The Smoking Gun: A Study of Destructive Juvenile Vaping and Drug Trends

by: Davis Keene, Eric Chen, Ethan Lee, Josh Fleckner, and Shaashwat Sharma

Table of Contents

Executive Summary 2

Question 1: Darth Vapor 3

2.1) Restatement of the Problem 3

2.2) Assumptions 3

2.3) Developing the Model 4

2.4) Validating the Model 8

2.5) Results 9

2.6) Strengths and Limitations 10

Question 2: Above or Under the Influence? 10

3.1) Restatement of the Problem 10

3.2) Assumptions 11

3.3) Developing The Model 11

3.4) Validating the Model 12

3.5) Results 14

3.6) Strengths and Limitations 14

Question 3: Ripples 15

4.1) Restatement of the Problem 15

4.2) Assumptions 15

4.3) Developing the Models 15

4.4) Validating the Models 18

4.5) Results 19

4.6) Strengths and Limitations 20

Conclusions 20

References 23

Appendix 27

The Smoking Gun: A Study of the Juvenile Vaping and Drug Trend

1. Executive Summary

The first task of our challenge asked us to predict nicotine usage over the next ten years. In attempting to find statistical backing for a graph, we turned to trends in the most similar epidemic in recent history: cigarette smoking. Our group projected that e-cigarette usage among children would increase over a number of years, but at a decreasing rate, so we settled on a logarithmic function graphing e-cigarette usage with respect to time. We expect that, as with cigarettes, it will take generations of e-cigarette users to gather data to form a strong claim about the health effects e-cigarettes. As a result, we expect that e-cigarette manufacturers will maintain their shield of doubt for at least the next decade.

For the second task, we looked at studies on the correlation between certain drug usage and external factors, such as race, mental health, gender, and sexual orientation. Using the probabilities from those studies, we coded a program which generates a mock random sample of high school seniors, and approximates the probabilities that each one will become addicted to each of four drugs (nicotine, opioids, alcohol, and marijuana) based on the aforementioned factors. We verfied the approximate accuracy of the simulation by observing that the total proportion of users of each of the four drugs out of the mock senior class roughly matched the national high school averages for each of those drugs.

For the third task, we were asked to evaluate the costs of usage of the four drugs mentioned in the second task. Assessing three categories (economic loss, environmental destruction, and death), we found that each drug was detrimental to each of the three categories in varying degrees using a model that took into account the fluctuating population of users of each of the substances. Using a Markov Chain, we were able to include into our model the long-term probability of a drug user transitioning to remission and vice-versa. Using this, we ranked each of these four substances based on their Destructive Impact Index, a score metric we made.

By answering each of these questions presented, we look forward to having an impact on the future of awareness on drugs, nicotine products, and alcohol.

2. Question 1: Darth Vapor

2.1) Restatement of the Problem

The problem asks to determine a model for the next ten years regarding the spread of nicotine due to e-cigarettes.

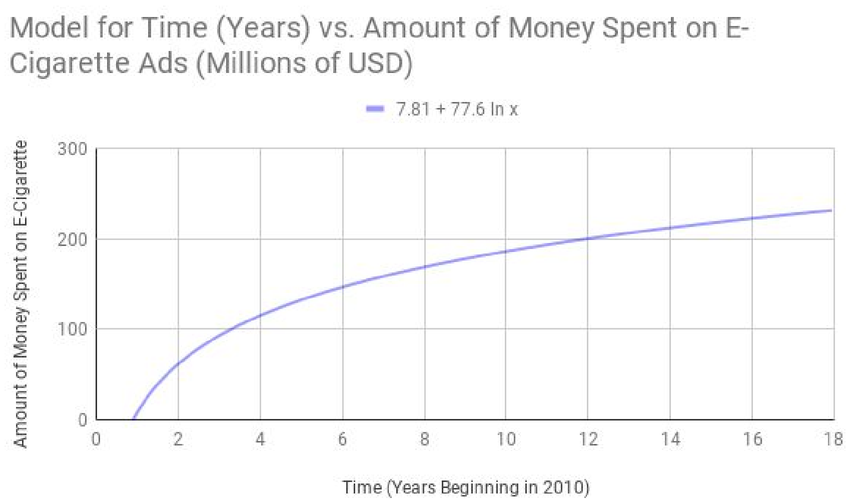
2.2) Assumptions

1. The percent of people who use electronic cigarette products will reach a plateau at a certain point in time.
    a. Justification: If there was no way for this value to plateau off at a certain percentage, then the percent of people using e-cigarette products would continue to grow until over 100% of the population is using them, which is not possible.
2. There will be no legislation created over the next ten years that effectively serves to significantly decrease the use of e-cigarettes, cigarettes, or other nicotine products.
3. The percentage of the population who use cigarettes will decrease over time but will not reach 0%; instead, it will plateau off to a certain value.
    a. Justification: As can be seen in other activities, a trend or habit is not ever completely eradicated in society, but instead decreases until it hits a certain value very close to 0.
4. Children from the age of infancy to middle school do not vape or smoke at statistically significant rates, and were not considered in the modeling for this question.
5. The value of the US dollar will not change drastically in the next ten years, nor has it in the last ten years.
6. The proportion of adults in the world has remained and will stay constant at 65% for the next ten years.
7. The population of the world for now and the next ten years is 6.87 billion people[35].

2.3) Developing the Model

In our model, we took into account the fact that many social norms and trends such as the smoking of cigarettes or e-cigarettes is often stimulated by the amount of advertising aimed at potential customers. This is because e-cigarette companies are looking to expand their market with the decline of cigarette usage. Thus, we used the amount of money spent on e-cigarette advertising to create a model predicting the percent of people in the world who use e-cigarette products.

