Early Lessons from Data-Driven Justice Pilot Sites

Every year, more than 11 million[1] people move through America's 3,100 local jails, many on low-level, nonviolent offenses, costing local governments approximately $22 billion a year in jail costs alone.[2] In local jails, 64 percent of people suffer from mental illness[3] and 68 percent have a substance use disorder.[4] With three times more people with mental health problems and other co-occurring issues housed in jails or prisons than in mental health treatment facilities[5], communities across the country face the reality that local police, whose involvement often leads to arrest and incarceration, have become the primary responders to people experiencing behavioral health crises.

Even more challenging is the small number of highly-vulnerable people, called frequent utilizers, who cycle repeatedly not just through local jails, but also through hospital emergency rooms, shelters, and other public systems due to underlying behavioral health, housing, and social service needs. The result is fragmented, uncoordinated, and costly care that fails to break this endless cycle or lead to improved stability for these highly vulnerable individuals. To face these challenges, a bipartisan coalition of 146 jurisdictions, representing more than 94 million people, has committed to tackling some of the root-causes driving, and keeping, people in our jails and emergency rooms by joining the Data Driven Justice (DDJ) initiative.

With the support of Arnold Ventures (AV), DDJ communities are using data-driven strategies to divert frequent utilizers away from the criminal justice and emergency health systems and into community-based treatment and services. The Philanthropy's funding to-date includes support to the National Association of Counties to provide assistance to the network of DDJ communities, along with a commitment of more than $10 million to fund evaluations of promising frequent utilizer interventions and emergency response programs.

AV is also funding three DDJ pilot sites in Middlesex County, MA, City of Long Beach, CA, and Johnson County, IA, to launch and test new methods of coordination between police, hospitals and social services. Pilot sites were selected because of their community's commitment to DDJ, demonstrated through strong leadership buy-in and a vision for how they could better leverage data to design diversion strategies, programs, and policies to improve outcomes for their local frequent utilizer population. Prior to AV's support, each site had little to no capacity to share, integrate, and analyze data. The DDJ sites each implemented the following steps to organize their work: (1) building stakeholder consensus; (2) understanding the people you hope to serve; (3) defining resource ecosystem and identifying gaps in service; (4) leveraging data to develop better policies, programs, and multi-system approaches to break the cycle of arrest and incarceration.

DDJ at a Glance:

Who: 146 communities nationwide

What: The DDJ is a data-centered multi-system collaboration that aims to improve outcomes for vulnerable populations by providing communities with two critical tools to drive reform: (1) data integration technology; and, (2) a playbook to guide implementation.

To Learn More:
visit https://www.naco.org/resources/signature-projects/data-driven-justice

How
1. Build stakeholder consensus;
2. Understand the people you hope to serve;
3. Establish a framework for data governance;
4. Document resources availability and identify gaps in service;
5. Build continual data use and information sharing into cross-system organizational operations and policy discussions.

Just one year into the pilot program, all three communities have made significant strides integrating data such as 911 calls for service, jail bookings, EMS transports, emergency department visits, and homeless shelter records. The valuable insights gained from this progress has already allowed the pilot sites to design new programs and policies, support stakeholders, and optimize the operations of agencies working to address the issue of frequent utilization.

The purpose of this brief is to share early lessons learned from integrating and analyzing cross-system data. While these lessons primarily rely on lessons from the DDJ sites, we will sometimes use case examples from the literature and/or from Camden, New Jersey, whose work with Arnold Ventures predated the DDJ initiative, and Allegheny County, PA, who is part of the broader DDJ network.

Lesson 1: Integrating administrative data is challenging and requires time and support. But it's worth it.

Data-Driven Justice efforts envision collaboration and data integration across criminal justice, health, and social services but many DDJ communities, like communities across the country, struggle to accomplish this vision. Integrating data is challenging. It requires, perhaps above all, political will to collaborate and integrate, but even with strong leadership, data integration can be stalled or even halted, in discussions about the legality, security, and cost of integration. An enduring lesson from the DDJ work is that communities need support to integrate and share data. Communities can benefit from clear practice guides for assembling and building trust among stakeholders and establishing the policies, procedures, and technology necessary to use integrated data to drive their collective efforts. While examples of these sorts of guides exist broadly, there are no products designed specifically for crisis system improvement efforts and developing these products would help advance the field. And, if jurisdictions are starting from scratch, we should acknowledge that data integration work is a project of its own and requires time, dedicated staff, and support to make it happen.

