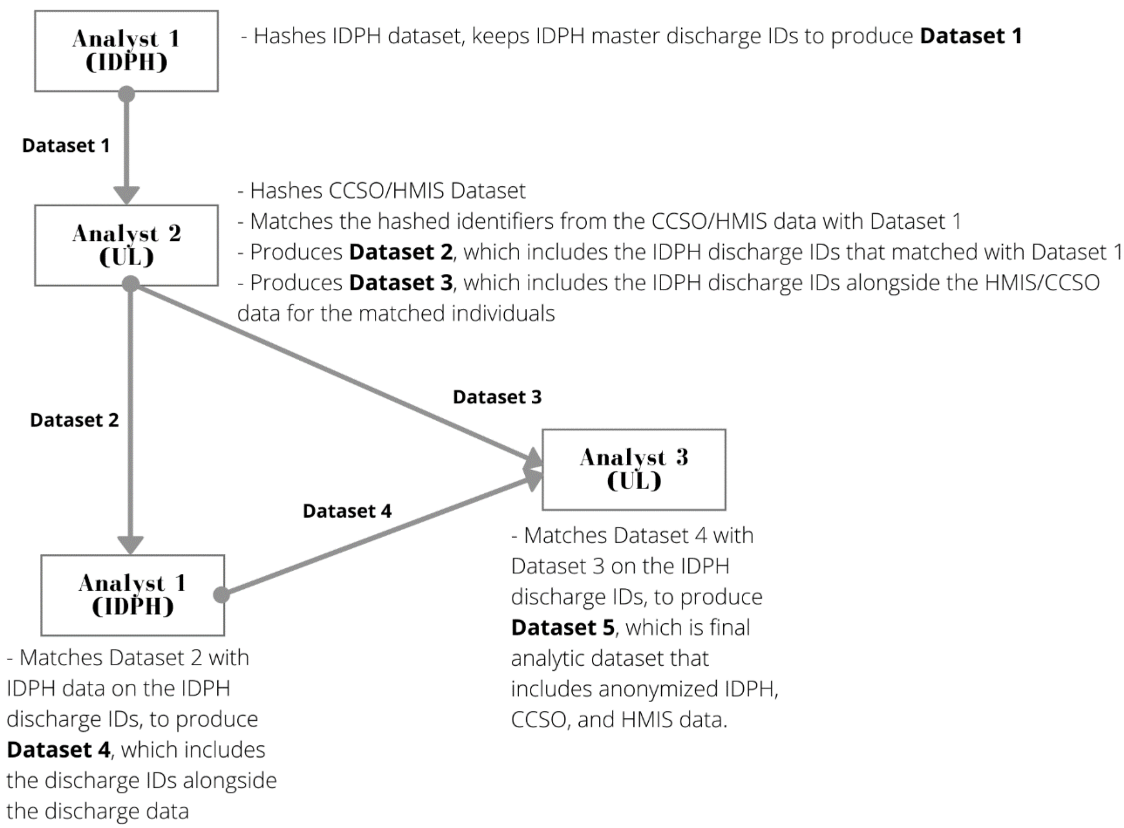
Appendix A: Hashing.

Three analysts were involved in the data transfer and matching process – two from the Urban Labs (labeled as UL in Supplemental Figure 1 below), and a single collaborator from IDPH. We detail the process in Supplemental Figure 1 below:

Supplemental Figure 1: Record linkage and data matching protocol



The undisclosed hash algorithm used in this paper takes a message input of arbitrary length as input and produces a message of an undisclosed bit length as output (Supplemental Figure 2). Although the implemented algorithm is not an encryption scheme in the strictest sense, we make use of the function's one-way direction to anonymize our datasets.

Supplemental Figure 2: Hashing the string 'John Smith' using an MD5 algorithm

As an additional guard against brute force attacks such as rainbow table attacks,[1] we included a key and a 'salting' procedure in our hashing process to render them more difficult. A rainbow table attack is where precomputed hash values (see examples in Supplemental Table 1) for commonly used strings - such as 'John Smith' in Supplemental Figure 2 above - are stored in large databases, so that they can be retrieved and used to decrypt digests created using a specific hashing algorithm.

