Better Foster Youth to Independence Data Speaks : What can be done?

Gandalf Tech

U.S. Department of Housing and Urban Development

FYI Foster Youth to Independence

UofSC South Carolina

Georgia Tech

Origin of Gandalf Tech…

In Middle Earth, Gandalf is a wizard who has great power, but works mostly by encouraging and persuading. He devotes his whole life protecting people from losing their homes, especially the four Hobbit youths.

Nowadays, however, there are still thousands of youths enduring homelessness across the country. This should never happen. **Therefore, we aim to be today's Gandalf, who works not only by persuading, but also utilizing business, policy and technology...**

Outline

Problem Framing

Theoretical Framework

Data Analysis

Insights and Discussion

Problem Framing: FYI Initiative and Policy Iterations

FUP

95% — Family 1990

Provides HCVs to **families involved in the child welfare system** for whom lack of housing is a primary barrier to reunification

5% — Youth (FYI Initiative) 2019

provides HCVs to **Youth** at least 18 years and not more than 24 years of age who left foster care or will leave foster care within 90 days
Application Process: a youth must be referred by a Public Child Welfare Agency (PCWA) to apply a FYI voucher at PHA. When exploring why PHAs were not serving youth, the main reason cited was lack of referral of youth from PCWAs.

FYI – Less PHA 07/2019

- Only available to PHAs that do not administer FUP vouchers
- Each PHA was capped at 25 vouchers in a year

FYI – More PHA 10/2020

- PHAs that participate in FUP are eligible to request FYI vouchers
- Funded from a set-aside of FUP funding of $20 million

Competitive FYI 01/2021

- Designated $20 million for FUP for youth, half of which was to be awarded competitively and half on an as-needed basis

Challenges in promoting FYI and corresponding research focus

	Challenge	Research focus
Policy Evaluation	Effectiveness of past policy iterations unknown	1 Prove effectiveness of policy iterations by data analysis
	Constraint in FYI budget	2 Argue whether the constraint in FYI budget is a major challenge
	Inadequate cooperation between PHA and PCWA	3 Provide data evidence & give suggestions of facilitating cooperation
Promotion Target	Unclear of important factors influencing PHA's FYI adoption level and number of homeless youth	4 Identify important factors
	Unclear of which counties to target with high priority for promoting FYI	5 Identify important features in locating promising counties for promoting FYI and provide a list of the counties

Outline

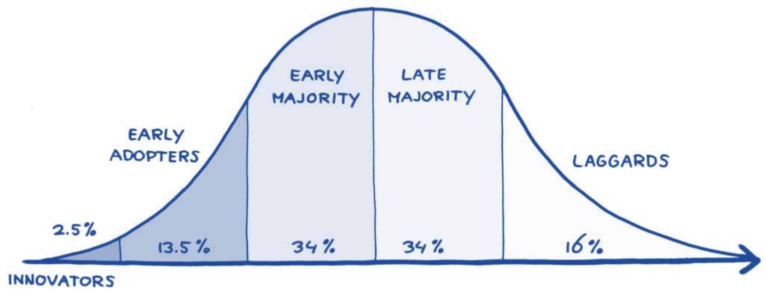
Problem Framing

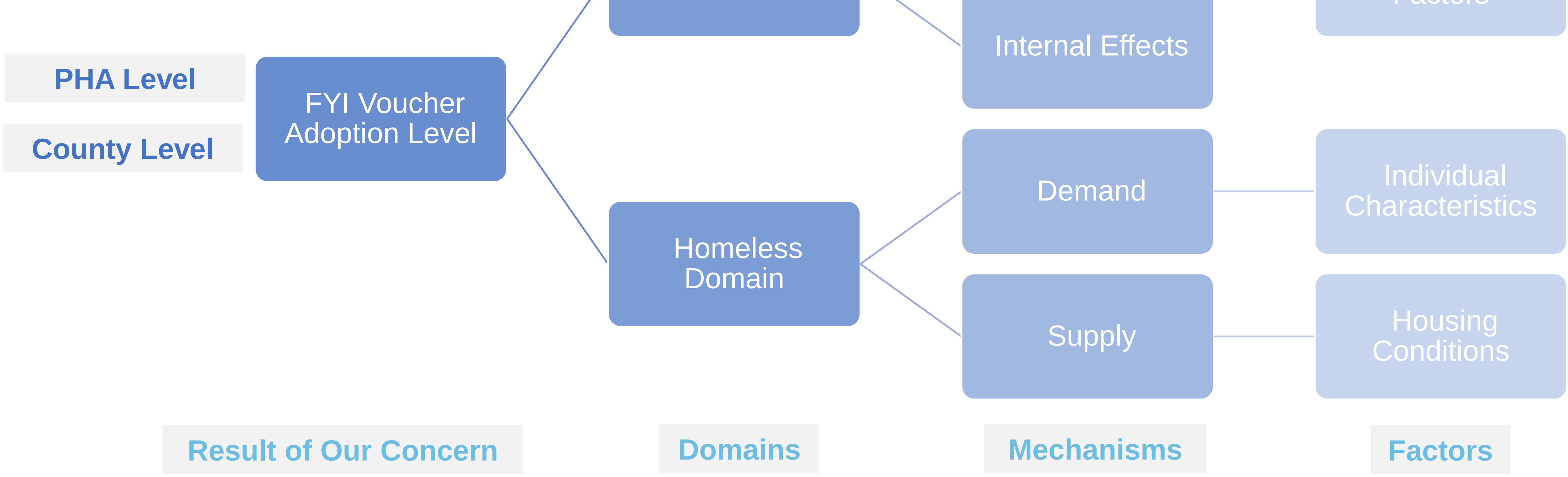
Theoretical Framework

Data Analysis

Insights and Discussion

Theoretical Framework: more than Diffusion of Innovation Theory





Graph source: https://militaryfamilieslearningnetwork.org/2017/01/11/assumptions-of-basing-our-work-on-diffusion-of-innovation/

Outline

Problem Framing

Theoretical Framework

Data Analysis

Data Acquisition

Exploratory Data Analysis

Modeling

Insights and Discussion

Data acquisition of PHA domain variables

Factor	Variable	Source
Internal Effects	PHA Code	MMC data
	PHA City	
	PHA County	
	PHA State	
	PHA FYI Funding	
	PHA First FYI Adoption Date	
	PHA FYI Application Count	
	PHA # Housing Inventory	Housing Inventory Count
	PHA Annual Budget	HUD Dashboard
	PHA Per Unit Per Month Cost	
	PHA # Total Vouchers	

Factor	Variable	Source
External Effects-Socioeconomic	State # Homeless Youth	Adoption & Foster Care Statistics
	County Median Rent	ACS 1-Year
	County Unemployment Rate	BLS
	County GDP	SAIPE
	County Gini Index	ACS 5-Year
	City Density	
External Effects-Demographic	City Population	Census
	County Race Percentage	
	County Age Percentage	
	County Gender Percentage	

Data acquisition of Homeless domain variables

Factor	Variable	Source
Demand (Homeless Individual Characteristics aggregated to county level)	Homeless Race/Ethnicity Distribution	Adoption and Foster Care Analysis and Reporting (AFCARS) Foster Care Annual File (2021) **Note: the county FIPS code is unified as "8" if there are fewer than 1000 homeless cases in the County. Our analysis therefore focuses only on 125 counties with 1000+ homeless cases.**
	Removal Reasons	
	Length Since Latest Removal	
	Average Rural/Urban Level	
	The Number of Places the Youth Has Lived	
	Foster Caretaker Race/Ethnicity Distribution	
	Social Security Act Benefits	
	Physical Condition	
	Mental Condition	
	Caretaker Characteristics	
	Others (see Appendix 1)	
Supply (Housing Supply Conditions)	County Housing Density	ACS 1-Year
	County Housing Price Index	ACS 1-Year

Outline

Problem Framing

Theoretical Framework

Data Analysis

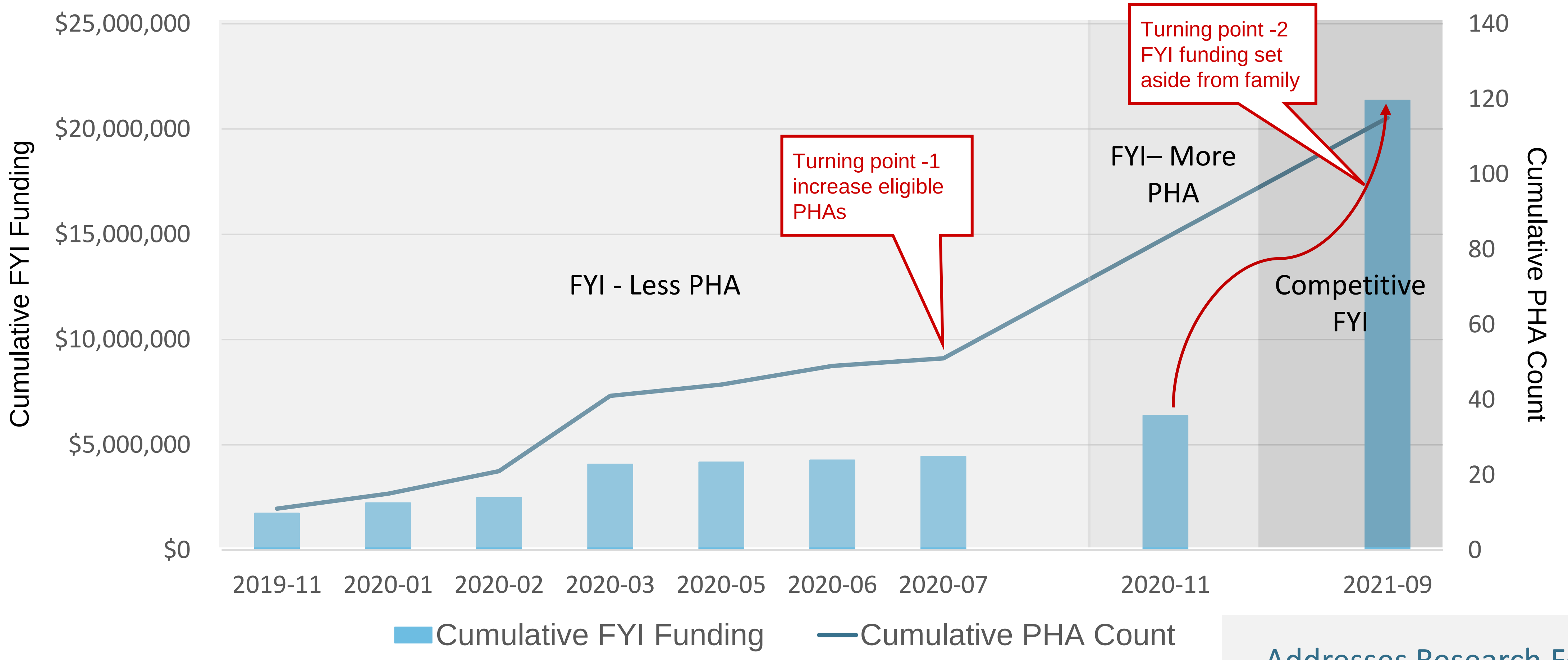
- Data Acquisition
- **Exploratory Data Analysis**
- Modeling

