workers' productivity as they spend less time getting to their workplace and have more time to focus on their work. By reducing their commuting times, this additional time that employees have can allow them to start their workday earlier, have more time to complete projects if needed, and possibly reduce stress associated with the struggles of getting to work. Additionally, as reduced commuting time gives people the chance to improve their work-life balance, people will feel more encouraged to be productive in their time at work.



This graph highlights how people are able to spend more time with their family when they get the opportunity to work remotely. In this dataset from 2020, we see that in the past year, people spend an average of 3.97 hours per day with their family when they commute to work, while in the last 3 months when they work remotely, they spend an average of 4.65 hours per day with their family. This suggests that spending more time with family has an effect on one's work-life balance. This is an interesting insight on work productivity because employees who have more time to recharge outside of work tend to feel more refreshed and focused when they have things to complete in their remote workplace. Additionally, many people can agree that strong family relationships provide workers with a source of encouragement and emotional support because people who feel valued and loved by their family back home might feel more encouraged to be more productive as family connections can positively influence job performance and satisfaction. Positive job performance can also lead to an increase in productivity and commitment to one's work goals.



data['productivity'] = (data['Time Spent Remote']/100)*((data['WFH - working'] - data['In person - time working']) + (data['In person - commute time '] - data['Remote - commute time']))

One of the most important things to highlight in this heat map is that in the top right, we see that there is a positive correlation between time spent working remotely and productivity, which is calculated by using the formula above. Another thing to note is that there is a positive correlation between time spent working remotely and one's domestic work time, where they have time to do other things at home besides working. This analysis is important because it highlights the benefits of remote work arrangements to foster an efficient and even more effective work environment for people. Also, in terms of domestic responsibilities, people who work remotely can find a better balance between their professional and personal responsibilities as their daily commutes are eliminated in this work environment. One can argue that finding this balance contributes to their well-being and satisfaction with their careers as illustrated in the heat map, which can also lead to increased productivity and good performances.

Previous Analyses/Similar Datasets (Provide References):

1) **Does Working from Home Work? Evidence from a Chinese Experiment | NBER**

"Does working from home work? Evidence from a Chinese experiment" by Nicholas Bloom was a study that followed several thousand employees at a Chinese travel agency corporation, employees who were a part of the study were randomly assigned to either work in person or from home for nine months. The study found significant performance improvements for remote employees. The authors found that this could largely be attributed to the increased convenience and more focus-inducing environment employees associated with working from home. The study also found that employees who were working from home were more likely to stay with the company longer term. Their analysis used subjective survey data along with objective metrics of employee performance from the company to better isolate the specific impacts that working from home had on employees.

Overall, Bloom's study and analyses on work productivity is still relevant in the modern world because by providing readers with empirical evidence of the positive effects on employee performance and working from home, one can realize the potential benefits of remote work in improving work productivity and organizational outcomes, both of which are important in many businesses and work settings. It prompts new discussions such as the argument on whether or not remote work can enhance one's job, satisfaction, and loyalty to their company as his previous research highlights how flexibility and autonomy from remote work allow people to value their careers more.

2) **How the Coronavirus Outbreak Has – and Hasn't – Changed the Way Americans Work**

The Pew Research Center conducted a similar study, they tried to assess the impact of the pandemic on work practices specifically within the US. Their study surveyed a large sample of working adults across the US from all types of jobs and industries, asking questions about how their work has been impacted by COVID-19. The study found that around 70% of people in the survey were working from home for the majority of their time during the pandemic, which was around 20% more than pre-pandemic.

Given the number of people working remotely, it was notable that the survey results found that a majority of people preferred to continue working from home post-pandemic, largely due to the increased flexibility and because they would no longer have to commute. Negatively, most employees shared they would likely feel less connected to their coworkers if they didn't work in person. This study utilized a general quantitative survey method using online and phone surveys.

Overall, this study found mixed results in terms of remote work's impact on productivity as efficiency was improved for some, but for others the difficulty associated with collaboration and maintaining work-life balance was detrimental. Although remote work offers many benefits such as more family time, increased flexibility, and less commuting time, we see that in this case study, remote work also comes with challenges, such as communication and less opportunities to have face-to-face interactions. This suggests that it is important for businesses to think of certain strategies and technologies that will provide support for both modes as we are living in a world

that is driven by technology, while also prioritizing their employees' health and well-being for them to have a healthy work-life balance.

