Ride Like the Wind Without Getting Winded: The Growth of E-bike Use

1 Executive Summary

Dear Transportation Secretary Pete Buttigieg,

Over the past 20 years, traffic congestion has grown in cities of all sizes in the United States. As cities continue to grow and become more densely populated, this issue will only get worse. Public transportation infrastructure has failed to keep up with these developments, and as a result, people have turned to electric bikes as a solution. E-bikes have many benefits: they provide bikers with a fast, reliable form of transportation and allows them to bypass the issues of timing public transit or worrying about parking. Moreover, using e-bikes reduces congestion on roads, allowing those who choose to drive to have a more pleasant experience. E-bikes can also be a more affordable mode of transportation for those who cannot afford a car. E-bikes are more environmentally friendly compared to their gas counterparts.

We first developed a mathematical model to predict the number of e-bikes that will be sold two and five years from now in the US. After analyzing the year-to-year number of e-bike sales in Europe, France, and China, we concluded that an exponential model is a good fit for e-bike sales in emerging markets such as those in Europe and France, while a logistic model is a good fit for e-bike sales in well-established markets like China. The exponential model effectively captures the "coolness" factor of e-bikes, as more people purchase e-bikes their peers will also follow.

Next, we analyzed the relative importance of different factors that may have contributed to the growth of e-bike sales. We looked at factors such as environmental awareness, gas prices, health and fitness, personal finances, and market sentiment. Our model computed the correlation between Google Search frequencies of important keywords such as popular stock exchanges, terms pertaining to finances (401(k), inflation, gas prices), environment, and wellness. Our findings show that people buy e-bikes not necessarily for the environment but for their own finances and health. By investing in e-bike infrastructure such as charging stations, bike lanes, racks, and trails, you can help build an emerging lucrative market while helping people live healthier, happier lives.

Finally, we quantify the impacts of the growth of e-bikes on carbon emissions, road safety, sound pollution, traffic congestion, and road repair costs. We fit a logistic curve to model the total number of e-bike sales in the distant future before calculating the effects. Our findings show that e-bikes can significantly reduce carbon emissions and sound pollution. Furthermore, e-bikes have the potential to save 7 billion dollars over the next 20 years in road repair costs. While short-term effects may include an increase in road-related accidents, the long-term effects are impossible to ignore. These figures from our preliminary research are extremely promising. E-bikes have the potential to be a revolutionary force in our everyday lives.

Mr. Buttigieg, we urge you to invest in developing more e-bike-related infrastructure, moving our country toward efficient, environmentally-friendly mediums of transportation. While we predict that the sales of e-bikes and the e-bike market will rapidly increase over the next 20 years, the proper infrastructure needs to be implemented. Without the proper safety mechanisms, the US could witness record levels of transportation accidents. It's imperative we ensure safe, reliable access to e-bikes for the next 20 years.

Contents

2 Part I: The Road Ahead

2.1 Restatement of the Problem

This problem asks us to predict the number of e-bikes that will be sold two and five years from now in the US.

2.2 Assumptions

- *No significant legislation changes will occur in the next 5 years.* It would be infeasible to predict what kinds of e-bike-related legislation will be passed because politics is ever-changing, so we will not include it in this model.

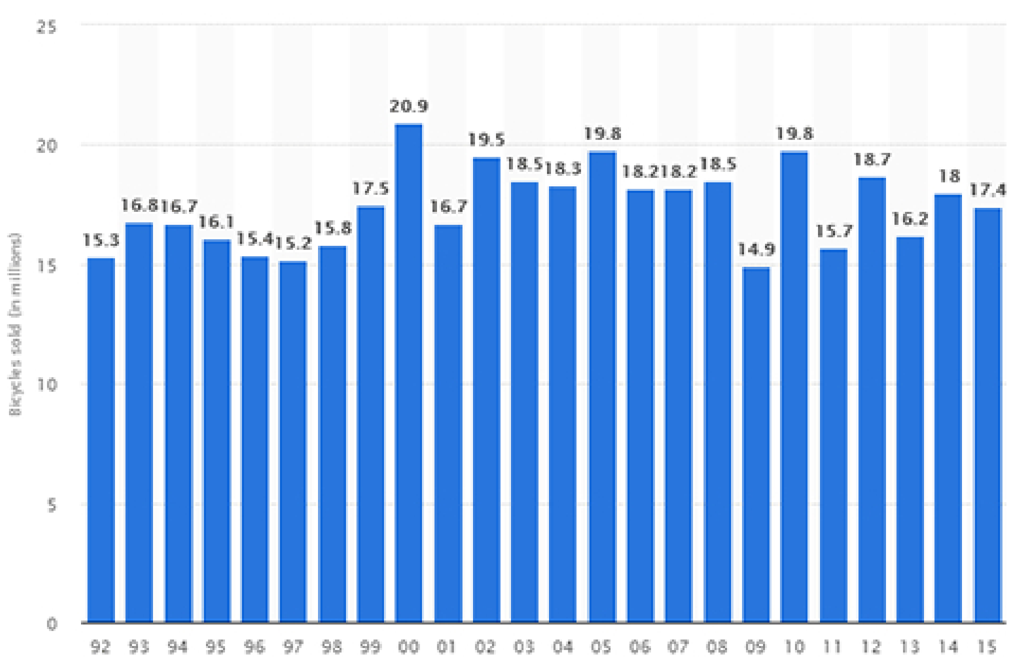
- *Gas prices will not significantly change in the next 5 years.* Gas prices fluctuate as a result of politics and many other factors, so it would be unreasonable to try and predict gas prices as a part of our model. We assume that the influence of gas prices on a consumer's decision to buy an e-bike will remain the same.

- *Technology and infrastructure for e-bikes will improve at roughly the same rate as other modes of transportation.* A breakthrough in automotive technology or public transportation infrastructure is hard to predict and would be unlikely in the next 5 years. In addition, any improvements in technology or infrastructure would likely improve all modes of transportation.

- *E-bike ridership is heavily concentrated in urban areas.* The most commonly cited reasons people provide for not using bikes more often are hills and lengthy distances. As a result, bikes are mostly used in urban areas [2].

- *The distribution of age will not significantly change in the next 5 years.* In 5 years, the age distributions in the US and in the UK are unlikely to change by even a percentage point [3] [4].

Figure 1: Number of bikes sold worldwide from 1992 to 2015.



[5]

2.3 Variables

Variables that are a function of time t are in terms of years.

Symbol	Definition	Unit
$p(t)$	Proportion of e-bikes over all bike sales	Percentage
r	The exponential growth rate	N/A
B	Total bike sales	Number of bikes
$y(t)$	The number of e-bikes sold	N/A

Table 1: Variables used for Problem 1.