Insights and Discussion

FYI Policy Revisions are proving effective as FYI Funding boost and more PHAs adopt in recent past

BUDGET/PHA

Cumulative FYI Funding & PHA Count vs Date

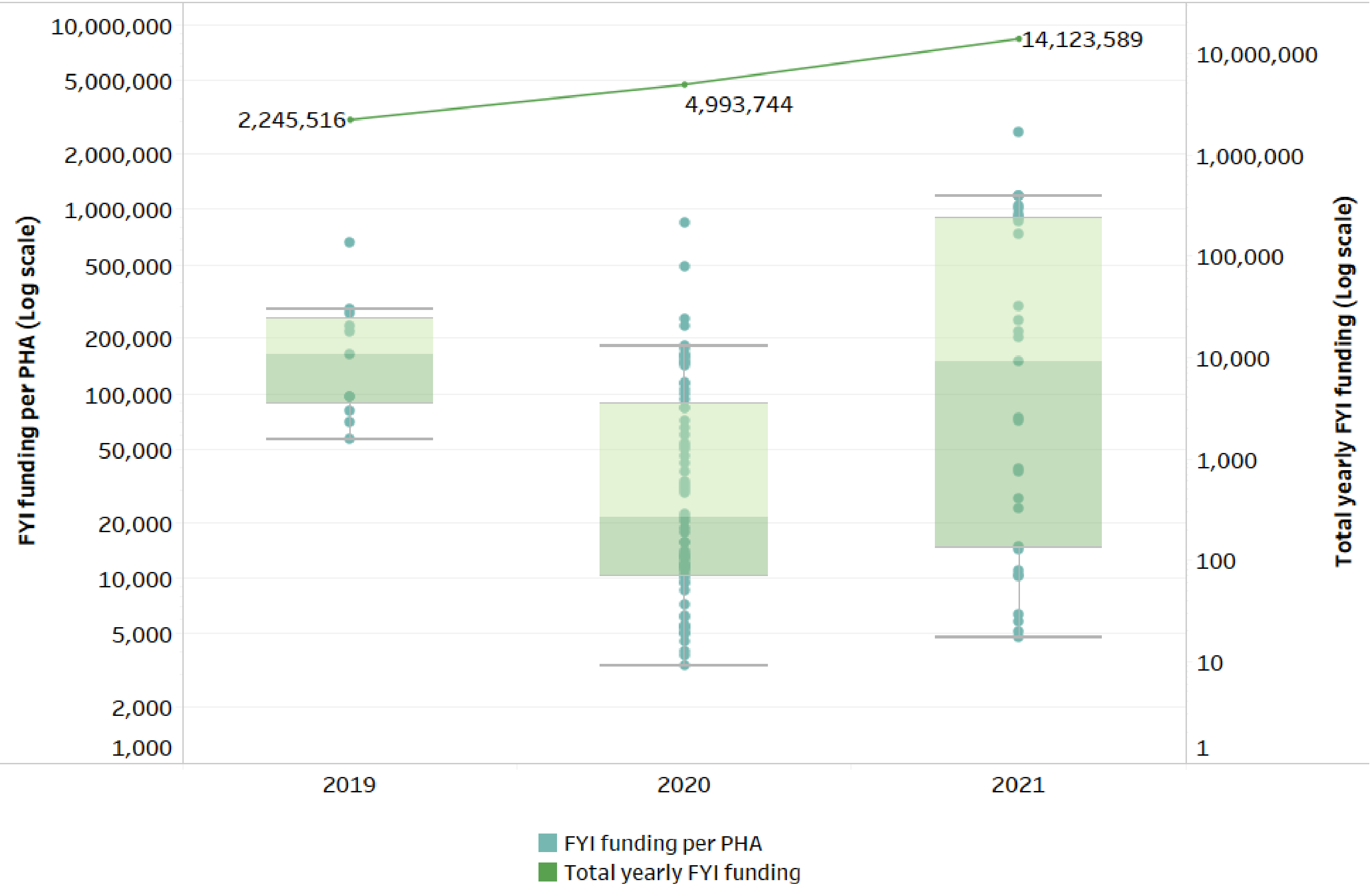


Addresses Research Focus ①

Despite the positive effect of budget separation on FYI adoption, the total available funding of $20M is still under utilized by about ~30%

BUDGET/PHA

Distribution of each PHA FYI budget in the year of adoption



Observations

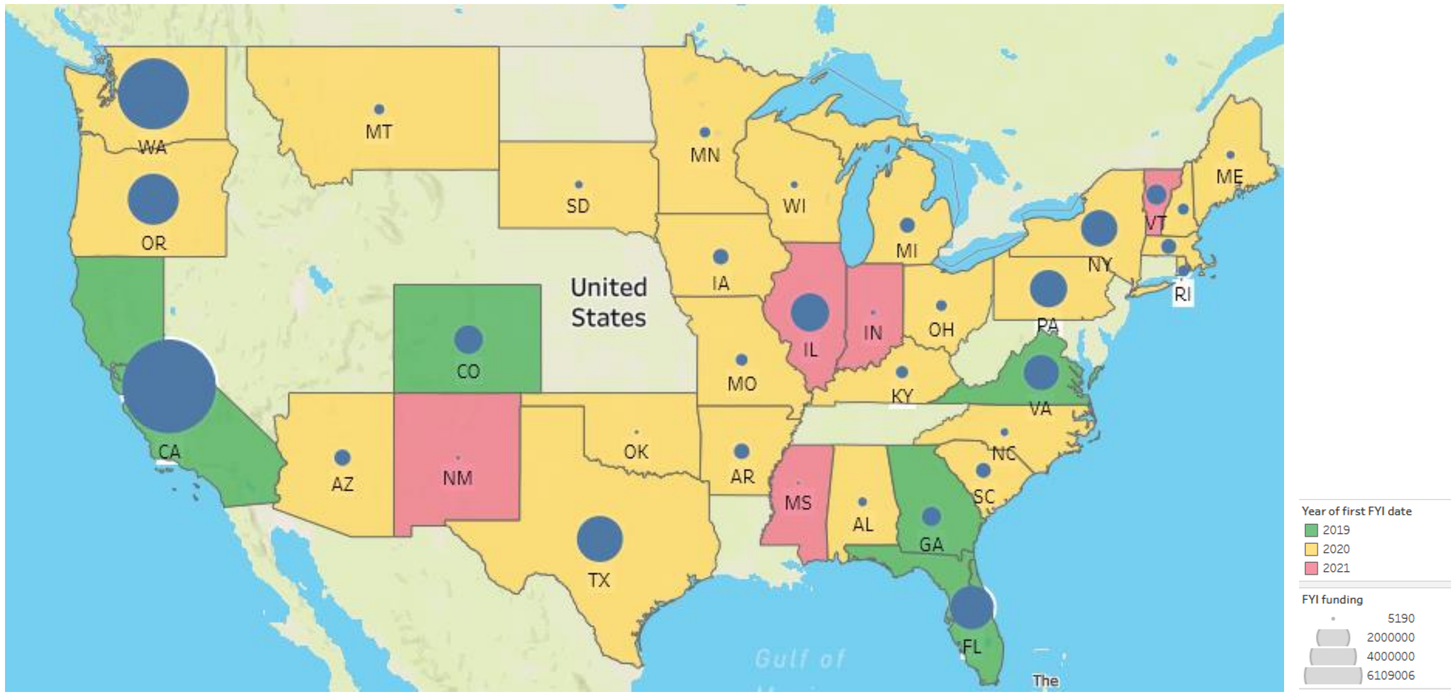
- PHA FYI spending has gone up significantly in 2021 to ~$14MM
- Despite increased PHA spend on FYI related activities, still the FYI budget is under utilized by ~30% for FY2021 (*available total $20MM*)
- There is a significant variation in FYI funding across different PHAs in 2021, indicating that recently varied size PHAs are also adopting the program