Analyses Performed:

1) **Hypothesis Testing:**

- **Data Cleaning and Preparation:**

  We only selected the relevant columns for analysis: 'WFH - working' (productivity hours while working from home) and 'In person - time working' (productivity hours while working in-person) for the hypothesis test. Additionally, we removed any rows with missing values in these two columns to ensure the accuracy and reliability of the statistical test.

  This is important in hypothesis testing because selecting relevant columns and removing rows with missing values are essential steps to conduct a hypothesis test as the results will ensure us that the test specifically illustrates the differences in productivity between working in-person and remotely. By removing the null values, we also reduce the possibilities of encountering unnecessary complexity as it provides us with a direct conclusion.

- **Difference Calculation:**

  We created a new column called 'productivity_difference' to capture the difference in productivity between working from home and working in-person for each respondent. By calculating the productivity difference between in-person and remote work, we now have a measurable variable that can form the basis of our hypothesis test to see if the observed differences are statistically significant.

  Furthermore, by having a new variable that focuses on the relationship between remote work and productivity, we can hypothesize that there is a significant positive difference in productivity between in-person workplaces and remote settings. As illustrated on the next page, we tested our hypothesis using a t-test once we decided on a variable that is appropriate for statistical testing.

- **Statistical Testing:**

  We conducted a paired t-test using the ‘’ttest_rel’ function from the scipy.stats module in Python, which compares the productivity scores of individuals when working from home versus in-person. The alternative hypothesis for this test was set to 'greater', indicating the test was one-tailed with the assumption that working from home might lead to higher productivity. By using the paired t-test, we can argue that the potential benefits of remote work in increasing productivity are valid if the test detects that productivity while working from home is greater than productivity while working in person.

2) **Correlation Matrix:**

- **Relationships:**

  We decided to use a correlation matrix as one of our tests in this project because a correlation matrix helps us understand how variables are correlated to one another. A positive correlation, in which the coefficient is closer to 1, suggests that as one variable increases, the other also tends to increase and a negative correlation, in which the coefficient is closer to -1,

suggests that as one variable increases, the other tends to decrease. Finally, a correlation coefficient of 0 suggests that there is no relationship between the two variables.

- **Variable Selection:**

  In data analysis, we use correlation matrices to look for certain variables that are highly correlated with one another because those are the ones that illustrate strong relationships in our datasets. These variables are important for predictive modeling because they suggest an important relationship in our argument, while variables with low correlation are not as important as the other ones. Usually, we try to reduce our interpretations of these variables since they are not as important and this will also enhance our argument by providing readers with relevant information.

3) Logit Regression:

- **Probabilities:**

  The logit regression model measures the probability that something will occur given the values of the independent variables. Using the logistic function, the logit regression model changes the predictors into probabilities that are between 0 and 1, as illustrated on the next page. We used this model of regression because it is a good model for regression analysis as it uses the probabilities to form a linear relationship with the variable of productivity.

- **Pseudo R-Squared:**

  This value is a measure of the goodness of fit for a logistic regression model. This value illustrates how well the model explains the variance in dependent variables by providing an estimate of the model's predictive power. This value ranges from 0 to 1, like the coefficients of this model, in which a value that is closer to 1 indicates a better fit and a value that is closer to 0 illustrates no correlation between the two variables.

- **LLR (Likelihood Ratio Test) P-Value:**

  The LLR P-value is a test statistic that highlights the significance of a logistic regression model. It tells us whether or not the logistic regression model, where all of its variables are included, provides a good fit to our dataset compared to a null model, in which all coefficients of predictors equal to zero. For example, a small LLR p-value that is less than the significance level suggests that the model tremendously improves the fit to the data, where at least one variable is significantly related to the outcome. We say that this value is significant as it is a source of evidence in favor of our logistic regression model's capability of explaining the variability in our dataset.

