MathWorks Math Modeling Challenge 2022

New Trier Township High School

Team #15522, Winnetka, Illinois

Coach: Bradley Kuklis

Students: Nika Chuzhoy, Nathan Liu, Aruni Chenxi, Max Hartman, Connor Lane

M3 Challenge RUNNER UP—$15,000 Team Prize

JUDGE COMMENTS

Specifically for Team # 15522 —Submitted at the Close of Triage Judging:

COMMENT 1: The assumptions were nicely explained and the sensitivity analysis was a nice in conclusion. Try spending a little more time in the discussions of each model especially where the result may be counterintuitive, for example with the negative percent changes in question 1. The comparison of the overall percentages and actual number of workers in question 3 was very nice.

COMMENT 2: Great job answering all parts of the Challenge! Not an easy feat! I especially appreciated the writing. I appreciated knowing your thought process.

COMMENT 3: Your paper is very well done. You provide a strong justification for the use of linear regression in the first part of the problem as well as for the use of weights in the second part and the need to adapt the Bass Diffusion Model in the third part. All parts have a strong section on strengths and weaknesses and a thorough sensitivity analysis. In the third part, the sensitivity analysis would benefit from a. written analysis of the results in addition to the table. Overall, your paper has a great level of detail- there is enough to follow all of your work but you (appropriately) skip over the nitty-gritty of your calculations.

COMMENT 4: Good leverage of data to drive model development, and sensitivity analysis was great to see.

Remote Work: Fad or Future

Team 15522

1 Executive Summary

For much of the world's white collar labor force, remote work has been a fact of life throughout most of 2020 and 2021. However, as the COVID-19 pandemic recedes, many workers are trickling back into offices, leaving employers and employees alike to wonder about the future of remote work. Companies must carefully weigh worker satisfaction, company culture, productivity, and safety to determine whether to ask workers to telecommute or work in person [1].

To begin the process of estimating the prevalence of remote work in the following years, we first determined the maximum number of jobs that could theoretically be done remotely. We called these jobs "remote-ready" and calculated the proportion of jobs that are remote-ready for a select group of cities (Seattle, WA; Omaha, NE; Scranton, PA; Liverpool, England; and Barry, Wales) for the years 2022, 2024, and 2027. To do so, we predicted the growth of each industrial sector in each city that comprises its total labor force, based on data from the Bureau of Labor Statistics and the UK Office of National Statistics. Then we multiplied the number of employees in each sector by the proportion that the National Bureau for Economic Research estimates can be done remotely, and considered the sum to be the total number of remote-ready jobs in that city.

However, it is important to keep in mind that not all remote-ready job positions have been done or will be done remotely. Employers cite the idea that "proximity boosts productivity" and the difficulty of encouraging collaboration [2] as reasons for bringing workers back to the office. Employees, on the other hand, seem to prefer working from home (WFH) [3]; 76% of both UK citizens and Americans want to WFH at least one day a week even after the pandemic is over. To estimate the probability of an individual wanting to WFH, we considered six attributes (gender, age, household income, education level, custody of children, and commute time) and weighted them according to importance. Then we combined that with the probability that their industry would allow the individual to WFH to determine one's final probability to actually WFH. For example, we found that a 45-year-old man who works in government, has a medium household income, is well educated, is a parent, and has a 50-minute commute has a 40% likelihood of actually working from home post-pandemic.

Finally, we used our previous predictions in order to calculate the proportion of people in our group of cities that will be working remotely in 2024 and 2027. To do so, we combined our projections for the growth of various industries with an estimate of how the fraction of people working remotely in each industry will increase over time. We arrived at this estimate by modifying an existing model for the adoption of new technology in a way that took into account the effects of the pandemic. Then, after accounting for the positive impacts to children of WFH parents [21] and the commute time saved, we ranked the cities by magnitude of impact that remote work would have on it in 2027: Seattle, WA in first; Omaha, NE in second; Liverpool, England in third; and Barry, Wales and Scranton, PA tied for fourth. Our forecasts about the future of remote work may be used to guide policy decisions regarding the political encouragement or curbing of remote work.

Contents

2 Part I: Ready or Not

2.1 Restatement of the Problem

Consider the following cities:

- Seattle, WA
- Omaha, NE
- Scranton, PA
- Liverpool, England
- Barry, Wales

Create a model to estimate the percentage of workers whose jobs are currently remote-ready. Then apply the model to the cities below to make predictions for the percentage of remote-ready jobs in 2024 and 2027, while accounting for how the inputs to the model will change over time.

2.2 Assumptions

- *"Remote-ready" jobs are defined as the set of occupations that could theoretically be done from home without significant loss in efficiency.*
- *Technological advancements will not significantly increase the proportion of remote-ready jobs.* Although we recognize that the proliferation of new technologies—such as virtual reality and fiber internet—may increase the proportion of remote-ready jobs in the future, we do not believe that they will develop fast enough within the next five years to have a substantial effect.
- *The number of employees in an industry increases or decreases linearly.* In the long term, the number of people employed in an industry generally changes according to an exponential curve, in accordance with population growth. However, we are only given 20 years of data and are tasked with considering the next 5 years. Because the data we are given is very short-term in comparison to industry patterns, which usually takes decades to change [4], we choose a linear regression over any other model to avoid over-fitting.
- *The percentage of employees in a specific industry who are working remotely is independent of the city that they work in.* The location that an employee works at does not influence the odds that they will work remotely. Rather, the odds depend on company policy, worldwide events, and technological advances.
- *The entire workforce of each city is the sum of the number of people employed in the following industries:*

- *Mining, logging, construction (MCL)*
- *Manufacturing*
- *Trade, transportation, and utilities (TTU)*
- *Information*
- *Financial activities (Financial)*
- *Professional and business services (ProfessionalBusiness)*
- *Education and health services*
- *Leisure and hospitality*
- *Government*
- *Other services*

This list of industries is comprehensive, and thus encompasses the ability of all jobs to be all industries in terms of remote-work ability. It also includes a category for other classifications, which guarantees that all industry types are being considered.

2.3 Model Development

Our model hinges upon the prediction of industry growth in the given cities. Broadly, as we assume that the fraction of remote-ready jobs in each industry is constant over the next five years, we can compute the future proportion of remote-ready jobs in a city by considering how the proportion of the city's population employed in different industries changes.

