Riding into the Future: A Mathematical Exploration
Team #16667

Executive Summary

As the world continues to shift towards more environmentally sustainable practices, electric vehicles, but particularly the electric bicycle, will become more widespread. However, mapping the causes of e-bike growth and its implications for the future presents a unique mathematical challenge. Thus, this report is an investigation into future e-bike usage in the next 2 and 5 years, its correlation with the most commonly cited causes including gas prices and disposable income, and its impacts on carbon emissions, traffic congestion, cultural perception, and lifespan in the United Kingdom.

First, we predicted future United Kingdom (UK) e-bike sales in the next 2 and 5 years using data from France. Since no e-bike sales for the UK were available, the purchases were approximated using France's sales. While both nations have similar populations, the UK has a population density 2.36 times that of France [1], and e-bikes are more prevalent in densely-populated areas. However, since France's e-bike population has grown dramatically and the government greatly supports switching to e-bikes financially [2], UK sales would not be expected to be extremely high relative to that of France. Thus, taking into account both of these factors, UK sales being 50% more than France sales was obtained. Then, using a divided differences model, it was predicted that 1.913 million e-bikes would be sold in 2024 and 3.236 million would be sold in 2027. Using a logistic model, it was predicted that 1.473 million e-bikes would be sold in 2024 and 1.774 million sold in 2027.

Next, we evaluated if gas prices or disposable income, two factors commonly deemed causes of e-bike growth, did indeed correlate or cause UK e-bike growth. First, using the Kendall Tau correlation coefficient, it was found that gas prices and disposable income were correlated with e-bike growth. However, after stabilizing the data with the Augmented Dickey-Fuller Test and applying the Granger Causality Model, it was determined that neither factor appeared to cause the growth directly. Therefore, we concluded that while these factors likely played a role, they were not major reasons for the growth.

Finally, to quantify the resulting impacts that buying more e-bikes has on carbon emissions, overall health, and traffic congestion over time, we made three models with varying parameters based on the polynomial model extrapolated in part 1. Firstly, it was predicted that the increased use of e-bikes in the UK would lead to a substantial decrease of carbon emissions, preventing around 3.5% of all transportation emissions by 2027. In a similar model, with different parameters, it was also predicted that the traffic congestion would be alleviated significantly, saving around 1.94 hours per person in the United Kingdom in 2027. Finally, the model predicted an increase in health and wellness, specifically through lifespan, estimating a 51.2 day longer average lifespan across all of the UK population by 2027.

Keywords: E-bike, Divided Differences, Logistic Model, Kendall Tau Correlation Coefficient, Augmented Dickey-Fuller Test, Granger Causality Model

Table of Contents

Background

With increasing innovation in the field of clean energy in recent years, the electric vehicle industry has dramatically increased in popularity, leading to the success of companies like Tesla. Using high-profile celebrities and directors, new galleries, and publicity stunts, electric vehicle brands have reached more mainstream audiences [3]. However, although electric automobiles have been the most publicized and sales of electric cars continue to grow exponentially, the most in-demand electric vehicle in the market has been the electric bicycle, commonly referred to as the e-bike.

E-bikes have been around for more than a century [4], but their convenience, especially in urban areas, has made them a reliable form of transportation. E-bikes allow people to travel quickly and efficiently without needing to worry about issues like public transit schedules, parking, or traffic congestion, and in low-income areas, they are an affordable form of transportation. Additionally, e-bikes provide environmental benefits, namely producing zero carbon emissions and air pollution, an issue that creates smog and plagues many urban areas. Since they run on electricity instead of gas or diesel, e-bikes prevent greenhouse gases from being emitted [5]. Furthermore, e-bikes help those with physical limitations, either from disability or age, since the electric motor reduces or eliminates the need to pedal [5].

Because of their utility and recent growth, companies have shown interest in e-bikes, and some, such as Amazon and FedEx, have already invested in them as part of a more efficient delivery system [6]. In addition, policymakers have wondered if e-bikes will become part of a sustainable energy movement to get cars off the roads, and some are even considering rebates for people who use electric bicycles [7]. To better understand and promote the advancement of transportation technology, it is essential to analyze the growth of e-bikes and its causes and implications for the future.

Global Assumption

Assumption	Justification
UK e-bike annual sales are 50% more than that of France.	The UK has a population density of 2.36 times that of France [1], and e-bikes are more prevalent in densely-populated areas. However, France's e-bike population has grown dramatically and the government greatly supports switching to e-bikes financially [2], so UK sales would likely not grow significantly faster than France sales. Fifty percent of the growth of France was a good middle ground for these conditions.

1 The Road Ahead

1.1 Problem Restatement

Question 1 asks us to model future growth in e-bike sales in either the United Kingdom (UK) or the United States (US) and predict the number of e-bikes that will be sold two and five years after 2022. We chose to model UK e-bike sales.

1.2 Assumptions and Justifications

Assumption	Justification
Trends in e-bike sales will continue, even if there is slight fluctuation in the data. No major developments in e-bike technology will occur, and no major events will impact its trend.	While we can take into consideration some of the factors related to e-bike growth, it is nearly impossible to foresee significant events that will dramatically affect the trend of the data. This also allows us to sufficiently simplify the situation for modeling.

1.3 Developing the Models

1.3.1 Divided Differences Model

To model future e-bike sales in the UK, we utilized the divided differences method, which helps to interpolate a set of numerical data. The end result of the divided differences method is an nth-degree polynomial equation, which we used to predict future values. The steps of the model are as follows:

1. Determine the appropriate degree of the polynomial. To do so, we take increasing numerical divided differences (first, second, third, etc.) until the values are close enough to zero, which is based upon qualitative judgment. We do not want to increase the divided difference infinitely, as it would make the polynomial susceptible to oscillation and sensitive to data error. The divided differences are calculated using the equation presented in figure 1.1.

x	$f(x)$	First Divided Difference	Second Divided Difference	Third Divided Difference
x_0	$f[x_0]$			
		$f[x_0,x_1]=\frac{f[x_1]-f[x_0]}{x_1-x_0}$		
x_1	$f[x_1]$		$f[x_0,x_1,x_2]=\frac{f[x_1,x_2]-f[x_0,x_1]}{x_2-x_0}$	
		$f[x_1,x_2]=\frac{f[x_2]-f[x_1]}{x_2-x_1}$		$f[x_0,x_1,x_2,x_3]=\frac{f[x_1,x_2,x_3]-f[x_0,x_1,x_2]}{x_3-x_0}$
x_2	$f[x_2]$		$f[x_1,x_2,x_3]=\frac{f[x_2,x_3]-f[x_1,x_2]}{x_3-x_1}$	
		$f[x_2,x_3]=\frac{f[x_3]-f[x_2]}{x_3-x_2}$		$f[x_1,x_2,x_3,x_4]=\frac{f[x_2,x_3,x_4]-f[x_1,x_2,x_3]}{x_4-x_1}$
x_3	$f[x_3]$		$f[x_2,x_3,x_4]=\frac{f[x_3,x_4]-f[x_2,x_3]}{x_4-x_2}$	
		$f[x_3,x_4]=\frac{f[x_4]-f[x_3]}{x_4-x_3}$		
x_4	$f[x_4]$			