Analyses Interpretation:

1) Hypothesis Testing: T-statistic = 8.89, P-value = about 0.00

The output of the previously described hypothesis test provided us with both a test statistic and the corresponding p-value. The test statistic measured the size of the difference relative to expected variation. The test statistic value is very high showing that there is a strong difference between productivity when working from home versus in person. It being a positive value demonstrates that the productivity is higher when working compared to working in the office. Similarly, the very small p-value which is well below the .05 alpha threshold shows that

the result is statistically significant. Showing that the likelihood of observing such a large difference between productivity given the null hypothesis of there being no difference is extremely small, and close to zero. This essentially means that there is very strong evidence against the null hypothesis, meaning there is little to no evidence that there is no difference between productivity when working from home versus in person, and the low p-value shows that this is significant and not something that can be attributed to chance. The data therefore shows that employees were significantly more productive under remote working conditions, this can be due to a number of factors including reduced commuting time, and better work environment.

For this reason, organizations should consider these results as reason to continue or being implementing flexible work policies, including giving employees the option to work from home, given that the data suggests their employees would actually be more productive remotely, and thus moving to remote work could both increase output from workers, while also potentially reducing overhead costs such as parking and rent. That being said, additional studies and surveys are likely required before definitively being able to say whether or not remote work is the ideal solution, especially considering the wide variety of industries and demographics that employees could be based in.

2) Correlation Matrix: Variables & Correlations

1. Time Spent Remote Working and Productivity: correlation = 0.43
    - There is a positive correlation between "Time Spent Remote" and "Productivity, which suggests that individuals who spend more time working remotely have higher productivity, illustrating how remote working environments might enhance productivity.
2. Company Encouraged Remote Work and Productivity: correlation = 0.09
    - There is a weak, but positive, correlation that illustrates how organizational support for remote working is associated with higher productivity. One can argue that encouraging remote work with good support can improve productivity.
3. WFH Collaboration Ease and Productivity: correlation = 0.16
    - The ease of collaboration while working from home has a moderate positive correlation with productivity. Effective communication and collaboration tools are crucial for maintaining productivity in a remote work setting.
4. WFH Personal Recommendation and Productivity: correlation = 0.20
    - There is a positive correlation between recommending remote work and productivity.
5. WFH Time Preference and Productivity: correlation = 0.28
    - There is a positive correlation between time preference and productivity, which suggests that employees who prefer working remotely tend to report higher productivity. One reason for this is that remote work allows for better work-life balance and personal satisfaction, both of which can enhance productivity.
6. In-person Commute Time and Productivity: correlation = 0.31
    - There is a positive correlation between commute time and productivity. This highlights how reducing commute times can improve productivity as people can allocate more time for work and personal lives.
7. WFH - Working Hours and Productivity: correlation = 0.28

- There is a positive correlation between the number of hours spent working from home and productivity.

8. Remote - Commute Time and Productivity: correlation = -0.26
    - There is a negative correlation between commute time and productivity, implying that longer commute times tend to lead to a decrease in productivity. Therefore, minimizing commute times is beneficial for workers as it allows them to maintain high productivity levels.

Interpretation:

Based on the correlation matrix in this test, one can argue that remote working leads to higher productivity. From this test, we can argue that organizations should promote support and collaboration in remote environments to maintain high productivity because as illustrated above, we had a negative correlation between commute time and productivity. This means that reducing commute times can enhance productivity as people have more time for their other commitments, such as personal and family tasks, while crafting time to do work in an encouraging manner.

3) Logit Regression on Productivity

1. WFH Collaboration Ease:
    - Coefficient = 0.1888, P-value = 0.033
    - We have a positive coefficient, which suggests that as collaboration increases in a remote work setting, the likelihood of higher productivity will increase as well.
2. WFH Personal Recommendation:
    - Coefficient = 0.2165, P-value = 0.027
    - The positive coefficient in logit regression illustrates how employees who recommend working from home tend to see higher productivity. This could be due to work style preferences and working conditions that they found suitable for themselves, leading to more efficient work.
3. WFH Time Preference:
    - Coefficient = 0.0236, P-value = 0.00
    - The coefficient and p-value =f 0 suggests a strong positive association between time preference and productivity, illustrating how employees who prefer flexible working hours while working remotely are more productive. This prompts flexibility in managing work hours because we see that it is crucial for one's productivity in a remote work setting.
4. Job Experience:
    - Coefficient = 0.0242, P-value = 0.759
    - Since our p-value is 0.759 > 0.05, we cannot conclude that there is a significant relationship between job experience and one's work productivity. The lack of significance suggests that job experience does not relate to productivity in the workplace. Regardless of how long an employee has been working, there are other factors that are more relevant to remote work and productivity levels.

