# A study into the relationship between cannabis legalisation and suicide and traffic-related fatalities

May 7th 2021


## SUMMARY

This project looks at the relationship between the legalisation of recreational cannabis and the traffic fatality rate for all ages, and the suicide rate for those between the ages of 21 and 34. To do so, focus was given to the USA, comparing the eight states plus the District of Columbia where cannabis has been legalised to the remaining 42 in order to conclude whether cannabis legalisation could be identified as a causal factor. This paper was inspired by the existing literature on the relationship between recreational cannabis use and its effects on depression, leading us to question whether there exists a relationship between the policy in question and the suicide rate. Moreover, a study regarding alcohol consumption and road traffic fatalities again directed the study to whether recreational cannabis use has an impact on the number of these types of deaths. As can be said for the existing literature, the analysis did find trends and correlation among the data, however the data was insufficient to establish a causal relationship between the policy and either of the supposed effects.


## INTRODUCTION

This paper aims to answer whether the legalisation of recreational cannabis use in certain US states led to an change in the number of traffic and suicide-related fatalities. In order to fully understand the effects that contribute to the number of fatalities, multiple control variables were included, for example: the population, amount of licensed drivers, lane-length, and land area. The ability to interpret the estimates as causal effects can be limited by some factors. For example, there may be unobservable differences between people with different mental health states and their physical ability to drive. However, the size of available datasets and control groups may well be sufficient enough to eliminate these issues. Control groups also need to be considered to allow clear assumptions to be made such as: the level of unemployment, poverty rates, and number of hospitals within each jurisdiction. This further enables the study to correctly compare the effects and differences in fatalities.

## LITERATURE REVIEW

The existing relevant literature here is broad and informative. Aydelotte et al. (2017) attempts to evaluate motor vehicle accidents in Washington and Colorado where recreational cannabis had been legalised. This paper uses public data to monitor the number of fatalities between 2009 and 2015 within control and non-control states. Hence, to correctly answer the research question, it was imperative to pick a time frame and to allocate the control states as the states that have not legalised cannabis. Again, Aydelotte et al. (2017) used approaches that controlled for variables such as traffic characteristics and population, which inspired the research question to consider the controlling factors such as: the number of licensed drivers, number of hospitals, and lane length in order to reduce the effect of the omitted variable bias.

There has also been extensive work carried out by Sewell et al. (2010) suggesting that the reason for the fatal accidents is due to the impairments caused by drugs i.e. the individuals reaction time. It was found that the increase in fatalities was mainly due to higher doses of cannabis than lower doses. Sewell et al. (2010) considered other drivers as a control group in order to see the full effect of drug use on traffic fatalities. However the validity of this was unstable since it depended upon careful selection of appropriate control groups for comparison. To avoid any difficulty in trying to compare drug drivers and non-drug drivers, it was important for the project to distinguish between the treatment and control group such as the level of employment within each state. It was imperative for Sewell et al (2013) to note that there is no one such causal effect of drug use and traffic fatalities, but the fact there are also other factors which contribute. This idea was important to consider throughout the project because the research must take into account that there is not one cause, but that there are correlations between the two events being tested.

Horwood et al. (2012) focus on the relationship between the frequency of cannabis and the severity of depression utilising four Australasian cohorts. It is relevant to this study as health organisations highlight that those diagnosed with severe depression are at a higher risk of death by suicide compared to those without depression (HHS, 2014). Some of the cohorts found a significant pathway from cannabis use to depression, however, when collated, no direction of causation was identified. These results led the research to focus on whether the legalisation of cannabis in some US states increased suicide rates. Despite the paper focusing on Australasian studies, it is a relevant piece of literature to this US analysis as cannabis usage having adverse mental health effects is a worldwide concern.