Supplemental Table 1: Some precomputed MD5 digests

Input String	MD5 Digest
John Smith	6117323d2cabbc17d44c2b44587f682c
Michel Foucault	e02e7f44283852199a522ccbe80ea694
Bertrand Russell	c7d13621805acbd3d253e594245ad933
Hannah Arendt	535ada6b77112d0fa796e7a8f0273543

Appendix B: Using Jenks Natural Breaks Optimization to Define High Use

[1] Tatlı Eİ. Cracking more password hashes With patterns. IEEE Transactions on Information Forensics and Security. 2015 Aug;10(8):1656–65.

We applied the Jenks Natural Breaks Optimization to our data as an atheoretical alternative method in discovering natural cutoffs in the data to define high use in each system. This clustering method minimizes within group variance while maximizing the distinction between groups. We performed the Jenks method varying the number of bins (classes) from 3 to 10. Selected results for 4 bins are detailed in Supplemental Table 2, and for 10 bins in Supplemental Table 3.

Supplemental Table 2: Jenks Natural Breaks Optimization for 4 bins

Bin/Sector	IDPH N = 301,869	HMIS N = 59,274	CCSO N = 195,655
Q1	• **n = 6** • 523 ED Visits • 808 ED Days • 163 IP Stays • 1243 ED + IP Days	• **n = 1** • 34 Shelter Stays • 1172 Shelter Days • 1532 St Outreach	• **n = 4** • 34 Jail Stays • 1460 Jail Days
Q2	• **n = 17,133** • 99 ED Visits • 159 ED Days • 21 IP Stays • 237 ED + IP Days	• **n = 747** • 3 Shelter Stays • 282 Shelter Days • 368 St Outreach	• **n = 3,720** • 6 Jail Stays • 846 Jail Days
Q3	• **n = 145,093** • 17 ED Visits • 28 ED Days • 4 IP Stays • 62 ED + IP Days	• **n = 2,141** • 1 Shelter Stays • 84 Shelter Days • 87 St Outreach	• **n = 31,887** • 2 Jail Stays • 207 Jail Days
Q4	• **n = 139,637** • 0 ED Visits • 0 ED Days • 0 IP Stays • 1 ED + IP Days	• **n = 56,386** • 0 Shelter Stays • 0 Shelter Days • 0 St Outreach	• **n = 160,044** • 1 Jail Stays • 1 Jail Days

