Drug Use Among Age Groups

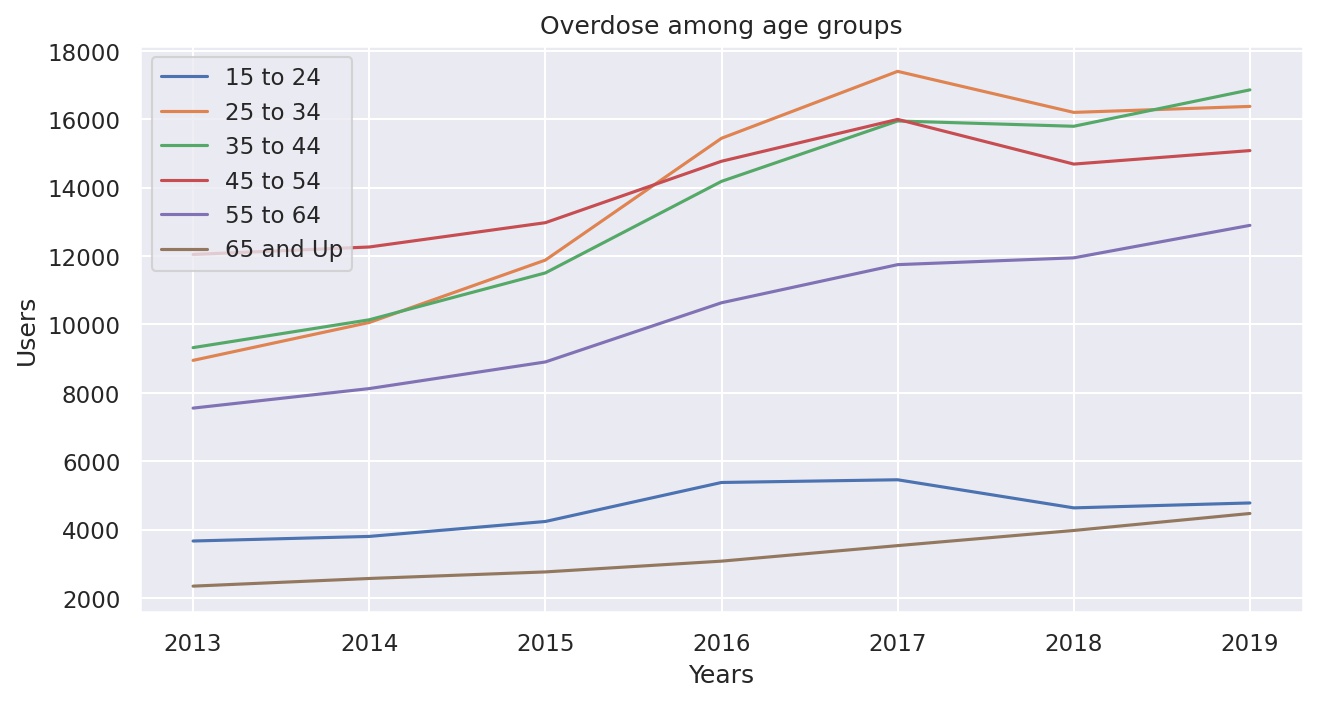
Final Report

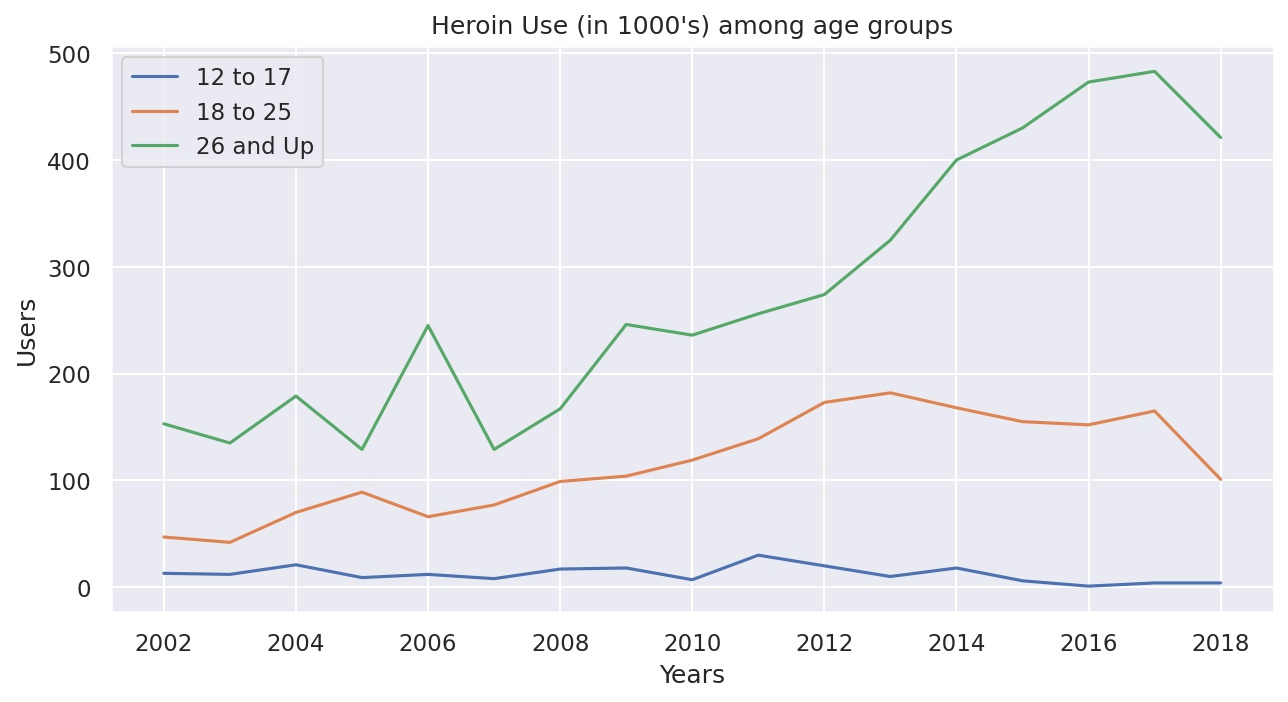
Introduction

For this project, I decided to focus on drug use by age. I wanted to see if younger people overdose more than older people.

Related Work

I replicated as best as I could from the original dataset. But really I just plotted the data since that’s what was shown in original replication.





