





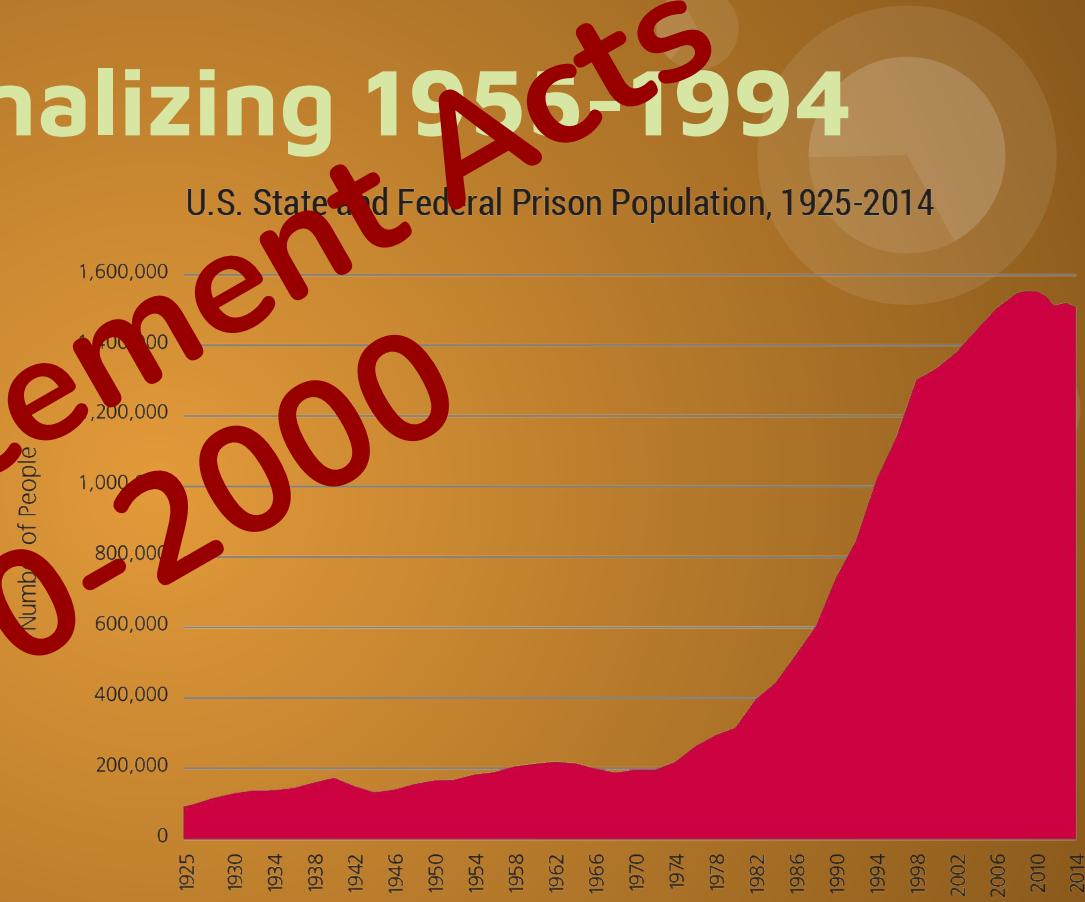
# Topic Modeling for GrassRoot Campaigns: Homelessness



Nextdoor.com content pertaining to homelessness by neighborhood in east Denver.

# (Re)-Institutionalizing 1955-1994

- 1955
- < 200,000 incarcerated
  - > 500,000 mental patients
  - 0.3% of pop. Mentally Ill  
(> 1.1 million in 2020)
- 2020
- ~2.1 million incarcerated
  - 10 beds per 100,000 ppl
  - 15% of pop. Mentally Ill  
(> 51.5 million)



Source: Bureau of Justice Statistics *Prisoners Series*.

<https://www.vera.org/downloads/publications/people-in-jail-and-prison-in-2020.pdf>