Supplemental Table 3: Jenks Natural Breaks Optimization for 10 bins

Bin/Sector	IDPH N = 301,869	HMIS N = 59,274	CCSO N = 195,655
Q1	• **n = 6** • 523 ED Visits • 808 ED Days • 163 IP Stays • 1243 ED + IP Days	• **n = 1** • 34 Shelter Stays • 1172 Shelter Days • 1532 St Outreach	• **n = 4** • 43 Jail Stays • 1460 Jail Days
Q2	• **n = 73** • 341 ED Visits • 493 ED Days • 82 IP Stays • 751 ED + IP Days	• **n = 15** • 19 Shelter Stays • 716 Shelter Days • 918 St Outreach	• **n = 1,010** • 20 Jail Stays • 1066 Jail Days
Q3	• **n = 206** • 206 ED Visits • 273 ED Days • 45 IP Stays • 500 ED + IP Days	• **n = 45** • 10 Shelter Stays • 531 Shelter Days • 687 St Outreach	• **n = 1,360** • 13 Jail Stays • 769 Jail Days
Q4	• **n = 741** • 117 ED Visits • 149 ED Days • 26 IP Stays • 315 ED + IP Days	• **n = 409** • 6 Shelter Stays • 395 Shelter Days • 522 St Outreach	• **n = 1,937** • 9 Jail Stays • 550 Jail Days
Q5	• **n = 2,498** • 63 ED Visits • 83 ED Days • 16 IP Stays • 200 ED + IP Days	• **n = 268** • 4 Shelter Stays • 288 Shelter Days • 373 St Outreach	• **n = 3,094** • 6 Jail Stays • 384 Jail Days
Q6	• **n = 6,696** • 33 ED Visits • 45 ED Days • 10 IP Stays • 127 ED + IP Days	• **n = 435** • 3 Shelter Stays • 202 Shelter Days • 247 St Outreach	• **n = 4,945** • 4 Jail Stays • 252 Jail Days
Q7	• **n = 16,741** • 18 ED Visits • 24 ED Days • 6 IP Stays • 78 ED + IP Days	• **n = 618** • 2 Shelter Stays • 130 Shelter Days • 151 St Outreach	• **n = 8,805** • 3 Jail Stays • 151 Jail Days
Q8	• **n = 35,980** • 10 ED Visits • 13 ED Days • 3 IP Stays • 45 ED + IP Days	• **n = 1151** • 1 Shelter Stays • 70 Shelter Days • 81 St Outreach	• **n = 14,042** • 2 Jail Stays • 79 Jail Days
Q9	• **n = 68,325** • 5 ED Visits • 6 ED Days • 1 IP Stays • 22 ED + IP Days	• **n = 811** • 0 Shelter Stays • 22 Shelter Days • 25 St Outreach	• **n = 25,134** • 1 Jail Stays • 26 Jail Days
Q10	• **n = 170,603** • 0 ED Visits • 0 ED Days • 0 IP Stays • 1 ED + IP Days	• **n = 55,521** • 0 Shelter Stays • 0 Shelter Days • 0 St Outreach	• **n = 135,324** • 1 Jail Stays • 1 Jail Days

We found the Jenks optimization method to be very sensitive to outliers and extreme values as exemplified by the breakpoints in the first and last bins. For example, with 4 bins, the bin with the highest utilization comprised only 10 individuals: those with 1,243 or more ED and inpatient days in total, or more than 1,000 days of interaction for both the jail and homelessness systems. Comparing the cutoffs obtained via the Jenks method to the original cutoffs that we obtained through considering similar percentiles from previous research and the interest of our partners, we observe that our high user cutoffs lie in the middle of the pack between the second and third quartiles in the 4-bin breakdown.

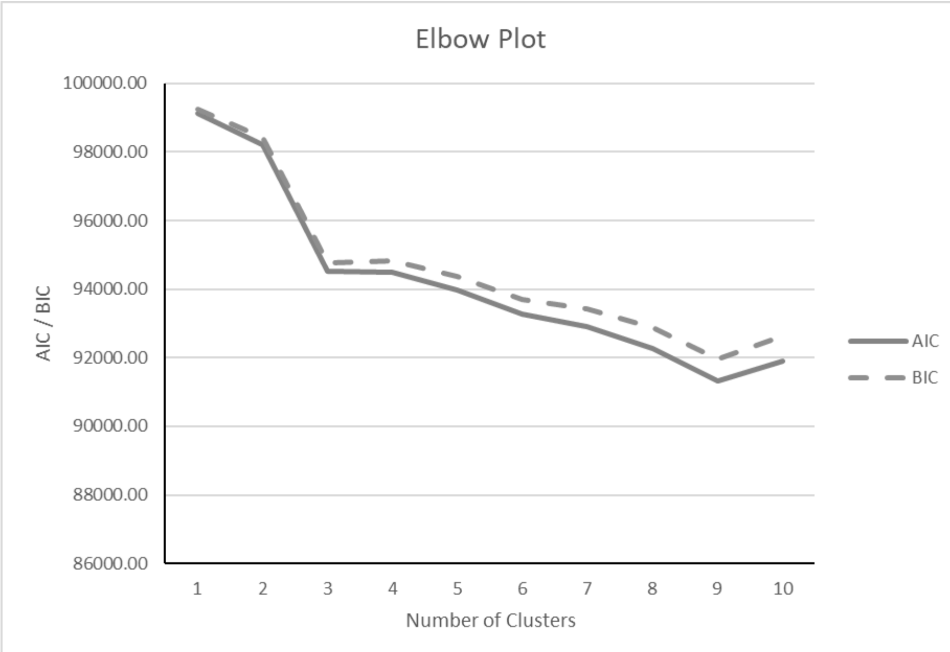
Appendix C: Latent Profile Analysis to Derive MHU Profiles