To begin, we used a dataset showing the amount of money spent by e-cigarette companies on advertisements (in millions of US dollars). Because the data only included statistics from 2011 to 2014, we created a linear regression model to predict what the amount of money spent by these companies on advertisements would be in 2015 to 2018, inclusive:



(Figure 1)

Using this model, the number of millions of dollars used every year for e-cigarette advertisements was calculated:

Year	**Money Used for E-Cigarette Advertisements (Millions of Dollars)**
2011 (1)	7.81
2012 (2)	61.59
2013 (3)	93.05
2014 (4)	115.36
2015 (5)	132.68
2016 (6)	146.83

2017 (7)	158.79
2018 (8)	169.15

Table 1

Next, we used datasets showing the percent of high schoolers who vape, as well as the percent of adults who vape[22]; these datasets were weighted based on the current proportion in the world of adults and high schoolers[34]. In order to sufficiently validate the model we produced, the two underlined data-points were omitted while creating the model to be used as a testing set later on.

Year (years since 2010)	**Percent of High School E-cigarette users**	**Percent of Adult E-cigarette users**	**Weighted Average of Percent of E-cigarette users**
2011 (1)	1.5	0.17	0.25
2012 (2)	2.8	0.22	0.39
2013 (3)	4.5	0.25	0.55
2014 (4)	13.4	0.43	1.40
2015 (5)	16	0.47	1.64
2016 (6)	11.3	0.50	1.28
2017 (7)	11.7	0.55	1.35
2018 (8)	20.8	0.83	2.29

Table 2

These values were then used to find a linear regression showing the correlation between money used on e-cigarette advertisements and the overall percent of E-cigarette users in the world. The equation of the regression is as shown, with x being equal to time in years where 2011 is represented as 1, 2012 is represented as 2. etc:

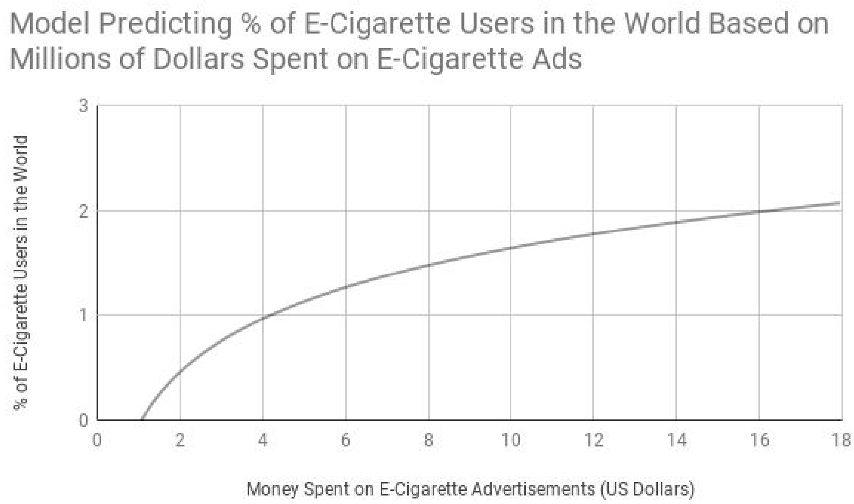


Figure 2

$$g(x) = 0.009461483 * f(x) + 0.0237499932$$

where f(x) is the amount of money spent on e-cigarette advertisements:

$$f(x) = -7.81 + 77.6 * ln(x)$$

The model to find the trends in cigarette spread over the next ten years was calculated by finding a linear regression model for the percent of people in the world using cigarettes. The equation calculated, where x is time in years since 2010, was:

$$c(x) = 15.21878157 * (0.9072916194)^x$$

Lastly, to find the amount of nicotine spread due to the use of vape and cigarette products, the following models were used for each one. First, for cigarette products, the average amount of nicotine per cigarette is 12 mg; thus, the equation to find the amount of nicotine consumed by smokers in the world, where w is the world population, is:

$$n1(x) = c(x) * w * 12 * 0.01 \text{ mg}$$

For the amount of nicotine in e-cigarettes, because e-cigarettes can be bought at vastly differing levels of nicotine (most commonly 0mg, 6mg, 12mg, 18mg, 24mg, and 36mg), a normal distribution model was used to find the probability that the level of nicotine in a vape is at a certain level[11]. The standard deviation used was 6, as that is the most common increment between the amounts of nicotine in an e-cigarette. Then, this model was used to find a rough estimate of the probability that a vape had a nicotine level over 0 mg:

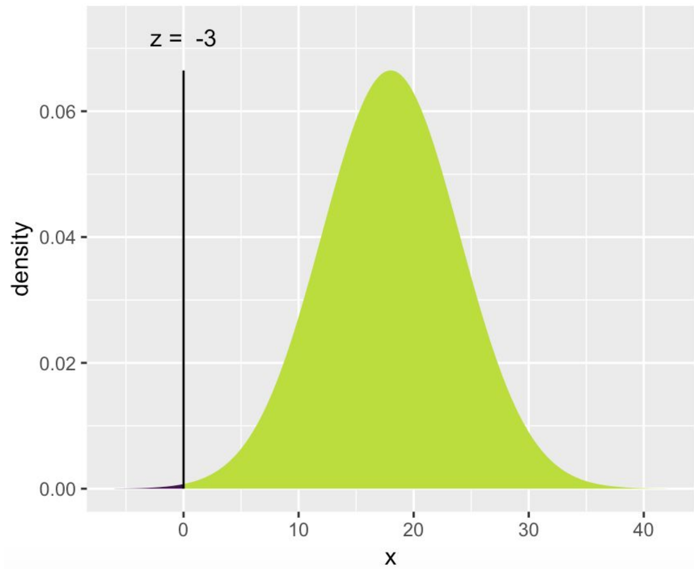


Figure 3

The probability of 0.9987 calculated from this approach was multiplied by the assumed average nicotine level of 18 mg in a vape, giving the following equation for the amount of nicotine spread due to vapes:

$$n2(x) = 0.9987 * 18 * w * g(x) * 0.01 \text{ mg}$$

2.4) Validating the Model

To validate our model, we used two data points (years 2017 and 2018) as a test set. We did not include these data points when creating the model. To test the accuracy of our model, we used it to calculate the percentage of E-cigarette users in 2017 and 2018. The results the model calculated were a 1.378% E-cigarette usage rate for 2017 and 1.476% in 2018. The residual for 2017 was -.028%, while the residual for 2018 was 0.814%. The extremely small residual for the 2017 data point shows the accuracy of the model, while the 2018 data point cannot be considered as very revealing of the trends in the data as it is a severe outlier. Thus, using these residuals contributes to the validation of the model.

Additionally, this model can be validated logically. Based on the shape of the graph of the model for e-cigarette user percentage we discovered, the percent of people in the world who use e-cigarette should plateau off to a certain percentage within a few hundred years. This is a reasonable conclusion, because the number of people in the world using such products would most likely not increase linearly or exponentially; this would cause this number to exceed the world population.

2.5) Results

Using our model for the amount of nicotine consumed in the world due to cigarettes and e-cigarettes over the next ten years, we found that $n1(19) = 2.21 * 10^9$ mg of nicotine would be consumed due to cigarettes in the world in 2029. Additionally, we found that $n2(19) = 2.89 * 10^9$ mg of nicotine would be consumed due to vapes in the world in 2029. Thus, it can be assumed that while the amount of nicotine consumed due to e-cigarette products will increase substantially over the next ten years, the amount of nicotine consumed due to cigarettes will decrease over the next ten years, with the impact of e-cigarettes overtaking that of cigarettes.

To compare the current growth of e-cigarettes with the growth of cigarettes over the 1900's, the following chart was used:

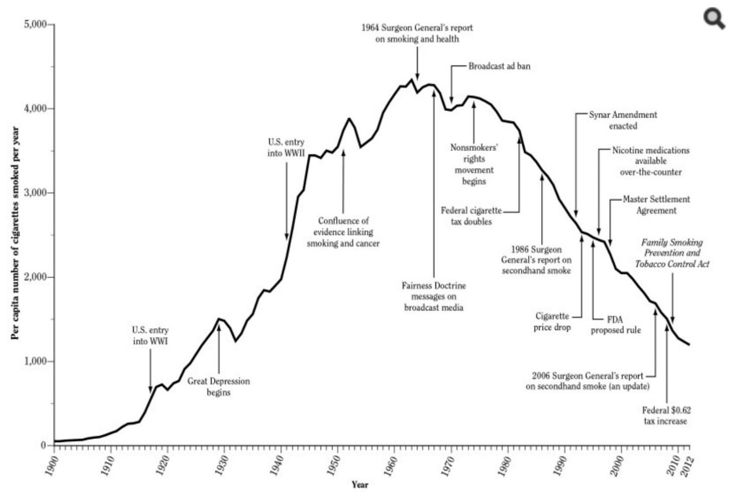


Figure 2.1 Adult* per capita cigarette consumption and major smoking and health events, United States, 1900–2012

Figure 4

Similarly to the early growth of cigarettes, e-cigarettes are currently undergoing a steep increase in popularity and sales. However, based on our model, the number of e-cigarettes will reach a certain peak and plateau off, just as cigarettes did between the years 1900 and 1980. Although the number of cigarettes smoked per year per capita did decrease overall, our model differs from the above graph in this way because we assume that there will not be another nicotine product that overtakes the e-cigarette in popularity.

Thus, our model shows that the current growth of e-cigarettes is similar to the previous growth of cigarettes.

2.6) Strengths and Limitations

The strengths of our model lies in the fact that it acknowledges a variety of specific criteria while being a function of one easily quantifiable variable. Firstly, it accounts for the fact that the percent of world population who smoke both e-cigarettes and cigarettes will plateau off at a certain point. This is because there is most likely a certain percent in the world that will not be using nicotine products permanently, while another percent will loyally use such products. Second, it shows the effect of social pressures through the amount of money spent on e-cigarette advertisements. Such advertisements are usually targeted at demographics most vulnerable to these nicotine products, meaning that they have a direct correlation to the percent of people using these products. In these ways, our model shows a strong predictor for vape and cigarette usage in the world.

There are several limitations in our model. First, the sample size used is very small; the datasets used to determine the percent of high schoolers and adults who use cigarettes and e-cigarettes may not have been sufficient to be representative of the world population. Second, there was no data available that specifically determined the total amount of nicotine consumed due to e-cigarettes and cigarettes, because of the varying levels of nicotine contained in these products. Thus, this was not accounted for in the model. Lastly, the world population is constantly changing, and the model was tested only on one specific value of the world population (6.87 billion people). Thus, finding some way to implement the dynamic nature of the population of the world would have improved the model.

3. Question 2: Above or Under the Influence?

3.1) Restatement of the Problem

The problem asks us to create a model that simulates the likelihood that a given individual will use a given substance (nicotine, alcohol, marijuana, and narcotics).

3.2) Assumptions

1. Assuming the simulation is taken at the end of high school, all high school seniors are 18 years of age.
2. All people sampled are not active drug users.

3. The law does not affect drug usage by a significant amount.
4. All the people in the sample are exposed to the same type and amount of advertisements.
5. All descriptive characteristics are equally important in determining one's likelihood of substance abuse.
6. The demographic makeup of the 300 students generated by the model are all randomly generated, and represent the greater US population.