Figure 1.1: Calculation for Divided Differences [8]

Using the data derived for the sales of e-bikes in the United Kingdom, we determined the third divided difference was the most appropriate. Thus, we chose to fit a third-degree polynomial to the data, with the equation

$$z(y) = ay^3 + by^2 + cy + d$$

where z(y) is the sales of e-bikes per year, *y* is the year after 2005, and *a*, *b*, *c*, and *d* are constants that we must find.

2. After finding the polynomial, determine the constants. Our goal for this model is to minimize S, where

$$S = \sum_{i=1}^{m}[z_i - (ay_i^3 + by_i^2 + cy_i + d)]^2$$

and *m* is the total number of years in question. In order to do so, we need to ensure that the partial derivative of each of the variables with respect to S is 0. In equation form,

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = \frac{\partial S}{\partial c} = \frac{\partial S}{\partial d} = 0.$$

Given that g_i must equal $(ay_i^3 + by_i^2 + cy_i + d)$ to achieve partial derivatives of 0, we set up the following systems of equations to determine our a, b, c, and d values.

$$(\sum y_i^3)a + (\sum y_i^2)b + (\sum y_i)c + md = \sum z$$

$$(\sum y_i^4)a + (\sum y_i^3)b + (\sum y_i^2)c + (\sum y_i)d = \sum y_i z_i$$

$$(\sum y_i^5)a + (\sum y_i^4)b + (\sum y_i^3)c + (\sum y_i^2)d = \sum y_i^2 z_i$$

$$(\sum y_i^6)a + (\sum y_i^5)b + (\sum y_i^4)c + (\sum y_i^3)d = \sum y_i^3 z_i$$

3. Use the given data to determine the constants and substitute the constants back into the polynomial equation. After substituting values from years 3-14 into the system in step

two, we achieved a, b, c, and d values of 0.5227, -5.984, 26.01, and -5.402, respectively. Therefore, the empiric cubic model is given by

$$z(y) = 0.5527x^3 - 5.984x^2 + 26.01x - 5.402.$$

Plotting the polynomial with the original e-bike sales data gives us the following graph, which has an R-squared value of 0.97315. In addition, from the model, we can calculate the predicted amount of e-bikes sold both in 2024 and 2027, which are shown in table 1.1.

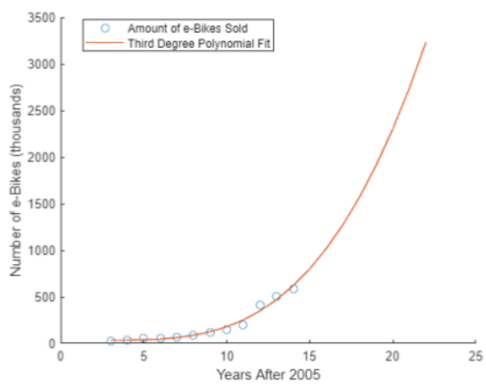

Figure 1.2: Divided Differences Model for UK e-bike Sales

Year	Number of e-Bikes Sold
2024	1.913 million
2027	3.236 million

Table 1.1: Predicted e-bike Sales After 2 and 5 years

1.3.2 Logistics Model for Population Growth

Although the divided differences model results in a prediction based on previous years, it is dependent on the idea that e-bike sales will experience uncontrollable growth. Given that a range of factors can disrupt e-bike sales, especially if demand for e-bikes grows faster than supply can keep up, we believe using a logistics model can help to provide a new lens to the problem. For this model, we defined the following variables:

Variable	Value
P	Population at time t
k	Compounded Annual Growth Rate (CAGR)
K	Maximum Number of e-bikes Sold in 1 Year
t	Time in Years

Table 1.2: Variables for Logistics Model

Before evaluating the differential equation for the model, we have to determine our k and K values. To calculate the k-value or CAGR, we used the following equation

$$CAGR = \left(\frac{e-bikes\ sold\ in\ 2019}{e-bikes\ sold\ in\ 2008}\right)^{(1/(2019-2008))} - 1$$

which resulted in k = 0.344. We utilized CAGR as our growth rate instead of the rate of proportionality in our logistics equation because it is often used in finance to compare past performances and predict future values. The rate of proportionality, in contrast, is dependent on the historic data, which may not accuratly represent changes in future growth rates. For our K-value, we estimated that a maximum of 2 million e-bikes could be sold in a given year. Given that roughly 3.1 million total bikes (including e-bikes) were purchased in the UK in 2020 [9], we find it reasonable to believe that if the UK could hit similar numbers, roughly 66% of those people could be purchasing an e-bike.

Since our initial data consists of e-bikes sold per year, the model, which is the rate of change of total e-bikes sold from year to year and scaled to thousands of e-bikes, is as follows:

$$\frac{d^2P}{dt^2} = kP(1 - \frac{P}{K})$$

$$\frac{d^2P}{dt^2} = 0.344 * P(1 - \frac{P}{2000})$$

In order to solve this function, we utilized the Matlab ‘ode45’ function, which solves nonstiff differential equations. Even though we had a second degree differential equation, we treated it like a first degree differential equation, which allowed us to input parameters of the differential equation, time t, and the number of e-bikes sold in 2008, which was 22.5. After solving the differential equation, we plotted the predicted value with the logistics model alongside the actual values. Our graph and resulting predictions for 2 and 5 years from 2022 are shown in figure 1.3 and table 1.3 below.

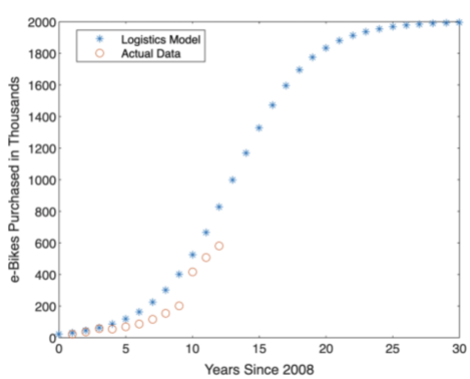

Figure 1.3: Logistics Model for UK e-bike Sales

Year	Number of e-Bikes Sold
2024	1.473 million
2027	1.774 million

Table 1.3: Predicted e-bike Sales After 2 and 5 years

Although these two models provide contrasting values, we believe that the divided differences model more accurately predicts future values because the exponential model is appropriate to use in the short term before the growth comes unrealistic. In contrast, the carrying capacity of the logistics model may have restricted the growth of e-bikes too much.