In addition to the k-means analysis, we conducted a latent profile analysis (LPA) on our dataset, assuming independence of each observation. Supplemental Table 4 shows the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for 1-10 classes. Supplemental Figure 3 shows a graph of the AIC and BIC against the number of clusters, giving a sense of how much additional useful information is captured by increasing the number of clusters. From the plot, moving from two to three clusters provides the greatest information gain, and additional clusters beyond that do not reduce AIC to as substantial an extent. The BIC in fact increases if we increase the number of clusters from three to four. Therefore, three clusters was selected as the optimal number of clusters in this data using LPA.

Supplemental Table 4: AIC/BIC for Latent Profile Analysis (10 classes)

Classes	AIC	BIC
1	99110.64	99233.43
2	98220.16	98409.47
3	94513.63	94769.45
4	94491.65	94813.98
5	93981.50	94370.35
6	93258.38	93713.73
7	92912.72	93434.60
8	92275.32	92863.71
9	91322.57	91977.47
10	91903.55	92624.96

Supplemental Figure 3: LPA Elbow Plot



Supplemental Table 5 shows cluster sizes by the number of clusters created in the data. The three-cluster solution would result in cluster sizes very similar to cluster sizes obtained with the original k-means cluster solution (Supplemental Table 6). There is a large overlap between the LPA and k-means models, with 95.29% of the individuals in the sample remaining in the same cluster groupings. The largest difference resides with Cluster 2. Of the 124 individuals grouped into Cluster 2 by the k-means analysis, 53 of them (42.7%) were assigned to Cluster 1 by LPA.

Supplemental Table 5: LPA Clusters, Breakdown of Cluster Sizes by Number of Clusters

Number of Clusters/Cluster Size	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
2	1021	211	-	-	-
3	1107	71	54	-	-
4	535	71	54	572	-
5	455	70	54	166	487

* Models with more than 6 clusters did not converge properly.

Supplemental Table 6: Cluster Overlaps, Comparing the LPA 3-Cluster Model versus K-means 3-Cluster Model

Cluster Overlap	K-means Cluster 1	K-means Cluster 2	K-means Cluster 3
LPA Cluster 1	1054	53	0
LPA Cluster 2	0	71	0
LPA Cluster 3	5	0	49

*Total overlap: 95.29% (1174 of 1232)

Supplemental Table 7 shows demographic, utilization, and other characteristics of individuals in each of the three LPA clusters. The characteristics of each cluster are very similar to the characteristics of corresponding MHU profiles obtained using k-means analysis (Appendix D), and the qualitative interpretation of each cluster remains the same. The largest differences are seen in Cluster 2, corresponding to the second MHU profile titled "Older singles with highest rates of behavioral health challenges." Using LPA, this cluster contained only 71 individuals, versus 124 individuals in the profile using k-means. However, this cluster remains the oldest of the three (average age 53.5 years using LPA, 52.8 years using k-means), with the highest rates of co-occurring behavioral health issues. When detained in jail, 59.2% required medical services (versus 62.9% using k-means), 53.5% required mental health services (63.7% using k-means), and 56.3% required detox upon entry (58.9% using k-means). This cluster continues to demonstrate the highest average number of days of cross-system engagement over the study period (366.0 days versus 344.6 days using k-means), with utilization predominantly being that of the jail. The two other clusters obtained via LPA showed similar overlaps in characteristics and interpretation with the two other profiles obtained via k-means. Therefore, we found that LPA produced similar clusters to the profiles we obtained using k-means, increasing our confidence that the original profiles obtained via k-means analysis are meaningful in our population.

