What question is the student trying to answer? Is this question clear? (14 pts)

**Define a problem of interest** (+2).

We believe demographic features such as poverty status, whether a person is middle aged, education level, employment status, and household income may be associated with higher incidences of opioid use in census tracts within Tempe. In our study, we extract these demographic features from the United States Census’ American Community Survey five-year (ACS5) estimates and compare them to daily measures of the population normalized mass load (PNML) of opiates in wastewater to determine if this relationship exists.

This project will study the concentration of opiates in urban watersheds to estimate opioid usage patterns in cities. We intend to measure the relationship between estimated opiate use and access to care within certain communities.

The individuals who will care about this project are city policy makers, city health officials, and emergency medical services. This project can help better allocate resources to areas that may not have been identified through common opioid overdose measure. We believe that a geographic analysis and mapping of this public health problem provides an easily interpretable framework for non-technical audiences to better identify opportunities to improve treatment and identify risk factors.

**Pick a question you would like to address relevant to this problem** (+2)

What relationships are present between opioid usage, access to treatment, and other demographic or geographic features? Can the presence of these relationships be used to form a testable hypothesis that can improve public health outcomes?

What we can do is to figure out the communities that are most affected and the traits that they have in common, which could be the basis for future research.

BLUF: Specify what, specifically, are the constructs that you are studying (+2)

The point of our analysis will be to detect the presence of certain relationships between opioid use, treatment levels, and certain community level features.

BLUF: **Specify how you will measure what you are studying** (+2)

BLUF: **Specify the logic relating the things that you are measuring**. (+2)

BLUF: **Specify the algorithm or technique that you will use to assess the strength of the relationship between these things**. (+2)

BLUF: **Tell us why it would matter if that relationship existed.** (+2)

if it is shown that there is a negative relationship between treatment and opioid usage, we could test the effectiveness of a public health intervention on reducing drug use in certain areas. If this is effective, then we can establish that the same type of intervention should be leveraged in other areas to improve health outcomes. If the intervention is ineffective, we can test for other confounding relationships amongst features in the community that may be contributing to the opioid problem.

Given their datasource, what is the student trying to measure?

Are they measuring what they think they’re measuring? (15 pts)

**What are the key constructs in your theory?** (+2)

* Opioid usage - because we cannot directly measure each person’s opioid usage, especially if they don’t report or go to the hospital, you have no way to know if someone uses opioids or not, if they do, how much.
* Low income - because we can not directly measure what exactly is considered “low” for income, and it might be subjective because everyone’s opinion about what low income should be could vary.
* Middle age - because we don’t exactly know what exactly is considered middle age especially when it could vary from region to region, from some demographics to other demographics and etc.

**What measures will you use to operationalize your constructs?** (+2)

**Are these measures nominal, ordinal, interval, or ratio? Why?** (+1)

Based on the construct of opioid usage, we want the features PNML and Number of Opioid Related EMS Calls to be ratio because they both have an absolute zero.

And for the features we want to extract from census data:

* We want the percentage of the population whose incomes are below the federal poverty level to be a ratio because the percentage of the population whose incomes are below the federal poverty level does have an absolute zero.
* We want the percentage of the population whose age is within middle age range to be ratio also because the percentage of the population whose age is within middle age range does have an absolute zero.

For the features that are not associated to a construct:

* We are using the features Site code, Geometry (Geographic Polygon), and Geometry (Geographic Point) as geographical data boundaries, therefore we are declaring them as nominal.
* We are not focused on a specific type of opioid so we did not create a construct for it. Therefore, for the sake of our project we want the feature Chemical name to be nominal.
* Since we are assuming that time is independant, we do not have a construct based on time. Therefore, to help control the time variance we are declaring the features Sample\_date and Call Date as nominal[[1]](#footnote-0).