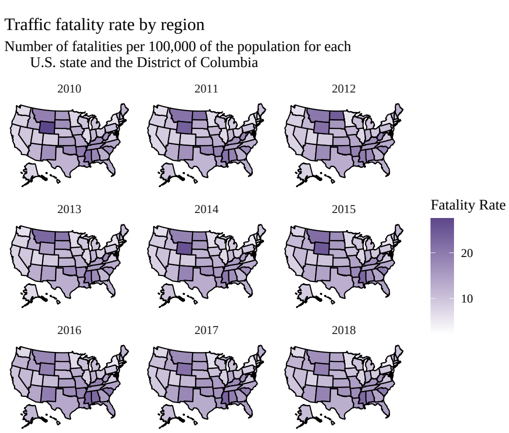
The use of our control variables was motivated by a German study which considered the relationship between socioeconomic status and social isolation (i.e. divorce/living alone) and their impact on suicide rates (Naher, Rummel-Kluge and Hegerl, 2020). The results of this paper highlight a positive relationship between socioeconomic status and social isolation in relation to suicide rates. This paper was important to consider as it sheds light on other potential causes of suicide which may be relevant to take into account throughout the analysis.

## DATA

Our data takes on a relatively simple structure, as follows: the data consists of 51 subjects (the 50 US states plus the District of Columbia), measured on an annual basis between the years of 2010 and 2018, giving a total of 459 observations. A number of variables for each observation are required to carry out the analysis, which revolves around two key variables:

- Traffic fatality rate – this refers to a population-adjusted figure taken from the number of traffic-related fatalities in each state for the given time period. Analysis has been carried out on both the total population, and a subset between the ages of 21 and 34, to reflect the unequal distribution of cannabis users between age groups.
- Suicide rate – this is a population-adjusted figure based on the number of suicides recorded among those between the ages of 21 and 34, again reflecting the distribution of cannabis use. A representation of these two variables over time is shown in figures 1 and 2, from which variation over time can be seen, without any strict pattern, thus warranting further analysis.