Getting started: The text box provides some key questions to answer as you bring together DDJ stakeholders to collaborate to examine frequent utilizers.

The lesson from these three DDJ sites is to start with data you have access to, have strong executive leadership for system and data integration, be clear about your goals (a goal to reduce frequent utilizers of jail services will require different data sets than a goal to improve crisis response to people living on the street), and seek data sets to help understand your population of interest.

SITE EXAMPLES

City of Long Beach, California: Leaders began their efforts towards data integration by convening key stakeholders including the Police Department, Fire Department, Health Department, and the City Prosecutor's Office to identify administrative data sets that could help them understand frequent utilizers of crisis services. Then, stakeholders went through a process of manually matching data and reporting back monthly at an established Multi-disciplinary Team (MDT) meeting. This process led the City of Long Beach to develop a deeper understanding of their own administrative data sets, a stronger sense of ownership for the project goals, and more buy-in to pursue automated data integration.

Middlesex County, Massachusetts: Augmenting quantitative analysis with case studies can help leaders understand the need to integrate data and can also humanize the project and focus the DDJ partners. One of the extreme frequent utilizers documented in Middlesex County is "Mary." Mary is a homeless woman with chronic alcoholism. She had 85 engagements with the police, 25 of which were arrests, and spent 872 days in jail. She required over 1,000 EMS transports over a 13-year period, amounting to over $600,000 in costs to Emergency Medical Services (EMS) alone. Understanding the need to improve their response to "Mary," leaders sought to build key data partnerships.

Example of an extreme frequent utilizer. "Mary" is a homeless woman with chronic alcoholism.

85	25	872	1,000	$600K
engagements with the police	arrests	days in jail	EMS transports over a 13-year period	costs to Emergency Medical Services

Lesson 2: Fully mine the data you have and examine all criminal justice data sets.

While the slow work to integrate cross-systems data sets at scale is happening, communities can and should start with the data they have in hand.[6] Mining the dispatch, police contact, or jail booking data is a necessary step in the process. Each criminal justice data set provides different information and each has its own strengths and limitations.

Dispatch data is the most encompassing criminal justice data set. It provides information on incidents and people never arrested by the police but who nonetheless may be frequent utilizers of crisis services. These data are not without their challenges. The first is that in many sites the data will reflect calls for service where first responders are dispatched to an address or location, not necessarily an individual, making it difficult to identify frequent utilizers. Second, by the very nature of this system, the incidents collected in it are unverified. A caller may suspect a drug overdose, call it in as such and it may turn out to be a heart attack. Lastly, frequent callers or frequent locations may turn out to be shelter administrators or social services agencies and thus say more about the location or the occupation of the caller than help to identify frequent utilizers. In short, dispatch data are critical to identify frequent utilizers who are un- or under-detected in other criminal justice data sets, but should be used with caution.

Police data can include any person level contact with the police or only data where there has been an arrest or citation. The more expansive the data set, the more likely it is to contain frequent utilizers who do not have substantial arrest or booking histories and may be the low-level offender that DDJ efforts aim to support. Sites should request the most expansive data set possible.

Jail booking data include the population booked in the jail awaiting disposition on a charge, a probation or parole violation hearing, or serving a short sentence. Jail bookings tend to include a population with more serious charges or more significant criminal history than the preceding criminal justice data sets and may have fewer very low-level offenders, but this depends on the jurisdiction. Focusing on this population may be important to policymakers because of the increased cost for housing people in the jail, beyond the already high costs of repeated arrests and 911 dispatches.

Ideally, all the criminal justice data sets are used to inform DDJ efforts. For example, once you have developed your list of frequent utilizers of the jail, if you receive 911 dispatch data you can use these data to see how much more likely your list of frequent utilizers is to have a 911 call than the broader group. It should hold that your frequent utilizers are (much) more likely to have 911 calls for service. This helps you to feel more confident in your analysis and helps demonstrate to your partners that effective interventions for this group can impact their operations. You could also use these data to further reduce your list of frequent jail utilizers by focusing on those with both high utilization of the jail AND high utilization of 911. And, you could look at the group with high 911 dispatches but low/no arrests or jail bookings. This may be the population that most fits your community's perception of frequent utilizer. With consistent effort, sites can bring in additional data sets to further specify their frequent utilizer population.