Our model predicts the total number of employees in a diverse group of industries in each city in 2022, 2024, and 2027 by applying a linear regression (using the Python library Scikit-learn) to past data compiled by the Bureau of Labor Statistics and the UK Office of National Statistics [5][6]. We choose to use a linear regression over any other fit because the time frame of the data we are given (the past 20 years) and the time frame we are predicting (the following 5 years) is sufficiently short enough to put any curved fit at risk of over-fitting. Essentially, we are applying the principal that any curve viewed at a small enough scale appears to be linear.

From there we multiply the number of people employed by each industry in each city by the percentage of jobs that can be done remotely, as specified by the University of Chicago white paper "How Many Jobs Can be Done at Home?" by Jonathan I. Dingel and Brent Neiman [7]. This calculates the number of remote workers in each industry of each city. The following equation was used to calculate the percentage of remote-ready jobs in each city in 2022, 2024, and

2027:

$$P_{RemoteReady} = \frac{\sum N_\alpha}{\sum N_\beta}$$
$$N_\alpha = RemoteEmployees_{Industry}$$
$$N_\beta = TotalEmployees_{Industry}$$

The calculated percentages of remote employees in 2024 and 2027 are then used to compute a percent change for each city.

$$P_{Change} = \left(\frac{P_\rho - P_\epsilon}{P_\epsilon}\right) * 100$$
$$P_\rho = PercentRemote_{2027}$$
$$P_\epsilon = PercentRemote_{2024}$$

This helps to visualize the change in remote workers from 2024 to 2027.

2.4 Results

We were able to compute the portion of jobs that will be remote-workable in Seattle, Omaha, Scranton, Liverpool, and Barry over the next two and five years.

Table 1: Proportion of Remote-Workable Jobs

Year	Seattle, WA	Omaha, NE	Scranton, PA	Liverpool, UK	Barry, Wales
2022	60.801%	60.279%	56.332%	46.394%	64.437%
2024	60.799%	60.324%	56.172%	46.353%	64.519%
2027	60.797%	60.390%	55.930%	46.295%	64.641%
%Δ from 2024 to 2027	-0.0033%	0.1094%	-0.4308%	-0.1251%	0.1891%

As shown, we did not find that the fraction of remote-workable jobs changed significantly over time, likely due to the short length of the time span.

2.5 Sensitivity Analysis

By adjusting the theoretical maximum percentage of work that can be done remotely in each industry by ±10%, the changes in future remote-work percentages of a city are altered. The following displays the alternative percent change between the proportion of people working remotely in each city in 2024 and 2027.

Table 2: Sensitivity Analysis: Percent Change in Remote Workers from 2024–27 in Seattle, Omaha, Scranton, Liverpool, and Barry

%Δ in Remote Workers, 2024 to 2027		Seattle, WA	Omaha, NE	Scranton, PA	Liverpool, UK	Barry, Wales
MCL	+10%	-0.0049%	0.1094%	-0.4307%	-0.1271%	0.1875%
	-10%	-0.0049%	0.1078%	-0.4309%	-0.1231%	0.1876%
Manufacturing	+10%	-0.0457%	0.0939%	-0.5149%	-0.0784%	0.1769%
	-10%	-0.0381%	0.1234%	0.3451%	-0.1736%	0.1983%
TTU	+10%	-0.0016%	0.0722%	-0.3889%	-0.0616%	0.1858%
	-10%	-0.0066%	0.1457%	-0.4759%	-0.1886%	0.1908%
Informational	+10%	N/A*	N/A	N/A	N/A	N/A
	-10%	-0.0530%	0.1380%	-0.3832%	-0.1323%	0.2176%
Financial	+10%	N/A	N/A	N/A	N/A	N/A
	-10%	0.0215%	0.0908%	-0.4345%	-0.1151%	0.1721%
Business Services	+10%	0.0322%	0.1252%	-0.3820%	-0.1388%	0.2123%
	-10%	-0.0436%	0.0929%	-0.4831%	-0.1089%	0.1620%
Education & Health	+10%	0.0162%	0.1421%	-0.3786%	-0.1484%	0.1954%
	-10%	-0.0266%	0.0723%	-0.4834%	-0.1017%	0.1810%
Leisure & Hospitality	+10%	-0.0097%	0.1102%	-0.4259%	-0.1452%	0.1713%
	-10%	0.0016%	0.1053%	-0.4358%	-0.1046%	0.20578%
Other Services	+10%	-0.0049%	0.1091%	-0.4371%	-0.1392%	0.1855%
	-10%	-0.0049%	0.1063%	-0.4245%	-0.1108%	0.1912%
Government	+10%	N/A	N/A	N/A	N/A	N/A
	-10%	0.0352%	0.1033%	-0.4137%	-0.0889%	0.1752%

*N/A indicates that increasing this variable by 10% would cause the fraction of theoretical remote workers in a given industry to be over 100%.

For the most part, the difference between 2024 and 2027 does not change significantly when the theoretical maximum of just one industry is changed. It largely depends on the prevalence of that industry in the considered city. Barry, Wales: Business Services; Liverpool, UK: TTU; Scranton, PA: Manufacturing; Omaha, NE: TTU; and Seattle, WA: Informational.

Over the past 20 years, the world has seen a steady increase in remote work. This was amplified by the COVID-19 pandemic, causing a sharp adoption of remote working. As a result, it makes intuitive sense that our sensitivity analysis displays some cases where the percentage of a city's remote workforce will decrease.

2.6 Strengths and Weaknesses

By taking into account a comprehensive list of industries, our model nearly eliminates the possibility of industries being erroneously unaccounted for. Our model also gives us the flexibility to apply it to additional cities, given only the prior employment data for each industry that composes its labor force. Furthermore, our sensitivity analysis shows that the results from our model are not dramatically affected by changes in the theoretical fraction of an individual industry that can be done remotely.

The model is also adaptable in the sense that observed changes in the future, such as a large increase in the number of remote-workable jobs due to the growth of new technology like virtual reality, can be re-factored into the model by increasing our constants of the theoretical fraction of each industry that could work from home.

Our model is based on linear growth over a short time period, and thus would not adapt well to more complex scenarios, such as a rapid exponential change in an industry, or under any long time periods. While our model assumes the number of workers increases linearly in each industry, new, disruptive technologies in areas such as software and automation have the potential to cause sudden, exponential growth in some industries and decay in others. This would cause non-linear growth in employees in these industries and cannot be accounted for immediately in our model.

Our model only uses a handful of data points to make future predictions. Small amounts of data mean that the future values we calculated for 2024 and 2027 are not necessarily accurate—they can only serve as rough estimates. This also means that we did not have enough data to apply more complex models, such as a higher-degree regression.