**Justify your selection of these measures in terms of their validity – what type of validity (face, content, predictive, etc.) can you demonstrate?** (+2)

The PNML estimates of opioid use are a face valid measure. It makes sense that these chemical compounds should be detectable in wastewater byproducts. The PNML estimates could also be considered content valid if considering some other subjective criteria. For example, the PNML measurement is almost impossible for individuals to fake, and represents an anonymized view of the average drug use in a concentration. This means it is likely to be unbiased.

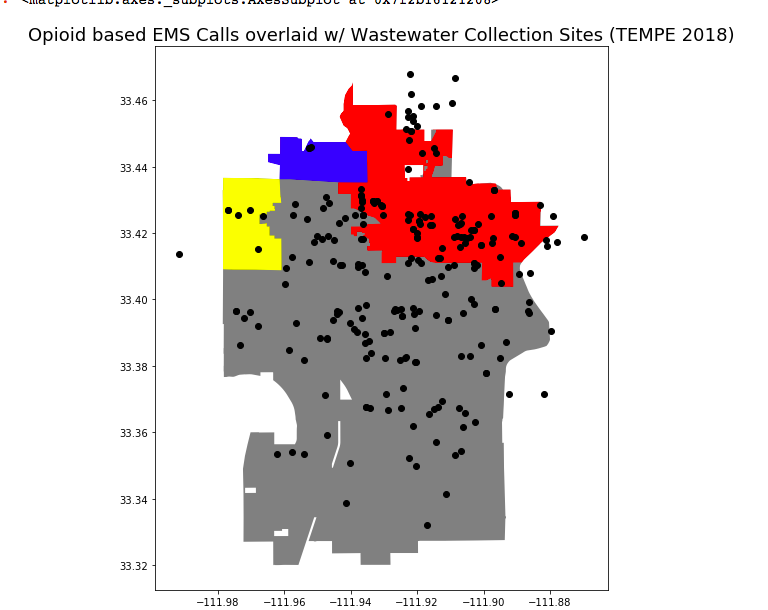
Our alternative measure, EMS calls, are also valid. It makes sense that one way to detect Opioid usage would be from overdoses reported by health care professionals. Content validity may be harder to establish in this case but could also be possible. EMS data is collected by healthcare professionals as they respond to calls for help, and provides a baseline for detection. However, determination about opioid use could be hard to measure in an emergency situation, and also represents a detection problem. Only those who actually call EMS are counted towards the total estimate of opioid use.

Because of these factors, we think that the PNML wastewater estimate may be a better, unbiased measure of the opioid use construct. To test this, we examine the convergent-discriminant validity of the two features below.

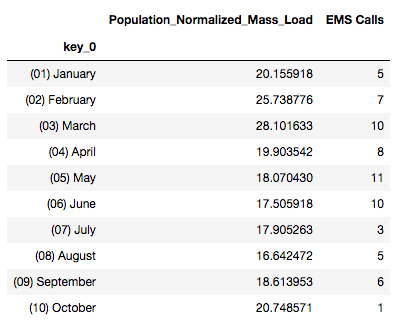
**How will you demonstrate the convergent and discriminant validity of your measures?** (+2)

**Demonstrate the validity of your measures on pilot data?** (+2)

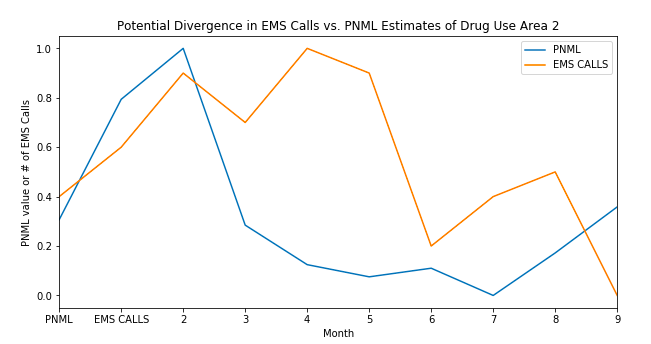
We tested the convergent-discriminant relationship between the PNML and the EMS calls in our pilot data by comparing the measure, correlation, and trend of each variable to one another to look for consistency. First, we restricted our analysis to only calls where opioid was detected in 2018 and within our site collection areas. Because of the low number of EMS calls in areas 4 and 5 in our pilot data, we focused only on area 2 for this preliminary analysis. A map showing the EMS calls in our collection areas is shown below. Collection area 2 is highlighted in red.



