Final Report: How Baby Boomers Get High

Hunter Baldwin

Baldw222

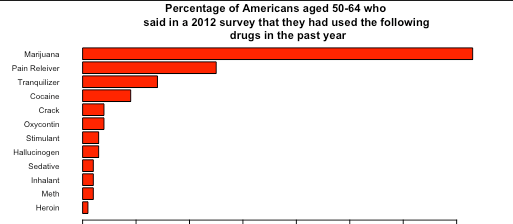
1. Introduction/ Problem Statement

Based on the fivethirtyeight website the main question or problem answered with this report was, how do baby boomers get high. At first I was confused why we would be analyzing baby boomers drug activity, when you normally relate drug use to a younger generation. But after reading I was genuinely surprised by the results of the report and by the number of baby boomers actually doing some of these drugs. Now they may not be leading the drive in recreational drug use but they are certainly not slowing it down. The question I wanted to answer was how do other ages compare when it comes to the types of drugs they use.

1. Related Work

This is obviously not the only dataset that analyzes drug use by age nor is it the best, but after searching for other ones to compare, this seemed to be the easiest dataset to work with. I found another data set from The National Institute on Drug Abuse where they did a study on the trends of recreational drug use but its only for 8th graders through 12th grade so there wasn’t any info on the older generations. I found other datasets while looking around but each one had it flaws and wouldn’t help on this project so I ended up sticking with the original.

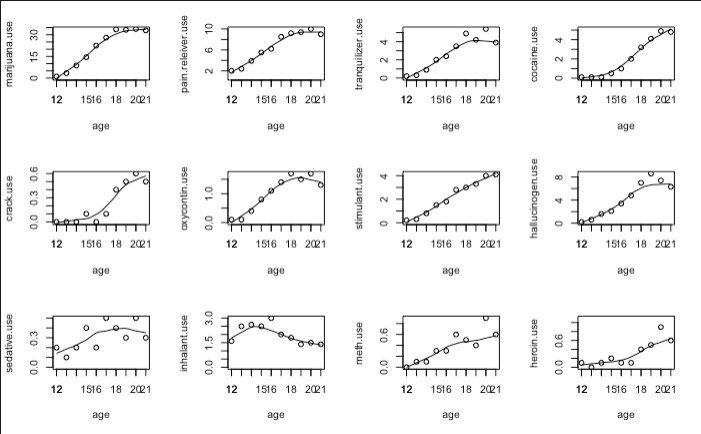
As guided I recreated the graph on the original report shown below:



1. Dataset/ Model

This dataset was from the fivethirtyeight github repository and given that there isn’t a ton of data that can be used to predict anything I wanted to see the distribution between each age group and each drug. The original model was only analyzing ages 50-64 which is certainly an interesting model but it would’ve been cool to visualize all drugs for all ages. I tried to get all the drugs analyzed for all ages on one bar chart but it was too much data and was unreadable so I didn’t include it. I then just ran the line of best fit through the ages and each drug.

Below are the graphs for the age groups against each of the drugs:



1. Results

There were some interesting results from the data just by graphing it, I can see that marijuana use is most popular among all ages, and stuff like crack, meth, and heroin were some of the less used drugs between the ages. It was also interesting to see the types of drugs 12 year olds were getting into like inhalants, sedatives, and most popular according to this data set, pain relievers.

I tried applying various models like svm to the dataset as well as glm, and spline but given the dataset the results weren’t really conclusive.

1. Conclusion

In conclusion this was a pretty interesting project, I was very surprised by some of the results I obtained especially the types of drugs kids were using. Given the lack of data and predictor there wasn’t much modeling you could do, but if I did this project again I would’ve been interested in finding a data set that had information on these drugs before some of them were legal, and after to see if there was any difference in the response. It would’ve also been interesting to see what drug had the most casualties, that way you could actually use predictions to analyze what would happen. But overall this was an interesting project and very eye opening.

1. Citations

* <https://fivethirtyeight.com/features/how-baby-boomers-get-high>
* <https://www.drugabuse.gov/drug-topics/trends-statistics/monitoring-future/monitoring-future-study-trends-in-prevalence-various-drugs>