PDS Opioid Project - Backwards Design Template

Matthew Holden, Katie Hucker, Lisa Wang

1 Topic

What problem are you (or your stakeholder) trying to address?

We want to address the increase in use and abuse of prescription opioids, specifically oxycodone and hydrocodone. The rise in opioid addiction from prescriptions led to a rise in both prescription overdose deaths, and from non-prescription opioids like heroin and fentanyl. This problem is mainly focused on the following stakeholders or groups: united states population, government regulators, and doctors/pharmacists.

2 Project Question

What specific question are you seeking to answer with this project?

Does prescription drug policy intervention limit the volume of opioids prescribed and drug overdose deaths?

We will be estimating the effectiveness of policy interventions designed to limit the over-prescription of opioids using the following outcome variables:

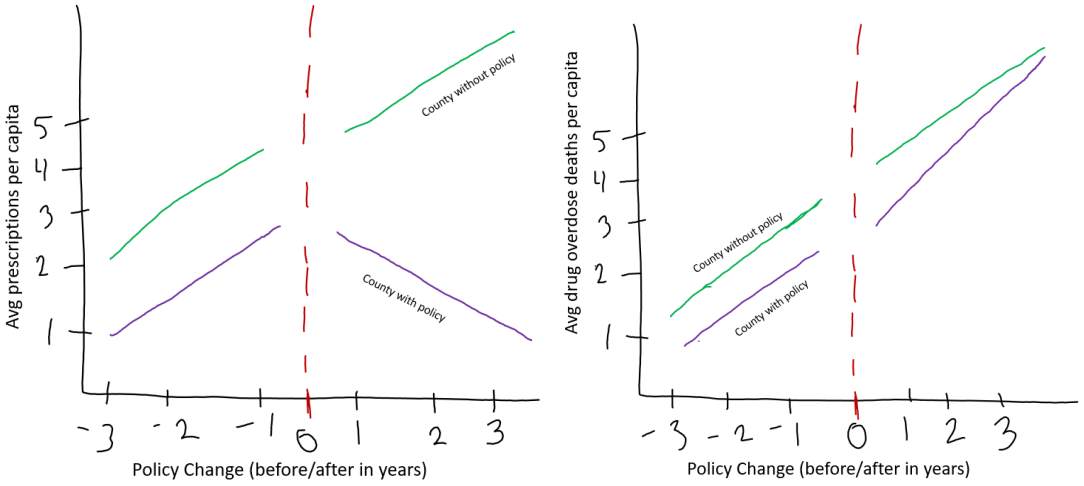
- The volume of opioids prescribed
- Drug overdose deaths

3 Project Hypothesis

What is your hypothesized answer to your question?

An increase in prescription opioid regulation will decrease the volume of opioids prescribed. However, the regulation increase will increase the number of drug overdose deaths. The quantity of prescribed opioids will decrease because this is how trackable opioids are acquired. Therefore, the policies will affect the already controlled method of receiving opioids. The rationale for the second hypothesis is that non-prescription opioids with unregulated dosages and side effects will flood the market to fill the gap left from reduced prescription opioids.

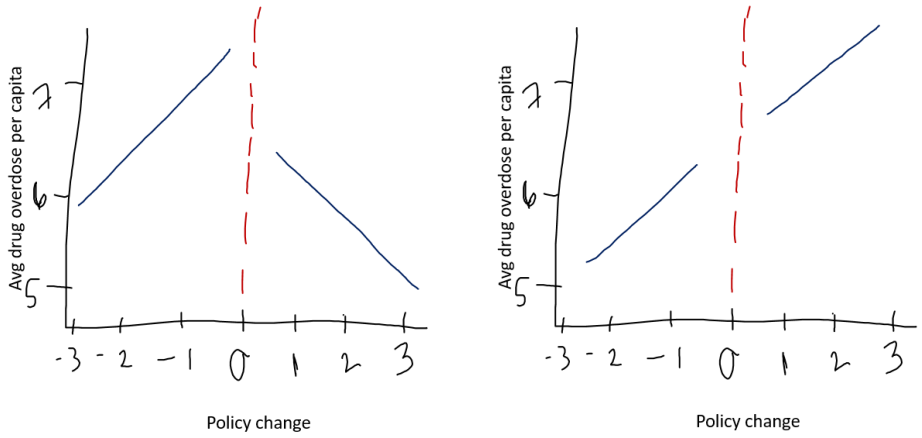
4 Model Results



- Left Figure: "Hypothesis Confirmed – Decreased Opioid Prescriptions"