3 Part II: Remote Control

3.1 Restatement of the Problem

Create a model that predicts whether an individual worker whose job is remote-ready will be allowed to and will choose to work from home.

3.2 Assumptions

1. *Being permitted to work from home is only dependent on a worker's industry and independent of demographics.* Theoretically, employers in a given industry will not differentiate between employees of different demographics. While we acknowledge that a person's status within an industry may be itself dependent on demographics, we overlook this nuance so as not to over-complicate the model.

2. *Wanting to work from home rests entirely on six attributes: age, household income, education level, custody of children, and commute time.* These six attributes cover many of the relevant conditions that could impact desire for at-home work.

3. *The six factors are weighted unequally.* As two of the factors, namely, commute time and having custody of children, are much more frequently [9] reported to be the main reasons to prefer telecommuting, they are accordingly weighted more heavily.

3.3 Model Development

We separate our model into two parts. First, we compute the probability that an employer will allow their employee to work from home. This is considered to be independent of the demographics of the worker. Then we compute the probability that a worker with certain demographic traits would like to work from home.

To calculate the probability that an employee will be allowed to work remotely, we simply consider the industry within which they are employed. By dividing the number of remote workers in late 2020 [5] in that industry by the theoretical maximum number of remote workers [11], we arrive at the probability that an individual within that industry would be allowed to work remotely, although they may not prefer to do so in the current late-pandemic environment.

To calculate the chance that an individual will prefer to work remotely, we would like to compute probabilities of the following form, where R is the event that someone is interested in remote work:

$$\begin{aligned} P(R|&\text{Gender} \wedge \text{Age} = a \wedge \text{Income} \\ &\wedge \text{Education Level} \wedge \text{Has Children} \\ &\wedge \text{Commute Time}) \end{aligned}$$

However, all we have is data of the following form:

$$\begin{aligned} &P(R|\text{Gender}) \\ &P(R|\text{Age}) \\ &P(R|\text{Income}) \\ &P(R|\text{Education Level}) \\ &P(R|\text{Has Children}) \\ &P(R|\text{Commute Time}) \end{aligned}$$

This presents a problem, as we cannot find a mathematically sound way to obtain the combined conditionals from the individual conditionals. This is not the same as having the ANDs on the left side of the conditional, which are relatively easy to handle. Even in a very simplified case, where we are given $P(X|Y), P(X|\neg Y), P(X|Z)$, and $P(X|\neg Z)$ and need to find $P(X|Y \wedge Z)$, we cannot find the solution.

Our best attempt involves writing equations like the following (with constants that we can find in datasets shown in red):

$$\color{red}{P(X|Y)P(Y)} \color{black}{= P(X|Y \wedge Z)}\color{red}{P(Y \wedge Z)} \color{black}{+ P(X|Y \wedge \neg Z)}\color{red}{P(Y \wedge \neg Z)}$$

While we can find four equations and four unknowns, this system is not solvable, as one equation would be a linear combination of the other three. As increasing the number of conditions only worsens the problem, we are forced to abandon attempts at computing the probabilities exactly.

Instead, we approximate the data using a weighted average. We use the names $C_1, C_2, C_3, C_4, C_5, C_6$ as shorthand for our six conditions:

$$P(R|C_1 \wedge C_2 \wedge C_3 \wedge C_4 \wedge C_5 \wedge C_6) = \frac{\sum_{i=1}^{6} \alpha_i P(R|C_i)}{\sum_{i=1}^{6} \alpha_i}.$$

We have chosen to select $\alpha_i = 2$ for having children and commute times, and $\alpha_i = 1$ for all other conditions. This is because children and commute times are the most influential aspects of whether or not an individual would like to stay home [9][19].

We used a previously conducted survey to get the values of $P(R|C_i)$ for each of our conditions, and it gave us this data:

Table 3: Data for Probability of Liking Remote Work

Gender	Age	Income	Education	Children	Commute Time
Male: 0.64	$a < 35 : 0.67$	Low: 0.62	Low: 0.59	$Yes : 0.68$	$\leq 60 : 0.433$
Female: 0.68	$35 \leq a < 50 : 0.65$	Mid:0.63	Mid:0.61	$No : 0.63$	$60 < t \leq 120 : 0.696$
	$50 \leq s : 0.63$	High:.69	High: 0.71		$121 < t : 0.869$

3.4 Results

Using our algorithm, we can compute the probability of an individual both being allowed to work remotely or choosing to work remotely given data on their age, gender, industry, household income, level of education, whether they have children, and the length of their commute.

As calculating every possibility would be too computationally intensive, we present three examples:

1 A 45-year-old man who works in government, has medium household income, is well educated, is a parent, and has a 50-minute commute: 61% chance of wanting to work from home, 66% chance of being allowed to work from home, and a 40% chance of actually working from home.

2 A 22-year-old woman who works in information, has high household income, has medium education, is not a parent, and has a 130-minute commute: 71% chance of wanting to work from home, 83% chance of being allowed to work from home, and 59% chance of actually working from home.

3 A 30-year-old woman who works in education, has low household income, has medium education, is a parent, and has a 20-minute commute: 60% chance of wanting to work from home, 89% chance of being allowed to work from home, and 53% chance of actually working from home.

3.5 Sensitivity Analysis

We perform a sensitivity analysis on our model for individual interest in remote work based of demographic factors, and see that, in general, even relatively large changes in the coefficients does not effect the probabilities too much, with most changes on the order of 0.1%. The largest changes were caused by changes to the parent and commute time factors; however, this makes sense, as they started as large coefficients, and equal percent changes to large coefficients lead to larger changes.

Table 4: Sensitivity Analysis: %Δ of the probability of liking remote work

coefficient being changed	percentage change	p1	p2	p3
Gender coefficient	+10%	0.064%	-0.067%	0.163%
Gender coefficient	−10%	-0.065%	0.054%	-0.166%
Age coefficient	+10%	0.105%	-0.083%	0.141%
Age coefficient	−10%	-0.14%	-0.084%	-0.109%
Income coefficient	+10%	0.046%	-0.048%	0.040%
Income coefficient	−10%	-0.046%	+0.048%	-0.040%
Education coefficient	+10%	0.207%	-0.186%	0.018%
Education coefficient	−10%	-0.212%	-0.191%	-0.020%
Parent coefficient	+10%	0.248%	0.147%	0.175%
Parent coefficient	−10%	-0.303%	-0.142%	-0.175%
Commute coefficient	+10%	-0.73%	-0.511%	0.45%
Commute coefficient	−10%	-0.73%	-0.537%	-0.45%

3.6 Strengths and Weaknesses

The advantage of our approach is that we can take a wide variety of individual factors into account. However, there are also disadvantages resulting from several of the assumptions we have to make. For one, the weighted averages we computed may not accurately represent the intersections of the demographic traits. Additionally, only considering the industry category in the probability of an individual being allowed to work overlooks the large variety of jobs present within the industry—for example, the leisure and hospitality sectors both have managerial jobs that can be done remotely, and physical hospitality jobs for which a person must be present. These jobs are likely to be strongly correlated with education level and household income.

