CS 2316 Final Project

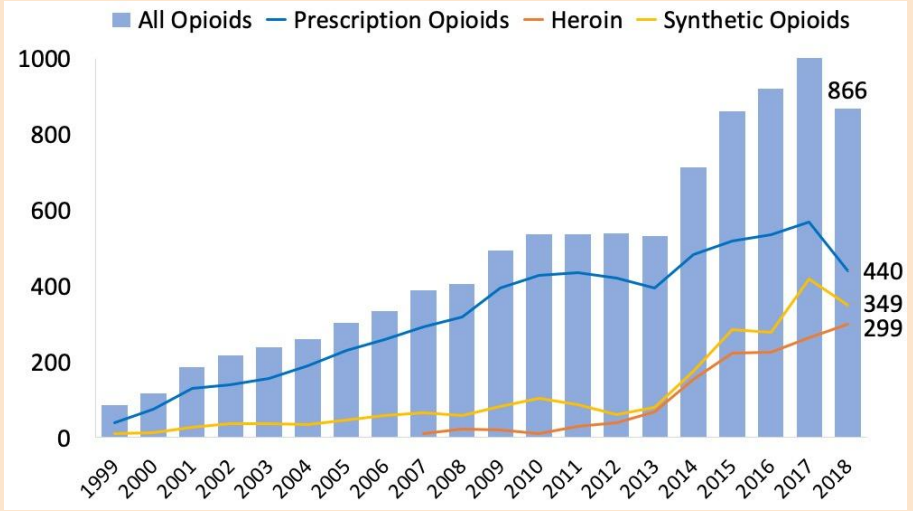
Georgia Opioid Epidemic

Anna Whitford & Neo Pak

Point of Interest and Purpose

- Georgia is in the middle of an **Opioid Epidemic**
- "In Georgia, over **60% of drug overdose deaths involved opioids** with 866 fatalities (a rate of 8.3) reported in 2018."
- **Purpose:** To pinpoint what factors (ex: unemployment rate) may help indicate a county with an opioid issue.

Number of drug and opioid-involved overdose deaths in Georgia



Source: National Institute on Drug Abuse

Point of Interest and Purpose

- We selected three sets of factors to consider when predicting opioid prescription rates:
  1. **Education Level** (how educated a county is)
     a. Percentage of people with less than 9th grade education
     b. Percentage of people with high school education
     c. Percentage with Bachelor's Degrees
  2. **Economic Factors** (how well is a county doing economically)
     a. Income-related factors (ex: Per Capita Net Income)
     b. Other factors (ex: Per Capita Unemployment Insurance Compensation, Per Capita Retirement)
  3. **Labor Statistics** (what percentage of people are working and not working)
     a. Unemployment Rate
     b. Employment Rate
     c. Not in Labor Force Rate

Data Collection Process

1. Downloaded Dataset: Source
   a. Economic data by county, for all counties in US from 1970 - current
   b. Used to gain insight on relationship between economic factors and the opioid prescription rate
2. Web Collection #1: Source
   a. Unemployment data by county in Georgia
   b. Used to gain insight about unemployment by county in Georgia
3. Web Collection #2 Source: Source (API Endpoint)
   a. Opioid data by provider (doctor that wrote prescription). Includes information about city of origin, medication, and number of prescriptions written
   b. Used to gain insight about the number of prescriptions written per county
4. Additional Dataset (CSV): Source
   a. Educational attainment data by county
   b. Used to gain insight about how educated the population in each county is
5. Additional Dataset (HTML, static file): Source
   a. Used for converting cities to counties
6. Additional Dataset (CSV, Added in Phase II): Source
   a. Used for converting counties to fips codes

Data Cleaning Process

1. All data needed to be in terms of county
    a. Opioid claims were aggregated into cities, then cities were converted to counties
        i. Many counties were not named correctly. Capitalization varied by dataset, and many datasets included GA after the county name.
    b. Extraneous data (states other than GA, years that were unnecessary) was removed
2. All data was saved as json files
3. For visualization, data was re-imported, converted to pandas dataframes and dictionaries, and manipulated for insights/visualizations
    a. 2-row 159-column pandas dataframe with counties as indexes
    b. Counties with no prescriptions were removed

Data Analysis, Insights, & Visuals

Insight 1: What is the relationship between the opioid prescription rates and the educational attainment in Georgia (less than 9th grade, high school graduate, bachelor's degree percentage)?