- Right Figure: "Hypothesis Confirmed – Increase Overdose Deaths"

If the hypothesis is true, the "difference-in-difference" graph will demonstrate a distinct pattern, with a substantial reduction in opioid prescriptions in the treatment group and a noticeable increase in overdose deaths per year. Alternatively, if the hypothesis is false, the plots could show a reverse pattern, with increased opioid prescriptions per year and a reduction in overdose deaths per year. The direction of the lines themselves may vary, but the key factor is the comparison with the trend of the comparison group.



- Left Figure: "Hypothesis Contradicted – Reduced overdose per capita trends"

- Right Figure: "Hypothesis Supported – Consistent overdose per capita trends or reversed patterns"

In the context of the "pre-post analysis," if the hypothesis is true, we expect to observe a notable reduction in opioid prescriptions and a corresponding increase, or no change, in overdose deaths per year following the implementation of the policy. This pattern would indicate that the policy effectively reduced opioid prescriptions but potentially contributed to higher overdose mortality. Conversely, if the

hypothesis is false, the graph would illustrate a different scenario, with a reduction in overdose deaths after the policy implementation. This outcome would suggest that the policy was more effective at addressing opioid overdose deaths than we are hypothesizing.

5 Final Variables Required

- Counties in Florida, Texas, and Washington State. In addition to two other states which no policies are in effect.
- FIPS County codes
- Drug overdose death counts per county in the years 2003-2015
- Total population for each year of each county
- Proportion of overdoses in per county in the U.S.
- Counts of opioid prescriptions
- Proportion of opioid prescriptions

Each row will contain an observation of a county on its month/year time frame. It will have all the variables listed above.

We think there will be 3 separate data frames at one point where we will have to use FIPS county codes to merge the separate data frames. The FIPS county codes will have to be applied on the individual data sets then merged based on the codes since counties can have the same name.

6 Data Sources

We will need several datasets to conduct our analysis on the impact of policy changes on opioid prescriptions and overdose deaths. These datasets should include information on opioid shipments, overdose mortality, and potentially other variables that can help us understand the broader context.

1. **Opioid Shipment Data:** We will primarily rely on the 'Prescription Opioid Shipment Data' obtained from The Washington Post, covering the years 2006 to 2019. This dataset provides information on the volume and distribution of opioid shipments, including details like the drug type, quantity, and distribution locations. We plan to merge this dataset with overdose mortality data to analyze its impact on overdose deaths. *Source: https://www.washingtonpost.com/national/2019/07/18/how-download-use-dea-pain-pills-database/?arc404=true*

