Predicting Risk of Youth Homeless in Los Angeles
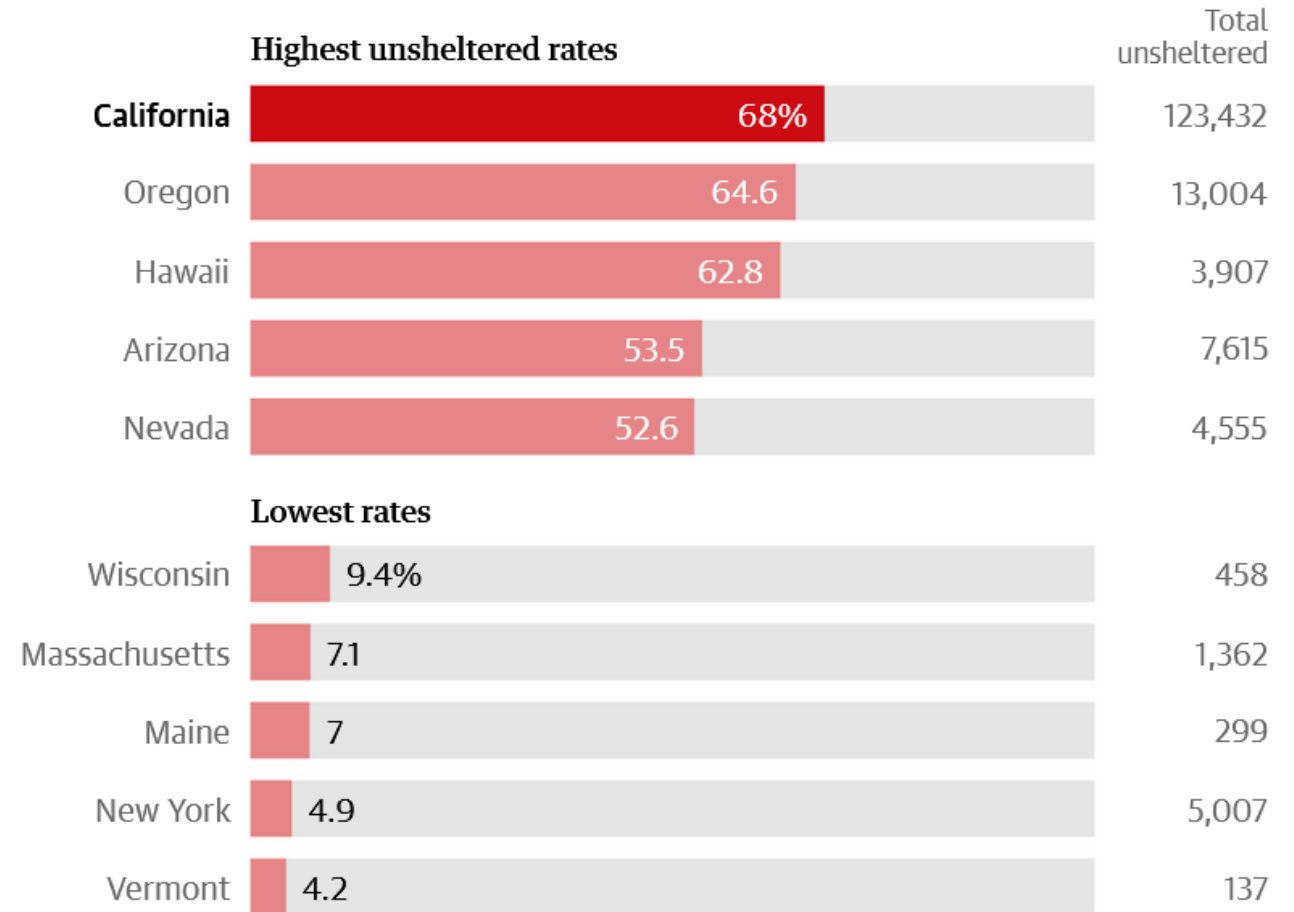
By