Figures 1 and 2



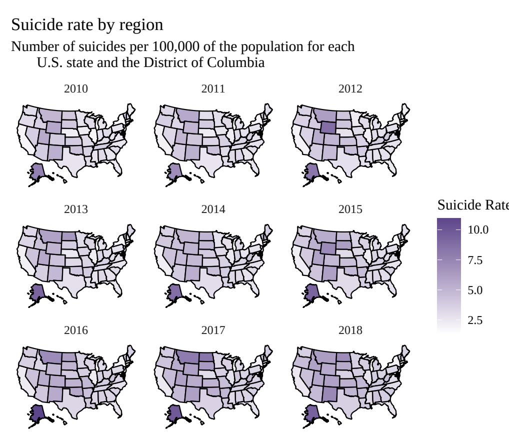
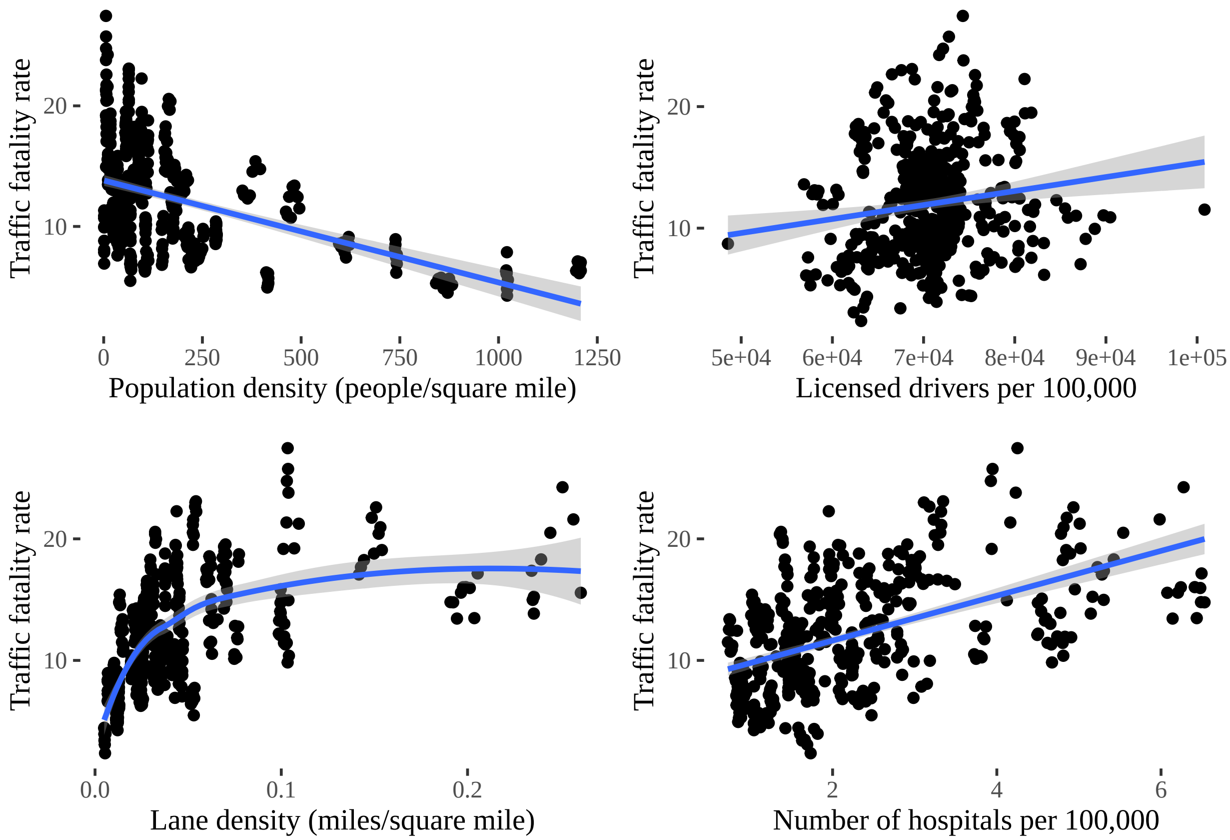
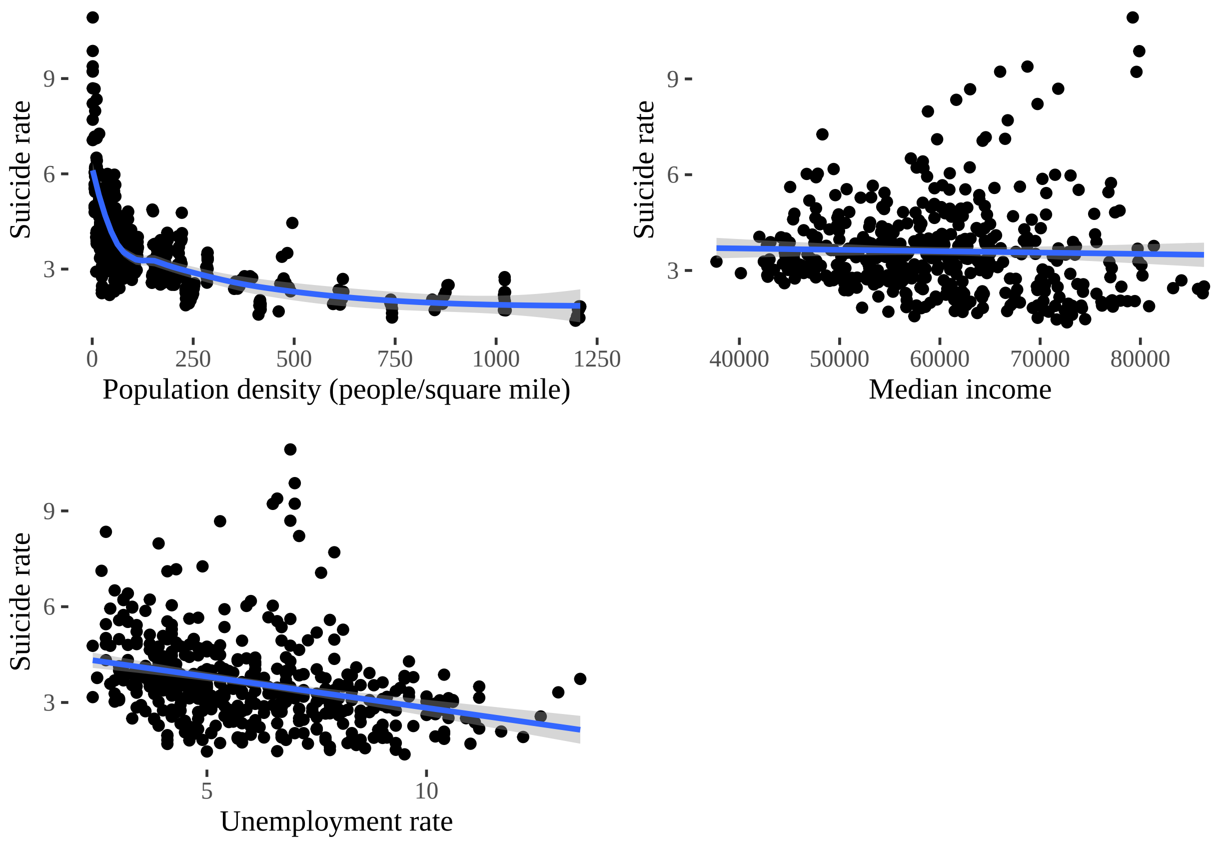