Overall Model Performance:

- Pseudo R-squared: 0.1024
    - The model explains about 10.24% of the variance in productivity. This highlights a potential limitation to our study because this is a moderate level of power, so

one can say that there are many other unobserved factors in this topic that may influence productivity.

- LLR p-value: 3.196e-26
    - Our p-value is definitely less than 0.05 and this highly significant LLR p-value indicates that the model in this project is statistically significant and strongly fits the data.

Interpretation:

The logit regression model illustrates the importance of remote working factors such as time preference, collaboration, and commuting time play in enhancing productivity. In terms of collaboration, we see that personal recommendation for remote work is a significant predictor of higher productivity as many people say that they are more satisfied with working remotely. Another interesting insight that we saw is that job experience does not play a significant role in this context. This can be due to the fact that many people are unfamiliar with technology and regardless of how many years they have been working, switching to a fully remote environment can be challenging for many people who do not work with technology on a daily basis. These insights can promote practices and a productive remote work environment as it illustrates how there are many factors that can increase productivity if they are enacted.

Potential Limitations/Shortcomings:

Although this project and analyses provides insight on the positive relationship between remote work and one's work productivity as people are given more time to spend with their family and spend less time commuting to work, there are some limitations to the analyses that we performed in this project. Factors such as generalizability, temporal effects, and other confounding variables might contribute to the limitations of our argument.

To start off, generalizability can provide us with some potential limitations because we only introduced a few variables that have an impact on work productivity. However, there are many possible factors, such as regional differences and various industries, that might influence the effectiveness of remote work practices and one's work productivity. For example, there is a possibility that individuals living in urban areas that have good internet connectivity might have a more pleasant work experience than people in rural areas that have low access to technology. There can be other factors such as government regulations, remote work funding, and how knowledgeable an industry is, that contribute to one's productivity with remote work.

Furthermore, since this project mainly considers the temporal effect of the COVID-19 pandemic, only using results in one year (2020) can lead to some shortcomings in this study. In the beginning of the pandemic, many businesses were forced to quickly implement remote work arrangements to promote safety for everyone. Since we only used a dataset from 2020, the year where transition to remote work brought challenges to technology, communication, and work policies, there could be outliers such as social restriction and limited sample size that result in the correlations that we found in this project. These outliers might have influenced people's perceptions and experiences of working remotely as many devastating events happened that year, so there is a possibility that the dataset that we worked with is limited. It is definitely possible for one to have different data analyses with different datasets if they consist of different sample sizes, whether it is less or more people.

Conclusion/Relation to Proposal:

The findings from our research align directly with our initial goal of developing a better understanding of the impacts of remote work, and how it impacts factors including productivity. By using several data analysis methods we found that there was a significant positive impact of remote work on an employee's productivity, which directly answers our initial question.

The various analysis methods used, especially the paired t-test provide strong evidence in support of the hypothesis that working from home can increase the productivity of employees compared to the traditional in-person work environment. This finding shows the relevance and importance of our research question, considering the immense impact that digital technology can now be expected to have on the traditional workplace.

As a result of our strong findings, it can be concluded that these conclusions could have direct implications for the work policies of the future. Companies and organizations should now more than ever more closely consider remote work options, as a part of their employee productivity efforts. This has multiple positive impacts including increased productivity, but our data also suggested increased happiness and employee satisfaction related to being able to work remotely.

Additionally, this project contributes to existing research on this topic, by running a wider variety of analyses on the dataset, providing several different pieces of evidence in support of the productivity benefits of remote work, during and after the pandemic.

The outcome of our study allows us to strongly recommend that companies move towards a remote work option for their employees. This can be seen in both the increased productivity of employees, as well as their increased happiness and ability to spend more time with their families. Similarly, employers can also gain by moving towards a remote work culture, as their overhead costs could be decreased without the need to pay for office space.

However, despite the insights gained in our analysis, there remains a need for ongoing research in this space. Future research is needed to better understand the long-term impacts of remote work on productivity and employee fulfillment. This would allow employers to better understand the most effective methods of sustaining productivity over the long term in remote work environments. Additionally, remote work includes a wide variety of different demographics and industries. For this reason, generalizing an optimal strategy for all situations is inherently limited. For this reason, tailored research to various different positions and industries would likely be required to maximize the benefits of remote work for the diverse sector of jobs available.