Executive Summary

In 2019, the world was shaken by the spread of COVID-19, leading to the movement of many jobs from in-person to online to allow work to continue despite lock downs to prevent the spread of the virus. In addition to slowing the spread of COVID-19, working online allowed for more flexible hours and no commute time.

However, certain drawbacks from working at home, including the lack of a cooperative environment and the inabilities for some jobs to fully shift may cause remote work to only be a temporary response to COVID instead of a permanent societal shift. Our team aims to conclude whether or not the trend towards remote work will be a permanent addition to society by looking at data trends from five locations located in the United States and United Kingdom.

Using polynomial regression on city employment data, we determined predicted occupational growth/decline in the five cities; using remote-ready job estimated percentages for each occupation, an approximation for the percentage of jobs that are remote-ready by 2024 and 2027 were shown. In 2024, the percentages were as follows: Seattle (44.2%), Omaha (49.0%), Scranton (44.3%), Liverpool (21.4%), and Barry (46.8%); in 2027, the percentages were as follows: Seattle (43.1%), Omaha (48.0%), Scranton (42.9%), Liverpool (21.3%), and Barry (46.6%).

We next ran a probabilistic model that took in inputs of age, education level, location, job, individual stance on remote work, and date of employment; these values would help determine the probability that a person would work from home. We lastly consolidated these findings and connected them to socioeconomic problems to validate the efficacy of our regression and probabilistic algorithms.

In order of remote working's impact from most to least, the cities are as follows: Barry, Omaha, Scranton, Seattle, and Liverpool. Wale's government push for remote work and Omaha's young and evolving population show these cities will be the biggest place for impact within the remote work industry.

Contents

1 Q1: Ready or Not

1.1 Workflow

First, the city employment data was normalized based on year, confining the values between zero and one. Python libraries, like *numpy* and *pandas*, were used to manipulate the data into a workable format for polynomial regression through additionally removing missing values.

	Job Activity	2000	2010	2015	2019	2020	2021
0	Mining, logging, construction	0.061074	0.050141	0.056706	0.061146	0.065666	0.064056
1	Manufacturing	0.127792	0.100162	0.099645	0.088317	0.085128	0.083109
2	Trade, transportation, and utilities	0.195532	0.180891	0.187642	0.190723	0.197301	0.194389
3	Information	0.047742	0.052600	0.051623	0.061530	0.067587	0.081239
4	Financial activities	0.061134	0.055239	0.050776	0.048591	0.050753	0.051198
5	Professional and business services	0.132417	0.132370	0.142214	0.144767	0.149479	0.162186
6	Education and health services	0.110317	0.138847	0.133054	0.135614	0.137549	0.130625
7	Leisure and hospitality	0.087557	0.093385	0.098057	0.099578	0.076130	0.077732
8	Other services	0.034711	0.037906	0.037168	0.037713	0.035942	0.034658
9	Government	0.141725	0.158460	0.143114	0.132020	0.134466	0.120807

Figure 1: Normalized Washington Employment Data

Special attention was given to the pandemic, which could have skewed employment scenarios; three datasets with particular annual ranges were formed: 2000-2019, 2019-2021, 2000-2021. Each of these datasets were run through polynomial regression for each city and job activity.

1.2 Polynomial Regression

Polynomial regression seeks to fit a two-variable scatterplot using a multivariable formula, in a quadratic ($ax^2 + bx + c$), cubic ($ax^3 + bx^2 + cx + d$), or n-th order equation. This is done through optimizing the mean squared error from the residuals of the predicted values ($\hat{y}$) to the true values (y), more directly defined as the cost function:

$$J = \frac{1}{n} \Sigma_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{1}$$

Looking at a quadratic formulation ($y = ax^2 + bx + c$), running gradient descent to arrive at a local minimum for J will produce the optimal a, b, c; through computing partial derivatives, we can update these variables accordingly: $a = a - \alpha \frac{\partial y}{\partial a} J$, $b = b - \alpha \frac{\partial y}{\partial b} J$, where α is the learning rate for optimization and c is the bias in the equation.

To fit the data provided, polynomial regression using the *sklearn* library was performed.

1.3 Visualizations

Shown below are graphical visualizations denoting city employment data from 2000-2019, split between America and Britain. These visualizations were also performed for 2019-2021 and 2000-2021 to better understand the data.

We see less definitive spikes throughout occupations pre-COVID; however, when looking specifically at COVID, there's a much heightened positive/negative impact for particular occupations. For example, we recognize a large decrease in informational services and other.