Daniel Kim

Current Homeless Situation in California

California had the highest rate of people experiencing homelessness who were unsheltered

Unsheltered refers to people living outside in tents, cars or other makeshift shelters

Highest unsheltered rates

State	Rate	Total unsheltered
California	68%	123,432
Oregon	64.6	13,004
Hawaii	62.8	3,907
Arizona	53.5	7,615
Nevada	52.6	4,555

Lowest rates

State	Rate	Total unsheltered
Wisconsin	9.4%	458
Massachusetts	7.1	1,362
Maine	7	299
New York	4.9	5,007
Vermont	4.2	137

Los Angeles City

Homeless Population Counts in Los Angeles County
All Los Angeles County & City of Los Angeles

Numbers for all Los Angeles County communities are included.

Year	City of Los Angeles			All Los Angeles County		
	Total	Unsheltered	Sheltered	Total	Unsheltered	Sheltered
2023	46,260	32,680	13,580	75,518	55,155	20,363
2022	41,980	28,458	13,522	69,144	48,548	20,596
2021	No count counducted this year due to the COVID-19 pandemic.					
2020	41,290	28,852	12,438	66,436	48,041	18,395
2019	36,300	27,221	8,944	58,936	42,471	13,786
Year	City of Los Angeles			All Los Angeles County		
	Total	Unsheltered	Sheltered	Total	Unsheltered	Sheltered
2018	31,285	22,887	8,398	52,765	39,396	13,369
2017	34,189	25,237	8,952	57,794	42,828	14,966
2015	25,686	17,687	7,999	44,359	31,018	13,341
2013	22,993	14,968	8,025	39,463	25,136	14,327
2011	-	-	-	39,135	20,517	19,176

Homeless Youth in Los Angeles

Homeless Young People
Point-in-Time Count, 2018 Through 2023

We still await the release of 2023 numbers for some categories and the City of Los Angeles.