Once the criminal justice data sets are available, they need to be analyzed. Below are some key steps and questions to answer when analyzing the criminal justice data sets.

KEY STEPS IN THE ANALYSIS:

1) Describe the full (dispatch, arrest, booking) file:

- Who (gender, race, ethnicity, age, residence)?
- What (charge seriousness, type, arresting agency, reason for dispatch)?
- How long (length of stay, release type)?
- How many (choose a threshold to identify frequent utilizers)?

2) Describe the frequent utilizers in the file:

- Who (gender, race, ethnicity, age, residence)?
- What (charge seriousness, type of charge, reason for dispatch)?
- How long (length of stay)?

3) Compare the two groups

- Look for demographic shifts, offense shifts, etc.

4) Write case studies

- Pull a few cases to examine in more detail and humanize the issues. While data may not be integrated at scale, agency partners might be willing to share data on a case by case basis

Once you have analyzed your starting data sets, a key question is what other data can inform your efforts. These data need not be administrative data, it could be intelligence and insights gained from stakeholders who can review your list of identified frequent utilizers and cross reference from their work or official databases and help understand if you are on the right track.

SITE EXAMPLES

Middlesex County, Massachusetts: The first job in Middlesex County was to bring together data across 53 police departments. The team worked to standardize data and variables requested, established validation procedures, and made recommendations to improve police data collection. For example, analysts in Middlesex County made recommendations on how to improve and standardize data for people in need of mental health crisis services across all police departments.

Johnson County, Iowa: When analysts added the dispatch data to their analysis of the jail bookings, they were able to understand how individuals interact with the crisis systems, even in situations where there are no criminal charges filed. This integration enhanced their understanding of the DDJ population. But there were limitations. About 30% of dispatch incidents did not identify any individual. Another 30% of the incidents that had names, were not names of people, but of businesses. This limited the utility of the file. Twenty people were identified as particularly high-contact, the top five of which had over 1,000 contacts amongst them. However, one of the people on the list was an elected official and respected local business owner. "EO" runs a business in the medical field, which includes in-home care. The type of services "EO" provides, combined with the location of the services and client characteristics, results in frequent calls for medical assistance or ambulance transportation for their patients and resulted in "EO" being on the top 20 list. This was not a result of "EO's" misuse of any system or any system failing to meet "EO's" needs and was a false positive on the list. This, and similar type confusions can happen with the dispatch data, so it is important to use it carefully.

Lesson 3: You don't need 10 years of data to support DDJ. Recent data is better for operational purposes.

Common wisdom suggests that we need years, even decades of data to inform our work. Many take pride in their analyses of the last five or 10 years of data. While there are obvious uses for long term data holdings, particularly when doing longitudinal analysis, operational work like that required for Data-Driven Justice benefit more from the most recent 18 to 24 months of data. Reducing to more recent data has three benefits. First, it reduces the number of people

with whom to intervene (since most jurisdictions only have a limited budget for any intervention). Second, people in a more recent time-period are more likely to appear (be booked, arrested, or require emergency transports) again in the future, when a targeted intervention can be offered. Lastly, some agencies don't feel comfortable or cannot extract historic data, so it might be easier to obtain recent data. In short, sites can feel free to request the last 10 years of data for their initial analysis but should consider only using the most recent 18-24 months of data for operational purposes.

Johnson County, Iowa: Johnson County had access to over a decade of jail booking data. When they analyzed jail bookings 2014-2017, only six percent of them had a booking in 2018 (the year that they would want to offer the housing first intervention). Many more of the frequent utilizers (people with five or more bookings) during this time period experienced a booking in 2018 (44%), but the results are better if the analysis is limited to those with bookings in the most recent two years (2016-2017), where they find that 17% of all those booked during the period had another booking in 2018, and 57% of the frequent utilizers for 2016-2017 were booked in 2018. If you want to intervene in 2018, you should prioritize frequent utilizers with bookings in the most recent period 18-24 months.

City of Long Beach, California: The same results were found in Long Beach, California. When they used five years of data, only 11% of people will have an arrest in the most recent 12 months but the results are more relevant for operational purposes using the most recent 18 months.

Lesson 4: Frequent utilizers require more system coordination because they are (much) more likely to be involved with multiple agencies.

