

## Task.M2.1:

### Data Dictionary

Variable Name	Datatype
County Code/FIPS	Integer
State	Char
Population	Integer
Opioid Dispensing rate	Float
Crude Rate	Float
Drug/Alcohol induced cause	object
Deaths	Int
Premature death	Float
Male/female Population(Age)	Int
Drug/Alcohol induced cause code	Int

**County Code/FIPS:** It is important to include in data dictionary because it is there in every dataset. It will be easy for us to merge on this common column.

**State:** It is important to take state to learn how the opioid pandemic is affecting over state. It will be useful while we want data to be sorted by state. Further it can be easily visualized which state is having highest rate of opioid epidemic in USA.

**Population:** Population is an important factor while dealing with this dataset. Because it will be used for statistical analysis how much population is there in that county or state. This will lead to easy calculation of calculating death rate that is number of deaths divided by population.

**Opioid Dispensing rate:** This will give us the Opioid drug dispensed in that county or a state. It will also give us how many people are addicted if the dispensing rate is more.

**Crude rate:** Crude death rate is defined by total number of deaths in each time interval. This will give us death rate in the particular time interval.

**Drug/Alcohol induced cause:** Drug/Alcohol induced cause will give us it was a suicidal condition, or it was due to overdose of drug.

**Premature death:** This will determine the premature death rate. We can know the number of age group in which the deaths are less or more.

**Deaths:** Death's column will give us number of deaths in the county and state.

**Male/female Population(Range of Age):** It will be easy to know with the age data range how age group is connected to mortality rate.

**Drug/Alcohol induced cause code:** This will give us the code for the particular cause it can be due to suicide or drug overdose or so on.