Accuracy Differences in Cannabis Retailer Information Ascertained from Webservices and Government-Maintained State Registries Across US States Legalizing the Sale of Cannabis in 2019

Michael Williams[1,2], Matt Mahlan[1], Connor Holmes[1], Magdalena Pankowska[1], Manjot Kaur[1], Aderonke Ilegbusi[2], Danielle F. Haley[1,3]

[1]Institute for Health Equity and Social Justice Research and Department of Health Sciences, Northeastern University
[2]Bouvé College of Health Sciences, Northeastern University
[3]Department of Community Health Sciences, Boston University School of Public Health

ABSTRACT

Cannabis retailer locations used to investigate geographic cannabis access are frequently ascertained from two sources: 1) webservices which provide locations of cannabis retailers (e.g., Yelp) or 2) government-maintained registries. Characterizing the operating status and location information accuracy of cannabis retailer data sources on a state-by-state level can inform research examining the health implications of cannabis legalization policies. This study ascertained cannabis retailer name and location from webservices and government-maintained registries for 26 states and the District of Columbia legalizing cannabis sales in 2019. Validation subsamples were created using state-level sequential sampling. Phone surveys were conducted by trained researchers for webservice samples (n=790, November 2019 - May 2020) and government-maintained registry (n=859, February - June 2020) to ascertain information about operating status and location. Accuracy was calculated as the percent agreement among subsample and phone survey data. For operating status and location, webservice derived data was 78% (614/790) and 79% (484/611) accurate, whereas government-maintained registry derived data was 76% (657/859) and 95% (622/655) accurate, respectively. Fifty-nine percent (15/27) of states and the District of Columbia had over 80% accuracy for operating status and 48% (13/27) states had over 80% accuracy for location information with both data sources. However, government-maintained registry derived information was more accurate in 33% (9/27) states for operating status and 41% (11/27) states for location information. Both data sources had similar operating status accuracy. Research using spatial analysis may prefer government-maintained registry derived data due to high location information accuracy, whereas studies looking at broad trends across states may prefer webservice derived. State level COVID-19 restrictions had minimal impact on ascertainment of cannabis retailer operating status and location information via phone survey derived from webservices and government-maintained registries.

Key words: = cannabis; dispensaries; cannabis retailers; cannabis outlets; phone survey

State legalization of the sale of cannabis for medical or adult use in the United States (US) is a rapidly evolving and highly variable landscape. By the end of 2019, 29 states and the District of Columbia (DC) have legalized and implemented the sale and consumption of cannabis for medical use and 9 states and DC have legalized and implemented the sale and consumption cannabis

Corresponding Author: Michael Williams, Ph.D., Institute for Health Equity and Social Justice Research, Department of Health Sciences; Bouvé College of Health Sciences, Northeastern University. Email: mpw144@gmail.com

for adult use (Legal Science, 2017a, 2017b; NORML Foundation, 2022). By 2022, an additional 5 states have legalized and implemented the sale and consumption of medical cannabis and 8 additional states have legalized and implemented the sale and consumption of cannabis for adult use. State approaches to the legalization of cannabis are not static. States which have passed cannabis legalization in some form frequently update cannabis related policies and infrastructure over time (Legal Science, 2017a, 2017b). For example, New York legalized the sale and consumption of cannabis for medical use in 2014. This policy was updated in 2015 to include restrictions on how close medical cannabis retailers could be to specific buildings, such as churches (Legal Science, 2017a). Furthermore, state approaches to cannabis legalization vary across states. For example, in New Hampshire there are no policies set regarding cannabis retailer proximity to churches (Legal Science, 2017a). Previous research has identified factors associated with cannabis retailer locations such as local supply and demand (Morrison et al., 2014), housing prices (Burkhardt & Flyr, 2019; Cheng Cheng et al., 2018; Conklin & Diop, 2017; Francesca Ortegren, 2021), and local zoning ordinances (Morrison et al., 2014). Furthermore, policy within and across US states has been suggested as a factor relating to state-level heterogeneity in geographic access to cannabis retailers (Chapman et al., 2016).

Beyond cannabis legalization and implementation policy, recent state policies emerging from the COVID-19 pandemic may have influenced the operating status of cannabis retailers. On March 13th 2020, a national emergency was declared in response to the novel coronavirus (COVID-19)(*COVID-19 Disaster Declarations | FEMA.Gov*, 2021; *Declaring a National Emergency Concerning the Novel Coronavirus Disease (COVID-19) Outbreak*, 2020). This declaration approved all 50 states and D.C. for major disaster declarations. Shortly thereafter, many state governments also declared a state of emergency which in most cases included provisions for mandatory shutdowns of all businesses except those considered essential businesses. In many states, cannabis retailers were considered essential businesses while other states did not consider cannabis retailers essential businesses and therefore these

businesses were temporarily shutdown (Julia Raifman et al., 2020). While the specific effects of COVID-19 related shutdown policies have on cannabis retailer operating status is unclear, recent research has illuminated that cannabis sales (Chong et al., 2022) and consumption (MacKillop et al., 2021; Schauer et al., 2021) increased during the pandemic in various parts of North America. These studies suggest specific aspects of COVID-19 policies as potentially related factors to the increase in cannabis sales and consumption (e.g., shelter in place orders). Altogether, this suggests that COVID-19 may have had an impact on cannabis retailer operating status.

Numerous studies have leveraged cannabis retailer location information to explore how geographic access to cannabis influences cannabis use and other health outcomes (Contreras, 2017; Freisthler & Gruenewald, 2014; Lankenau et al., 2019; Lipperman-Kreda et al., 2014; Mair et al., 2015, 2021; Shi et al., 2016, 2018; Shih et al., 2019; Tabb et al., 2018). Geographic access to cannabis retailers is frequently measured through the number of cannabis retailers per geographic unit (e.g., count of cannabis retailers in a census tract) or relative to a geographic point of interest (e.g., count of cannabis retailers within 3 miles of a school) (Amiri et al., 2019; Freisthler et al., 2017; Freisthler & Gruenewald, 2014; Morrison et al., 2014; Shi et al., 2016, 2018; Tabb et al., 2018; Thomas & Freisthler, 2016). These measures are used to investigate how access to cannabis and cannabis retailers impact use and consumption of cannabis among key populations, (Freisthler & Gruenewald, 2014; Shi et al., 2016). However, the relationship between geographic access to cannabis retailers and cannabis use is unclear. For example, two studies conducted in Los Angeles, California assessed frequency of cannabis use as a function of geographic access to medical cannabis retailers in young adults (Lankenau et al., 2019; Shih et al., 2019). One study conducted in 2016-2017 found that living near a medical cannabis retailer was positively associated with frequency of use (Shih et al., 2019) while the other study conducted in 2014-2015 found that there was no association (Lankenau et al., 2019). Both studies used a different combination of techniques to generate lists of cannabis retailers. In addition to other confounding factors, it is possible that varying

data sources used to create measures of geographic access to cannabis retailer may play a role in the divergent relationships between cannabis retailer access and frequency of use.

