Fanning the Fumes: The Forecast of Substance Abuse in the USA

Team 11729

March 2019

0   Executive Summary

In the United States today, drug legalization and regulation stands as a major political wedge issue. For local, state, and national governments to most effectively regulate and/or restrict substance use and abuse, a thorough understanding of the future of substance industries and the impact of substance use is critical on individuals, societies and schools.

To begin our exploration, we modeled the spread of nicotine due to vaping over the course of the upcoming decade. We defined spread by the expansion of new nicotine users rather than the net change in nicotine usage because the new generation of smokers will determine the long term trend of American nicotine consumption. Due to the fact that the majority of vaping growth can be attributed to youth demand and that nearly 9 out of 10 adult smokers began in their teenage years[5], we approximate the growth rate in nicotine use to be proportional to the high school usage at a given time. Using a logistic model for high school e-cigarette usage, we find that nicotine usage growth rate from vaping will gradually increase before leveling off in the mid 2020s at a factor of about 1.23 times the current rate in 2028. We find that the rate of nicotine usage spread from traditional cigarettes will fall linearly before leveling off at 3%. Overall, nicotine usage growth rate will increase approximately logistically, leveling off at a rate approximately 1.24 times the current rate prior to 2028. Additionally, the proportion of nicotine spread attributed to vaping has increased drastically over the past decade and will continue its ascent before leveling off, accounting for approximately 91.5% of all nicotine spread by the mid 2020s, as the impact of vaping continues to overtake that of traditional cigarette usage.

We then shifted our focus to determining which factors/demographics had the greatest influence on the likelihood of an individual using marijuana, alcohol, unprescribed opiates, or nicotine. We identified the four most influential categories as: age (<18. 18-25, 26+), race (White, African-American, Hispanic/Latino, Asian), gender (Female, Male), and poverty level(<100%, 100-199%, 200+%). These were used to create a weighted scale, with each subcategory having a multiplier anchored in sample proportions for substance use in the entire USA. The product of the multipliers from each demographic category is used to calculate how much more or less inclined an individual is to use a given substance relative to the population parameter, as well as a concrete percent likelihood of substance usage. We then tested our model on a sample of 300 US high school seniors, predicting that 66 seniors will use marijuana, 169 seniors will consume alcohol, 2 seniors will abuse opiates, and 87 seniors will consume nicotine in some form.

Finally, with our new-found expertise relating demographics and substance use, we utilized a social cost/benefit analysis to quantify the societal impact of the use of each of nicotine, marijuana, alcohol, and unprescribed opiates in the US. This monetary value, labeled Total Impact(TI) consists of monetary amalgamations of both financial and non-financial factors represented by 3 factors delineated as Total Market Value (M), Humanitarian Costs (H) and External Costs (E). Overall, we found that the net Total Impact (TI) for each of the four drugs were: -$7.67 \cdot 10^{12}$, -$1.14 \cdot 10^{12}$, -$5.78 \cdot 10^{11}$, and -$1.33 \cdot 10^{12}$ respectively for each of Nicotine, Marijuana, Alcohol and Unprescribed Opiates. This results in a final ranking in decreasing order of impact magnitude of Nicotine, Unprescribed Opiates, Marijuana, and Alcohol, indicating that Nicotine is the greatest detriment to society.

Contents

1   Part 1: Darth Vapor

1.1 Restatement of Problem

We are tasked with the following:

- Develop a mathematical model for the spread of nicotine use due to vaping over the next ten years.

- Analyze how nicotine use growth from e-cigarettes such as JUUL compares with that of traditional cigarettes.

1.1.2 Refinement of Problem Statement

The prompt directs us to predict the comparative growth of nicotine use from vaping and traditional cigarettes. According to data from the Department of Health and Human Services, nearly 90 percent of adult smokers began before age 18 [5]. Thus the vast majority of new nicotine users can be attributed to high school aged students. Additionally, over 73% of e-cigarettes sales in the US in 2018 can be attributed to youth consumers[7]. Therefore we assume magnitude of new nicotine users is directly proportional to the number of high school students using nicotine products. Note that we have chosen to calculate growth based on the number of new nicotine users, rather than the net change because the new generation of users will determine the long term trends of nicotine use. Also in our model we calculate the relative magnitude of nicotine use growth for each type (traditional and e-cigarettes). This allows for comparisons between growth across years and between the two types.

1.2 Assumptions and Definitions

- Definition: We define current drug use as use of a given drug within the past month (30 days). This definition is used throughout all sections of our paper.

  Justification: 30 days is the standard for current drug use for the National Youth Tobacco Survey, CDC, and other government agencies. Furthermore survey data for use within the past 30 days is readily available.

- Global Assumption: The terms e-cigarette use and vaping can be used interchangeably.

  Justification: Vaping refers to the use of e-cigarettes.

- Assumption: High school e-cigarette use will follow a relatively logistic model of growth.

  Justification: The modern marketing of vaping devices to teens through products such as JUUL which have a USB-like design intended is fairly new. This has lead to explosive growth in popularity which is roughly exponential, consistent with the initial growth phase of a logistic model. However, once the product becomes normalized in the market, vaping will level off; it cannot continue to grow indefinitely. The leveling or stabilization point is analogous to the carrying capacity of the drug. Thus, a logistic model fits the growth of vaping.

- Assumption: Traditional cigarette use among high school students will not drop below 3%.

  Justification: Though teenage cigarette use has had a consistent downward trend for multiple decades, it is bounded because the percent cannot drop below 0. Even as e-cigarettes continue to grow in

popularity and traditional cigarette use continues to dwindle, it is reasonable that the use will stabilize at some low percentage because some students will continue to smoke cigarettes. We approximate the stabilization point to be around 3%.

- Assumption: The magnitude of cigarette and vaping usage of high school is proportional to the magnitude of the growth of new nicotine users.

  Justification: See 1.1.2

1.3 Variables

Variables (Fig 1.1)		
Symbol	Definition	Units
t	Years after 2011 (t=0 in 2011)	years
P(t)	The proportion of high school students who use e-cigarettes as a function of t	%
p(t)	The estimated proportion of high school students who use e-cigarettes as a function of t based on our model	%
K	Long term percent of high school students using e-cigarettes (analogous to carrying capacity in logistic model)	%
r	Constant in logistic model	
T	The estimated proportion of high school students who use traditional cigarettes based on our model	%

1.4 Data for the Logistic Model

Using results from the National Youth Tobacco Survey, we compiled the following data for the percentage of high school students using e-cigarettes in each year from 2011 to 2018. Consistent with our global assumption, the statistics are based on the percentage who reported using e-cigarettes in the prior 30 days.