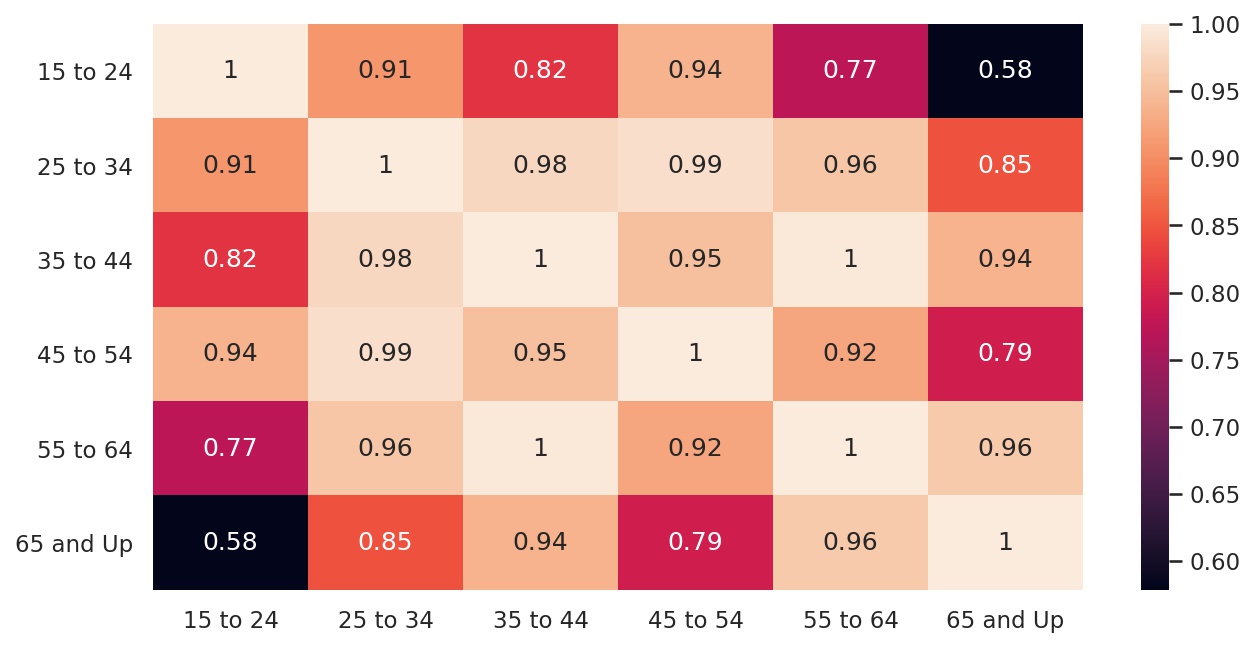
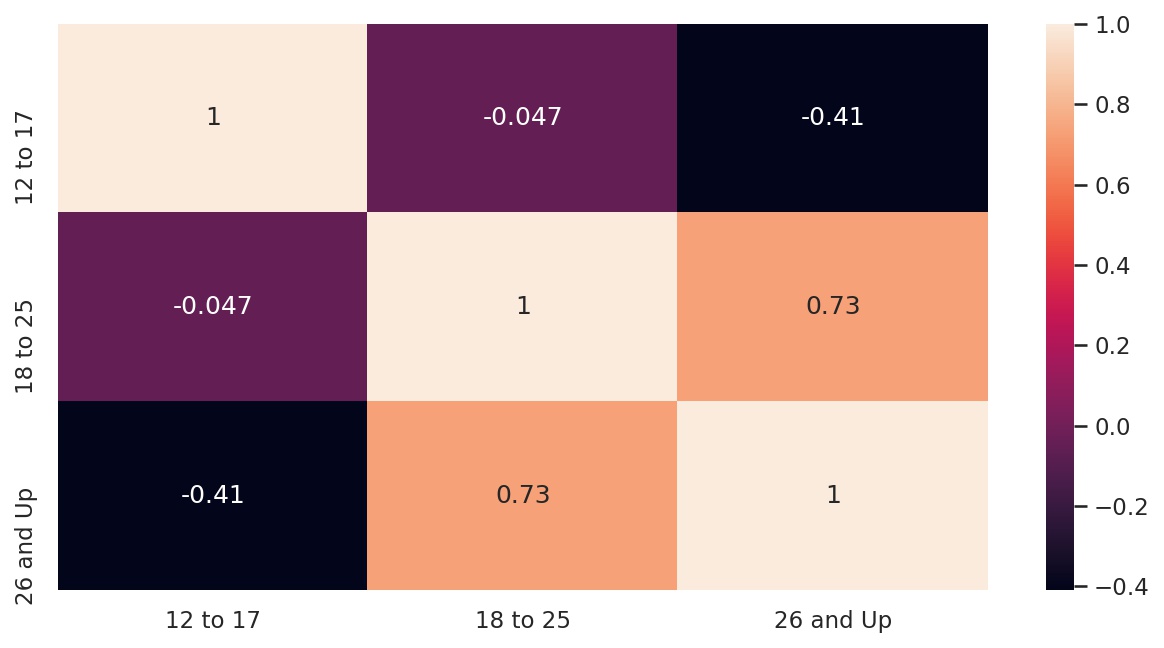
Dataset

I got my dataset from statista. It focuses on drug overdose from 2013 to 2019 in the US. There are 6 age categories.