# Continuum Of Care:

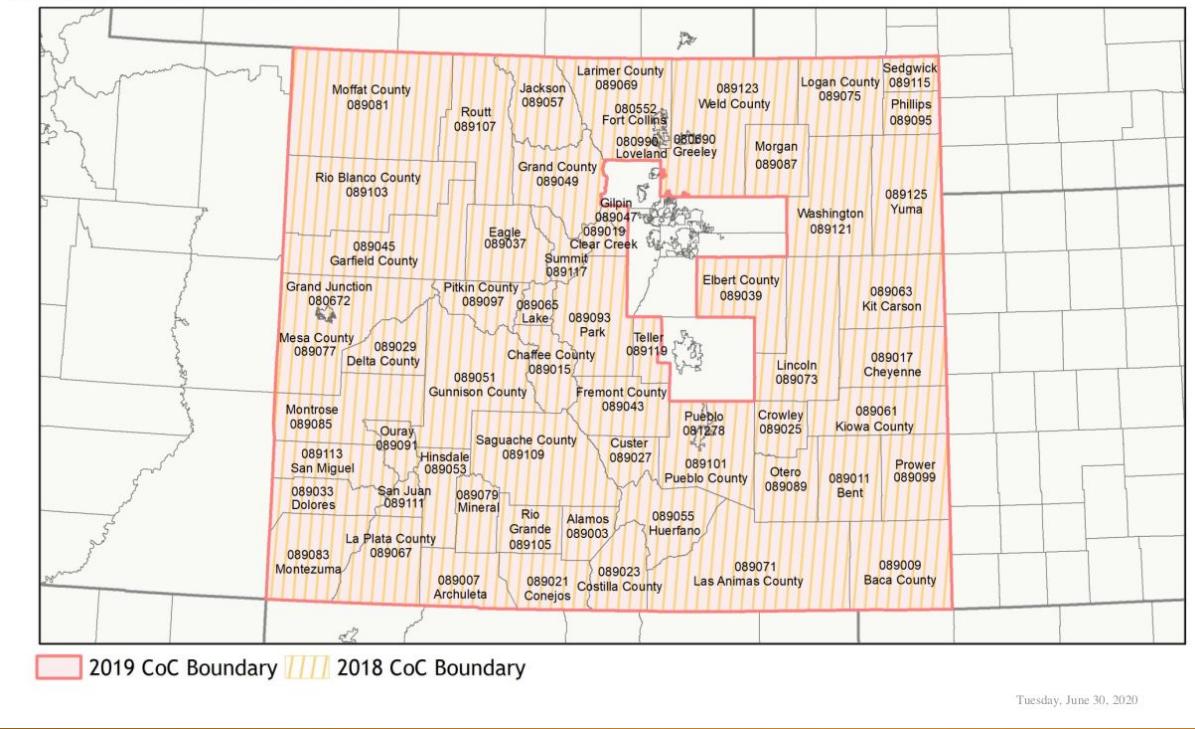
National Annual  
Homeless Population &  
Housing Inventory  
Report.