The cannabis retailer data sources used in research leverage real-world data which are often not constructed with research purposes in mind. Across the US, cannabis retailer locations are frequently ascertained from two primary sources: 1) webservices which provide the locations of cannabis retailers (e.g., Yelp, Leafly)(Freisthler et al., 2017; Freisthler & Gruenewald, 2014; Morrison et al., 2014; Pedersen et al., 2018, 2020; Shi et al., 2018; Thomas & Freisthler, 2016; Unger et al., 2020) or 2) government-maintained registries (e.g., California Bureau of Cannabis Control licensing directory) (Amiri et al., 2019; Freisthler et al., 2017; Freisthler & Gruenewald, 2014; Shi et al., 2016; Tabb et al., 2018; Thomas & Freisthler, 2016; Unger et al., 2020). Both data sources have different considerations for use in research. For example, data sourced from webservices can be unreliable due to listing incorrect or incomplete address information (e.g., only listing the town and state) (Pedersen et al., 2018). Additionally, cannabis licensing is handled by a variety of government bodies within and across states (e.g., departments of public health, departments of revenue), making both the licensing process and record keeping heterogenous across states. For example, some states provide separate registries for medical and adult use cannabis retailers, provide business contacts instead of the contact information for the storefront, or provide historical data for locations which have since closed (Klieger et al., 2017).

Two studies conducted by Pedersen and colleagues examining the accuracy of cannabis retailer data sources found that the accuracy of webservices with respect to operating status varied highly between webservice providers (Pedersen et al., 2018, 2020). Since both studies were conducted in Los Angeles, California, findings may not be generalizable to other states. Furthermore, neither of these studies evaluated the accuracy of retailer address information. Due to differences in cannabis legalization policy and cannabis infrastructure, cannabis retailer information such as location and operating status may vary across states and within states when derived from different sources (e.g., webservices or government-maintained registries). Without

characterizations of the accuracy of the primary data sources from which cannabis retailer information is derived, policy makers and researchers alike are potentially drawing incomplete or wrong conclusions surrounding the health implications of access to cannabis retailers.
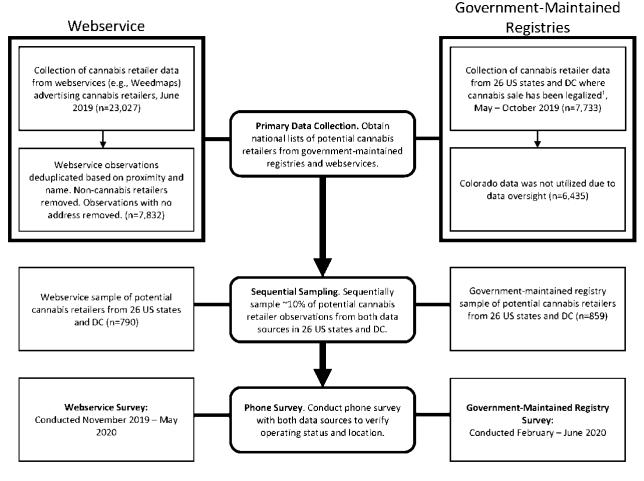
This study examines the accuracy of cannabis retailer operating status and location information across states and data sources by ascertaining cannabis retailer information obtained from webservices and government-maintained registries across 26 states and DC in 2019. We conducted a follow-up phone survey with a sample from each to confirm their operating status and location. There are two primary aims of this study. The first aim is to quantify the accuracy of observations from both webservices and government-maintained registries with respect to operating status and location across 26 states and DC in 2019. The second aim is to provide considerations for utilizing webservices and government-maintained registries for cannabis research. These aims stand to explicate the accuracy of cannabis retailer information ascertained from webservices and government-maintained registries which in turn will help policy makers and researchers alike contextualize research utilizing these data sources by providing considerations per data source and on a state-by-state basis.

METHODS

This study ascertained cannabis retailer information from webservices and government-maintained registries and used phone surveys to quantify the accuracy of operating status and location information in each. This study consisted of four stages: 1) primary data collection, 2) sequential sampling of retailer lists, 3) phone validation survey, 4) analysis (Figure 1). This research was reviewed by the University of North Carolina at Chapel Hill Institutional Review Board and was determined to be non-human subjects research.

Primary Data Collection

Cannabis retailer information (name, location) was ascertained from two data sources: 1) webservices which advertise cannabis retailer information (i.e., Yelp, Weedmaps, Leafly, canorml,

Figure 1. *Data Collection and Sampling of Cannabis Retailer Information from Webservices and Government-Maintained Registries for Use in Phone Survey to Assess Accuracy of Operating Status and Location Information in United States, 2019*

Webservice

Government-Maintained Registries

Collection of cannabis retailer data from webservices (e.g., Weedmaps) advertising cannabis retailers, June 2019 (n=23,027)

Webservice observations deduplicated based on proximity and name. Non-cannabis retailers removed. Observations with no address removed. (n=7,832)

Primary Data Collection. Obtain national lists of potential cannabis retailers from government-maintained registries and webservices.

Collection of cannabis retailer data from 26 US states and DC where cannabis sale has been legalized[1], May – October 2019 (n=7,733)

Colorado data was not utilized due to data oversight (n=6,435)

Webservice sample of potential cannabis retailers from 26 US states and DC (n=790)

Sequential Sampling. Sequentially sample ~10% of potential cannabis retailer observations from both data sources in 26 US states and DC.

Government-maintained registry sample of potential cannabis retailers from 26 US states and DC (n=859)

Webservice Survey: Conducted November 2019 – May 2020

Phone Survey. Conduct phone survey with both data sources to verify operating status and location.

Government-Maintained Registry Survey: Conducted February – June 2020

1. Alaska, California, Colorado, Connecticut, Washington D.C., Delaware, Florida, Hawaii, Illinois, Louisiana, Massachusetts, Maine, Maryland, Michigan, Minnesota, North Dakota, New Hampshire, New Jersey, New Mexico, New York, Nevada, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, Vermont, Washington

Kushguide) and 2) state government-maintained registries from all states and DC which had legalized and implemented cannabis for adult and/or medical use in 2019 (Figure 1).

Webservices

In June 2019, we obtained the name, location, and geocoordinates of all listed cannabis retailers from five different webservices (n=23,027). Due to the large volume of observations, an algorithmic deduplication procedure was utilized to eliminate duplicate observations across webservices. Two or more observations were considered duplicates if more than 50% of the characters in the name matched (position and alphabetical character) and the latitude and longitude were within 200 feet of

each other. This procedure identified 3,399 observations with one or more duplicates across webservices, which once removed resulted in a dataset of 16,559 observations. After deduplication, 6,010 observations with no address information and 2,717 observations which corresponded to self-advertised non-cannabis retailers were removed from the dataset. We extracted flags used internally by each webservice to denote properties of the facilities advertised on their platform (e.g., doctor, delivery, online ordering). These flags were used to remove non-cannabis retailers (e.g., doctor's offices) from the final dataset. This resulted in a dataset where each observation was a unique potential cannabis retailer (n=7,832 cannabis retailers). This dataset was used for sampling as described below.

Government Maintained Registries

Throughout May to October 2019 the 29 states and DC which had legalized the sale of cannabis for adult or medical use as of October 2019 were contacted to solicit registries of cannabis retailers (Figure 1). Of the 29 states and DC contacted, 27 provided registries for cannabis retailers in their state. Arizona and Montana did not provide registries. The address associated with each observation in each registry was geocoded to obtain the latitude and longitude of each cannabis retailer. These registries were then cleaned by uniformly formatting column names and column data types across registries then removing non-cannabis retailers (e.g., cannabis growers or testing labs). Each registry was then merged into a dataset containing all states' cannabis retailer information including name, location, and geocoordinates (latitude and longitude) (n=7,733). Due to a data oversight, Colorado registry information was not incorporated during this step. This dataset comprising 26 states and DC was used for sampling described below.