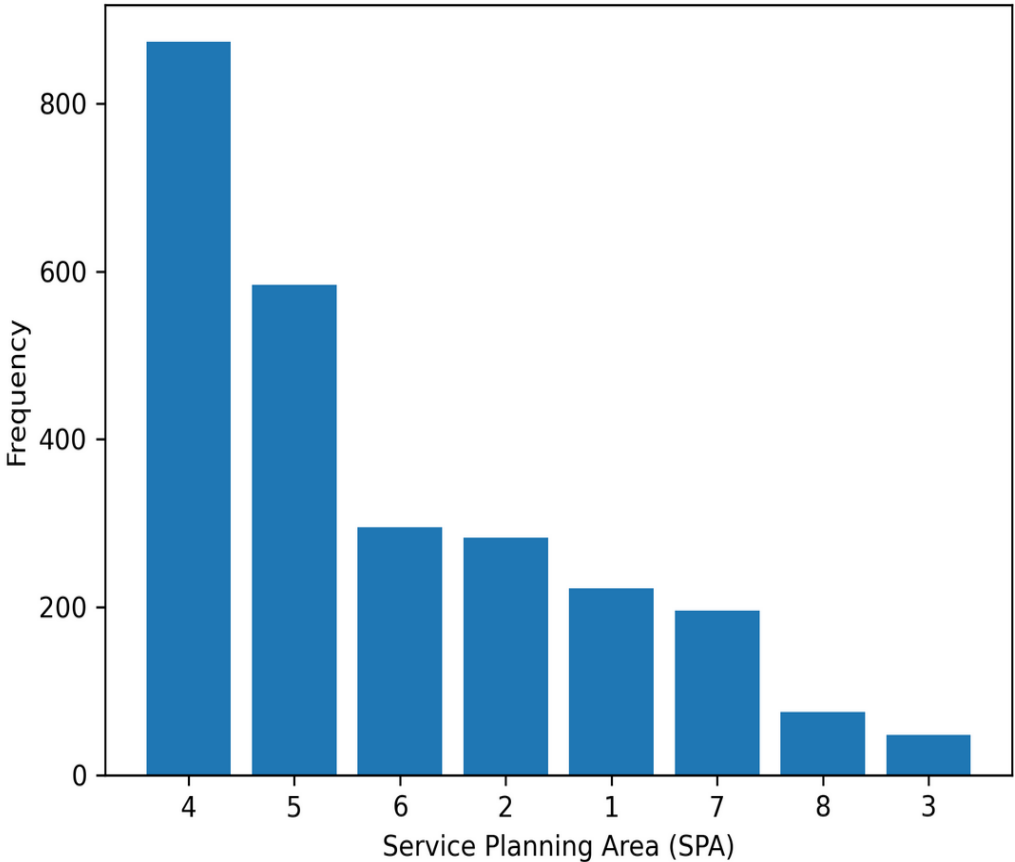| Youth Age Group	City of Los Angeles					All of Los Angeles County*				
	2023	2022	2020	2019	2018	2023	2022	2020	2019	2018
Transitional Age Youth (18-24): **Total**	---	**1,037**	**2,083**	**1,833**	**1,573**	**3,718**	**2,786**	**3,029**	**2,818**	**2,274**
Transitional Age Youth (18-24): **Sheltered**	---	574	613	569	576	**1,593**	1,747	1,005	931	864
Transitional Age Youth (18-24): **Unsheltered**	---	463	1,470	1,237	997	**2,125**	1,039	2,075	1,887	1,410
Unaccompanied Minors (Under 18): **Total**	---	**90**	**17**	**54**	**58**	---	**121**	**69**	**66**	**63**
Unaccompanied Minors (Under 18): **Sheltered**	---	88	13	21	31	---	116	25	21	35
Unaccompanied Minors (Under 18): **Unsheltered**	---	2	4	33	27	---	5	44	45	28

* Numbers do not include those for Glendale, Long Beach and Pasadena.

Problem Statement

California leads the country in [youth homelessness](). To help address this growing issue in Los Angeles, this project's aim is to use classification models leveraging public data on youth homelessness surveys to predict who is at risk of becoming homeless based on family conditions and geographic consideration. The goal of this project is to help LA city government gain a better understanding of who might be most at risk by utilizing binary classification models. To enhance the models' accuracies, I will incorporate domain expertise and external research to conduct feature selection, identifying relevant factors influencing the likelihood of becoming homeless.

EDA - SPA

Number of Youth Homeless Survey Participants by SPA (2017-2019)
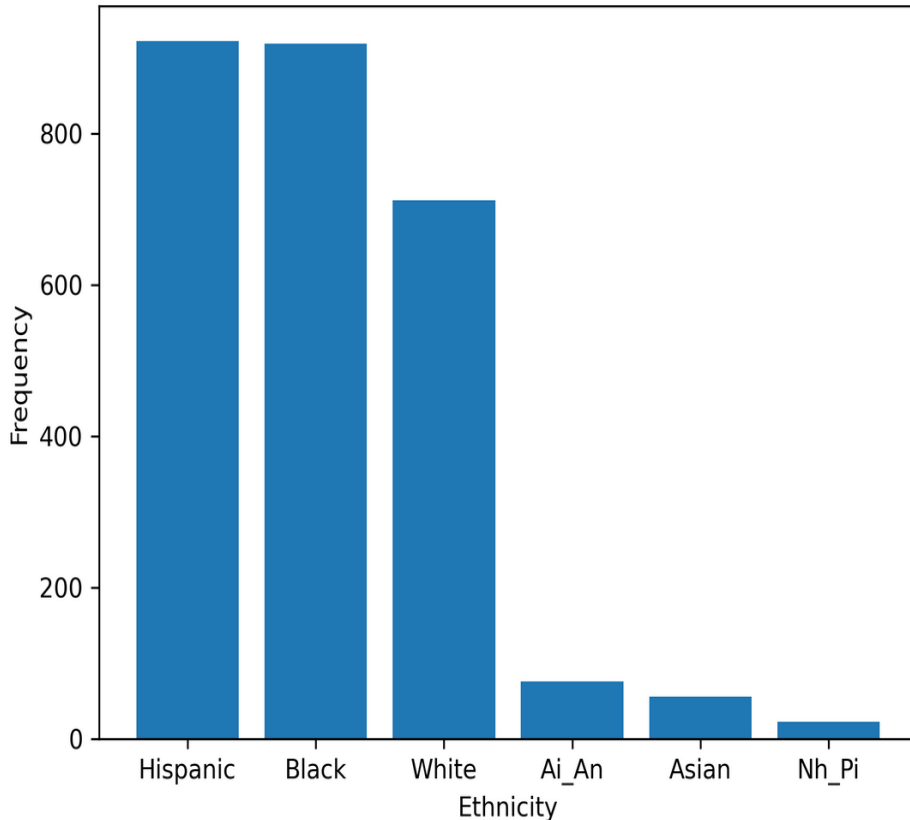
Homeless Population by Geographic Area
By Service Planning Areas (SPAs)

