? Homelessness in the US

By Megan Ma

Outline?

Intro

Goal: Looking at different housing cost and housing affordability measures to determine which ones are most correlated to rates of homelessness

Literature Review

Zillow Research Group:

[Inflection Points in Community Level Homelessness Rates](#)

[Rent Affordability Presentation](#)

[Rent Affordability Blog Post](#)

Book about homelessness and rent/housing

About Homelessness Rate Data

HUD Exchange ([Source](#))

- Point-in-Time Estimates by State (2007-2021)

Census Population Estimates ([Source](#))

- Annual Estimates of the Resident population for the United States

$$\text{Proportion Overall Homeless* =} \frac{\text{PIT Estimates}}{\text{Census State Population Estimate}}$$

*For each state in 2020

might be good to use homeless per 10,000 residents. You can just multiply your propotion by 10,000.

About Housing Data

Zillow Home Value Index (ZHVI): Source
A smoothed, seasonally adjusted measure of the typical home value across a given region and housing type.

- Bottom-tier ZHVI: reflects the typical value for homes in the 5th to 35th percentile range.
- Mid-tier ZHVI: reflects the typical value for homes in the 35th to 65th percentile range.
- Top-tier ZHVI: reflects the typical value for homes in the 65th to 95th percentile range.
- Condo/coops ZHVI: typical value for all condo/coops in a given region
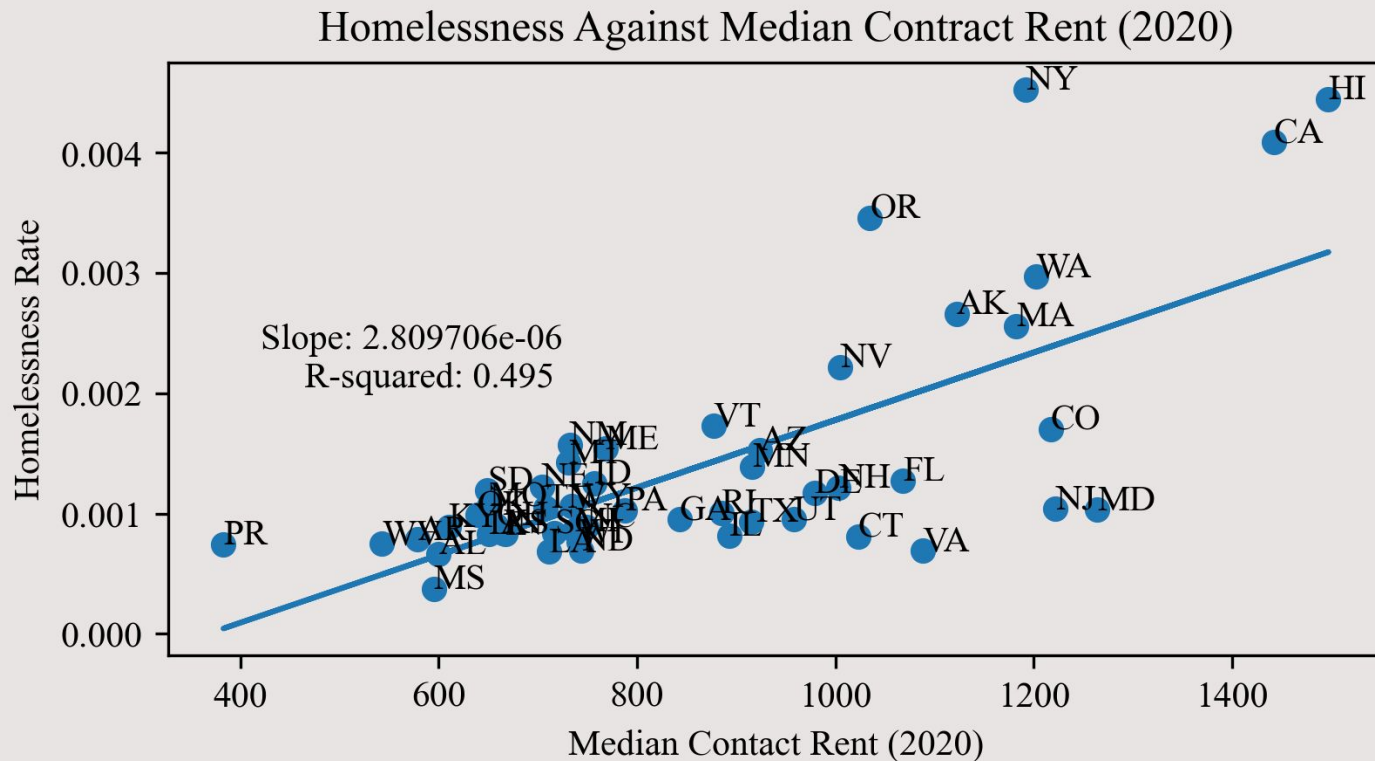
    Additional Information: overview of ZHVI and a deep-dive into its methodology.

