Glossary and Data Dictionary

Here is a glossary defining the key data elements (Dimensions, Measures, KPIs, Calculated Fields, Parameters, and Filters) used across the dashboards in this project, explaining what each represents and its role.

1. Dimensions

Dimensions are qualitative values that typically cannot be aggregated in a meaningful way. They are used to categorize, segment, and reveal the granularity of data.

- **State**: The U.S. state where a hospital or drug use incident occurred. Used for geographic mapping and filtering.
- **SETTING**: The patient care environment (e.g., "ED" for Emergency Department, "IP" for Inpatient). Used to categorize drug use by location within a facility.
- **Quarter**: The three-month period within a year (e.g., "2023 - Q1"). Used for temporal trends, derived from START_TIME.
- **INDICATOR**: The specific drug type or category of drug use (e.g., "All Drugs," "Opioids," "Cannabis," "Stimulants"). Used to categorize drug use incidents and as a primary filter.
- **GROUP**: Broad demographic categories (e.g., "Age," "Sex," "Total," "Urban-Rural"). Used to group demographic breakdowns.
- **SUBGROUP**: Specific sub-categories within demographic groups (e.g., "0-15 Years," "Female," "Rural Areas"). Used for granular demographic analysis.
- **City/Town**: The city or town where a hospital is located. Used for geographic granularity on maps and for filtering.
- **Hospital Name**: The name of the healthcare facility. Used for identifying specific hospitals, ranking, and drill-down filtering.
- **Address**: The physical street address of the hospital. Used for detailed hospital information.
- **ZIP Code**: The postal code of the hospital's location. Used for detailed hospital information.
- **Year**: The year of the drug use record or socioeconomic data. Used for filtering and temporal analysis.

2. Measures

Measures are quantitative, numerical values that can be aggregated (summed, averaged, counted, etc.).

- **VALUE**: The core metric representing the volume or count of drug use incidents or specific drug-related events. It is aggregated as SUM(VALUE) across various dimensions.
- **Median Income**: The median household income for a given state and year. Used as a socioeconomic indicator.

- **Population**: The population of a given state for a specific year. Used as a socioeconomic indicator and to size points in scatter plots.
- **Unemployment Rate**: The unemployment rate for a given state and year. Used as a socioeconomic indicator.
- **Poverty Rate**: The poverty rate for a given state and year. Used as a socioeconomic indicator.
- **Bachelors Degree Or Higher (%)**: The percentage of the population with a bachelor's degree or higher for a given state and year. Used as a socioeconomic indicator.

3. Key Performance Indicators (KPIs)

KPIs are specific measures chosen to reflect the critical success factors of a business or project. In this project, they are displayed as single, aggregated numerical values.

- **No. of Hospitals Reporting**: The total count of unique hospitals that submitted data.
- **Avg Duration of Stay**: The average length of time patients stayed at a hospital for drug-related incidents.
- **Total Drug Use Volume**: The grand total sum of VALUE across the entire dataset, representing overall drug usage.

4. Calculated Fields

Calculated fields are new fields created using formulas based on existing dimensions, measures, or parameters.

- **Drug Use Rate per Capita**: A calculated ratio (e.g., SUM(VALUE) / SUM(Population)) used to normalize drug use volume by population size, providing a more comparable measure across different regions.
- **Avg. Dynamic Socioeconomic X-Axis**: This is a placeholder for the average of the socioeconomic measure currently selected by the Select X-Axis Measure parameter (e.g., AVG(Median Income), AVG(Unemployment Rate)). It drives the X-axis of the scatter plot and lines in the socioeconomic trend chart.
- **State Rank**: A ranking calculated using RANK_UNIQUE(SUM(VALUE), 'desc') to assign a unique rank to each state based on its total drug use value, in descending order. Used for the Top N filtering.

5. Parameters

Parameters are dynamic values that can be controlled by the user. They allow for interactive changes in calculations, filters, or visualization elements.

- **Select X-Axis Measure**: A string parameter that allows users to select which socioeconomic measure (Median Income, Unemployment Rate, Bachelors Degree Or Higher, Poverty Rate, Population) will be used on the X-axis of the scatter plot and in the socioeconomic trends chart.

- **Top N States**: An integer parameter that allows users to dynamically select the number (N) of top states to display (e.g., 5, 10, 15, 20), thereby filtering relevant visualizations.

6. Filters

Filters are used to restrict the data shown in a visualization or dashboard based on specific criteria. They can be applied globally (to multiple sheets) or locally (to a single sheet).

- **Year**: A global filter allowing users to select one or more years for analysis, affecting relevant visualizations across multiple dashboards.
- **State**: A global filter allowing users to select one or more states for analysis, affecting relevant visualizations across multiple dashboards.
- **Setting**: A global filter allowing users to select specific patient settings (e.g., "ED," "IP") for analysis.
- **Top N States by Drug Use Set**: This is a set, driven by the Top N States parameter, that acts as a filter. It identifies and includes only the states that fall within the top N based on SUM(VALUE). When applied as a filter, it restricts the data displayed to only those top states.
- **Indicator**: Used as a filter, often driven by a dashboard action, to narrow down analysis to specific drug types.