## Ways to Improve?

- Existing Partnerships
- Self Reporting
- Seasonal Count
- Senior Data

CoC Number: CO-500

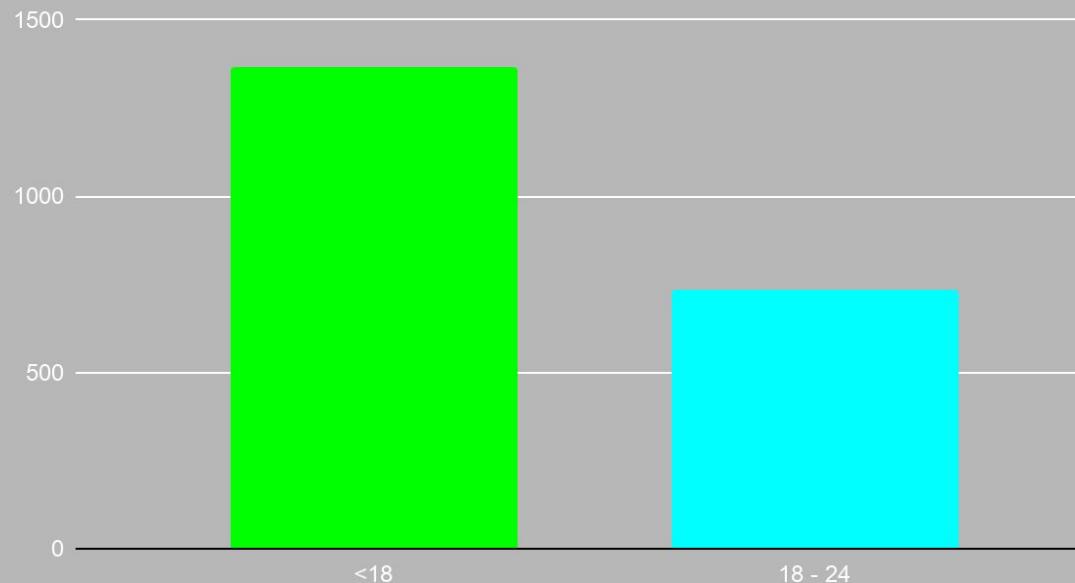
CoC Name: Colorado Balance of State CoC



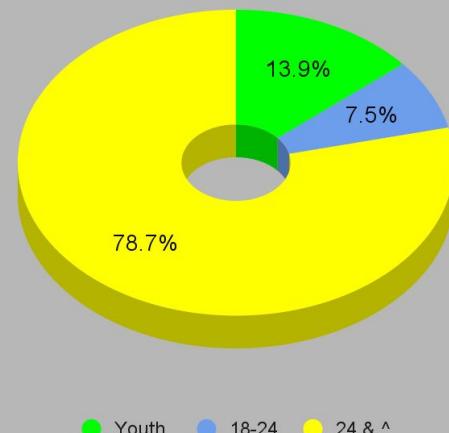
Tuesday, June 30, 2020



## 2020 Unhoused by Age



All Ages

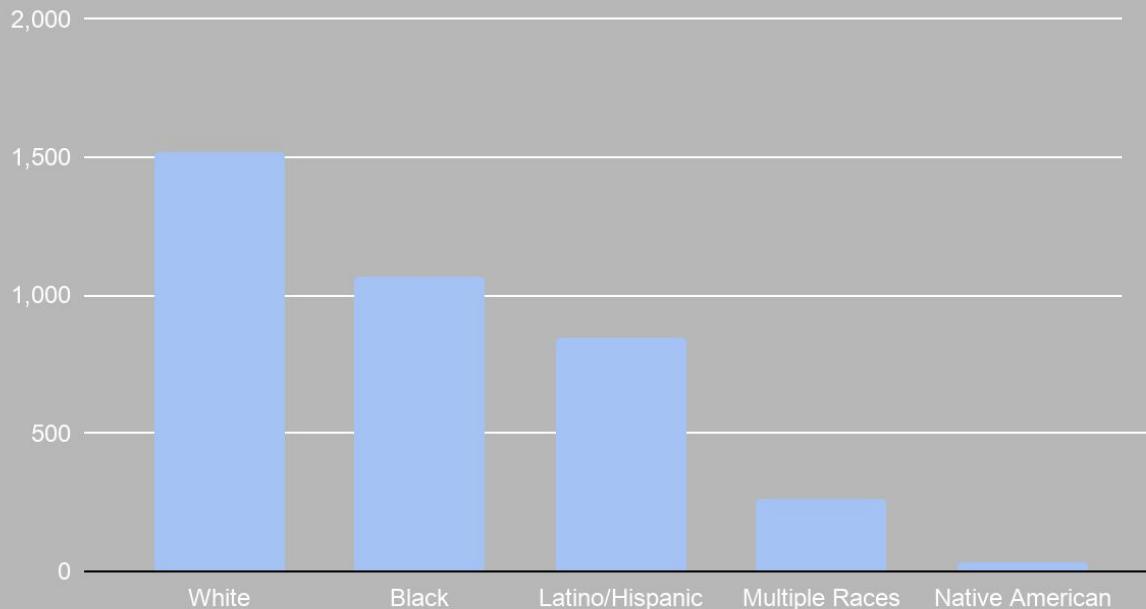




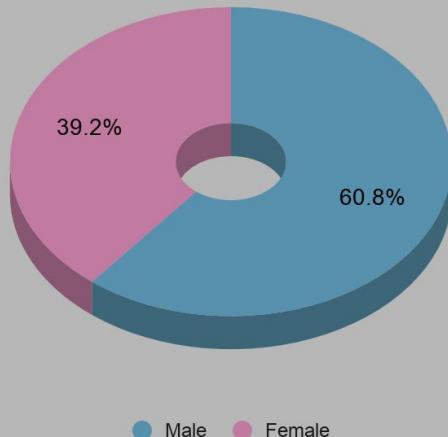
# 10,000 Unhoused in Colorado

## Who are they ?

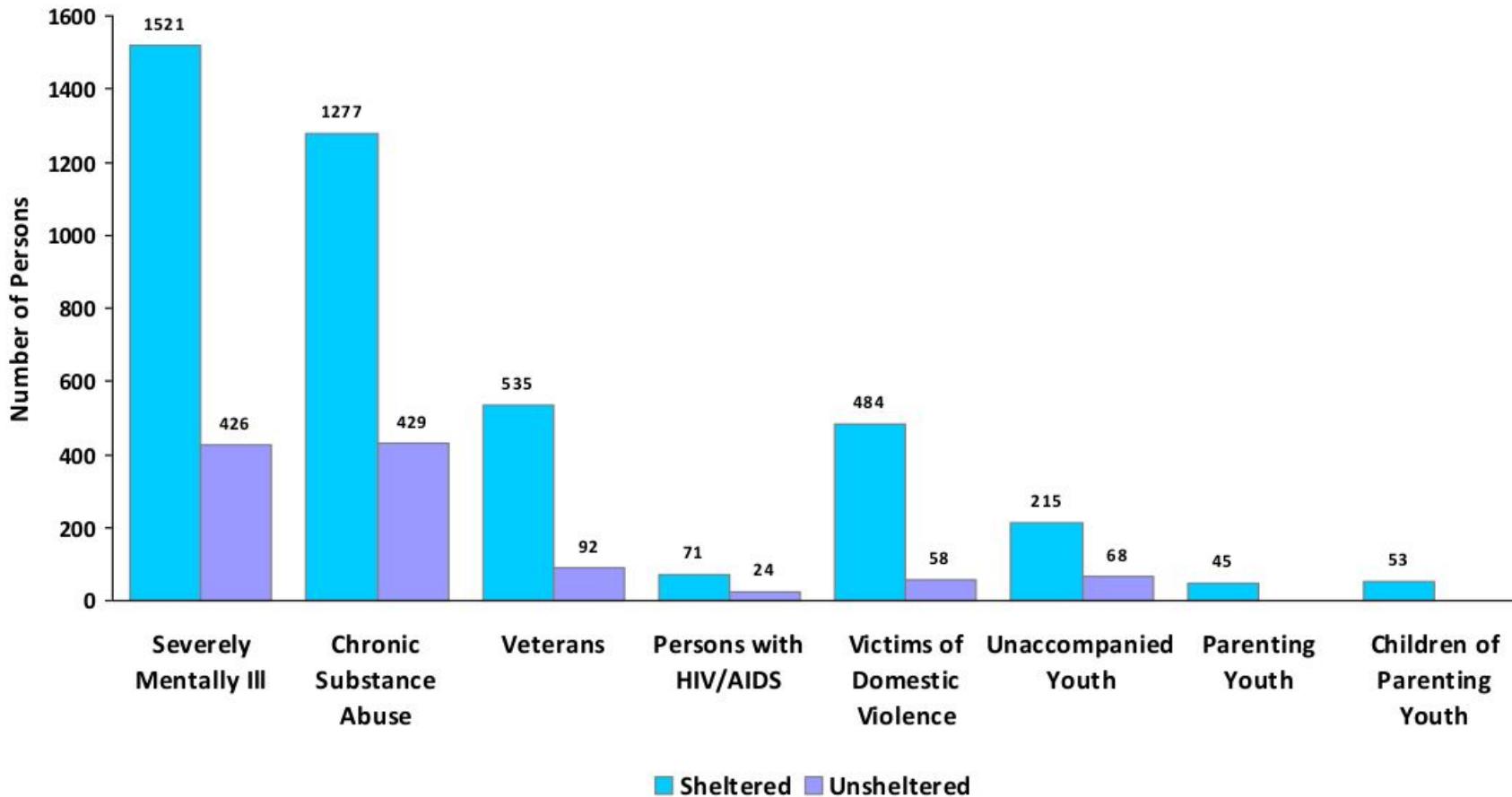
Unhoused by Race

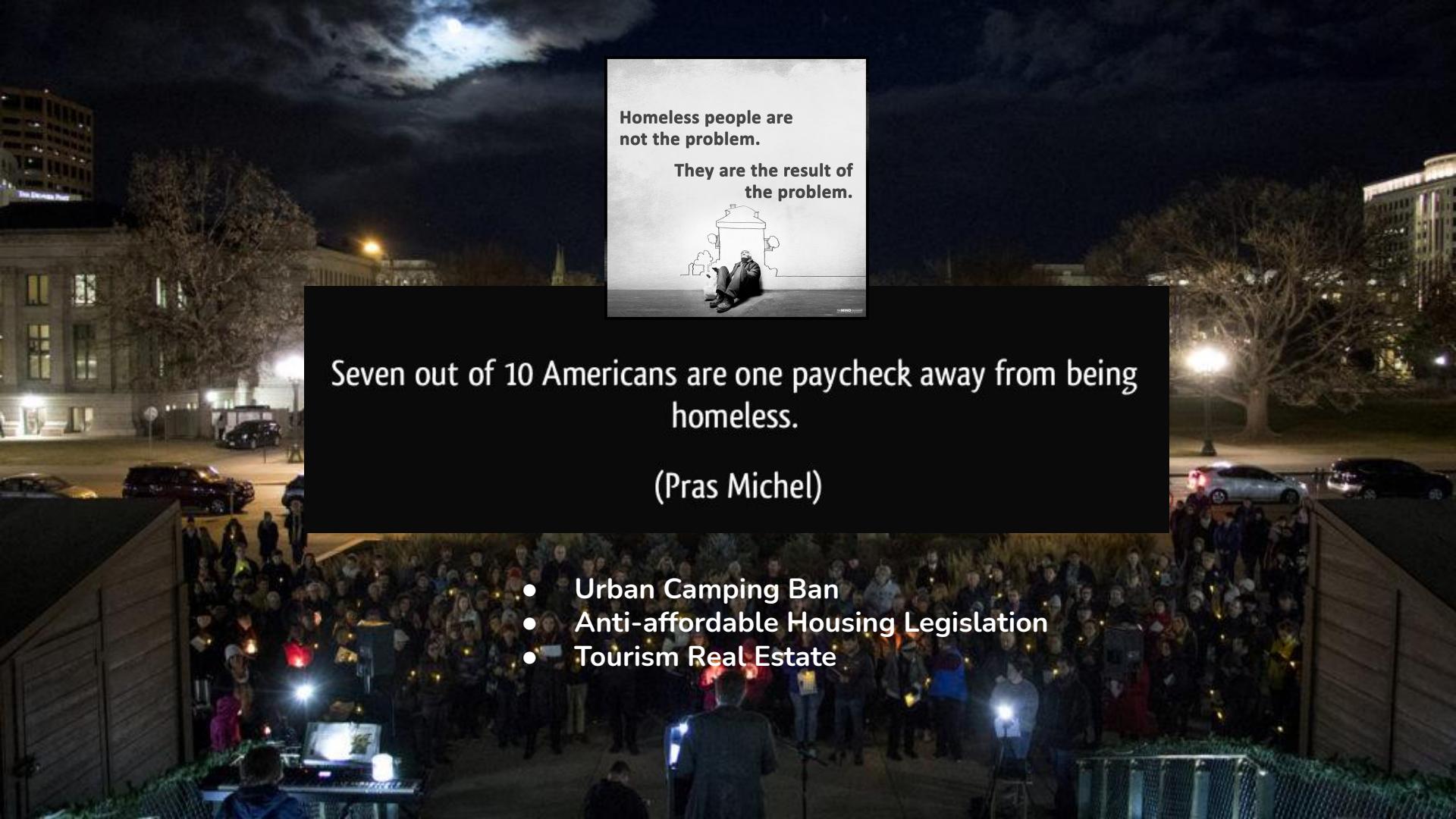


By Sex



## 2019 Point in Time Count Summarized by Sub-Population





Seven out of 10 Americans are one paycheck away from being homeless.

(Pras Michel)

- Urban Camping Ban
- Anti-affordable Housing Legislation
- Tourism Real Estate

**Homeless Tent City in North Park Hill? When were we able to voice our say on this? To whom it may concern (and it should concern everyone) in our wonderful North Park Hill neighborhood:**

I was literally floored and became sick to my stomach once I heard the news that

Was just presented to me by a friend. Honestly I can't see how this location would be considered an appropriate location for housing homeless population. This was apparently approved "under the radar" and was in the works without many knowing about it! Question: Who is going to police this and keep our children's safety top priority? Who is going to keep the area sanitized and clean of debris that we already see all over the city where these tent cities currently exist? Not to mention our property values will likely diminish as a result of this horrible decision.....