Sampling Method

A state-level sequential sampling method was used to enable a wide range of geographic regions within each state to be sampled and for each state to provide a critical mass of observations to the overall validation study. This sampling method also aligns with the goals of a parent grant-funded study which investigates geographic access to cannabis retailers (NIDA 5K01DA046307). For each state in both the webservice dataset and the government-maintained registry dataset, the data were first sorted by census tract and then ~10% of the data were sequentially sampled. In states where the total number of observations was less than 10, all observations from that state were sampled. In states where different registries were provided for medical and adult use cannabis retailers, the sampling was done without replacement so that the same retailer could not be sampled twice from both registries. This resulted in a sample of 790 cannabis retailers from webservices and 859 cannabis retailers from government-maintained registries. Colorado was dropped from the government-maintained registries sample due to a data oversight, and thus was excluded from the registry validation samples,

leaving 29 states plus DC included in the webservice sample and 26 states plus DC in the registry sample.

Phone Survey

Phone surveys were conducted by trained project staff for both webservice (November 2019 to May 2020) and government-maintained registry (February to June 2020) samples to ascertain accuracy of operating status and location. Dedicated project staff were trained on how to conduct phone surveys using a detailed step-by-step protocol. This protocol walked project staff through the use of a tracking spreadsheet to track call attempts and meta information about calls (such as call time), the procedures for making the calls, and how to collect information during the phone surveys. Google Voice, a digital phone service, was used to standardize the phone number used for each call and maintain a centralized call log. Qualtrics was used by project staff to capture data regarding operating status and location. Survey scripts were standardized for different contingencies when calling potential cannabis retailers (attendant answered the phone, reaching a dial-based menu, and no answer). If there was no answer or the phone survey was incomplete (e.g., call dropped) then research staff were instructed via the protocol to call up to six times total. These repeat call attempts were spread across the project duration.

Measures and Accuracy

Operating status was determined by phone surveyors. If the storefront they contacted sold cannabis products which contained THC for adult-use or medical consumption, this location was considered a cannabis retailer. Storefronts which only sold CBD cannabis products (e.g., a naturopathic medical store) or sold no cannabis products (e.g., a grocery store) were considered non-cannabis retailers. When surveyors completed six or more call attempts which resulted in an inability to determine the operating status of a given observation, these observations were considered unconfirmed.

Location status was determined by phone surveyors by directly asking the attendant if the address on file is correct or using the information from a dial-based menu to compare against our

data. The address was piped directly into the Qualtrics survey from the respective data source and phone surveyors could respond to this question by confirming the address is correct, confirming the address is incorrect, or responding that they could not confirm the address information either way (unconfirmed).

Accuracy of both operating status and location information was operationalized as the percentage of agreement among observations between the primary data source and the phone validation survey for the same data source. With this operationalization, the percent agreement was conceived similarly to a measure of intercoder reliability (ICR), where two coders label observations on the nominal scale (O'Connor & Joffe, 2020). ICR is an analytic framework used across a variety of disciplines in qualitative research including public health and social psychology (O'Connor & Joffe, 2020). This framework was applied to our analysis, where the primary data source and follow-up phone survey were considered two different attempts to code the underlying population of cannabis retailers. In qualitative literature, a benchmark of 80% agreement is frequently used a minimum benchmark for acceptability of coding reliability (O'Connor & Joffe, 2020). For this analysis, the same 80% agreement benchmark is used to characterize acceptable accuracy for states and data sources.

Operating status accuracy was calculated using the entirety of each sample and location accuracy was calculated using a subset of each sample comprised of confirmed operating cannabis retailers. This was to avoid counting accurate locations of non-cannabis retailers in the accuracy percentage. To examine operating status and location accuracy across states, each state was categorized depending on whether both data sources had acceptable accuracy (greater than 80% agreement), one data source had acceptable accuracy while the other did not, or both data sources had lower than acceptable accuracy. Further, among the cases where one data source had acceptable accuracy and the other did not, we categorized the degree of the difference between the two data sources as less than 10% or greater than 10%.

Due to a portion of phone surveys being conducted after COVID-19 related restrictions were put into place in many states, we examined accuracy of operating status in both data sources pre- and post- COVID-19 related lockdowns to assess the degree to which these lockdowns may have impacted our analysis. A national emergency was declared in the US on March 13th 2020 (*Declaring a National Emergency Concerning the Novel Coronavirus Disease (COVID-19) Outbreak*, 2020) regarding COVID-19 and we used this date to determine pre- and post- COVID-19 related lockdowns. States in which mandatory shutdowns of cannabis retailers were implemented were considered in this analysis (Julia Raifman et al., 2020). Analysis was carried out using Python with Numpy and Pandas libraries (*NumPy*, 2022; *Pandas - Python Data Analysis Library*, 2022).

RESULTS

Overall

Of the 790 potential cannabis retailers sampled from webservices (Table 1), 614 (78%) were confirmed operating, 123 (16%) were confirmed to not be operating, and 53 (7%) were unconfirmed. Three observations were confirmed operating via phone survey, but the call was dropped before surveyors could confirm location information and subsequent phone call attempts were unsuccessful. Of these 611 remaining confirmed operating cannabis retailers derived from webservices, 484 (79%) were confirmed to be at the location listed in the web service data source, 108 (18%) confirmed location of retailer was incorrect, and 19 (3%) could not be confirmed either way (Table 1).

Of the 859 potential cannabis retailers sampled from government-maintained registries (Table 1), 657 (76%) were confirmed operating, 97 (11%) were confirmed to not be operating, and 105 (12%) were unconfirmed. Two observations were confirmed operating but did not have location information as the call was dropped before surveyors could ascertain that information. Subsequent call attempts were unsuccessful. Of the 655 remaining confirmed operating cannabis retailers collected via government-maintained registries, 622 (95%) confirmed location, 24 (4%) confirmed the location of retailer was incorrect, and 9 (1%) could not be determined either way (Table 1).

Table 1. *Operating Status and Location Accuracy as Determined through Phone Validation Survey of Cannabis Retailer Information Solicited From Web Services and Government-Maintained Registries*

Operating Status	Web Services (n=790) n (%)	Government-Maintained Registries (n=859) n (%)
Confirmed Operating, n (%)	614 (78%)	657 (76%)
Confirmed To Not Be Operating, n (%)	123 (16%)	97 (11%)
Could Not Confirm Either Way, n (%)	53 (7%)	105 (12%)
Location[1]	Web Services (n=611)[2]	Government-Maintained Registries (n=655)[3]
Confirmed Location, n (%)	484 (79%)	622 (95%)
Confirmed Location Is Incorrect, n (%)	108 (18%)	24 (4%)
Could Not Confirm Either Way, n (%)	19 (3%)	9 (1%)

[1]Location is calculated for observations which were confirmed to be operating cannabis retailers
[2]Three observations in the phone survey were excluded due to the call being dropped before the location question was reached and further attempts were not able to reestablish a connection with either a person or a dial-based menu
[3]Two observations in the phone survey were excluded due to the call being dropped before the location question was reached and further attempts were not able to reestablish a connection with either a person or a dial-based menu