Supplemental Table 7: Demographic, Utilization, and other Characteristics of LPA Clusters

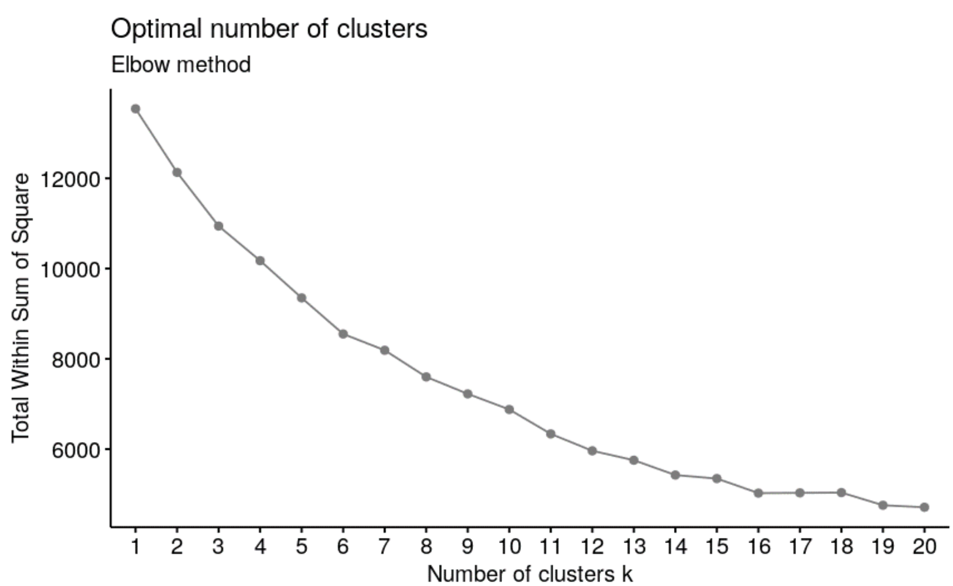
Variable	Cluster 1	Cluster 2	Cluster 3	*p-value*
N	1107	71	54	-
ED Stays (Mean/SD)	9.65 (9.92)	7.49 (7.05)	15.24 (20.01)	**<0.001***
ED Days (Mean/SD)	14.42 (17.96)	10.65 (11.62)	23.87 (32.47)	**<0.001***
Ever ED Stay (%/n)	94.58% (1047)	91.55% (65)	94.44% (51)	0.560
IP Admits (Mean/SD)	3.60 (5.32)	2.38 (4.55)	6.94 (10.26)	**<0.001***
IP Days (Mean/SD)	25.16 (46.97)	15.82 (34.51)	47.83 (71.22)	**0.001***
Ever IP Stay (%/n)	82.57% (914)	77.46% (55)	90.74% (49)	0.150
Emergency Shelter (Mean/SD)	108.73 (147.70)	48.51 (62.15)	108.91 (166.12)	**0.003***
Ever Shelter Stay (%/n)	77.78% (861)	83.10% (59)	61.11% (33)	**0.008***
Street Outreach (Mean/SD)	28.64 (105.77)	25.59 (96.12)	12.89 (64.64)	0.544
Ever in Outreach (%/n)	17.16% (190)	12.68% (9)	14.81% (8)	0.571
Permanent Housing (Mean/SD)	85.00 (270.29)	110.21 (283.15)	189.15 (443.80)	**0.025***
Ever in Permanent Housing (%)	14.18% (157)	18.31% (13)	25.93% (14)	**0.044***
Jail Stays (Mean/SD)	2.74 (2.05)	8.20 (5.40)	2.26 (1.59)	**<0.001***
Jail Days (Mean/SD)	135.85 (174.72)	291.93 (181.49)	80.26 (115.51)	**<0.001***
Ever Jail Stay (%/n)	100% (1107)	100% (71)	100% (54)	-
Cross-System Engagement (Mean/SD)	273.22 (229.76)	366.03 (190.60)	202.06 (179.72)	**<0.001***