2.4 Exponential Growth

E-Bikes have taken over the world in popularity. Because the adoption of e-bikes in each country is different, we will establish two sides of the spectrum on the adoption of e-bikes. There are well-established markets like China, which has recorded over 30 million sales in the past 10 years, and there are emerging markets for e-bikes in Europe and America. For well-established markets, anyone who wants an e-bike would have already gotten one, so the growth is minimal. For emerging markets, however, we see exponential growth as countries adopt technology.
Most of the major commonly cited reasons for buying an e-bike are existing bike infrastructure such as bike paths or charging stations, favorable policies such as subsidies or tax benefits, and the adoption by people in their community.

Exponential growth can be explained for a few reasons:

- The more people who use e-bikes, the more pressure (through votes, petitions, etc) they will put on the government to provide more bike-friendly infrastructure.
- With more people using e-bikes, there will also be more pressure on the government to add incentives for buying e-bikes.
- If each person convinces some number of people on average to buy an electric bike, at a small percentage of the population, the general adoption in the population will model an exponential curve. Later on, as more people adopt the e-bike, the yearly sales will model a logistic curve.

For these reasons, at an emerging market scale like the US currently, we can use an exponential regression to estimate the short-term projected sales.

2.5 Model Development

We want to take into consideration the idea that most e-bike users come from using bikes. This leads us to model both the proportion of e-bikes among all bikes sold p and the total number of bike sales B. By multiplying these two quantities, we get the total number of e-bikes sold in a given year s.

$$s = B \cdot p$$

Based on our research, we found that an exponential model would likely be a good fit for modeling the proportion of e-bikes out of all bikes [6]. Usually, we would model this proportion with a logistic curve since the proportion is bounded above at 1, but since the proportion of e-bikes is so low (less than 5%), an exponential model works well in this instance. Since the growth of e-bikes sold is continuous, we chose to use e, Euler's number, as the base for our exponential function:

$$p = e^{rt}.$$

Therefore, our final equation is $s = B \cdot e^{rt}$.

Figure 2: Linear and exponential regression of Europe E-Bike sales per year, in thousands of units. The exponential curve has an R^2 fit of 0.9918. The linear regression is charted here to demonstrate the lack of fit for the supposed "line of best fit".

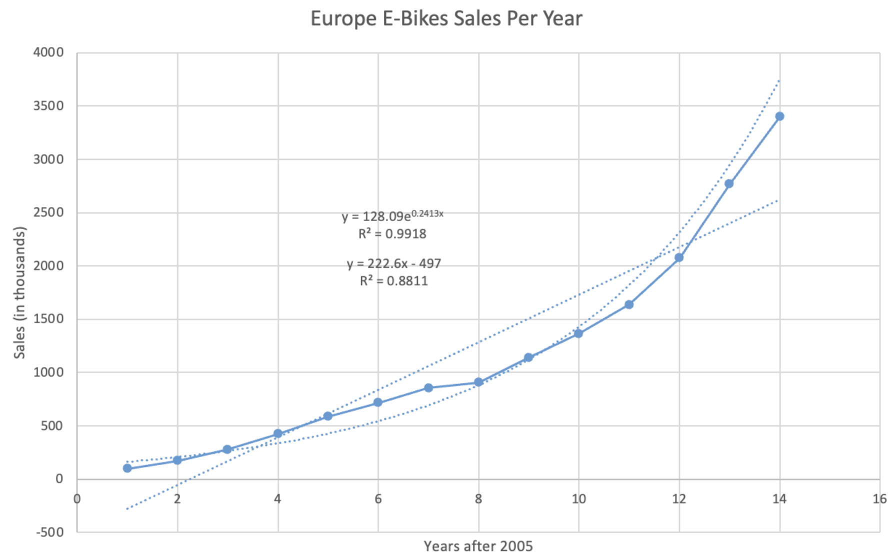


Figure 3: Linear and exponential regression of France E-Bike sales per year, in thousands of units. The exponential curve has an R^2 fit of 0.9896.

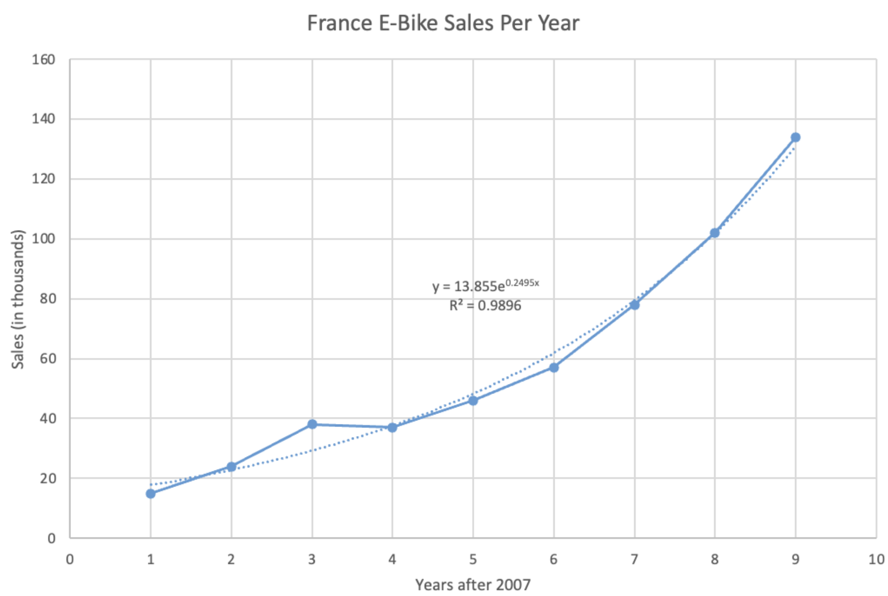
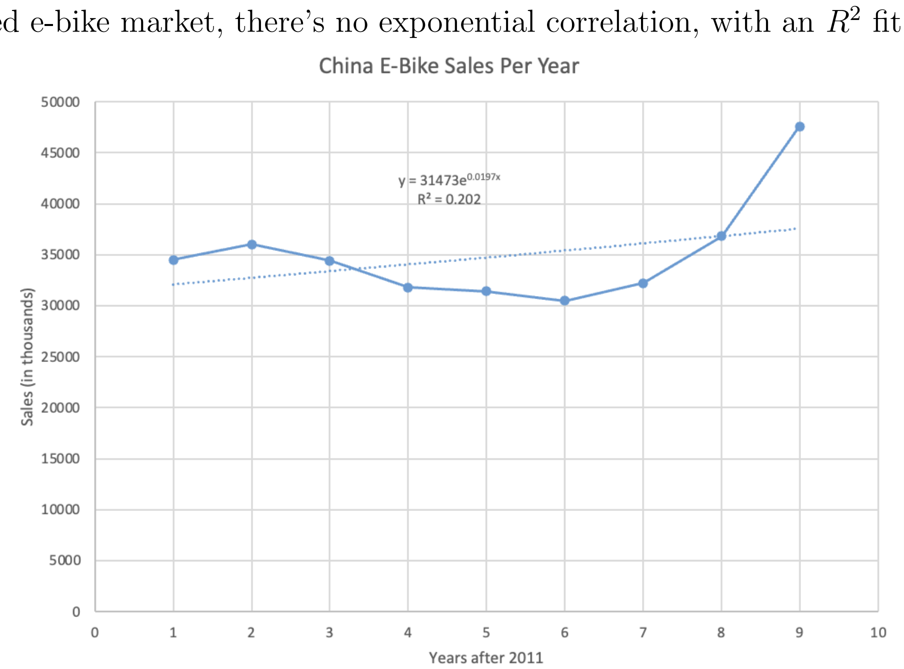