Addresses Research Focus 2

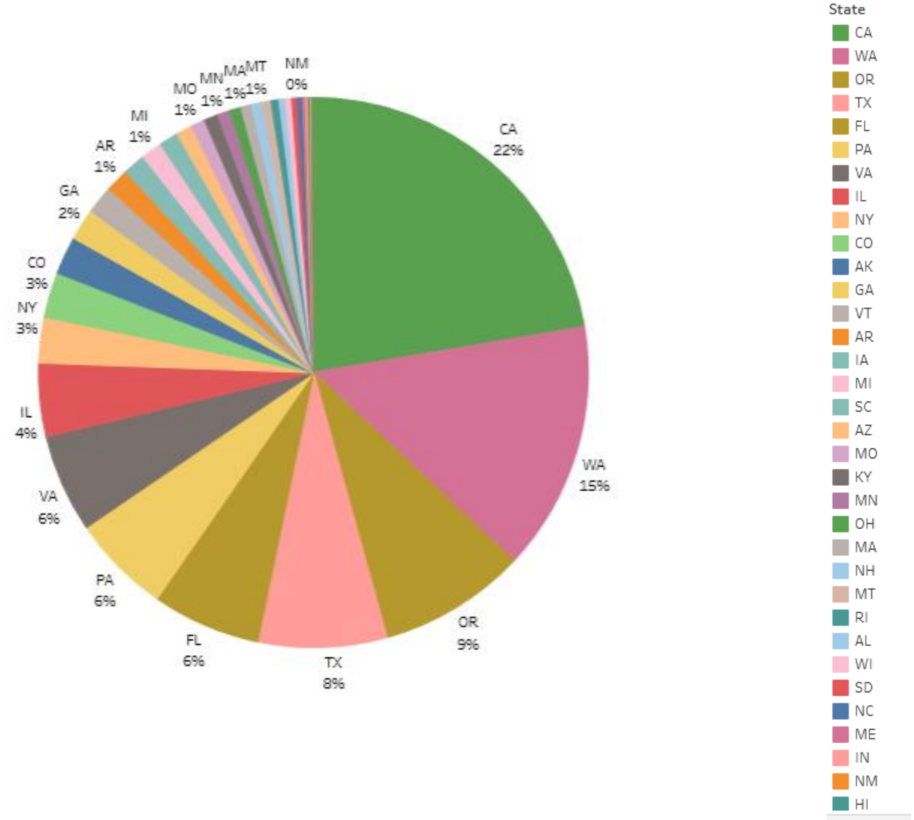
80% FYI spending is constituted by top 8 states that contribute to ~73% of FYI voucher unit counts

STATE

State level representation of earliest FYI adoption and funding observed



State level share of FYI voucher units count



KPI Alert! ☺

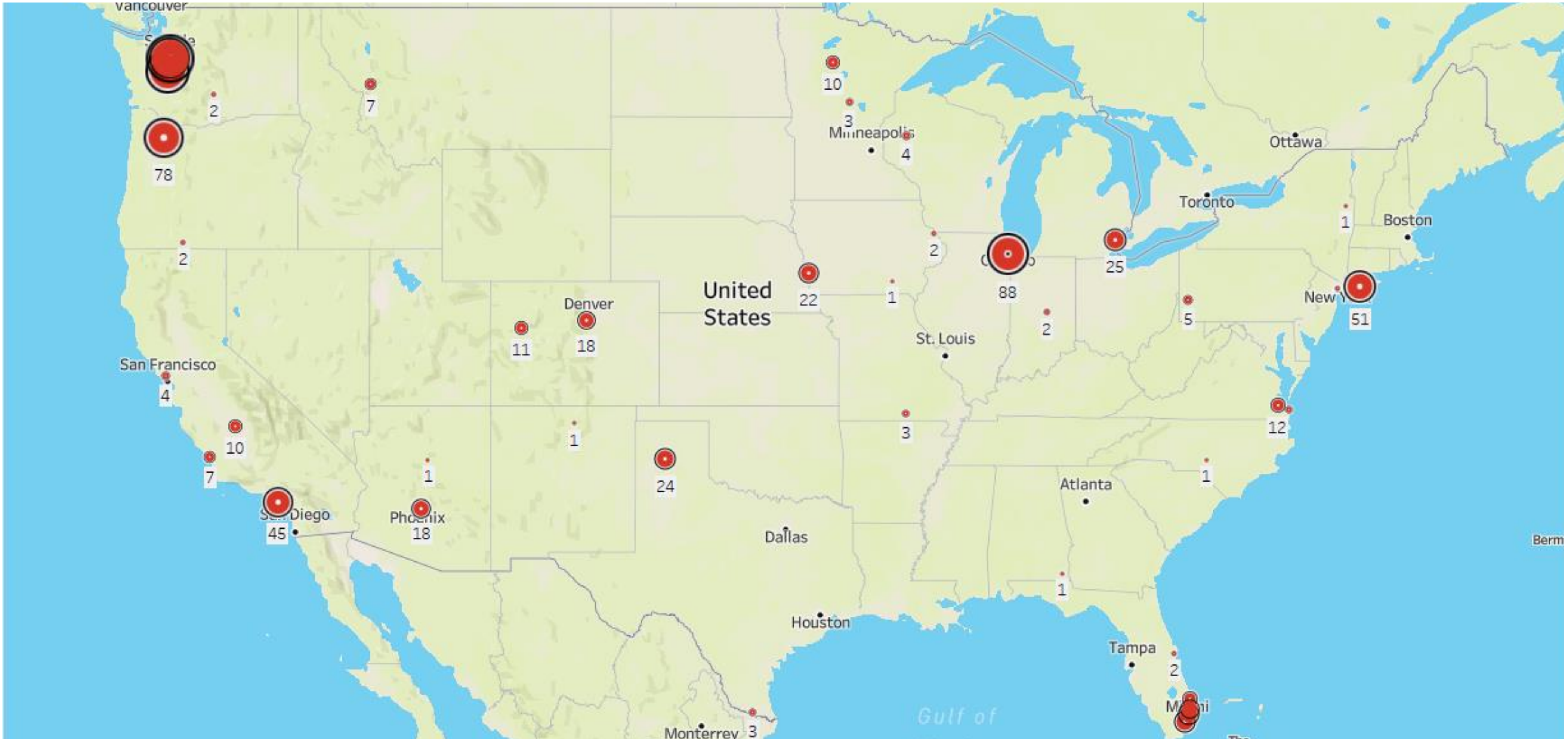
$$\text{FYI voucher units count}^{**} = \frac{FYI\ Funding\ (\$)}{12\ *\ per\ unit\ per\ month\ cost\ (\$)}$$

**It's a directional KPI, absolute values may slightly differ

Cities like Seattle, Chicago, Santa Ana have high FYI voucher units count

PHA-CITY

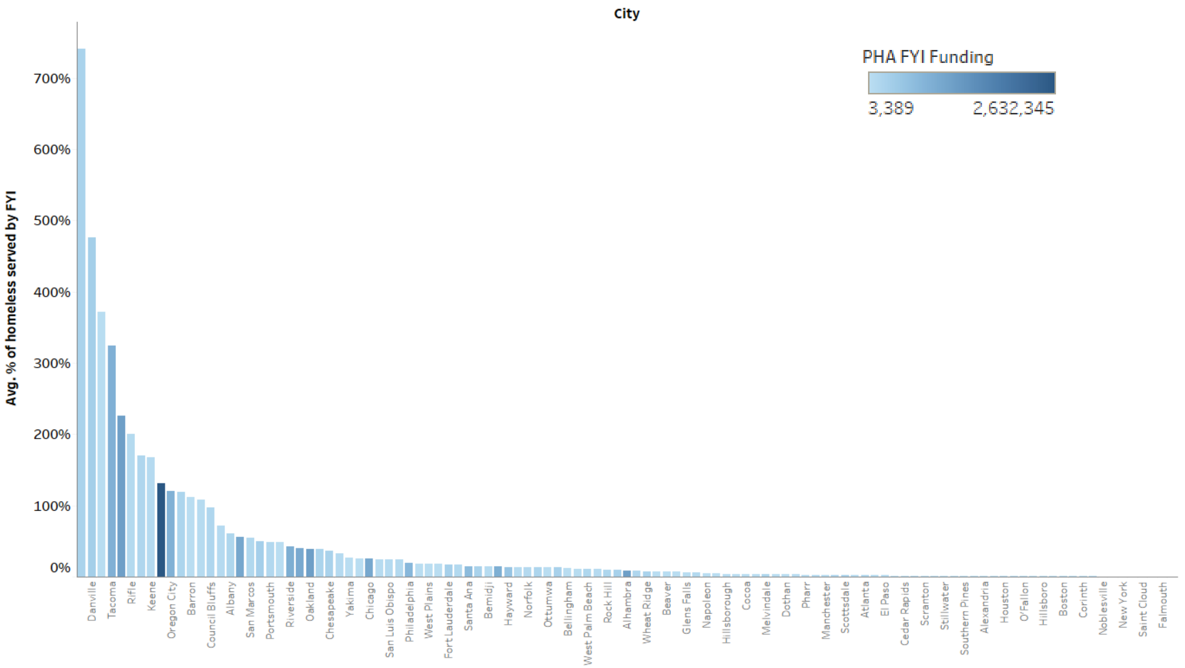
Map of cities with adopted PHAs showcasing FYI voucher units count



~70% of PHA-cities with more than average FYI spending are struggling to meet even 50% of estimated homeless youth demand

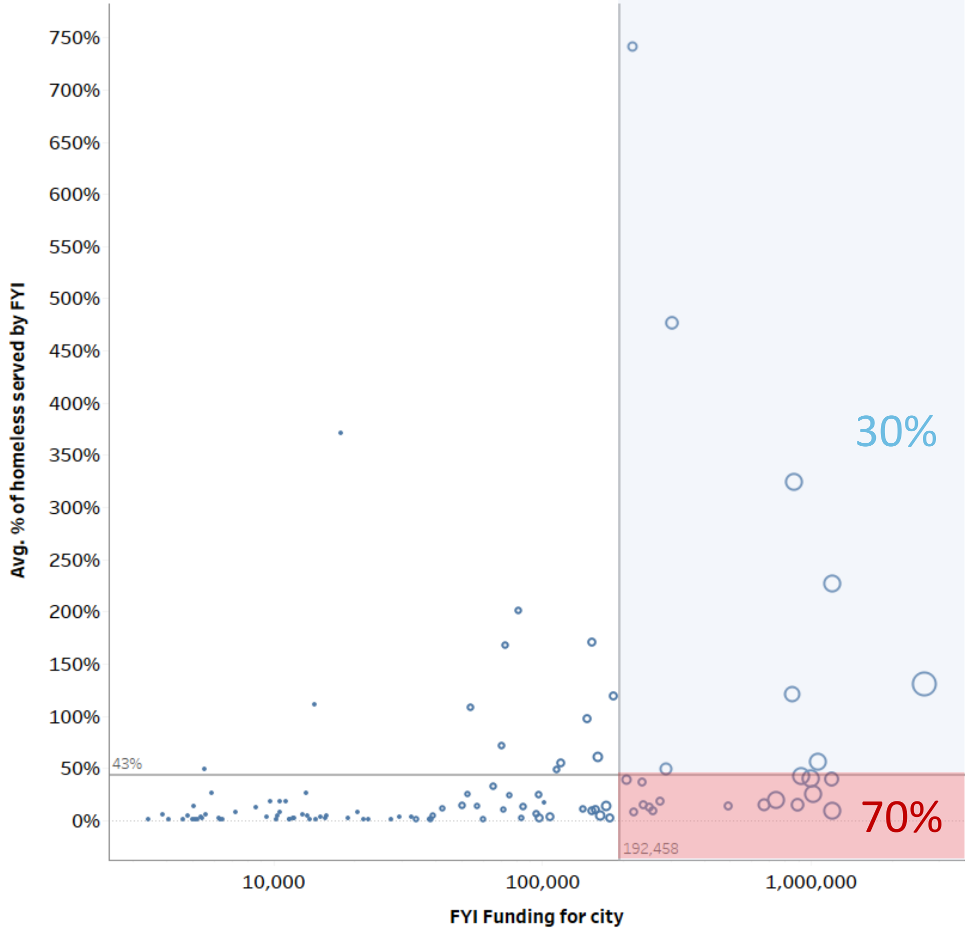
PHA - CITY

Est. avg. homeless youth count served by FYI voucher units per city**



$$\text{Avg. \% of homeless youth demand served by PHA} = \left(\frac{\text{\# FYI voucher units count}}{\text{\#est. avg. homeless youth count}}\right)$$