Data supports what many stakeholders have come to know anecdotally: The typical frequent utilizer is someone who is experiencing homelessness and may be suffering from mental health, substance use disorder, and/or chronic health conditions. We know we need to be working better together across these crisis response systems. Perhaps a bit surprising, though, is the extent to which frequent utilizers are also encountering multiple law enforcement agencies. So, in addition to needing to work more collaboratively with health and social services agencies, these findings call for an examination of the frequency of contacts across local law enforcement agencies, and for more cooperation across police departments and with other first responders for frequent utilizers.

To examine the other health and social services systems that frequent utilizers might encounter, Allegheny County pulled key cross-systems data for two criminal justice data sets: (1) people issued an arrest or citation that was filed by law enforcement in Allegheny County during the years 2016 to 2017; (2) people booked into the Allegheny County Jail during years 2016 to 2017. The health and social services data examined included physical health (emergency department and inpatient stays), behavioral health (all mental health services, mental health crisis services, substance use disorder services), homelessness services and supports, receipt of public benefits, and child welfare involvement (as a parent).[7] More detail on these data sets can be found in Appendix B.

The analysis examined the likelihood that people with one, two, three, four, or five or more arrests would have these service involvements in the year prior to the first arrest in the period. Complimentary analyses were also conducted examining involvement in the year post arrest. The same analyses were conducted for people with one, two, three, four,

or five or more jail bookings and can be found, along with additional details, in Appendix B.

The results are summarized in the Table 1 below. While the most likely service to be accessed by arrestees is public benefits, because it is so common, it may not say very much about the differences between someone who was arrested once in the two-year period and someone who was arrested five or more times during the same period. It could, however, be a key intervention opportunity since so many arrestees access the benefit. Similarly, use of emergency departments is very common. The services where there is the biggest difference between those with five or more arrests and those with only one arrest are mental health crisis services (frequent arrestees are 3.5 times more likely to access the service than people with one arrest), substance use services, and homeless services.

Table 1: People arrested in Allegheny County in 2016-2017 and their selected health and social services involvement during the 12 months prior to the incident.

Number of Arrests	Emergency Department Encounter	Physical Health Inpatient	Mental Health Services	Mental Health Crisis	Substance Use Services	Homeless Services	Public Benefit Receipt	Child Welfare Parent
1	28%	6%	19%	4%	11%	2%	53%	5%
2	38%	9%	28%	7%	19%	3%	66%	7%
3	43%	10%	34%	9%	24%	3%	74%	10%
4	49%	13%	43%	11%	29%	5%	80%	11%
5 or more	53%	14%	47%	14%	35%	6%	83%	12%
Average	32%	8%	24%	6%	15%	2%	58%	6%
Ratio of 5 or more arrests to one arrest	1.9	2.3	2.5	3.5	3.2	3.0	1.6	2.4

SITE EXAMPLES

Johnson County Iowa: A significant percentage of the frequent jail utilizers have been booked by more than one of Johnson County's five law enforcement agencies. Only 7% of frequent jail users had all of their bookings initiated by one agency; 73% had two or three different agencies making their bookings; and 20% had four or more agencies making bookings for the individual. This finding speaks to the need to coordinate across police departments.

Literature: Across the jurisdictions studied by the City University of New York, the average percentage of frequent utilizers (of the jail) with mental health conditions was 40% compared to just 16% for the rest of the jail population.[8] The results were similar for substance use disorders. Across five sites, the average percentage of frequent utilizers with substance use concerns was 35% compared to 13% for the general jail population.

Lesson 5: Choose thresholds carefully.

There is a lot to consider when identifying a threshold for examining the frequent utilizer population. This is an art and not a science and these decisions have consequences. Multiple factors can go into selecting thresholds, including: (1) how many individuals could be served with identified interventions; (2) the number of individuals needed to support a rigorous evaluation of newly developed interventions; and, (3) decisions about whether the resulting list of people really meets the jurisdiction's definition of frequent utilizer (has face validity). A couple of themes emerged.

If you choose a low threshold for inclusion, say the top 15% (or people with two or more bookings during the period), then you may face a few issues. First, you may have more people in the intervention than you can possibly serve. Picking a group with thousands of people will (at a minimum) limit the intensity of the intervention you are able to offer. In addition, the lower the threshold the more likely people in the group are to desist and not return (to jail or to police contact or to the hospital) to be engaged. Finally, the lower the threshold, the more diverse (heterogeneous) the group will be, meaning that you'll have to think carefully about what intervention will work for all or how to do further assessment to specify which intervention will work for which subgroups within the frequent utilizer population.