Numbers include those for Glendale (SPA 2), Long Beach (SPA 8) and Pasadena (SPA 3)

Year	Antelope Valley (SPA 1)	San Fernando Valley (SPA 2)	San Gabriel Valley (SPA 3)	Metro LA (SPA 4)	West LA (SPA 5)	South LA (SPA 6)	East LA County (SPA 7)	South Bay/ Harbor (SPA 8)
2023	4,686	10,443	5,009	18,531	6,669	12,995	6,511	7,775
2022	4,598	9,604	4,661	17,820	4,604	14,598	4,781	4,445
2021	No count counducted this year due to the COVID-19 pandemic.							
2020	4,755	9,277	5,082	17,121	6,009	13,012	4,586	6,594
2019	3,293	7,730	4,489	16,436	5,262	9,543	5,095	4,409
2018	3,203	7,738	4,282	14,218	4,401	8,343	4,569	6,011
2017	4,559	7,459	3,552	15,393	5,511	9,243	5,189	4,282
2015	2,818	5,216	3,093	11,681	4,276	7,513	3,571	3,006
2013	2,113	4,836	2,794	10,472	3,667	7,045	2,429	2,168
2011	1,412	3,720	2,499	10,086	3,512	7,541	3,208	2,644

EDA - Demographic

Ethnic Composition of Survey Respondents (2017-2019)

Race/Ethnicity of Homeless
Annual Point-in-Time Count, 2018 through 2023

We still await the release of 2023 numbers for the City of Los Angeles.

The category **Hispanic/Latino** is actually an ethnic rather than a racial category. Persons identified as Hispanic/Latino may be from any racial group.

Race/ Ethnic Group	City of Los Angeles					All of Los Angeles County*				
	2023	2022	2020	2019	2018	2023	2022	2020	2019	2018
American Indian/ Alaska Native (not Hispanic/ Latino)	---	474 (1.1%)	430 (1.0%)	530 (1.5%)	324 (1.0%)	723 (1.0%)	610 (0.9%)	686 (1.1%)	975 (1.7%)	565 (1.1%)
Asian (not Hispanic/ Latino)	---	434 (1.0%)	502 (1.2%)	299 (0.8%)	352 (1.1%)	1,212 (1.7%)	598 (0.9%)	774 (1.2%)	455 (0.8%)	545 (1.1%)
Black/ African American (not Hispanic/ Latino)	---	13,814 (32.9%)	15,622 (37.8%)	13,629 (37.7%)	12,235 (39.1%)	22,606 (31.7%)	19,523 (28.2%)	21,509 (33.8%)	18,719 (33.3%)	17,825 (35.7%)
Hispanic/ Latino	---	17,470 (41.6%)	13,424 (32.5%)	12,403 (34.3%)	10,813 (34.6%)	30,350 (42.6%)	28,940 (41.9%)	23,005 (36.1%)	20,523 (36.5%)	17,540 (35.1%)
Native Hawaiian/ Pacific Islander (not Hispanic/ Latino)	---	96 (0.2%)	76 (0.2%)	100 (0.3%)	77 (0.2%)	389 (0.5%)	142 (0.2%)	205 (0.3%)	314 (0.6%)	221 (0.4%)
White (not Hispanic/ Latino)	---	8,327 (19.8%)	10,293 (24.9%)	8,223 (22.7%)	6,978 (22.3%)	13,826 (19.4%)	13,661 (19.8%)	16,208 (25.4%)	13,871 (24.7%)	12,538 (25.1%)