PHA demand serving standing against FYI funding



KPI Alert! ☺

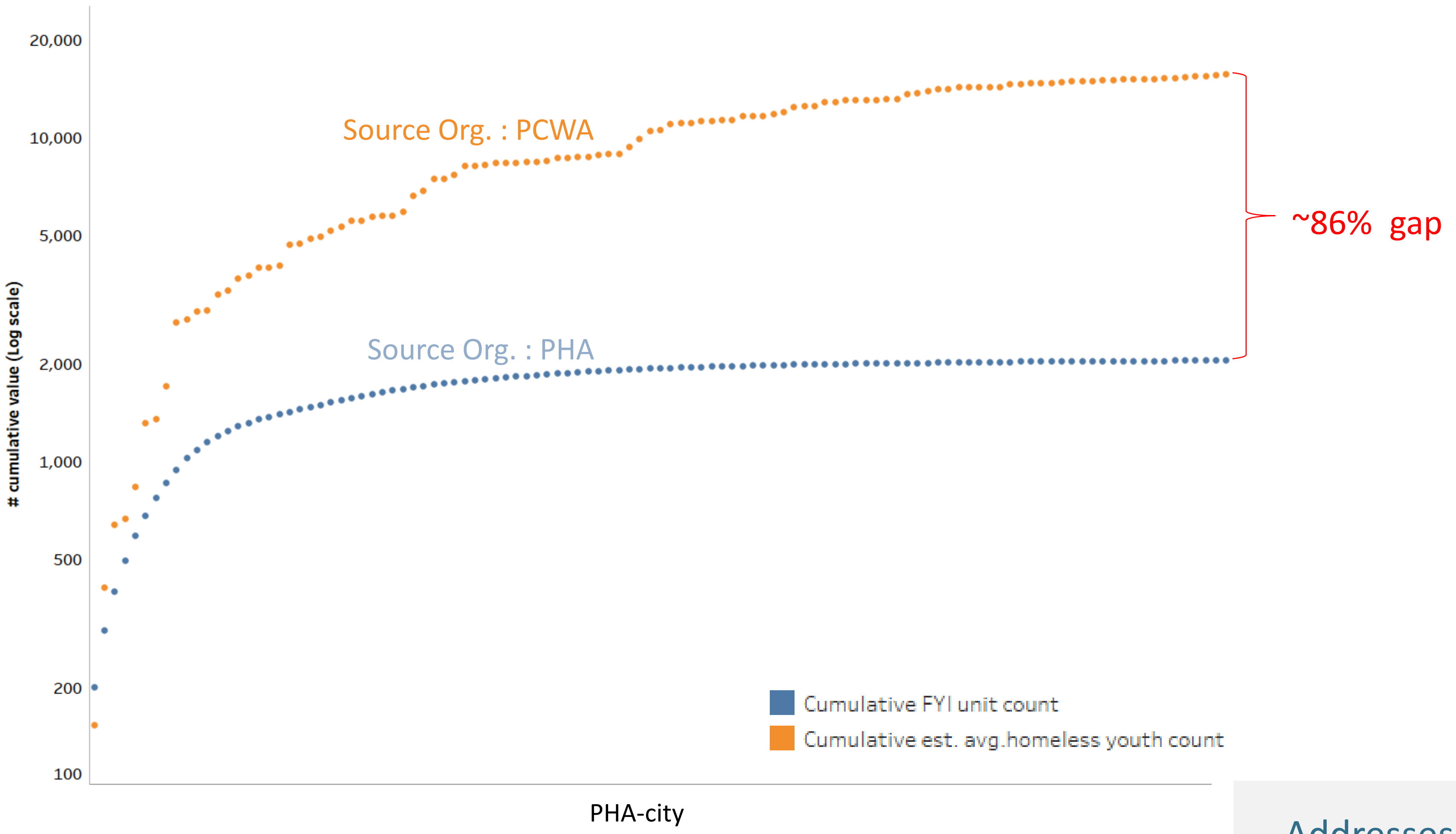
$$\mathit{Est. Avg.}\text{ homeless youth} = \#\ \mathit{of\ state\ total\ homeless\ youth} \times \frac{\mathit{city\ population}}{\mathit{state\ popopulation}}$$

Note : **Estimated avg. homeless youth number is obtained at State level (adoption agency DB) and further distributed using city population w.r.t. state as weight

Overall FYI-adopted PHAs are underserving their homeless youth demand by 86%, indicating a potential friction between PHA and PCWA

PHA-CITY

Cum. log-scale values for each PHA, Est. avg. homeless youth count & FYI voucher units counts

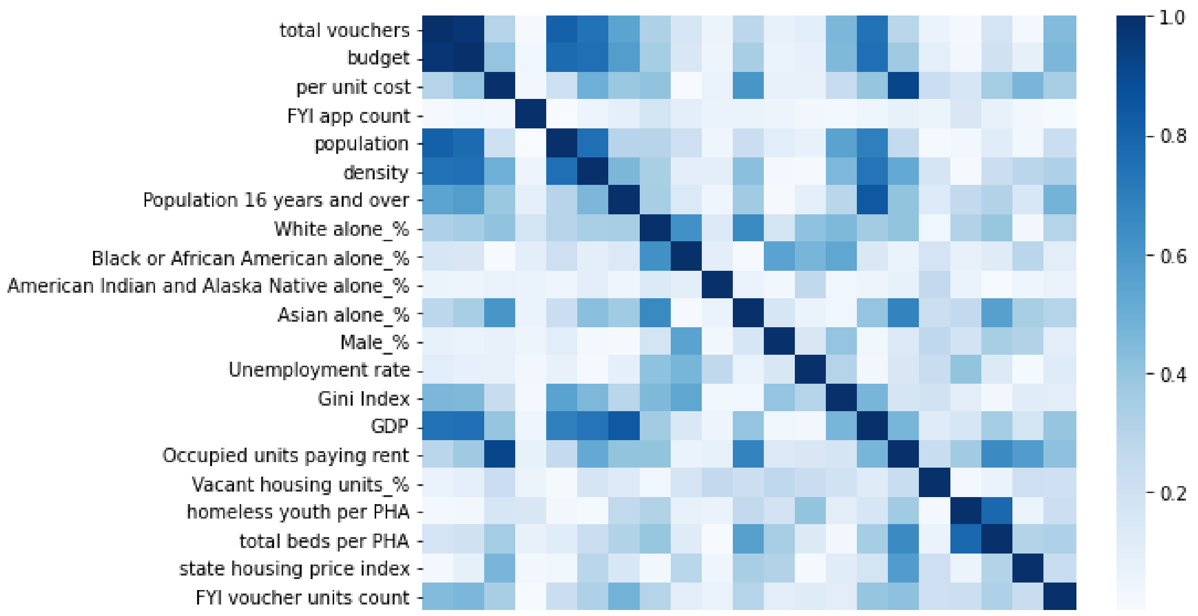


Addresses Research Focus ③

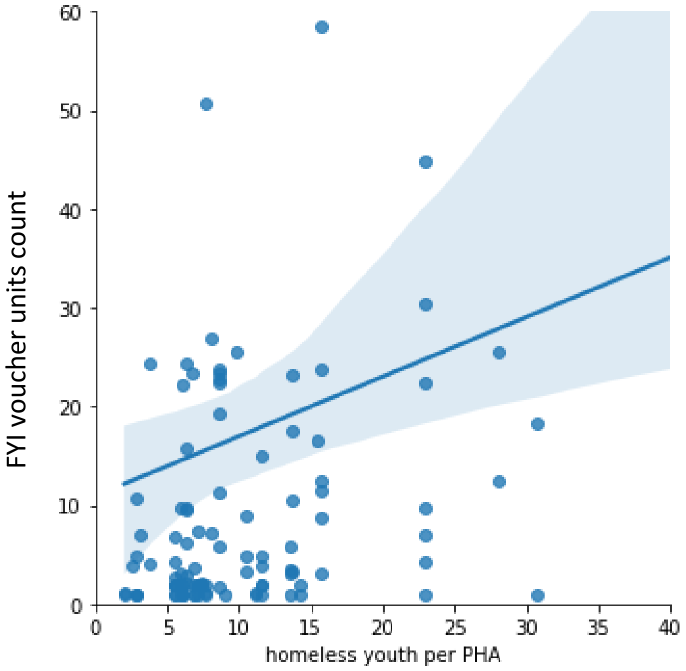
Weak correlation between FYI voucher units count (Y1) and est. homeless youth (Y2) further emphasizes on friction in communication between PHA and PCWA

PHA-CITY

Correlation chart of FYI voucher units count where coef. > 0.01



Weak correlation is observed between Y



Observations

- FYI voucher units count is highly correlated (>0.4) with city population, PHA budget, All kind voucher count, Median rent, GDP
- Possibility of multicollinearity is observed across several features and apt treatment should be chosen