Proportion of high school age students using e-cigarettes by year [1][2][3] (Fig 1.2)								
Year	2011	2012	2013	2014	2015	2016	2017	2018
% usage by high schoolers	1.5	2.8	4.5	13.4	16.0	11.3	11.7	20.8

1.5 Developing the Logistic Model for E-Cigarettes

Based on our assumption regarding the spread of e-cigarette use, our model will predict e-cigarette use using a logistic model. As defined in the above, t is the number of years after 2011 and p(t) represents the projected proportion of e-cigarette use among high school students in the year represented by t. Furthermore, K represents the proportion that our logistic model will asymptotically approach. Assuming logistic growth with these variables we have:

$$\frac{dp}{dt} = r \cdot (p) \cdot (1 - p/K)$$

Below we show the separation of variables and integration used to find an explicit function for e:

$$\frac{dp}{p(1 - \frac{p}{K})} = r \cdot dt$$

$$\frac{K \cdot dp}{p(K - p)} = r \cdot dt$$

$$\int_{p(0)}^{p(t)} \left(1/p + 1/(K-p)\right)dp = \int_0^t r \cdot dt$$

$$\left(\ln(p) - \ln(K-p)\right)\Big|_{p(0)}^{p(t)} = rt\Big|_0^t$$

$$\ln\left(\frac{K-p}{p}\right)\Big|_{p(0)}^{p(t)} = -rt\Big|_0^t$$

$$\ln \frac{\frac{K-p(t)}{p(t)}}{\frac{K-p(0)}{p(0)}} = -rt$$

$$\frac{K-p(t)}{p(t)} = \frac{K-p(0)}{p(0)e^{-rt}}$$

Analytically we find:

$$p(t) = \frac{K}{1 - \frac{K-p(0)}{p(0)} \cdot e^{-rt}}$$

We notice that there are 3 unknowns in this model (K, r, and p(0)). We do not have the computational power to run three variable logistic regression. Instead we will simplify the regression by letting $p(0) = P(0) = .015$ from the table. Now are equation becomes:

$$p(t) = \frac{K}{1 - \frac{K-.015}{.015} \cdot e^{-rt}}$$

For simplicity of calculations, we apply the substitution $e^r = d$. We proceed to do least squares regression analysis. To find the sum of the squares of the residuals, we compute:

$$RSS = \sum_{t=0}^{7}(P(t) - p(t))^2$$

In order to minimize RSS and thus create the optimal regression equation we must find K and d such that:

$$\frac{\partial RSS}{\partial K} = 0 \text{ and } \frac{\partial RSS}{\partial d} = 0$$

Due to lack of computing power, we instead calculated the optimal d value for various K values in the range of 25-40% which we consider to be reasonable based on our assumption about the long run behavior. We find d to be approximately 1.75 for each of the K values, so for simplicity we use $d = 1.75$. Thus our equation for p(t) becomes:

$$p(t) = \frac{K}{1 - \frac{K-.015}{.015} \cdot 1.75^{-t}}$$

The graph shown on the following page displays the logistic models for long run user proportions from 25% to 37.5%. Using least squares regression analysis, we find the best fit to be when $K = 32.5\%$ so we will use that logistic curve for our model. Thus:

$$p(t) = \frac{.325}{1 - \frac{.325-.015}{.015} \cdot 1.75^{-t}}$$

From the graph, we can see that p(t) and consequently nicotine spread rate from vaping will continue to grow in the next ten years but at a decreasing rate, and less rapidly than the growth that has been observed in the previous decade. $\frac{p(17)}{p(8)} = 1.23$ therefore the rate of nicotine usage growth will increase by a factor of approximately 1.23 in the next decade.

Percent of High School Students Vaping Over Time
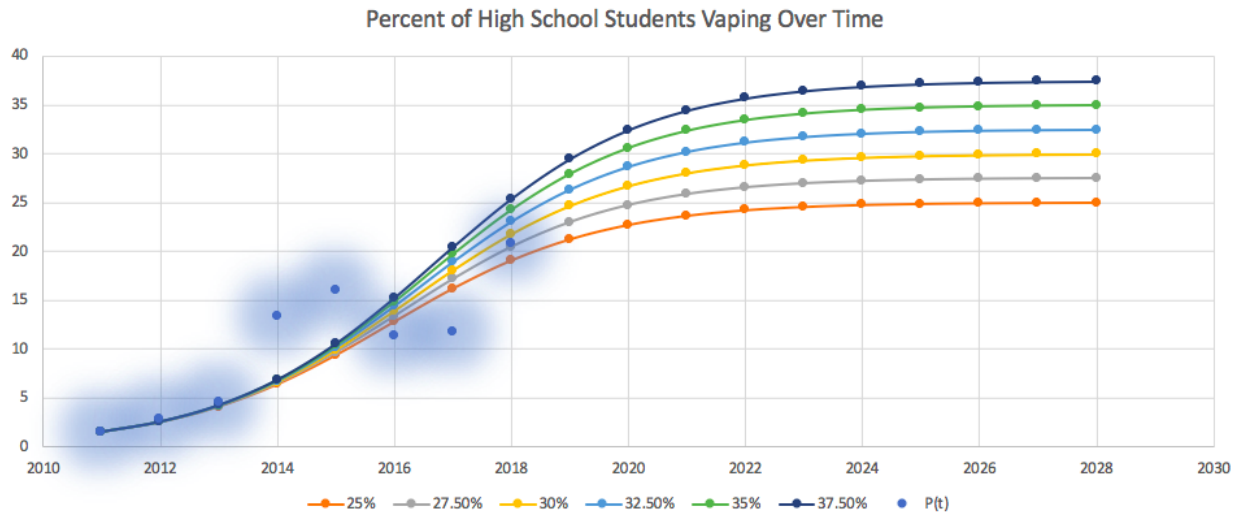
1.6 Modeling Traditional Cigarette Use

We must compare growth of vaping to traditional cigarettes. We use traditional cigarette data from the National Youth Tobacco Survey, the same source as was used for vaping, to maintain consistency[5]. The data is compiled in the table below.

Proportion of high school age students using traditional cigarettes by year [5] (Fig 1.3)						
Year	2000	2002	2004	2006	2009	2011
% usage by high schoolers	28.0	22.5	21.7	19.8	17.2	15.8
Year	2012	2013	2014	2015	2016	2017
% usage by high schoolers	14.0	12.7	9.2	9.3	8.0	7.6

Shown graphically:

% of High School Students Using Cigarettes by Year

$y = -1.1339x + 2294.6$
$R^2 = 0.9722$
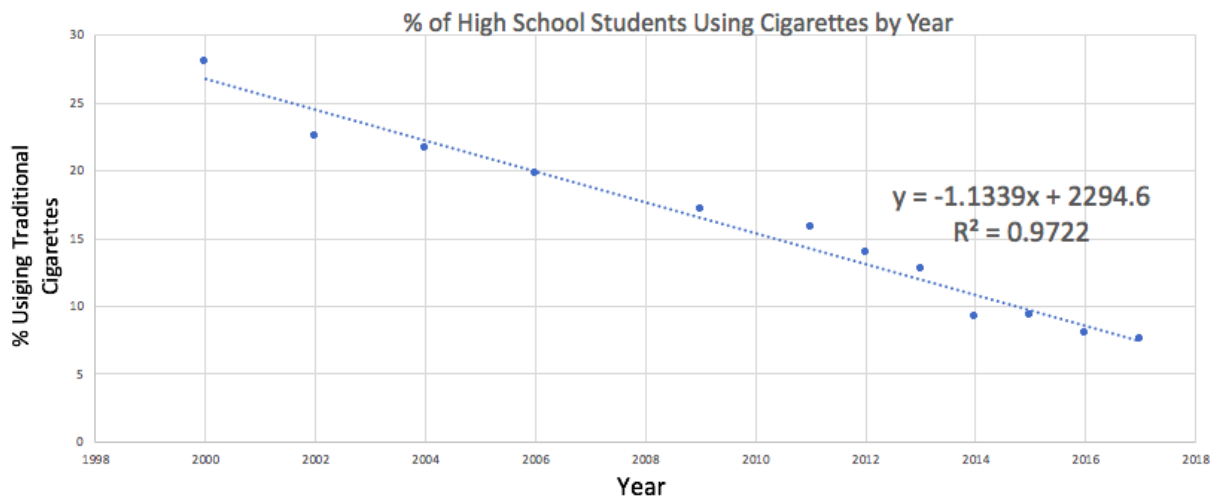
Noting the linear appearance of the scatter plot, we apply least squares linear regression. We find the r value to be -.986 which suggests a strong negative linear relationship. Also $R^2 = 97.22\%$ so 97.22% of the variability in the percentage of high school students using traditional cigarettes can be explained by the linear model relating the proportion of use to the time in years. We therefore feel confident in the fit of the

linear model. As defined in the table of variables, we let T represent the percentage of high school students using traditional cigarettes. Using our regression, we find:

$$\hat{T} = -1.1339(t + 2011) + 2294.6 = -1.1339t + 14.3$$

However, if we project this linear pattern into the future, the negative slope will lead to a negative value of $\hat{T}$ after a certain point. We refer to our assumption that traditional cigarette use will not drop below 3 percent and apply a lower bound of 3 to $\hat{T}$.

1.7 Putting it all together

The prompt directs us to compare the addition of new e-cigarette users to cigarette users. We decide to calculate $p(t) + \hat{T}$ for each of the next ten years.

These values should represent the relative magnitude of new nicotine users. We show the trend in the graph on the following page which displays that the relative magnitude of new users will increase due to the rise in vaping popularity but level off around 2024 once the market stabilizes. This logistic looking graph is logical because vaping use is growing at an increasing rate currently, but eventually should level off when the market readjusts. We calculate $\frac{p(17) + T(\hat{17})}{p(8) + \hat{T}(8}) = 1.12$ to be the ratio between the projected growth rate of nicotine usage in 2028 to the current growth rate.

Relative Magnitude of New Nicotine Users By Year
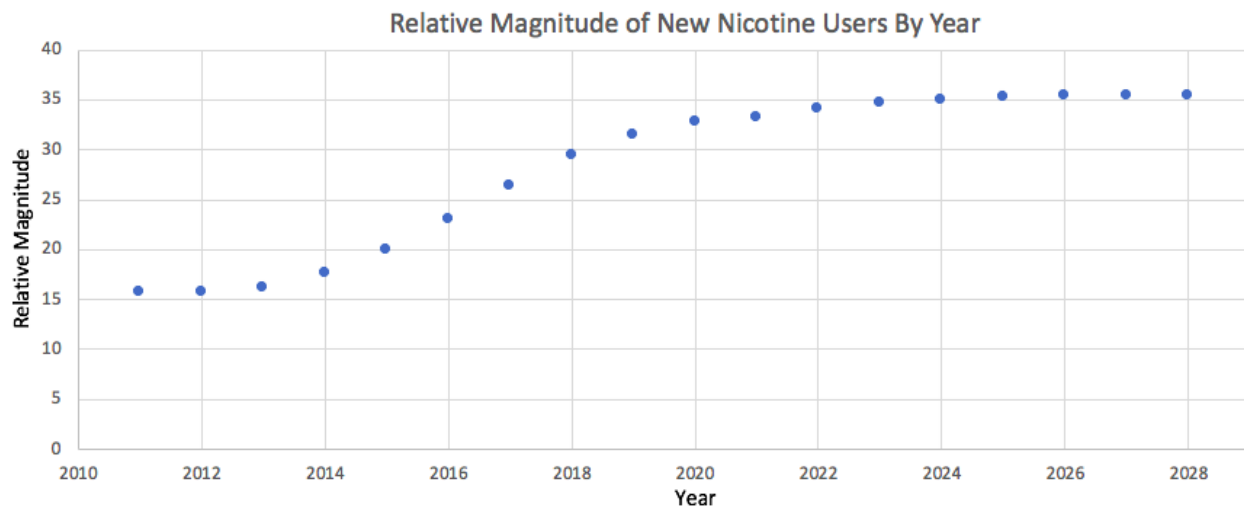
We also compute $\frac{p(t)}{p(t) + \hat{T}}$ which represents the proportion of new nicotine users that come from vaping use. The values are shown in the below table:

Proportion of new nicotine use from e-cigarettes (Fig 1.4)						
Year	2011	2012	2013	2014	2015	2016
% new use from vaping	9.49	16.16	25.85	38.05	50.96	62.50
Year	2017	2018	2019	2020	2021	2022
% new use from vaping	71.60	78.35	83.42	87.50	90.96	91.21
Year	2023	2024	2025	2026	2027	2028
% new use from vaping	91.36	91.44	91.49	91.51	91.53	91.54

As is apparent from the table, the proportion of new nicotine users from vaping increased from 9.49% in 2011 to 50.96% in 2015 to 83.42% in 2019 and will continue to increase until leveling out around 91.5% in

the mid 2020s. Thus with the vast majority of new smokers choosing vaporization options into the 2020s, vaping will take up an increasing proportion of the nicotine market as older cigarette smokers die and new vapors replace them.

1.8 Commentary and Assessment of Model

1.8.1 Compact Restatement of Model

The estimated percentage of high school students using e-cigarettes t years after 2011 is:

$$p(t) = \frac{.325}{1 - \frac{.325 - .015}{.015} \cdot 1.75^{-t}}$$

The estimated percentage of high school students using traditional cigarettes is:

$$\hat{T} = \max\{-1.1339t + 14.3, 3\}$$

$p(t)$ is proportional to the growth of new e-cigarette users and $\hat{T}$ is proportional to the growth of new traditional users. $p(t) + \hat{T}$ can be used to model the relative magnitude of total new nicotine use. $\frac{p(t)}{p(t) + \hat{T}}$ represents the proportion of new nicotine use that can be attributed to vaping.

1.8.2 Strengths

The logistic model used to model the growth of e-cigarette use among high school students makes sense in the context of the current vaping situation. Vaping had been growing exponentially due to the rapid increase in youth directed marketing and popularity but the growth rate cannot be sustained (otherwise over 100% of people would vape, a mathematically impossible situation). Instead it must level off at its "carrying capacity". Based on our logistic model, high school e-cigarette use will level off at 32.5%. This seems reasonable because high school cigarette use leveled off around 33% during the 1970s before falling due to greater awareness of health impacts. The linear model for traditional cigarette use is very strong with an R^2 value over 97%. Finally the conclusion that E-cigarettes will make up around 91.5% of new nicotine use by the mid 2020s is consistent with the trend of vaping rapidly taking over the nicotine market share.