Table 2. *State Level Accuracy of Cannabis Retailer Operating Status Information from Webservices and Government-Maintained Registries*

State	Webservice (n=790)					Government-Maintained Registries (n=859)				
	Total Obs. n	Sampled Obs. n	Operating n (%)	NOT Operating n (%)	Cannot Determine n (%)	Total Obs. n	Sampled Obs. n	Operating n (%)	NOT Operating n (%)	Cannot Determine n (%)
AK	94	11	10 (91%)	0 (0%)	1 (9%)	102	10	9 (90%)	1 (10%)	0 (0%)
AZ[1]	225	26	17 (65%)	9 (35%)	0 (0%)	--	--	--	--	--
CA	1525	154	98 (64%)	33 (21%)	23 (15%)	1336	133	105 (79%)	17 (13%)	11 (8%)
CO[2]	1072	104	93 (89%)	7 (7%)	4 (4%)	--	--	--	--	--
CT	24	10	9 (90%)	1 (10%)	0 (0%)	18	9	9 (100%)	0 (0%)	0 (0%)
DC	7	7	6 (86%)	0 (0%)	1 (14%)	6	6	6 (100%)	0 (0%)	0 (0%)
DE	5	5	5 (100%)	0 (0%)	0 (0%)	4	4	4 (100%)	0 (0%)	0 (0%)
FL	213	24	16 (67%)	8 (33%)	0 (0%)	122	9	9 (100%)	0 (0%)	0 (0%)
HI	24	10	6 (60%)	4 (40%)	0 (0%)	7	7	7 (100%)	0 (0%)	0 (0%)
IL	89	9	7 (78%)	1 (11%)	1 (11%)	55	10	9 (90%)	1 (10%)	0 (0%)
LA	8	8	1 (12%)	7 (88%)	0 (0%)	9	9	8 (89%)	1 (11%)	0 (0%)
MA	128	12	12 (100%)	0 (0%)	0 (0%)	21	5	4 (80%)	1 (20%)	0 (0%)
MD	99	11	11 (100%)	0 (0%)	0 (0%)	75	10	10 (100%)	0 (0%)	0 (0%)
ME	129	13	13 (100%)	0 (0%)	0 (0%)	8	8	8 (100%)	0 (0%)	0 (0%)
MI	217	23	14 (61%)	8 (35%)	1 (4%)	70	10	8 (80%)	1 (10%)	1 (10%)
MN	14	10	9 (90%)	1 (10%)	0 (0%)	8	8	8 (100%)	0 (0%)	0 (0%)
MT[1]	115	12	8 (67%)	3 (25%)	1 (8%)	--	--	--	--	--

State	Sampled Obs. n	Correct Location Information n (%)	Incorrect Location Information n (%)	Cannot Determine n (%)	Sampled Obs. n	Correct Location Information n (%)	Incorrect Location Information n (%)	Cannot Determine n (%)
ND	2	2	2 (100%)	0 (0%)	0 (0%)	7	7	7 (100%)	0 (0%)	0 (0%)
NH	6	6	6 (100%)	0 (0%)	0 (0%)	5	5	5 (100%)	0 (0%)	0 (0%)
NJ	14	10	6 (60%)	4 (40%)	0 (0%)	6	6	6 (100%)	0 (0%)	0 (0%)
NM	103	11	6 (55%)	2 (18%)	3 (27%)	98	9	9 (100%)	0 (0%)	0 (0%)
NV	147	9	7 (78%)	1 (11%)	1 (11%)	240	21	17 (81%)	3 (14%)	1 (5%)
NY	122	12	5 (42%)	5 (42%)	2 (17%)	114	10	8 (80%)	1 (10%)	1 (10%)
OH	70	10	8 (80%)	1 (10%)	1 (10%)	18	10	10 (100%)	0 (0%)	0 (0%)
OK	857	87	73 (84%)	8 (9%)	6 (7%)	1492	153	96 (63%)	15 (10%)	42 (27%)
OR	1076	107	91 (85%)	12 (11%)	4 (4%)	2060	205	135 (66%)	36 (18%)	34 (17%)
PA	75	10	9 (90%)	1 (10%)	0 (0%)	45	10	10 (100%)	0 (0%)	0 (0%)
RI	3	3	3 (100%)	0 (0%)	0 (0%)	3	3	3 (100%)	0 (0%)	0 (0%)
VT	8	8	8 (100%)	0 (0%)	0 (0%)	5	5	5 (100%)	0 (0%)	0 (0%)
WA	652	66	55 (83%)	7 (11%)	4 (6%)	1665	177	142 (80%)	20 (11%)	15 (8%)

[1] Registry request was denied after multiple attempts including formal FOIA request.
[2] Data oversight in government-maintained registry resulted in no data from CO.

Table 3. *State Level Accuracy of Cannabis Retailer Location Information from Webservices and Government-Maintained Registries[1]*

State	Webservice (n=611)[2]				Government-Maintained Registries (655)[3]			
	Sampled Obs. n	Correct Location Information n (%)	Incorrect Location Information n (%)	Cannot Determine n (%)	Sampled Obs. n	Correct Location Information n (%)	Incorrect Location Information n (%)	Cannot Determine n (%)
AK	10	10 (100%)	0 (0%)	0 (0%)	9	9 (100%)	0 (0%)	0 (0%)
AZ[4]	17	14 (82%)	1 (6%)	2 (12%)	--	--	--	--
CA	97	79 (81%)	17 (18%)	1 (1%)	105	103 (98%)	1 (1%)	1 (1%)
CO[5]	93	78 (84%)	11 (12%)	4 (4%)	--	--	--	--
CT	9	7 (78%)	2 (22%)	0 (0%)	9	9 (100%)	0 (0%)	0 (0%)
DC	6	3 (50%)	2 (33%)	1 (17%)	5	5 (100%)	0 (0%)	0 (0%)
DE	5	5 (100%)	0 (0%)	0 (0%)	4	4 (100%)	0 (0%)	0 (0%)
FL	16	12 (75%)	3 (19%)	1 (6%)	9	9 (100%)	0 (0%)	0 (0%)
HI	6	6 (100%)	0 (0%)	0 (0%)	7	7 (100%)	0 (0%)	0 (0%)
IL	7	5 (71%)	1 (14%)	1 (14%)	9	7 (78%)	0 (0%)	2 (22%)
LA	1	1 (100%)	0 (0%)	0 (0%)	8	8 (100%)	0 (0%)	0 (0%)
MA	12	12 (100%)	0 (0%)	0 (0%)	4	4 (100%)	0 (0%)	0 (0%)
MD	11	7 (64%)	4 (36%)	0 (0%)	10	10 (100%)	0 (0%)	0 (0%)
ME	13	8 (62%)	5 (38%)	0 (0%)	8	7 (88%)	1 (12%)	0 (0%)
MI	14	9 (64%)	4 (29%)	1 (7%)	8	8 (100%)	0 (0%)	0 (0%)
MN	9	6 (67%)	1 (11%)	2 (22%)	8	8 (100%)	0 (0%)	0 (0%)
MT[4]	8	4 (50%)	3 (38%)	1 (12%)	--	--	--	--

(table continues)