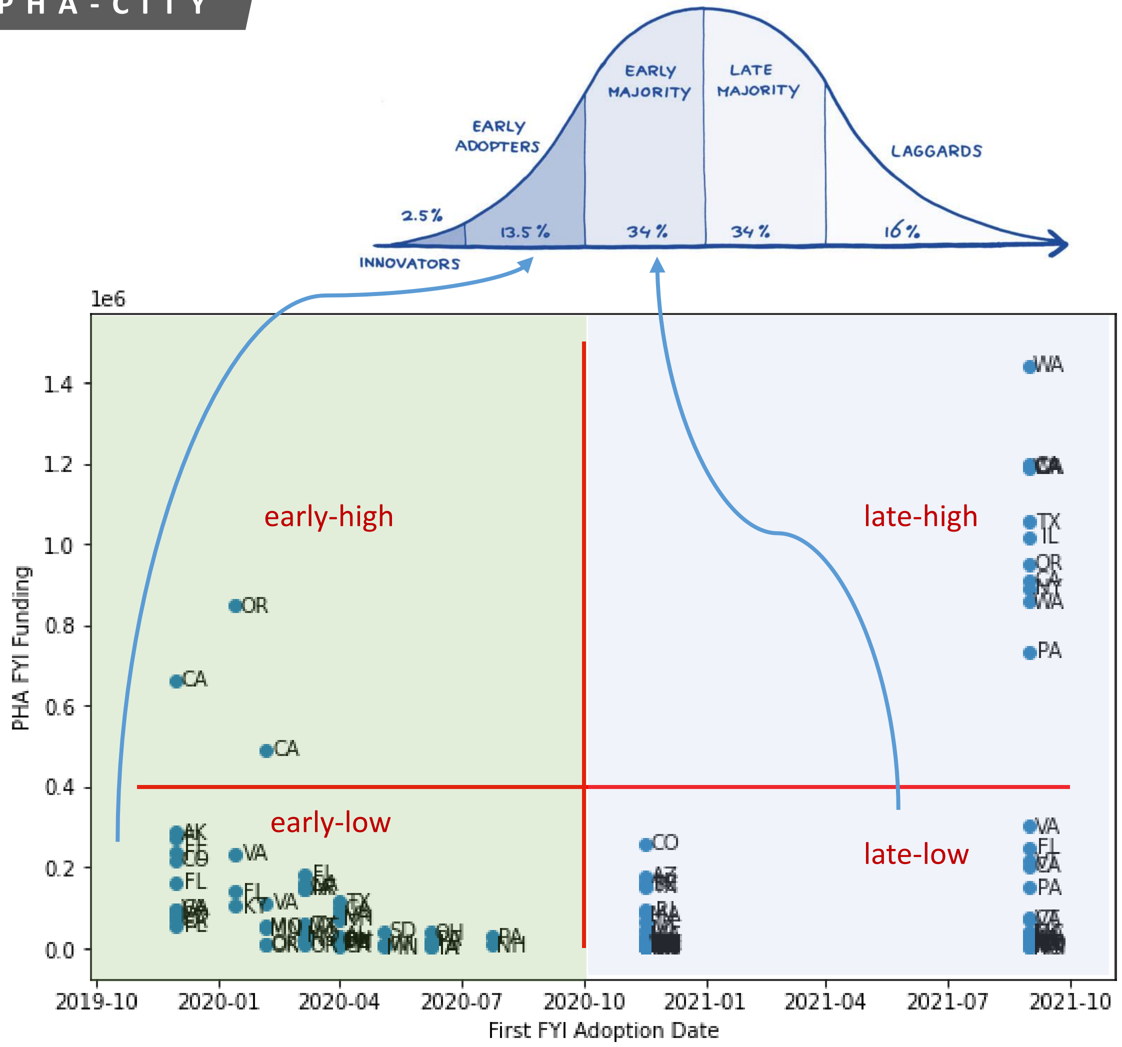
Observations

- Estimated avg. homeless youth count is weakly correlated (0.22) with FYI voucher units count

Addresses Research Focus 3

Diffusion theory can be used to explain the FYI policy penetration among PHAs

PHA-CITY



PHA label	FYI funding	Count
early-high	2,003,919	3
early-low	3,798,784	48
late-high	12,640,418	12
late-low	2,919,727	52

Observations

- Early adopters and early majority have a varied intensity of adoption (FYI funding $)
- About 94% Early adopters with low intensity of FYI contribute to 65% of early FYI total spending
- Important to investigate the factors that contribute to this difference of intensity

Outline

- Problem Framing
- Theoretical Framework
- Data Analysis
  - Data Acquisition
  - Exploratory Data Analysis
  - Modeling
- Insights and Discussion

Models on PHA and counties using machine learning and statistical methods

Methods

PHA Prediction

Regression

Variants of multiple regression will be used to investigate linear dependency of internal and external factors on FYI voucher units count and Est. avg. homeless youth

Ensemble

Models such as **Random Forest** and **Gradient Boosting** are used to best predict the cases where non-linear relationship exists between X & Y

County Target

Unsupervised

To identify promising counties that can adopt FYI, unsupervised methods based on similarities such as **Spectral Clustering** and **PCA** will be used

Statistical

Statistical methods including **hypothesis testing** and **PSM** will be used to find promising counties on homeless characteristics

Tech Stack

python, jupyter, CO, scikit learn, seaborn, H2O AutoML, RStudio, pandas, NumPy, PyTorch, tableau

PHA prediction: after feature engineering, 15 features are finalized to train the model

Data Combination:
Acquired data is filtered, cleaned (JSON extract), restructured, combined and validated

Missing Value Treatment:
Selected niche of dataset is enriched using techniques like MICE, Statistical descriptors

Multicollinearity Check and Treatment:
Conducted VIF test to identify and remove high multicollinearity features

Encoding & Feature development:
Integrate Agency adoption rate related variables using imputed dataset that will help improve statistical inferencing & model prediction

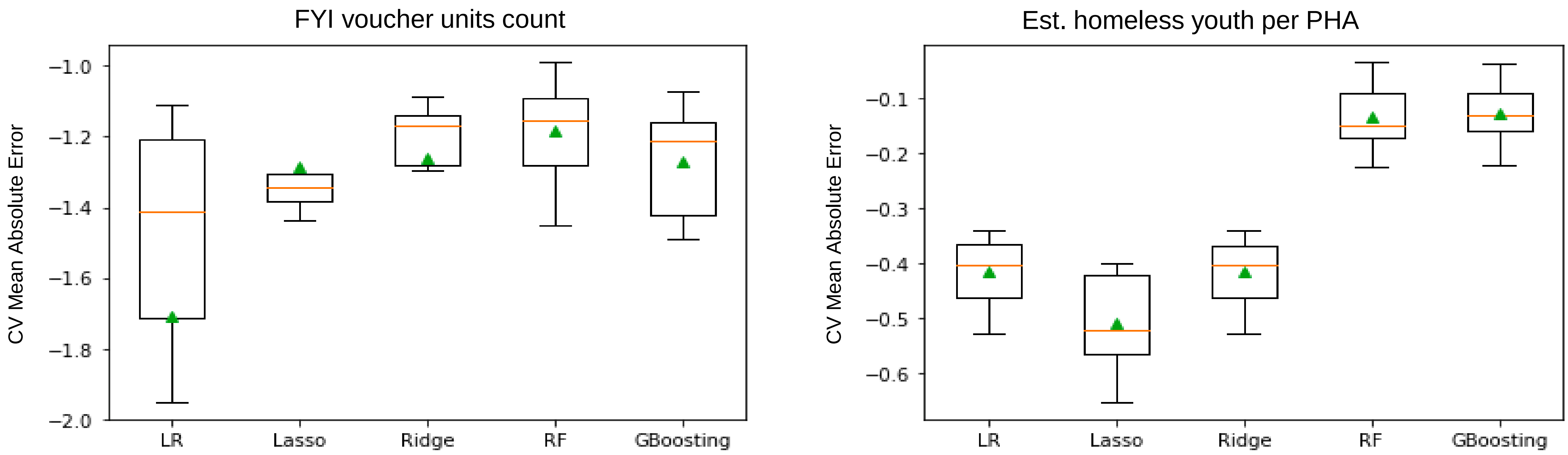
Normalize & Scale:
Experiment with normalization features and scaling to improve model fit

Final Features
PHA Budget
PHA FYI Application Count
PHA # Housing Inventory
PHA # Homeless Youth
City Population
City Density
White Race Percentage
Black or African American Race Percentage
American Indian and Alaska Native Race Percentage
Asian Race Percentage
Unemployment Rate
GDP
Median Rent
Vacant Housing Percentage
Housing Price Index

Random Forest achieves smallest MAE in predicting FYI voucher units count and est. homeless youth per PHA

P H A

Performance of models predicting FYI voucher units count and Est. homeless youth per PHA



Observations

- A Stratified K-fold cross validation of Random Forest model and Gradient Boosting resulted in lowest MAE with 5 features selected using Recursive Feature Elimination method
- Random forest model outperforms as :
  - it works well with non-linearly related feature set
  - it takes advantage of bagging (variance reduction) and boosting (bias reduction)