American Communities Survey 2020 (ACS)
- Median Contract Rent

is it important to say "Contract" in every slide. if not, maybe just say median rent.

Correlation with Homelessness Rate

Homelessness Against Median Contract Rent (2020)

Correlation with Homelessness Rate

Median Home value plot

Correlation with Homelessness Rate

	Proportion State Overall Homeless,2020
condo_coop_2020	0.821
top_tier_2020	0.810
mid_tier_2020	0.775
crowded_pct	0.759
Median Contract Rent (2020)	0.710
bottom_tier_2020	0.690
renters_pct	0.646
senate_dem_pct	0.561
house_dem_pct	0.557
med_hh_income	0.460
col_grad_pct	0.327
unemploymed_pct	0.110
pop_density	0.045
gini	0.026
households_with_children_pct	-0.040
poverty_rat_under_50_pct	-0.212
governer_party_2020	-0.297

try to use names like this for all variables →

- High correlation between homelessness rate and housing measures compared to other variables
- Condo/coop and top tier housing have highest correlation with homelessness

Reasonable to expect higher correlation between homelessness and bottom tier/rent prices, so what is going on here?

Possible Explanation: Gentrification

An increase in wealthier, usually white, people arriving in an existing urban neighborhoods causes low income residents to be priced out of their housing, which explains the correlation between high cost housing and homelessness.

Indicators of Gentrification
- Increase in rent/home prices
- Increase in median income rates

Insert time series visualizations
Insert literature about gentrification

Multivariable Linear Model

Fit a multiple regression model to get a better handle on what may be going on with those measures with unexpected correlations (like high-tier pricing)

(Insert what we expect to see from the model)

About Additional Data

American Communities Survey 2020 (ACS)
- percentage of renters
- percentage of college grads
- percentage of people with incomes at half or less than half of the poverty threshold
- unemployment percentage
- median household income
- Gini coefficient of income inequality
- population density
- percentage of crowded households (units occupied by 1.01 persons or more per room)
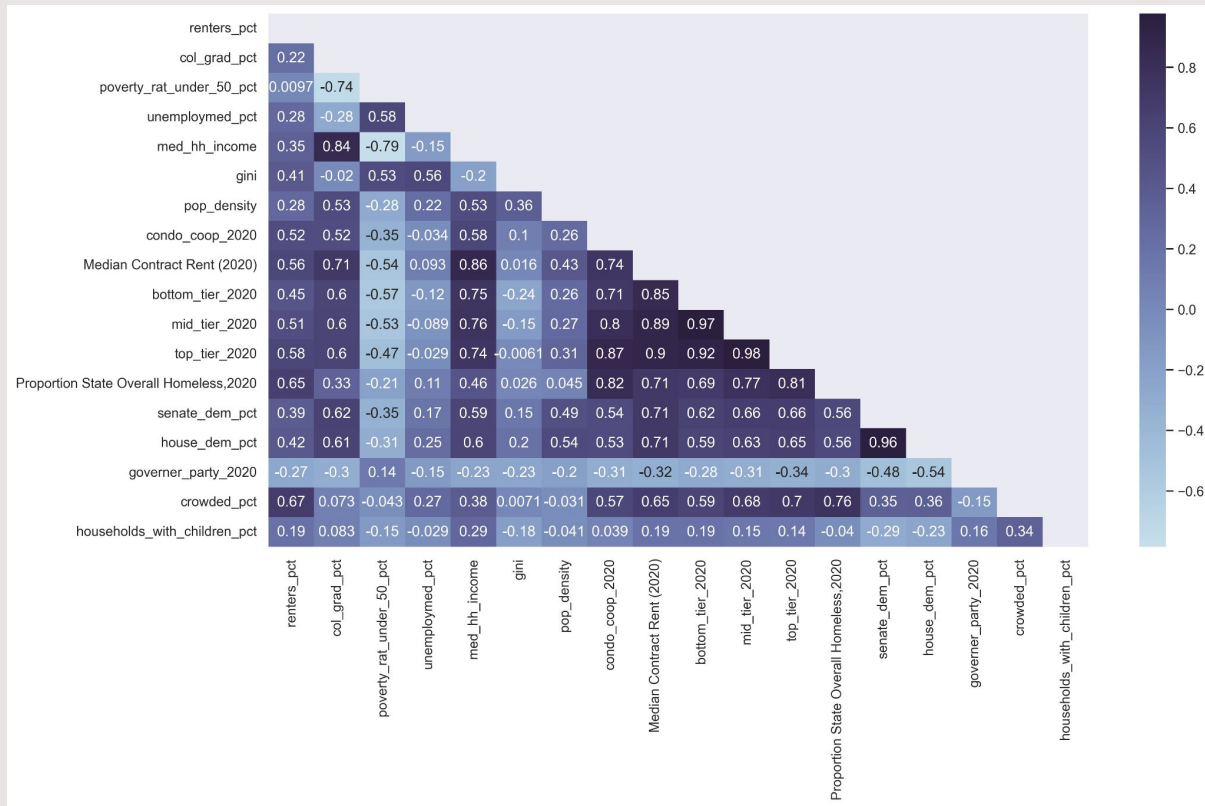- percentage of households with children)