1.8.3 Limitations and Areas for Future Research

The model calculates the relative magnitude of new nicotine use, but not the absolute magnitude. In essence this allows for year to year comparison of the growth of nicotine use allowing us to model how the growth rate will change over time. It also allows for comparison between the nicotine spread due to e-cigarette use in comparison to traditional cigarette use. A stronger model would be able to quantify the increase by a metric such as the change in the proportion of people using nicotine directly rather than just making unit-less comparisons from year to year and e-cigarette to traditional. Additionally our model will become increasingly unreliable the further we extrapolate into the future do to potential policy changing surrounding vaping options. With the surge of illegal teenage use, there are already efforts being made to strengthen regulations on companies such as JUUL which would hinder growth.

2   Part 2: Above or Under the Influence?

2.1 Restatement of Problem

We are tasked with the following:

- Create a model to predict the probability that an individual uses one of marijuana, alcohol, non-prescribed opiates, and nicotine based on various demographic and social factors.

- Apply our model to a sample of 300 high schoolers with varying characteristics to determine use of aforementioned drugs.

2.2 Assumptions

- Assumption: The influence of individual factors is independent.

  Justification: If males are more likely than females to use drug Z in general, it is reasonable to assume that white males are more likely users than white females, African-American males are more likely users than African American females, etc. with the same approximate ratio of use. That same logic can be applied to any demographic or social factor to justify the above assumption.

- Assumption: Influence of the characteristics of the drug itself is embedded in the usage statistics of Fig 2.3.

  Justification: Fig 2.3 establishes the general population parameters for usage of the studied drugs. Factors such as accessibility are inherently included in how many people are using the substance. For example, if drug Y is more accessible than drug Z, the disparity in accessibility will be fully accounted for by the disparity in usage.

- Assumption: High school seniors fall in the 18-25 year old bracket.

  Justification: A typical American high school senior turns 18 during their senior year, so they can be categorized in the 18-25 year old bracket.

- Assumption: The sample of 300 US high school seniors will be representative of the racial, gender, and poverty-level demographics of the United States stated in Figure 2.1.

  Justification: Since the prompt asks us to consider seniors with varying characteristics, we can assume that the sample will be diverse and representative of the US population.

2.3 Table of Percentages of Demographics that Use Drugs

Based on the detailed reporting of substance use by age, race, gender, and poverty-level in SAMHSA's annual national drug use reports, we determined that those four factors were the most influential in determining substance use [30]. Additionally, these four factors take into account both internal factors such as biological differences in age groups and genders, as well as external factors such as poverty level. The following data table summarizes substance use percentages for age, gender, race, and poverty level demographics used in subsequent calculations.

Compiled Data of Drug Use Statistics (Fig 2.1) [8][9][10][12][13][16][17][18]						
Demographic	Subgroups	% Total Population	% Use of Marijuana	% Use of Alcohol	% Use of Opiates	% Use of Nicotine
Age	0-17	24	6.5	9.9	0.9	4.9
	18-25	9	22.1	56.3	2.0	29.1
	26+	67	7.9	55.8	1.2	23.4
Race	White	60.7	9.8	56.8	1.4	24.6
	African-American	13.4	11.6	43.0	1.1	23.7
	Asian	5.8	3.7	38.4	0.5	9.1
	Hispanic/Latino	18.1	8.1	44.7	1.3	16.7
Gender	Male	49.2	11.9	55.5	1.5	28.6
	Female	50.8	7.3	48.1	1.1	16.6
Poverty-Level	< 100%	11	13.0	35.8	1.8	31.5
	100 − 199%	17	10.8	39.9	1.6	27.0
	> 200%	72	8.4	58.8	1.1	19.1

[4][30] The poverty levels represent a persons income as a percentage of the poverty line. A person with a 100% poverty level is at the poverty line. Below 100% signals poverty, and above 100% signals excess money (not in poverty).

2.4 Developing the Model

Using the table above, we created a weighted metric system to determine whether each factor makes an individual more or less inclined to use a given drug. In our assumptions, we established that every factor that we are assessing as an impact of drug usage can be assumed an independent impact from every other factor. For example, if males are twice as likely to use drug X as females, then a white male is twice as likely to use drug X as a white female. Thus we can assess the impact of every factor individually. The goal is to calculate a multiplier for each factor such that the product of the multipliers yields an index value, I. If I is greater than 1, then an individual of that particular circumstance is I times more likely to use that substance than the average person. If I is less than 1, then an individual of that particular circumstance is I times less likely to use that substance than the average person. % Substance Use is given by Fig 2.3. The multiplier can be calculated with the following process:

- Multiply the percentage of the population in the subgroup by the percentage of the subgroup that uses the drug.

- Repeat the above step for each subgroup within the demographic.

- Sum each value calculated from the above two steps (This is the proportion of the total population that uses the drug)

- Divide the percentage of people within the subgroup who use the drug by the sum from step above.

The multiplier system ensures both that the usage ratio between subgroups of a given demographic is independent of other factors. Furthermore, the process described above ensures that the sample proportion

of Americans using each drug (Fig 2.3) is maintained.

Example: To find the multiplier for each race for marijuana usage, we scale the average probability of use to the probability of use given that race using data from Fig 2.1. Average probability of marijuana use is:

$$0.24 \cdot 0.065 + 0.09 \cdot 0.221 + 0.67 \cdot 0.079 = 0.08842$$

The multiplier for an individual using marijuana while under the age of 18 is:

$$\frac{.065}{.08842} = .735$$

The multipliers for each factor have been calculated in the same fashion and are displayed in the table below:

Multipliers per Drug and Demographic (Fig 2.2)					
Demographic	Subgroup	Marijuana	Alcohol	Opiates	Nicotine
Age	0-17	.735	.2208	.75	.25163
	18-25	2.499	1.25588	1.6667	1.494
	26+	.893	1.2447	1	1.2017
Race	White	1.067	1.1235	1.1098	1.1358
	African-American	1.263	.8505	.872	1.094
	Asian	.40289	.7595	.396	.420
	Hispanic/Latino	.882	.884	1.0305	.771
Gender	Male	1.244	1.0727	1.1567.	1.2709
	Female	.7633	.9296	.8482	.7376
Poverty -Level	< 100%	1.3957	.6747	1.4263	1.4445
	100 − 199%	1.1595	.752	1.2678	1.2381
	> 200%	.9019	1.1082	.8716	.8759

National Substance Use (Fig 2.3) [4][30]				
	Marijuana	Alcohol	Opiods	Nicotine
% of Americans	6.7	56	.3	15.4

Finally, we can take the product of the multipliers for the different demographic subgroups represented by an individual and multiply by the proportion of the overall population that uses the drug to commute the probability that the given individual uses the drug.