3.3) Developing The Model

In our model, we took into account that each one of the multiple factors contributed differently to the probability of a certain person abusing that specific drug. To compare the likelihood of an individual with specific characteristics would use illicit substances, we first needed to develop a way to simulate a random sample of a population of teenagers. To compute the probability that a specific person uses illegal substances, we had to gather probabilistic data about certain demographics and how likely each demographic was to regularly use each substance, denoted by P_j, where j is the substance. Then, to compute the overall likelihood, we calculated the probability, ζ, that a person regularly used either alcohol, nicotine, marijuana, and opioids.

We used the following parameters in our analysis of the problem: a person's ethnicity, sexual orientation, the gender with which they identify, and their overall happiness. Since we couldn't find a census or sample with the data that we wanted, we simulated taking an SRS from the population of the United States and see how well our model compared to official US data reports [1,7,9,13,15,19,23,25,27,30,36].

Parameters	$P_{\text{nicotine use}}$	$P_{\text{marijuana use}}$	$P_{\text{opioid use}}$*	$P_{\text{alcohol use}}$
Black	.168	0.14	.047	.22
White	.166	0.11	.047	.26
Hispanic	.101	0.166	.047	.3
Asian	.07	.21	.047	.18
Gay	.206	.4	.047	.25
Straight	.149	.166	.047	.13
Male	0.167	0.17	.047	.23

Female	0.136	0.15	.047	.19
Depression	.28	.18	.047	.165

Table 3

**Note: $P_{opioid\ use}$ remains .047[23] across all parameters because there was no data for opioid use among high schoolers split by demographics. The .047 number comes from the overall average rate of high school opioid use over all demographics.*

For each individual, the probability that they are using a certain drug *i* can be modeled as such:

$$\varsigma_i = \frac{\sum_{j=0}^{n} P_j}{n},$$

where n is the number of parameters used and j is each parameter's index. The proportion of a sample of size *s* that regularly uses a substance then, is as follows:

For all individuals in the population:

$$if : Random\ Integer(1, 99) < \zeta * 100$$

$$\supset$$

$$then : \ number\ of\ individuals\ that\ regularly\ use\ substance\ +\ 1$$

$$sample\ proportion : \ \frac{total\ number\ of\ students\ that\ regularly\ use\ substance}{s}$$

3.4) Validating the Model

To test this model, we compared the results we got from generating 300 high school students and calculating how many will regularly use alcohol, nicotine, and marijuana to the true proportion of high schoolers that regularly use alcohol, nicotine, and marijuana.

Our model calculated that $\frac{49}{300} \approx .163$ of the high schoolers regularly use alcohol

Our model calculated that $\frac{53}{300} \approx .176$ of the high schoolers regularly use nicotine

Our model calculated that $\frac{48}{300} = .16$ of the high schoolers regularly use marijuana

Our model calculated that $\frac{12}{300} = .04$ of the high schoolers regularly use opioids

True proportions:

The true proportion of high schoolers who regularly use alcohol is .154

The true proportion of high schoolers who regularly use nicotine is .177

The true proportion of high schoolers who regularly use marijuana is .166

The true proportion of high schoolers who regularly use opioids is .047

The mean error of the calculated values is -1.25%, which provides strong evidence supporting the validity of our model.

Probability of Usage among our Team

To see if this model can accurately predict the chance of drug usage for real-world individuals, we decided to try it on ourselves. Since none of us are addicted to any of the drugs above, we were all qualified control subjects to test this model. The results are as shown below. Davis & Josh (white) are both similar in traits which explains their near-identical probabilities, followed by Ethan (asian) then Shaashwat and Eric (other).

Drug Person	Davis	Josh	Ethan	Shaashwat	Eric
Alcohol	.1850	.1856	.1592	.1480	.1485
Nicotine	.160	.160	.1286	.158	.158
Marijuana	.1486	.1486	.182	.168	.168
Opioids	.047	.047	.047	.047	.047

Table 4

3.5) Results

Our results from generating 300 high school students and calculating how many will regularly use alcohol, nicotine, opioids, and marijuana gave us these results:

1) 49 out of 300 will regularly use alcohol

2) 53 out of 300 will regularly use nicotine
3) 48 out of 300 will regularly use marijuana
4) 12 out of 300 will regularly use opioids

These results can draw a few different conclusions. First, the proportions of the population abusing alcohol, nicotine, and marijuana were very similar. Second, the proportion of the population abusing opioids was much smaller, showing a potentially significant difference in the popularity of these substances.

3.6) Strengths and Limitations

One of the strengths of this model is that it's relatively robust: we got these statistics from general sample data across the web for each demographic. Since we are simulating populations with random individuals and averaging the probabilities from different demographics (ethnicities, genders, etc.), this provides a somewhat accurate model for predicting an individual's likelihood of abusing a specific substance in the diverse and current population. Additionally, the simulation run in this question accounts for the fact that one student could be abusing two or more different substances at the same time, making it a more complex model.

One of the major weaknesses of this model is that there was very limited data available for opioid usage statistics by demographic, and therefore we had to consider that the opioid usage across our simulated population (and all demographics) was equal to the national average, which is 4.7%[23]. Having this sort of limitation, we were not able to tailor the data to directly represent the different demographics of the population. But otherwise, using the national average of opioid use instead of different demographic data points would still be able to represent the population itself.

4. Question 3: Ripples

4.1) Restatement of the Problem

After analyzing the probability that any individual could become a user of a specific substance, we are now asked to find the financial and non-financial "ripples" that these drug users cause.