Important factors and interpretation of Random Forest model

PHA

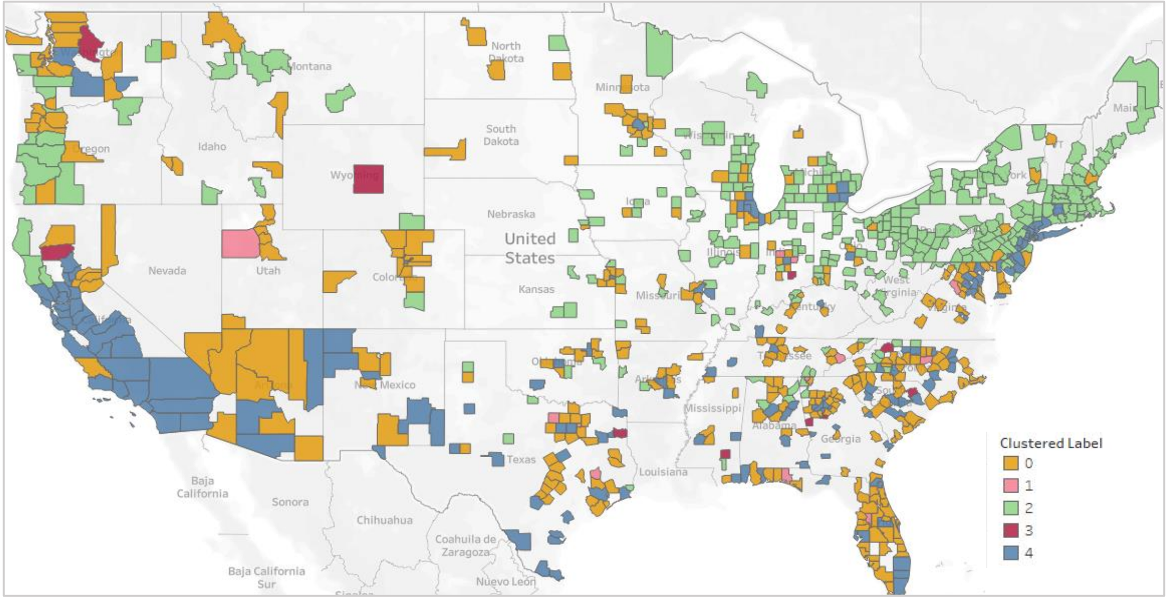
Y	Top 5 Important Features	Interpretation
FYI Voucher Units Count	PHA Budget County GDP City Density Est. Homeless Youth per PHA County Median Rent	• FYI adoption level is highly associated with community economic levels: PHA Budget, County GDP, City Density and County Median • Inadequate cooperation between PHA and PCWA: # Homeless Youth Per PHA, the data from PCWA, however, is expected to rank higher but only ranks 4
Est. Homeless Youth per PHA	PHA # Housing Inventory State Housing Price Index County Unemployment Rate City Density County American Indian and Alaska Native alone %	• PHA housing supply is highly related with its Est. homeless youth • Socioeconomic factors including Housing price, urbanicity (density) and unemployment rate are good predictors for homeless • Race structure also impacts homeless condition

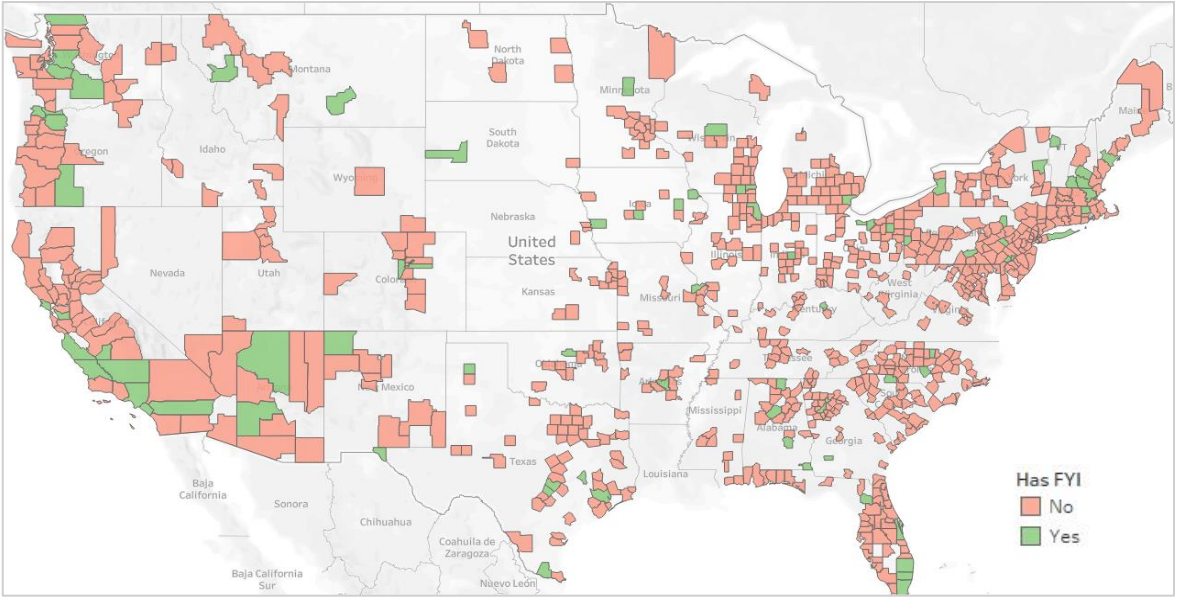
Addresses Research Focus ④

Potential county target: counties clustered in 5 groups based on 34 unique socioeconomic factors

COUNTY

Performance of models predicting FYI vouchers unit count and Est. homeless youth





Clustered Label	Count of County	County with FYI	Adoption %
0	309	23	7%
1	12	0	0%
2	297	28	9%
3	10	0	0%
4	184	34	18%

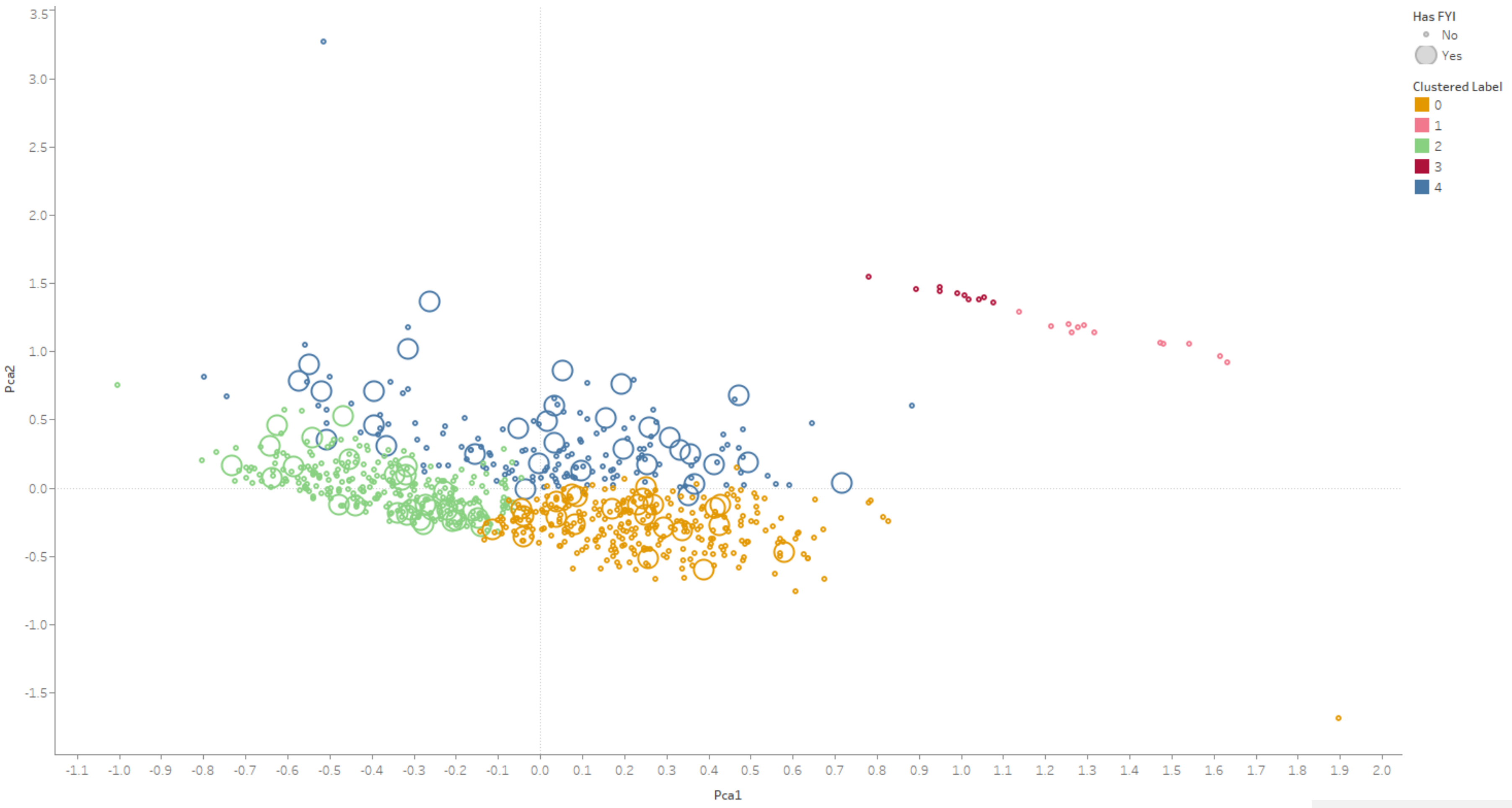
Observations

- Clustering based on Socioeconomic factors result in 5 groups
- Cluster 4 counties have highest adoption rate
- Cluster 4 counties are highly developed areas or metro cities
- Since Cluster 4 has highest adoption rate the member counties of this cluster can be the next potential adopters

Visualizing next potential adopter counties: Cluster 4 (Blue)

COUNTY

Principle Component Analysis to visualize clusters



Addresses Research Focus 5

Important characteristics of homeless influence decisions in adopting FYI

COUNTY

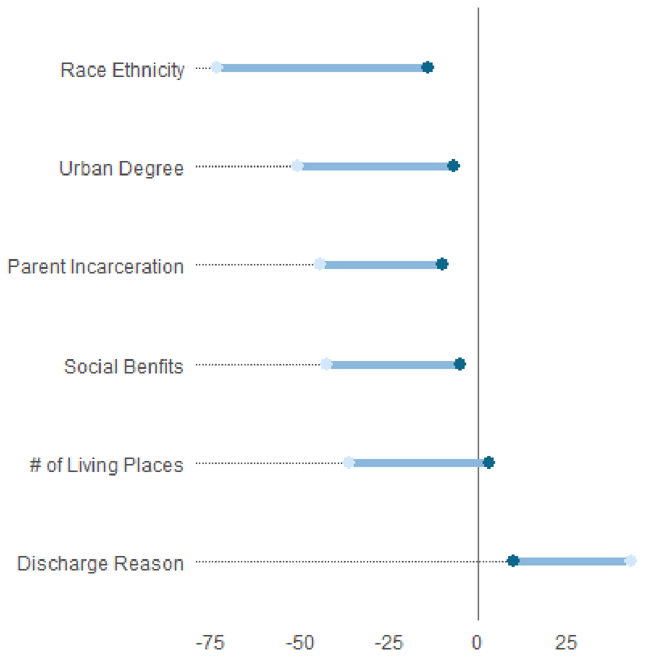
Using dataset AFCARS, and comparing 33 counties that have adopted FYI and 92 counties that do not have adopted FYI, we found these two groups of counties are SIMILAR except:

Variables	Mean of 33 counties (1)	Mean of 92 counties (0)	P-value
Avg. rate of **homeless youth's race being white**	0.44	0.54	0.0373
Average rate of **homeless youth being white and non-Hispanic**	0.25	0.37	0.000418
Avg. rate of **homeless youth being in Metro areas**	0.53	0.73	0.0108
Avg. rate of **homeless youth receiving support under Title XVI** or other Social Security Act titles	0.038	0.052	0.0343
Avg. number of no case plan goal has yet been established other than the care and protection of the youth	65.39	52.93	0.0433
Avg. rate of removal reason as parent incarceration	0.048	0.069	0.0131
The number of places the youth has lived, including the current setting, during the current removal episode	55	59	0.0346
Avg. rate of homeless youth reaching majority according to the law by virtue of age, marriage	0.04	0.03	0.0405

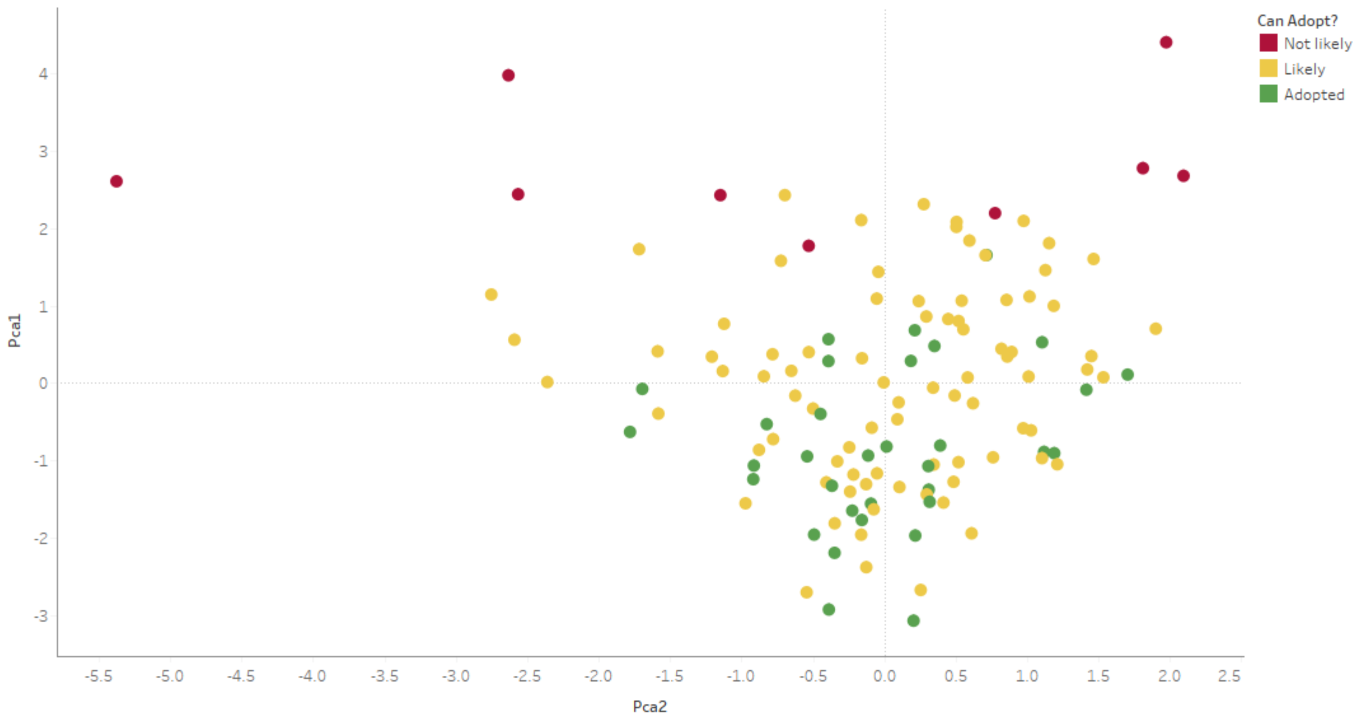
Use propensity score matching (PSM) to identify potential counties adopting FYI

COUNTY

Bias reduction in variables that distinguish groups



Three groups of counties based on the likelihood of FYI adoption



Observations

- The selection bias is largely reduced after propensity score matching
- Based on that, three groups of counties are effectively identified: 33 counties that have already adopted FYI vouchers, 83 that are likely to adopt, and 9 who are not likely to adop.

Addresses Research Focus 5

Outline

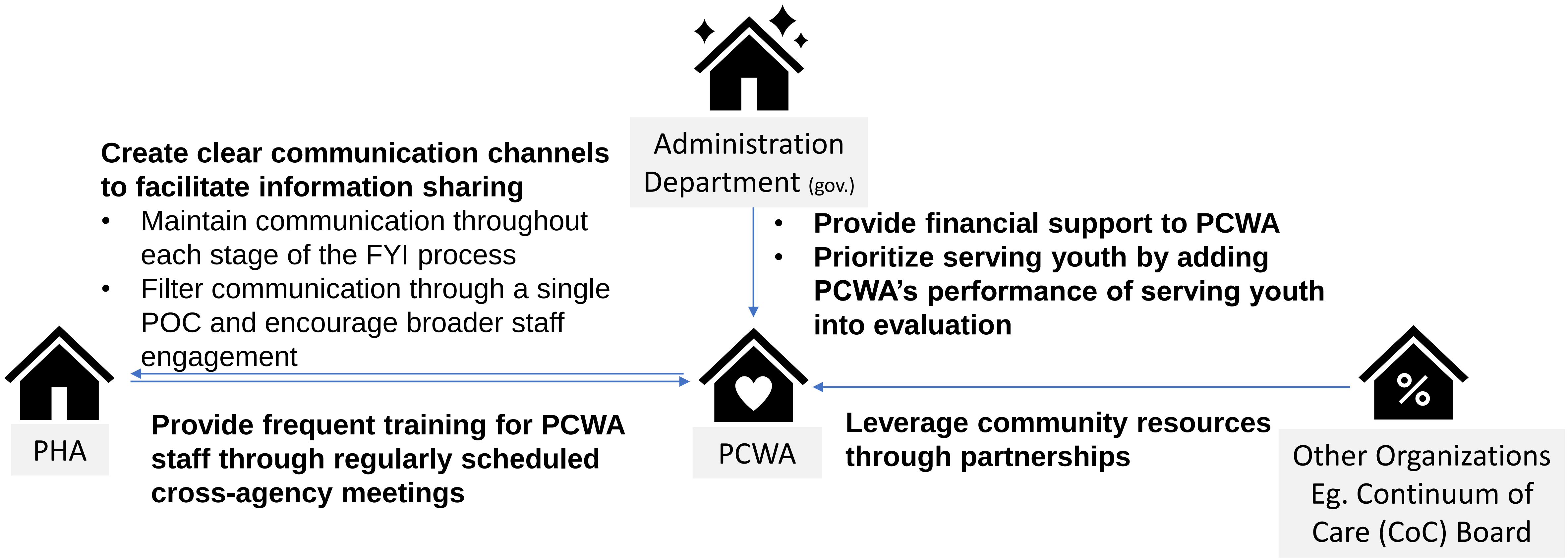
- Problem Framing
- Theoretical Framework
- Data Analysis
- **Insights and Discussion**

Policy Evaluation: effectiveness proved, and suggestions provided

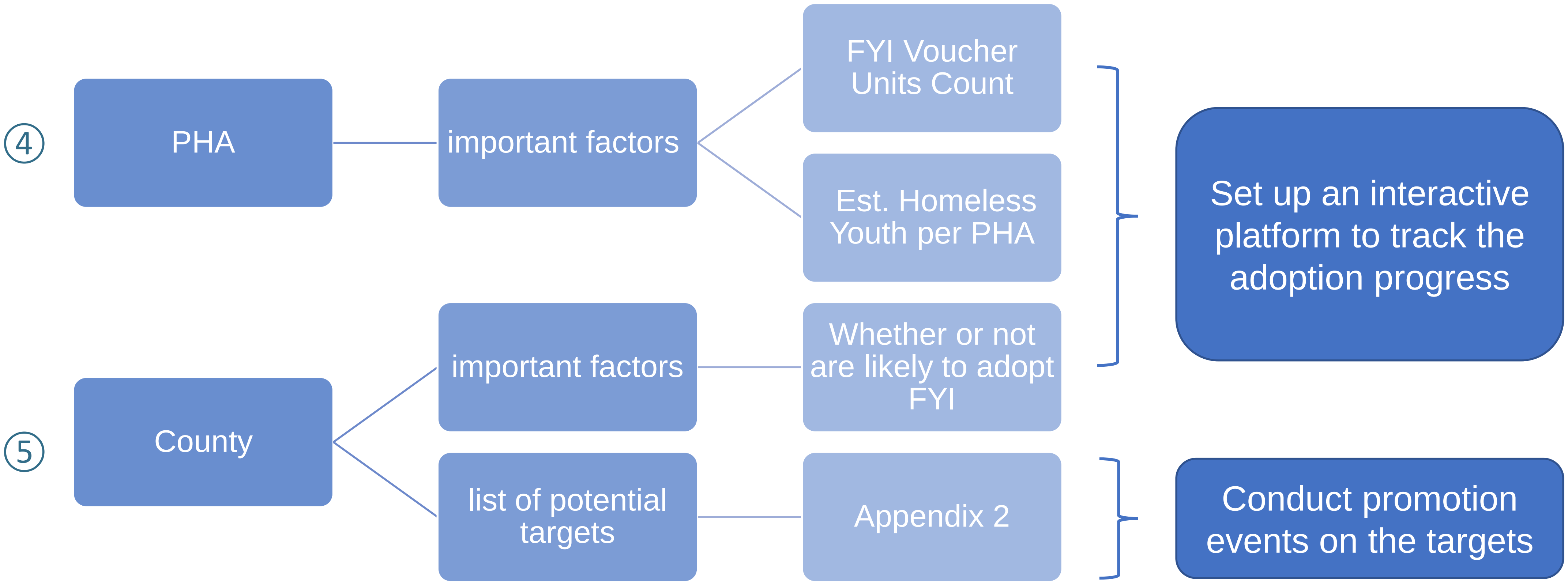
① FYI Policy iterations are proving effective as FYI Funding boost and more PHAs adopt in recent past