Figure 4: Exponential regression of China E-Bike sales per year, in thousands of units. As an established e-bike market, there's no exponential correlation, with an R^2 fit of 0.202.

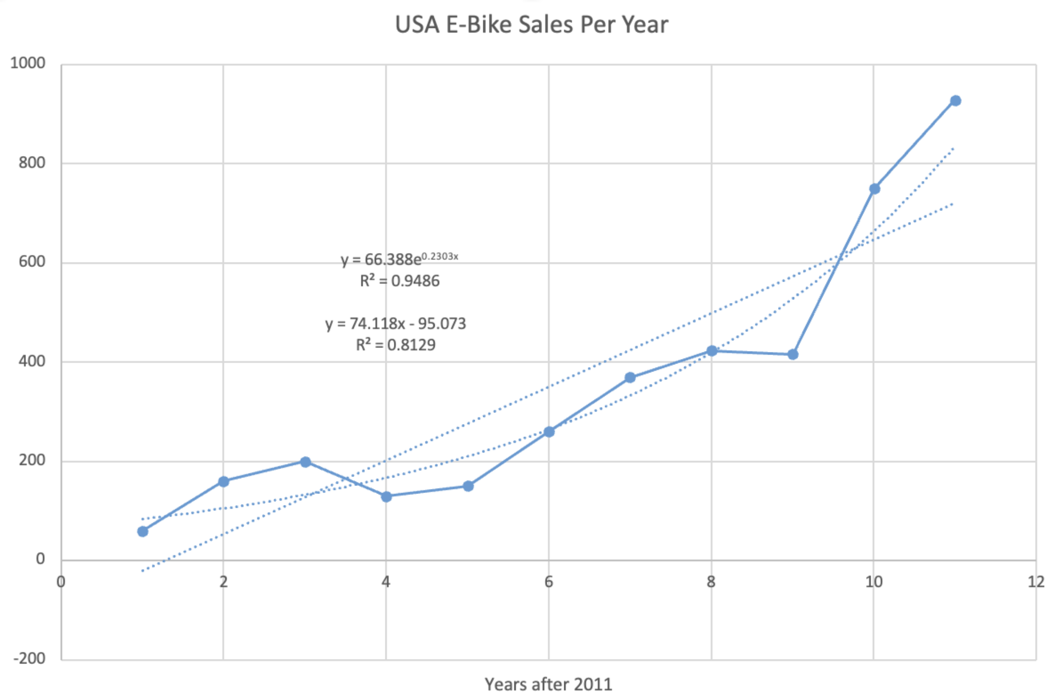


2.6 Results

We tested our theory with the data provided. For Europe and France, our exponential regression had an extremely high correlation. We tested the linear regression for Europe, which did not have a strong correlation. This confirmed our theory that emerging markets would match exponential functions. We also tested it on China's market, which is on the very established side of the e-bike market. As our theory suggested, once these markets establish growth would slow, and would no longer track an exponential. It's important to note that our exponential function only works for new and emerging e-bike markets.

We then tested our models on historical data of e-bike sales in the US [7].

Figure 5: Linear and exponential regression of USA E-Bike sales per year, in thousands of units. The exponential regression fits the US well, with a R^2 fit of 0.9486.



USA:

Year	Predicted (in thousands)	Actual (in thousands)
2018	333	369
2019	419	423
2020	528	416
2021	664	750
2022	836	928
2025	**1669**	N/A
2028	**3330**	N/A

Table 2: Predicted number of e-bike sales using our model.

To answer the original question, we project 1,669,000 e-bike sales in 2025, and 3,330,000 e-bike sales in 2028.

2.7 Sensitivity Analysis

As a regression model, this model is dependent on the past data points. Since it depends on the past years data, each individual year will not affect the regression by that much. This helps our model against potential variables that would cause uncertainty, such as a year where sales were particularly low due to an economic recession or an unfavorable policy. With so few variables needed in this model, the variability depends entirely on the historical data.

2.8 Strengths and Weaknesses

Strengths:

- This exponential model has a very strong fit for the emerging markets in Europe, France, and the USA.
- This model takes into consideration the distinction between emerging e-bike markets and well established markets.
- This model is incredibly simple, and only needs the historical data to forcast the next few years.
- It produces results that make sense in the context of the larger bike market.

Weaknesses:

- This model does not incorporate other information, so cannot adapt for policy changes, age group populations, urban and rural population changes.
- It cannot predict long term results.
- The model does not work for well-established e-bike markets.

2.9 Conclusion

Exponential growth has been a staple of e-bike's growth over the past 20 years. In the US and Europe, E-bike sales have grown over 25% compounded annually for the past ten years. Its adoption as a new era technology is undeniable, and countries like China have already implemented and established infrastructure, manufacturers, and markets for the e-bike. With similar trends hitting Europe and the US, we can expect to see e-bikes begin to become much more prevalent in urban areas.

3 Part II: Shifting Gears

3.1 Restatement of the Problem

In this problem, we are asked to mathematically model the correlation between e-bike growth and various factors such as personal finance, gas prices, environmental awareness, and fitness.

3.2 Assumptions

- *Google inquiries correspond to e-bike sales.* People who google e-bikes are likely interested in the benefits associated with e-bikes. By modeling the Google search traffic containing "electric bike," we can determine how likely e-bikes are to be sold and the market size.
- *A majority of e-bike sales are online.* With increasing emphasis on digital e-commerce in the 21st century, more and more sales take place through online mediums. Thus, measuring the Google search volume of electric bikes will provide meaningful insight on e-bike sales.