1.4 Sensitivity Analysis

Although the divided differences model provides limited room for sensitivity analysis, we performed sensitivity analysis on the logistics model by changing both K, the maximum number of e-bikes sold in a year, and k, the CAGR. For the maximum number of e-bikes sold in a year, we choose two arbitrary numbers: one above the carrying capacity and one below. For the growth rate, we calculated a new k-value using the following equation, which estimates the slope of the line between the initial and final values:

$$k = \frac{2019-2008}{(\#\ of\ e-bikes\ sold\ in\ 2019)-(\#\ of\ e-bikes\ sold\ in\ 2008)} = 0.01966.$$

In addition, we used an arbitrary high k-value of 0.5 to compare it with our low k-value of 0.01966. The results of the sensitivity analysis are shown in table 1.4 below.

	Number of E-bikes Sold			
Year	K = 1000; k = 0.344	K = 3000; k = 0.344	K = 2000; k = 0.01966	K = 2000; k = 0.5
2024	849,792	1.949 million	30,689	1.943 million
2027	940,789	2.517 million	32,523	1.987 million

Table 1.4 : Sensitivity Analysis for Logistics Model

The sensitivity analysis corroborates our expectations as we see that decreasing the maximum number of e-bikes sold within a single year decreases our predictions for 2024 and 2027 and vice versa. When the carrying capacity of a logistics model is limited, it forces values to essentially scale down, resulting in a positive correlation between the K value and the e-bikes sold after a certain number of years. The k-value displayed similar trends because when the growth rate was extremely limited, the number of e-bikes sold per year barely grew from its starting point of 22,500 e-bikes. However, when the growth rate was higher than 0.344, the number of e-bikes sold grew rapidly at first but eventually stagnated due to the carrying capacity of 2 million e-bikes.

1.5 Strengths and Weaknesses

1.5.1 Strengths

- **Multiple Methods.** We used two models (Divided Differences and Logistics) to address this, which reflects the breadth and depth of our considerations.
- **Adaptable.** Both models that were used to predict e-bike sales can be applied to other sets of data. In addition, as e-bike sales can be tracking the coming years, the models can be re-applied to future data to determine accuracy.

1.5.2 Weaknesses

- **Lack of Complete Consensus Among Models.** The two models that we used gave varying results, which may be an indication that the parameters and assumptions we made were not reflective of the data. We could have improved this by conducting a more in-depth error analysis.
- **Accuracy Testing.** Given more time, we could have performed multiple out-of-sample tests to verify the accuracy of our logistics models, which could have highlighted flaws in the modeling technique.

2 Shifting Gears

2.1 Problem Restatement

Question 2 asks us to consider one or more factors that may have contributed to e-bike growth and use mathematical modeling to determine whether those factors were significant causes for the growth in e-bike usage. For this problem, we examined gas prices and disposable income, the amount of money remaining after tax deductions and other mandatory charges.

2.2 Assumptions and Justifications

Assumption	Justification
Gas prices and disposable income are the two most feasible and relevant factors to e-bike growth.	Based on the environmental responses to the UK YouGov poll, at most 30% of the respondents believe “climate change and the environment” is an important issue [10]. Therefore, we conclude that people in the UK would likely be more motivated to purchase e-bikes if they had more money to spend or more financial reasons to switch. Although environmental consciousness is a possibility we considered, it did not have enough data overlap with the e-bike data to justify its usage in this section.
Using a smaller set of empirical data for e-bike sales that we have is preferable to using a broader set of data that could be derived from calculated e-bike sales in Question 1.	Although larger sets of data may be preferable for interpolation, using the data given for e-bike sales (adjusted in our global assumption) helps to reduce errors in case our predicted results data from Question 1 are highly inaccurate. In addition, using calculated data from Question 1 would only allow us to add 3 more years of data for gas prices and 2 more years for disposable income, which is not worth the risk.

2.3 Our Modeling Process

For this problem, we determined that to evaluate where a factor significantly contributed to the growth in e-bike usage, we must consider both correlation and causation. To consider both of these factors, we used the process displayed in figure 2.1.

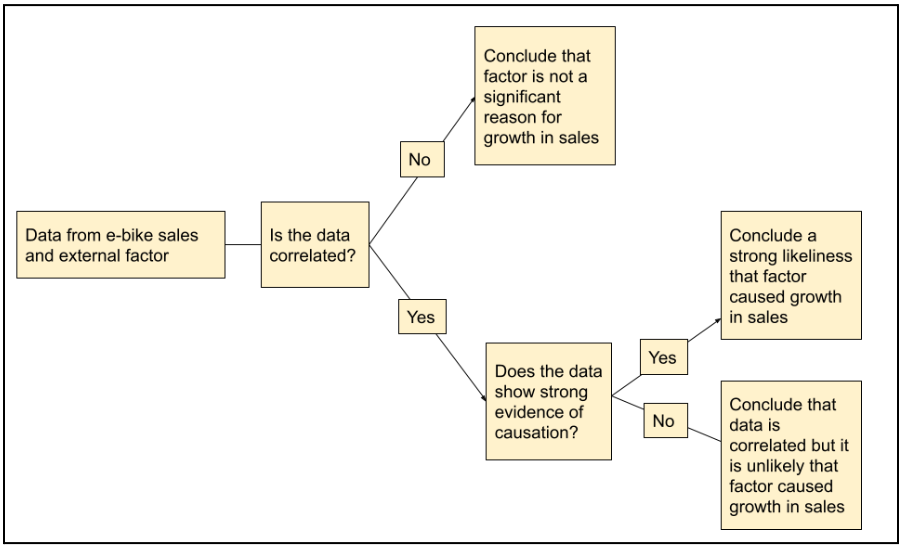

Figure 2.1: Flowchart for Determining Significance of Factor

2.4 Correlation

To determine correlation, we calculated the Kendall Tau Correlation Coefficient, or τ-coefficient, which measures the strength and direction of an association between two variables. It is often used when the necessary qualifications to use the Pearson Correlation Coefficient fails. In this case, since our data does not follow a normal distribution, we are unable to use the Pearson Correlation Coefficient. The Kendall Tau Correlation Coefficient is calculated as follows:

$$\tau = \frac{n_c - n_d}{n(n-1)/2}$$

where n_c is the number of concordant pairs, n_d is the number of discordant pairs, and n is the data size. This method assumed that no values are tied in the data set. To apply this coefficient to our data, we developed a Matlab code that ranks the values using the ‘tiedrank’ function, and then compares the values using a series of “for” loops and “if” statements. We believe that this method is preferable to a calculator because it helps to efficiently input a given set of data and perform all the necessary evaluations and calculations in a sequence. For the correlation between the number of e-bikes sold and gas prices in the UK from 2008 to 2019, we determined that $\tau = .0909$, and for the correlation between the number of e-bikes sold and disposable income in the UK from 2008 to 2019, we determined that $\tau = 0.9091$. These results reveal that while there is strong evidence of the correlation between the number of e-bikes sold and disposable income, there is extremely weak evidence of correlation between the number of e-bikes sold and gas prices in the UK. From here, we can conclude that gas prices are not a significant reason for the growth in e-bike sales. Disposable income requires further investigation.

