# Homelessness in California

Springboard

Data Science Cohort 2024

Capstone Project

Ofri Oren

# outline

- Abstract
- Background
- Data
- Exploratory Data Analysis
- Processing and Modelling
- Forecasting and Results
- Bibliography and Datasets

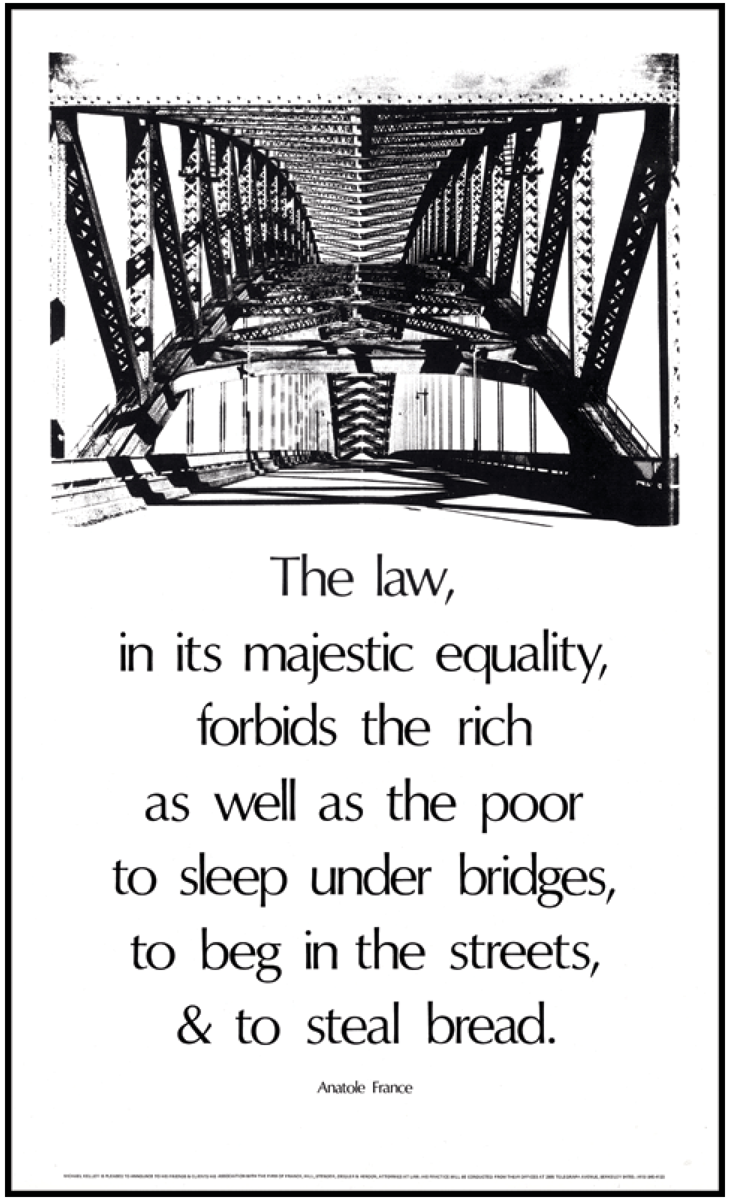
# abstract

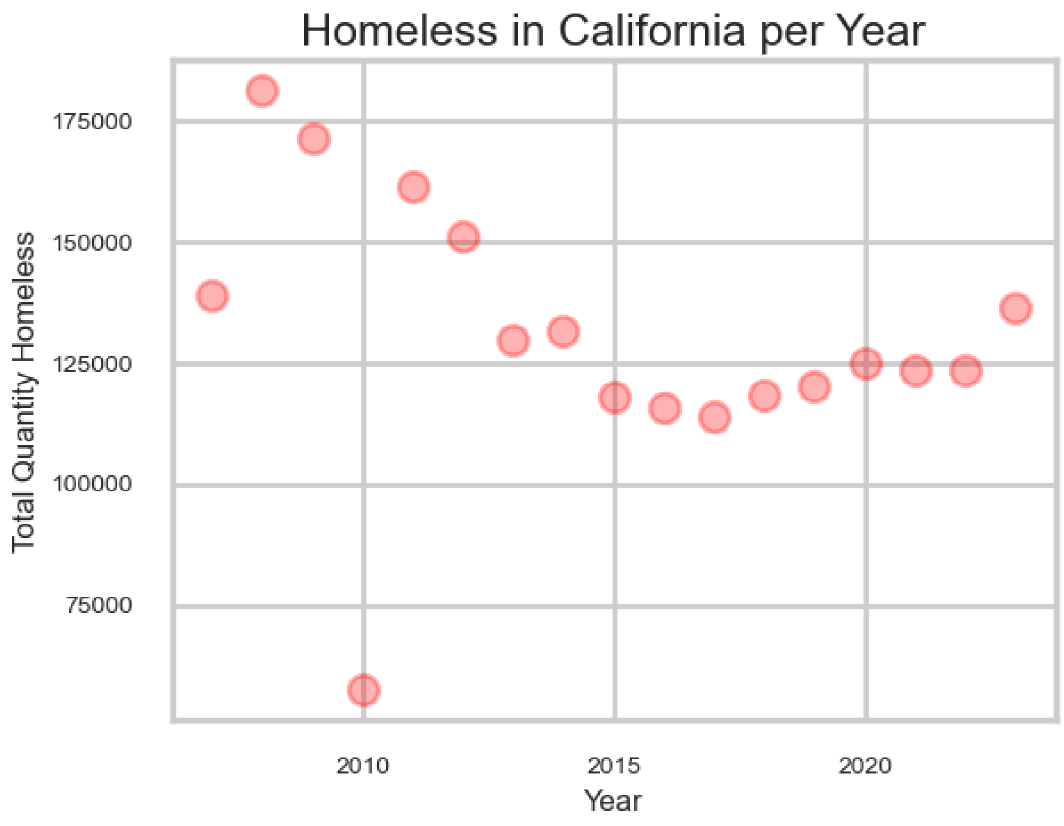
WHAT CAUSES HOMELESSNESS?