Alternatively, if you choose a very tight threshold for inclusion, say the top 2% of the population (or people with 10 or more jail bookings in the period), you may run into different challenges. For example, you will have some of the hardest and most complex individuals to serve and thus will need an appropriately intense and collaborative intervention. Or, the size of the population may not be enough to justify the intervention or support a rigorous evaluation. That said, it is more likely that this group will have face validity (will feel like prototypical frequent utilizer) and it is more likely that the group will be amenable to the similar intervention.

Given these lessons, analysts should not make decisions about thresholds on their own; instead they should provide several different options for cut-offs and share counts, demographics, and other summary information on the groups derived from the various thresholds chosen. This information can then be digested by first responders (to see if the list meets their definition of a frequent utilizer), leaders designing and funding the interventions, and evaluation partners.

SITE EXAMPLES

Johnson County, Iowa: Size of the group — Policymakers in Johnson County settled on a threshold of six or more bookings during a two-year period. Over the two years, these 144 people represented a small enough group to try to engage in their Housing First intervention and also represented a sizable number of bookings and jail bed days.

City of Long Beach, California: Evaluation — Another key factor is choosing a threshold that will provide sufficient sample size for an evaluation. Long Beach carefully considered their intervention and estimated the engagement rate into a multi-disciplinary team to be no more than 20%. They then performed power calculations with two possible metrics of interest, 12-month recidivism rate and 12-month arrest-count to identify the population size and the engagement rate needed to attain sufficient samples for the evaluation. This analysis led them to focus on the population with two or more arrests.

All sites: Demographics — It's worth taking a look at the demographics of the whole population compared to frequent utilizers. One site found African American men overrepresented in this group and others saw no differences by race.

Frequent utilizers may also be older than the average of the whole population. And some other perhaps unexpected results can happen. For example, frequent utilizers of the emergency department are more likely to be women than men, that is until you cross this with frequent utilizers of they criminal justice system, and then it flips to majority male.

Camden, New Jersey: Heterogeneity — The Camden Coalition linked all-payer claims data from four regional hospital systems with Camden County police arrest records. Analysts focused in on the top five percent of individuals based on numbers of emergency department visits and arrests. Understanding that even within such a small group there would be variability in patterns of cross-system engagement, the analysts explored the data using cluster analysis to identify subgroups of individuals and better understand their needs. Four subgroups were found in the data: (1) nonviolent, medically complex individuals primarily arrested for drug offenses; (2) nonviolent individuals with behavioral health complexity who are arrested predominantly for petty crimes; (3) assault victims with mental health challenges and addictions arrested for crimes against other persons; (4) young men arrested for a wide array of offenses, including drug trafficking, property crime, and violent crime, who have few hospitalizations and a comparatively low prevalence of mental illness and addiction. This analysis helped the Camden Coalition customize their intervention strategies.

Lesson 6: The 80/20 rule applies.

In some ways the DDJ work relies on the 80/20 rule, or Pareto Principle. The 80/20 rule theorizes that a large percentage of the problem is concentrated in a small percentage of people or places. The actual percentages may not be 80/20, but the theory still holds. In criminal justice, we find the Pareto Principle in play frequently.

The DDJ work envisions identifying those most frequently touching the criminal justice system (calls for service, arrests, jail bookings) as well as other service systems (emergency departments, homeless services, behavioral health services) and finding a different approach to treating them. Across the three sites, we find the 80/20 rule in play with a small group of people (less than 5%) responsible for 15-20% of all bookings/police contacts. Sites should consider running this analysis early on in their work. It can help make the case for the impact of the work on their partners operations, but it can also ensure that goals and expectations are aligned.

SITE EXAMPLES

Johnson County, Iowa: 250 people accounted for a total 1,689 bookings during 2016-2017. In other words, 4% of the people booked in the jail during this period accounted for 16% of the bookings.

Middlesex County, Massachusetts: Approximately 2% of arrestees in some communities in Middlesex County were responsible for 13-15% of all police contacts. Frequent utilizers had an average of nine contacts each, ranging from 5 to 95 contacts.