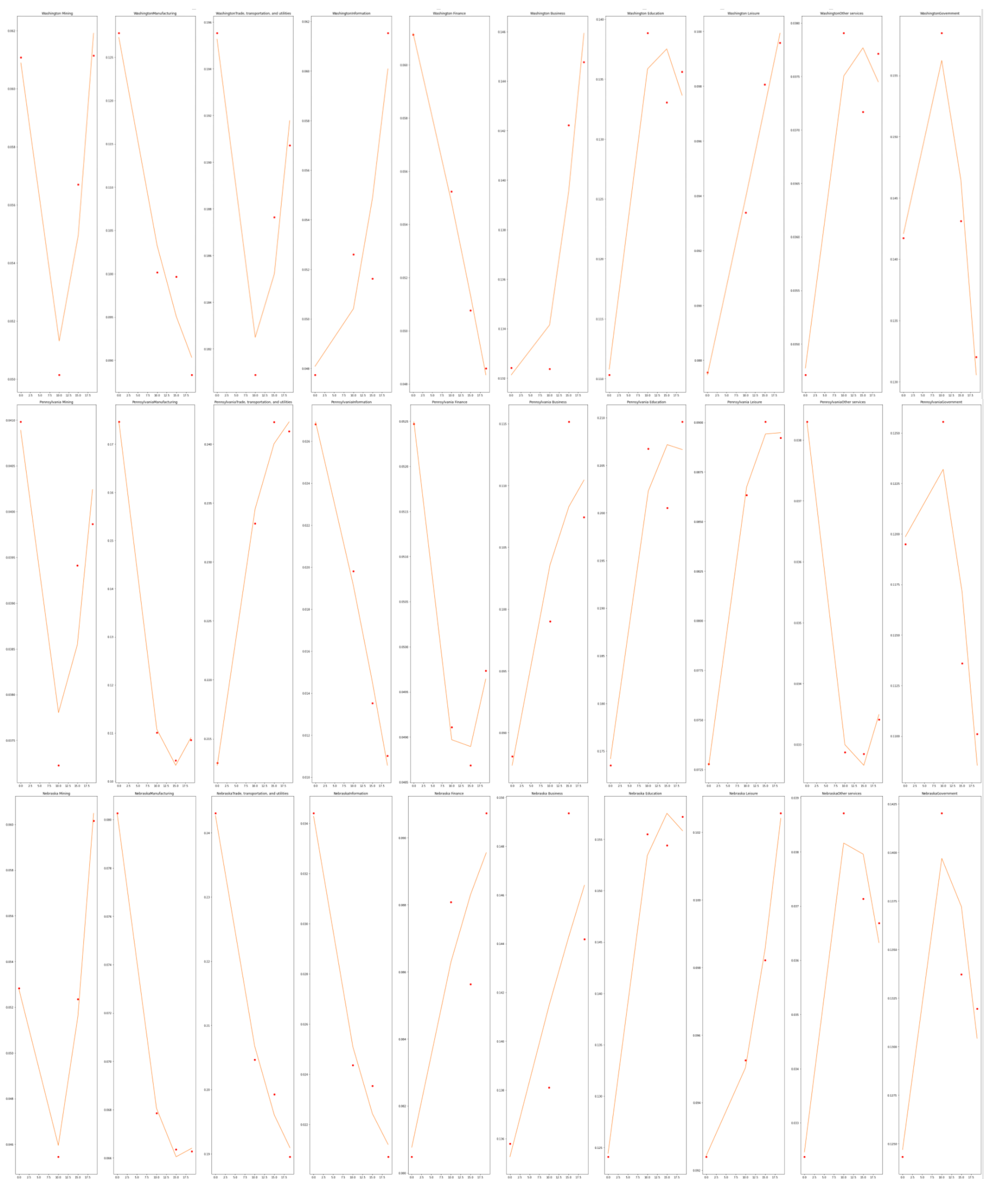


Figure 2: Washington, Pennsylvania, Nebraska Regression

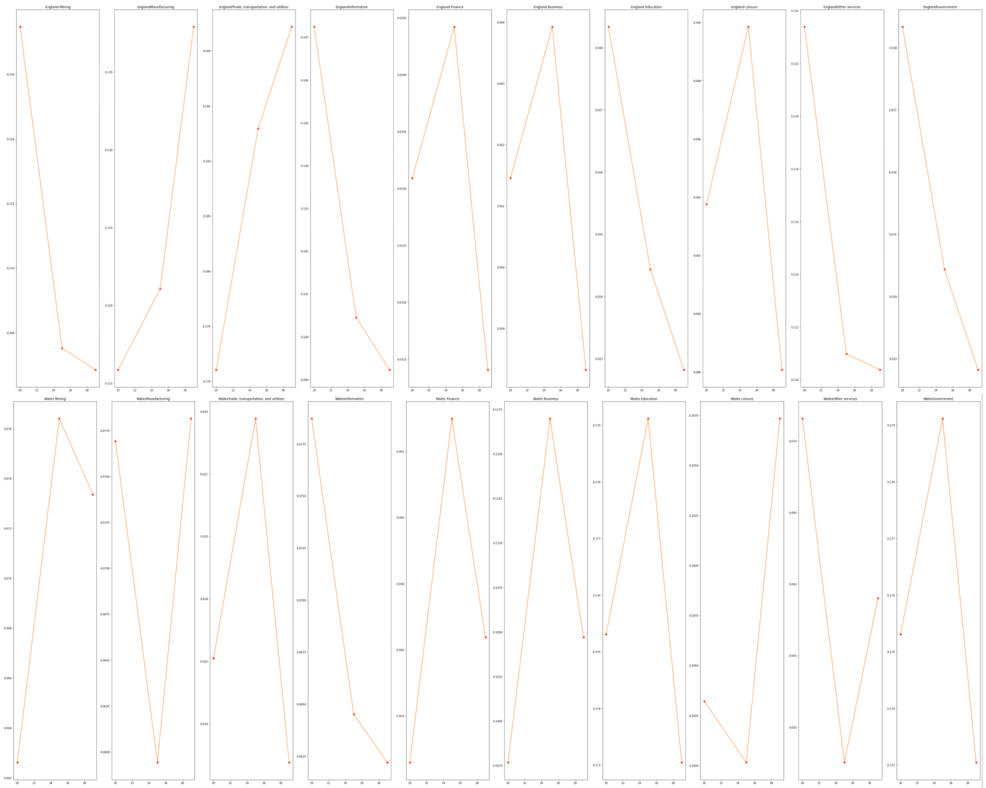


Figure 3: Britain, Wales Regression

1.4 Predictions (2024, 2027)

Polynomial regression estimated the percentage of workers in each occupation in 2024 and 2027; this regression analysis was performed for each of the three datasets. Note that the 2024/2027 predicted values were re-normalized to ensure that the percents would add to one.

Through inspection, some of the values predicted from 2019-2021 (COVID) dataset were negative; this was probably because the model was unable to generalize the variability of job unemployment from COVID. However, for the other data, most the values were positive, with each occupational percentage increase/decrease being different for the five cities.

Using remote work data from the National Bureau for Economic Research, we can estimate the amount of remote-ready jobs (2024) by multiplying the percentage of remote-ready jobs (from D3) and the predicted 2024 value (from D1). For example, looking at Omaha values: $(0.00)(0.073161) + (0.01)(0.07347) + (0.03)(0.188777) + (0.28)(0.19589) + (0.88)(0.092116) + (0.88)(0.146680) + (0.98)(0.156859) + (0.26)(0.093502) + (0.25)(0.034439) + (0.65)(0.124530)$

After multiplying these values for each occupation, we arrive at $\sim$48.96% of workers in

Omaha whose jobs are currently remote-ready; similar analysis can be performed with other cities and other yearly points. This pipeline can be utilized for other cities, given job activity over year data; therefore, this model is easily transferable and applicable for macroeconomic research and future initiatives for economic prediction.

	Job Activity	0	10	15	19	20	21	24	27
0	Mining, logging, construction	0.052833	0.045455	0.052354	0.060158	0.062590	0.062058	0.073161	0.084087
1	Manufacturing	0.080261	0.067856	0.066356	0.066272	0.067943	0.067718	0.070347	0.073850