Where I live, we have a different problem. Most of the people in the street tents are mostly young men (and some women) who are hard core drug addicts who do heroin and meth all day. I have never seen a single child in any of these tents in our neighborhood, and I see them in the alley right outside my bedroom window every day. I also met with the Denver Police this week, and the police confirmed they are "service resistant" people meaning that they are offered mental health, drug treatment, and housing support every day and they refuse. They instead want the "freedom" to live in a tent, do drugs, and not have to change. People (many from churches) bring them food to eat every day. I saw a van full of nuns do this last week. The drug tents remain in neighborhoods who are terrorized by these people every day with open drug use, public deification, completely blocked sidewalks, urination, stolen bicycles, car theft, prostitution, yelling in the middle of the night, etc.. I have been the victim of this crime with 7 of these illegal drug tent "communities" within 1 block of me for over the last year. Trust me, you do not want to have to live with this! Last year someone was stabbed right on my block next to a child care center in daylight hours.

Who is going to clean up? The squatter camp has disappeared in the night. I'm glad. This wasn't the place for it. Now I'm wondering which neighbor who talked about homeless rights is going to clean up the mess left behind? With rights come responsibility.

Housing Assistance? Hi neighbors. There is an older gentleman in our neighborhood who is losing his home that he has been renting for 13 years because it has been purchased by developers. This is happening while he is going through chemotherapy. Since he is in no shape to actively find housing and has no family, I am trying to help. Who has a lead for an emergency situation like this? Thanks in advance for any help you can provide.

Extreme Need for the Homeless. Helping Hands for the Homeless Denver is hosting its monthly event on the 16th. We will be serving a full baked potato bar, fruit and more. We are collecting donations of the most needed items beforehand.

We are running extremely low on the following:  
Blankets  
Jackets  
Hygiene products (soap, shampoo, conditioner, toothpaste, toothbrushes, deodorant, Combs)  
Hand/foot warmers

If you have any items that you no longer need or use and would be willing to donate to the homeless, please let me know.

It's very hard for someone living on the streets and the cold is not ones friend. We take most items (including some furniture) and it gets distributed to those on the streets and those getting just into a home (furniture, etc). For a full list of items we can use for those on the street, I will post them in the comments.

- Selenium
  - Web Scraping JavaScript.
- Nextdoor.com
  - Community focused Social Media Outlet.
- Spacy
  - English Word Embeddings
- Latent Dirichlet Allocations
  - Topic Modeling
- VADAR
  - Sentiment Content Analysis
- Text2Emotions
  - Emotional Content Analysis

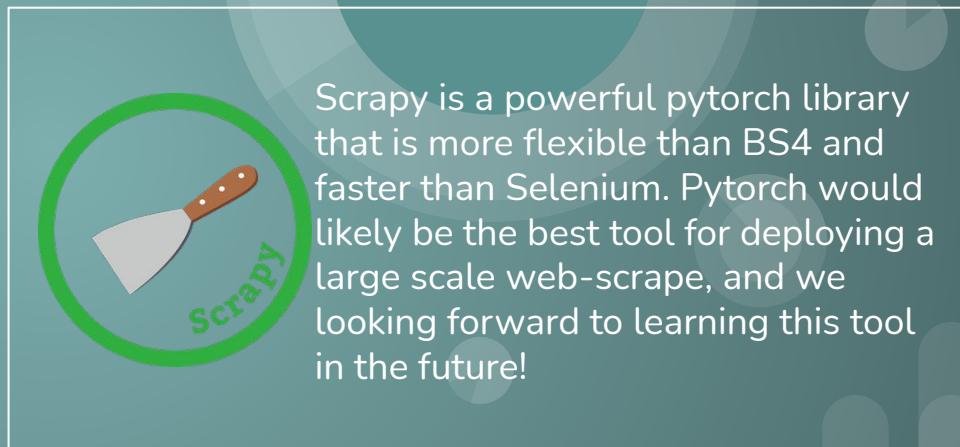
# Machine Learning Tools to Change the Narrative:

Can NLP Help Our Communities Heal?