2.5 Granger Causality

While it is impossible to use only data to demonstrate causation, the Granger Causality model can show if one variable Granger-causes another, meaning it can be used to predict the future value of another variable, thus suggesting that it at least influences the other variable.

To apply the test, the data must first be stationary, meaning that its mean, variance, and autocorrelation structure do not change significantly over time. Time series data must be stationary to make each individual point independent (having no trend). This allows the Granger Causality model to determine how closely one variable resembles another and how many lags are needed. To differentiate, we apply the Augmented Dickey-Fuller Test on the data between 2008 and 2019, starting with the null hypothesis that the time series is non-stationary and the p-value as the level of marginal significance. Only when the p-value is less than 0.05 do we reject the null hypothesis and conclude that the time series is stationary. Otherwise, we differentiate the data until we arrive at a stationary series.

The p-value of the original e-bike time series is 0.9988, so we take the first difference of the data. This yields a p-value of 0.3184, which is still not enough to reject the null hypothesis, so we take the second difference of the data. This results in a p-value of 0.0341, so the data is stationary. Repeating this process with UK disposable income, the original p-value is 0.9861 and its first difference p-value is 4.280×10^{-7}. Since the number of data points is different among the variables due to their different degrees of differencing, we remove the earliest 2 values from UK gas prices and the earliest value from disposable income, which allows all variables to represent the same time frame.

Now, since all the data is stationary, we apply the F-test from the Granger Causality. Again, we start with the null hypothesis that time series x does not Granger-cause time series y, or in other words, knowing the value of time series x at a certain lag is not useful for predicting the value of time series y at a later time period. The F-test produces a p-value, and similarly, a p-value of fewer than 0.05 means we reject the null hypothesis and conclude that time series x Granger-causes time series y. However, due to the small data size, only lags of up to 2 years can be applied. In addition, it is impossible to have 0 lags because it implies that one variable causes another event that is happening at the same time. In this context, we assumed people who earn more money after taxes will not immediately go to buy a new e-bike. The following table shows the lag order and p-value for determining if disposable income Granger-causes e-bike usage.

Tested Granger Causing Variable	Lags	P-Value
UK Disposable Income	1	0.6285
	2	0.9046

Table 2.1: Granger Causality Results

As shown by the table, UK disposable income does not Granger-cause e-bike usage. Therefore, although this variable is shown to be strongly correlated with e-bike growth, it does not have sufficient evidence to support its causality or influence with it.

2.6 Strengths and Weaknesses

2.6.1 Strengths

- **Process of Elimination Algorithm.** We used a process to immediately eliminate variables that did not meet the most basic criteria for causality, which is correlation. This made our modeling process quicker and more efficient.
- **Objective.** Both criteria that we used to determine causality did not require any creation of constants or coefficients, meaning that they are objective and not subject to modeling bias.

2.6.2 Weaknesses

- **Inability to Produce Direct Causation.** Due to the nature of data analysis, our models were unable to completely determine if one variable caused another. Experiments would have to be conducted to determine direct cause-and-effect relationships.
- **Insufficient Considered Factors.** Given more time, we could have investigated more factors related to e-bike usage. We likely would have determined a factor that strongly influenced e-bike popularity, which would have strengthened our results, instead of only finding factors that did not cause e-bike usage.

3 Off the Chain

3.1 Problem Restatement

Question 3 asks us to evaluate the impact of e-bike popularization on the usage of other modes of transportation, such as gas vehicles and regular bikes. In addition, we are asked to quantify the effect of e-bike growth on carbon emissions, traffic congestion, health, and any other factors. For this question, we investigated the impacts on carbon emissions, traffic congestion, and lifespan.

3.2 Assumptions and Justifications

Utilizing the polynomial model determined in Question 1: The Road Ahead and other parameters, the impacts that e-bike sales have on carbon emissions across the United Kingdom can be modeled and quantified. In order to predict the resulting influence of e-bikes, we opted for a proportional model that incorporates the amount of e-bikes sold per year, the vicinity to bike lanes, and cultural perception of climate change.

Assumption	Justification

Around 20 percent of all buildings in the United Kingdom are at most 400 meters away from a protected bike lane, and that number grows at a linear rate of around 1.75 percent each year.	As a part of Greater London's Strategic Cycling Plan revealed before the pandemic, the city plans to introduce a large increase in the number of bike lanes throughout the city, with 19 percent of buildings being within 400 meters of a large network of bike lanes in late 2021 [11]. Such plans continue as the city pledges to have 35 percent by 2025 and 70 percent by 2041. Given that there is not much past data on this statistic, we assumed that the amount of bike lanes has been linearly growing, starting from 3 percent in 2008. We also extrapolated this data from the Greater London Area to represent all of the United Kingdom, due to the similarities between the many urban areas and the impossibility of measuring the proportion of bike roads to all buildings in the UK.
Each car in the United Kingdom emits an average of 1.682 metric tons of CO2 per year.	Based on the data given by the United Kingdom, the average car produces 221.4 grams of CO2 per mile [12]. The same source also documented that cars in the UK average 7600 miles each year. This leads to an overall production of around 1.682 metric tons of CO2 per car.
Being 400 meters or closer to a bike lane increases the use cycling for transportation by 70%.	A study from Transport for London (TfL) shows that a close vicinity to the large network of bike lanes has led to an increase of 70% in cycling levels [11].
Cultural Perception of Climate Change is exponential in the short term and can be modeled as such.	As the given data depicts, there has been a sharp increase in the concerns of climate change throughout the United Kingdom as heatwaves, drastic natural disasters, and storms have become much more commonplace throughout the world. From this, we assume that the population of the United Kingdom will continue to express increased concern as temperatures continue to rise, at least in the short run. The data that best fits the model is an exponential one. Although it is not sustainable in the long run, it demonstrates a realistic rise in concern for at least the next 5 years.
People will use e-bikes around half of the time.	Due to weather, longer trips, and the variability of human nature, it cannot be expected that anyone who owns an e-bike will use it for all trips, forgoing all public or private transportation. In this, we assume that around half of the time e-bikes will be used. The other half is dedicated for longer trips, bad weather, and other extraneous circumstances.
Those who are concerned about climate change are 10% more likely to use an	Although many have claimed to express concern for climate change, ultimately many will not choose to spend upwards of a thousand dollars to reduce their own footprint. In addition, the model also should take into account those who have an e-bike