Example: A fictional 18-25 year old African American female at a 100-199% poverty level named Nicole using nicotine. The overall population drug use probability is 15.4%.

$$I = 1.494 \cdot 1.094 \cdot 0.7376 \cdot 1.2381 = 1.4926$$

Thus, the probability that Nicole uses nicotine given her circumstances is:

$$1.4926 \cdot 15.4\% = 22.9\%$$

2.5 Predicting Number of Drug Users out of 300 Seniors

Since we are assuming that the 300 seniors are demographically representative of the US, we multiply 300 by the percentages from age group 18-25 in Fig 2.1 to determine the number of students using each substance.

Thus, when rounding to the nearest whole number:

- 66 of the seniors will use marijuana

- 169 of the seniors will use alcohol

- 6 of the seniors will misuse or abuse unprescribed opiates

- 87 of the seniors will use nicotine

2.6 Summary and Assessment of Model

2.6.1 Sensitivity Analysis

To perform a sensitivity analysis, we examine the sensitivity of the multipliers in Figure 2.2. For all four drugs presented, people of ages 0-17 have a lower multiplier than people ages 18+. The lower multipliers signify that children consume drugs *under* the national averages. This can be explained by reduced access to the drugs for this age group compared to the others. For nicotine and alcohol, this is consistent with the fact that they are legal only for adults and not for children, thus adult usage should be significantly higher. For drugs illegal across all ages, it is still true that adults have far greater access, especially when compared to the youngest children.

The alcohol category has the greatest sensitivity because it has the highest sample percent usage. For example, when varying the age groups, it is predicted that $(.56)(.2208)$, or 12.3648%, of children ages 0-17 use alcohol, while $(.56)(1.25588)$, or 70.33%, of people age 26 or over use alcohol. While the alcohol category is most sensitive, our model is still valid since the difference in usage between age groups is reasonable with a legal drinking age of 21.

2.6.2 Dimensional Analysis

When the unit-less multipliers are multiplied by the percentage of the entire population that uses a drug, then the final unit is a percentage. The percentage can then be used to calculate the proportion of a group that may use a drug. Often dimensional inconsistency or confusion between decimals and percentages can create illogical values. All values calculated in our model yield positive percentages below 100% but not unreasonably small, so unit errors have not been shown to have occurred.

2.6.3 Strengths and Limitations

The model takes into account multiple factors (age, race, gender, and poverty level) to signify that no one factor can completely determine drug usage. Moreover, the multipliers of our weighted scale for these factors are anchored in and compared to reliable data from the NASDUH's annual substance use report, leaving very little room for error in our calculations.

Possible aspects to improve: One area of improvement would be to increase the scope of our model by adding other categories to our weighted scale, such as family history of substance use. However, due to the constraints of time and available research, we decided not to include categories beyond the present four to preserve the accuracy of our current model. Another possible improvement to the scope of our model would be including Native Americans as one of race subcategories due to their extremely high rate of youth alcoholism [46]. Again, due to a lack of published research on other substance use, we decided to leave them out.

2.6.4 Computation and Verification

The purpose of our computational program was twofold: to calculate all possible combinations of categories and to verify that our results are reasonable. There are 288 possible combinations to categorize an individual: 4 races, 4 drugs, 3 income levels, 3 age brackets, and two genders. $4 \cdot 4 \cdot 3 \cdot 3 \cdot 2 = 288$. To find the index value for a particular individual with varying characteristics, five multipliers would need to be multiplied together. To be more efficient, we created a program (in Java) that generates all 288 indices and labels the characteristics of each product. Computing 288 values with a calculator would have taken more time than a computer program. At the end of the document is the code used as well as the results that the code outputted. However, it is simple and time-effective to calculate the usage probability for any singularly combination of factors.

2.6.5 Verifying the Code

To verify the code, we tested five cases. We computed five index values and compared it to the values that the computer program generated. The results are shown below:

- Male, age under 18, African American, 200+% poverty level, marijuana. The calculator generated an index value of 1.0415. The computer program also generated an index value of 1.0415.

- Female, age 18-25, White, 100-199% poverty level, alcohol. The calculator generated an index value of 0.986. The computer program also generated an index value of 0.986.

- Male, age 26+, Asian, 100-199% poverty level, nicotine. The calculator generated an index value of 0.7941. The computer program also generated an index value of 0.7941.

- Female, age 18-25, Hispanic, <100% poverty level, opiates. The calculator generated an index value of 2.07785. The computer program also generated an index value of 2.07785.

- Female, age under 18, African American, 200+% poverty level, nicotine. The calculator generated an index value of 0.17785. The computer program also generated a value of 0.17785.

The computer program is verified accurate for five possible combinations; thus, it is most likely accurate for the other 283 combination characteristics.

3 Part 3: Ripples

3.1 Restatement of Problem

Given the potential monetary and societal impacts of nicotine, marijuana, alcohol, and unprescribed opiates we are tasked with the following:

- Develop a metric to quantify the impact of substance abuse which considers both financial and non-financial factors.

- Rank nicotine, marijuana, alcohol, and unprescribed opiates based on their respective impacts.

3.2 Assumptions

- Assumption: An individual addicted to any of the four aforementioned drugs will only be addicted to one of the four.

  Justification: There is little to no research precise or general enough to determine the effects of multiple drug addictions.

- Assumption: The proportion of the value of statistical life (VSL) per year is approximately the same at $122,000 per year

  Justification: A majority of those experimenting with drugs are within the range of a healthy adult, and thus can be thought of as relatively similar. The value of statistical life is approximately $9.6 million[27] and an average American lives 78.69[28] years so $\frac{\$9.6 \cdot 10^6}{78.69} = \$122,000/yr$.

- Simplification: We will constrain the impact index to the United States of America.

  Justification: A majority of drug use/abuse research has centered around the United States, so it is logical to only take into consideration the U.S. when analyzing the impacts of the drugs.

- Assumption: We did not consider the comorbidity of the drugs or the effect they have on inducing the usage of the others.

  Justification: There are a lack of concrete statistics relating the induced effect of "gateway" drugs on the abuse of different drugs.

- Assumption: Workers with SUDs (Substance Use Disorder) are more likely to miss a greater volume of work days under typical excuses such as illness/injury.

  Justification: Substance abuse leads to erratic or endangering behavior of individuals and decline of their physical health, and thus those who are affected by SUDs are more prone to injury or illness than the non-addicted workers.

- Assumption: The average income of a worker in America per year can be represented by the annual median personal income.

  Justification: The distribution of incomes of workers is skewed left and the median is more resistant to skewness and thus gives a more accurate sense of an average worker's personal income.

- Assumption: Lost tax revenues from people who die directly from drugs are insignificant.

  Justification: The money lost from taxes will be negligible in comparison to the tax revenue generated nationwide and thus it is insignificant.

- Assumption: All those who use nicotine products are employed.

  Justification: A majority of those who are able to afford quantities of nicotine products should be within the workforce. There was also no relevant data concerning this statistic.

- Assumption: There is negligible cost of criminal justice of nicotine usage [44].

  Justification: The effect of usage of nicotine is mostly concentrated around the impacts of direct medical care and lost productivity.

