Minimizing Crime Rates Through Budgetary Decisions

Logan DeSmet

Problem: Crime and Budget Decisions

- Thinking that crime is a way of measuring police effectiveness is only addressing a fraction of what really influences crime

> *FBI: "It's important to consider the various factors that lead to crime activity and crime reporting in a community before interpreting the data. Without these considerations the available data can be deceiving. Factors to consider include population size and density, economic conditions, employment rates, prosecutorial, judicial, and correctional policies, administrative and investigative emphases of law enforcement, citizens' attitudes toward crime and policing, and the effective strength of the police force."*

- Crime is influenced by population profiles and statistics
- Population profiles and statistics are influenced by budgetary decisions

- **Goal: Optimizing budget decisions to minimize crime rates**
  - *So future budgetary decisions can be made with data driven insights*

Data Sourcing

Crime data: FBI: Crime Data Explorer

Population Profile Data: ACS Population Profiles (2010-2021, no 20)

Budget Data: 2020 Budget Data - census.gov

How Can This Be Optimized? (Key Assumptions)

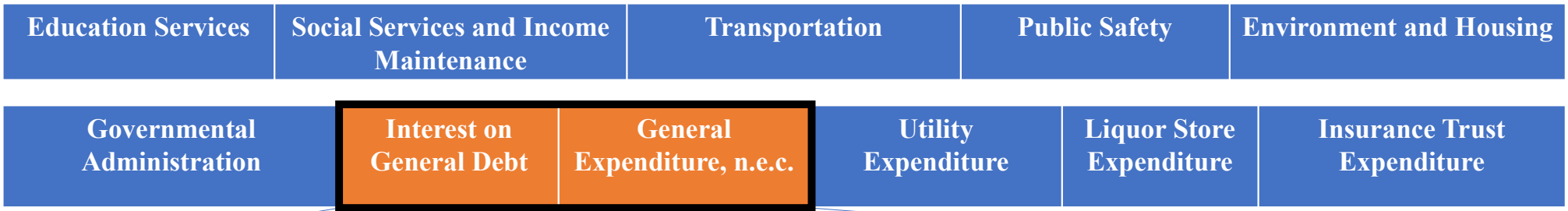
- Budget Data from Census.gov is broken into these official categories:

Education Services	Social Services and Income Maintenance	Transportation	Public Safety	Environment and Housing
Education, Libraries	Public Welfare, Hospitals, Health, Employment Security, Veterans' services	Highways, Air Transportation, Parking Facilities, Sea and Inland Port Facilities	Police Protection, Fire Protection, Correction, Protective inspection and regulation	Natural Resources, Parks and Recreation, Housing and Community development, sewerage, solid waste management

Governmental Administration	Interest on General Debt	General Expenditure, n.e.c.	Utility Expenditure	Liquor Store Expenditure	Insurance Trust Expenditure
Financial Administration, Judicial and Legal, General Public Buildings, Other governmental administration	Interest on General Debt	Miscellaneous Commercial Activities, Other and unallocable	Water supply, electric power, gas supply, transit	Liquor store expenditure	Unemployment Compensation, Employee Retirement, Workers Compensation, Other insurance trust

(i.e. Education Services = Education + Libraries)

How Can This Be Optimized?



- "Interest on General Debt" and "General Expenditure, n.e.c" are not adjustable - that part of the budget is predetermined
- The model combined these into a variable called "Other Expenditures", and must take the value as a constraint

Calculations In the Model

- All the data was combined from 2015-2021 (no 2020)
- Crime data = Violent Crimes + Property Crimes
- Budget data was converted to proportions of total budget
- Crime data was converted to crimes per resident
- Population Profile data was converted to proportions of population

- Key Assumption: State residents act in similar ways (crime) with similar population profile data
  - *Model does not account for the difference culturally*

Calculations in the Model (cont.) - *Python*

- A Neural Network was used to calculate feature importance among the population profile data points to predict crime rate
  - The top 10 most important features were selected
- The selected population profile features were then used in a Linear Regression Model to predict violent crime rate and property crime rate separately
- The budget category features were then used in a Linear Regression Model to predict each of the selected population profile features separately
- The weights (coefficients and intercepts) from each model were then extracted and put into functions to use for the optimization model
  - *With scaling when necessary*

The model performance

- The output of the machine learning model crime rate predictions were impressively extremely accurate
  - All crime rate predictions were within 0.001 of the expected true rate
  - Suggesting that culture does not make too large of an impact