4.2) Assumptions

1. Assuming the simulation is taken at the end of high school, all high school seniors are 18 years of age.
2. All people sampled are not active drug users.
3. The law does not affect drug usage by a significant amount.
4. All the people in the sample is exposed to the same type and amount of advertisements.
5. The economic costs and tolls of each drug/substance will stay the same each year.

4.3) Developing the Models

To address this issue, we first used the solutions produced in Part 2 and interpreted the value ζ as an addiction factor–essentially, the probability that a person who regularly uses a certain substance will continue to use that substance in the future. Then, in order to clarify the model of substance use over a lifetime, we created a Markov chain that displays the transitions between the two states of being a user of drugs and not a user of drugs, as shown in the visualization to the right:

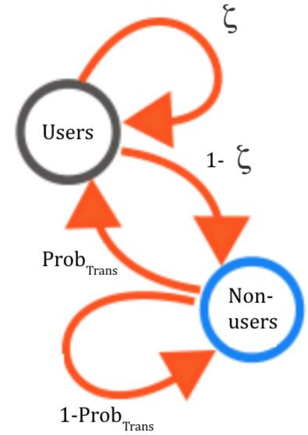


Figure 5

We applied this Markov chain later on when trying to account for the fluctuations in the number of users and non-users of these substances. Then, we took into account three different costs of these four drugs/substances, created a severity index for each one, and then combined these to find an overall cost for each drug/substance.

A) FINANCIAL

The primary financial cost we considered was the accumulative economic cost of each drug/substance per year.

Economic Costs:

Type of Substance	Most Recent Economic Cost per Year (millions of $)	% of Total of Economic Costs Caused by the Four Substances
Alcohol	250,000	45.5%
Marijuana	-58	-0.00011%

Nicotine	300,000	54.5%
Opioids	78.5	0.00014%

Table 5

These percentages were used in the final model later incorporated with the Markov chain.

B) NON-FINANCIAL

Health:

To factor in the impact on health caused by each substance, the average number of deaths in the US due to these types of substances was calculated between 2010 and 2014.

Type of Substance	**Average Number of Deaths in US from 2010-2014**	**% of Total of Avg. # of Deaths Caused by the Four Substances**
Alcohol	2100[1]	0.420%
Marijuana	237[14]	0.00047%
Nicotine	480,000	96.1%
Opioids	17,000	3.40%

Table 6

These percentages were used in the final model later incorporated with the Markov chain.

Environmental Impacts:

To determine the environmental effects of each substance, the amount of carbon dioxide produced during the production of 1 g of each of them was calculated.

Type of Substance	**Amount of Carbon Dioxide Produced per Gram of Substance Produced (g)**	**% of Total Carbon Dioxide Produced by the Four Substances**
Alcohol	0.96*	0.021%
Marijuana	4600[3]	98.7%
Nicotine	59.7[8,29]	1.28%

Opioids	2.04[12]	0.044%

*This number was calculated based on the chemical reactions involved in fermentation, and does not factor in additional, and likely far greater, carbon emissions from the mechanisms involved in brewing various alcoholic beverages.

These percentages were used in the final model later incorporated with the Markov chain.

The last data collected to create the final model was the lifetime transitional probabilities[28] of each substance (denoted by Γ). In other words, we found the probabilities that an individual would transition from being a non-user to a user of a substance during their lifetime; this would be used in the final implementation of the Markov chain:

Type of Substance	Γ
Alcohol	0.227
Marijuana	0.089
Nicotine	0.675
Opioids	0.053

C) Markov Chains to Simulate Future Impact:

The data that we examined earlier only refers to current environmental, health and financial impacts. In order to fully assess the ranking of each illegal substance, we have to predict the change in the number of users in the future. Every individual currently exists in some finite state relating to a narcotic: a user (User) or a non-user (Non-User). Therefore, if we know the lifetime transitional probabilities, we can model the associating Transition Matrix T_i for substance *i* with entries q_{AB} that an individual will move from state A to state B. The matrix *T* can be written as follows:

$$\begin{bmatrix} UU & UN \\ NU & NN \end{bmatrix}$$

with entries UU, UN, NU, NN, defined below:

UU = 1- ς	UN = ς
NU = Γ_i	NN = 1- Γ_i

Thus, using this transitional matrix, the probabilities were implemented in the final model, which gave a Destructive Impact Index (DII) in the following way, for each type of substance i with the number of users u, the number of non-users n, % of financial impact f for the substance, % of death impact d for the substance, and % of environmental impact e for the substance:

$$DII = ((UU - UN) * u + (NU) * n) * (f + d + e)/100$$

4.4) Validating the Models

Validating the transitional probabilities was done with Python and NumPy to simulate a markov chain probability distribution. The following transitional matrices were generated by calculating a one step transitional probability average over 10,000 iterations:

$T_{nic} = \begin{pmatrix} .846 & .155 \\ .084 & .909 \end{pmatrix}$ $T_{opioids} = \begin{pmatrix} .841 & .157 \\ .087 & .909 \end{pmatrix}$

$T_{marijuana} = \begin{pmatrix} .8411 & .1636 \\ .087 & .915 \end{pmatrix}$ $T_{alcohol} = \begin{pmatrix} .838 & .159 \\ .089 & .910 \end{pmatrix}$

These transitional matrices are consistent with the ones hand-calculated in our DII results above. Since the number of iterations is so large, the standard deviation of the transitional matrix elements is very small.

The financial, health, and environmental impacts in the tables above can be validated as follows. First, the financial impacts are most likely to remain roughly constant over the next decades because they were calculated in percentages of the total economic impacts, meaning that these proportions are not likely to alter much over time. The health impacts are also most likely to remain constant because these most likely depend on the population of people using the product, which we accounted for in the Markov chains; this is because the health risks depend on the chemical compositions of each substance overall. Lastly, the environmental impacts will remain constant over long periods of time because these depend on the amount of carbon dioxide produced while these substances are created; since this creation process will most likely remain the same, this value will not vary.