3.3.1 Developing the Model

Our model is a social cost-benefit analysis that consists of 3 major factors - Total Market Value (M), Humanitarian Costs (H) and External Costs (E) - to quantify the total impact of a drug as a monetary amount(TI). The M factor refers to the total monetary market value of the drug, Humanitarian Costs encompass the social cost of deaths and addiction due to the perspective drug, and External Costs refer to the summation of the societal costs of lost productivity, criminal justice and social welfare/health care. With this method, we assigned monetary values to typically non-financial factors which are easily comparable and do not require the creation of arbitrary coefficients to compare and contrast the financial and non-financial factors.

3.3.2 Total Market Value

The total market values of the four drugs were gathered from Euromonitor, Statista and Grand View Research [19][20][21][22].
They are as follows:

Total Market Values of Each Drug in the US (2017) (Figure 3.1)	
Drug	Total Market Value(M)
Nicotine	$\$1.20 \cdot 10^{11}$
Marijuana	$\$7.97 \cdot 10^{9}$
Alcohol	$\$2.31 \cdot 10^{11}$
Unprescribed Opiates	$\$2.30 \cdot 10^{10}$

3.3.3 Humanitarian Costs

To account for the social costs associated with drug-induced death and/or drug dependence, we scaled the annual direct fatality rates and addiction rates with the value of statistical life (VSL) approximated as \$9.6 million per person [27] or \$122,000 per person per year. We represent this as

$$H = p \cdot \$9.6 \cdot 10^{6} + A$$

where H represents the total annual social cost of addiction and death for each drug, p represents the number of direct deaths due to the drug, and A represents the calculated "loss of life" due to addiction to the drug.

A is further defined as:

$$A = P \cdot DW \cdot \$122,000$$

Where P is the number of people addicted to the drug and DW is the estimated disability weight due to addiction. Disability Weights are official numbers which reflect the severity of a disease on a scale from 0

(perfect health) to 1 (equivalent to death)[45]. The constant $122,000 is multiplied to quantify the proportion of the value of statistical Life wasted annually, due to an individual bringing themselves closer to death by the DW proportion. The product is finally multiplied by the number of people addicted to the specific drug to extrapolate the individualistic statistic to a nationwide issue.

The statistics used in our calculations are as follows:

Annual Death and Addiction Statistics from 2017 (Figure 3.2) [23][24][25][26][30]		
Drug	Number Addicted(people)	Direct Fatalities(people/yr)
Nicotine	$5 \cdot 10^7$	480,317
Marijuana	$3.5 \cdot 10^7$	0
Alcohol	$1.5 \cdot 10^7$	34,865
Unprescribed Opiates	$1.16 \cdot 10^7$	30,571

Disability Weights (DW) from 2016 (Figure 3.3)[29]	
Drug	Disability Weight
Nicotine	0.486 (0.329-0.637)*
Marijuana	0.266 (0.178-0.364)
Alcohol	0.134 (0.122-0.137)
Unprescribed Opiates	0.697 (0.51-0.843)

Note*: Amphetamines were used to approximate the disability weight of someone severely affected by nicotine as their are currently no disability weights for nicotine based problems. Amphetamines and nicotine are both stimulants and thus they are suitable substitutes.

Using these values, we calculated A($) and H($) for each of Nicotine, Marijuana, Alcohol, and Unprescribed Opiates.

Sample Calculation:
$$A_{Nicotine} = 5 \cdot 10^7 \cdot 0.486 \cdot \$122,000 = \$2.96 \cdot 10^{12}$$
$$H_{Nicotine} = 480,317 \cdot \$9.6 \cdot 10^6 + 2.96 \cdot 10^{12} = \$7.57 \cdot 10^{12}$$

Calculated Values (Figure 3.4)			
Drug	p $\cdot \$9.6 \cdot 10^6$	A($)	H($)
Nicotine	$\$4.61 \cdot 10^{12}$	$\$2.96 \cdot 10^{12}$	$\$7.57 \cdot 10^{12}$
Marijuana	$0	$\$1.14 \cdot 10^{12}$	$\$1.14 \cdot 10^{12}$
Alcohol	$\$3.35 \cdot 10^{11}$	$\$2.45 \cdot 10^{11}$	$\$5.80 \cdot 10^{11}$
Unprescribed Opiates	$\$2.93 \cdot 10^{11}$	$\$9.85 \cdot 10^{11}$	$\$1.278 \cdot 10^{12}$

3.3.4 External Costs

The first factor we considered for the monetary magnitude external costs from the aforementioned drugs was the reduction of productivity in workers caused by substance usage. The following table analyzes the rate of absenteeism induced as a result of each specific drug.

Average Absenteeism Rates of Workers Due to Substance Use Disorders (Figure 3.5) [32]	
Drug	Average Days Missed Per Worker Per Year
Nicotine	18.2
Marijuana	15.4
Alcohol	14.1
Unprescribed Opiates	29.0

To calculate the monetary value of the lost potential productivity in the U.S. per year, we used the following equation:

$$P = (D_2 - D_1) \cdot \frac{I}{365} \cdot N$$

where P is the productivity lost per year in dollars, D_2 is the average amount of days lost per year for workers with SUDs, D_1 is the average amount of days lost per year for workers not affected by SUDs, I is the median income of workers in dollars per year, and N is the amount of workers using the substance. The average rate of absenteeism for workers who aren't affected by SUDs (D_1) is 10.5 days per year [32] and the median income per year (I) is \$31,099 [35].

Number of users who are workers for each drug are listed in the table below:

Amount of Workers Using a Drug (Figure 3.6) [32][33][34]	
Drug	Workers
Nicotine	$7.47 \cdot 10^7$
Marijuana	$5.98 \cdot 10^6$
Alcohol	$1.43 \cdot 10^8$
Unprescribed Opiates	$6.84 \cdot 10^6$

Sample Calculation:

$$P_{Nicotine} = (18.2 - 10.5) \cdot \frac{31099}{365} \cdot 7.47 \cdot 10^7 = \$4.9 \cdot 10^{10}$$

Monetary values of lost productivity for each substance are delineated in the table below:

Productivity Lost in Dollars Per Year (Figure 3.7)[35]	
Drug	Lost Productivity
Nicotine	$\$4.9 \cdot 10^{10}$
Marijuana	$\$2.48 \cdot 10^9$
Alcohol	$\$4.4 \cdot 10^{10}$
Unprescribed Opiates	$\$1.08 \cdot 10^{10}$

In addition to the lost productivity per year, we also identified the effect of criminal justice costs and social welfare costs as crucial components of the total annual external cost of a drug. The following table details the criminal justice costs and social welfare costs for each drug:

Total Cost of Criminal Justice and Social Welfare (Figure 3.8)[36][37][38][39][40][41][42][43]		
Drug	Criminal Justice	Social Welfare
Nicotine	$0	$1.70 \cdot 10^{11}$
Marijuana	$7.7 \cdot 10^9$	$9.5 \cdot 10^7$
Alcohol	$2.5 \cdot 10^{10}$	$1.60 \cdot 10^{11}$ *
Unprescribed Opiates	$1.04 \cdot 10^{10}$ **	$5.18 \cdot 10^{10}$ **

*We included the costs of traffic incidents in the total cost of social welfare for effects of alcohol.