ND	2	2 (100%)	0 (0%)	0 (0%)	7	7 (100%)	0 (0%)	0 (0%)
NH	4	4 (100%)	0 (0%)	0 (0%)	5	3 (60%)	2 (40%)	0 (0%)
NJ	6	6 (100%)	0 (0%)	0 (0%)	6	5 (83%)	1 (17%)	0 (0%)
NM	6	5 (83%)	1 (17%)	0 (0%)	9	9 (100%)	0 (0%)	0 (0%)
NV	7	6 (86%)	1 (14%)	0 (0%)	17	17 (100%)	0 (0%)	0 (0%)
NY	5	4 (80%)	1 (20%)	0 (0%)	8	7 (88%)	1 (12%)	0 (0%)
OH	8	5 (62%)	3 (38%)	0 (0%)	10	9 (90%)	0 (0%)	1 (10%)
OK	73	55 (75%)	18 (25%)	0 (0%)	96	89 (93%)	7 (7%)	0 (0%)
OR	91	78 (86%)	9 (10%)	4 (4%)	134	127 (95%)	6 (4%)	1 (1%)
PA	9	5 (56%)	4 (44%)	0 (0%)	10	7 (70%)	0 (0%)	3 (30%)
RI	3	2 (67%)	1 (33%)	0 (0%)	3	3 (100%)	0 (0%)	0 (0%)
VT	8	3 (38%)	5 (62%)	0 (0%)	5	5 (100%)	0 (0%)	0 (0%)
WA	55	43 (78%)	11 (20%)	1 (2%)	142	136 (96%)	5 (4%)	1 (1%)

[1]Location is calculated for observations which were confirmed to be operating cannabis retailers
[2]Three observations in the phone survey were excluded due to the call being dropped before the location question was reached and further attempts were not able to reestablish a connection with either a person or a dial-based menu
[3]Two observations in the phone survey were excluded due to the call being dropped before the location question was reached and further attempts were not able to reestablish a connection with either a person or a dial-based menu
[4]Registry request was denied after multiple attempts including formal FOIA request.
[5]Data oversight in government-maintained registry resulted in no data from CO.

Table 4. *State-Level Accuracy Acceptability Categories for Cannabis Retailer Operating Status and Location Information from Webservices and Government-Maintained Registries[1]*

State Accuracy Benchmark Categories	Registry >= 80%, Web Service >= 80%	Registry >= 80%, Web Service < 80%		Registry < 80%, Web Service >= 80%		Registry < 80%, Web Service < 80% Accuracy
		Difference Between Data Sources < 10%	Difference Between Data Sources >= 10%	Difference Between Data Sources < 10%	Difference Between Data Sources >= 10%	
Operating Status	Alaska, Connecticut, Washington D.C., Delaware, Massachusetts, Maryland, Maine, Minnesota, North Dakota, New Hampshire, Ohio, Pennsylvania, Rhode Island, Vermont, Washington	Nevada	Florida, Hawaii, Illinois, Louisiana, Michigan, New Jersey, New Mexico, New York		Oklahoma, Oregon	California

Location Information	Alaska, California, Delaware, Hawaii, Louisiana, Massachusetts, North Dakota, New Jersey, New Mexico, Nevada, New York, Oregon, Rhode Island	Connecticut, Washington D.C., Florida, Maryland, Maine, Michigan, Minnesota, Ohio, Oklahoma, Vermont, Washington	New Hampshire	Illinois, Pennsylvania

[1]Colorado, Arizona, Montana, Rhode Island are not categorized due to only having webservice data for those states.

Table 5. *Operating Status Prior to and Post Covid-19 Related Shutdowns determined through Phone Validation Survey of Cannabis Retailer Information Solicited From Web Services and Government-Maintained Registries[1]*

Source	Status Prior to Covid-19 Related Shutdowns				Status Post Covid-19 Related Shutdowns			
	Obs.	Operating	NOT Operating	Cannot Determine	Obs.	Operating	NOT Operating	Cannot Determine
	n	n (%)	n (%)	n (%)	n	n (%)	n (%)	n (%)
Webservices (n=64)	47	39 (83%)	7 (15%)	1 (2%)	17	13 (76%)	2 (12%)	2 (12%)
DC	4	4 (100%)	0 (0%)	0 (0%)	3	2 (67%)	0 (0%)	1 (33%)
MA	11	11 (100%)	0 (0%)	0 (0%)	1	1 (100%)	0 (0%)	0 (0%)
ME	10	10 (100%)	0 (0%)	0 (0%)	3	3 (100%)	0 (0%)	0 (0%)
ND	1	1 (100%)	0 (0%)	0 (0%)	1	1 (100%)	0 (0%)	0 (0%)
NJ	8	5 (62%)	3 (38%)	0 (0%)	2	1 (50%)	1 (50%)	0 (0%)
NY	7	2 (29%)	4 (57%)	1 (14%)	5	3 (60%)	1 (20%)	1 (20%)
VT	6	6 (100%)	0 (0%)	0 (0%)	2	2 (100%)	0 (0%)	0 (0%)
Government-Maintained Registries (n=40)	23	23 (100%)	0 (0%)	0 (0%)	17	14 (82%)	2 (12%)	1 (6%)
DC	4	4 (100%)	--	--	1	1 (100%)	0 (0%)	0 (0%)
MA	0	--	--	--	5	4 (80%)	1 (20%)	0 (0%)
ME	8	8 (100%)	--	--	0	--	--	--
ND	0	--	--	--	1	1 (100%)	0 (0%)	0 (0%)
NJ	6	6 (100%)	--	--	0	--	--	--
NY	0	--	--	--	10	8 (80%)	1 (1%)	1 (1%)
VT	5	5 (100%)	--	--	0	--	--	--

[1]March 3rd, 2020 is used to designate pre- and post- COVID-19 related lockdowns.

By State

The accuracy of operating status and location information within webservice and government-maintained cannabis retailer data was compared across states. Tables 2 and 3 display each states operating status and location information accuracy, respectively, in both webservice and government-maintained registry samples. With respect to operating status, 15 states and DC (59%) observed acceptable accuracy with both data sources (Table 4). In 9 (33%) states, acceptable accuracy was observed for government-maintained registries only and in 2 (7%) acceptable accuracy was observed for webservices only (Table 4). California was the only state where both data sources simultaneously were considered unacceptable (Table 4). In most states where one data source observed acceptable agreement and the other did not, the difference in accuracy exceeded 10% with a range of 10% to 77%. In Nevada, where government-maintained registries achieved acceptable accuracy and webservices did not, this difference was less than 10% (Table 4).

Of the 26 states and DC included in this study, 13 (48%) states demonstrated acceptable accuracy in both data sources for location. In 11 (41%) states, acceptable accuracy was observed from government-maintained registries only and New Hampshire was the only state where acceptable accuracy was observed in web services only. In 2 (7%) states, both data sources observed below acceptable accuracy (Table 4). For all 11 states where acceptable accuracy was achieved in government-maintained registries only, the accuracy disparity between webservice and government-maintained registry derived cannabis retailer information accuracy ranged from 10% to 62% (Table 4).