* Accumulated numbers for all Service Planning Areas, not including numbers for Glendale, Long Beach and Pasadena.

Modeling Process Set Up

- Classification Models
- Logistic Regression
- KNN Neighbors
- Random Forest
- Ada Boost
- SVC
- Train Test Split
- Test size is 0.20
- Random state is 42
- Stratify is y

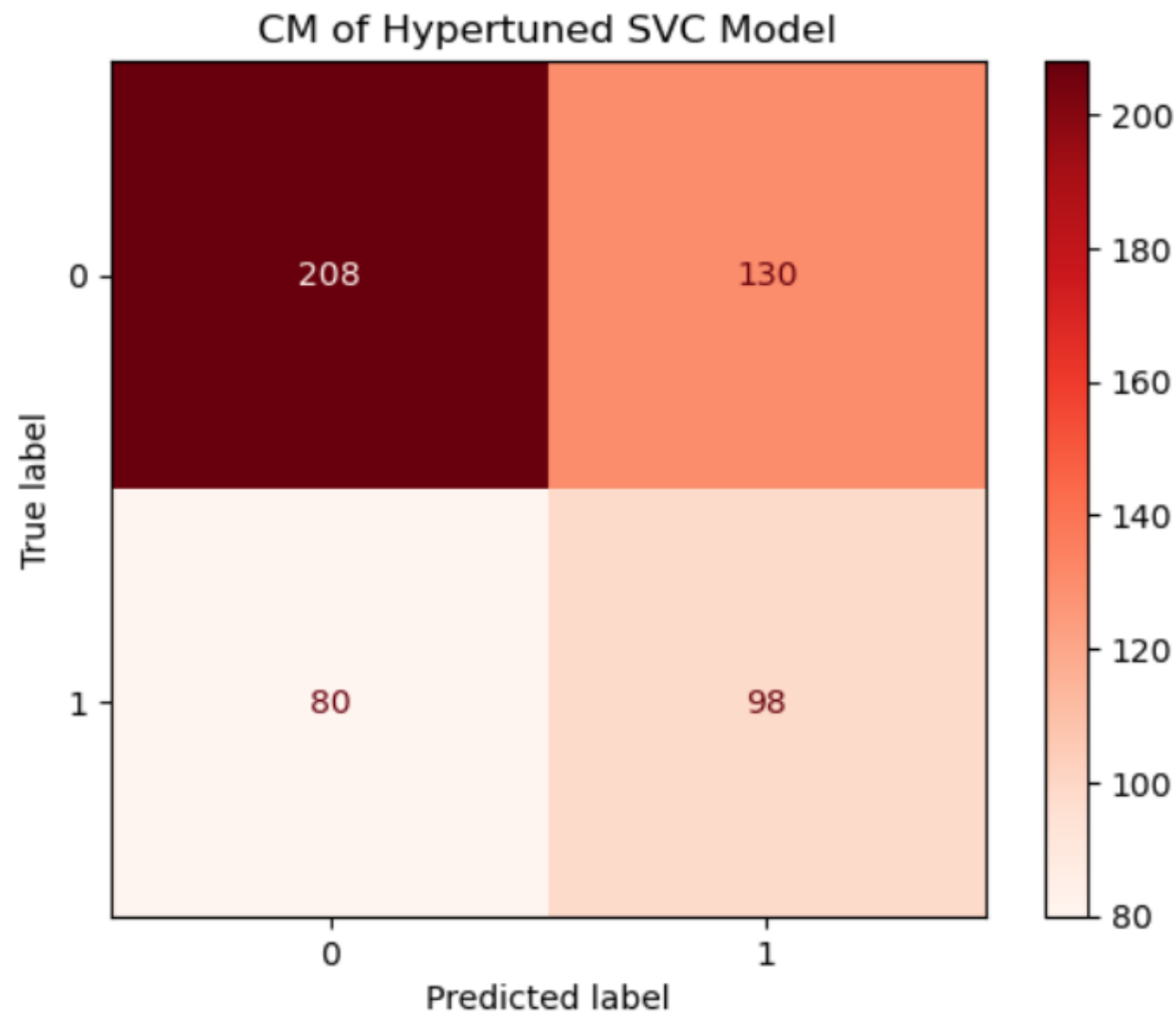| Variable	Data Type	Value Count	Description
hmlsmorethan1Yr	int64	2577	0 - No, 1 - Yes: homeless more than 1 year this time
dv_neglect	int64	2577	0 - No, 1 - Yes, 2 - R, 3 - S, 4 - D, 5 - C, 6 - N: faced neglect
dv_physical	int64	2577	0 - No, 1 - Yes, 2 - R, 3 - S, 4 - D, 5 - C, 6 - N: faced physical abuse
dv_physical_rel	int64	2577	0 - No, 1 - Yes, 2 - R, 3 - S, 4 - D, 5 - C, 6 - N: faced physical abuse by parents
dv_sexual_rel	int64	2577	0 - No, 1 - Yes, 2 - R, 3 - S, 4 - D, 5 - C, 6 - N: faced sexual abuse by parents
subsabuse	int64	2577	0 - No, 1 - Yes, 2 - M, 3 - Y, 4 - A: substance abuse problem of long duration (18+ only)
drugabuse	int64	2577	0 - No, 1 - Yes, 2 - M: Drug Abuse
SPA	int64	2577	1 - 8: Service Planning Areas