Symbol	Definition	Unit
$B(t)$	Google Search Volume for "e-bikes"	Normalized Volume
$E(t)$	Google Search Volume for "Environment"	Normalized Volume
$C(t)$	Google Search Volume for "Charging Station"	Normalized Volume
$F(t)$	Google Search Volume for "401(k)"	Normalized Volume
$I(t)$	Google Search Volume for "Inflation"	Normalized Volume
$N_a(t)$	Google Search Volume for "NASDAQ"	Normalized Volume
$N_y(t)$	Google Search Volume for "NYSE"	Normalized Volume
$H(t)$	Google Search Volume for "Health, Fitness and Wellness"	Normalized Volume
$P(t)$	Google Search Volume for "Pandemic"	Normalized Volume
$G(t)$	Google Search Volume for "Gasoline"	Normalized Volume
$S(t)$	Google Search Volume for "SIPP"	Normalized Volume
$L(t)$	Google Search Volume for "LSE"	Normalized Volume

Table 3: Variables analyzed in our model.

- *A positive correlation between the volume of Google searches of e-bikes and another topic is indicative of that topic being a driver of e-bike demand.* While we make a relatively big assumption, high search volume for a certain term such as "environment" likely means this concept is prevalent in a consumer's mind. Thus, if there is a high correlation between factors that can impact a consumer and e-bike sales, it is likely that these factors motivate the consumer to buy. A corollary to this assumption is that the same people who are searching for e-bikes are also searching for the other topic.

3.3 Variables

Variables that are a function of time t are in terms of weeks.

3.4 Model Development

After obtaining data from Google Trends for the past 10 years (1/1/2013 - 3/5/2023), we computed the Pearson product-moment correlation coefficients, a metric that can measure the correlation between the arrays.

The calculation for the correlation coefficient matrix R is as follows:

$$R_{i,j} = \frac{C_{i,j}}{\sqrt{C_{i,i}C_{j,j}}},$$

where C is the covariance matrix. In this case, C is a 2 by 2 matrix showing the pairwise correlation between searches for "e-bikes" and searches for a different factor. The covariance between two variables x and y, each with N samples, is

$$\frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{N-1}.$$

The value of R varies between -1 (strong negative correlation) and 1 (strong positive correlation).

For our purposes, we are looking for $R_{1,2}$, where the first variable x is the search volume for "e-bikes" and the second variable y is the search volume for the other factor. Note that $C_{1,1}$ and $C_{2,2}$ are just the variances of x and y, respectively. Therefore,

$$R_{1,2} = \frac{C_{1,2}}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}} = \frac{\frac{\sum(x_i-\overline{x})(y_i-\overline{y})}{N-1}}{\sigma_x\sigma_y}.$$

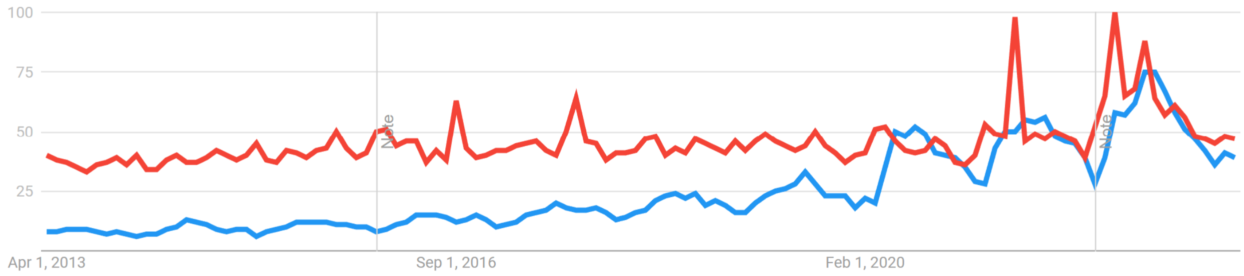


Figure 6: Gasoline (red). Electric bike (blue). Data shown for the past ten years for US.

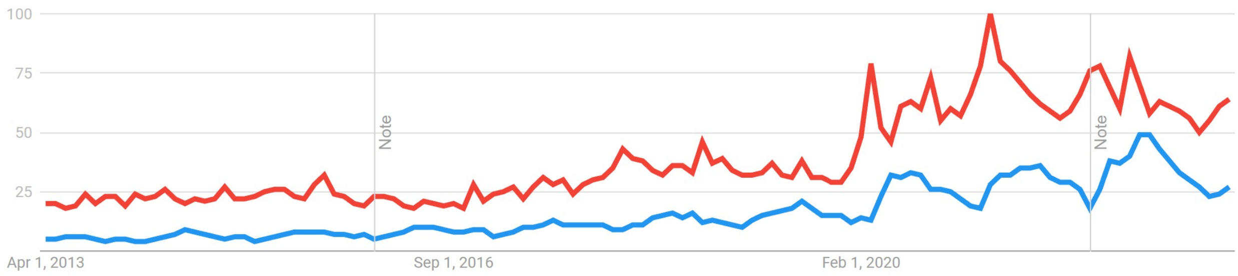


Figure 7: NASDAQ (red). Electric bike (blue). Data shown for the past ten years for US.

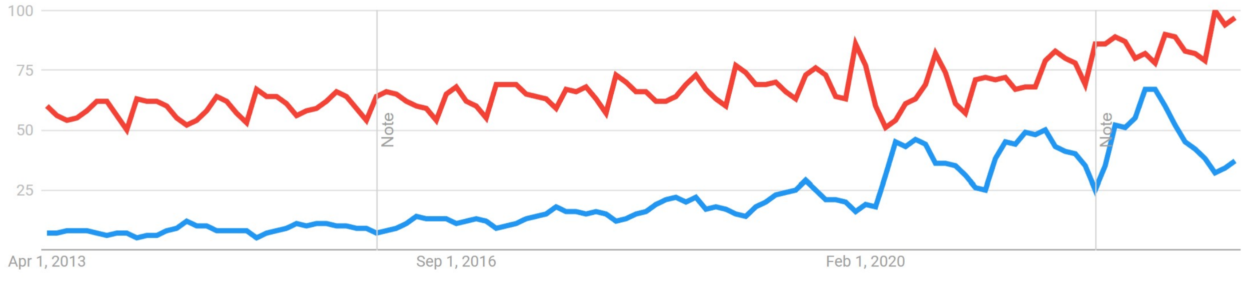


Figure 8: Health, Fitness, and Wellness (red). Electric bike (blue). Data shown for the past ten years for US.