Lesson 7: Canaries in the coal mine

What if the experiences of people cycling through the criminal justice system and other emergency systems said more about our policies or system responses to vulnerable populations than it said about that people themselves? It's worth looking at the most frequent crimes of the whole population you are examining as well as the crimes of the frequent utilizers. Maybe these are the types of incidents that could be rethought/decriminalized altogether, and policy and system levers could impact a significant number of arrests and jail bookings without any individual level interventions. These kinds of changes require community discussions and rely on a community's comfort level with these offenses/social harms.

SITE EXAMPLES

In one site where frequent arrestees were examined, the person with the highest number of arrests was a person repeatedly charged with graffiti. This seemed like an anomaly until the next few people had the same pattern. Perhaps there is another way to deal with this behavior, other than arrest.

Johnson County, Iowa: Public intoxication was the most common charge for frequent utilizers (and the second most common charge for all people booked in the jail). This finding led leaders to open a sobering center as a key DDJ intervention. If the county diverts to the sobering center over jail, they could conservatively save over 200 bookings per year, nearly 50% of all individuals booked with public intoxication as their only charge.

Similarly, in another site, the most common charges among frequent utilizers were possession of unlawful paraphernalia, parks/beach loitering, and public consumption of alcohol. Perhaps the community would consider a different response to such activities.

We hope that the lessons learned in the DDJ pilot sites will help other jurisdictions as they seek to identify and support frequent utilizers of crisis services.

Appendix A: Resources for data integration and data analysis

The following resources can help guide communities through the process of integrating data, creating ongoing data sharing infrastructure, and analyzing key criminal justice data sets:

Resources for integrating data and creating ongoing data sharing infrastructure:

Actionable Intelligence for Social Policy (2016). AISP Innovation Expert Panel Reports. Available at https://www.aisp.upenn.edu/resources/aisp-innovation-expert-panel-reports/

Christenson, C. et al. (2019). Collaboration and Cross-Sector Data Sharing to Create Healthier Communities. *2-1-1 San Diego*. Available at https://documentcloud.adobe.com/link/track?uri=urn%3Aaaid%3Ascds%3AUS%3Ade70141e-eabe-4db5-a3ce-4e53306649c9

Corporation for Supportive Housing (2019). Data Integration Across Jail & Homeless Services Blueprint for Success. Available at https://d155kunxf1aozz.cloudfront.net/wp-content/uploads/2019/03/CSH-Data-Integration-Blueprint_20190321_FINAL.pdf

Doar, Robert (2018). Data Access and Integration. *Evidence-Based Policymaking Collaborative*. Available at https://www.evidencecollaborative.org/toolkits/data-access-and-integration

Justice and Health Connect. The Justice and Health Connect Toolkit. Available at http://www.jhconnect.org/toolkit#module-1-identify-and-convene-stakeholders

The National Center for Complex Health and Social Needs (2017). Cross-sector Collaboration for Data Sharing. Available at https://www.nationalcomplex.care/research-policy/resources/webinars/data-sharing/

Petrila, J. & Fader-Tower, H. (2010). Information Sharing in Criminal Justice-Mental Health Collaborations: Working with HIPAA and Other Privacy Laws. *Council of State Governments Justice Center*. Available at https://www.bja.gov/Publications/CSG_CJMH_Info_Sharing.pdf

Stewards of Change & National League of Cities (2014). Sharing Data for Better Results: A Guide to Building Integrated Data Systems Compatible with Federal Privacy Laws. Available at https://www.nlc.org/sites/default/files/users/user75/Data%20Sharing%20for%20Better%20Results.pdf

Resources for analyzing dispatch data:
RTI International (2019). CFS Analytics. Available at https://www.rti.org/impact/cfs-analytics

Resources for analyzing police data:
Clarke, R. & Eck, J. (2005). Crime Analysis for Problem Solvers in 60 Small Steps. *U.S. Department of Justice Office of Community Oriented Policing Services*. Available at https://www.popcenter.org/sites/default/files/library/reading/PDFs/60Steps.pdf

Fridell, L. (2005). Understanding Race Data From Vehicle Stops. *Police Executive Research Forum*. Available at https://cops.usdoj.gov/pdf/publications/understanding_Race_Data.pdf

Institute for Intergovernmental Research. (2019). Crime Analysis Toolkit: Analytical Techniques. Available at https://it.ojp.gov/CAT/?t=7&q=

Resource for analyzing jail data:
Elias, G. (2007). How to Collect and Analyze Data: A Manual for Sheriffs and Jail Administrators. *U.S. Department of Justice National Institute of Correction*. Available at https://www.prearesourcecenter.org/sites/default/files/library/howtocollectandanalyzedataamanualforsheriffsandjailadmn.pdf