2	Trade, transportation, and utilities	0.243031	0.204654	0.199269	0.189546	0.189006	0.190216	0.188777	0.189857
3	Information	0.034397	0.024358	0.023539	0.020710	0.020383	0.019810	0.019589	0.019391
4	Financial activities	0.080486	0.088082	0.085633	0.090730	0.093679	0.089145	0.092116	0.093060
5	Professional and business services	0.135791	0.138104	0.149351	0.144181	0.145975	0.145341	0.146680	0.147020
6	Education and health services	0.124101	0.155502	0.154424	0.157199	0.160593	0.160906	0.156859	0.152902
7	Leisure and hospitality	0.092401	0.095259	0.098214	0.102564	0.089150	0.096018	0.093502	0.091126
8	Other services	0.032374	0.038712	0.037135	0.036686	0.036442	0.036992	0.034439	0.032092
9	Government	0.124326	0.142018	0.133726	0.131953	0.134239	0.131797	0.124530	0.116615

Figure 4: Omaha Predicted Values (2024, 2027) from 2000-2021 data

0	1	2	3	4	5	6	7	8	9	10	11	12	13
Education, training and library	Legal	Business and financial operations	Management	Arts, design, entertainment, sports and media	Office and administrative	Architecture and engineering	Life, physical and social science	Community and social service	Sales and related	Personal care and service	Protective service	Healthcare practitioners and technical	Transportation and material moving
98.00%	97.00%	88.00%	87.00%	76.00%	65.00%	61.00%	54.00%	37.00%	28.00%	26.00%	6.00%	5.00%	3.00%

Figure 5: Remote-Ready Job Estimated Percentages

Looking at these images, computing the dot-product of the values of Figure 4 and values of Figure 5 will provide the percentage of workers whose jobs could be remote-ready. The generalized model M is (with P being predicted values and RRP being remote-ready percents):

$$M(P_i, RRP_i) = \Sigma_{i=1}^{10}(P_i) \cdot \Sigma_{i=1}^{10}(RRP_i); P_i = \hat{y}_i(2024) = a_i x^2 + b_i x + c_i \tag{2}$$

The percentages for Seattle, Scranton, Liverpool, and Barry for jobs that are remote-ready, based on the model, in 2024 are **44.2%, 44.3%, 21.4%, and 46.8%** respectively.