Category	Search Term	Correlation
Personal Finance	401(k)	0.451
Personal Finance	NASDAQ	0.831
Personal Finance	NYSE	-0.051
Personal Finance	Inflation	0.728
Health Benefits	Health, Wellness and Fitness	0.607
Infrastructure	Charging Station	0.851
Infrastructure	Bike Rack	0.443
Environmental Awareness	Environment	0.138
COVID-19	Pandemic	0.358
Misc	Gasoline	0.596

Table 4: Correlation values for searches within the US

Category	Search Term	Correlation
Personal Finance	SIPP	0.282
Personal Finance	LSE	0.781
Personal Finance	Inflation	0.497
Health Benefits	Health, Wellness and Fitness	0.626
Infrastructure	Charging Station	0.632
Infrastructure	Bike Rack	0.555
Environmental Awareness	Environment	-0.269
COVID-19	Pandemic	0.361
Misc	Gasoline	0.605

Table 5: Correlation values for searches within the UK

3.5 Results

Tables 4 and 5 show our correlations for various search terms. The results show that environmental awareness (0.138 in US, -0.269 in UK) is consistently the lowest motivation for the e-bike market growth. The pandemic, health, infrastructure, and personal wellness are all major proponents of the e-bike market. Personal finance is a relatively big proponent for e-bike growth, specifically "NASDAQ", a stock exchange consisting of several big tech companies. This shows that when the tech industry is robust people are more likely to invest in e-bikes. Both charging stations and bike racks have high correlations with the e-bike market. This shows a positive feedback loop that continually propels the market forward. When people buy e-bikes, there is more incentive to invest in infrastructure which prompts further sales. Gasoline's high correlations with e-bikes show that gas prices are also a significant factor in e-bike sales. Health benefits (0.607 in US, 0.626 in UK) are also one of the biggest advantages of the e-bike.

3.6 Sensitivity Analysis

Given that the model is strictly based on Google Trends, there are no parameters directly within the model. However, one important factor to consider is the time frame utilized. We use a relatively large time window (10 years) which means that data from the past two years does not have that much impact. Correlations can be very different depending on the time window used.

3.7 Strengths and Weaknesses

Our model is good at analyzing information from online inquiries.
Strengths:

- Effectively captures the correlations between google search terms
- Incorporates search data from the past ten years
- Can be easily modified to analyze trends in different countries (Google provides data for every country)

Weaknesses:

- Correlation between search terms and e-bike market may not be completely statistically significant (i.e. trendy topics could overlap but not be necessarily related)
- Limited search items tested
- Dependent on the time frame window of the search queries

3.8 Conclusion

For Q2, we develop a model that calculates the correlations of various search terms from the past ten years. We utilize the Pearson product-moment correlation coefficients which show us that consumers have little to no regard for the environment. Rather, they focus on their own finances and wellness. Our model efficiently leverages big data from the past ten years. However, our model does make assumptions that correlations between search terms are indicative of the driving forces behind the e-bike market.

4 Part III: Off The Chain

4.1 Restatement of the Problem

For this problem, we will quantify the impact of the growth of e-bikes not in terms of market value, but instead in the net added societal benefit. Specifically we will quantify this in terms of the change in carbon emissions, road safety, noise pollution, and traffic congestion, categories we think are most significantly impacted by the change in e-bikes in the time-frame of the next 20 years.

4.2 Assumptions

E-bike buyers are those that want to use it for transportation. Though there is a small tourism and leisure market for E-bikes, they are primarily used in cities and for transportation.
Once someone buys an e-bike, they will continue using it until it breaks down. This assumption just means that they will not revert to using a normal bike, because tracking the percentage is unfeasible.
If someone's e-bike breaks down or stops functioning, they will buy a new one instead of reverting to biking. This assumption is fairly reasonable, and it ensures that when we observe the total number of users of e-bikes, users that have previously bought an e-bike will continue using one.
More e-bikes are the reason for the creation of more bike lanes. Everywhere in the US, the bike infrastructure is hardly enough, and in some cases, nowhere near enough, to sustain the number of riders that use them. As a result, as more people take to biking, we assume they will push for the creation of more lanes.

4.3 Assessing Number of E-Bikes Long-Term

To quantify the impact of E-Bikes on society, we need to know how many E-bikes there will be in the US over time. This calls for our model from part 1.

In Question 1, we gave a short-term exponential model for a country with an emerging e-bike market. However, because the percentage of e-bikes are so low currently for these emerging countries, they will not have a significant impact in the short term. Instead, we want to chart the long-term impact, as countries transition from the e-bike emerging market phase to the established market phase, which takes somewhere in the ballpark of 10-20 years for most countries.
To do this, we can no longer use an exponential curve, but rather a logistic curve fits this purpose best. The choice of a logistic curve is because there is a maximum to the number of e-bike users, and that as we near this maximum, the growth of e-bike users will decrease and eventually taper off. A more technical way of imagining this is similar to a virus, where

each person influences an average of x new people, but because as time goes on more people will already have an e-bike, eventually the growth tapers off. This is where the logistic curve comes in, which can also be used for charting a virus's total cases.

To create a logistic curve, we first consider the viable population that would potentially own an e-bike to be 45 million, which is the number of Americans that have used a bike for transportation in the last year. Since E-bikes are primarily used for transportation, the 45 million Americans would be the potential customer base. As the population change in America is a relatively small change compared with the rate of change of e-bikes, we will assume the viable customer bases stays at 45 million for simplicity.

Using this information, we can finally properly quantify the amount of e-bikes. First, we started with the total users in the most recent year, 2022, by summing the past sales of e-bikes in the US. With the short time-frame we are looking at, we can assume that all of these sales are first time sales, so we get about 3600 thousands of e-bike users currently. We then are able to find the next years recursively. Let P = previous years total users, s = the exponential rate, and N = current year's total users. Because at a small rate, the sales track an exponential, we found $s = e^{0.2303}$, or 1.26 with our linear regression. This means that every year, there is a 26 percent increase in e-bike users. So with this we can find:

$$N = P + P \cdot .26 \cdot \frac{45000 - P}{45000}$$

This equation stems from the fact that P is the previous value, then the added value for the current year will be: (old value)·(percent change)·(proportion of total population that is viable).