Pre and Post COVID-19 Related Restrictions

Of the 26 states and DC represented in both webservice and government-maintained registry data sources, only 7 states had state mandated shutdowns (DC, Massachusetts, Maine, North Dakota, New Jersey, New York, Vermont). Table 5 represents operating status accuracy prior to and after COVID-19 related business shutdowns for the 7 states in which cannabis retailers would have been affected. Of the 64 observations in COVID-19 shutdown affected states in the webservice sample,

47 (73%) observations were completed prior to COVID-19 related shutdowns and 17 (27%) were completed post COVID-19 related shutdowns. Of the 47 observations completed prior to COVID-19 related shutdowns, 39 (83%) were confirmed operating and of the 17 observations completed post COVID-19 related shutdowns, 13 (76%) were confirmed operating. Similarly, Of the 40 observations in COVID-19 shutdown affected states in the government-maintained registry sample, 23 (58%) observations were completed prior completed prior to COVID-19 related shutdowns and 17 (42%) were completed post COVID-19 related shutdowns. Of the 23 observations completed prior to COVID-19 related shutdowns, 23 (100%) were confirmed operating and of the 17 observations completed post COVID-19 related shutdowns, 14 (82%) were confirmed operating.

DISCUSSION

As the sale of cannabis for medical and adult use is legalized in more and more US states, there is a growing need to build a nuanced understanding of the cannabis retailer data sources used to generate knowledge about the impacts of access to cannabis on health. In this study we used a standardized phone survey to compare two cannabis retailer data sources (webservice and government-maintained registry) with respect to the accuracy of operating status and location information across 26 states and DC legalizing the sale of cannabis for medical or adult use in 2019. Additionally, the potential impact of state level COVID-19 related restrictions on ascertainment of phone survey data was evaluated on a state-by-state basis. Our findings demonstrate that two main features of cannabis retailer data sources inform research practices and interpretation in this domain: data source and geographic location. Furthermore, state level COVID-19 restrictions were uncommon in states with cannabis legalization policies and in the states which had COVID-19 restrictions in place during the study period, few showed any negative effects on the capability of phone surveys to ascertain operating status and location information in both data sources.

Overall, we found that both data sources observed similar accuracy with respect to the proportion of retailers confirmed as open and

operating (78% in webservice, 76% in government-maintained registries). However, differences were observed between the data sources with respect to the proportion of confirmed non-cannabis retailers and the proportion of unconfirmable observations. Roughly 4% more observations sampled from webservices were confirmed as non-cannabis retailers (e.g., a grocery store) as compared to the government-maintained registry data source. The webservice data cleaning procedures in this study determined that 16% of the observations were self-advertised non-cannabis retailers and thus removed from the deduplicated webservice data prior to conducting the phone surveys. Another 16% of the observations in the webservice sample were confirmed as non-cannabis retailers via phone survey. These findings align with research by Pedersen et al. 2020 which suggest that webservice data retailer lists tend to overestimate the number of cannabis retailers.[24] The results of our study support this claim. We found that webservice based data sources, such as Weedmaps, often contained advertisements for retailers that did not sell cannabis, but advertise on the platform (e.g., dental offices). This suggests that webservice based approaches to generating cannabis retailer lists should invest resources into culling the relatively large number of identifiable non-cannabis retailers present in these data sources.

Roughly half the observations sampled from webservices were unconfirmable compared to government-maintained registries. Government-maintained registries often included establishment information associated with a corporate office, manager, or business owner. It is possible these individuals were less likely to answer a call from an unknown number than a retail store, leaving these observations unable to be confirmed as cannabis or non-cannabis retailers. This suggests, government-maintained registry information may not be suitably verifiable through phone-based methods alone due to the contact information not necessarily corresponding to the brick-and-mortar storefront of the retailer.

Among confirmed operating cannabis retailers, the accuracy of location information derived from government-maintained registries was very high (95%), while the accuracy of location information derived from web services is bordering on acceptable (79%). While prior work has characterized the accuracy of license and operating status information of cannabis retailer data

sources, to our knowledge, this work is the first to quantify the accuracy of location information. Due to the high accuracy of location information in government-maintained registries, these data sources may be more suitable for research in which specific location is a fundamental component (e.g., spatial analysis). In comparison, webservice based cannabis retailer lists may prove useful when conducting analyses across numerous states, as the data will be uniform across state lines. Additionally, webservice based data sources may be acceptable when the outcome of interest is a low-resolution density measure, such as the number of cannabis retailers in a state or county.

The accuracy of location and operating status information deviated on a state-by-state basis. In most states, both webservice and government-maintained registry-based data sources demonstrated acceptable accuracy or government-maintained registry performed better with respect to operating status. In California, where numerous research studies examining links between access to cannabis retailers and health outcomes have been conducted, (Bachhuber et al., 2014; Freisthler et al., 2013, 2016; Freisthler & Gruenewald, 2014; Mair et al., 2015; Morrison et al., 2014; Shi et al., 2018a; Thomas & Freisthler, 2016; Unger et al., 2020) neither webservice nor government-maintained registry based data sources were more than 80% accurate with respect to operating status. We were able confirm 64% of our webservice-based observations as open across California as a whole which aligns with previous work by Pedersen et al. (2018) which found that 54% of cannabis retailers were confirmed operating using information derived from web-services in Los Angeles (Pedersen et al., 2018). Similarly, we confirmed 79% of observations sampled from government-maintained registries which aligns with Cao et al. who found that 77% of observations taken from government-maintained registries in California were confirmed operating (Cao et al., 2020). The observed low accuracy may be due to the high turnover of cannabis retailers in California, which has been suggested to be the result of Proposition D, a regulatory measure which set zoning limitations and capped the number of cannabis retailers at 135 (Thomas & Freisthler, 2017). Due to California being a focal point of access to cannabis research, these findings add to a body of literature suggesting that generalizability of findings derived from California

may be limited. Oklahoma and Oregon were the only two states where webservices provided superior accuracy with respect to operating status and exceeded the 80% benchmark. The differences observed in these states may be due to a variety of reasons including registry and licensure procedures (e.g., in Oklahoma, the online registry auto-populates as license information is filled out, (*List of Licensed Businesses*, 2022) which may be potentially prone to errors) or potential low regulation in rural areas of these states (Fertig, 2022). These findings suggest that in these two states, webservice based approaches to generating cannabis retailer lists may be more suitable.

Our study used phone validation methods to verify the operating status and location of cannabis retailers. Previous research seeking to validate cannabis retailer lists has utilized physical site visits, virtual site visits (via Google Street View), social media analysis, phone surveys, and government-maintained registry information (Cao et al., 2020; Pedersen et al., 2018, 2020). Previous research suggests a mix of online and observational methods are best practices for ascertaining reliable cannabis retailer information (Pedersen et al., 2018, 2020). This is due to an increase in accuracy of determining operating status provided by combining both methods. However, these studies were limited to California. Additionally, Pedersen et al. (2020) noted the infeasibility of using ground-truthing methods such as walkabouts to verify operating status of cannabis retailers, even on a smaller scale (Pedersen et al., 2020). Similar to Cao et al., this study utilized phone survey methods to verify cannabis retailers. However, we verified roughly three quarters of the potential retailers as operating sampled across 26 states and DC between November 2019 and June 2020. This extends Cao et al.'s work and suggests that phone validation may offer a feasible approach for verifying the accuracy of retailer lists, even on a large geographic scale.