CAN WE PREDICT IT AND HOW?

# background

Homelessness is not only a humanistic issue reflecting inequity in distribution of wealth but also a legal crime.





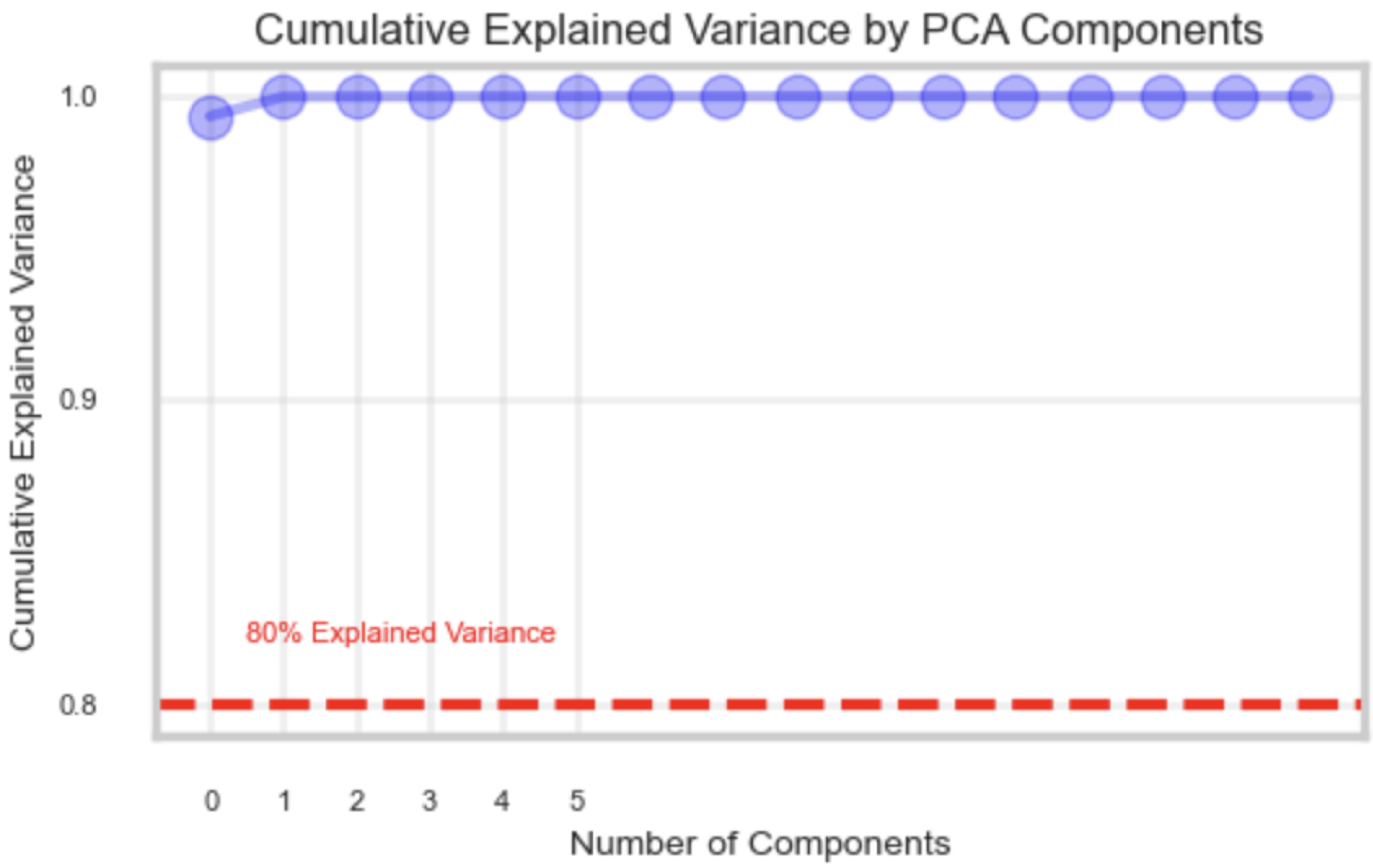

# data

What influences homelessness in this model?