This tracks the logistic curve, and the results are plotted below:

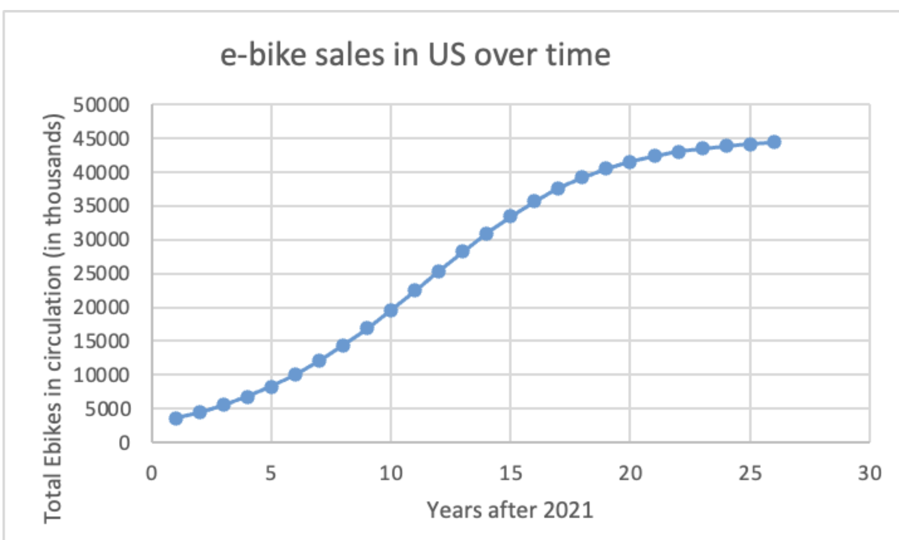


Figure 9: Logistic regression curve for the e-bike sales in US over time.

4.4 Model Development

We got the recursive equation, but to get the total e-bikers at any given time, we need a closed-form equation. To do this, we used a program to find the logistic regression of this equation. The results we found are below:

$$y = \frac{1417.033}{0.031 + 0.504 \cdot e^{-x \cdot 0.252}}$$

Using our model, in 2035 there will be around

$$\frac{1417.033}{0.031 + 0.504 \cdot e^{-14 \cdot 0.252}} = 29.5 \text{ million}$$

e-bike sales. This is a significant fraction of the total bike sales in the US.

4.5 Results

4.5.1 Net CO2 Emission Change

The effectiveness of e-bikes on net emissions depends on whether people transfer from a bike or car. Including manufacturing and energy emissions, an average car uses 229 g CO_2/km, where an ebike uses 14.8 g CO_2/km [8]. The number for E-bikes includes the CO_2 impact of the food needed to fuel the cyclist per km, which interestingly enough accounts for 43 percent of emissions. Regardless, an e-bike is over 15 times more energy efficient than a car. An e-bike is surprisingly more CO_2 efficient than a normal bike as well. Because of the amount of caloric energy that the body needs to cycle, and the CO_2 emissions needed to grow/raise our food, a normal bike is averaged at 21 g CO_2/km, making the e-bike the most energy-efficient form of transportation.
Since we used the assumption in the logistic curve that the e-bike would replace normal bikes and not cars for transportation, we will continue with that assumption here. To find the net CO_2 gain from e-bikes, we must find the number of km traveled. Using the census bureau's data, we find that the average cyclist's work commute is 3.5 miles or 5.6 km. Assuming they work for 240 days in a year after holidays and vacations, that gives us a total of 1344 km in a year. This means that the difference in CO2 emissions for switching to an e-bike for an average American is $1344 \cdot (21 - 14.8) = 8332$ g CO_2, or 8.3 kg CO_2.
Now we can graph the net CO_2 impact over time as e-bikes are implemented, and to do this we just multiply the number of bikers by the average saved CO_2 emissions to get the yearly total saved emissions. This gives us:

$$y = 8.332 \cdot \frac{1417.033}{0.031 + 0.504 \cdot e^{-x \cdot 0.252}}$$

where y is in kg CO_2 emissions saved.

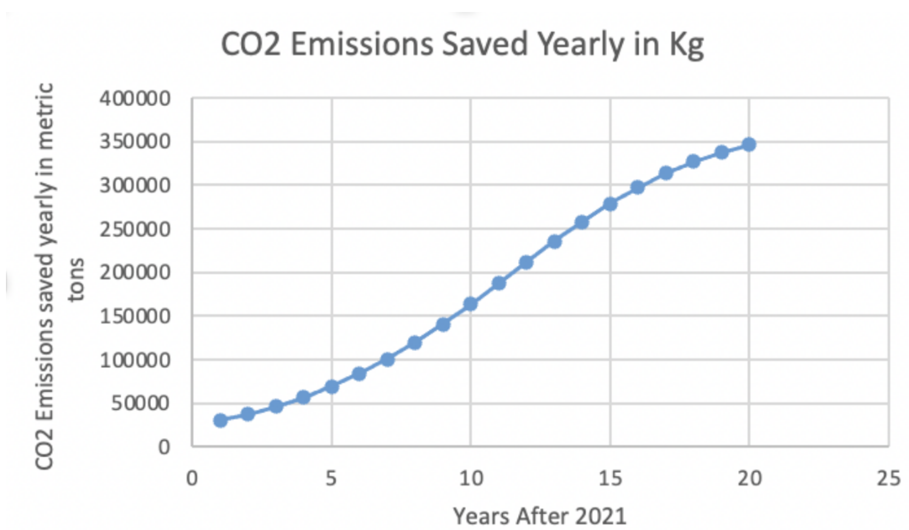


Figure 10: Saved CO_2 Emissions from E-bikes in the US yearly.

From now until 2040, e-bikes will save us 2.96 million metric tons of CO_2. Then after that, we will save 360,000 metric tons of CO_2 per year since e-bikes will nearly be fully rolled out. 360,000 metric tons of CO_2 is equivalent to 40,509,000 gallons of gasoline used, or 77600 gas-powered cars for a year, which is a pretty significant contribution [9]. Another factor for emissions will come from the conversion from cars to bikes, which is a 15 times reduction in emissions per mile. This number is unpredictable and hard to quantify, because it involves a lifestyle change, and data is hard to find.

4.5.2 E-bike safety

For the section, we make a couple of specific assumptions that do not apply to the other subproblems:
The number of e-bikes sold is roughly equal to the number of e-bike riders. People buy new e-bikes to ride them, so it is reasonable to assume that most new e-bikes bought are used. However, there may be e-bikes bought in previous years that are still being used, but e-bikes can break down, so some people who bought e-bikes in the past must buy a new e-bike in

the future. In addition, not all new e-bikes will be used, which balances out the fact that not all e-bikes being used are bought in the given year.
No new safety precautions such as creating more bike lanes are implemented. We make this assumption in order to show the effects of not adapting to this new and fast-growing mode of transportation.