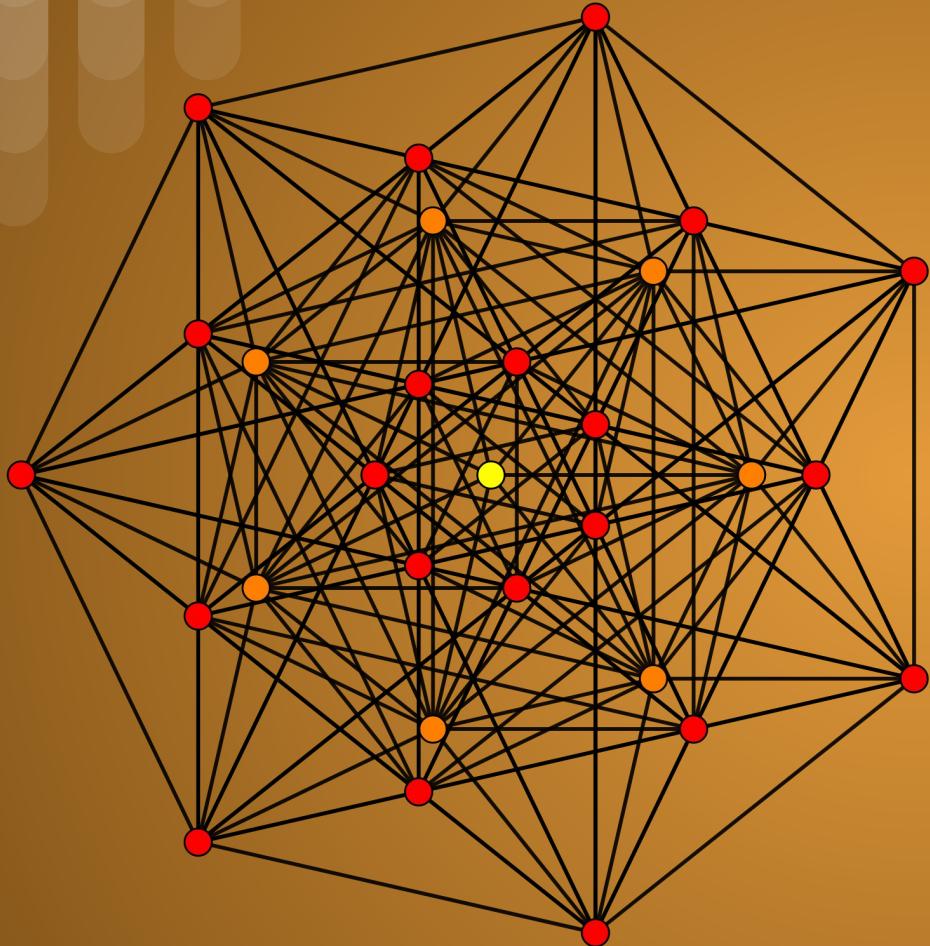


# Web Scraping with Selenium.

- JavaScript Interactive
- Slow
- Semi- Supervised Scrape



Scrapy is a powerful pytorch library that is more flexible than BS4 and faster than Selenium. Pytorch would likely be the best tool for deploying a large scale web-scrape, and we looking forward to learning this tool in the future!



Latent  
Dirichlet  
Allocation

Documents

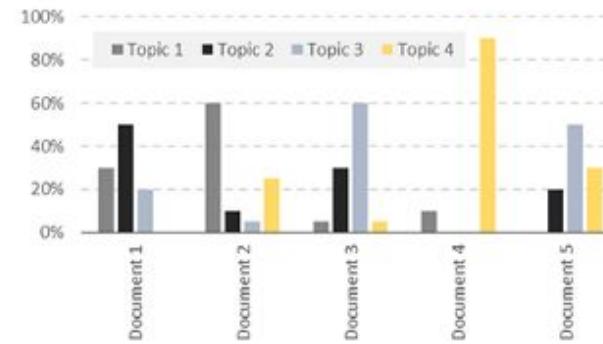


LDA

Creation of topics

	weight	words
Topic 1	3%	flower
	2%	rose
	1%	plant
...		
Topic 2	2%	company
	1%	wage
	1%	employee

Topics allocation to documents



# Tuning Method 1:

Alpha ~ Per Doc Topics  
Beta ~ Per Word Topics

Since our data is social media content, and our posts are relatively short, starting with a low Alpha & Beta parameter proved to be good intuition

# Tuning Method 2:

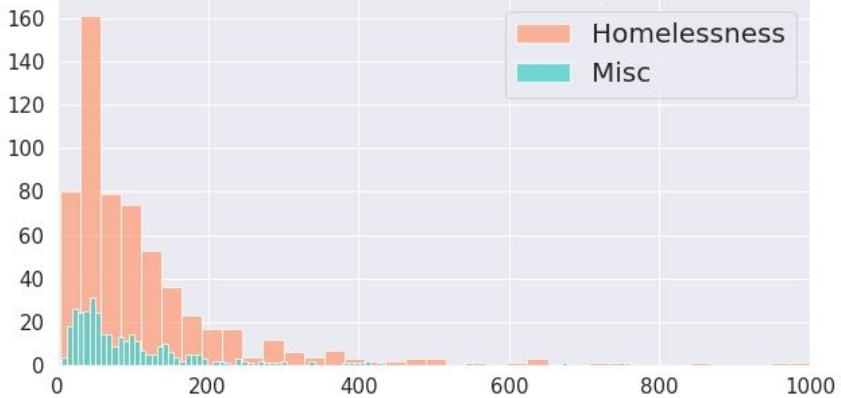


K\_topics ~ How many topics are their  
Chunksize ~ n documents to train on each pass

Unlike other machine learning models,  
the batch size for LDA is much more than  
a matter of time and computational  
resources.