Age (Mean/SD)	50.83 (11.58)	53.52 (9.48)	34.85 (13.19)	**<0.001***
<18	0.27% (3)	0.00% (0)	5.56% (3)	-
18 - <25	6.41% (71)	2.82% (2)	22.22% (12)	-
25 - <35	6.59% (73)	5.63 % (4)	29.63% (16)	-
35 - <45	5.96% (63)	1.41% (1)	14.81% (8)	-
45 - <55	38.66% (428)	40.85% (29)	18.52% (10)	-
55 - <65	38.21% (423)	45.07% (32)	9.26% (5)	-
>= 65	3.88% (43)	4.23% (3)	0.00% (0)	-
Gender – Male (%/n)	87.53% (969)	87.32% (62)	44.44% (24)	**<0.001***
Gender – Female (%/n)	12.20% (135)	12.68% (9)	55.56% (30)	**-**
Gender – Trans/Non-Conform (%/n)	0.27% (3)	0% (0)	0% (0)	**-**
Race – Black (%/n)	83.56% (925)	88.73% (63)	94.44% (51)	0.178
Race – Hispanic (%/n)	4.07% (45)	2.82% (2)	5.56% (3)	**-**
Race – White (%/n)	9.94% (110)	7.04% (5)	0% (0)	**-**
Race – Other (%/n)	2.44% (27)	1.41% (1)	0% (0)	**-**
Survived Domestic Violence (%/n)	7.86% (87)	5.63% (4)	25.93% (14)	**<0.001***
Disabling Condition (%/n)	65.85% (729)	73.24 (52)	42.59% (23)	**0.001***
Received Benefits (%/n)	65.67% (727)	63.89% (46)	70.37% (38)	0.740
Health Insurance (%/n)	68.29% (756)	74.65% (53)	61.11% (33)	0.904
Behavioral Health Issues (%/n)	32.79% (363)	32.39% (23)	27.78% (15)	**<0.001***
Medical Services in Jail (%/n)	23.04% (255)	59.15% (42)	0% (0)	**<0.001***
Mental Health Services in Jail (%/n)	33.42% (370)	53.52% (38)	20.37% (11)	**<0.001***
Detox upon Entry to Jail (%/n)	30.80% (341)	56.34% (40)	18.52% (10)	**<0.001***