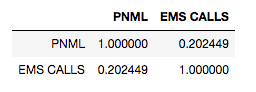
Then, we compared the level of EMS estimated opioid use to the PNML estimated opioid use in area 2 each month. The count of opioid related EMS calls in a given month represents that feature’s ability to measure the level of drug use. Because PNML represents a concentration of drug use per 1000 and is measured daily, we used the monthly average as our measure of detected opioid use.



Examining the relationship between the two measures of opioid use shows that they are somewhat divergent. They are only lightly correlated as shown in the chart below, and only loosely follow the same trend over time. This suggests that the ways in which they detect opioid use are different. Because we think that there is bias in the EMS detection of opioid use, this gives a foundation to test our hypothesis.



**Correlation of PNML vs EMS Calls**



**How will you demonstrate the reliability of your measures?** (+2)

For the PNML data, the researchers include a quantitative measure of the detected measurement error of each sample they collect. We assume that this a reliable measure collected by experts in their field that shows the relative margin of error of each load in the waste water. At this point in time, we are currently evaluating how to translate this chemical deviation into a statistical one, such as a standard error or confidence interval.

The measurement error of the EMS call data is not easily quantifiable because it is only reported by the emergency responders. A potential way to measure error in this case could be to compare the number of opioid related EMS calls to the number of reported overdoses within a collection area and calculate a mean squared error. However, this could be hard to implement because of the likely divergence between these two measures. Some individuals may overdose in hospitals without ever having contacted EMS.

For all demographic features extracted from the census, we can leverage the sampling weights and collection methods of the ACS that are published by the Census Bureau to get an estimated variance for each measure. This is likely a worst case scenario as well, as the census publishes the estimated margin of error for these estimates for a large number of features.

**Demonstrate your approach to reliability on pilot data** (+2)

We think for construct of opioid use we’ll use internal consistency because both features/measurements for opioid use is ratio, which means that they are continuous variable, and we should use correlation for continuous variables. We will assign each area as the item and within each item we will look at the measurement of opioid use from PNML and EMS Call center Opioid report and see how they correlate.

For construct of middle age, there really is not another way of measuring age, so we think we don’t need to have multiple measures for construct of middle age.

For construct of low income, we’ll use internal consistency as well because both features/measures are ratio. Now we have the federal poverty level and we want to categorize households whose incomes are below the federal poverty level as low income households, however, we can also categorize low income household in another way, say, the household whose incomes are within the lowest 10 percentile of incomes in Tempe, AZ as low income households. In this case, we will assign each area as the item, and look at the percentage of the household whose incomes are below the federal poverty level and the percentage of the household whose incomes are within the lowest 10 percentile of all incomes in Tempe and see how they correlate

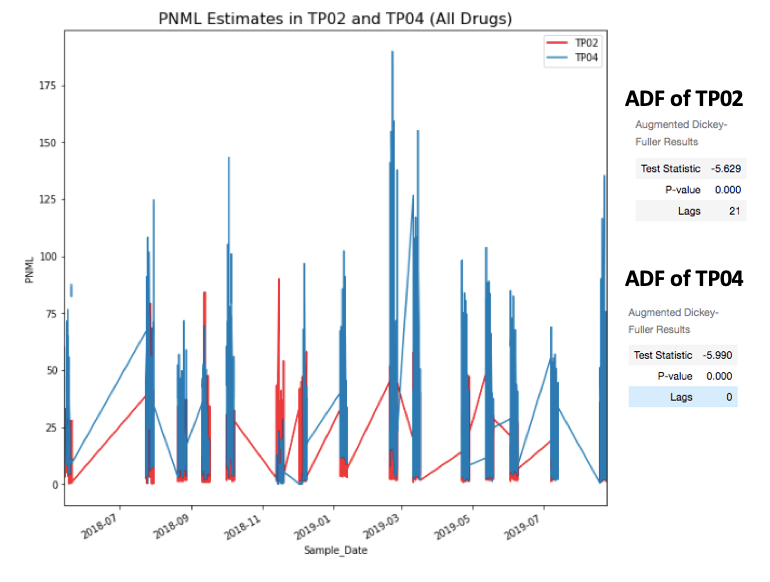
**How will the students know that their answer to their question is true?**