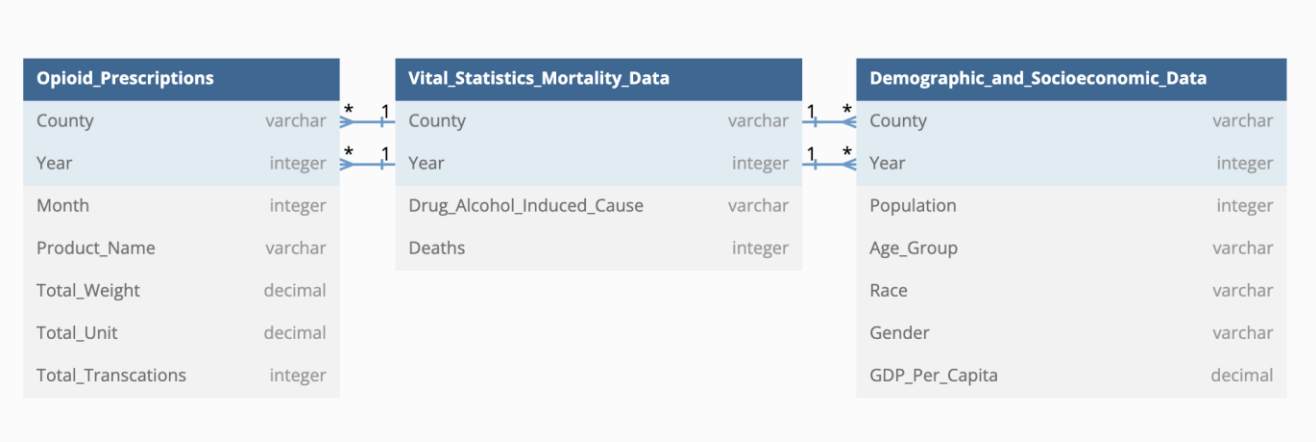
2. **Overdose Mortality Data:** The 'Vital Statistics Mortality Data' for drug and non-drug-related causes for every U.S. county from 2003 to 2015 will be our primary source for overdose mortality statistics. We will aggregate this data to an annual level to match the opioid shipment data. *Source: https://www.dropbox.com/s/kad4dwebr88l3ud/US_VitalStatistics.zip?dl=0*

3. **Demographic and Socioeconomic Data:** To account for potential confounding variables, we may also need demographic and socioeconomic data. These could include variables like county-level income, education, and population density. We will need to find and merge datasets that contain this information to control for these factors. *Source: https://www.census.gov/data/tables/time-series/demo/popest/2010s-counties-total.html#par_textimage*

In terms of relating and merging these datasets, we anticipate linking them based on shared variables such as county, year, and potentially drug type or specific opioid names. The 'county' variable is expected to be the primary key for merging, as it allows us to match overdose mortality data with opioid

shipment data for the same geographic regions. Additionally, we may need to standardize and aggregate the data to ensure a consistent time frame for the analysis.

A Draft for our final cleaned database design



7 Division of Labor

Now that you have identified what needs to be done, how do you plan to divide those tasks among team members?

Here's a plan on how to divide tasks among team members:

1. **Data Gathering:** Assign one team member to be responsible for gathering and downloading the necessary data. This includes the prescription opioid shipment data from The Washington Post and the vital statistics mortality data. This person will be in charge of ensuring that the most recent and accurate data is obtained and maintained.

2. **Data Cleaning and Organization:** As data cleaning is a significant focus, we will distribute this task among team members to handle different aspects of cleaning and organizing the data. For example, one team member can focus on handling missing values, while another can address formatting issues. Ensure that the team members maintain clear documentation of the cleaning process and the decisions made to maintain data integrity.

3. **Data Merging:** Assign a team member to merge the various datasets that we have obtained. This can be a complex task, as it involves matching and combining data from different sources. Ensuring that variables are correctly linked is critical to the analysis.

4. **Analysis - Pre-Post and Difference-in-Difference:** Divide the analysis into two parts, one for the "pre-post analysis" and one for the "difference-in-difference analysis." Assign team members to each part based on their expertise and familiarity with these analysis methods.

5. **Data Presentation and Visualization:** Designate a team member responsible for creating visualizations and graphs that illustrate the results of the analysis. Effective visualization is key to conveying the findings to both your instructor and the imaginary policy maker.

6. **Documentation and Reports:** Allocate the task of preparing the project reports to a team member who can clearly and concisely communicate the project's motivation, data sources, summary statistics, analysis, and interpretation.

7. **Code Review and Version Control:** Every team member should actively utilize GitHub for project management, code review, and to maintain version control. This collaborative approach ensures code quality, identifies errors, and promotes consistency in the project's codebase.

Here are our initial assignments, this may change later for time constraints or help needed:

Tasks	**Initial**	**Reviewer**
Data Gathering	Matt	Katie
Data Cleaning	Each member cleans a different table	Round robin review
Data Merging	Lisa	Matt
Analysis	Katie	Matt
Data Visualizations	Katie	Lisa
Documentation / Reports	All members will report on their section for the report	All members
Version Control Practices	All members	All members