# Word Count Distributions

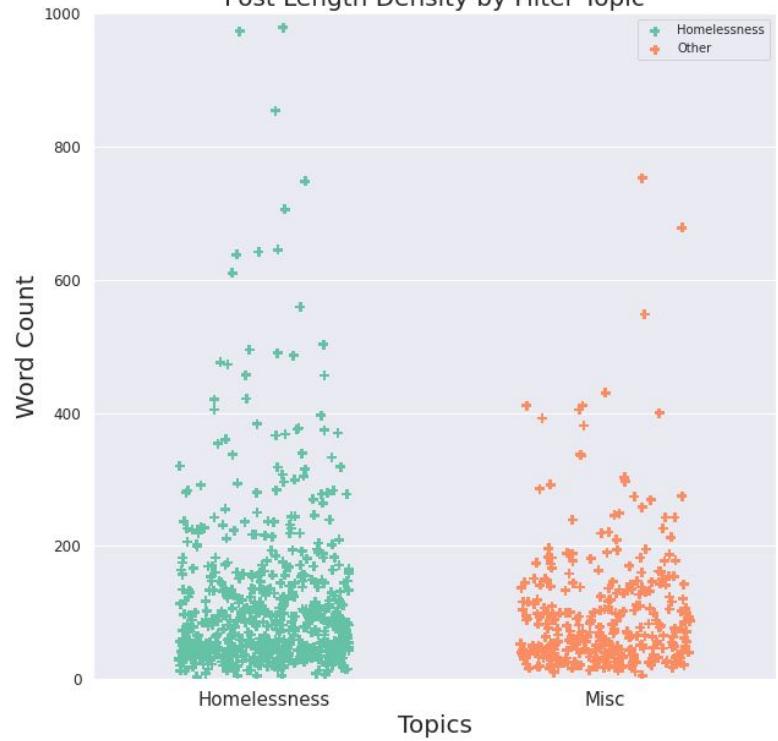
Distribution of Post Length



Log Distribution of Post Length



Post Length Density by Filter Topic



## Sentence Topic Coloring for Documents: 7 to 13

**Doc 0:**happen second strange walk park bring afternoon alley bit lose past ready recycling ...

**Doc 1:**emergency go happen help homeneighbor try housing family situation thank year neighborhood purchase ...

**Doc 2:**th away hope know morning run say catch bike door garage act hair person ...

**Doc 3:**look thank chip donate event food leave lot open place recommendation soda water weekend ...

**Doc 4:**camp love city claim homeless reason refuse lot accomplish actually anti base citation consideration ...

**Doc 5:**let away help know neighbor week couple great homeless safe place beloved bench cold ...

**Doc 6:**camp go squatter come neighbor homeless night leave place clean disappear glad mess responsibility ...

# Let's Visualize!

- Simplexes
- Filters Topics
- Final Topics

location	Word Count	Angry	Surprise	Sad	Fear	Happy	Polarity
92nd and Grove	37.000000	0.000000	0.110000	0.220000	0.330000	0.330000	0.876000
Allendale	21.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Arvada Columbine	164.000000	0.150000	0.220000	0.220000	0.190000	0.190000	0.907600
Aurora	59.666667	0.026667	0.125000	0.091667	0.495000	0.095000	0.686733
Aurora Arts District	129.200000	0.036000	0.398000	0.122000	0.388000	0.104000	0.401360
Aurora Cultural Arts District	145.166667	0.045833	0.185833	0.158333	0.415833	0.108333	0.856608
Barnum	49.000000	0.000000	0.400000	0.100000	0.400000	0.100000	0.908000
Barnum West	128.000000	0.050000	0.210000	0.160000	0.370000	0.210000	-0.818300
Beeler Park	145.000000	0.000000	0.166667	0.113333	0.416667	0.306667	0.955600
Berkeley-Regis	383.000000	0.160000	0.140000	0.180000	0.390000	0.140000	-0.724200
Bluff Lake Aurora	74.000000	0.000000	0.270000	0.090000	0.360000	0.270000	0.888100
Central Park	69.685714	0.052571	0.184571	0.200857	0.299714	0.03714	0.652954
Cheesman Park	31.000000	0.000000	0.330000	0.170000	0.500000	0.000000	0.750600
Cherry Creek North	126.062500	0.064375	0.246250	0.245000	0.318750	0.123750	0.343781
City Park	91.181818	0.036363	0.131818	0.218182	0.526354	0.090009	0.597699
Congress Park	116.660714	0.057857	0.220536	0.244821	0.311250	0.147321	0.618389
Crestmoor/Mayfair Park	167.428571	0.032381	0.271905	0.160000	0.313333	0.171905	0.838557
East Colfax	111.100000	0.029667	0.140333	0.225333	0.302333	0.135667	0.529513
Fort Logan	65.000000	0.000000	0.000000	1.000000	0.000000	0.000000	-0.153100
Gov	109.747126	0.039000	0.232414	0.161494	0.418500	0.113563	0.313345
Hale	186.782609	0.034348	0.173478	0.193913	0.438696	0.158696	0.422170
Highlands	54.000000	0.000000	0.140000	0.290000	0.290000	0.290000	0.962600
Hilltop	247.586207	0.021034	0.228276	0.231379	0.338276	0.144483	0.403083
Historic Montclair	82.250000	0.008750	0.268750	0.172500	0.416250	0.132500	0.396637
Hoffman Heights	71.307692	0.051538	0.173846	0.142308	0.445385	0.185385	0.463431
Holly Hills	42.000000	0.000000	0.500000	0.000000	0.500000	0.000000	0.668800
Hudson Street	28.000000	0.000000	0.500000	0.000000	0.500000	0.000000	0.000000
Locust St	70.750000	0.000000	0.152500	0.302500	0.135000	0.160000	-0.048325
Lowry	80.860465	0.045581	0.207674	0.161628	0.393721	0.143023	0.549219
Mayfair	161.166667	0.064333	0.156333	0.188667	0.328000	0.229667	0.438057
Memorial Park	126.000000	0.230000	0.080000	0.150000	0.080000	0.460000	0.971000
Montbello	84.333333	0.033333	0.227222	0.210000	0.302222	0.115556	0.535667
Morris Heights	85.000000	0.127500	0.280000	0.247500	0.282500	0.062500	0.327400
North Aurora/Fitzsimons	98.312500	0.043125	0.226875	0.200000	0.440625	0.088125	0.485931
North Park Hill	83.076923	0.019231	0.208462	0.144615	0.422300	0.127692	0.576538
Northeast Park Hill	144.678571	0.031071	0.235000	0.213214	0.338929	0.145000	0.609157
Old Thornton	305.000000	0.060000	0.350000	0.120000	0.320000	0.150000	0.992400
Parkfield	24.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.458800
RiNo Art District	20.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
Skyland	169.000000	0.090000	0.270000	0.090000	0.230000	0.320000	0.904700
Sloans/HIGHLANDS	106.666667	0.000000	0.263333	0.216667	0.303333	0.216667	0.429233
South Park Hill	113.178571	0.067143	0.210714	0.217143	0.345893	0.121429	0.543504
Thornton	151.000000	0.045000	0.325000	0.085000	0.325000	0.220000	0.953350
Vasquez Blvd	115.000000	0.016667	0.230000	0.253333	0.306667	0.196667	0.965233
Wandering View	2705.000000	0.030000	0.280000	0.150000	0.400000	0.150000	0.999300
Wash Park West	38.000000	0.000000	0.140000	0.430000	0.430000	0.000000	0.734500
Westwood	13.000000	0.000000	0.500000	0.250000	0.000000	0.250000	0.751200
Westy Hills	88.000000	0.000000	0.600000	0.000000	0.300000	0.100000	0.727400
Whittier	50.000000	0.080000	0.170000	0.420000	0.330000	0.000000	-0.190100
Windsor Gardens East	59.000000	0.000000	0.400000	0.200000	0.200000	0.200000	0.911100
Wyo Park	226.000000	0.000000	0.140000	0.140000	0.670000	0.050000	-0.655300