4 Part III: Just a Little Home-Work

4.1 Restatement of the Problem

We are asked to synthesize the first two models to create a model which, for a given city, estimates the percentage of workers who will actually work remotely in 2024 and 2027. We then use these predictions to rank the cities in terms of the magnitude of impact that remote work will have on the city.

4.2 Assumptions

1. *Adoption of remote work spreads through a combination of word of mouth and technological improvement.* By "word-of-mouth," we refer to the tendency of companies to imitate other members of their industry, while by "technological improvement," we refer to the increasing efficacy of remote work.

2. *The theoretical number of jobs that can reasonably be done remotely will not change significantly in the next five years.* We believe that a five-year time span is not enough time for workforce technology to change greatly, barring unusual circumstances. While we acknowledge that sudden disruptions or developments can result in change, accounting for this possibility introduces too much complexity into the model.

3. *The pandemic will last for approximately one more year.* We base this admittedly somewhat arbitrary assumption on the reports of epidemiologists [18].

4. *The largest impacts on a city will be the positive impacts of having more commute time and more time with one's children.* As not to complicate our model, we choose to only consider the two most important reasons people want to work from home: commute time and having children [9]. Having a work-from-home option brings benefits to both the mental health of the parent and allows them to better care for their children [21].

4.3 Model Development

In Part I, we predict the total number of jobs that can be done remotely in the given cities in 2024 and 2027. However, not every job that can be done remotely in the future will be done remotely. This is due to a multitude of reasons, including societal perceptions of remote work and limiting factors such as individuals lacking internet or with complicated home situations.

To predict the adoption of remote learning, we use the Bass Diffusion Model. The standard Bass Diffusion Model is used to predict the adoption of new products or technologies and is defined using the following differential equation:

$$\frac{f(t)}{1 - F(t)} = p + qF(t),$$

where $F(t)$ is the base fraction of individuals who have adopted the new technology (in this case, remote work), $f(t)$ is the change in the base fraction over time (that is, $f(t) = \frac{d}{dt}F(t)$), p is the coefficient of innovation, and q is the coefficient of imitation.

Unfortunately, this model is not perfectly suited to our purposes, because it fails to take into account the effects of the pandemic. Factors such as lockdowns and the desire of individuals to avoid contracting Covid-19 resulted in additional incentives to go remote, which we would not expect to be in place post-pandemic.

For this reason, we adapt the Bass-Diffusion model by replacing the constant "coefficient of innovation" p with a logistic curve $p_{\alpha,k}(t)$, such that $p_{\alpha,k}(t) = \frac{P_f}{1+e^{-k(t-\alpha)}}$, where α is the expected length of the pandemic, k is the rate of transition out of pandemic conditions, and p_f is the maximum coefficient of innovation. The choice to modify the coefficient of innovation in particular is motivated by the assumption that innovations related to remote work, such as improved videoconferencing software, would likely slow post-pandemic due to a decrease in consumer demand.

Our final equation is defined as follows:

$$\frac{f(t)}{1-F(t)} = \frac{P_f}{1+e^{-k(t-\alpha)}} + qF(t)$$

For the constants P_f, q, and k we use the typical values of 0.03, 0.38, and 2, respectively [17]. For α, we make the assumption that the pandemic will last for one year [18].

To implement the model, we use two sets of data: firstly, the fraction of people employed remotely in various industries in 2020 [5], and secondly, the theoretical maximum fraction of people who could be employed remotely in those industries, as compiled by the National Bureau for Economic Research [10]. Thus, for each industry, we are able to compute the initial base fraction $F(0)$ by dividing the number of remote workers in that industry by the theoretical maximum. Then we can use the Python package Odeint to solve for the fraction of remote-workable jobs that are actually remote after 2 and 5 years. Finally, by multiplying the respective fractions by the number of employees in each city's industry derived in Part I, we find the projected total number of remote workers in each city.

Then we determine the total impact by including the two most important factors considered in Q2: commute time and having a child. Using average commute times from TUC and University of Nebraska Omaha [13][14], we weight them by multiplying them by the fraction of workers expected to be remote in each city at 2027 and scale it from 0-1.

Table 5: One-Way Traffic (min)

	One-Way (min)	% in 2027	Converted to a 0-1 Scale
Seattle, WA	31.1	62	1
Omaha, NE	20.5	59	0.627
Scranton, PA	22.6	59	0.689
Liverpool, England	27.4	55	0.782
Barry, Wales	25.4	54	0.710

We factor in the positive impact of WFH parents being present for children similarly. From data collected by City Population and the United States Census Bureau [15][16], we calculate the percentage of children in each population and weight them by multiplying them by the fraction of workers expected to be remote in each city at 2027 and scale it from 0-1.

Table 6: % of Children in the Population Today vs. 2027

	% of Children in the Pop.	% in 2027	Converted to a 0-1 Scale
Seattle, WA	15.0	62	0.628
Omaha, NE	25.1	59	1
Scranton, PA	20.5	59	0.818
Liverpool, England	19.7	55	0.730
Barry, Wales	21.8	54	0.797

Finally, we add the Commute Time 0-1 score to the Children 0-1 score to determine the final impact score for each city and rank them.

4.4 Results

Overall, we find that the city with the highest proportion of remote workers in both 2024 and 2027 is Barry, Wales.

Table 7: Fraction of Remote Workers by City

Ranking	City	Fraction Remote in 2024	Fraction Remote in 2027
1	Barry, Wales	58%	62%
2	Seattle, WA	55%	59%
3	Omaha, NE	54%	59%
4	Scranton, PA	50%	55%
5	Liverpool, England	42%	45%

However, Seattle's large population results in it having the greatest total number of remote workers.