Figure 3




A number of other variables have been necessary in our analysis including control variables; in the case of the traffic fatality this includes population density, the population-adjusted number of licensed drivers, lane density (miles of road lane per square mile), and the number of hospitals per 100,000 of the population; in the case of the suicide rate this includes population density, median income, and unemployment rate. Graphical representations of these variables plotted against their respective dependent variable are shown in figures 3 and 4. Some display clear relationships, while others are less clear; in all cases linear hypothesis testing of the variables and the existing literature warranted their inclusion in the analysis.

Figure 4




Finally, some variables are used simply to facilitate our analysis, these are:

- FIPS state code – a unique two-digit code used to identify states as part of the Federal Information Processing Standard (FIPS)
- Legal – a value used to represent the intensity of the policy in question in a given year (greater detail is given to this within the methodology)
- Population – used to create adjusted variables, allowing comparison between states
- Land area – similarly used for creating adjusted variables. Some summary statistics of all numeric variables necessary for analysis can be seen in Table 1. Detailed information regarding data sources is available in the data appendix.

## METHODOLOGY

Our methodology can be summarised as a multi-state, multi-period, difference-in-difference (DiD) model, differing from the basic DiD model (two states across two time periods) as the treatment in question has

Table 1:

| Statistic | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| FIPS state code | 459 | 28.961 | 15.694 | 1 | 56 |
| Legal | 459 | 0.066 | 0.245 | 0 | 1 |
| Year | 459 | 2,014.000 | 2.585 | 2,010 | 2,018 |
| Population | 459 | 6,240,204.000 | 7,036,145.000 | 564,487 | 39,461,588 |
| Land area | 459 | 69,426.220 | 84,731.360 | 61.050 | 570,640.900 |
| Traffic fatalities (21-34) | 459 | 181.031 | 205.691 | 3 | 1,163 |
| Traffic fatalities | 459 | 678.625 | 721.998 | 15 | 3,837 |
| Lane miles | 459 | 170,339.400 | 117,062.100 | 3,413 | 680,981 |
| Licensed drivers | 459 | 4,254,374.000 | 4,530,297.000 | 384,940 | 27,039,400 |
| Suicide fatalities (21-34) | 455 | 188.974 | 176.150 | 10.000 | 1,059.000 |
| Median Income | 459 | 59,651.160 | 9,709.131 | 37,714 | 86,345 |
| Unemployment rate | 459 | 6.063 | 2.206 | 2.400 | 13.500 |
| Total hospitals | 459 | 98.078 | 80.332 | 7 | 528 |
| Population density | 459 | 406.337 | 1,495.081 | 1.251 | 11,491.350 |
| Traffic fatality rate | 459 | 11.971 | 4.569 | 2.362 | 27.459 |
| Traffic fatality rate (21-34) | 459 | 3.177 | 1.333 | 0.472 | 9.050 |
| Lane density | 459 | 0.045 | 0.046 | 0.005 | 0.261 |
| Licensed drivers per 100,000 | 459 | 70,606.220 | 5,902.718 | 48,544.830 | 100,815.400 |
| Suicide rate (21-34) | 455 | 3.606 | 1.408 | 1.375 | 10.924 |
| Hospitals per 100,000 | 459 | 2.184 | 1.297 | 0.725 | 6.530 |