**Have they ruled out alternative explanations?** (16 pts)

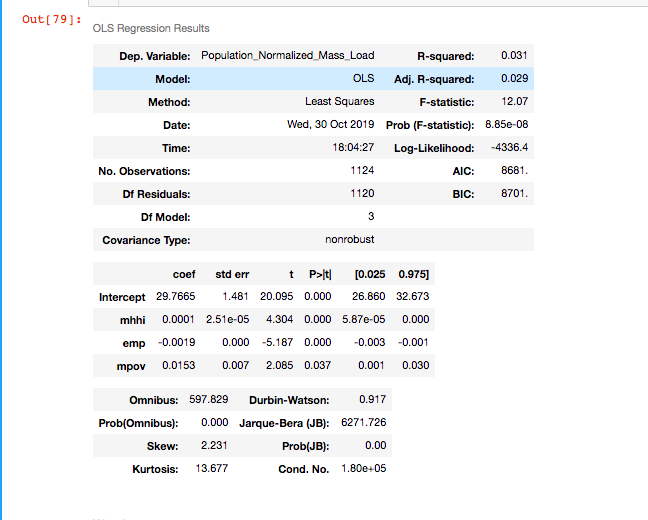
**Describe in detail the causal theory underlying your analysis. What are the proposed causes and what are the proposed effects?** (+2)

**Show pilot data demonstrating some statistical relationship, including temporal precedence and covariance between your proposed cause and effect**. (+2)

Beyond daily volatility, we observe a near constant mean in the PNML estimates across the two years of measurement we have in our sample. The ACS5 variables are held constant over this period, so we are assuming that temporal precedence is not an applicable factor in this analysis. As such, we expect relationships between predictive demographics (or other predictors that are shown to be time-constant over this period) to have a consistently higher correlation and/or impact on drug use compared to other factors. We do observe some cyclicality in the PNML that could be worth investigating further in the future. However, we find that the entire series is time stationary.

****

In terms of an actual statistical relationship goes, we find that the hypothesis we originally proposed does not have an impact on the PNML measures of Opioid uses in our data. This is highly concerning, but after a further view of the literature not unexpected. Researchers have generally struggled to find consistent demographic correlates with opioid use due to the various measurement confounds discussed in the previous sections.



The Arizona Department of Public Health has published some preliminary findings suggesting that features such as gender, mental health factors or a history of substance abuse may have significant relationships with opioid usage. We have requested access to a 2018 release by the Arizona Public Department of Public Health that contains some of these statistics. We are also in the process of identifying mental health measures for our analysis but are having difficulty finding these measures at the city or census tract level.

We have also obtained geo-spatial point estimates of several different behavioral treatment and health care centers within tempe that we plan to include as spatial regressors (i.e. number of X centers within Y miles of the tract centroid) in our analysis. We expect these to have a negative impact on opiate use.

**What are the threats to the internal validity of your analysis?** (+2)

Instrumentation Confound:

The instrumentation confound is a major threat to our analysis because we do not know the detection threshold for the tools that the researchers are using to get the PNML estimates. In other words, PNML estimates may miss low quantities or certain types of drug use based on the way they are measured. Because we do not directly control the source data and measurements there is an inherent and consistent instrumentation confound in our analysis.