Table 8: Total Remote Workers by City

Ranking	City	Total Remote Workers in 2024	Total Remote Workers in 2027
1	Seattle, WA	1,083,084	1,189,925
2	Liverpool, England	312,490	345,573
3	Omaha, NE	276,258	302,382
4	Scranton, PA	127,062	135,204
5	Barry, Wales	34,671	37,916

Table 9: Final Impact Score by City

Ranking	City	Final Impact Score (0-2)
1	Seattle, WA	1.628
2	Omaha, NE	1.627
3	Liverpool, England	1.512
4	Scranton, PA	1.507
4	Barry, Wales	1.507

4.5 Sensitivity Analysis

Table 10: Sensitivity Analysis

Constant	Δ	% Change in 2024 Fraction Remote	% Change in 2024 Fraction Remote
p_f	+10%/-10%	+0.24%/-6.4%	-0.21%/-2.5%
k	+10%/-10%	+0.1%/-6.1%	+0.1%/-2.1%
q	+10%/-10%	+1.7%/-8.5%	+1.8%/-4.5%
α	+10%/-10%	-2.9%/-5.7%	-0.11%/-2.2%

4.6 Strengths and Weaknesses

As shown by the sensitivity analysis, adjusting the constants by 10% does not result in a significant change in the final remote fraction. The fact that it is difficult to arrive at an exact value for our constants decreases the accuracy of our results. However, it would be nearly impossible to predict how the world will emerge from the pandemic to a high degree of certainty. To that end, an advantage of our model is that the constants can be refined upon observation of how remote work grows during the next few years.

Another weakness of our model is that it only accounts for two of the effects that remote work has on a city. We acknowledge that housing prices have been extremely affected by the rise of remote work [20], with many others that scientists haven't yet predicted sure to come.

5 Conclusion and Further Studies

In Part I, our team applied a linear regression model to project the future total employment of given industries [5] in the select cities. We found that, in the present, 60.801% of jobs in Seattle, Washington, are remote-ready. The years 2024 and 2027 will be 60.799% and 60.797% remote-ready, respectfully. As for Omaha, Nebraska, 60.279% of jobs are currently remote-ready, 60.324% will be remote-ready in 2024, and 60.797% will be remote-ready in 2027. Scranton, Pennsylvania, currently has 56.332% of jobs remote-ready, with that figure decreasing to 56.172% and 55.930% by 2024 and 2027. Across the pond, Liverpool currently has only 46.394% remote-ready jobs, and will have 46.353% and 46.295% remote-ready jobs in 2024 and 2027. Barry, Wales, presently has 64.437% remote-ready jobs, with an increase to 64.519% and 64.641% remote-ready jobs by 2024 and 2027.

If we found more employment data of the select industries in the given cities (such as monthly data)—as stated 2.5.3—we could have applied advanced machine learning algorithms to precisely predict future employment levels. As a result, rather than assuming all industries changed their unemployment numbers linearly, we could have researched the growth or decline rates of the select industries to accurately follow the industries' employment trend. Given this insight, it would have been manageable to mathematically assess the optimal constants in a given best-fit line and its degree. This would prevent the over-fitting and under-fitting of our model.

In Part 2, we wrote a function that would compute a weighted average of demographic data and industry data in order to output the probability that an individual will be allowed to work from home and will choose to work from home. Given a much larger dataset, we may have been able to solve this problem using more sophisticated methods, such as a Random Forest Model. Otherwise, given more time, we would prefer to refine our equation in order to arrive at a more robust method of computing the intersection of our probabilities.

In Part 3, we used a modified Bass Diffusion Model to predict the future fraction of jobs that will be performed remotely in a given industry, and combined those predictions with our industry employment projections per city from Part 1 in order to arrive at a fraction of workers who are remote in 2024 and 2027. Then, by weighting the fraction with the relative commute time length and population percentage of children, we were able to arrive at a final impact score. Here, we see the greatest room for improvement by the calculation of more accurate constants and the development of our definition of "impact." Other factors that we would like to consider would be housing prices, increasing urban sprawl, and climate [20].

6 References

[1] - https://m3challenge.siam.org/node/559.

[2] - https://www.shrm.org/resourcesandtools/hr-topics/employee-relations/pages/drawbacks-to-working-at-home-.aspx

[3] - https://www.ipsos.com/sites/default/files/ct/news/documents/2021-07/Global%20Advisor%20-%20Return%20to%20Workplace%20Survey.pdf

[4] - https://hbr.org/2004/10/how-industries-change

[5] - https://www.bls.gov/sae/

[6] - https://www.nomisweb.co.uk/

[7] - https://bfi.uchicago.edu/wp-content/uploads/BFI_White-Paper_Dingel_Neiman_3.2020.pdf

[8] - https://himalayas.app/advice/fully-remote-companies

[9] - https://digital.com/the-real-reasons-people-want-to-keep-working-from-home/

[10] - https://www.nber.org/papers/w26948

[11] - https://news.gallup.com/poll/355907/remote-work-persisting-trending-permanent.aspx

[12] - https://wfhresearch.com/wp-content/uploads/2022/01/Barrero-AEA-NABE-Jan22.pdf

[13] - https://www.tuc.org.uk/news/annual-commuting-time-21-hours-compared-decade-ago-finds-tuc

[14] - https://www.unomaha.edu/college-of-public-affairs-and-community-service/governing/stories/commute-times.php

[15] - https://www.census.gov/quickfacts/

[16] - https://citypopulation.de/

[17] - https://en.wikipedia.org/wiki/Bass_diffusion_model

[18] - https://www.reuters.com/business/healthcare-pharmaceuticals/country-by-country-scientists-eye-beginning-an-end-covid-19-pandemic-2021-11-03/

[19] - https://hubblehq.com/blog/impact-of-commute-time-on-work-preferences

[20] - https://www.planetizen.com/news/2022/01/115783-how-remote-work-could-reshape-american-cities

[21] - https://www.benefitnews.com/news/better-wfh-policies-are-critical-to-retaining-working-parents-study-shows