e-bike, in addition to other variables.	not out of necessity, but because of their concern for the environment. Those who are more aware will ride their e-bike more often than the general population. This 10% proportion takes into account both of these values.
The average driver in the UK loses 80 hours each year from traffic congestion.	A study done by INRIX Inc. showed that for 2022, the average driver lost around 80 hours in traffic [13].
Consistent bikers will, on average, live 3 years longer than the average global lifespan.	An 18-year-long study done by Danish researchers through the Copenhagen City Heart Study found that for men who cycled consistently with average intensity gained 2.9 years of life, but those who cycled with fast intensity lived 5.3 years longer [14]. Women gained 2.2 and 3.9 respectively. Using this, we make the assumption that on average, those who bike on normal bikes consistently will increase their lifespan by approximately 3 years.
In correlation to lifespan, on average, those who specifically own an e-bike will receive 95% of the 3 years that consistent bikers will receive.	This is due to the fact that not all who buy an e-bike will exercise regularly with it, nor will get the full benefit as its electric capabilities may take away the exercise.

3.3 Modeling the Impacts on Carbon Emissions

3.3.1 Exponential Model for Concerns on Climate Change

With the assumptions above, we defined the following variables.

Variable	Value
P	e-bikes sold
H	Proportion of homes near bike lanes
k	Proportion of e-bike use
C	Proportion who are climate change concerned
I	Proportion of those that are concerned and willing to act with e-bikes

Table 3.1: Variables for Carbon Emissions Impact Model

Before we can properly quantify the impact that e-bikes have on emissions, we must extrapolate the data given for concern of climate change to the broader range of 2008-2027. To do so, we created a simple exponential regression model. Since the range of given data was only from 2011 to 2022, this model allows us to find the proportion C for years 2008-2027. Using the regression, we get the equation

$$y = 0.01947e^{0.1631x}$$

where x is the years since 2005 and y is the proportion of the UK population that express concern for climate change. The purpose for modeling this statistic is to demonstrate the increasing concern as it relates to the amount of e-bikes that will be used and with what frequency. Plotting this equation against the given data gives us an R^2 value of 0.867, which tells us that roughly 86.7% of the variability is explained and resolved by the model. The results are represented in figure 3.1 below.

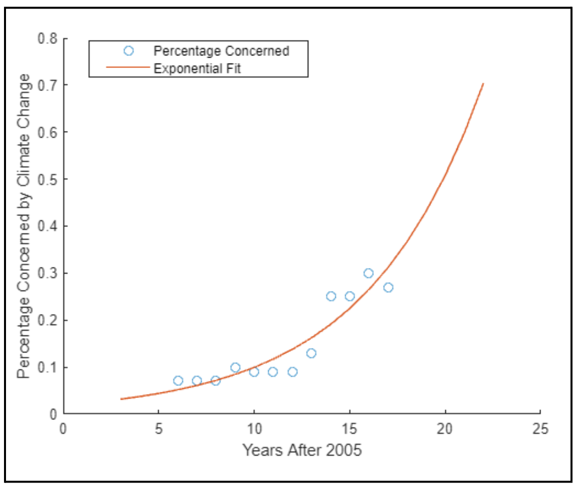

Figure 3.1: Exponential Model of Proportion of UK Population Concerned With Climate Change

3.3.2 Emissions Saved Model

Using data from 2008-2027, we developed a MatLab Code that combines the modeled data of e-bike sales with concerns for climate change, while also including the parameter for the vicinity of bike lanes, which plays an important role in urban settings. The code then plots the resulting function showing how any increase in bike sales leads to less emissions. The equation is as follows:

$$E(P) = 1.682kP + 0.7HP + ICP$$

To display in a meaningful way, we found that the total vehicle emissions in the United Kingdom for 2021 was equal to approximately 107.5 million metric tons [15]. From this, we found the percentage of the total amount of emissions that would be prevented by traveling with e-bikes instead. The results for 2008-2027 are represented in figure 3.2 below, with table 3.2 presenting values from 2022, 2024, and 2027.

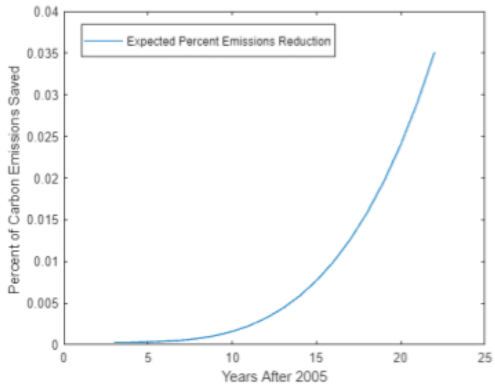

Figure 3.2 Percent of Emissions Prevented Over Time

Year	Percent of Emissions Prevented
2022	1.26
2024	1.96
2027	3.51

Table 3.2: Percent of Emissions Prevented for 2022, 2024, 2027

It is pertinent to note that if we expand our graphical analysis beyond 2027, we see that the model continues to grow exponentially, which is unrealistic. This is due to the fact that we extrapolated this model using the polynomial fit, and thus the model should only be looked at for short term values. However, in the years that are documented here, it shows a reasonable amount of emissions that can be prevented with e-bikes. In 2027, if bike lanes continue to be developed, concerns for climate change increase, and sales continue to rise, a 3.51% of all transportation carbon emissions will be prevented by e-bikes.

3.4 Modeling the Impacts on Traffic Congestion

Similar to the model on the impact that e-bikes have on overall carbon emission, increased use of e-bikes will have a similar impact on traffic congestion, specifically demonstrating a decrease in the amount of time the average UK resident will waste in traffic. We used similar parameters, but instead of using the proportion for carbon emissions, we quantified the traffic congestion in average number of hours wasted. We then produced this equation:

$$T(P) = 80kP + 0.7HP + ICP$$

In this model, the parameters of bike lanes and concern for climate change stay constant, but we replaced the average carbon emissions prevented with average hours in traffic prevented. After calculating the hours gained through e-bike travel, we then divided that value by the overall population of the United Kingdom, 67.33 million, showing the average gain in hours for each inhabitant of the UK. This format allows the reader to better understand the amount of time being saved through different means of transportation. The results for 2008-2027 are represented in figure 3.3 below, with table 3.3 presenting value from 2022, 2024, and 2027.