4.5) Results

By using the DII model, we determined the following DIIs for the four substances of alcohol, marijuana, nicotine, opioids, as shown below:

$DII(alcohol)$ = 42,138,507.04

$DII(marijuana)$ = 48,967,347.1

$DII(nicotine)$ = 332,503,290

$DII(opioids)$ = 921,720.7468

Ranking (from greatest to lowest dangerous impact):

1. Nicotine
2. Marijuana
3. Alcohol
4. Opioids

Based on our model, nicotine is the most destructive drug by far (both financially and non-financially), followed by marijuana, then alcohol and finally opioids.

4.6) Strengths and Limitations

One strength of this model were that the use of the Markov chain provided a comprehensive view of how much of a factor addictiveness as well as transition probability are for each substance. Because of this, a nearly accurate population of users and non-users of each of the substances could be calculated. Additionally, the calculation of the DII was such that it used values for the financial, health, and environmental impacts of each substance that would most likely not change over time, thus making it a relatively permanent model.

One weakness of this model was that the datasets for the impacts of these substances didn't have very recent data; for example, the average number of deaths caused by each substance was calculated with data between the years 2010 and 2014. To account for this, an improved dataset could have been used for each calculation. Another weakness was the fact that there could be changes in the future that could alter the values of each lifetime transitional probability; however, this could not be accounted for except for by changing the model in the future. Lastly, the transitional probability of opioids used in the study was not a lifetime transitional probability, but a short-term transitional probability due to lack of available data. This could be accounted for by finding the actual statistic for the opioid transitional probability.

5. Conclusions

Through this study, we discovered the various trends and models regarding the use of e-cigarettes, nicotine products, alcohol, marijuana, and opioids. Various significant trends with implications for the future of drug use in the United States and the world were discovered.

In the first question, we created several models in the process of finding the amount of nicotine spread due to e-cigarettes and cigarettes for the next ten years. Some of these were a linear regression to find the amount of money spent on advertisements for e-cigarettes over the next ten years, an exponential model to find the percent of the population who would smoke cigarettes over the next ten years, a normal distribution to find the proportion of nicotine levels in e-cigarette products, and the final model that combined all of these factors to compare the amounts of nicotine consumed due to cigarettes and e-cigarettes in the future. This final model allowed us to find that the amount of nicotine consumed due to e-cigarettes would exceed that of cigarettes in the future, by $2.89 * 10^9$ mg to $2.21 * 10^9$ mg. The strengths of this final model was that it provided clear evidence for correlation between the amount of money spent on e-cigarette advertisements and the percent of the population using e-cigarette advertisements, while some weaknesses were the shortage of available data.

In the second question, we simulated the use of alcohol, marijuana, nicotine, and opioid products among 300 high school students. Among the factors we considered in our simulated population were gender, race, sexual orientation, and depression/anxiety level. We used national averages to create a weighted random student generator, which we used to populate our class of 300. We used the proportion of high school students belonging to each factor that regularly use each of the substances in order to create a probability model that is able to consider how each factor affects the chance that an individual is addicted to alcohol, marijuana, nicotine, and/or opioids. We found that out of the 300 students we generated, 49 out of 300 will regularly use alcohol, 53 out of 300 will regularly use nicotine, 48 out of 300 will regularly use marijuana, and 12 out of 300 will regularly use opioids.

In the third question, we calculated the Destructive Impact Index to find a way to rank alcohol, marijuana, nicotine, opioids in terms of how much of a detrimental impact each has on society. The types of impacts considered were financial, health, and environmental risks. Each of these were calculated by using previous datasets showing how much money was spent in the US due to these substances, how many deaths were caused by these substances on average, and how much carbon dioxide was produced as a result of the creation of each of these substances. To account for the fact that the number of people using each substance will change over time, we created a Markov chain model to multiply by each of the substances' risks. This value was the Destructive Impact Index, which allowed us to rank the four substances, in order of highest to lowest danger to society, as nicotine, marijuana, alcohol, and opioids.

This study allowed us to create various mathematical statistical models that showed how certain drugs, nicotine products, and alcohol substances would affect larger populations over time. Future research that could be done to extend the studies covered in this paper are how other factors could affect their drug usage. These factors could include the personality of the person, the childhood of the person, the social circles of that person, and other factors of their everyday lives. We hope that this study has had a significant impact on tackling the many issues related to drug and alcohol abuse in the US and the world.

