**[Deliverable #1 - Project Proposal](https://blackboard.umbc.edu/webapps/assignment/uploadAssignment?content_id=_3970468_1&course_id=_52580_1&group_id=&mode=view)**

**This is an individual project. Find an issue that you are passionate about or are interested in as a topic for this exploratory data analysis project.**

**Make sure there are publicly available data about this issue to perform exploratory analysis.**

**The analysis should include descriptive statistics and data visualization, and optionally statistical inference.**

**Write and submit a research proposal in MS Word format that answers the following questions:**

1. **What is your issue of interest?**
2. **Why is this issue important?**
3. **What questions do you try to answer?**
4. **Where do you get the data to help answer your questions?**
5. **What will be your unit of analysis (for example, patient, organization, or country)**
6. **What variables/measures will you use in your analysis?**
7. **What kinds of techniques you will be using (for example,****summary statistics, scatter plot, bar chart, chi-squared test)?**

Individual Research Proposal

**1. Topic of interest: Healthcare**; Opioid Epidemic Analysis by US States.

**2. Importance:** The abuse and misuse of addictive substances are more common than ever in the world today. This probably has to do with the ease of acquisition, increase in the rate of school dropout, poor life styles, unemployment rates and other social factors.

**3. The opioid crisis**. According to the CDC, approximately 115 Americans are dying each day from opioid-related overdoses. Through integrated models of care, healthcare organizations are in a unique position to identify and treat individuals who may be suffering from a substance misuse disorder.

4.**Data Source:** factfinder.census.gov, https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas, <https://catalog.data.gov/dataset?tags=unemployment-rate>, <https://wonder.cdc.gov/>, <https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html#par_textimage_70769902>, <https://collegescorecard.ed.gov/data/>

5.Unit of analysis will by States in the US.

6. **Variables considered:** ID, State, County, Population, Land Area Type, County by Type, Total Count, Non-US Born, Education Type, Education Count, Commute Type, Commute Count, Income, GDP Type, GDP, Unemployment, High School Graduation Rate, College Graduation Rate.

7. Techniques to be used:summary statistics, scatter plot, bar chart & chi-squared test