Multi-group threats:

We are not assigning people to the groups/areas, and Tempe is relatively small as a city which also has a major problem of opioid abuse, in this case we might consider the population in Tempe normally and randomly distributed among the three areas that we are studying and the multi-group threats should have little impact

Selection-X threats:

We do not anticipate selection-history, selection-maturation, selection testing, instrumentation, or regression threats in our study. As discussed previously, the time-constant mean of the PNML measures suggests that each collection area does not differ much between the high frequency observations in terms of history or maturation. We do not face a selection-testing threat because we are not introducing a specific intervention into the dataset. Likewise, we do not face a selection-instrumentation threat because all measures in our dataset are affected by the instrumentation confound into the dataset.

The only potential selection-x threat we could face is selection-mortality. Subjects in high opioid use collection areas for example may have a greater incentive to move out of the sampling frame during our study than subjects in areas that do not face these issues.

The few threats that our analysis might be susceptible to are:

Instrumental Confound, Experimental Mortality and Selection-X Threats(specifically Selection-Mortality threat)

**Which of these threats are plausible?** (+2)

The instrumentation confound is a plausible threat to our analysis that we are unable to eliminate. Because our study is solely based on the measurements gathered by the ASU researchers and their methodology is not published, we are unable to state the degree of bias we face from the instrumentation confound. Therefore, we have to accept this risk when conducting our study. This instrumentation confound is at least constant across all PNML estimates, and would similarly be present in other opioid measures such as emergency services calls. As such, we consider this risk acceptance to be appropriate because we have no other options for a cleaner measurement.

While we do think that the experimental mortality and the selection-Mortality threat could be plausible, we do not expect it to have a significant impact on our study. Unless there is a major population shift from one area to another or a major health event (i.e. overdose spike) we do not expect our sample to change much during our study. We consider this particularly unlikely because we are looking at only a few years of data where we do not expect sample characteristics to change much absent a significant event.

As for Selection-X threats as a whole, the areas are closely located and we can expect some spillover between and among areas, in which case the groups/areas will not have significantly different characteristics themselves, in which case, the difference in opioid concentration in different areas is not likely caused by the characteristics internal of the groups. We plan to address these effects by including spatial regressors in our analysis.

**Choose an experimental design to rule out plausible threats, assuming you can manipulate all units of analysis and conduct random assignment. Your design should allow you to rule out all plausible confounds.** (+2)

In a perfect Tempe World, we would have all of the PNML for every tract. Then we would randomly assign which tract we would measure opioid usage based on PNML and randomly assign which tract we would measure opioid usage based on EMS Opioids Calls.

The experimental design that could be done in a Perfect Tempe World where we could conduct random assignment for which census tracts are being measured with PNML and by EMS opioid call would be a Posttest-Only Control Group Design. The setup for this experimental design would be:

|  |  |  |
| --- | --- | --- |
| Randomly Assigned Tracts (R) | Opioid Usage Data from Tracts Measured by PNML (New Opioid Measurements) Being Published (X) | Observe Community and City Policy Effects (O) |
| Randomly Assigned Tracts (R) | Opioid Usage Data from Tracts Measured by EMS Opioid Call Data Being Published (Control because they have been publishing this data longer) | Observe Community and City Policy Effects (O) |

We can have more proxies for our opioid use construct to rule out potential instrumental confound, and if we see the same or similar trend over time in the same area and across areas we could potentially mitigate the instrumental confound that we have, but we do not have an efficient way to completely eliminate instrumental confound.

**Acknowledging any manipulations that are not feasible, which threats to internal validity are you unable to address? (+2)**

Because we are not randomly assigning a subsample to a specific intervention, any association we observe is assumed to be merely suggestive and not causal. This is because we cannot rule out potential self-selection of the population into the census tracts in Tempe we are observing. This may bias our estimates. Similarly, we are also susceptible to potential spillover effects between watersheds in the collection areas that may over or understate the effects of a given characteristic on opiate use in the census tracts we are studying. We are also unable to make the distinction between prescribed and unprescribed use of opioids in our PNML estimates because we are only observing total usage in the city wastewater.

