SCRUM

Week 1: Feb 1 - Feb 7

- Worked on the project proposal
    - Discussed different ideas and picked one that we found the data to work with and found interesting to work on
        - Impact of marijuana legalization on drug related deaths
    - Developed our research questions
        - 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
    - Found datasets that would work
    - Defined various steps to make progress over the semester
        - **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.

Week 2: Feb 8 - Feb 14

- Got the project approved
- Updated the project proposal to include feedback from Lance
    - Shifted focus to one state (Colorado) rather than all US states
- Identified limitations of the project
    - Drug abuse statistics are underreported
        - But will this actually sway our data?
    - Missing data from states that have legalized in recent years

Week 3: Feb 15 - Feb 21

- Deliverable 1: https://docs.google.com/document/d/1Efu8MXFL9zmSdQsZuoTw_lLE7CSLuQd1pjbZzbzHyK0/edit?usp=sharing
    - Looked at year on year change in rate of deaths in Colorado before legalization of Marijuana, after medical legalization, after recreational legalization
    - Found that the rate decreased after both points of legalization

Week 4: Feb 22 - Feb 28

- Modified Project Proposal
- New Steps
    - **Step 1**: Check one variable for Colorado (D1 - drug deaths)
    - **Step 2**: Check other variables for Colorado (discussed below)
    - **Step 3**: Build model for Colorado (D2)
    - **Step 4**: Use model on test state - California (D3)
    - **Step 5**: If the model works, use it to predict how legalization will impact those variables in a state where marijuana has not been legalized for recreational use (D4)

Week 5: Mar 1 - Mar 7

- Meeting with Lance
- Identified variables to help build our model for Colorado
    - Suicide Rates
    - Cigarette Sales
    - Alcohol Consumption
    - Alcohol-Related Death
    - DUI and DWI
    - Admissions to substance abuse treatment facilities
    - Black Market Marijuana Sales
    - Drug-related Crime
    - Unemployment
    - Drug Tests by Private Companies
    - Underage Marijuana Usage
    - Income per Capita
    - Tax Revenue per Capita

Week 6: Mar 8 - Mar 14

- Meeting with Yifu
    - Introductions
    - Confirmed if the variables identified work for the project
    - Things to consider
        - Factors like the economy are very complex and influenced by many factors. It's hard to say if it's because of legalization or other factors.
        - As a data science project, the number of features may be less, only one feature.
        - If you apply the model to a new state that just legalized it, how can you verify whether your model is precise or not?
        - Good direction to start in but dig into the area to find more interesting aspects of this direction to answer some questions.
            - Eg: after the state legalized the drug, which community is affected by the most by this legalization, which race, income level, gender,

etc. has been impacted the most by this change. Demographic analysis of who was affected the most by legalization.

Week 7: Mar 15 - Mar 21

- Found data on the new list of variables we plan to work with
    - GDP per capita
    - Tax Revenue per capita
    - Unemployment Rates
    - Cigarette Sales
    - Consumption of Alcoholic Beverages
    - Admission to Rehabilitation Services
    - Suicide Rates
    - Alcohol-related Driving Fatalities
- Meeting with Yifu
    - Introductions
    - Things to consider
        - What is the plan if you do not find any useful results from this data?

Week 8: Mar 22 - Mar 28

- Meeting with Yifu
    - Things to consider
        - Compare before and after of the variables
        - What about finding the correlation between legalization and drug deaths?
            - Suppose drug deaths increased after legalization, how can we tell if the increase in other states is related to legalization.
        - We need to define a correlation or similarity function - changes of the curve are highly correlated to that. Let the data make the conclusion.
        - Since we changed our proposed questions, address that in the deliverable. Don't change in the proposal. Have another ongoing record of our questions.
- Deliverable 2: https://docs.google.com/document/d/1SN-VcXgGdj7Ge7rE3RWtUwrYpRBylNotOebw9Wl7k7U/edit?usp=sharing
- Ran regression analysis on the multiple variables identified
- Compared how those variables have changed over time in the US vs in Colorado
- Concluded if legalization has been beneficial in Colorado or not
    - But how can this be tied to legalization
    - Try correlation next

Week 9: Mar 29 - Apr 4

- Meeting with Lance
    - Focus on correctly understanding the data from one state
    - Pay more attention to Data Visualization
    - Do not worry about trying to build a model

- Meeting with Yifu
    - Plot long-term trend
    - Predict trend
    - Plot trend after legalization
    - Plot short term trends for ~5 years before legalization and compare to data after legalization
    - Plot actual data points (easier to read)
- Work on Deliverable 3: https://docs.google.com/document/d/1vBT3bt5JTaG2DYfKxBt-Ja-7_pI7DLPsevWPlQ_-mt8/edit?usp=sharing
    - Plot better graphs for the socio-economic variables from the previous deliverable

Week 10: Apr 5 - Apr 11

- Early Insight Presentations: https://docs.google.com/presentation/d/1wtEuSGlU9n4MJUCGdCeWRhyY-4gr8373OuUPvkyCb4Q/edit?usp=sharing
    - Background and Motivation
    - Findings
    - Limitations and challenges
    - Next steps
- Complete Deliverable 3: https://docs.google.com/document/d/1vBT3bt5JTaG2DYfKxBt-Ja-7_pI7DLPsevWPlQ_-mt8/edit?usp=sharing
    - Graphs plotted only for Colorado

Week 11: Apr 12 - Apr 18

- Meeting with Yifu
    - Deliverable 4 looks good
    - Add more information for the final report that might have been missed in Deliverable 4
    - Not building a model for lack of time
- Deliverable 4: https://docs.google.com/document/d/11IcDyQvVc9-Fo7u-uj7ppfoN4-06xfrHqdomLjB6Dbs/edit?usp=sharing
    - Graphs plotted for Colorado
    - Graphs plotted for National Data

TIMELINE GOING FORWARD

Week 12: Apr 19 - Apr 25

- Meeting with Yifu
- Final Project Report

Week 13: Apr 26 - May 2

- Presentations
- Final Project Report