The main difference between e-bikes and bikes is that e-bikes can go faster and usually do not require the rider to pedal. Other e-bikes are pedal-assisted, meaning that they require the rider to pedal in order to generate extra energy when going uphill. However, data has shown that e-bikes can be more dangerous than regular bikes. According to [10], e-bike victims have a 17% chance of internal injury, while pedal-bike accident victims have a 7.5% chance of internal injury. This means that e-bike victims have $\frac{17\%}{7.5\%} = 2.267$ times the chance of having an internal injury.
Currently, the total number of bikers in the US is roughly 52 million [11], and for the sake of simplicity, we assume that this number will remain constant. This number agrees with our logistic regression, which predicts that the number of e-bike sales (which we assume is roughly the peak number of e-bikers) will plateau at around $\frac{1417.033}{.031} = 45.7$ million; at this point, there would be a small minority of roughly 6.3 million bikers using regular bikes.
In 2022, there were $928,000$ e-bikes sold, which corresponds to roughly $928,000$ e-bike users, and roughly $52,000,000 - 928,000 = 51,072,000$ pedal-bike users. Suppose that the chance of any given biker having an internal injury is p. Then, we can reasonably assume that the chance of any given e-biker having an internal injury is $2.267p$ from our previous calculations. Then, the total risk comes out to be

$$928,000 \cdot 2.267p + 51,072,000 \cdot p = 53,175,776p.$$

Using our predictions for the year 2035 we project that there will be $29,500,000$ e-bikers and $52,000,000 - 29,500,000 = 22,500,000$ million bikers. **If the risks of riding e-bikes and riding bikes stay the same, then the increase in e-bike usage could be disastrous.** Using the same risk of internal injury as in the previous scenario, the total risk comes out to be

$$29,500,000 \cdot 2.267p + 22,500,000 \cdot p = 89,376,500p.$$

The percent increase in risk is

$$\frac{89,376,500p - 53,175,776p}{53,175,776p} = 68\%.$$

4.5.3 Traffic Congestion and Noise Pollution

Between 2007 and 2014, New York added over 31 miles of bike lanes. In these seven years, traffic congestion remained the same for a variety of reasons. These results end the stigma that adding more bike lanes worsens traffic congestion. On Columbus and 8th avenue, two major roads in New York in different areas of Manhattan, traffic congestion decreased by 14% after the addition of bike lanes. This amazing decrease in an area so crucial to the city's infrastructure is a direct result of car commuters switching to bikes and the fact that the bike lanes only take one lane away from the road. In order to accommodate an entire bike lane, as well as a boundary between it and the road itself, lanes were simply made thinner, allowing for only one lane to be taken away. The loss of one lane on Columbus and 8th results in a major decrease in congestion, and in the entire city as a whole, there is only a <1% increase in congestion using this method of integrating the 31 miles of bike lanes. Therefore, as we follow our model into 10 and 15 years after 2021, the increase in traffic congestion will be virtually negligible despite the addition of bike lanes to accommodate for their demand with all the new e-bikes on the road [12].
When looking at almost purely cycling-only areas like Central Park and Prospect Park that are still within the city, the average decibel level is about 36 dB. It is far less compared with the average decibel levels of the entire cities of Montreal and Paris at 73.4 dB and 70.7 dB, two cities at the front of becoming bike-able, but with only 2% and 3% share of transportation. However, when compared with Copenhagen, a city that has a history of having so many bikes, its decibel level is still 68dB despite bikes being 38% of its urban transportation. Cars are a huge percentage of this noise, especially in the vast majority of

the city, which is a significant distance from any major sound producer, which we found to be ports and factory districts in particular. These results reinforce our statement that traffic congestion remains the same, despite the addition of bikes, since the vast majority of car noise pollution comes from angry drivers honking and yelling, along with cars stopping and starting in bumper-to-bumper traffic. Yet, once cities reach a point where cars become so obsolete, there will be a significant change in noise data. Our data suggest slightly under 350,000 bike sales in 2041 which is not enough to significantly decrease noise pollution [13].

4.5.4 Analyzing Government Infrastructure Savings

Thesis: Since electric bikes damage roads significantly less than cars, the government needs to spend considerably less repaving roads that have a large proportion of total traffic as electric bike users compared to similarly trafficked roads that have fewer bikes and more cars.
Subproblem-specific assumptions:
An electric bike does not damage roads at all. Given that the damage from vehicles is proportional to the weight and that the average car weighs 4,000 lbs and an e-bike weighs 80 lbs on the high end (0.2%), the damage caused by e-bikes is insignificant compared to the damage caused by cars. Given time constraints we assume that public transportation such as buses also does not damage roads.
The total number of e-bikes and passenger-vehicle passenger-miles increases linearly over time. This simplification allows us to project the passenger-miles years into the future.
All bicycles in the future are e-bikes. This assumption is clearly wrong, but it is reasonable to assume that over time as the price decreases most consumers will choose e-bikes over regular bicycles as they will want better e-bike performance.
40% of all e-bike miles traveled would have been traveled with a car in the absence of a bicycle [14]. This statistic is calculated with data from a study conducted in the Netherlands in 2015. We acknowledge there are likely cultural differences between traveling in the US and the Netherlands, but given time constraints this is the best number we could find.

This section aims to define a function $c(t)$ which contains the total costs saved in road repair t years after 2023. We start by calculating linear regressions for the $m_b(t)$ and $m_c(t)$, the number of bike and car miles in the US for t years after 2023. Linear regressions show that

$$m_b(t) = 0.8975t + 40.211$$
$$m_c(t) = 64.489t + 5810.965$$

Next, we want to find the total bike and car miles l years after 2023. This was achieved via the following integral:

$$B_l = \int_0^l m_b(t)dt$$
$$C_l = \int_0^l m_c(t)dt$$

The automotive miles saved s is defined as follows:

$$s = B_l \cdot 0.4$$

Dividing this quantity by the total number of automotive miles C_l and multiplying by the annual government spending on road repairs 110 billion yields $c(t)$ with $c(t)$ in millions of dollars

$$c(t) = \frac{s}{C_l} \cdot 110000$$

Table 6 shows the amount of money saved through e-bike adoption over time. Approximately 300 million government dollars are saved yearly by the government in road repairs. This number appears low because the average electric biker only travels 5.7 miles [15] whereas the average car trips are around 40 miles [16]. Thus, the total mileage saved is quite low since cars possess the ability to travel such long distances. However, in general, e-bikes can save a potentially significant amount.

Years After 2023	Money Saved (M)
1	306.172
2	615.705
3	928.544
4	1244.634
5	1563.924
6	1886.360
7	2211.895
8	2540.476
9	2872.058
10	3206.591
11	3544.031
12	3884.330
13	4227.446
14	4573.334
15	4921.953
16	5273.259
17	5627.213
18	5983.774
19	6342.903
20	6704.562

Table 6: Total amount of money saved in road repairs l years after 2023.

4.6 Sensitivity Analysis

Our model in (4.4) is dependent on the total number of bikers possible, as we assume that almost all bikers will be converted to e-bikers.