**The most recent statistic found for costs of unprescribed Opiates was in 2011.

The total annual external costs (E) was found as from the summation of the cost of lost productivity, criminal justice and social welfare costs for each drug. This is detailed in the table that follows:

Total External Costs (Figure 3.9)	
Drug	External Cost
Nicotine	$2.19 \cdot 10^{11}$
Marijuana	$1.03 \cdot 10^{10}$
Alcohol	$2.29 \cdot 10^{11}$
Unprescribed Opiates	$7.3 \cdot 10^{10}$

3.4 Summary, Assessment and Implications of Model

To complete the Social Cost-Benefit Analysis, we assigned each factor (M, H and E) as either a cost, or a benefit. We assigned M as a benefit, since market value contributes directly to the GDP of the nation, which helps drives along the job market. We assigned both H and E as costs, since the impact of deaths/addiction and extra spending due to the drugs is a detriment to society's overall well being. This resulted in the following formula for TI, the total annual impact of a drug:

$$TI = M - H - E$$

Sample Calculation:

$$TI_{Nicotine} = 1.2 \cdot 10^{11} - 7.57 \cdot 10^{12} - 2.19 \cdot 10^{11} = -\$7.67 \cdot 10^{12}$$

Total Social Cost-Benefit Analysis Impact (Figure 3.10)				
Drug	M($)	H($)	E($)	Total($)
Nicotine	$1.2 \cdot 10^{11}$	$7.57 \cdot 10^{12}$	$2.19 \cdot 10^{11}$	$-\$7.67 \cdot 10^{12}$
Marijuana	$7.97 \cdot 10^{9}$	$1.14 \cdot 10^{12}$	$1.03 \cdot 10^{10}$	$-\$1.14 \cdot 10^{12}$
Alcohol	$2.31 \cdot 10^{11}$	$5.80 \cdot 10^{11}$	$2.29 \cdot 10^{11}$	$-\$5.78 \cdot 10^{11}$
Unprescribed Opiates	$2.30 \cdot 10^{10}$	$1.28 \cdot 10^{12}$	$7.3 \cdot 10^{10}$	$-\$1.33 \cdot 10^{12}$

According to the total social cost-benefit impact of each of the four drugs, nicotine, marijuana, alcohol, and unprescribed Opiates, we rank the severity of impact of each of them in decreasing order of magnitude starting from the most impactful: Nicotine, Unprescribed Opiates, Marijuana, and Alcohol.

When compared with various sources, our value of societal cost (the negative sign depicts cost) is significantly larger. However, this is explainable due to our increased emphasis on the value of the humanitarian perspective and its respective social costs, which total cost estimates do not often take account of.

Additionally, the final ranking of Nicotine > Unprescribed Opiates > Marijuana > Alcohol is justifiable from a societal perspective. Alcohol is by far the most socially accepted drug and thus it logically should fall below the others in terms of cost. While the amount of people who have used alcohol within the past 30 days may be deceptively large, the true number of addicted users in actuality is a much smaller amount [30], which is taken into account in the humanitarian costs, which also holds the most weight in the total social cost-benefit impact. Additionally, the stark decline and battle against the tobacco industry justifies the first

place rank of Nicotine. Assuming our model as true, there are multiple possible courses of action. Firstly, the United States government should focus their efforts on lowering rates of Nicotine consumption due to its much greater monetary cost towards society. Secondly, although Marijuana has the second lowest monetary, our model would urge caution towards legalization due to it having greater than double the monetary cost of alcohol, society's most tolerable drug.

In terms of the dimensional stability of our model, the total costs of criminal justice and social welfare are inherently measured in dollars and are thus correct. On the other hand, the productivity, as we defined as lost income in dollars, can be verified through dimensional analysis. Our equation: $P = (D_2 - D_1) \cdot \frac{I}{365} \cdot N$ uses units of days for D_2 and D_1, dollars per year per person for I, days per year for the 365 days in a year, and amount of people for N, which indeed results in units of dollars for P.

3.5 Strengths and Limitations

Strengths: One strength of our model is that it accounts for financial and non-financial factors of impact. The model addresses varied statistics such as annual death rate and annual lost productivity which limits the sensitivity of the model towards any certain variable. Another proficiency is that our impact model is split into 3 categories which are easily applicable to differing situations and concerns.

Limitations: One area for growth lies in our generalization of the emotional and mental struggles of human life as the value of statistical life. One could argue that a human is unquantifiable, and as such, should not be given a dollar value. Furthermore, our model is heavily weighted towards the Humanitarian index. However, given the personal pull of the humanist perspective, and the extreme gravity of the loss of life to addiction or death, this is generally acceptable.

4 References

1 The 2016 Surgeon General's Report: E-Cigarette Use Among Youth and Young Adults (Chapter 2, Figure 2.2b)

2 https://www.cdc.gov/mmwr/volumes/66/wr/mm6623a1.htm

3 https://www.fda.gov/downloads/TobaccoProducts/PublicHealthEducation/ProtectingKidsfromTobacco/UCM625955.

4 https://www.census.gov/quickfacts/fact/table/US/RHI125217RHI125217

5 https://www.hhs.gov/ash/oah/adolescent-development/substance-use/drugs/tobacco/trends/index.html

6 https://www.tobaccofreekids.org/assets/content/press_office/2018/2018_06_06_yts_infographic_pr.png

7 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3662471/

8 https://www.statista.com/statistics/737849/share-americans-age-group-smokes-marijuana/

9 https://www.drugabuse.gov/publications/research-reports/marijuana/what-scope-marijuana-use-in-united-states

10 https://www.childstats.gov/americaschildren/tables/pop1.asp

11 https://www.nami.org/learn-more/mental-health-by-the-numbers

12 https://www.census.gov/content/dam/Census/library/visualizations/2018/demo/p60-263/figure4.pdf

13 https://www.kff.org/other/state-indicator/distribution-by-fpl/?currentTimeframe=0selectedDistributions=under-100percentselectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7DsortModel=%7B%22colIc

14 https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/alcohol-facts-and-statistics

15 https://www.deseretnews.com/article/900004001/12-percent-of-americans-say-a-family-member-is-addicted-to-opioids-new-survey-shows.html

16 https://www.addictioncenter.com/nicotine/

17 https://www.drugabuse.gov/publications/research-reports/heroin/scope-heroin-use-in-united-states

18 https://www.addictioncenter.com/nicotine/

19 https://www-statista-com.eu1.proxy.openathens.net/statistics/491248/alcoholic-drinks-united-states-market-value/

20 https://www.statista.com/statistics/491709/tobacco-united-states-market-value/

21 https://www-statista-com.eu1.proxy.openathens.net/chart/12406/us-marijuana-market-the-grass-is-getting-greener/

22 https://www.grandviewresearch.com/industry-analysis/opioids-market

23 https://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/tobacco_related_mortality/index.htm

24 https://www.cdc.gov/mmwr/volumes/67/wr/mm675152e1.htm?s_cid=mm675152e1_w

25 https://www.addictioncenter.com/nicotine/

26 https://www.washingtonpost.com/news/wonk/wp/2017/04/19/11-charts-that-show-marijuana-has-truly-gone-mainstream/?utm_term=.4e3bfb08d53a

27 https://cms.dot.gov/sites/dot.gov/files/docs/2016

28 https://www.usatoday.com/story/news/nation/2014/10/08/us-life-expectancy-hits-record-high/16874039/

29 http://ghdx.healthdata.org/ihme-data/global-burden-disease-study-2016-gbd-2016-disability-weights

30 https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHDetailedTabs2017/NSDUHDetailedTabs2017.