**Propose a combination of quasi-experimental designs might you use to mitigate the remaining threats to internal validity. Discuss how it would rule out other measure the size of these remaining threats.** (+2)

We propose using an equivalent materials quasi-experimental design to test the effects of demographic features on opioid use in Tempe. We will test the effects of the same demographic predictors on different measures of opioid use such as the PNML wastewater estimates and the EMS call data. While there is no explicit control group here, a determination of a consistent relationship between the same group of certain community factors and different proxies for opioid can help bolster our claim that these risk factors are enduring.

For this reason, we would use the Equivalent Materials Design mixed with a Patched Design.

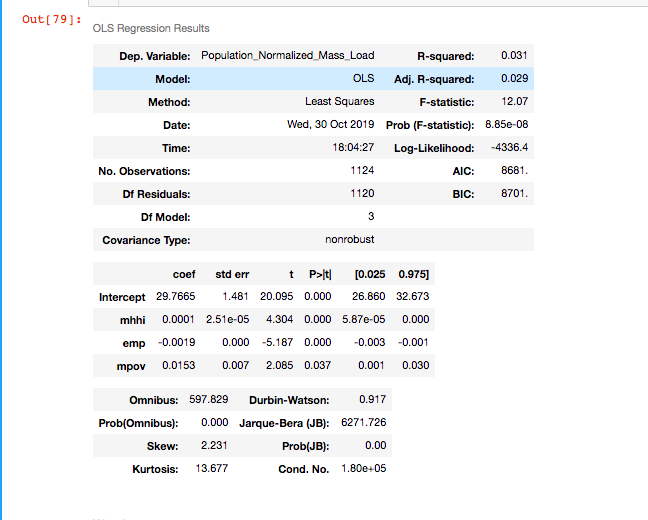
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Census Tracts within Wastewater Collection Sites  (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by PNML  (MaX1) | Observe which age group has higher PNML based on location within the wasterwater collection site.  (O) |  |  |  |  |
| Census Tracts within Wastewater Collection Sites  (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX1) | Observe which age group has higher EMS calls based on location within the wasterwater collection site.  (O) |  |  |  |  |
|  |  |  |  | Census Tracts **not** in Wastewater Collection Sites (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX0) | Observe which age group has higher EMS calls based on location **not** in the wasterwater collection site.  (O) |

**Demonstrate your proposed experimental or quasi-experimental design on pilot data.** (+2)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Census Tracts within Wastewater Collection Sites  (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by PNML  (MaX1) | Observe which age group has higher PNML based on location within the wasterwater collection site.  (O) |  |  |  |  |
| Census Tracts within Wastewater Collection Sites  (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX1) | Observe which age group has higher EMS calls based on location within the wasterwater collection site.  (O) |  |  |  |  |
|  |  |  |  | Census Tracts **not** in Wastewater Collection Sites (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX0) | Observe which age group has higher EMS calls based on location **not** in the wastewater collection site.  (O) |

Again here is our plan. We have to re-work which factor that we want to do but for the sake of this demonstration we will stick with age.

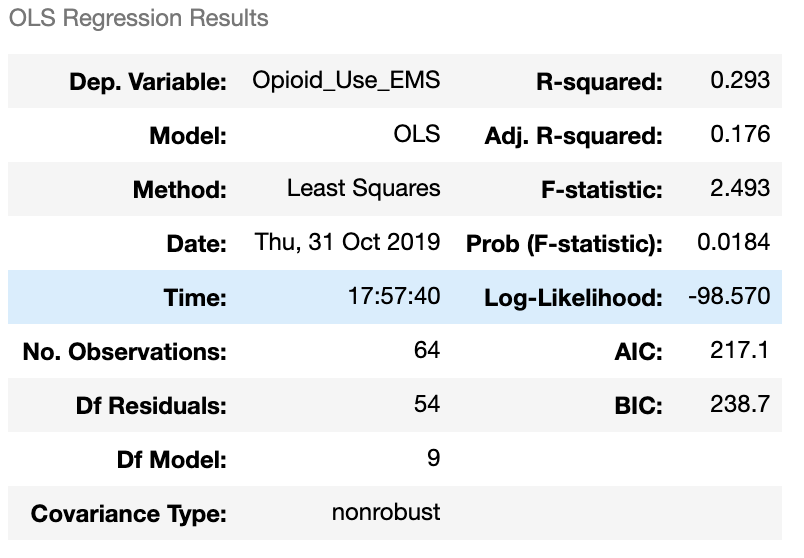
For PNML within Collection Sites:



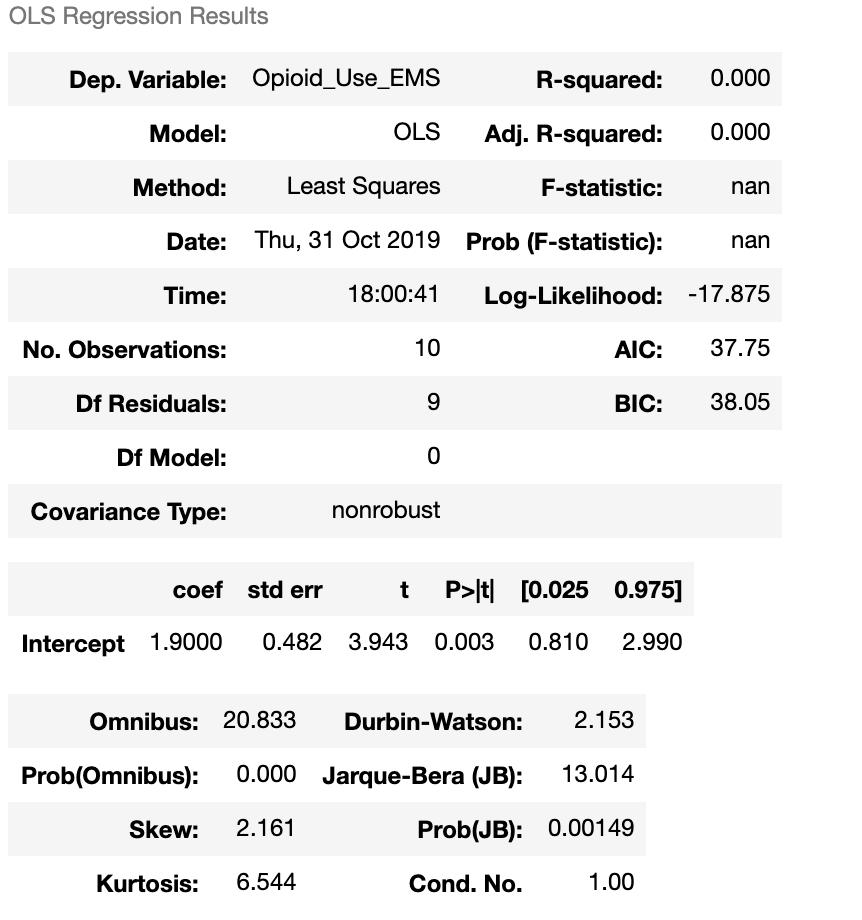
For EMS Opioid Calls within Collection Sites:

|  |  |  |  |
| --- | --- | --- | --- |
| Census Tracts within Wastewater Collection Sites  (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX1) | Observe which age group has higher EMS calls based on location within the wastewater collection site.  (O) |