The percentages for Seattle, Scranton, Omaha, Liverpool, and Barry for jobs that are remote-ready in 2027 are **43.1%, 42.9%, 48.0%, 21.3%, and 46.6%** respectively.

The resulting code for Problem 1, as well as the other two, can be found in the project Github:

https://github.com/karthikm15/MathWorks-Modeling-Competition-2022. This contains the code, visualizations, and tables that were used to produce this paper.

2 Q2: Remote Control

2.1 Inputs

The following inputs can be used for data about an individual worker to determine whether they are willing and will be allowed to work at their job:

- **Age:** The age of the individual, blocked into the following categories (16-24, 25-54, 55-64, 65+).
- **Educational Level:** The educational level based on the NQF (National Qualifications Framework), blocked into the following categories (less than high school diploma, high school, some college, bachelor's degree, post-graduate degree).
- **Location:** The location of work (currently includes Seattle, Omaha, Scranton, Liverpool, and Wales).
- **Job:** The type of job, blocked into the following categories (mining, logging, and construction; manufacturing; trade, transportation, and utilities; information; financial activities; professional and business services; education and health services; leisure and hospitality; other services; government).
- **Individual Opinion:** What the individual opinion about the problem is, blocking into the following categories (exclusively remote, hybrid, in-person).
- **Date:** The date that the person is looking for the job, in the format MM/YYYY (US) and YYYY (UK).

2.2 Assumptions

To simplify the model, the following assumptions were made:

1. *The model assumes a hybrid and remote working-from-home model and does not assume differentiation between the two.* The model assumes a binary classification of in-person or not in-person.

2. *The aforementioned five inputs (age, educational level, location, job, individual opinion) are all independent from one another.* Two events are independent if the occurrence of one does not affect the probability of the occurrence of the other. Thus, for a simple case of two events:
$$P(A \cap B) = P(A)P(B) \tag{3}$$
and for a more complex case of n events:
$$P(A_1 \cap A_2 \dots A_n) = P(A_1)P(A_2)\dots P(A_n) \tag{4}$$

3. *The data provided in the challenge (MathWorks Math Modeling Challenge 2022) is accurate and substantively rigorous for ensuring the validity of our model.* Inaccurate data will provide a skewing of results, as the model heavily relies on probabilistic estimates derived from challenge data.

4. *The complexity of a probability model with six priors requires simplification based on time constraints.* Note that a complete probability model is P (remote | age, educational level, location, job, individual opinion, date). However, this model is unnecessarily complex, since even for $P(A \mid B, C)$:
$$P(A \mid B, C) = \frac{P(B|A, C)P(A|C)}{P(B|C)} \tag{5}$$
Therefore we analyze these separately and consider P(remote | age), P(remote | educational level) . . ., P(remote | date).

5. *A weighted percentage is calculated based on the six percentages calculated from Assumption 4 based on arbitrary weights.* The weights, chosen based on prior schema and research is as follows: age (10%), educational level (10%), location (10%), job (40%), individual opinion (15%), date (15%).

2.3 Model

We aim to use a probabilistic model as the primary method of determining an individual's remote status. Consider an individual in the following categories (Age: 25-54, Educational

Level: bachelor's degree, Location: Omaha, Job: information, Individual Opinion: exclusively remote, Date: 03/2021).

- P(remote | 25-54) = 2051/17176 = 11.94% - These values are taken from D3.
- P(remote | bachelor's degree) = 15226/37539 = 40.76% - These are taken from D3.
- P(remote | Omaha, information) = 2% - Lack of data presumes that we must use multiple priors to calculate the probability of remote given the individual plans to work in Omaha. See the Q1 regression model and corresponding figure for details of this calculation.
- P(remote | information) = 72% (see 1.2.4 for information about this percentage)
- P(remote | exclusively remote) = 37% - data was taken from here.
- P(remote | 03/2021) = 31% + 17% = 48% - These values are taken from D4.

We then calculate the weighted percentage:

$$(0.1*0.1194)+(0.1*0.4076)+(0.1*0.02)+(0.4*0.72)+(0.15*0.37)+(0.15*0.48) = 0.4702 \quad (6)$$

Therefore, the probability that given the above assumptions, the probability a person will work remotely from home is **47.02%**.

2.3.1 Age

Given data from D3, we note that the overall percentage that work from home is 25.7% with the following breakdowns for age: 16-24 (11.94%), 25-54 (28.84%), 55-64(24.8%), and 65+ (22.28%). We can discover almost no relationship between age and percentage of remote workers from this analysis (besides a sharp drop of 16-24 because of minimum-wage jobs).

We go deeper into the analysis by using cities as an intermediary; first, the a row-wise percentage calculation on 2021 values is calculated to determine which city relies most heavily on a particular job. Then, using open-source Bureau of Labor data and past analysis, the percentage of remote work was calculated: mining, logging, and construction (19%); manufacturing (1%); trade (3%), transportation, and utilities; information (72%); financial activities (76%); professional and business services (80%); education and health services (83%); leisure and hospitality (71%); other services (33%); government (35%).