6. References

1. “Alcohol Facts and Statistics.” *National Institute on Alcohol Abuse and Alcoholism*, U.S. Department of Health and Human Services, Aug. 2018, www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/alcohol-facts-and-statistics.
2. Arteaga, Irma, et al. “Childhood Predictors of Adult Substance Abuse.” *Children and Youth Services Review*, vol. 32, no. 8, 2 May 2010, pp. 1108–1120., doi:10.1016/j.childyouth.2010.04.025 .
3. Burns-Edel, T. (2016). Environmental Impacts of Illicit Drug Production. *Global Societies Journal*, 4. Retrieved from https://escholarship.org/uc/item/4w64g29s
4. Cantrell, Jennifer, et al. “Rapid Increase in e-Cigarette Advertising Spending as Altria's MarkTen Enters the Marketplace.” *Tobacco Control*, BMJ Publishing Group Ltd, 1 Apr. 2016, tobaccocontrol.bmj.com/content/25/e1/e16.
5. “CDC - Historical NYTS Data and Documentation - Smoking & Tobacco Use.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 9 May 2018, www.cdc.gov/tobacco/data_statistics/surveys/nyts/data/index.html.
6. “CDC Features.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 13 July 2018, www.cdc.gov/features/costsofdrinking/index.html.
7. Center for Tobacco Products. “Youth & Tobacco - Youth Tobacco Use: Results from the National Youth Tobacco Survey.” *U.S. Food and Drug Administration*, Center for Drug Evaluation and Research, 11 Feb. 2019, www.fda.gov/TobaccoProducts/PublicHealthEducation/ProtectingKidsfromTobacco/ucm405173.htm.
8. “Cost of Cigarettes 'Must Rise to Reflect Environmental Damage from Tobacco Industry'.” *The Independent*, Independent Digital News and Media, 2 Oct. 2018, www.independent.co.uk/environment/cost-cigarettes-rise-tobacco-industry-smoking-environment-who-climate-change-pollution-a8564796.html.
9. Dews, Fred. “Chart of the Week: Marijuana Use by Race.” *Brookings.edu*, The Brookings Institution, 11 Aug. 2017, www.brookings.edu/blog/brookings-now/2017/08/11/charts-of-the-week-marijuana-use-by-race/.
10. “Drugs, Brains, and Behavior: The Science of Addiction.” *PsycEXTRA Dataset*, 2007, doi:10.1037/e596722007-001.
11. “E-Cigarette Ads and Youth | VitalSigns | CDC.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 23 Mar. 2017, www.cdc.gov/vitalsigns/ecigarette-ads/index.html.

12. "Economic Toll of Opioid Crisis in U.S. Exceeded $1 Trillion Since 2001." *Altarum*, 6 Dec. 2018, altarum.org/news/economic-toll-opioid-crisis-us-exceeded-1-trillion-2001.
13. "Ethnicity and Health Disparities in Alcohol Research." *National Institute on Alcohol Abuse and Alcoholism*, U.S. Department of Health and Human Services, pubs.niaaa.nih.gov/publications/arh40/152-160.htm.
14. "Fatal Marijuana Overdose Is NOT a Myth." *The Psychological Basis of Addiction*, 15 Sept. 2016, www.rehabs.com/pro-talk-articles/fatal-marijuana-overdose-is-not-a-myth/.
15. Feldman, L, et al. "Alcohol Use Beliefs and Behaviors among High School Students." *Current Neurology and Neuroscience Reports.*, U.S. National Library of Medicine, Jan. 1999, www.ncbi.nlm.nih.gov/pubmed/9890365.
16. Hanson, Carl Lee, et al. "An Exploration of Social Circles and Prescription Drug Abuse Through Twitter." *Journal of Medical Internet Research*, vol. 15, no. 9, 2013, doi:10.2196/jmir.2741.
17. "How Illicit Drug Use Affects Business and the Economy." *National Archives and Records Administration*, National Archives and Records Administration, 2008, https://obamawhitehouse.archives.gov/ondcp/ondcp-fact-sheets/how-illicit-drug-use-affects-business-and-the-economy
18. "How Much CO2 Is Produced from Brewing?" *Life, Fermented*, 16 Dec. 2014, lifefermented.wordpress.com/2014/01/22/how-much-co2-is-produced-from-brewing/.
19. Jamal A, Gentzke A, Hu SS, et al. Tobacco Use Among Middle and High School Students —United States, 2011–2016. MMWR Morb Mortal Wkly Rep 2017;66:597–603. DOI: http://dx.doi.org/10.15585/mmwr.mm6623a1.
20. Jones, Lora. "Vaping - the Rise in Five Charts." *BBC News*, BBC, 31 May 2018, www.bbc.com/news/business-44295336.
21. Kuwana, Ellen. "Cigarette Ads - A Broken Promise." Neuroscience For Kids - Brain vs. Computer, 26 Sept. 2001, https://faculty.washington.edu/chudler/smokead.html.
22. LaVito, Angelica. "As Vaping Surges, Teen Cigarette Smoking Ticks up after Decades of Decline." *CNBC*, CNBC, 23 Oct. 2018, www.cnbc.com/2018/10/22/teen-cigarette-smoking-ticks-up-as-vaping-surges.html.
23. McCabe, Sean Esteban et al. "Medical and nonmedical use of prescription opioids among high school seniors in the United States" *Archives of pediatrics & adolescent medicine* vol. 166,9 (2012): 797-802.
24. National Institute on Drug Abuse. "Overdose Death Rates." *NIDA*, 29 Jan. 2019, www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates.
25. National: National Youth Tobacco Survey, 2011-2014; Kim, et al (2014), Truth Initiative (2015)