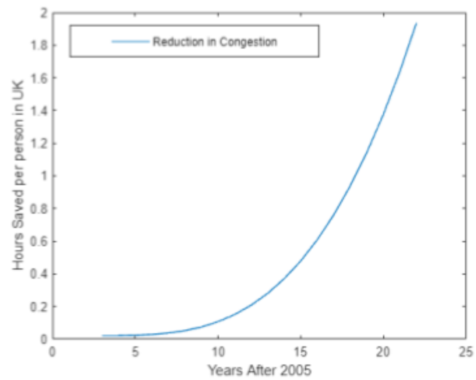

Figure 3.3: Hours Saved Per Person Over Time

Year	Amount of Hours Saved Per Person
2022	0.76
2024	1.14
2027	1.94

Table 3.3: Hours Saved Per Person for 2022, 2024, 2027

Similar to 3.3, we note that if we expand our graphical analysis beyond 2027, we see that the model continues to grow unrealistically. This is due to the fact that we extrapolated this model from a polynomial fit. However, in the years that are documented here, it shows that in 2027, 1.94 hours per person will be saved due to e-bikes serving as an expanded mode of transportation.

3.5 Modeling e-Bike Impact on Lifespan

Much like the other models, the impact that e-bikes have on health and wellness, specifically lifespan, can be modeled using many of the same parameters. However the abundance of bike lanes has no real impact on life expectancy as the relative frequency of biking does not substantially affect the health benefits, especially if one already owns an e-bike. By using the assumption made earlier that consistent cycling adds on an average of 3 years to one's lifespan, we can model the total increase of lifespan in the United Kingdom with this equation:

$$L(P) = 3(0.95)P + ICP$$

Another change that was made to this model was the removal of the *k* constant for 0.95. Since this model measures lifespan, more than half of the people who own an e-bike will be able to get the full potential and increased lifespan. Even if one does not travel only on an e-bike, they will still reap many of the health benefits. The equation was then plotted on MatLab, but similar to the traffic congestion model, it was divided by the whole United Kingdom population to show the average increase in lifespan per person in the UK. For easier comprehension, the data was converted from years to days by multiplying by 365 days, showing the amount of days added to the average lifespan. The results for 2008-2027 are represented in figure 3.4 below, with table 3.4 presenting values from 2022, 2024, and 2027.

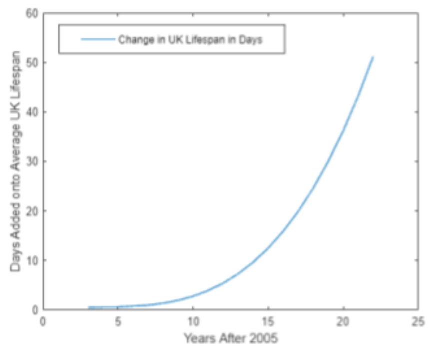

Figure 3.4: Days Added Per Person Over Time

Year	Amount of Days Added to Average Lifespan
2022	19.9
2024	30.0
2027	51.2

Table 3.4: Days Added Per Person for 2022, 2024, 2027

From this model, we found that in the present year, the amount of days that are added to the lifespan of the entire UK population due to e-biking is equal to 19.9 days. It is predicted to be equal to 30 days and 51.2 days in 2024 and 2027 respectively.

3.6 Sensitivity Analysis

Firstly, with the carbon emissions, if we show k to be 0.25, 0.75, and 1, we achieve the following results for 2022, 2024, 2027.

	Percent of Transportation Emissions Prevented		
Proportion of e-bike use (k)	2022	2024	2027
.25	0.76	1.21	2.24
.5	1.26	1.96	3.51
.75	1.76	2.71	4.77
1	2.26	3.46	6.04

Table 3.5: Sensitivity Analysis Results for Carbon Emission

From our sensitivity analysis, we found that as k increases, so does the amount of carbon emission prevented, which fits with our intuitive assumption. If everyone who had bought an e-bike solely rode on it, the effect on the emissions by 2027 would be much more pronounced, at around 6% of all transportation emissions being prevented.We can also see how as the years pass, the predicted values begin to deviate farther from a realistic expectation, as our models are based on a polynomial function.

For the model on traffic congestion, we adjusted the rate at which bike lanes were built each year, starting from 1.75%, the actual amount, to 3%, 4%, and 5%.

	Hours Saved from Traffic Congestion Per UK Citizen		
Linear Growth of Bike Lanes	2022	2024	2027
.0175	0.7619	1.1443	1.9382
.03	0.7643	1.1483	1.9462
.04	0.7661	1.1515	1.9526
.05	0.7680	1.1547	1.9590

Table 3.6: Sensitivity Analysis Results for Traffic Congestion

This sensitivity analysis shows the relatively small impact that the vicinity of bike lanes near buildings has for traffic congestion. Even with 5% of the total bike lanes being built each year, where in 2027, 98% of all houses will be in a close vicinity to a bike lane, the change in hours saved is minimal. We would see a similar lack of change in the sensitivity analysis for concern for climate change, based on the assumption that even those who are concerned would show varying degrees of commitment to e-biking.

Finally, for the last model, we compared the average increase in lifespan for those who cycle consistently to see how the average UK lifespan would change.

	Days Added to Average UK Lifespan		
Additional Years	2022	2024	2027
1	6.784	10.304	17.902
3	19.921	30.016	51.236
5	33.058	49.727	84.570
7	46.194	69.439	117.903

Table 3.7: Sensitivity Analysis Results for UK Lifespan

This sensitivity analysis shows that the change in years is relatively linear to the amount of days added to the lifespan. The amount of days added to average lifespan if cycling adds 3 years is around 3 times as much as the data in the first row, where cycling gives 1 more year, although not exactly. This is due to the other parameters in the equations.

3.7 Strengths and Weaknesses

3.7.1 Strengths

- **Variety of In Depth Assumptions.** The assumptions made to create this model were in depth and backed up by a variety of historical data regarding e-bikes, carbon emissions, traffic congestion, and health and wellness. This helps to increase the precision of the models.
- **Connection to Other Questions.** Problem 3 drew upon analysis done in question 2 to make assumptions and used the model found in question 1 to show the effects of exponentially increasing e-bike consumption on the rest of society. Cohesion between models helps to make a more complete understanding of the problem.