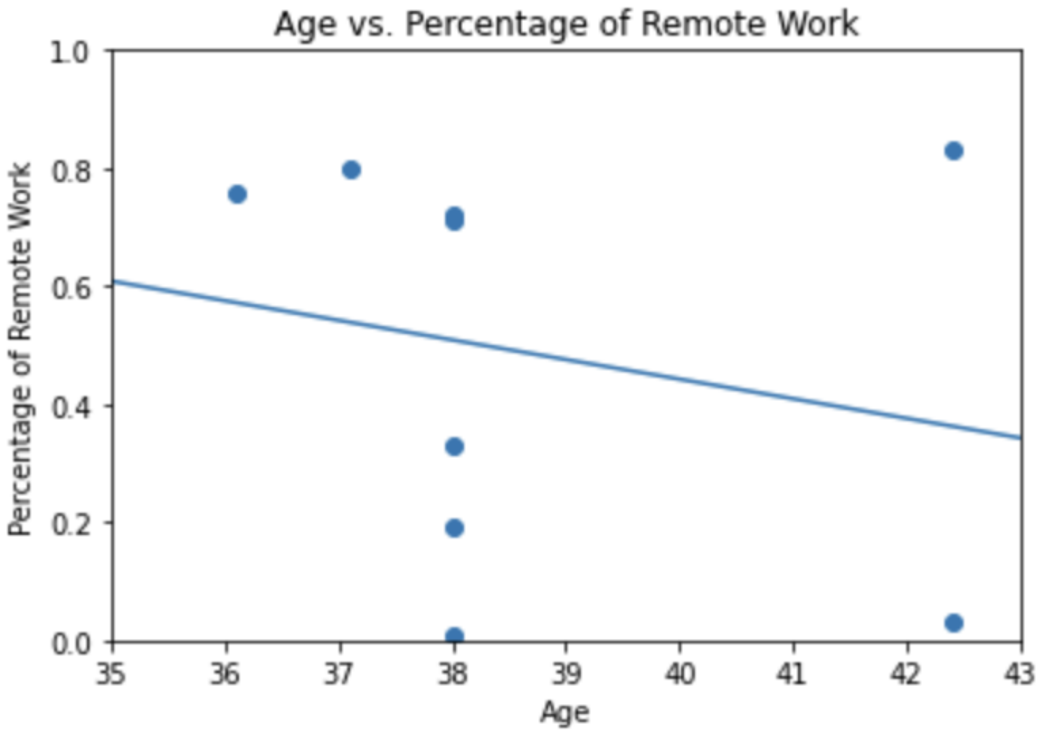


Figure 6: Analysis of Age on Percentage of Remote Individuals

Running an analysis, we note that the percentage of remote work and age have a slight negative relationship with another using polynomial regression with degree one as seen in Figure 6.

2.3.2 Educational Level

A similar analysis was run with educational level. Given data from D3, the following breakdowns for educational level are these: less than high school diploma (3.32%), high school (8.78%), some college (16.9%), bachelor's degree (40.56%), post-graduate degree (54.4%). We can infer a positive relationship which is supported by a graphical analysis as seen on Figure 7.

2.3.3 Month and Year

To account for global trends induced by the COVID-19 pandemic that began in early 2019 and started to spike in late 2020, a regression model was used on United States and United