Strengths & Limitations

This study is a first of its kind comparison of the two widely utilized cannabis retailer data sources across on a national scale, shedding light on the considerations of each data source in different geographic locations. However, this study includes some limitations. Notably, our research identified variability in the acceptability of data sources within and across states. We are unable to make conclusions about the performance of registry data in Arizona, Colorado, and Montana, as these data were not included in our sample. Additionally, it is possible that COVID-19 and the mandatory closure of cannabis retailers in some states may have impacted our ability to verify the operating status of retailers contacted after March 13th, 2020. However, the impact of COVID-19 related shutdowns seems minimal. Firstly, only 6 states and DC experienced any government-mandated cannabis retailer related shutdowns out of the 26 states and DC represented in the study. Secondly, these states only contributed 64 of the total 790 observations in the webservice sample (8%) and 40 of the total 859 observations in the government-maintained registry sample (5%). Thirdly, in many states in both samples amongst the 6 states and DC, the operating status accuracy did not change after COVID-19 related shutdowns. In the webservice sample, 3 states operating status changed after COVID-19 related lockdowns. In one of those states (New York), the operating status accuracy improved from 29% to 60%. Similarly, in the government-maintained registry sample, DC was the only geographic area to have phone surveys conducted before and after COVID-19 related shutdowns and observed 100% accuracy before and after. The other 6 states represented in the government-maintained registry sample which experienced government-mandated COVID-19 related shutdowns all observed acceptable accuracy. Furthermore, the phone survey protocol instructs surveyors to re-call unconfirmed observations throughout the project duration which means that any observations which may have been potentially compromised due to COVID-19 related shutdowns would have the entire project duration to potentially be verified. Finally, if COVID-19 related shutdowns significantly impacted confirmation capabilities via the phone survey, one would expect to see a bias towards non-confirmation. However, in the 6 states and DC which had mandatory shutdowns, operating status accuracy is grossly acceptable with many states at 100% accuracy in both data sources.

This study also has numerous strengths. This study fulfills a critical literature gap – characterizing the accuracy of the two most widely used data sources for cannabis retailers lays the groundwork for future work in this area in addition to researchers and policy makers alike utilizing

this work as a point of reference to interpret the results of other access to cannabis retailer studies. This study used phone surveys to verify over 1,500 cannabis retailers across 26 states and DC, as opposed to relying on government-maintained registries as the ground truth. This allowed us to not assume one of the data source's veracity and instead study this empirically.

Conclusions

Data source and geographic location help to contextualize past research and inform future research using cannabis retailer lists. With respect to data source, research where specific cannabis retailer locations are a key element may prefer to use information extracted from government-maintained registries. However, research which aims to estimate the number of cannabis retailers in a broad geographic region (such as a large state or across multiple states) may prefer to use data extracted from webservices as this data will be uniform across geographic regions and be easier to filter down to only operating cannabis retailers at a large scale. Furthermore, location may also further inform choices with respect to methodology and data source. Using a carefully cleaned webservice-based data source may provide a more accurate depiction of cannabis retailers in Oregon and Oklahoma, whereas other states may be better suited utilizing data from government-maintained registries. While prior work has developed methods for assessing the accuracy of operating status within a webservice-derived cannabis retailer dataset using a variety of observational methods (e.g., social media, site visits), these methods also include using government-maintained registry information as a verification method. Given that in our study we observed an overall similar rate of accuracy between both datasets, it may not be prudent to use one data source as a ground-truth for the other without considering the state-specific accuracy of each source. Instead, future research should aim to build on these findings and the methods proposed by Pedersen et al (2018, 2020). to estimate the degree to which webservices may be overestimating the number of cannabis retailers (either through unlicensed retailers or through non-cannabis retailers) and the degree to which government-maintained registries may be underestimating the true population of cannabis retailers (Pedersen et al., 2018, 2020).

REFERENCES

Amiri, S., Monsivais, P., McDonell, M. G., & Amram, O. (2019). Availability of licensed cannabis businesses in relation to area deprivation in Washington state: A spatiotemporal analysis of cannabis business presence between 2014 and 2017. *Drug and Alcohol Review, 38*(7), 790–797. https://doi.org/10.1111/dar.12987

Bachhuber, M. A., Saloner, B., Cunningham, C. O., & Barry, C. L. (2014). Medical cannabis laws and opioid analgesic overdose mortality in the United States, 1999-2010. *JAMA Internal Medicine, 174*(10), 1668. https://doi.org/10.1001/jamainternmed.2014.4005

Burkhardt, J., & Flyr, M. (2019). The effect of marijuana dispensary openings on housing prices. *Contemporary Economic Policy, 37*(3), 462–475. https://doi.org/10.1111/coep.12414

Cao, Y., Carrillo, A. S., Jankowska, M. M., & Shi, Y. (2020). Validation of secondary data sources for enumerating marijuana dispensaries in a state commercializing marijuana. *Drug and Alcohol Dependence, 215*, 108183. https://doi.org/10.1016/j.drugalcdep.2020.108183

Chapman, S. A., Spetz, J., Lin, J., Chan, K., & Schmidt, L. A. (2016). Capturing heterogeneity in medical marijuana policies: a taxonomy of regulatory regimes across the United States. *Substance Use & Misuse, 51*(9), 1174–1184. https://doi.org/10.3109/10826084.2016.1160932

Cheng Cheng, Walt Mayer, & Yanling Mayer. (2018). The effect of legalizing retail marijuana on housing values: evidence from Colorado. *Economic Inquiry*, 56(3), 1585–1601. https://doi.org/10.1111/ecin.12556

Chong, W. W.-Y., Acar, Z. I., West, M. L., & Wong, F. (2022). A scoping review on the medical and recreational use of cannabis during the COVID-19 pandemic. *Cannabis and Cannabinoid Research, 7*(5), 591–602. https://doi.org/10.1089/can.2021.0054

Conklin, J., & Diop, M. (2017). Contact high: the external effects of retail marijuana establishments on house prices. *Real Estate Economics*, 48(1), 135–173. https://dx.doi.org/10.2139/ssrn.2885017

Contreras, C. (2017). A block-level analysis of medical marijuana dispensaries and crime in the city of Los Angeles. *Justice Quarterly, 34*(6),

1069–1095.
https://doi.org/10.1080/07418825.2016.1270346

COVID-19 Disaster Declarations | FEMA.gov.
(2021, August 20).
https://www.fema.gov/disaster/coronavirus/disaster-declarations

Declaring a National Emergency Concerning the Novel Coronavirus Disease (COVID-19) Outbreak. (2020, March 18). Federal Register.
https://www.federalregister.gov/documents/2020/03/18/2020-05794/declaring-a-national-emergency-concerning-the-novel-coronavirus-disease-covid-19-outbreak

Fertig, N. (2022, January 21). *'Talk About Clusterf---': Why Legal Weed Didn't Kill Oregon's Black Market.* POLITICO.
https://www.politico.com/news/magazine/2022/01/14/oregon-marijuana-legalization-black-market-enforcement-527012

Francesca Ortegren. (2021, July 12). 2021 Study: how legalizing recreational marijuana impacts home values. *Real Estate Witch.*
https://www.realestatewitch.com/marijuana-study-2021/

Freisthler, B., Gaidus, A., Tam, C., Ponicki, W. R., & Gruenewald, P. J. (2017). From medical to recreational marijuana sales: marijuana outlets and crime in an era of changing marijuana legislation. *The Journal of Primary Prevention*, *38*(3), 249–263. https://doi.org/10.1007/s10935-017-0472-9