7 Appendix

7.1 Ready or Not

```python
import numpy as np
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt

#Same year array for all occupations and cities. Constant except for Liverpool and Barry.
timePoints = np.array([2000,2005,2010,2015,2019,2020,2021]).reshape((-1, 1))
def fittingModel(Employees, time = timePoints):
  Model = LinearRegression().fit(time, Employees)
  y_pred = Model.predict(np.array([2024, 2027]).reshape(-1,1))
  return y_pred

def totalModel(MLCEmployees, ManufacturingEmployees, TTUEmployees, InformationEmployees, FinancialEmployees, BusinessServicesEmployees, EducationHealthEmployees, LeisureHospitalityEmployees,
               OtherServicesEmployees,GovernmentEmployees, time = timePoints):
  #Fits linear models to each industry being considered.
  MLC = fittingModel(MLCEmployees, time)
  Manufacturing = fittingModel(ManufacturingEmployees, time)
  TTU = fittingModel(TTUEmployees, time)
  Information = fittingModel(InformationEmployees, time)
  Financial = fittingModel(FinancialEmployees, time)
  BusinessServices = fittingModel(BusinessServicesEmployees, time)
  EducationsHealth = fittingModel(EducationHealthEmployees, time)
  LeisureHospitality = fittingModel(LeisureHospitalityEmployees, time)
  OtherServices = fittingModel(OtherServicesEmployees, time)
  Government = fittingModel(GovernmentEmployees, time)

  Year2024Total = (MLC[0] + Manufacturing[0] + TTU[0] + Information[0] + Financial[0] + BusinessServices[0] + EducationsHealth[0] + LeisureHospitality[0]
                   + OtherServices[0] + Government[0])

  Year2024Remote = (MLCTheoreticalPercentRemote*MLC[0] + ManufacturingTheoreticalPercentRemote*Manufacturing[0]
                    + TTUTheoreticalPercentRemote*TTU[0] + InformationTheoreticalPercentRemote*Information[0] +
                    FinancialTheoreticalPercentRemote*Financial[0] + BusinessServicesTheoreticalPercentRemote*BusinessServices[0]
                    + EducationHealthTheoreticalPercentRemote*EducationsHealth[0] + LeisureHospitalityTheoreticalPercentRemote*LeisureHospitality[0]
                    + OtherServicesTheoreticalPercentRemote*OtherServices[0] + GovernmentTheoreticalPercentRemote*Government[0])

  Year2027Total = (MLC[1] + Manufacturing[1] + TTU[1] + Information[1] + Financial[1] + BusinessServices[1] + EducationsHealth[1]
```

```
    + LeisureHospitality[1]
                    + OtherServices[1] + Government[1])

  Year2027Remote = (MLCTheoreticalPercentRemote*MLC[1] +
    ManufacturingTheoreticalPercentRemote*Manufacturing[1]
                    + TTUTheoreticalPercentRemote*TTU[1] +
    InformationTheoreticalPercentRemote*Information[1] +
                    FinancialTheoreticalPercentRemote*Financial[1]
    + BusinessServicesTheoreticalPercentRemote*BusinessServices[1]
                    + EducationHealthTheoreticalPercentRemote*
    EducationsHealth[1] +
    LeisureHospitalityTheoreticalPercentRemote*LeisureHospitality
    [1]
                    + OtherServicesTheoreticalPercentRemote*
    OtherServices[1] + GovernmentTheoreticalPercentRemote*
    Government[1])
  return2024 = round((Year2024Remote/Year2024Total)*100,3)
  return2027 = round((Year2027Remote/Year2027Total)*100, 3)
  return [return2024, return2027, ((return2027-return2024)/
    return2024)*100]


#Theoretical Percent Remote by Industry
MLCTheoreticalPercentRemote = 0.02
ManufacturingTheoreticalPercentRemote = 0.59
TTUTheoreticalPercentRemote = 0.35
InformationTheoreticalPercentRemote = 1.00
FinancialTheoreticalPercentRemote = 0.94
BusinessServicesTheoreticalPercentRemote = 0.8
EducationHealthTheoreticalPercentRemote = 0.53
LeisureHospitalityTheoreticalPercentRemote = 0.50
OtherServicesTheoreticalPercentRemote = 0.34
GovernmentTheoreticalPercentRemote = 0.95


#Seattle, Washington
#Mining, logging, construction
MLCSeattleEmployees =
    [101700,104700,83600,107100,127600,129900,109600]

#Manufacturing
ManufacturingSeattleEmployees = [212800,  171300, 167000, 188200,
    184300, 168400, 142200]

#Trade, transportation, and utilities
TTUSeattleEmployees = [325600,  313200, 301600, 354400, 398000,
    390300, 332600]

#Information
InformationSeattleEmployees = [79500, 77700,  87700,  97500,
    128400, 133700, 139000]

#Financial activities
FinancialSeattleEmployees = [101800,  106700, 92100,  95900,
    101400, 100400, 87600]

#Professional and business services
```

```
BusinessServicesSeattleEmployees = [220500, 214400, 220700, 268600,
     302100, 295700, 277500]

#EducationHealth
EducationHealthSeattleEmployees = [183700,  198400, 231500, 251300,
     283000, 272100, 223500]

#LeisureHospitality
LeisureHospitalitySeattleEmployees = [145800, 152500, 155700,
    185200, 207800, 150600, 133000]

#OtherServices
OtherServicesSeattleEmployees = [57800, 61800,  63200,  70200,
    78700,  71100,  59300]

#Government
GovernmentSeattleEmployees = [236000, 252100, 264200, 270300,
    275500, 266000, 206700]


#Omaha, Nebraska
#Mining, logging, construction
MLCOmahaEmployees = [23500, 25700,  20900,  25800,  30500,  30400,
     30700]

#Manufacturing
ManufacturingOmahaEmployees = [35700, 32900, 31200 ,32700 ,33600
    ,33000  ,33500]

#Trade, transportation, and utilities
TTUOmahaEmployees = [108100,  99700,  94100,  98200,  96100,
    91800,  94100]

#Information
InformationOmahaEmployees = [15300, 13300,  11200,  11600,  10500,
     9900, 9800]

#Financial activities
FinancialOmahaEmployees = [35800, 37600,  40500,  42200,  46000,
    45500,  44100]

#Professional and business services
BusinessServicesOmahaEmployees = [60400,  61700,  63500,  73600,
    73100,  70900,  71900]

#Education and health services
EducationHealthOmahaEmployees = [55200, 61200,  71500,  76100,
    79700,  78000,  79600]

#Leisure and hospitality