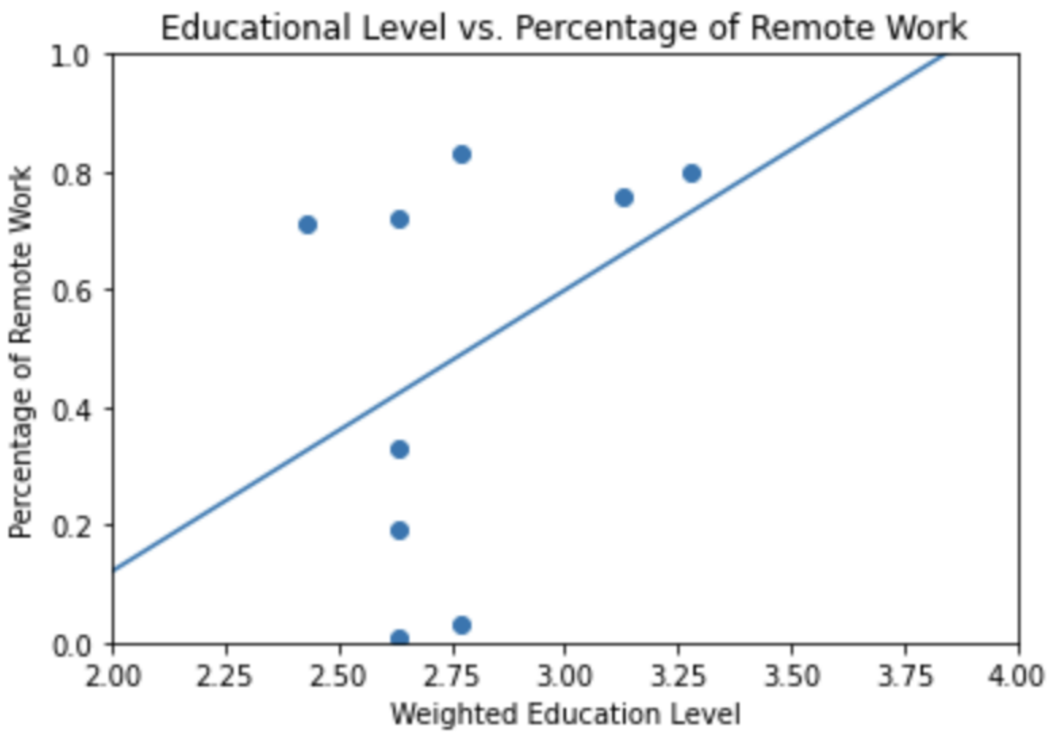


Figure 7: Analysis of Educational Level on Percentage of Remote Individuals

Kingdom workers. The workers were blocked into two categories of part and full time remote workers.

The coefficient of determination or R^2 is:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \tag{7}$$

United Kingdom Polynomial Regression:

- Applying a polynomial regression with a power of 5 to Figure 8, it is revealed that the amount of total remote workers is positively correlated with time, with a coefficient of determination of approximately 0.93.
- Applying a polynomial regression with a power of 5 to Figure 9, it is revealed that the amount of part time remote workers is positively correlated with time, with a coefficient of determination of approximately 0.81.
- Applying a polynomial regression with a power of 5 to Figure 10, it is revealed that

the amount of exclusively remote workers is positively correlated with time, with a coefficient of determination of approximately 0.99.

Because the coefficient of determination for all of the figures are well above 0.70, the threshold for a strong correlation, it can safely be assumed that the polynomial regression model with a power of 5 accurately represents the data in the model, and that both blocks of part time and exclusively remote workers exhibit a positive correlation with time.

United States Polynomial Regression:

- Applying a polynomial regression with a power of 5 to Figure 11, it is revealed that the amount of total remote workers is negatively correlated with time, with a coefficient of determination of approximately 0.92.
- Applying a polynomial regression with a power of 5 to Figure 12, it is revealed that the amount of part time remote workers is positively correlated with time, with a coefficient of determination of approximately 0.43.
- Applying a polynomial regression with a power of 5 to Figure 13, it is revealed that the amount of exclusively remote workers is negatively correlated with time, with a coefficient of determination of approximately 0.93.

Because the coefficient of determination for Figure 8 and Figure 10 are well above 0.70, the threshold for a strong correlation, it can safely be assumed that the polynomial regression model with a power of 5 accurately represents the data in the model, and both the exclusive workers and remote workers exhibit a negative correlation with time. Despite the coefficient of determination of Figure 9 being below 0.70, the coefficient of determination of Figure 9 still is above 0.40, indicating a moderate positive correlation between the part time remote workers and time.

3 Q3: Just a Little Home-work

Seattle, Washington

From Census Reporter in 2021, we observe that the people living in Seattle experienced an average commute time of 31.6 hours, higher than any other city. With the largest travel time to get to work, individuals are likely to change to remote working with the advantage of not spending hours on the road and wasting energy.

According to D2, Seattle experienced the third highest percent of people who worked at home with 6.5%. This number is gradually increasing as more individuals are given to option to work from home. For example, with Amazon's new remote work policy, the option to work from home will come down to the team leaders. This will lead to an increase in people choosing to work remotely from home in Seattle.

Omaha, Nebraska

From D2, we notice that Omaha has the lowest median age of 36.1 and the highest percentage of some college education of 31%; thus, the Omaha population is young and slightly experienced. A young and experienced population is likely to have more liberal ideals, and will be more receptive to change than more conservative ideals to remote working.

From 1.3.1, a negative relationship between age and percentage of remote work is found, and since Omaha's age group is relatively young, this explains why nearly 49% of workers (a high percentage compared to the other four cities) will be remote ready by 2024.

According to Livability's 2016 list of Best Cities for Entrepreneurs, Omaha is booming in high-growth startups within e-commerce, financial services, and software-as-a-service companies, all of which fall under the computer and mathematical occupation category (100% can be remote as seen in D2). Paired with its low cost of living, Omaha is a clear contender for disruption in the remote industry.

Scranton, Pennsylvania

From D2, we notice that Scranton has a high percentage of the population below 20 years old, with 22.5% of the population below 20 years old. Additionally, Scranton has one of the highest percentages of high school education with 39%. Thus the population of Scranton, Pennsylvania is relatively young and educated. A young population is more likely to have more liberal ideals, and will be more receptive to changes in working conditions.

From 1.3.1, a negative relationship between age and percentage of remote work is found, and since Scranton's age group is relatively young, this explains why nearly 49% of workers (a high percentage comparing to the other four cities) will be remote ready by 2024.

According to Indeed's job listings for Scranton, Pennsylvania, jobs have been shifting from low pay, sales workers to more white collar work such as Business and Financial Operations and Management, both of which are almost 90% remote, as seen in D2. This explains why Scranton, Pennsylvania, is predicted by the model to be one of the leaders in remote work.

