

# Testing the Effects of Marijuana Legalization on Crime and Other Drug Usage in Seattle

*Benjamin Gaiser and Jeremy Russell*

*11 November 2016*

**OVERVIEW:** Building on our work since the second assignment and given time constraints, we decided to focus solely on the city of Seattle, Washington for this third assignment. A full comparative analysis of Seattle and Denver is to come in the final assignment. As mentioned, residents of nine other states in addition to Colorado, Washington, Oregon and Alaska recently voted on measures to legalize either recreational or medicinal cannabis use. California, Maine, Massachusetts and Nevada voted to legalize recreational use, while the outcome in Arizona is still forthcoming. Florida, Arkansas and North Dakota voted to legalize the medicinal use of cannabis, while the outcome in Montana is also forthcoming.

**REPOSITORY:** This repository contains the R code for web scraping, cleaning and analyzing our data. Currently we have data for marijuana production and distribution in the state of Washington, and crime and marijuana stats for the city of Seattle. We added U.S. census data to account for socio-economic factors.

**ASSIGNMENT:** In the third pair assignment you will gather web based data from at least two sources, merge the data sets, conduct basic descriptive and (some) inferential statistics on the data to address a relevant research question and briefly describe the results including with dynamically generated tables and figures. Students are encouraged to access data and perform statistical analyses with an eye to answering questions relevant for their Collaborative Research Project. Deadline 11 November, the write up should be 1,500 words maximum and use literate programming, 10% of final grade.

**DATA PROCESSES:** We began by scraping the data for cannabis retailers and producers separately, using phantomjs server, which coincidentally is not compatible with a Mac. So we made a simple .txt file. After cleaning the data, we made variables for tax, sales and recorded month. There was a problem with the names of retailers not uniquely identifying retail shops because there were two different URLs for 365 cannabis use. Next, we subset the data on retailers from Washington into King County; downloaded the Seattle crime data with jsonlite; subset this data into narcotics and/or alcohol related data, since we wanted to focus on drug-related crimes; and wrote this information into a zip file. We would like to note that we are considering adding data on the effects of legalization on more serious crimes (property and violent crimes) in addition to our current analysis of drug related crimes to more accurately specify the effects of legalization. Then we scraped U.S. census data and determining which variables to include in our analysis; we decided to include (1) median household income, (2) age, (3) race and (4) education. We used information from the following blog post to do this: <http://zevross.com/blog/2015/10/14/manipulating-and-mapping-us-census-data-in-r-using-the-acstigris-and-leaflet-packages-3/>. **HERE EXPLAIN WE DROPPED DUPLICATE DATA FROM 502 AND KUSHTOURISM SITES / EXPLAIN WHERE DATA COMES FROM ON THESE SITES.** We then saved the geolocation data as a .txt file; created parameters to assess the minimum distance between latitude/longitude points; and made a binary variable representing when a retailer was established (for a total of fifteen retailers). Our assumption for the U.S. census data on district-specific crimes is that a perpetrator of some crime is probably more likely to commit that crime in their own neighborhood than to travel far away to do so, but we recognize the limitations of this logic.

**DESCRIPTIVE STATISTICS:** We used a simple bar chart to show the growth of producers/processors and retailers of cannabis in Seattle between the years 2014-16. Between 2014-15, producers grew rapidly. They surged by 201% (from 185 to 574); between 2015-16, they grew more slowly - by 53% (from 574 to 881). Retailers grew by 130% (from 82 to 189) and 69% (from 189 to 319) respectively between 2014-15 and 2015-16. **SHOW BAR CHART**

Data on sales and tax revenue also show a massive increase over these two years. By 2015, sales revenue leaped by 960% from \$29.21 million to \$309.86 million, and in 2016, by 77% from \$309.86 million to \$549.93

million. Tax revenue in 2015 grew by an astounding 1040% from \$9.74 million to \$110.97 million, but slowed down a bit in 2016, increasing by 83% from \$110.97 million to \$203.47 million. SHOW CHART

Next we created a map showing all narcotics-related crimes in green, all retailers in red, and wealth as a heat map where median household income was \$200,000+ per year. Crime was concentrated in the central area of Seattle, and was not clustered around retailers. SHOW MAP

Lastly, we plot crime development in Seattle, and narcotics, alcohol and weed do not seem to show any trend, though weed has a higher variation than the other two variables.

**OUTCOMES:** Univariate regression Multivariate regression and comparison with Denver to follow in the final analysis

## References

---