② The utilization of FYI budget is only ~70%, indicating FYI budget is not a major limit

③ Multiple data evidences indicate inadequate cooperation between PHA and PCWA. Suggestions:



Promotion Target: Put more effort on PHAs and counties with addressed features



Future Mission of Gandalf Tech…

- Conduct surveys and informational interviews in local agencies to identify important factors related to internal administrative and organizational behaviors;
- Employ natural language processing and sentiment analysis on stakeholders' comments to evaluate the performance of FYI program;
- Explore the further evaluation of FYI program in a broader way, focusing not only the voucher adoption level of local agencies but also the final changes in homeless youth as the end-user of the policy.
- Gandalf Tech will always be devoted at **Tech for Social Good**

THANK YOU! ☺

Appendix 1: all county-level variable names from AFCARS

#9 Child's Date of Birth
#10 Child Sex
#11 Child Race American Indian or Alaska Native
#12 Child Race Asian
#13 Child Race Black or African American
#14 Child Race Native Hawaiian or Other Pacific Islander
#79 Title IV-E Adoption Assistance
#80 Title IV-A TANF Payment
#81 Title IV-D Child Support Funds
#82 Title XIX Medicaid
#83 SSI or Social Security Act Benefits
#84 Only State or Other Support
#85 Monthly Foster Care Payment
#86 Length (Days) Since Latest Removal
#87 Length (Days) in Current Placement Setting
#88 Length (Days) of Previous FC Stay
#89 Total Days Stay in FC, All Episodes
#90 Age on the First Day of the Fiscal Year
#91 Age at Most Recent Removal/Entry into Foster Care
#92 Age of Child at the End of FFY, or at Exit
#93 Child was in FC at the Beginning of the FFY
#94 Child was in FC at the End of the Fiscal Year
#95 Entered Foster Care During the Fiscal Year
#96 Child was Discharged from Foster Care During the Fiscal Year
#97 Child was in at Start or Entered FC During the FY
#98 Child is Waiting for Adoption
#99 Parents Rights Have Been Terminated
#100 Youth is No Longer Eligible for Foster Care Due to Age
#101 Derived Child Race and Ethnicity
This variable is derived from the five variables
#102 Derived Race
This variable is derived using the five variables for the child's race
#103 Rural Urban Continuum Code
#104 State Foster Care ID

#15 Child Race White
#16 Child Race Unable to Determine
#17 Child Hispanic or Latino Ethnicity
#18 Child Has Been Clinically Diagnosed with Disability
#19 Mental Retardation
#20 Visually or Hearing Impaired
#21 Physically Disabled
#22 Emotionally Disturbed
#23 Other Medically Diagnosed Condition Requiring Special Care
#24 Child Has Previously Been Adopted
#25 Age on Date of Legal Adoption
#26 Date of First Removal
#27 Total Number of Removals from Home
#28 Discharge Date for Previous Removal
#29 Date of Latest Removal from Home
#30 Removal Transaction Date
#31 Begin Date for Current Placement Setting
#32 Number of Placement Settings in Current FC Episode
#33 Removal Manner
#34 Removal Reason-Physical Abuse
#35 Removal Reason-Sexual Abuse
#36 Removal Reason-Neglect
#37 Removal Reason-Alcohol Abuse Parent
#38 Removal Reason-Drug Abuse Parent
#39 Removal Reason-Alcohol Abuse Child
#40 Removal Reason-Drug Abuse Child
#41 Removal Reason-Child Disability
#42 Removal Reason-Child Behavior Problem
#43 Removal Reason-Parent Death
#44 Removal Reason-Parent Incarceration
#45 Removal Reason-Caretaker Inability Cope
#46 Removal Reason-Abandonment

#47 Removal Reason-Relinquishment
#48 Removal Reason-Inadequate Housing
#49 Current Placement Setting
#50 The Current Placement Setting is Outside the State
#51 Most Recent Case Plan Goal
#52 Caretaker Family Structure
#53 1st Principal Caretaker Year of Birth
#54 2nd Principal Caretaker Year of Birth
#55 Termination Date of Parental Rights-Mom
#56 Termination Date of Parental Rights-Dad
#57 Date of Parents Loss of Parental Rights
#58 Foster Family Structure
#59 1st Foster Caretaker Year of Birth
#60 2nd Foster Caretaker Year of Birth
#61 1st Foster Caretaker Race American Indian or Alaska Native
#62 1st Foster Caretaker Race Asian
#63 1st Foster Caretaker Race Black or African American
#64 1st Foster Caretaker Race Native Hawaiian or Other Pacific Islander
#65 1st Foster Caretaker Race White
#66 1st Foster Caretaker Race Unable to Determine
#67 1st Foster Caretaker Hispanic or Latino Ethnicity
#68 2nd Foster Caretaker Race American Indian or Alaska Native
#69 2nd Foster Caretaker Race Asian
#70 2nd Foster Caretaker Race Black or African American
#71 2nd Foster Caretaker Race Native Hawaiian or Other Pacific Islander
#72 2nd Foster Caretaker Race White
#73 2nd Foster Caretaker Race Unable to Determine
#74 2nd Foster Caretaker Hispanic or Latino Ethnicity
#75 Date of Discharge from Foster Care
#76 Date that the Discharge Was Recorded
#77 Discharge Reason
#78 Title IV-E Foster Care Payments

Appendix 2: list of counties in three groups

FIPSCode	county	state	label
1073	Jefferson	AL	2
2020	Anchorage Municipality	AK	2
4013	Maricopa	AZ	2
6001	Alameda	CA	2
6029	Kern	CA	2
6037	Los Angeles	CA	2
6059	Orange	CA	2
6065	Riverside	CA	2
8059	Jefferson	CO	2
12009	Brevard	FL	2
12011	Broward	FL	2
12086	Miami-Dade	FL	2
12099	Palm Beach	FL	2
13121	Fulton	GA	2
15003	Honolulu	HI	2
17031	Cook	IL	2
19153	Polk	IA	2
21067	Fayette	KY	2
25017	Middlesex	MA	2
25025	Suffolk	MA	2
26163	Wayne	MI	2
27007	Beltrami	MN	2
30111	Yellowstone	MT	2
36029	Erie	NY	2
36061	New York	NY	2
41051	Multnomah	OR	2
42101	Philadelphia	PA	2
44007	Providence	RI	2
48201	Harris	TX	2
48215	Hidalgo	TX	2
48453	Travis	TX	2
53033	King	WA	2
53053	Pierce	WA	2

FIPSCode	county	state	label
4019	Pima	AZ	1
4021	Pinal	AZ	1
6013	Contra Costa	CA	1
6019	Fresno	CA	1
6067	Sacramento	CA	1
6071	San Bernardino	CA	1
6073	San Diego	CA	1
6075	San Francisco	CA	1
6077	San Joaquin	CA	1
6085	Santa Clara	CA	1
6099	Stanislaus	CA	1
6107	Tulare	CA	1
6111	Ventura	CA	1
8001	Adams	CO	1
8031	Denver	CO	1
8041	El Paso	CO	1
9003	Hartford	CT	1
9009	New Haven	CT	1
11001	District of Columbia	DC	1
12033	Escambia	FL	1
12057	Hillsborough	FL	1
12081	Manatee	FL	1
12083	Marion	FL	1
12095	Orange	FL	1
12103	Pinellas	FL	1
12127	Volusia	FL	1
13089	DeKalb	GA	1
17143	Peoria	IL	1
17163	St. Clair	IL	1
17167	Sangamon	IL	1
17201	Winnebago	IL	1
18003	Allen	IN	1
18089	Lake	IN	1
18095	Madison	IN	1
18097	Marion	IN	1
21111	Jefferson	KY	1

24510	Baltimore city	MD	1
25005	Bristol	MA	1
25009	Essex	MA	1
25013	Hampden	MA	1
25023	Plymouth	MA	1
25027	Worcester	MA	1
26081	Kent	MI	1
26125	Oakland	MI	1
27053	Hennepin	MN	1
27123	Ramsey	MN	1
27137	St. Louis	MN	1
28047	Harrison	MS	1
29095	Jackson	MO	1
29189	St. Louis	MO	1
31055	Douglas	NE	1
32003	Clark	NV	1
34007	Camden	NJ	1
34013	Essex	NJ	1
35001	Bernalillo	NM	1
37051	Cumberland	NC	1
39035	Cuyahoga	OH	1
39049	Franklin	OH	1
39061	Hamilton	OH	1
39095	Lucas	OH	1
39113	Montgomery	OH	1
39153	Summit	OH	1
40109	Oklahoma	OK	1
40143	Tulsa	OK	1
42003	Allegheny	PA	1
42011	Berks	PA	1
45045	Greenville	SC	1
47037	Davidson	TN	1
47093	Knox	TN	1
47157	Shelby	TN	1
48027	Bell	TX	1
48029	Bexar	TX	1
48113	Dallas	TX	1
48121	Denton	TX	1
48303	Lubbock	TX	1
48309	McLennan	TX	1
48439	Tarrant	TX	1
49035	Salt Lake	UT	1
53011	Clark	WA	1
53061	Snohomish	WA	1
53063	Spokane	WA	1
54039	Kanawha	WV	1
55079	Milwaukee	WI	1

FIPSCode	county	state	label
5131	Sebastian	AR	0
12031	Duval	FL	0
12071	Lee	FL	0
12101	Pasco	FL	0
12105	Polk	FL	0
18163	Vanderburgh	IN	0
20173	Sedgwick	KS	0
32031	Washoe	NV	0
41039	Lane	OR	0

Reference

- HUD announced FYI
- FYI introduction
- Market Predictors of Homelessness
- FYI initiative webinar series data analytics