26. "Preventing Chronic Disease: April 2009." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, Apr. 2009, www.cdc.gov/pcd/issues/2009/apr/08_0039.htm.
27. "State Reports From the 2016 NSDUH." *Stages of Community Readiness | SAMHSA*, Substance Abuse and Mental Health Services Administration, www.samhsa.gov/data/nsduh/state-reports-NSDUH-2016
28. Szigeti, Balázs, and Péter Sárosi. "Home." *Drugreporter*, 12 Aug. 2016, drogriporter.hu/en/dose-of-science-the-probability-and-predictors-of-substance-abuse/.
29. "The Environmental and Health Impacts of Tobacco Agriculture, Cigarette Manufacture and Consumption." *World Health Organization*, World Health Organization, 30 Nov. 2015, www.who.int/bulletin/volumes/93/12/15-152744/en/.
30. "Tobacco Use in Racial and Ethnic Populations." *American Lung Association*, Medical Editorial Review Panel, 14 Feb. 2019, www.lung.org/stop-smoking/smoking-facts/tobacco-use-racial-and-ethnic.html.
31. "Today's Heroin Epidemic Infographics | VitalSigns | CDC." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, July 2015, www.cdc.gov/vitalsigns/heroin/infographic.html#infographic.
32. Truth Initiative. "Tobacco Use in LGBT Communities." *Truth Initiative*, Truth Initiative, 13 Feb. 2018, truthinitiative.org/news/tobacco-social-justice-issue-smoking-and-lgbt-communities.
32. Whitesell, Mackenzie, et al. "Familial, Social, and Individual Factors Contributing to Risk for Adolescent Substance Use." *Journal of Addiction*, vol. 2013, 20 Mar. 2013, pp. 1–9., doi:10.1155/2013/579310.
33. Wills, T. A., et al. "Risk Factors for Exclusive E-Cigarette Use and Dual E-Cigarette Use and Tobacco Use in Adolescents." *Pediatrics*, vol. 135, no. 1, 2014, doi:10.1542/peds.2014-0760.
34. "World Age Structure." *Zambia GDP - per Capita (PPP) - Economy*, www.indexmundi.com/world/age_structure.html.
35. "World Population by Age and Region 2018 | Statistics." *Statista*, Statista, 2018, www.statista.com/statistics/265759/world-population-by-age-and-region/.
36. "Youth and Tobacco Use | CDC." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 5 Feb. 2019, www.cdc.gov/tobacco/data_statistics/fact_sheets/youth_data/tobacco_use/index.htm.
37. "Youth Risk Behavior Surveillance – United States, 2017." *Morbidity and Mortality Weekly*

Report, vol. 67, no. 8, 15 June 2018, doi:10.13072/midss.361.

38. Zhang, Mona. "Legal Marijuana Is A Boon To The Economy, Finds Study." *Forbes*, Forbes Magazine, 13 Mar. 2018, www.forbes.com/sites/monazhang/2018/03/13/legal-marijuana-is-a-boon-to-the-economy-finds-study/#791bc6dcee9d.

7. Appendix

```
alc_dict = {
    'black':.22,
    'white':.26,
    'hispanic':.3,
    'asian':.18,
    'homosexual':.25,
    'straight':.13,
    'male':.167,
    'female':.136,
    'unhappy':.165
}

opium = .047

marijuana_dict = {
    'black':.14,
    'white':.11,
    'hispanic':.166,
    'asian':.21,
    'homosexual':.4,
    'straight':.166,
    'male':.17,
    'female':.15,
    'unhappy':.18
}

nic_dict = {
    'black':.168,
    'white':.166,
    'hispanic':.101,
    'asian':.07,
    'homosexual':.206,
    'straight':.149,
    'male':.167,
    'female':.136,
    'unhappy':.28
}
```

This is a screenshot of the python dictionaries containing the percentages of high schoolers that regularly use the various drugs mentioned.

```python
def create_person():
    arr = []
    # https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755735/
    gender = 'male' if rd.uniform(0,1) >= .5 else 'female'
    sexuality = 'straight' if rd.uniform(0,1) <= .66 else 'other' #
    age = 18 # assumption for model
    rand_socio = rd.uniform(0,1) # https://www.childtrends.org/ind
    if rand_socio <= .13:
        socio_bg = 'unsafe'
    elif rand_socio >= .43:
        socio_bg = 'somewhat safe'
    else:
        socio_bg = 'safe'
    happy = 'happy' if rd.uniform(0,1) >= .2 else 'unhappy' # http
    raceStat = rd.uniform(0,1)
    if raceStat < .73:
        race = 'white' # white
    elif raceStat >.73 and raceStat > .856:
        race = 'black' # black
    elif raceStat > .856 and raceStat < .906:
        race = 'asian' # asian
    else:
        race = 'other' # other
    arr = [gender,sexuality,age,socio_bg,happy,race]
    return arr


def get_chance(dic):
    if dic == .047:
        return .047
    count = 0
    sum = 0
    for i in create_person():
        if i in dic:
            sum = sum + dic[i]
            count = count + 1
    sum = sum/count
    return sum
```

This screenshot reveals the python code that was used to create a high school senior using weighted random probabilities for demographic factors. It also shows the function which returns the probability that a person will be addicted to alcohol based on their demographic factors.

```python
import random

totalnic = 0;
totalalc = 0;
totalmj = 0;
for i in range(5):

    arr_nic = []
    arr_alc = []
    arr_mj = []

    nicusers=0
    alcusers=0
    mjusers=0
    count = 0
    for i in range(300):
        count += 1
        print("Student " + str(count) + ":")
        nic_chance = get_chance(nic_dict)
        arr_nic.append(nic_chance)
        alc_chance = get_chance(alc_dict)
        arr_alc.append(alc_chance)
        mj_chance = get_chance(marijuana_dict)
        arr_mj.append(mj_chance)
        if(random.uniform(0,99)>nic_chance*100):
            print("not abusing nicotine")
        else:
            print("abusing nicotine")
            nicusers=nicusers+1
        if(random.uniform(0,99)>alc_chance*100):
            print("not abusing alcohol")
        else:
            print("abusing alcohol")
            alcusers=alcusers+1
        if(random.uniform(0,99)>mj_chance*100):
            print("not abusing marijuana")
        else:
            print("abusing to marijuana")
            mjusers=mjusers+1
        print("")

    print("Number of alcohol abusers: "+ str(alcusers))
    print("Number of nicotine abusers: "+str(nicusers))
    print("Number of marijuana abusers: "+str(mjusers))
    totalnic = totalnic + nicusers
    totalalc = totalalc + alcusers
    totalmj = totalmj + mjusers
```

This screenshot shows the python code which simulates the creation of our class by looping the creation of 300 people using the random weighted probability model. We looped this process 5 times in order to create 5 classes (to ensure that it was working properly).