Output:

```
Coefficient of Determination (x-axis: Less than 9th Grade Percentage, y-axis: Opioid Perscription Rate): 0.03336514847139982
Coefficient of Determination (x-axis: High School Graduate Percentage, y-axis: Opioid Perscription Rate): 0.04817887735844517
Coefficient of Determination (x-axis: Percentage with Bachelor's Degree, y-axis: Opioid Perscription Rate): 0.0765730963631368
```

Data Analysis, Insights, & Visuals

Insight 2: What is the relationship between the opioid prescription rates and the unemployment, employed, and not in labor force percentage in Georgia?

Output:

```
Coefficient of Determination (x-axis: Unemployment Rate, y-axis: Opioid Perscription Rate): 0.020513816801276064
Coefficient of Determination (x-axis: Not in Labor Force Percentage, y-axis: Opioid Perscription Rate): 0.1086445214982632
Coefficient of Determination (x-axis: Employed Percentage, y-axis: Opioid Perscription Rate): 0.12193428268656659
```

Data Analysis, Insights, & Visuals

Insight 3: What is the maximum, minimum, and average opioid prescription rate (# of Prescriptions Per Person) out of all the counties of Georgia?

Output:

	County	# of Prescriptions Per Person
Maximum	Schley	11.9545
Minimum	Columbia	0.000084
Average	N/A	0.800909

Data Analysis, Insights, & Visuals

Insight 4: What is the relationship between opioid prescription rates and economic variables such as unemployment insurance, income maintenance benefits, and supplements to salaries?

Output:

```
Coefficient of Determination between Per Capita Personal Current Transfer Receipts (Dollars) and Opioid Perscription Rate: 0.012255768469679573
Coefficient of Determination between Per Capita Income Maintenance Benefits (Dollars) and Opioid Perscription Rate: 0.02975089490958338
Coefficient of Determination between Per Capita Unemployment Insurance Compensation (Dollars) and Opioid Perscription Rate: 0.0026756756187700237
Coefficient of Determination between Per Capita Retirement And Other (Dollars) and Opioid Perscription Rate: 0.007924394799971979
```

Data Analysis, Insights, & Visuals

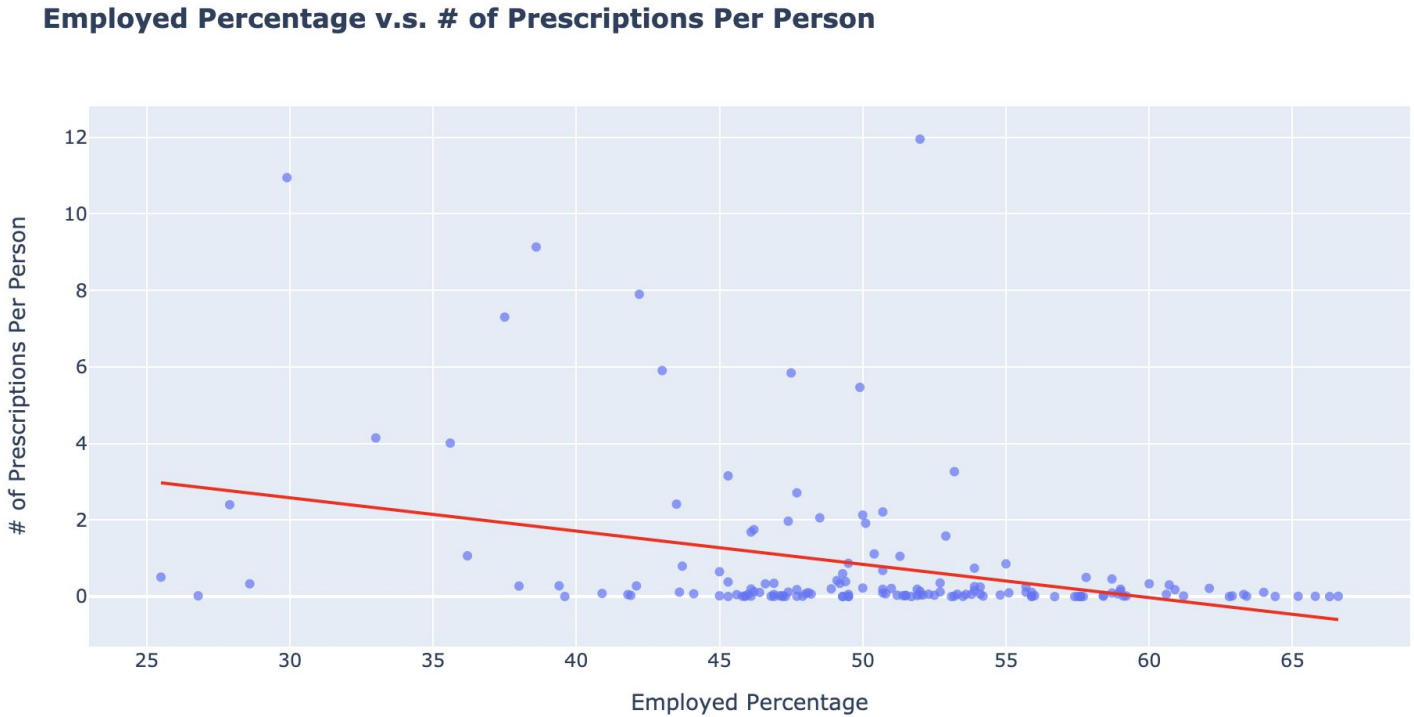
Insight 5: What is the relationship between opioid prescription rates and economic variables such as unemployment insurance, income maintenance benefits, and supplements to salaries?