3.7.1 Weaknesses

- **Similar Approach for All Three Factors.** The equation used to model all three factors was quite similar, which resulted in similar exponential trends. Using a different modeling approach, such as a Monte Carlo simulation for a condensed population, could have revealed results that corroborate or challenge our findings.
- **Lack of Cross Analysis.** Although we synthesized each of these three models separately in our modeling process, we did not engage in analysis between them to determine the impact of one area or another. This could be performed through multivariable analysis.

6 References

[1] *Country comparison*. (n.d.). WorldData. Retrieved March 4, 2023, from https://www.worlddata.info/country-comparison.php?country1=FRA&country2=GBR

[2] Thompson, H. (2021, June 4). *France in top 5 for cycling in the world, Google Maps data shows*. The Connexion. Retrieved March 4, 2023, from https://www.connexionfrance.com/article/French-news/France-in-top-5-for-cycling-in-the-world-Google-Maps-data-shows-for-World-Bike-Day

[3] Swant, M. (2022, January 2). *Why Auto Giants And Startups Are Marketing Electric Vehicles To More Mainstream Audiences*. Forbes. Retrieved March 4, 2023, from https://www.forbes.com/sites/martyswant/2022/01/02/why-auto-giants-and-startups-are-marketing-electric-vehicles-to-mainstream-audiences/?sh=55a2077f4f5a

[4] *The History of The Electric Bike*. (n.d.). Rize Bikes. Retrieved March 4, 2023, from https://rizebikes.com/blogs/blog/the-history-of-the-electric-bike

[5] Aarstol, S. (n.d.). *Environmental Impact of Electric Bicycles*. Tower Electric Bikes. Retrieved March 4, 2023, from https://www.towerelectricbikes.com/environmental-impact

[6] Ralston, W. (2022, September 21). *It's Time for Cities to Ditch Delivery Trucks–for Cargo Bikes*. Wired. Retrieved March 4, 2023, from https://www.wired.co.uk/article/cargo-bikes-greener-quicker

[7] Cobler, N. (2023, January 13). *E-bike popularity soars as Austin Energy rolls out pilot program*. Axios Austin. Retrieved March 4, 2023, from https://www.axios.com/local/austin/2023/01/13/e-bike-popularity-austin-energy-rebates

[8] *Numerical Analysis/Divided differences*. (2021, March 31). Wikiversity. https://en.wikiversity.org/wiki/Numerical_Analysis/Divided_differences
[9] *UK Bicycle Supply In 2021*. (2021, May 25). Bicycle Association. Retrieved March 4, 2023, from https://www.bicycleassociation.org.uk/news-press/uk-bicycle-supply-in-2021/#:~:text=Bicycle%20Association%20statement%3A%20UK%20bicycle%20supply%20in%202021&text=Last%20year%20over%203.1%20million,of%20the%20last%20five%20years.
[10] Ride Like the Wind, MathWorks Math Modeling Challenge 2023, https://m3challenge.siam.org/node/596.
[11] Buehler, R., & Pucher, J. (2022, October 11). *Cycling Through COVID-19: London*. StreetsBlogUSA. Retrieved March 4, 2023, from https://usa.streetsblog.org/2022/10/11/cycling-through-covid-london-united-kingdom/
[12] Yurday, E. (2022, April 14). *Average CO2 Emissions per Car in the UK*. NimbleFins. Retrieved March 4, 2023, from https://www.nimblefins.co.uk/average-co2-emissions-car-uk
[13] *INRIX 2022 Global Traffic Scorecard: London Tops List as Most Congested City, U.S. Cities Inch Closer*. (2023, January 10). INRIX. Retrieved March 4, 2023, from https://inrix.com/press-releases/2022-global-traffic-scorecard-uk/#:~:text=The%20average%20driver%20in%20the,lost%20%C2%A3707%20in%20time
[14] Schnohr, P., Marott, J. L., Jensen, J. S., & Jensen, G. B. (2012). Intensity versus duration of cycling, impact on all-cause and coronary heart disease mortality: the Copenhagen City Heart Study. *European journal of preventive cardiology*, *19*(1), 73–80. https://doi.org/10.1177/1741826710393196
[15] *Transport and environment statistics 2022*. (2022, October 20). Gov.UK. Retrieved March 4, 2023, from https://www.gov.uk/government/statistics/transport-and-environment-statistics-2022/transport-and-environment-statistics-2022#greenhouse-gas-emissions-from-transport

7 Appendix

7.1 Code for Question 1

7.1.1 Matlab Code for Divided Differences Graph

```
% Assigning variables from given data

france_table = readtable("TCP23_data.xlsx","Range","D10:D21");

france_bike = france_table{:,:}.';

france_years = linspace(3,14,12);

% Plotting polynomial for UK

uk_bike = 1.5*france_bike; % Assumption based off population density and bike population

scatter(france_years,uk_bike)
```

```
hold on

future_years = linspace (3,22,20);

bikes_sold =
(0.5227*future_years.^3)-(5.984*future_years.^2)+(26.01*future_years)-5.
402; % Modeled Equation of Bikes Sold per Year

plot(future_years,bikes_sold)

hold off

xlabel("Years After 2005")

ylabel("Number of e-Bikes (thousands)")

legend("Amount of e-Bikes Sold","Third Degree Polynomial Fit")

legend("Position", [0.17202,0.84008,0.35714,0.082143])
```

7.1.2 Matlab Code for Logistics Model

```
% Growth Rate obtained using CAGR

k = .344;

% Maximum number of e-bikes that can be sold in a year

K = 2000;

% Time to run the code through

T = 0:1:30;

% Defining the differential Equation

h = @(t,y)[k.*y(1).*(1-(y(1)/K))]

% Utitilize the ode45 function to solve

[T za] = ode45(h,T,[22.5])

% Graphing the Logistics Function

plot(T,za(:,1),'*')

hold on

% Data for the Original UK Data (Adjusted from France Data by x1.5)

x = [1 2 3 4 5 6 7 8 9 10 11 12];

y = [22.5 36 57 55.5 69 85.5 117 153 201 417 507 582];
```

```
% Graphing Original Data on the same plot
plot(x,y,'o')
hold off
% Graph labels
xlabel('Years Since 2019');
ylabel('e-Bikes Purchased');
legend('Logistics','Original Data');
```