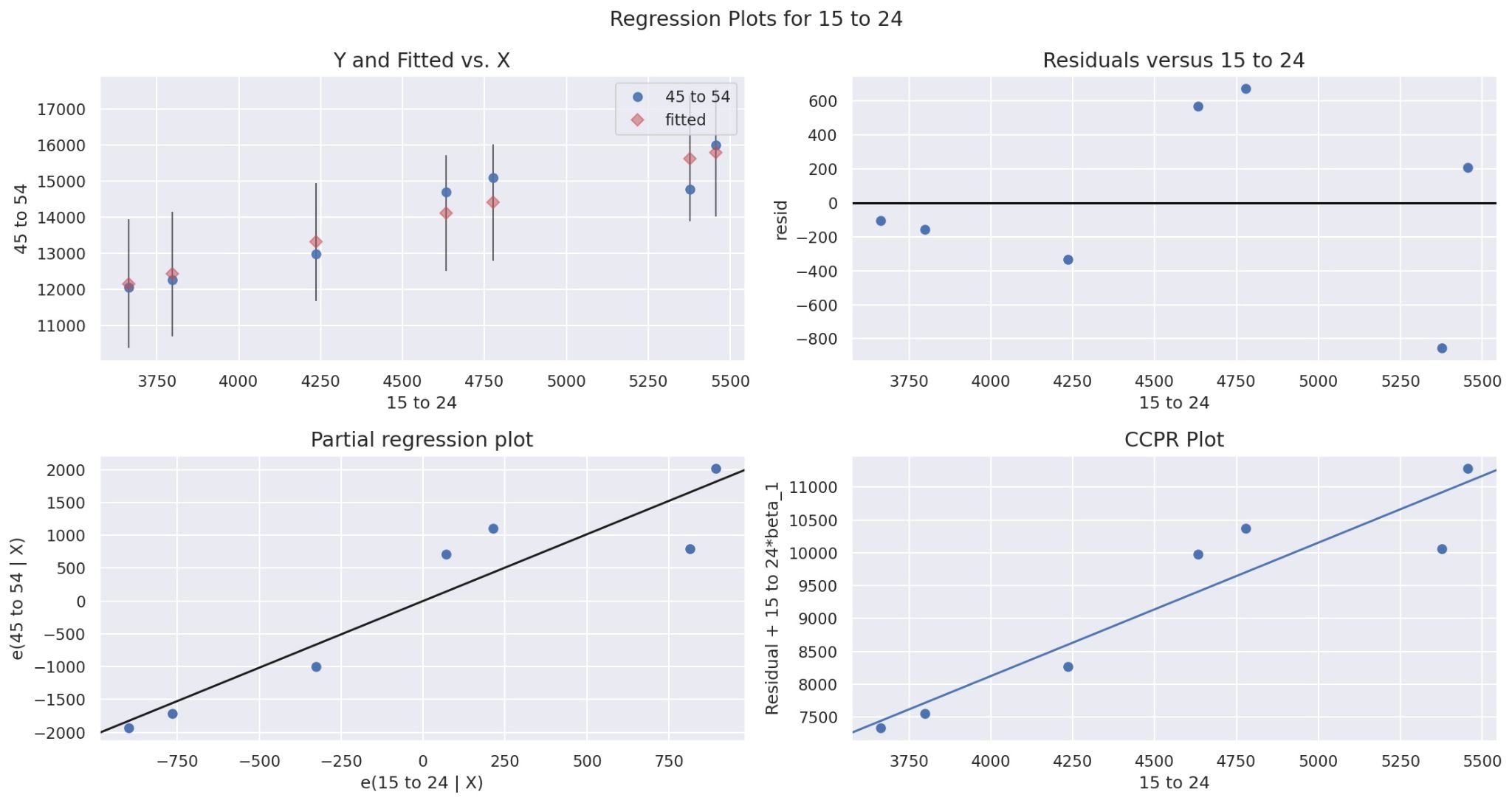
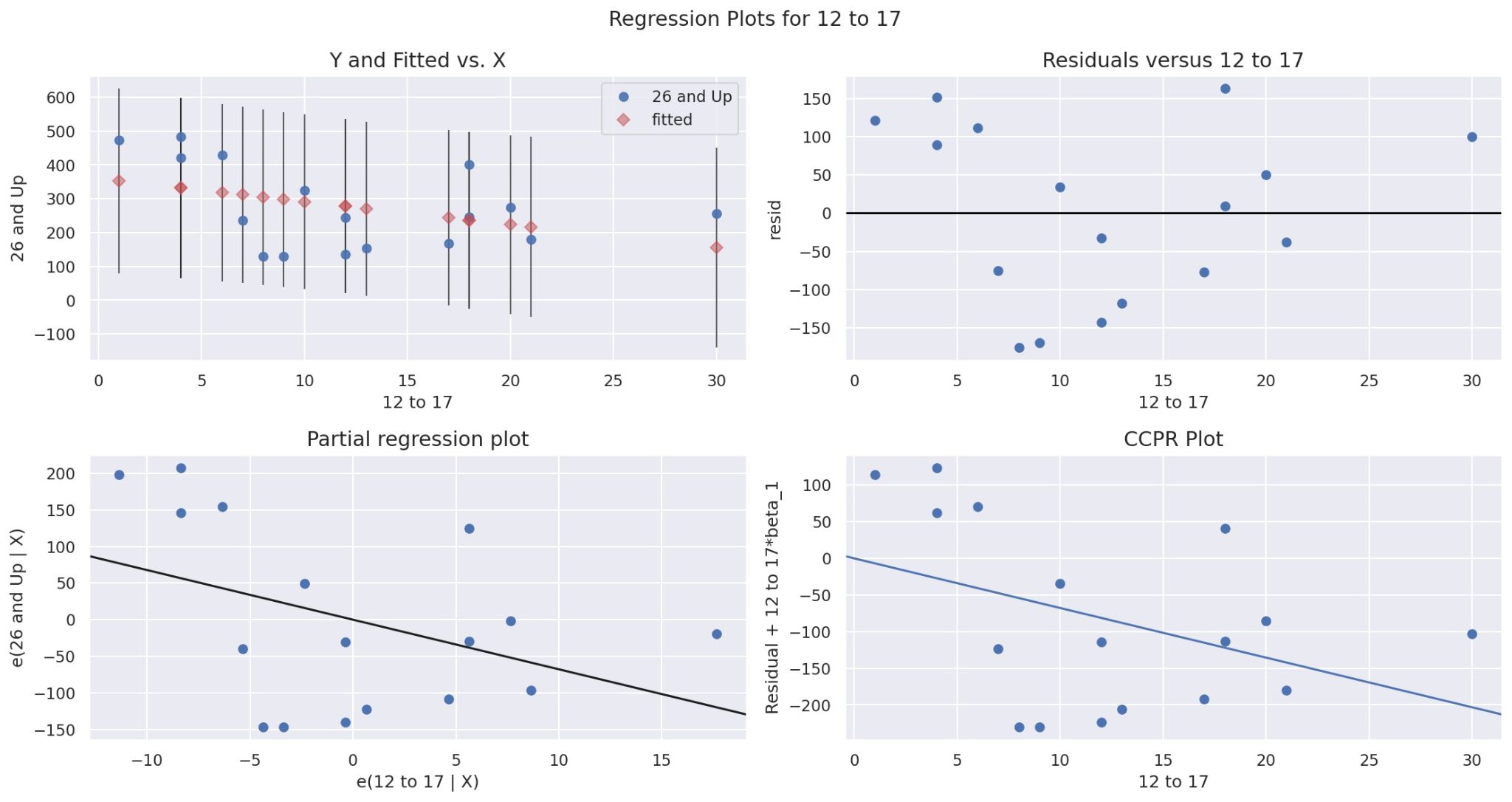
I also used another dataset from statista. It focused on heroin use from 2002 to 2018. It focuses on younger age groups but provides interesting information