Results and Findings

balanced_accuracy: 0.5829732065687122
recall: 0.550561797752809
precision: 0.4298245614035088
f1 score: 0.4827586206896552

CM of Hypertuned SVC Model

Streamlit App

User Input Features

Example CSV input file

Upload your input CSV file

> **Drag and drop file here**
> Limit 200MB per file • CSV
>
> Browse files

Neglect

6

0   6

Physical Abuse

6

0   6

Physical Abuse By Parent

6

0   6

Sexual Abuse By Parent

6

0   6

Homeless Youth in LA Prediction App

This app predicts whether a young individual will become homeless more than 1 year.

Data obtained from the LA County Homeless Count Data Library.

User Input features

Awaiting CSV file to be uploaded. Currently using example input parameters (shown below).

	dv_neglect	dv_physical	dv_physical_rel	dv_sexual_rel	subsabuse	drugabuse	SPA
0	6	6	6	6	6	2	8

Prediction

value
No

References

- https://www.youthvoices.live/the-neglect-of-skid-row/
- https://www.theguardian.com/us-news/2023/dec/19/california-us-street-homelessness-youth-unsheltered-annual-report
- http://www.laalmanac.com/social/so14.php
- https://oclawin.org/wp-content/uploads/2015/08/FactSheet_GeneralCauses_and_challenges_2015.pdf
- https://github.com/dataprofessor/streamlit_freecodecamp/tree/main/app_8_classification_penguins