been applied multiple times across different treatment groups; by including all treatment groups (for which sufficient data was available) and increasing our sample size, we reduce the risk that our results are skewed by any external factors (outliers in the data, measurement error, etc.), and strengthen our analysis. To reflect the staggered rollout of the policy, we rely on the creation of the 'legal' variable, so as to incorporate all groups into a single model; this reflects varying intensity of the treatment over time, taking a value between zero and one, a group not exposed to the policy whatsoever being attributed zero, and a group fully exposed to the policy being given a value of one. As our data is recorded annually, it has sometimes been necessary to attribute a value between zero and one, reflecting the fact that the policy may only have been brought into effect partway through the year, with the policy varying in intensity in a given year as a result. If the policy was brought into effect within the first third of the month, we defined the treatment as occurring for the entire month; if brought in within the second third, it was defined as a half month; if brought in within the final third we considered the treatment not to have been in effect for that month. It was felt that this simplification would simplify our methodology without significantly impacting our results. To facilitate our analysis, we employ three different ordinary least squares (OLS) regression models for each dependent variable, including both simple and multiple regression. Our first model is a simple regression of the dependent variable against the legal variable, with the regression coefficient as our output. In the second model we include variables to represent state-fixed effects and time-fixed effects in order to control for any variation between states and across time that aren't related to the policy, and a variable to control for state-specific time-trends. In the final model, we include multiple different control variables that we believe to be correlated with the dependent variable in question. Each model is represented by the following, respectively:

$$Y = \beta_0 + \beta_1 LEGAL + e$$

$$Y = \beta_0 + \beta_1 LEGAL + \beta_2 STATE + \beta_3 YEAR + \beta_4 STATE : YEAR + e$$

$$Y = \beta_0 + \beta_1 LEGAL + \beta_2 STATE + \beta_3 YEAR + \beta_4 STATE : YEAR + \beta_5 CONTROL_1 + ... + \beta_n CONTROL_n + e$$

## RESULTS

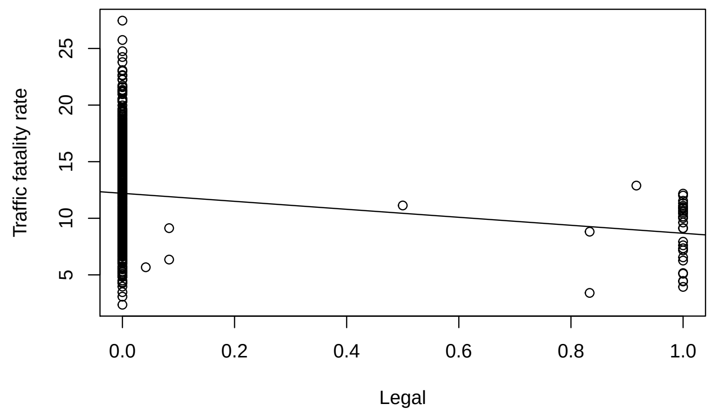
Table 2:

| | *Dependent variable:* | | |
|---|---|---|---|
| | Traffic fatality rate | | |
| | (1) | (2) | (3) |
| Legal | $-3.529^{***}$ | $-0.035$ | $-0.064$ |
| | (0.857) | (0.483) | (0.486) |
| Observations | 459 | 459 | 459 |
| $R^2$ | 0.036 | 0.957 | 0.957 |
| Adjusted $R^2$ | 0.034 | 0.943 | 0.943 |
| Residual Std. Error | 4.492 (df = 457) | 1.088 (df = 349) | 1.091 (df = 345) |
| F Statistic | $16.975^{***}$ (df = 1; 457) | $70.952^{***}$ (df = 109; 349) | $67.996^{***}$ (df = 113; 345) |
| *Note:* | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

The first important result (shown in Table 2) is the statistically significant coefficient estimate of legalization policy and its contribution to traffic fatalities, which implies that a policy of cannabis legalisation is associated with a drop in the traffic fatality rate of approximately 3.529, meaning 3.529 fewer deaths per 100,000 of the population as a result of the policy, shown in Figure 5. However, its associated R^2 number is very low at 0.034, suggesting that only 3.4 % of the observed variation can be explained by the model inputs. In other words, the inputs to the model are not responsible for explaining the variation of the data, even though but the relationship is considered statistically significant. Model 2 has a far greater associated $R^2$ number of 0.943, suggesting that the new inputs do explain the variation within the model, however also produced a high p-value, such that the correlation is no longer statistically significant. Therefore, the new objective was set to find the statistically significant variables and evaluate their relevance in the multiple regression model

3. As per column 3, none of the control variables appear to be statistically significant since they do meet the p-values of .01 or 0.05 or even 0.1 used to establish significance, and they have not impacted the overall coefficient estimate for the legal variable.

Figure 5.



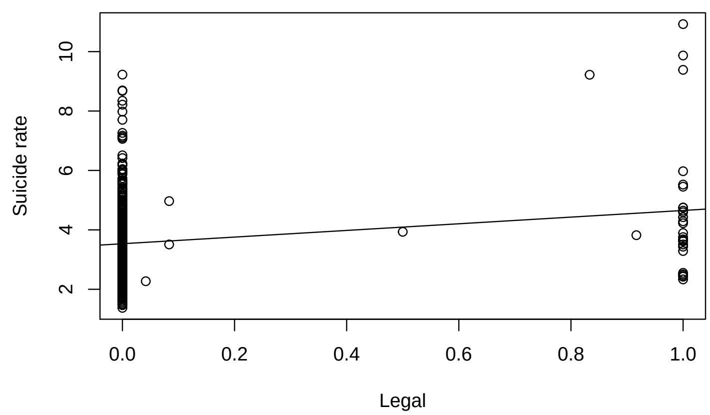

For the second part of our analysis (Table 3), considering the relationship between the legalisation of recreational cannabis and the suicide rate, similar results are found: policy implementation is associated with an approximate increase of 1.183 suicide fatalities per 100,000 population, shown in Figure 6.

Table 3:

| | *Dependent variable:* | | |
|---|---|---|---|
| | Suicide rate | | |
| | (1) | (2) | (3) |
| Legal | 1.120*** | −0.071 | −0.073 |
| | (0.267) | (0.198) | (0.199) |
| Observations | 455 | 455 | 455 |
| $R^2$ | 0.037 | 0.927 | 0.927 |
| Adjusted $R^2$ | 0.035 | 0.903 | 0.903 |
| Residual Std. Error | 1.383 (df = 453) | 0.438 (df = 345) | 0.439 (df = 342) |
| F Statistic | 17.621*** (df = 1; 453) | 39.921*** (df = 109; 345) | 38.593*** (df = 112; 342) |
| *Note:* | | | *p<0.1; **p<0.05; ***p<0.01 |