Compound Coherence : 47%  
Perplexity : -7.56846

author	Word Count	Angry	Surprise	Sad	Fear	Happy	Polarity
Abby Leeper	54.000000	0.000000	0.250000	0.250000	0.250000	0.250000	0.937900
Abeth Parisian	78.000000	0.000000	0.000000	0.670000	0.000000	0.330000	0.784000
Adam Oderberg	51.000000	0.000000	0.200000	0.000000	0.600000	0.200000	0.676700
Alex Williams	87.000000	0.100000	0.100000	0.300000	0.400000	0.100000	0.168000
Alexandra Kyles	160.000000	0.050000	0.140000	0.380000	0.330000	0.100000	0.982100
Alexis Mera	13.000000	0.000000	0.250000	0.250000	0.500000	0.000000	0.824100
Ali Rodriguez	138.000000	0.090000	0.450000	0.090000	0.360000	0.000000	0.484000
Allison Beasley	748.000000	0.010000	0.170000	0.270000	0.430000	0.120000	0.998300
Allison O'keeffe	34.000000	0.000000	0.500000	0.000000	0.250000	0.250000	0.815600
Allison Rawls	92.000000	0.000000	0.120000	0.380000	0.380000	0.120000	0.301800
Amanda Biffle	120.000000	0.150000	0.150000	0.310000	0.380000	0.000000	0.955600
Amanda Toney	172.000000	0.000000	0.260000	0.370000	0.370000	0.000000	-0.619700
Amber Vail	35.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.666200
America Bateman	65.000000	0.000000	0.250000	0.000000	0.000000	0.750000	0.717000
America PÃ©rez	36.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000
Amy Bachmeier	157.000000	0.000000	0.260000	0.050000	0.580000	0.110000	0.955700
Amy Pomante	188.000000	0.120000	0.240000	0.080000	0.320000	0.240000	0.990700
Andrew Sweet	94.000000	0.000000	0.090000	0.640000	0.270000	0.000000	0.337600
Angel Rodriguez	90.000000	0.090000	0.270000	0.270000	0.270000	0.090000	0.314500
Ann Burke	99.000000	0.290000	0.000000	0.290000	0.430000	0.000000	0.848800
Ann Smith	20.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000
Anna McCaffery	41.000000	0.000000	0.330000	0.330000	0.170000	0.170000	-0.361400
Anne Bouise Ward	284.000000	0.090000	0.260000	0.090000	0.140000	0.430000	0.992300
Arlene Burciaga	50.000000	0.080000	0.170000	0.420000	0.330000	0.000000	-0.190100
Audrey Burquest	59.000000	0.140000	0.140000	0.290000	0.140000	0.290000	0.796900



# Conclusions & Next Steps

1. LDA has promising applications in Targeted Campaign's
2. The CoC Annual Homeless Count Desperately need's improved.
3. Developed LDA model comparison methodology.
4. Constructing a full API for Nextdoor for use in topic analysis would be a good use of time and resources for the Denver Area.
1. Semantic Word Embedding with lda2vec
2. Expand text2emotion library
3. Explore alternative data sources
4. Refining a labeled dataset for use in training Binary classifiers.
5. Gib's Sampling Method