Barry, Wales

Barry has the lowest education level out of the five cities with 25.2% of its citizens not having a high school diploma. From 1.3.2, due to the negative relationship between education level and percentage of remote workers, it makes sense that a lower education level would elicit higher amounts of remote workers.

The higher level of remote workers is explained by the Welsh government's ambition to "see around 30% of Welsh workers working from home or near from home, including after the threat of COVID-19 lessens" due to reduced urban congestion and travel times for workers. Wales additionally has fewer roads than larger hubs like Seattle or Liverpool, so they'll likely be less affected by changes in businesses.

Liverpool, England

Liverpool has the highest percentage of the population with a high school education out of all five cities with 39% of its citizens having a high school diploma. From 1.3.2, a negative correlation was found between education level and percentage of remote workers. Since Liverpool has a large population with a high school education, it makes sense that the city will have the lowest prediction of remote workers of 21.3% by 2024.

According to Indeed's job listings for Liverpool, England, many of the jobs in the area are related to sales and other related store jobs, many of which cannot be worked remotely. According to data from D3, only approximately 28% of sales and related work is able to be replaced remotely, explaining why only around 21.3% of the population, the lowest out of the five cities, will be remote ready by 2024.

3.1 Results

From our previous analysis, we can determine the magnitude of impact that remote work will have on each of the cities in 2024 and 2027 (from most to least impact): Barry, Omaha, Scranton, Seattle, Liverpool.

Barry's move towards remote work is heavily supported by government regulations, allowing them to achieve the highest predicted growth rates towards remote work. Omaha's small and young population size also allows for a swift move towards virtual work as high-growth startups flock to start new financial and software services, jobs that are able to be done remotely.

Scranton is a transitioning economy that is shifting slowly to higher skilled jobs simultaneously with the transition to remote work, allowing them an edge in moving in-person to remote jobs. Seattle's long commute hours prompts a young population to adapt and move toward remote work so that they can save time and energy. Lastly, Liverpool's shift towards sales jobs and large population indicates a need for in-person work.

4 Further Research

Integrating employment data from years before 2000 will increase the accuracy of our polynomial regression model as it will have more points to use for its gradient descent optimization function. Further, testing out different n-degree functions (cubic, quartic, etc.) and determining their subsequent residuals, we can reduce testing set error. Note that higher degrees would over fit the data, decreasing the accuracy of future predictions; therefore, when researching complex models, there should be a trade off between generalization and accuracy.

Lastly, understanding remote conditions beyond just these five cities will broaden the impact of our mathematical pipeline; by verifying this model with experts and gathering more city data, we'll further the validity of this paper. The cities outlined above are based in first-world countries; however, it will be interesting to test model performance in third world-countries, which are facing debilitating conditions in light of COVID.

5 Appendix

5.1 Figures

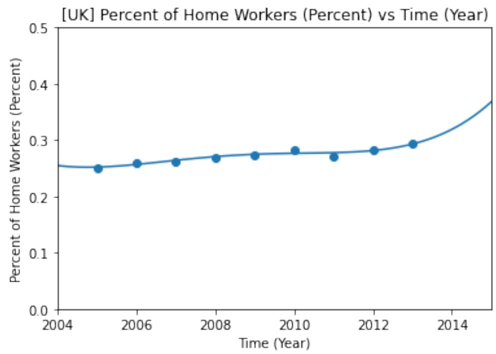


Figure 8: United Kingdom Remote Workers versus Time

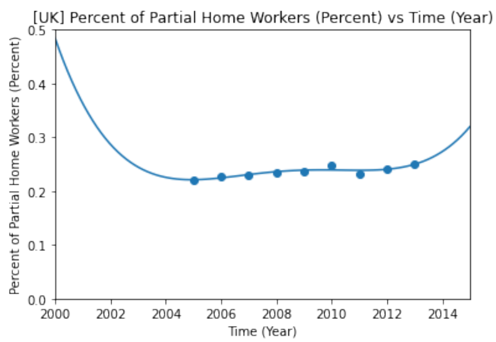


Figure 9: United Kingdom Partial Time Remote Workers versus Time

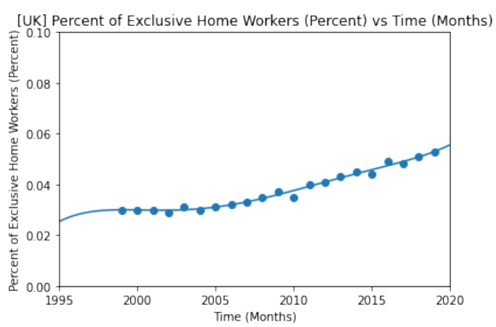
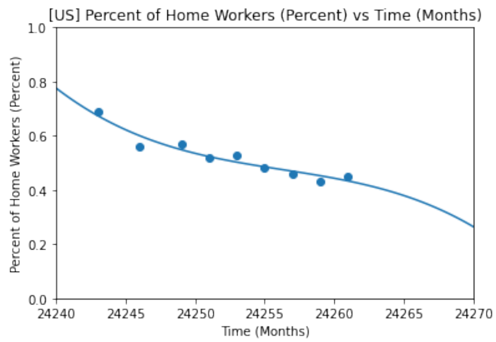
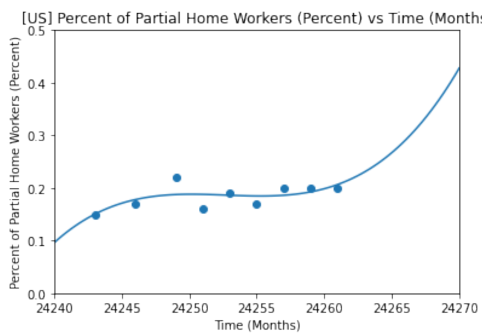


