**Research Question**

How have drug overdose death rates evolved across different U.S. states over the period from 2015 to 2023, and what patterns or trends can be observed for state-wise and year-wise variations in overdose deaths?

**Data Sources**

To answer this question, I will use the dataset “VSRR\_Provisional\_Drug\_Overdose\_Death\_Counts.csv,” which contains overdose death counts from 2015 to 2023 across different U.S. states. The dataset was obtained from the Centers for Disease Control and Prevention’s National Vital Statistics System. The dataset contains columns for State, Year, and Data Value, which is the total number of drug overdose deaths reported. Additional geographic data will be loaded using the GeoJSON web file hosted on Plotly’s GitHub repository for state visualizations.

* **Primary Dataset**: VSRR\_Provisional\_Drug\_Overdose\_Death\_Counts.csv

(Source: CDC – <https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm>)

* **Geospatial Data**: GeoJSON file from Plotly for U.S. states (Source: GitHub – <https://raw.githubusercontent.com/plotly/datasets/master/geojson-counties-fips.json>)

**Final Deliverable**

The final deliverable will consist of:

1. **Data Visualizations**: A series of visualizations that illustrate the trends in overdose deaths across the U.S.:
   * Bar Plot showing the total overdose deaths per year (2015-2023).
   * Line Plot depicting the trend of total overdose deaths over time.
   * Box Plot showing the distribution of overdose deaths by state over the years.
   * Choropleth Map displaying overdose deaths across U.S. states by year, to visually represent the geographic distribution and variations.
   * Top States Plot for 2022, which will highlight the states with the highest overdose deaths for that year.
2. **README.md File**: The README.md file will document the entire analysis process, including:

* Project overview and objectives
* Data source description
* Steps for data cleaning and processing
* Description of visualizations created
* Instructions on how to run the analysis using the provided Jupyter notebook
* Summary of key insights from the analysis

1. Requirements.txt File: A requirements.txt file will be provided, listing all the necessary Python libraries to run the analysis. The contents of the file will include:

* pandas
* seaborn
* matplotlib
* plotly
* numpy
* kaleido (for saving Plotly visualizations)

The combination of the visualizations and analysis will provide a comprehensive view of the evolution of drug overdose death rates across the U.S. from 2015 to 2023, allowing for further exploration of the factors contributing to these trends.