LeisureHospitalityOmahaEmployees = [41100,  42200,  43800,  48400,
     52000,  43300,  47500]

#Other services
OtherServicesOmahaEmployees = [14400, 16400,  17800,  18300,
    18600,  17700,  18300]
```

```
#Government
GovernmentOmahaEmployees = [55300, 59900, 65300, 65900, 66900,
     65200, 65200]


#Scranton, Pennsylvania
#Mining, logging, construction
MLCScrantonEmployees = [10700, 10600, 9400, 10200, 10500, 9800,
    10300]

#Manufacturing
ManufacturingScrantonEmployees = [45600, 34900, 27800, 27000,
    28600, 26900, 27200]

#Trade, transportation, and utilities
TTUScrantonEmployees = [55600, 58500, 58900, 62600, 63500,
    61900, 63900]

#Information
InformationScrantonEmployees = [7000, 6300, 5000, 3500, 2900,
    2600, 2500]

#Financial activities
FinancialScrantonEmployees = [13700, 13400, 12400, 12600, 13100,
      13000, 13000]

#Professional and business services
BusinessServicesScrantonEmployees = [23000, 23400, 25000, 29800,
      28300, 25500, 26100]

#Education and health services
EducationHealthScrantonEmployees = [45300, 49100, 52200, 51900,
    55200, 51500, 50500]

#Leisure and hospitality
LeisureHospitalityScrantonEmployees = [19000, 22000, 21800,
    23300, 23500, 17800, 18200]

#Other services
OtherServicesScrantonEmployees = [10000, 10000, 8300, 8500, 8800,
    7500, 7700]

#Government
GovernmentScrantonEmployees = [31200, 31700, 31700, 29400,
    29000, 27900, 28300]


#Liverpool, England
timePointsLiverpool = np.array([2005,2010,2015,2019,2020,2021]).
    reshape((-1, 1))

#Mining, logging, construction
MLCLiverpoolEmployees = [141000, 138700, 138000, 150300, 153500,
    146240]

#Manufacturing
ManufacturingLiverpoolEmployees = [80200, 73900, 80500, 100200,
```

```
    107500, 103120]

#Trade, transportation, and utilities
TTULiverpoolEmployees = [92000, 109100, 128300, 146900, 145800,
    146100]

#Information
InformationLiverpoolEmployees = [59900, 68400,  66800,  72100,
    73300,  73120]

#Financial activities
FinancialLiverpoolEmployees = [22400, 21105,  22890,  22820,
    20160,  25592]

#Professional and business services
BusinessServicesLiverpoolEmployees = [41600,  39195,  42510,
    42380,  37440,  47528]

#EducationHealth
EducationHealthLiverpoolEmployees = [29000, 24450,  22900,  23850,
      21450,  23900]

#LeisureHospitality
LeisureHospitalityLiverpoolEmployees = [69800,  59800,  66400,
    64000,  69600,  69700]

#OtherServices
OtherServicesLiverpoolEmployees = [70400, 78700,  73800,  80200,
    75000,  73120]

#Government
GovernmentLiverpoolEmployees = [29000,  24450,  22900,  23850,
    21450,  26560]


#Barry, Wales
timePointsBarry = np.array([2005,2010,2015,2019,2020,2021]).reshape
    ((-1, 1))

#Mining, logging, construction
MLCBarryEmployees = [4100,  3500, 4500, 4600, 3300, 4100]

#Manufacturing
ManufacturingBarryEmployees = [5700,  4300, 3500, 4900, 4800, 5700]

#Trade, transportation, and utilities
TTUBarryEmployees = [1400,  900,  1400, 800,  1200, 1400]

#Information
InformationBarryEmployees = [4000,  4400, 3800, 3900, 3600, 4000]

#Financial activities
FinancialBarryEmployees = [3045,  2940, 3710, 3535, 4095, 3045]

#Professional and business services
BusinessServicesBarryEmployees = [5655, 5460, 6890, 6565, 7605,
    5655]
```

```

#EducationHealth
EducationHealthBarryEmployees = [9700,  9800, 10550,  10850,
    11550,  9700]

#LeisureHospitality
LeisureHospitalityBarryEmployees = [9500, 10800,  11200,  13000,
    8000, 9500]

#OtherServices
OtherServicesBarryEmployees = [2400,  4000, 2800, 3700, 3100, 2400]

#Government
GovernmentBarryEmployees = [9700, 9800, 10550,  10850,  11550,
    9700]

returnSeattle = totalModel(MLCSeattleEmployees,
    ManufacturingSeattleEmployees, TTUSeattleEmployees,
    InformationSeattleEmployees, FinancialSeattleEmployees,
                           BusinessServicesSeattleEmployees,
    EducationHealthSeattleEmployees,
    LeisureHospitalitySeattleEmployees,
    OtherServicesSeattleEmployees, GovernmentSeattleEmployees)
print("Seattle, Washington")
print(returnSeattle)

returnOmaha = totalModel(MLCOmahaEmployees,
    ManufacturingOmahaEmployees, TTUOmahaEmployees,
    InformationOmahaEmployees, FinancialOmahaEmployees,
                           BusinessServicesOmahaEmployees,
    EducationHealthOmahaEmployees, LeisureHospitalityOmahaEmployees
    , OtherServicesOmahaEmployees, GovernmentOmahaEmployees)
print("Omaha, Nebraska")
print(returnOmaha)

returnOmaha = totalModel(MLCScrantonEmployees,
    ManufacturingScrantonEmployees, TTUScrantonEmployees,
    InformationScrantonEmployees, FinancialScrantonEmployees,
                           BusinessServicesScrantonEmployees,
    EducationHealthScrantonEmployees,
    LeisureHospitalityScrantonEmployees,
    OtherServicesScrantonEmployees, GovernmentScrantonEmployees)
print("Scranton, Pennsylvania")
print(returnOmaha)

returnLiverpool = totalModel(MLCLiverpoolEmployees,
    ManufacturingLiverpoolEmployees, TTULiverpoolEmployees,
    InformationLiverpoolEmployees, FinancialLiverpoolEmployees,
                           BusinessServicesLiverpoolEmployees,
    EducationHealthLiverpoolEmployees,
    LeisureHospitalityLiverpoolEmployees,
    OtherServicesLiverpoolEmployees, GovernmentLiverpoolEmployees,
                           time=timePointsLiverpool)
print("Liverpool, UK")
print(returnLiverpool)

returnBarry = totalModel(MLCBarryEmployees,
```

```
    ManufacturingBarryEmployees, TTUBarryEmployees,
    InformationBarryEmployees, FinancialBarryEmployees,
                               BusinessServicesBarryEmployees,
    EducationHealthBarryEmployees, LeisureHospitalityBarryEmployees
    , OtherServicesBarryEmployees, GovernmentBarryEmployees,
                               time=timePointsBarry)
print("Barry, Wales")
print(returnBarry)
```