- The optimization model had some key assumptions:
  - Only allowed budgetary decisions within the range of possible values in the dataset → this way the budget was spread around
  - Accepted a minimum crime rate of 0.01 → it would be unreasonable to predict a crime rate of 0
  - The budget total input into the model should not include Direct Expenditures (which was not one of the categories listed on slide 4) → the reason being this funding is split between national and state funding, and does not change much by state
    - *So with total budget of $1million, and direct expenditures of $200k → model input is $800k*

Key Challenges When Optimizing

- Inorder to scale the decision variables and work through the Machine Learning algorithms, the optimization model required some intense coding – something outside of this course
- In Summary: The decision variables needed to be set to 0 at the start of the model, and then optimized with the constraints mentioned on the last slide
  - The model was then solved and optimized using ‘cbc’
  - Then instances of the model were made for each ‘client’ and a constraint was added for said client’s ‘Other Expenditures’ category (Interest on General Debt” + “General Expenditure, n.e.c”)

The (Hypothetical) Client

- The optimization model was used for every state's 2021 budget
  - With each state's corresponding "other expenditures" being considered a constraint
  - 2022 and 2023 are simply not available yet (2022 is available Jan 1)
- Every state was optimized and added to a dictionary
- A user can call the function at the end of the python model specifying the state and get both the state's optimized budgetary proportions and dollar amounts
  - The user simply inputs the datafile where specified and the model does the rest
  - Summarized on the next slide...

Optimized State Budgets (for 2021)

- The full output of optimized budgets is in the optimization model results (python file)
  - The output for 3 states →

- Interesting Finds:
  - Every State maxed out ‘Education’, ‘Public Safety’, and ‘Environment and Housing’
    - The other budget categories all varied by state
    - Suggests the most important things for crime rate
  - Suggests that more funding should be put into these categories beyond what is currently the maximum found in the data

```
-------------
State: Vermont
Expected Crime Rate: 1.0%
Budget:
   Education services: $4557636.9444126
   Social Services and Income Maintenance: $5264046.7617716
   Transportation: $669390.0266754731
   Public Safety: $659789.9549124839
   Environment and Housing: $4177880.5651165214
   Governmental administration: $942148.85997807
   Utility expenditure: $687574.777706425
   Liquor store expenditure: $0.0
   Insurance trust expenditure: $3671667.635874819
   Other expenses: $8321244.038305901
TOTAL: $14099290
-------------
State: Rhode Island
Expected Crime Rate: 1.0%
Budget:
   Education services: $4557636.9444126
   Social Services and Income Maintenance: $5655590.435495
   Transportation: $669390.0266754731
   Public Safety: $659789.9549124839
   Environment and Housing: $4177880.5651165214
   Governmental administration: $942148.85997807
   Utility expenditure: $687574.777706425
   Liquor store expenditure: $0.0
   Insurance trust expenditure: $3671667.635874819
   Other expenses: $7616811.755874244
TOTAL: $18649466
-------------
State: New Mexico
Expected Crime Rate: 1.0%
Budget:
   Education services: $4557636.9444126
   Social Services and Income Maintenance: $6001437.695667921
   Transportation: $1187186.296494951
   Public Safety: $659789.9549124839
   Environment and Housing: $4177880.5651165214
   Governmental administration: $942148.85997807
   Utility expenditure: $687574.777706425
   Liquor store expenditure: $0.0
   Insurance trust expenditure: $3671667.635874819
   Other expenses: $6506712.572911967
TOTAL: $35609923
```

Conclusions and Projected Benefits

- ‘Education’, ‘Public Safety’, and ‘Environment and Housing’ are the most influential factors on expected crime rates
  - Every state maxed out the amount they could allocate to these budgetary categories
    - This suggests that the government currently is spending too little on these categories and more funding should be put in place
      - With less funding limitations the optimization model should be rerun to produce better results
- Of the 10 most important features identified from the Neural Network: ‘health insurance coverage’ and ‘Proportion of Population above 25 with no Highschool degree’ both align with these most influential factors
  - Suggesting that these should be heavily considered when making budgetary decisions

Limitations and Next Steps

- As discussed throughout the presentation, the most important limitation is the lack in the model to identify cultural impact on crime rates – a metric that varies state to state
  - Considering accuracy of the model, it likely minimizes this cultural effect

- Next Steps:
  - Further incorporate model with less maximum budgetary constrictions → allowing for budget to be allocated higher than what we saw in the dataset
    - This requires research on what the proper minimums are for each budget category
  - Add a way for users to highlight and constrain certain population profile metrics, not selected by Neural Network
    - i.e. Higher Education Percentage

Supporting Files

- Transforming the data from the raw data files into data that was applicable for the model took 20+ hours of coding and tableau prep work (*a lot of trial and error*)
  - The supporting folder includes read me files which walk through the steps of converting the data
    - Specifying the appropriate python and tableau prep file for each step