Legalization and Drug Deaths	
Contact	Ashley Sieper, Tristan Marchand, Sonam Varaiya, Annie Shi, Saanchitha Kottige asieper@bu.edu, tmarch@bu.edu, sonamv@bu.edu, annieshi@bu.edu, skottige@bu.edu 813-403-8933, 504-657-7327, 857-264-8138, 410-599-8603, 857-205-6470
Organization	N/A
Organization Description	N/A
Project Type	Data Science
Project Description	Our project seeks to determine the effect of cannabis/marijuana legalization (both medicinal and recreational) on drug-related deaths. Do we see an increased rate due to increased drug use, or do we see a decreased rate due to cannabis products being used as substitutes to other drugs, more resources for people with drug-related issues, and other effects of cannabis legalization. We intend to use the outcomes of states with already legalized marijuana to predict the outcomes in other states with similar statistics.

Data Sets	DATA DESCRIPTION: Data drug poisonings, measured as the number of drug-related overdoses both unintentional and intentional and economic conditions per 50 U.S. states and the District of Columbia over a time span of 1999-2016 (1999 - 2018).
	https://www.kaggle.com/andrewgatchalian/drug-deaths-economic-conditions-19992016?select=drug_abuse_data.csv
	(2017 and 2018 by Data Scraping)
	https://www.cdc.gov/drugoverdose/data/statedeaths.html
	
	(US Bureau of Labor Stats -for unemployment, min wage)
	https://www.bls.gov/cps/
	
	Attributes:
	- US States
	- Years: 1999 - 2018
	- Drug classification
	- Legalization data (medical use and recreational for every state)
	  - Drug accessibility (ties in with legalization)
	- Poverty: average cost of living in each state or some other metric
	  - Minimum wage
	  - GDP Per capita
	  - CPI

	- Demographics
	  - Population
	  - Age
	  - Gender
	  - Income (tied to poverty)
	  - Religion
	  - Race

Suggested Steps	**Step 1** - Find the effect we see on drug deaths in Colorado after medicinal legalization and after recreational legalization. **Step 2** - Analyze if we see this same pattern in other states and run the same experiment for states with legalization in some capacity. **Step 3** - After finding the pattern in several states with already legalized marijuana, use clustering to group states by income/capita, demographics, population, etc. so we can make predictions based on state type. **Step 4** - At this point we will have groups of states with similar statistics. We will use the outcomes of states with already legalized marijuana to predict the outcomes in other states with similar statistics. - Add 2017 and 2018 to the data set (Data Scraping): https://www.cdc.gov/drugoverdose/data/statedeaths.html
Questions to be answered in Analysis	- What is the effect of cannabis legalization on drug-related deaths? - Are there any outliers? If so, what have they done differently? - Through analysis of previously collected data based on various attributes - predict rate of drug overdoses in future years **Deliverable Questions:** - What effect do we see on drug deaths in Colorado after medicinal legalization and after recreational legalization? - Do we see the same pattern we saw in Colorado in other states? - Use the clustering algorithm to make predictions about other states' drug deaths based on state characteristics (such as income/capita, demographics, population, etc.). - Will states have similar drug death trends to those which have similar characteristics to it (using the outcomes of states with already legalized marijuana)?
Additional Information	- Project can be expanded in the future to look at how the Covid-19 crisis-affected drug-related deaths and the future legalization of drugs. - Completion of the project will draw conclusions about the rates of drug death in states that are yet to legalize marijuana medicinally and recreationally. - Future studies should look at the affect of federal decriminalization of marijuana, decriminalization of all drugs (ie. Oregon in 2020), and if our predictions post-legalization were accurate.

Limitations	- Drug abuse statistics are underreported - But will this actually sway our data? - Missing data from states that have legalized in recent years