Appendix B: Expanded analysis of social services involvement of people arrested or booked in the jail (in support of lesson 4)

Table 2: Allegheny County health and social services data sets and definitions

Health & Social Services Data Sets	Definition
Emergency department encounter	Individuals receiving Medicaid-funded emergency room services.
Inpatient hospitalization	Individuals receiving Medicaid-funded inpatient hospital services.
Mental health services	Individuals receiving a publicly-funded (Allegheny County or Medicaid managed care) mental health service. Includes both clinical services, such as individual and group therapy, and non-clinical services such as case management and peer support.
Mental health crisis encounters	Subset of all mental health services, individuals who received phone, mobile, residential crisis care.
Substance use disorder services	Individuals receiving a publicly-funded (Allegheny County or Medicaid managed care) substance use disorder (i.e. drug and alcohol) service. Excludes assessments.
Homelessness services	Individuals and families receiving prevention services, support services and/or housing who are homeless or at risk of becoming homeless. Services include shelter, street outreach, and transitional housing.
Public benefits utilization	Individuals or families receiving any of the following public benefits: SSI, SSDI, TANF, SNAP.
Parent on a child welfare case	Parents Associated with a Child Welfare Case (their child has an open child protective services case).

Tables 3 and 4 provide the full results for the analysis of individuals arrested in Allegheny County and their selected health and social services involvement. Table 3 expands the details displayed in Table 1 and looks at service involvement in the year prior to the initial arrest examines. Table 4 looks at service involvement in the year after the initial arrest.

Table 3: People arrested in 2016-2017 Allegheny County and their (selected) health and social services involvement during the 12 months prior to the incident:

# of Arrests	Distinct Individuals	Distinct Individuals (%)	% of Arrests	% Male	%Black/African American	Median Age	Emergency Dept. Encounter	Physical Health Inpatient	Mental Health Services	Mental Health Crisis	Substance Use Services	Homeless Services	Public Benefit Receipt	Child Welfare Parent
1	25,813	70%	43%	69%	41%	33	28%	6%	19%	4%	11%	2%	53%	5%
2	6,427	17%	22%	73%	47%	32	38%	9%	28%	7%	19%	3%	66%	7%
3	2,444	7%	13%	75%	49%	31	43%	10%	34%	9%	24%	3%	74%	10%
4	1,156	3%	8%	74%	47%	31	49%	13%	43%	11%	29%	5%	80%	11%
5 or more	1,228	3%	14%	73%	43%	32	53%	14%	47%	14%	35%	6%	83%	12%
Average	37,068	100%	100%	70%	43%	33	32%	8%	24%	6%	15%	2%	58%	6%
Ratio of 5 or more arrests to one arrest							1.9	2.3	2.5	3.5	3.2	3.0	1.6	2.4

Table 4: People arrested in Allegheny County in 2016-2017 and their (selected) health and social services involvement during the 12 months after the incident:

# of Arrests	Distinct Individuals	Distinct Individuals (%)	% of Arrests	% Male	%Black/African American	Median Age	Emergency Dept. Encounter	Physical Health Inpatient	Mental Health Services	Mental Health Crisis	Substance Use Services	Homeless Services	Public Benefit Receipt	Child Welfare Parent
1	25,813	70%	43%	69%	41%	33	25%	6%	20%	4%	14%	2%	60%	6%
2	6,427	17%	22%	73%	47%	32	38%	9%	33%	6%	25%	3%	73%	9%
3	2,444	7%	13%	75%	49%	31	45%	12%	42%	7%	34%	4%	80%	11%
4	1,156	3%	8%	74%	47%	31	52%	12%	49%	10%	37%	7%	86%	12%
5 or more	1,228	3%	14%	73%	43%	32	56%	14%	57%	12%	47%	8%	89%	15%
Average	37,068	100%	100%	70%	43%	33	30%	7%	26%	5%	19%	3%	65%	7%
Ratio of 5 or more arrests to one arrest							2.2	2.3	2.9	3.0	3.4	3.0	1.5	2.5

Tables 5 and 6 provide the results for the analysis individuals booked in the Allegheny County Jail and their selected health and social services involvement. Table 5 looks at service involvement in the year prior to the initial booking. Table 6 looks at service involvement in the year after the initial booking.