Figure 10: United Kingdom Exclusive Remote Workers versus Time



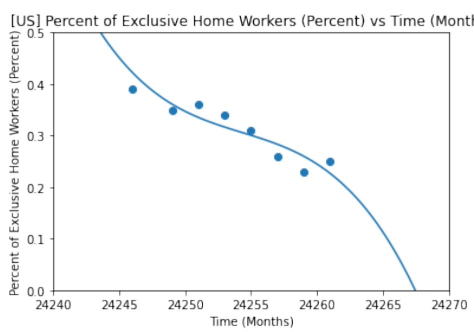
10cm

Figure 11: United States Remote Workers versus Time



.5

Figure 12: United States Part-Time Remote Workers versus Time



.5

Figure 13: United States Exclusive Remote Workers versus Time

```python
for city in range(len(employment_2000_2021)):
    fig, (ax1, ax2, ax3, ax4, ax5, ax6, ax7, ax8, ax9, ax10) = plt.subplots(1, 10, figsize=(15, 15))
    axes = [ax1, ax2, ax3, ax4, ax5, ax6, ax7, ax8, ax9, ax10]

    fig.set_size_inches(50,20)

    column_2024 = []
    column_2027 = []
    for professions in (range(len(employment_2000_2021[city].T.columns))):
        if (professions != "Job Activity"):
            x_train = np.array(rows_2000_2021[city]).reshape(-1, 1)
            y_train = np.array(employment_2000_2021[city].T[professions].values[1:]).reshape(-1, 1)
            naming = employment_2000_2021[city].T[professions].values[0]

            poly_reg = PolynomialFeatures(degree=2)
            X_poly = poly_reg.fit_transform(x_train)
            pol_reg = LinearRegression()
            pol_reg.fit(X_poly, y_train)

            correct_axes = axes[professions]

            correct_axes.scatter(x_train, y_train, color='red')
            correct_axes.plot(x_train, pol_reg.predict(poly_reg.fit_transform(x_train)), 'tab:orange')

            if (naming == "Mining, logging, construction"): correct_axes.set_title(cities[city] + " Mining")
            elif (naming == "Trade, transportation, utilities"): correct_axes.set_title(cities[city] + " Trade")
            elif (naming == "Financial activities"): correct_axes.set_title(cities[city] + " Finance")
            elif (naming == "Professional and business services"): correct_axes.set_title(cities[city] + " Business")
            elif (naming == "Education and health services"): correct_axes.set_title(cities[city] + " Education")
            elif (naming == "Leisure and hospitality"): correct_axes.set_title(cities[city] + " Leisure")
            elif (naming == "Other Services"): correct_axes.set_title(cities[city] + " Other")
            else: correct_axes.set_title(cities[city] + naming)

            column_2024.append(pol_reg.predict(poly_reg.fit_transform([[24]]))[0][0])
            column_2027.append(pol_reg.predict(poly_reg.fit_transform([[27]]))[0][0])
            # axs[city,professions].xlabel('Year')
            # axs[city,professions].ylabel('Percentage')

    print(column_2024)
    employment_2000_2021[city]['24'] = column_2024
    employment_2000_2021[city]['27'] = column_2027
    plt.tight_layout()
```

Figure 14: Code for Visualizing City Employment Data (Q1)

```python
#Create Parttime Column for workers only working part time from home
pt = df['Percentage of workers who worked from home partially.']
for i in range(len(pt)):
  j = float(pt[i][:-1])/100
  pt[i] = j
df['PartTime'] = pt

partialModel = numpy.poly1d(numpy.polyfit(list(df['MonthYear']), list(df['PartTime']), 8))
partialModelLine = numpy.linspace(24240, 24270, 100);
plt.scatter(list(df['MonthYear']), list(df['PartTime']))
plt.plot(partialModelLine, partialModel(partialModelLine))
plt.xlim((24240, 24270))
plt.ylim((0, 0.5))
plt.title("[US] Percent of Partial Home Workers (Percent) vs Time (Months)")
plt.ylabel("Percent of Partial Home Workers (Percent)")
plt.xlabel("Time (Months)")
plt.show()
ptR2 = r2_score(list(df['PartTime']), partialModel(list(df['MonthYear'])))
print(ptR2)
```

Figure 15: Code for Understanding Part-Time Work (Q2)