31 https://truthinitiative.org/news/juul-sales-increase-more-600-year-underscoring-popularity-among-teens

32 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5671784/

33 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2598486/

34 $https://www.cdc.gov/mmwr/volumes/67/wr/mm6714a1.htm?s\_cid = mm6714a1\_w$

35 https://fred.stlouisfed.org/series/MEPAINUSA672N

36 https://www.americanprogress.org/issues/criminal-justice/reports/2018/06/27/452819/ending-war-drugs-numbers/

37 https://altarum.org/news/economic-toll-opioid-crisis-us-exceeded-1-trillion-2001

38 https://www.asam.org/docs/advocacy/societal-costs-of-prescription-opioid-abuse-dependence-and-misuse-in-the-united-states.pdf

39 https://www.cdc.gov/chronicdisease/about/costs/index.html

40 https://www.verywellmind.com/impact-on-society-63268

41 https://health.costhelper.com/emergency-room.html

42 https://www.politifact.com/punditfact/statements/2018/jul/06/patrick-kennedy/has-marijuana-caused-emergency-room-visits-skyrock/

43 https://www.madd.org/blog/stats/drunk-driving-costs-the-united-states-132-billion-a-year/

44 https://www.drugabuse.gov/publications/research-reports/tobacco-nicotine-e-cigarettes/what-scope-tobacco-use-its-cost-to-society

45 https://www.who.int/healthinfo/global_burden_disease/daly_disability_weight/en/

46 https://pubs.niaaa.nih.gov/publications/arh40/152-160.htm

5   Code Appendix in Java

```java
/**
 * The Multiplier class calculates the multiplier for all 288
 * combinations of possible gender, income, age, race, and drug.
 * The arrays of individual multipliers are instantiated, and
 * then through a series of nested for loops, one entry from each
 * of four arrays is multiplied to calculate an index. The index is
 * then used in the main paper to find the number of students, out
 * of 300, that would use each drug.
 * @author (Team #11729)
 * @date (March 3, 2019)
 */
/**
 */
public class Multiplier
{
    public static void main (String[] args)
    {
        double[] genderM = new double[]{1.244,0.7633};
        //male, then female multipliers for gender and marijuana
        double[] incomeM = new double[]{1.3957,1.1595,0.9019};
        //<100% poverty, then 100-199%, then 200+%
        //multipliers for income and marijuana
        double[] raceM = new double[]{1.067,1.263,0.40289,0.882};
        //white, then African American, then Asian, then
        //Hispanic multipliers for race and marijuana
        double[] ageM = new double[]{0.735,2.499,0.893};
        //<18 years old, then 18-25 years old, then 26+ years old
        //multipliers for age and marijuana

        double[] genderA = new double[]{1.0727,0.9296};
        //same order as marijuana, except with alcohol
        double[] incomeA = new double[]{0.6747,0.752,1.1082};
        //alcohol and income
        double[] raceA = new double[]{1.1235,0.8505,0.7595,0.884};
        //alcohol and race multipliers
        double[] ageA = new double[]{0.2208,1.25588,1.2447};
        //alcohol and age multipliers

        double[] genderO = new double[]{1.1567,0.8482};
        //opiates multipliers
```

```java
double[] income0 = new double[]{1.4263,1.2678,0.8716};
double[] race0 = new double[]{1.1098,0.872,0.396,1.0305};
double[] age0 = new double[]{0.75,1.6667,1.0};

double[] genderN = new double[]{1.2709,0.7376};
//nicotine multipliers
double[] incomeN = new double[]{1.4445,1.2381,0.8759};
double[] raceN = new double[]{1.1358,1.094,0.42,0.771};
double[] ageN = new double[]{.25163,1.494,1.2017};

String gender = " "; String income = " ";
String race = " ";String age = " ";
//creates empty strings, which will be filled in the
//nested for loop


for (int genderCounter=0;genderCounter<2;genderCounter++)
//goes through for loop for gender
{
    if (genderCounter==0)
    {
        gender = "Male, ";
        //the first entry of the gender arrays is for males
    }
    else
    {
        gender = "Female, ";
        //second entry of the gender arrays is for females
    }
    for (int incomeCounter=0;incomeCounter<3;incomeCounter++)
    //goes through for loop for income
    {
        if (incomeCounter==0)
        {
            income = "<100% poverty level, ";
            //first entry of income arrays
        }
        else if (incomeCounter==1)
        {
            income = "100-199% poverty level, ";
```

```
        //second entry of income arrays
}
else
{
    income = "200+% poverty level, ";
    //third entry of income arrays
}
for (int raceCounter=0;raceCounter<=3;raceCounter++)
//race for loop
{
    if (raceCounter==0)
    {
        race = "White, ";
        //first entry of race arrays
    }
    else if (raceCounter==1)
    {
        race = "African American, ";
        //second entry of race arrays
    }
    else if (raceCounter==2)
    {
        race = "Asian, ";
        //third entry of race arrays
    }
    else
    {
        race = "Hispanic/Latino, ";
        //fourth entry of race arrays
    }
    for (int ageCounter=0;ageCounter<3;ageCounter++)
    //age for loop
    {
        if (ageCounter==0)
        {
            age = "Age <18, ";
            //first entry of age arrays
        }
        else if (ageCounter==1)
        {
```

```
        age = "Age 18-25, ";
        //second entry of age arrays
}
else
{
        age = "Age 26+, ";
        //third entry of age arrays
}
System.out.println(gender + income + race +
age + "Marijuana Multiplier: "
+ genderM[genderCounter]*incomeM[incomeCounter]*
raceM[raceCounter]*ageM[ageCounter]);
//Multiplies multipliers together,
//using the marijuana arrays, to calculate an index


System.out.println(gender + income + race +
age + "Alcohol Multiplier: "
+ genderA[genderCounter]*incomeA[incomeCounter]*
raceA[raceCounter]*ageA[ageCounter]);
//Multiplies multipliers together,
//using the alcohol arrays, to calculate an index


System.out.println(gender + income + race +
age + "Opiates Multiplier: "
+ genderO[genderCounter]*incomeO[incomeCounter]*
raceO[raceCounter]*ageO[ageCounter]);
//Multiplies multipliers together,
//using the opiates arrays, to calculate an index


System.out.println(gender + income + race +
age + "Nicotine Multiplier: "
+ genderN[genderCounter]*incomeN[incomeCounter]*
raceN[raceCounter]*ageN[ageCounter]);
//Multiplies multipliers together,
//using the nicotine arrays, to calculate an index
```