4.7 Strengths and Weaknesses

Strengths:

- CO_2 Emission model: Our model is simplistic and uses the previous model we developed for the number of e-bikes sold yearly.
- Infrastructure model: Our model is flexible and can easily calculate the governmental savings in any year in the future while our assumptions hold.

Weaknesses:

- E-bike safety prediction: Our prediction is clearly a rough estimate and is not super precise due to the assumptions we had to make. However, we believe that the model still provides meaningful value, as it shows what could happen if safety precautions are not taken.
- Infrastructure model: Some of our assumptions are presently incorrect and will likely skew the results for years soon in the future.

4.8 Conclusion

Overall, we show that an increase in the number of electric bikes will have a significant effect on society. Using our logistic curve model to predict the number of e-bike sales over the next 20 years, we are able to quantify the change that the surplus in e-bikes will create in the near future. Currently, riding e-bikes are particularly dangerous with the current infrastructure that exists. Accidents on e-bikes are at an all-time high and significantly more common than on regular bikes. With more bikes and assuming that there is no change in our current bike infrastructure, these injuries will only increase, and e-biking will become unsafe. If improvement in bike infrastructure continues at its current rate and with current methods, there will be no change in traffic congestion and noise pollution. When it comes to government saving on infrastructure, e-bikes are far less damaging to roads. Therefore, with more e-bikes and fewer cars on the road, the government will be spending 80 million dollars less on road infrastructure as a whole.

5 Conclusion

In Part I, we determined that future sales of e-bikes in the United States can be modeled with an exponential function. We used exponential regression on sales from 2018-2022 to predict the sales of e-bikes in 2025 and 2028.

In Part II, we then used Google Trends data to find the correlation between searches for e-bikes and various other terms. We found that personal finance and concerns over one's health and wellness are the largest determinants of e-bike sales while environmental concerns are not a major driver of sales.

In Part III, we looked at the consequences of the growth of e-bikes: CO_2 emissions, road safety, traffic congestion, noise pollution, and government infrastructure savings. We found that e-bikes would have a positive impact in these areas if safety measures were taken to reduce the risks of increased e-bike ridership.

As battery technology gets better over time and e-bikes become more affordable for the average American, ridership will continue to increase over time. Recognizing that e-bikes can replace gas-powered vehicles is critical as policymakers can craft legislation to accelerate this trend and lead the nation into a healthier, environmentally-friendly future.

References

[1] "Traffic Congestion and Reliability: Trends and Advanced Strategies for Congestion Mitigation", https://ops.fhwa.dot.gov

[2] MacArthur et al., "A North American Survey Of Electric Bicycle Owners", https://www.calbike.org/wp-content/uploads/2019/02/A-North-American-Survey-of-Electric-Bicycle-Owners.pdf

[3] "Age distribution in the United States from 2011 to 2021", https://www.statista.com/statistics

[4] "United Kingdom: Age distribution from 2011 to 2021", https://www.statista.com/statistics

[5] Balton, Jeff, "Bike Statistics & Facts [Of 2023]", https://www.bicycle-guider.com/bike-facts-stats/

[6] "Projections for the global electric bike market between 2018 and 2030", https://www.statista.com/statistics/1261084/global-e-bike-market-forecast/

[7] Ride Like the Wind, MathWorks Math Modeling Challenge 2023, https://m3challenge.siam.org/node/596.

[8] "What is the carbon footprint of an ebike?", https://modmo.io/blogs/news/what-is-the-carbon-footprint-of-an-ebike

[9] "Greenhouse Gas Equivalencies Calculator", https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator#results

[10] "Are E-Bikes Safer than Regular Bikes?", https://www.peoplepoweredmovement.org/are-e-bikes-safer-than-regular-bikes/

[11] McMahon, Conor, "28 BICYCLE INDUSTRY STATISTICS [2023]: MARKET SIZE, SHARE, GROWTH, AND TRENDS", https://www.zippia.com/advice/bicycle-industry-statistics/

[12] Stromberg, Joseph, "Bike lanes have actually sped up car traffic in New York City", https://www.vox.com/2014/9/8/6121129/bike-lanes-traffic-new-york

[13] "CyclistsÉxposure to Road Traffic Noise: A Comparison of Three North American and European Cities", Apparicio, Philippe and Gelb, Jeremy, https://www.mdpi.com/2624-599X/2/1/6

[14] Lee et al., "Electric Bicycle Use and Mode Choice in the Netherlands", https://journals.sagepub.com/doi/abs/10.3141/2520-01?journalCode=trra

[15] Sutton, Mark, "Electric bike owners progressively use cars less, finds study", https://cyclingindustry.news/electric-bike-owners-progressively-use-cars-less-finds-study/

[16] Gauthier, Michael, "QOTD: How Far Do You Typically Drive Each Day?", https://www.carscoops.com/2022/04/qotd-how-far-do-you-typically-drive-each-day/

6 Appendix

correlation.py

```python
import pandas as pd
import numpy as np
import os

filenames = ['environment.csv', 'health_fitness_wellness.csv',
'charging_station.csv', 'gas.csv']

for filename in filenames:
    data = pd.read_csv(os.path.join('data', filename), skiprows=2)
    arr = np.array(data.iloc[:, 1:3]).T
    corr = np.corrcoef(arr)[0][1]
    print(f'{filename}: {corr}')
```

LogisticRegressionFit.py

```python
import numpy as np
import scipy.optimize as opt
import matplotlib.pyplot as plt

x = np.array(range(1,27))

y = np.array([3629.76966, 4493.87434, 5541.34286, 6799.56312, 8294.24472, 10046.1489, 12


def f(x, a, b, c, d):
    return a/(b + c*np.exp(-x*d))

(a_, b_, c_, d_), _ = opt.curve_fit(f, x, y)

y_fit = f(x, a_, b_, c_, d_)
fig, ax = plt.subplots(1, 1, figsize=(6, 4))
ax.plot(x, y, 'o')
ax.plot(x, y_fit, '-')
```

BikeMileRegression.py

```python
import numpy as np
import sklearn.linear_model
import pandas as pd
import matplotlib.pyplot as plt
def LinearRegression(x, y):
  model = sklearn.linear_model.LinearRegression()
  model.fit(x,y)
  print(model.coef_, model.intercept_)
  plt.plot(x, y)
```

```
  plt.plot(x, model.coef_[0]*x + model.intercept_)
  plt.xlabel('Years after 2023')
  plt.ylabel('Total Bike Miles')
  df = pd.read_csv('TransportationMiles.csv')
columns = [1, 3]
for column in columns:
  col = df[['Year', df.columns[column]]]
  col = col.dropna()
  x = np.array(col.iloc[:,0])
  x = np.expand_dims(x-2023, axis=1)
  y = np.array(col.iloc[:,1])
  LinearRegression(x,y)
```