Governor political party 2020, 0 for dem 1 for republican (governer_party_2020)
Percent of 2020 senate members from each state that are democratic (senate_dem_pct)
Percent of 2020 house of representative members from each state that are democratic (house_dem_pct)

Checking for Multicollinearity

Model Results

Condo/Coop vs natural log of Homelessness Rate
("a one percentage point increase in people with college edu. leads to Y percent change in the homelessness rate")

$\beta \times 100\%$

```
                          OLS Regression Results
==============================================================================
Dep. Variable:     ln Proportion State Overall Homeless,2020   R-squared:                   0.576
Model:                                              OLS        Adj. R-squared:              0.567
Method:                                   Least Squares        F-statistic:                 63.88
Date:                                 Fri, 12 Aug 2022         Prob (F-statistic):       2.61e-10
Time:                                          14:20:50        Log-Likelihood:             23.400
No. Observations:                                    49        AIC:                        -42.80
Df Residuals:                                        47        BIC:                        -39.02
Df Model:                                             1
Covariance Type:                              nonrobust
==============================================================================
                     coef      std err          t       P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const             -3.3219        0.054      -60.964      0.000      -3.432      -3.212
condo_coop_2020  1.746e-06     2.18e-07        7.993      0.000    1.31e-06    2.19e-06
==============================================================================
Omnibus:                        4.130        Durbin-Watson:                 2.288
Prob(Omnibus):                  0.127        Jarque-Bera (JB):              3.009
Skew:                           0.488        Prob(JB):                      0.222
Kurtosis:                       3.721        Cond. No.                    6.21e+05
==============================================================================
```

Model Results

With additional variables

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     ln Proportion State Overall Homeless,2020   R-squared:                0.798
Model:                                               OLS   Adj. R-squared:              0.745
Method:                                    Least Squares   F-statistic:                 15.04
Date:                                   Fri, 12 Aug 2022   Prob (F-statistic):       2.32e-10
Time:                                           14:33:47   Log-Likelihood:             41.600
No. Observations:                                     49   AIC:                        -61.20
Df Residuals:                                         38   BIC:                        -40.39
Df Model:                                             10
Covariance Type:                               nonrobust
==============================================================================
                          coef     std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                  -1.5939       0.753     -2.117      0.041      -3.118      -0.069
condo_coop_2020       1.135e-06    2.67e-07      4.251      0.000    5.94e-07    1.67e-06
renters_pct             0.0168       0.007      2.377      0.023       0.002       0.031
col_grad_pct            0.0156       0.010      1.620      0.113      -0.004       0.035
poverty_rat_under_50_pct -0.0567     0.042     -1.351      0.185      -0.142       0.028
unemploymed_pct         0.0611       0.032      1.894      0.066      -0.004       0.126
med_hh_income         -1.27e-05    5.81e-06     -2.187      0.035    -2.45e-05   -9.45e-07
gini                   -4.3199       1.964     -2.200      0.034      -8.295      -0.344
pop_density            -0.0002       0.000     -1.400      0.170      -0.000     6.87e-05
house_dem_pct           0.2824       0.148      1.907      0.064      -0.017       0.582
crowded_pct             3.1524       2.306      1.367      0.180      -1.515       7.820
==============================================================================
Omnibus:                   3.022   Durbin-Watson:               2.291
Prob(Omnibus):             0.221   Jarque-Bera (JB):            2.124
Skew:                     -0.485   Prob(JB):                    0.346
Kurtosis:                  3.318   Cond. No.                 3.64e+07
==============================================================================
```

Caveats

Correlation vs causation

Mention which variables you chose to exclude and why

Other Descriptive Statistics

More visualizations