Output:

```
Coefficient of Determination between Per Capita Net Earnings (Dollars) and Opioid Perscription Rate: 0.05576972342873121
Coefficient of Determination between Net Earnings By Place Of Residence (Thousands Of Dollars) and Opioid Perscription Rate: 0.012411240840660143
Coefficient of Determination between Per Capita Personal Income (Dollars) and Opioid Perscription Rate: 0.05934571554798629
```

Data Analysis, Insights, & Visuals

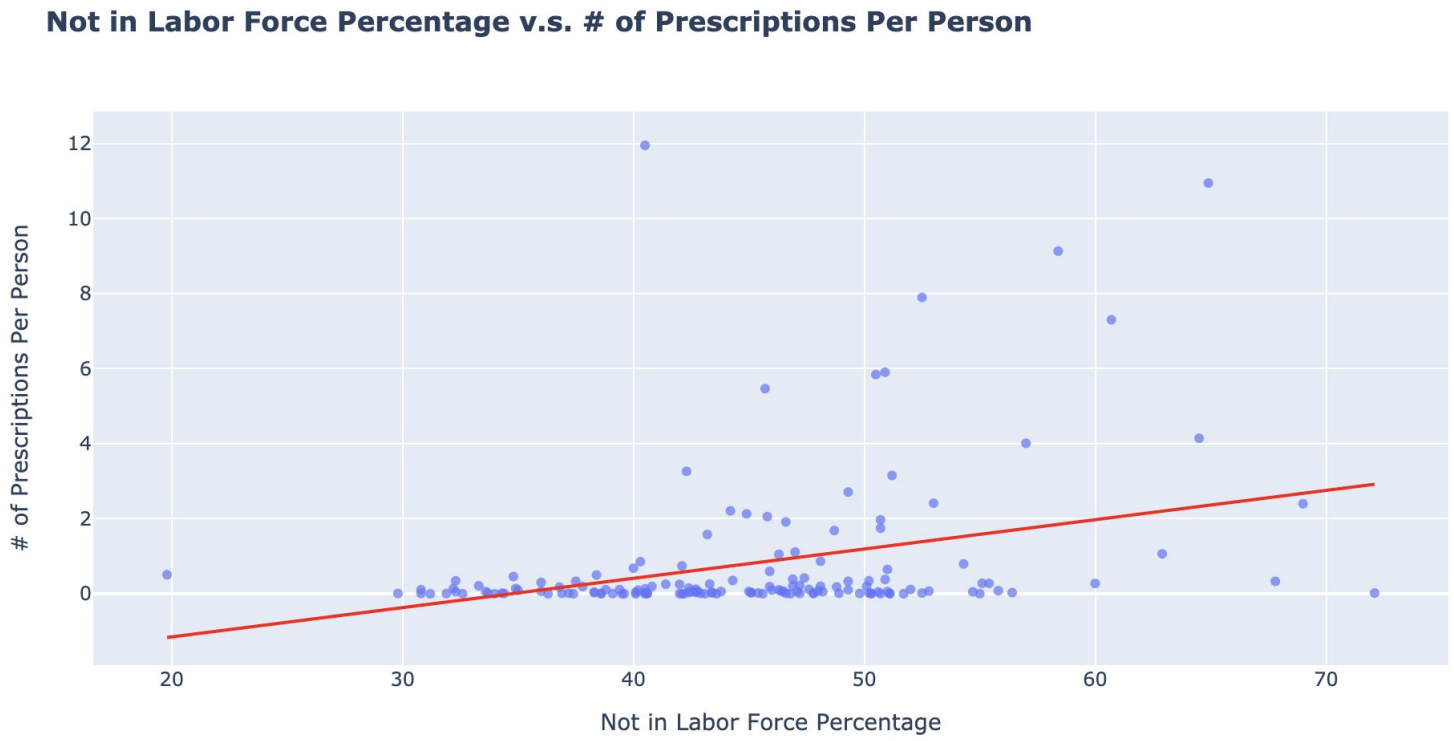
Data Visualization #1



Coefficient of Determination:
0.12193428268656659

Data Analysis, Insights, & Visuals

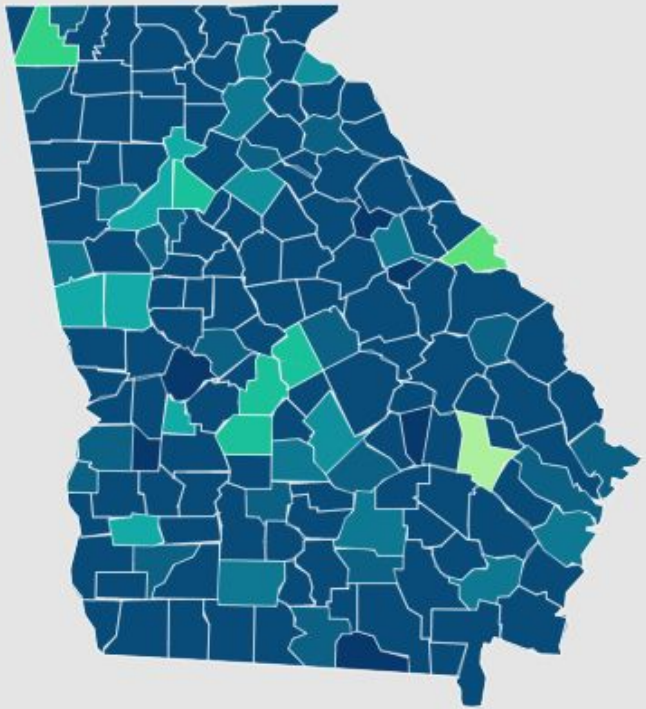
Data Visualization #2



Coefficient of Determination:
0.1086445214982632

Data Analysis, Insights, & Visuals

Data Visualization #3



Overall Results

Factor Compared	Coefficient of Determination
Unemployment Rate	0.020513816801276064
Not in Labor Force Percentage	**0.1086445214982632**
Employed Percentage	**0.12193428268656659**
Less than 9th Grade Percentage	0.03336514847139982
High School Graduate Percentage	0.04817887735844517
Percentage with Bachelor's Degree	0.0765730963631368
Per Capita Personal Income (Dollars)	0.05934571554798629

Factor Compared	Coefficient of Determination
Per Capita Personal Current Transfer Receipts (Dollars)	0.012255768469679573
Per Capita Income Maintenance Benefits (Dollars)	0.02975089490958338
Per Capita Unemployment Insurance Compensation (Dollars)	0.0026756756187700237
Per Capita Retirement And Other (Dollars)	0.007924394799971979
Per Capita Net Earnings (Dollars)	0.05576972342873121
Net Earnings By Place Of Residence (Thousands Of Dollars)	0.012411240840660143

Conclusion

- Relatively, the *"Not in Labor Force Percentage"* and the *"Employed Percentage"* were the top 2 factors that correlated to "*Opioid Prescription Rate"*
- However, the *Coefficient of Determination* is low and negligible in the larger scheme
- Research online states that less education, increasing economic challenges, and low social status correlate with higher opioid addiction rates
- Combination of many factors lead to addiction rather than one
- Further analysis:
  - Look at only counties with above-average claim rates
  - Multivariate analysis

Sources

1. https://www.drugabuse.gov/drug-topics/opioids/opioid-summaries-by-state/georgia-opioid-involved-deaths-related-harms
2. https://www.drugabuse.gov/about-nida/noras-blog/2017/10/addressing-opioid-crisis-means-confronting-socioeconomic-disparities