*Statistically significant at the 5% level. Note: Statistically significant differences in continuous variables were assessed with ANOVA and in categorical variables were assessed with chi-square tests.

Appendix D: Additional Details on K-Means Clustering Results

To determine the optimal number of clusters, we utilized the elbow method, graphing the number of clusters against the within-cluster sum of squares (WCSS). WCSS is the sum of squared distance between each point and the centroid in a cluster. As the number of clusters increases, the WCSS value correspondingly starts to decrease. Looking at the graph produced from the k-means analysis (Supplemental Figure 4), a slight elbow occurs at both three and six clusters. A model with six clusters contained some clusters of very small size; therefore, a three-cluster solution was chosen. While the presence of a clear elbow is not readily present, the number of clusters was nevertheless chosen to maximize the amount of variance explained by the model, but at the same time model interpretability and parsimony.

Supplemental Figure 4: K-means Elbow Plot



Supplemental Table 8 shows demographic, utilization, and other characteristics of individuals in each of the three MHU profiles, plus tests of statistically significant differences between the three profiles. Differences in continuous variables were assessed with ANOVA and differences in categorical variables were assessed with chi-square tests.

Supplemental Table 8: Demographic, Utilization, and other Characteristics of MHU Profiles (k-means)

Variable	Older singles with complex needs	Older singles with highest rates of behavioral health challenges	Younger individuals with families	p-value
N	1059	124	49	-
ED Stays (Mean/SD)	9.74 (9.81)	7.61 (9.14)	15.92 (20.89)	**<0.001***
ED Days (Mean/SD)	14.72 (18.25)	10.48 (12.86)	22.80 (32.20)	**<0.001***
Ever ED Stay (%/n)	94.71% (1003)	91.94% (114)	93.88% (46)	0.439
IP Admits (Mean/SD)	3.65 (5.37)	2.42 (4.24)	7.47 (10.63)	**<0.001***
IP Days (Mean/SD)	25.54 (50.18)	16.06 (34.28)	51.39 (73.78)	**<0.001***
Ever IP Stay (%/n)	83.47% (884)	71.77% (89)	91.84% (45)	**0.001***
Emergency Shelter (Mean/SD)	111.70 (149.86)	47.23 (61.81)	113.02 (171.30)	**<0.001***
Ever Shelter Stay (%/n)	77.81% (824)	80.65% (100)	59.18% (29)	**0.006***
Street Outreach (Mean/SD)	29.07 (107.43)	22.04 (81.02)	14.18 (67.79)	0.501
Ever in Outreach (%/n)	16.81% (178)	17.74% (22)	14.29% (7)	0.861
Permanent Housing (Mean/SD)	85.68 (270.83)	99.52 (282.05)	184.88 (446.76)	0.051
Ever in Permanent Housing (%)	14.45% (153)	15.32% (19)	24.49% (12)	0.155
Jail Stays (Mean/SD)	2.52 (1.72)	7.72 (4.58)	2.27 (1.63)	**<0.001***
Jail Days (Mean/SD)	129.60 (172.84)	275.30 (175.39)	83.10 (120.21)	**<0.001***
Ever Jail Stay (%/n)	100% (1059)	100% (124)	100% (49)	-
Cross-System Engagement (Mean/SD)	270.37 (230.98)	344.56 (193.52)	210.31 (184.97)	**<0.001***
Age (Mean/SD)	50.79 (11.61)	52.80 (9.92)	32.99 (12.36)	**<0.001***
<18	0.28% (3)	0.00% (0)	6.12% (3)	-
18 - <25	6.42% (68)	4.03% (5)	24.49% (12)	-
25 - <35	6.80% (72)	4.03 % (5)	32.65% (16)	-
35 - <45	6.04% (64)	2.42% (3)	16.33% (8)	-
45 - <55	38.05% (403)	45.97% (57)	14.29% (7)	-
55 - <65	38.62% (409)	38.71% (48)	6.12% (3)	-
>= 65	3.78% (40)	4.84% (6)	0.00% (0)	-
Gender – Male (%/n)	87.63% (928)	85.48% (106)	42.86% (21)	**<0.001***
Gender – Female (%/n)	12.09% (128)	14.52% (18)	57.14% (28)	**-**
Gender – Trans/Non-Conform (%/n)	0.28% (3)	0% (0)	0% (0)	**-**
Race – Black (%/n)	84.70% (897)	77.42% (96)	93.88% (46)	**0.006***
Race – Hispanic (%/n)	4.06% (43)	3.23% (4)	6.12% (3)	**-**
Race – White (%/n)	8.78% (93)	17.74% (22)	0% (0)	**-**
Race – Other (%/n)	2.46% (26)	1.61% (2)	0% (0)	**-**
Survived Domestic Violence (%/n)	7.74% (82)	8.87% (11)	24.49% (12)	**<0.001***
Disabling Condition (%/n)	65.82% (697)	70.16 (87)	40.82% (20)	**<0.001***
Received Benefits (%/n)	66.57% (705)	58.06% (72)	69.39% (34)	0.089
Health Insurance (%/n)	68.27% (723)	72.58% (90)	59.18% (29)	0.676

Behavioral Health Issues (%/n)	30.50% (323)	51.61% (64)	28.57% (14)	**<0.001***
Medical Services in Jail (%/n)	20.68% (219)	62.90% (78)	0% (0)	**<0.001***
Mental Health Services in Jail (%/n)	31.16% (330)	63.71% (79)	20.41% (10)	**<0.001***
Detox upon Entry to Jail (%/n)	29.27% (310)	58.87% (73)	16.33% (8)	**<0.001***

*Statistically significant at the 5% level.