7.2 Code for Question 2

7.2.1 Matlab Code Kendall Tau Correlation Coefficient

```
% Data from UK Biking (Adjusted From France Data)
% and UK Income/UK Gas Prices(2008-2019)
X = [22.5 36 57 55.5 69 85.5 117 152 201 417 507 582];
% Income Data; Could be wwapped with gas prices data
Y = [38188 37814 38282 38769 39732 38947 40118 41383 41821 42699 43886
44644];
% Rank the values of each variable
Rx = tiedrank(X);
Ry = tiedrank(Y);
% Calculate the number of concordant and discordant pairs
n = length(X);
C = 0;
D = 0;
for i = 1:n-1
   for j = i+1:n
       if (X(i) < X(j) && Y(i) < Y(j)) || (X(i) > X(j) && Y(i) > Y(j))
           C = C + 1;
       elseif (X(i) < X(j) && Y(i) > Y(j)) || (X(i) > X(j) && Y(i) <
Y(j))
```

```
            D = D + 1;
        end
    end
end
% Calculate Kendall's tau value
tau = (C - D) / (n * (n - 1) / 2);
% Display the result
fprintf('Kendall''s tau rank correlation coefficient = %.4f\n', tau);
```

7.2.2 Python Code for Granger Causality

```
#Gas UK
#Import packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
%matplotlib inline
#Read in excel spreadsheet and create data frame
gas_bike_csv = pd.read_csv('M3Granger.csv')
gas_bike_df = gas_bike_csv.iloc[:,[0,1]].dropna()
#Augmented Dickey Fuller Test
from statsmodels.tsa.stattools import adfuller
test_result = adfuller(df['UK Gas'])
def adfuller_test(Gas):
   result = adfuller(Gas)
   labels = ['ADF Test Statistic','p-value','#Lags Used','Number of
Observations']
   for value, label in zip(result,labels):
```

```
        print(label+':'+str(value))

#Print results of Adfuller for UK Gas

print('Original Series: UK Gas')

adfuller_test(df['UK Gas'])

gas_bike_df['First Order'] = gas_bike_df['UK Gas'] - gas_bike_df['UK
Gas'].shift(1)

gas_bike_df['Second Order'] = gas_bike_df['First Order'] -
gas_bike_df['First Order'].shift(1)

print('First Order Differencing')

adfuller_test(gas_bike_df['First Order'].dropna())

print('Second Order Differencing')

adfuller_test(gas_bike_df['Second Order'].dropna())

#Adfuller for UK Bikes

print("Original Series: UK Bikes")

adfuller_test(df['UK Bikes'])

gas_bike_df['First Order'] = gas_bike_df['UK Bikes'] - gas_bike_df['UK
Bikes'].shift(1)

gas_bike_df['Second Order'] = gas_bike_df['First Order'] -
gas_bike_df['First Order'].shift(1)

print('First Order Differencing')

adfuller_test(gas_bike_df['First Order'].dropna())

print('Second Order Differencing')

adfuller_test(gas_bike_df['Second Order'].dropna())

#Read in differenced gas and bike data

gas_bike_csv_dif = pd.read_csv('GasBike.csv')

gas_bike_df_dif = gas_bike_csv_dif.iloc[:,[0,1]].dropna()

#Granger causality test for gas and bike

from statsmodels.tsa.stattools import grangercausalitytests

grangercausalitytests(gas_bike_df_dif, maxlag = 2, verbose = False)
```

```
#Disposable Income

income_bike_csv = pd.read_csv('M3Income.csv')

income_bike_df = income_bike_csv.iloc[:,[0,1]].dropna()

#Adfuller for income

print('Original Series')

adfuller_test(income_bike_df['UK Disposable Income'])

income_bike_df['First Order'] = income_bike_df['UK Disposable Income'] -
income_bike_df['UK Disposable Income'].shift(1)

income_bike_df['Second Order'] = income_bike_df['First Order'] -
income_bike_df['First Order'].shift(1)

print('First Order Differencing')

adfuller_test(income_bike_df['First Order'].dropna())

print('Second Order Differencing')

adfuller_test(income_bike_df['Second Order'].dropna())

#Read differenced income and bike data

income_bike_csv_dif = pd.read_csv('IncomeBike.csv')

income_bike_df_dif = income_bike_csv_dif.iloc[:,[0,1]].dropna()

grangercausalitytests(income_bike_df_dif, maxlag = 2, verbose = False)
```

7.3 Code for Question 3

```
% CO2 Emissions Reduction From More eBikes

% Different Parameters

bike_lane =.03:.05:.98; % Assumed Linear Increase in Amount of Homes
near Bike Lanes

perception = readtable ("TCP23_data.xlsx","Sheet","Q2 Contributing
Factors","Range","J70:J81");

uk_perception = perception{:,:}.'./100;

perception_years = linspace (6,17,12);

scatter(perception_years,uk_perception)

hold on
```

```
modeled_perception = 0.01947*exp(0.1631*future_years); % Perception
change over time

plot (future_years, modeled_perception)

xlabel("Years After 2005")

ylabel("Percentage Concerned by Climate Change")

legend("Percentage Concerned","Exponential Fit")

legend("Position", [0.17202,0.84008,0.35714,0.082143])

hold off

% Total Emission Calculation

emissions =
1.682*.5*bikes_sold*1000+.7*bikes_sold*1000.*bike_lane+0.1*bikes_sold*10
00.*modeled_perception; % Emissions removed by eBike

total_emissions = 107.5*10^6;

percentage_emissions = emissions./total_emissions;

plot(future_years,percentage_emissions)

xlabel("Years After 2005")

ylabel("Percent of Carbon Emissions Saved ")

legend("Expected Percent Emissions Reduction")

legend("Position", [0.16309,0.81389,0.46786,0.082143])

% Traffic Congestion

hours_lost = 80; % Average Amount of hours lost per year from congestion
in UK

london_hours = 156; % Average number of hours lost per year in London

population = 67.33*10^6;

% Shows the hours gained from all people who switch to biking

total_hours_gained = .5*bikes_sold*1000*hours_lost +
.7*bikes_sold*1000.*bike_lane+0.1*bikes_sold*1000.*modeled_perception;

hours_per_person = total_hours_gained/population; % Average number of
hours given back due to less congestion

plot(future_years,hours_per_person)
```

```
xlabel("Years After 2005")

ylabel("Hours Saved per person in UK")

legend("Reduction in Congestion")

legend("Position", [0.16309,0.81389,0.46786,0.082143])

% Health and Wellness

increased_lifespan = 3;

lifespan =
.95*bikes_sold*1000*increased_lifespan+0.1*bikes_sold*1000.*modeled_perc
eption;

years_added = lifespan/population % Amount in years added on average to
the lifespan of the population of UK

days_added = years_added*365 % Amount in days added to the average
lifespan in all of UK

plot(future_years, days_added)

xlabel("Years After 2005")

ylabel("Days Added onto Average UK Lifespan")

legend("Change in UK Lifespan in Days")

legend("Position", [0.16309,0.81389,0.46786,0.082143])
```