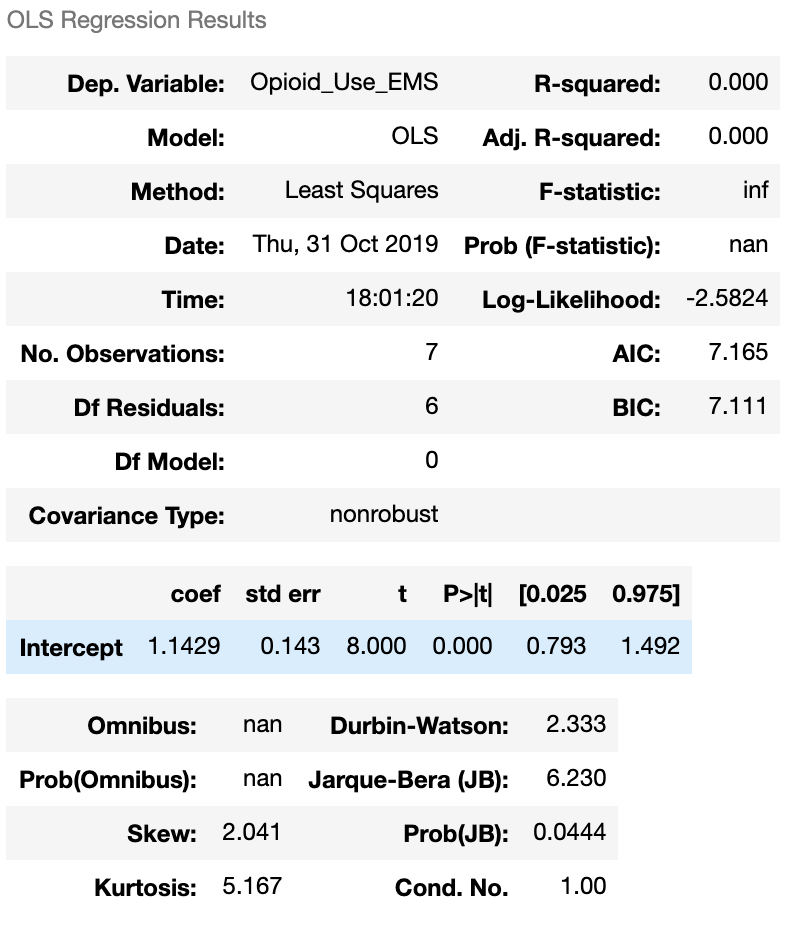
Area2:



Area 4:

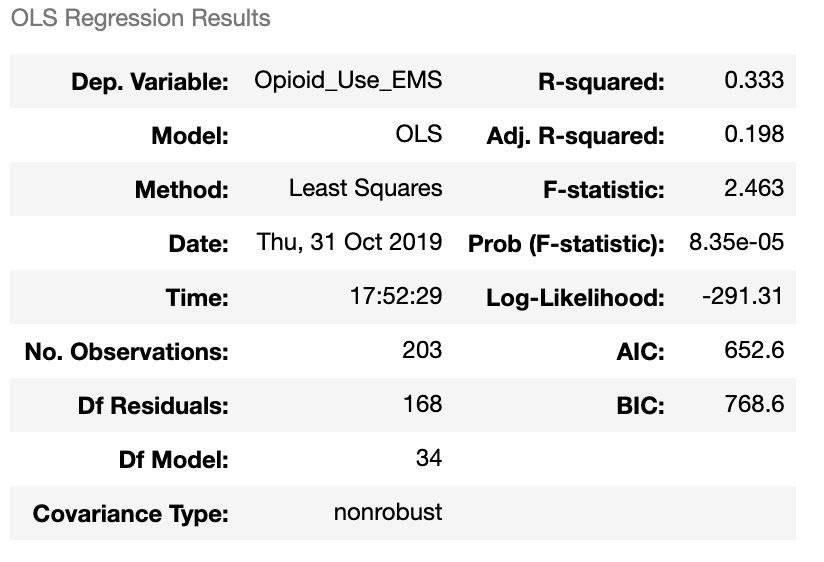


Area 5



For EMS Opioid Calls in Non-Collection Sites:

|  |  |  |  |
| --- | --- | --- | --- |
| Census Tracts **not** in Wastewater Collection Sites (N) | Observe breakdown of selected Factors (ages for now)  (O) | Opioid Usage Measured by EMS Opioid Related Calls  (MbX0) | Observe which age group has higher EMS calls based on location **not** in the wastewater collection site.  (O) |



1. Source: https://statisticalanalysisconsulting.com/is-time-nominal-ordinal-interval-or-ratio-is-it-categorical-or-continuous/ [↑](#footnote-ref-0)