Table 5: People booked in Allegheny County Jail 2016-2017 and their (selected) health and social services involvement during the 12 months prior to the incident:

# of Bookings	Distinct Individuals	Distinct Individuals (%)	% of Arrests	% Male	%Black/African American	Median Age	Emergency Dept. Encounter	Physical Health Inpatient	Mental Health Services	Mental Health Crisis	Substance Use Services	Homeless Services	Public Benefit Receipt	Child Welfare Parent
1	13,704	70%	43%	69%	41%	33	28%	6%	19%	4%	11%	2%	53%	5%
2	4,037	17%	22%	73%	47%	32	38%	9%	28%	7%	19%	3%	66%	7%
3	1,479	7%	13%	75%	49%	31	43%	10%	34%	9%	24%	3%	74%	10%
4	527	3%	8%	74%	47%	31	49%	13%	43%	11%	29%	5%	80%	11%
5 or more	327	3%	14%	73%	43%	32	53%	14%	47%	14%	35%	6%	83%	12%
Average	20,074	100%	100%	70%	43%	33	32%	8%	24%	6%	15%	2%	58%	6%
Ratio of 5 bookings to one booking							1.9	2.3	2.5	3.5	3.2	3.0	1.6	2.4

Table 6: People booked in Allegheny County Jail 2016-2017 and their (selected) health and social services involvement during the 12 months after the incident:

# of Bookings	Distinct Individuals	Distinct Individuals (%)	% of Arrests	% Male	%Black/ African American	Median Age	Emergency Dept. Encounter	Physical Health Inpatient	Mental Health Services	Mental Health Crisis	Substance Use Services	Homeless Services	Public Benefit Receipt	Child Welfare Parent
1	13,704	68%	45%	77%	49%	34	32%	8%	29%	7%	16%	3%	64%	7%
2	4,037	20%	27%	80%	51%	33	39%	10%	36%	10%	24%	4%	74%	11%
3	1,479	7%	15%	81%	49%	32	45%	12%	42%	12%	28%	5%	79%	11%
4	527	3%	7%	78%	48%	33	49%	15%	49%	14%	28%	7%	81%	12%
5 or more	327	2%	6%	75%	51%	33	53%	16%	50%	16%	36%	9%	84%	17%
Average	20,074	100%	100%	78%	50%	33	36%	9%	32%	8%	19%	3%	68%	8%
Ratio of 5 or more bookings to one booking							1.7	2.0	1.7	2.3	2.3	3.0	1.3	2.4

Endnotes

1   Wagner, Paul. (2015). Jails Matter. But who is Listening? Prison Policy Initiative. Available at https://www.prisonpolicy.org/blog/2015/08/14/jailsmatter/.

2   Kyckelhahn, Tracey. (2013). Local Government Corrections Expenditures, FY 2005-2011. Bureau of Justice Statistics. Available at https://www.bjs.gov/content/pub/pdf/lgcefy0511.pdf.

3   Bronson, J., & Berzofsky, M. (2017). Indicators of Mental Health Problems Reported by Prisoners and Jail Inmates, 2011-12. Bureau of Justice Statistics. Available at https://www.bjs.gov/content/pub/pdf/imhprpji1112.pdf.

4   Bronson, J., et al. (2017). Drug Use, Dependence, and Abuse Among State Prisoners and Jail Inmates, 2007-2009. Bureau of Justice Statistics. Available at https://www.bjs.gov/content/pub/pdf/dudaspji0709.pdf

5   Treatment Advocacy Center and National Sheriffs' Association. (2010). More Mentally Ill Persons Are in Jails and Prisons than Hospitals: A Survey of the States. Available at https://www.treatmentadvocacycenter.org/storage/documents/final_jails_v_hospitals_study.pdf

6   This paper does not describe technical issues the DDJ sites experienced when integrating data. Their issues ranged from knowing what data to request, standardizing data across multiple (police) departments, problems querying data from legacy data collection systems, and technical integration challenges. Sites that used a third-party technical partner to integrate data were able to integrate faster but still faced some challenges integrating and using data.

7   This analysis assumes that the site is starting with a criminal justice data set and layering other data sets on top of it. This is the way we imagined most DDJ sites would start.

8   Unpublished manuscript. For more information on the project visit: https://islg.cuny.edu/sites/our-work/frequent-utilizers-of-the-criminal-justice-health-and-social-service-systems/