Freisthler, B., & Gruenewald, P. J. (2014). Examining the relationship between the physical availability of medical marijuana and marijuana use across fifty California cities. *Drug and Alcohol Dependence*, *143*(1), 244–250.
https://doi.org/10.1016/j.drugalcdep.2014.07.036

Freisthler, B., Kepple, N. J., Sims, R., & Martin, S. E. (2013). Evaluating medical marijuana dispensary policies: Spatial methods for the study of environmentally-based interventions. *American Journal of Community Psychology*, *51*(1–2), 278–288. https://doi.org/10.1007/s10464-012-9542-6

Freisthler, B., Ponicki, W. R., Gaidus, A., & Gruenewald, P. J. (2016). A micro-temporal geospatial analysis of medical marijuana dispensaries and crime in Long Beach, California. *Addiction (Abingdon, England)*, *111*(6), 1027–1035.
https://doi.org/10.1111/add.13301

Julia Raifman, Kristen Nocka, David Jones, Jacob Bor, Sarah Ketchen Lipson, Jonathan Jay, & Philip Chan. (2020). *COVID-19 US state policy database.*

Klieger, S. B., Gutman, A., Allen, L., Pacula, R. L., Ibrahim, J. K., & Burris, S. (2017). Mapping medical marijuana: State laws regulating patients, product safety, supply chains and dispensaries, 2017. *Addiction (Abingdon, England)*, *112*(12), 2206–2216.
https://doi.org/10.1111/add.13910

Lankenau, S. E., Tabb, L. P., Kioumarsi, A., Ataiants, J., Iverson, E., & Wong, C. F. (2019). Density of medical marijuana dispensaries and current marijuana use among young adult marijuana users in Los Angeles. *Substance Use & Misuse*, *54*(11), 1862–1874.
https://doi.org/10.1080/10826084.2019.1618332

Legal Science. (2017a, February 1). PDAPS - Medical Marijuana Dispensaries.
https://pdaps.org/datasets/dispensaries-medical-marijuana-1501611712

Legal Science. (2017b, February 1). PDAPS - Recreational Marijuana Laws.
https://pdaps.org/datasets/recreational-marijuana-laws

Lipperman-Kreda, S., Lee, J. P., Morrison, C., & Freisthler, B. (2014). Availability of tobacco products associated with use of marijuana cigars (blunts). *Drug and Alcohol Dependence*, *134*(1), 337.
https://doi.org/10.1016/j.drugalcdep.2013.10.022

List of Licensed Businesses. (2022). Oklahoma Medical Marijuana Authority.
https://oklahoma.gov/omma/businesses/list-of-businesses.html

MacKillop, J., Cooper, A., & Costello, J. (2021). National retail sales of alcohol and cannabis during the covid-19 pandemic in canada. *JAMA Network Open*, *4*(11), e2133076.
https://doi.org/10.1001/jamanetworkopen.2021.33076

Mair, C., Freisthler, B., Ponicki, W. R., & Gaidus, A. (2015). The impacts of marijuana dispensary density and neighborhood ecology on marijuana abuse and dependence. *Drug and Alcohol Dependence*, *154*, 111–116.
https://doi.org/10.1016/j.drugalcdep.2015.06.019

Mair, C., Sumetsky, N., Kranich, C., & Freisthler, B. (2021). Availability of medical cannabis dispensaries and cannabis abuse/dependence-related hospitalizations in California. *Addiction*,

116(7), 1908–1913.
https://doi.org/10.1111/add.15420

Morrison, C., Gruenewald, P. J., Freisthler, B., Ponicki, W. R., & Remer, L. G. (2014). The economic geography of medical cannabis dispensaries in California. *International Journal of Drug Policy*, *25*(3), 508–515. https://doi.org/10.1016/j.drugpo.2013.12.009

NORML Foundation. (2022). *State Laws—NORML.* https://norml.org/laws/

NumPy. (2022). https://numpy.org/

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, *19*, 1609406919899220. https://doi.org/10.1177/1609406919899220

pandas—Python Data Analysis Library. (2022). https://pandas.pydata.org/

Pedersen, E. R., Firth, C., Parker, J., Shih, R. A., Davenport, S., Rodriguez, A., Dunbar, M. S., Kraus, L., Tucker, J. S., & D'Amico, E. J. (2020). Locating medical and recreational cannabis outlets for research purposes: online methods and observational study. *Journal of Medical Internet Research*, *22*(2). https://doi.org/10.2196/16853

Pedersen, E. R., Zander-Cotugno, M., Shih, R. A., Tucker, J. S., Dunbar, M. S., & D'Amico, E. J. (2018). Online methods for locating medical marijuana dispensaries: practical considerations for future research. *cannabis (Research Society on Marijuana)*, *1*(2), 22. https://doi.org/10.26828/cannabis.2018.02.003

Schauer, G. L., Dilley, J. A., Roehler, D. R., Sheehy, T. J., Filley, J. R., Broschart, S. C., Holland, K. M., Baldwin, G. T., Holmes-Chavez, A. K., & Hoots, B. E. (2021). Cannabis sales increases during COVID-19: Findings from Alaska, Colorado, Oregon, and Washington. *The International Journal on Drug Policy*, *98*, 103384. https://doi.org/10.1016/j.drugpo.2021.103384

Shi, Y., Cummins, S. E., & Zhu, S.-H. (2018). Medical marijuana availability, price, and product variety, and adolescents' marijuana use. *Journal of Adolescent Health*, *63*(1), 88–93. https://doi.org/10.1016/j.jadohealth.2018.01.008

Shi, Y., Meseck, K., & Jankowska, M. M. (2016). Availability of medical and recreational marijuana stores and neighborhood characteristics in Colorado. *Journal of Addiction*, *2016*. https://doi.org/10.1155/2016/7193740

Shih, R. A., Rodriguez, A., Parast, L., Pedersen, E. R., Tucker, J. S., Troxel, W. M., Kraus, L., Davis, J. P., & D'Amico, E. J. (2019). Associations between young adult marijuana outcomes and availability of medical marijuana dispensaries and storefront signage. *Addiction*, *114*(12), 2162–2170. https://doi.org/10.1111/add.14711

Tabb, L. P., Fillmore, C., & Melly, S. (2018). Location, location, location: Assessing the spatial patterning between marijuana licenses, alcohol outlets and neighborhood characteristics within Washington state. *Drug and Alcohol Dependence*, *185*, 214–218. https://doi.org/10.1016/j.drugalcdep.2018.01.004

Thomas, C., & Freisthler, B. (2016). Examining the locations of medical marijuana dispensaries in Los Angeles. *Drug and Alcohol Review*, *35*(3), 334–337. https://doi.org/10.1111/dar.12325

Thomas, C., & Freisthler, B. (2017). Evaluating the change in medical marijuana dispensary locations in Los Angeles following the passage of local legislation. *The Journal of Primary Prevention*, *38*(3), 265–277. https://doi.org/10.1007/s10935-017-0473-8

Unger, J. B., Vos, R. O., Wu, J. S., Hardaway, K., Sarain, A. Y. L., Soto, D. W., Rogers, C., & Steinberg, J. (2020). Locations of licensed and unlicensed cannabis retailers in California: A threat to health equity? *Preventive Medicine Reports*, *19*. https://doi.org/10.1016/j.pmedr.2020.101165