Methods

I first read in my data using the pandas module. I then plotted out my data using matplot and seaborn module. After doing that I plotted out heatmaps for both the overdose and heroin datasets.

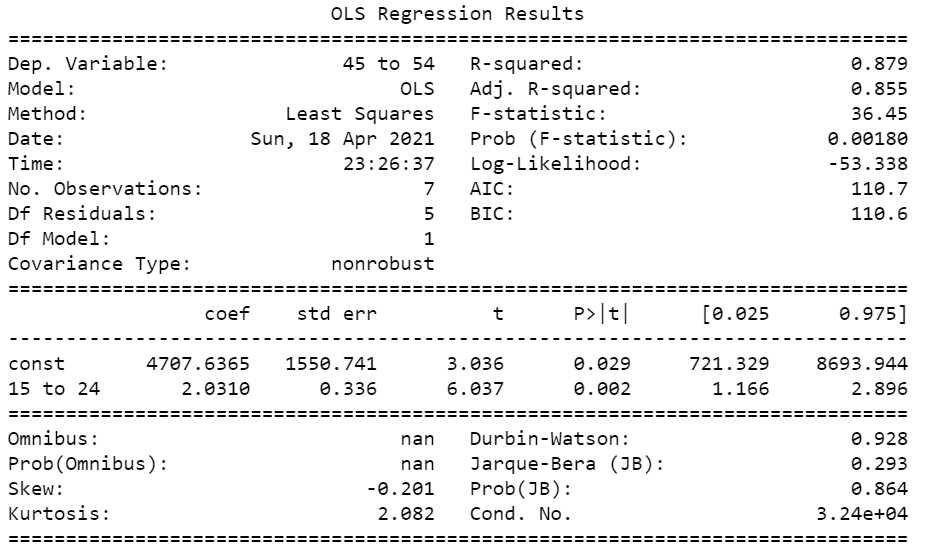
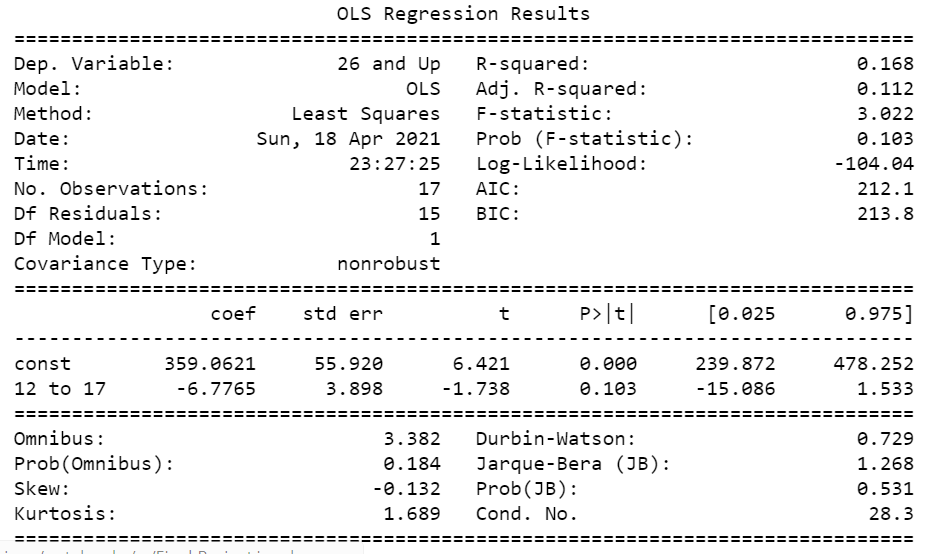
* 
  + This heatmap maps out all the different age categories for overdoses. You can tell younger age groups have a strong correlation to one another. However categories like “15 to 24” and “65 and Up” have a weak correlation
* 
  + This heatmap maps out all the different age categories for heroin use. Here the correlation between “12 to 17” and “26 and Up” is weak.

I then tried to see if I can find a linear model between the age group “18 to 25” and the older age groups. My thought process is that if a relationship exists between these two different age groups then I hope they’re inversely correlated, and I’ll later explain why.

* 
  + Given the overdose info for age group “15 to 24”, we tried to predict the overdoses for age group “45 to 54”.
* 
  + Given the overdose info for age group “15 to 24”, we tried to predict the overdoses for age group “45 to 54”.

Results

For my results, I printed out the summary of the Ordinary Least Squares model that I used from the statsmodel module.

* 
  + You can see there is a low p-value for the model, indicating that the predictor “15 to 24” is not weak.
* 
  + Here you can see that the P-value is above the .05 threshold, indicating that the predictor “12 to 17” is too weak to provide any useful information.

Conclusion

My conclusion is that younger age groups may or may not be more likely to be at risk for drug abuse. Dataset number 1 verifies that a correlation between younger age groups and older age groups exists. However, surprisingly if you take a look at the first graphs found in the related work section, you’ll notice that younger groups are at less risk than older groups for drug overdose. Dataset number 2 is a bit of a lost case. The age groups aren’t diverse enough to say that younger people are at more risk for heroin use. In addition, the models fail to show a relationship between the youngest age group “12 to 17” and the oldest age group “26 and Up”. Overall, this project didn’t plan out as I thought it would. I would’ve preferred more complex methods and a much larger dataset.

References

* Elflein, John. “Number of Drug Overdose Deaths in the U.S. from 2013 to 2019, by Age.” *Statista*, Mar. 2021, www-statista-com.proxy2.cl.msu.edu/statistics/611017/drug-overdose-deaths-number-in-the-us-by-age/.
* Elflein, John. “Number of Individuals with a Heroin Dependence or Abuse during the Past Year in the United States from 2002 to 2019, by Age.” *Statista*, Sept. 2020, www-statista-com.proxy2.cl.msu.edu/statistics/476285/number-of-individulas-with-a-heroin-dependence-or-abuse-by-age-us/.