Similarly to the simple regression model of traffic fatalities against the legal variable, it produces a very low $R^2$ number of $R^2 = 0.035$. When further regression analysis is carried out with the inclusion of state-fixed effects, time-fixed effects, state time-trends, and relevant control variables, the effect of the policy implementation is reduced and is found to be statistically insignificant. As with our first analysis, the relationship was found to be statistically insignificant, however the associated $R^2$ value increased to 0.903, suggesting that the new inputs explain the variation of data. Thus the model indicates that a causal relationship can not be established between the legalisation of recreational cannabis and either the traffic fatality rate or the suicide rate, as we expected based on the existing literature.

Figure 6.




## CONCLUSION

This paper contributed to the existing literature on the relationship between the legalisation of cannabis for recreational use and both traffic- and suicide-related fatalities. After carrying out this analysis, no statistically significant relationship was identified, hence, currently, the two cannot be linked, thus cannot be used as a legitimate political leverage. The lack of a statistically significant causal relationship between the variables in question repeats findings found across the existing literature. Future research in the same area may be able to take advantage of more extensive data, for example for more recent years than 2018, or with greater level of detail about the nature of fatalities, and may be able to prove a causal relationship where this paper was not able to.

## DATA APPENDIX

- Land area: sourced from the US Census Bureau, available from the 2010 US Census, Table 18, p. 42: https://www.census.gov/prod/cen2010/cph-2-1.pdf

- Lane miles: sourced from the FHWA's annual report (Table HM-60), a sample table available at: https://www.fhwa.dot.gov/policyinformation/statistics/2013/hm60.cfm
- Licensed drivers: sourced from the NHTSA's Annual Report Tables, available at: https://cdan.nhtsa.gov/tsftables/tsfar.htm
- Median income: sourced from the US Census Bureau's Current Population Survey , available at: tinyurl.com/npk7dy4e
- Population: sourced from the US Census Bureau's population estimates for 2010-2019, available at: https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html#par_textimage_1574439295
- Suicides (age 21-34): sourced from the CDC's WONDER Database, available at: https://wonder.cdc.gov/
- Traffic fatalities: sourced from the NHTSA's Annual Report Tables, available at: https://cdan.nhtsa.gov/tsftables/tsfar.htm
- Unemployment rate: sourced from the US Bureau of Labour's historical annual series of data, available at: https://www.icip.iastate.edu/tables/employment/unemployment-states

## BIBLIOGRAPHY

Bibliography Aydelotte, J.D., Brown, L.H., Luftman, K.M., Mardock, A.L., Teixeira, P.G. R., Coopwood, B. and Brown, C.V. R. (2017). Crash Fatality Rates After Recreational Marijuana Legalization in Washington and Colorado. American Journal of Public Health, [online] 107(8), pp.1329–1331. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5508149/.

HHS.gov. (2015). Does depression increase the risk for suicide? [online] Available at: https://www.hhs.gov/answers/mental-health-and-substance-abuse/does-depression-increase-risk-of-suicide/index.html.

Horwood, L.J., Fergusson, D.M., Coffey, C., Patton, G.C., Tait, R., Smart, D., Letcher, P., Silins, E. and Hutchinson, D.M. (2012). Cannabis and depression: An integrative data analysis of four Australasian cohorts. Drug and Alcohol Dependence, [online] 126(3), pp.369–378. Available at: https://www.sciencedirect.com/science/article/pii/S0376871612002190.

Naher, A.-F., Rummel-Kluge, C. and Hegerl, U. (2020). Associations of Suicide Rates With Socioeconomic Status and Social Isolation: Findings from Longitudinal Register and Census Data. [online] Frontiers.

Available at: https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00898/full.

Sewell, R.A., Poling, J. and Sofuoglu, M. (2009). The Effect of Cannabis Compared with Alcohol on Driving. American Journal on Addictions, [online] 18(3), pp.185–193. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2722956/.