7.2 Remote Control

Code to compute probability of liking remote work given factors:

```
def prob_choose_remote(gender, age, household_income, education,
    parent, commute):
    # Coefficients for each probability
    alpha = [1,1,1,1,2,2]
    # the factor each probability adds to the sum
    gender_prob = 0
    age_prob = 0
    income_prob = 0
    education_prob = 0
    parent_prob = 0
    commute_prob = 0
    # finding all the probabilities
    if gender == "female":
        gender_prob = alpha[0] * 0.68
    elif gender == "male":
        gender_prob = alpha[0] * 0.64
    if age == "under_35":
        age_prob = alpha[1] * 0.67
    elif age == "35-49":
        age_prob = alpha[1] * 0.66
    elif age == "50-74":
        age_prob = alpha[1] * 0.63
    if household_income == "low":
        income_prob = alpha[2] * 0.62
    elif household_income == "medium":
        income_prob = alpha[2] * 0.63
    elif household_income == "high":
        income_prob = alpha[2] * 0.69
    if education == "low":
        education_prob = alpha[3] * 0.59
    elif education == "medium":
        education_prob = alpha[3] * 0.61
    elif education == "high":
        education_prob = alpha[3] * 0.71
    if parent == "yes":
        parent_prob = alpha[4] * 0.68
    elif parent == "no":
        parent_prob = alpha[4] * 0.63
    if commute == "0-60":
        commute_prob = alpha[5] * 0.433
    elif commute == "61-120":
        commute_prob = alpha[5] * 0.696
    elif commute == "over_121":
```

```
        commute_prob = alpha[5] * 0.869
    #return the weighted mean
    probability_vector = (gender_prob, age_prob, income_prob,
    education_prob, parent_prob, commute_prob)
    return sum(probability_vector) / (alpha[0] + alpha[1] + alpha
    [2] + alpha[3] + alpha[4] + alpha[5])
```

7.3 Just a Little Home-work

```
from scipy.integrate import odeint
import numpy
import matplotlib.pyplot as plt
import math

#define the model
def model(F,t):
    p_f = 0.05
    k = 4.5
    a = 0.9
    p = p_f/(1 + math.exp(-k*(t-a)))
    q = 0.38
    dFdt = (1-F) * (p + q*F)
    return dFdt

#define the timespan in years
t = np.linspace(0,5,6)

#assign a starting base fraction
y0 = 0.46

#solve the differential equation
sol = odeint(model, y0, t)

#imported data from part I
seattle_data = np.array([[122281.62083936, 125811.28798842],
[157607.67004342, 152571.27351664],
[378726.91751085, 387780.31837916],
[140115.34008683, 149242.54703328],
[93709.98552822, 92558.82778582],
[303718.23444284, 316276.98986975],
[275938.49493488, 287018.08972504],
[170387.04775687, 172775.8683068 ],
[72090.44862518, 73726.33863965],
[255363.6758321, 256007.74240232]]).T

omaha_data = np.array([[30905.42691751, 32014.58031838],
[32617.00434153, 32452.35166425],
[91252.60492041, 89585.99855282],
[9023.01013025, 8313.82054993],
[46911.07091172, 48321.74384949],
[75092.11287988, 77036.14327062],
[84683.35745297, 88201.9536903 ],
[49254.92040521, 50273.552822  ],
[19164.83357453, 19651.5195369 ],
[68481.83791606, 69853.87120116]]
).T
```

```

scranton_data = np.array([[10003.03907381,  9946.16497829],
[22881.25904486, 20656.98263386],
[64760.85383502, 65854.16063676],
[1730.24602026, 1049.9276411 ],
[12727.42402315, 12646.34587554],
[28167.72793054, 28785.96237337],
[53975.32561505, 54826.3748191 ],
[20523.44428365, 20448.9869754 ],
[7491.60636758, 7170.11577424],
[27912.73516643, 27381.54848046]]).T

liverpool_data = np.array([[150978.58085809, 153096.99669967],
[108079.9669967,  113805.51155116],
[160455.61056106, 171263.03630363],
[75736.56765677, 78003.20132013],
[23212.13861386, 23451.35148515],
[43108.25742574, 43552.50990099],
[21308.82838284, 20325.66006601],
[67672.77227723, 68047.02970297],
[76852.73927393, 77402.54125413],
[22463.25082508, 21717.11221122]]).T

barry_data = np.array([[4061.22112211, 4076.07260726],
[4785.47854785, 4775.08250825],
[1143.23432343, 1129.8679868 ],
[3753.96039604, 3688.61386139],
[3739.62871287, 3854.5049505 ],
[6945.02475248, 7158.36633663],
[10953.13531353, 11151.40264026],
[10333.33333333, 10333.33333333],
[3097.85478548, 3108.25082508],
[10953.13531353, 11151.40264026]]
).T

#fractions calculated using the diffusion model
Industry_Fractions_2024 =
    [0.014,0.5,0.32,0.92,0.83,0.77,0.5,0.49,0.28,0.78]
Industry_Fractions_2027 =
    [0.009,0.57,0.34,0.98,0.91,0.79,0.52,0.5,0.32,0.9]

#calculate 2024 totals
seattle_2024 = np.dot(Industry_Fractions_2024,seattle_data[0])
omaha_2024 = np.dot(Industry_Fractions_2024,omaha_data[0])
scranton_2024 = np.dot(Industry_Fractions_2024,scranton_data[0])
liverpool_2024 = np.dot(Industry_Fractions_2024,liverpool_data[0])
barry_2024 = np.dot(Industry_Fractions_2024,barry_data[0])

#calculate 2024 fractions
seattle_2024_fraction_remote = seattle_2024 / seattle_data[0].sum()
omaha_2024_fraction_remote = omaha_2024 / omaha_data[0].sum()
scranton_2024_fraction_remote = scranton_2024 / scranton_data[0].
    sum()
liverpool_2024_fraction_remote = liverpool_2024 / liverpool_data
    [0].sum()
barry_2024_fraction_remote = barry_2024 / barry_data[0].sum()

```

```
#calculate 2027 totals
seattle_2027 = np.dot(Industry_Fractions_2027,seattle_data[1])
omaha_2027 = np.dot(Industry_Fractions_2027,omaha_data[1])
scranton_2027 = np.dot(Industry_Fractions_2027,scranton_data[1])
liverpool_2027 = np.dot(Industry_Fractions_2027,liverpool_data[1])
barry_2027 = np.dot(Industry_Fractions_2027,barry_data[1])

#calculate 2027 fractions
seattle_2027_fraction_remote = seattle_2027 / seattle_data[1].sum()
omaha_2027_fraction_remote = omaha_2027 / omaha_data[1].sum()
scranton_2027_fraction_remote = scranton_2027 / scranton_data[1].
    sum()
liverpool_2027_fraction_remote = liverpool_2027 / liverpool_data
    [1].sum()
barry_2027_fraction_remote = barry_2027 / barry_data[1].sum()
```