Who measures it?

# exploratory data analysis

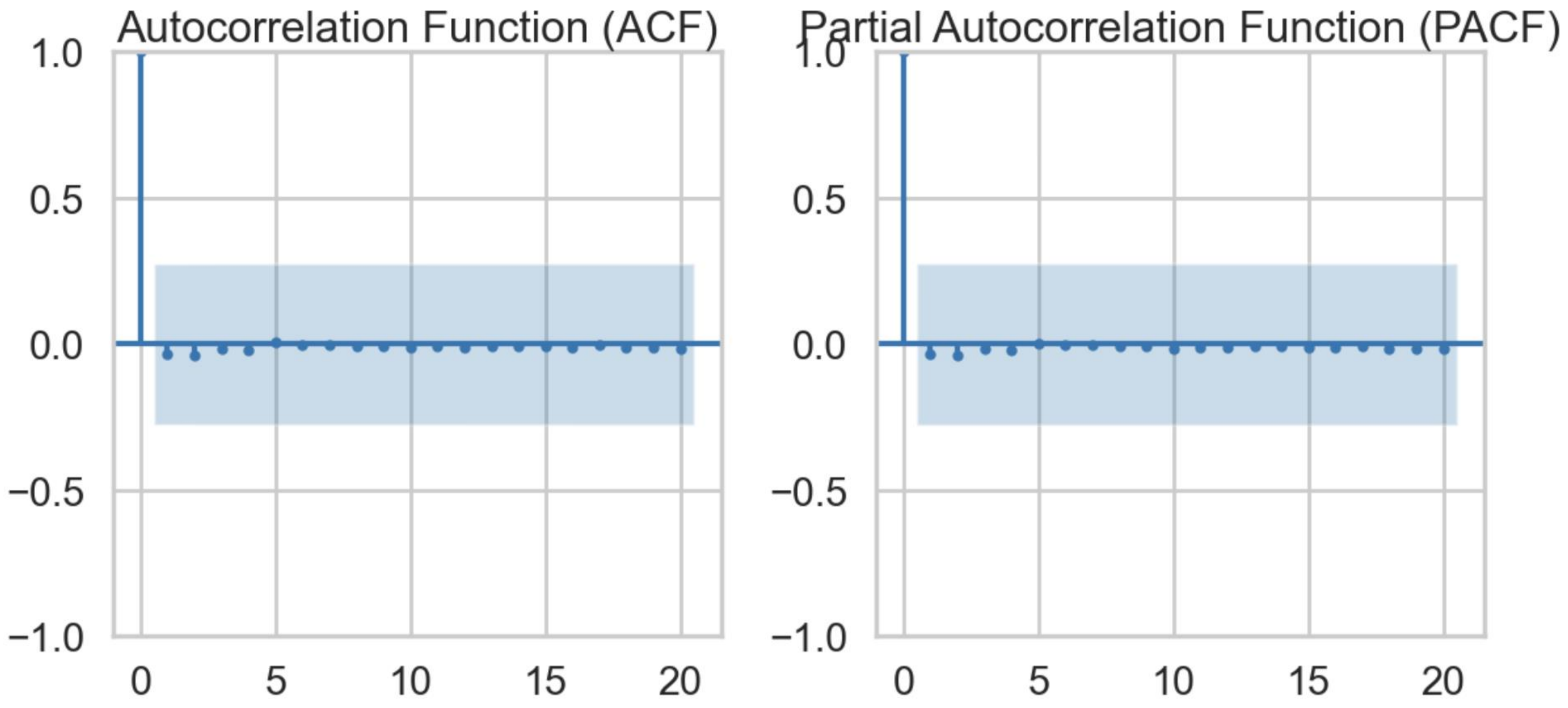
processing and modelling




Number of components explaining at least 80% variance: 1

# processing and modelling

ADF Statistic: -7.168337364030032
p-value: 2.84742280796459e-10



# processing and modelling

| Model | Metric 1 | Metric 2 |
|---|---|---|
| **Linear Regression (without PCA)** | MSE : 8809317.564591292 | $R^2$: -1.090373675217983 |
| **ARIMA 1 (one lag)** | MSE Scores: [0.4920852804015535, 0.16591506808298923, 0.10058916388114512, 0.585171870603099, 0.04961969883671817] | Average MSE: 0.278676216361101 |
| **ARIMA 2 (three lags)** | ARIMA2 MSE Scores: [1.0379242869687637, 0.21626883455424678, 0.11406775642131431, 0.5857232559016025, 0.05372395624812124] | Average MSE: 0.4015416180188097 |
| **VaR** | VAR MSE Scores: [354806659460.1925, 3258731.234438983, 235845996.07972065, 11479199.778941542] | Average MSE: 88764310846.8214 |
| | | |
| **ARIMA 1 (one lag)** | AIC: 291.742 | BIC: 297.538 |
| **ARIMA 2 (three lags)** | AIC: 291.742 | BIC: 297.538 |
| **VaR** | AIC: -144.141 | BIC: -132.773 |

# processing and modelling

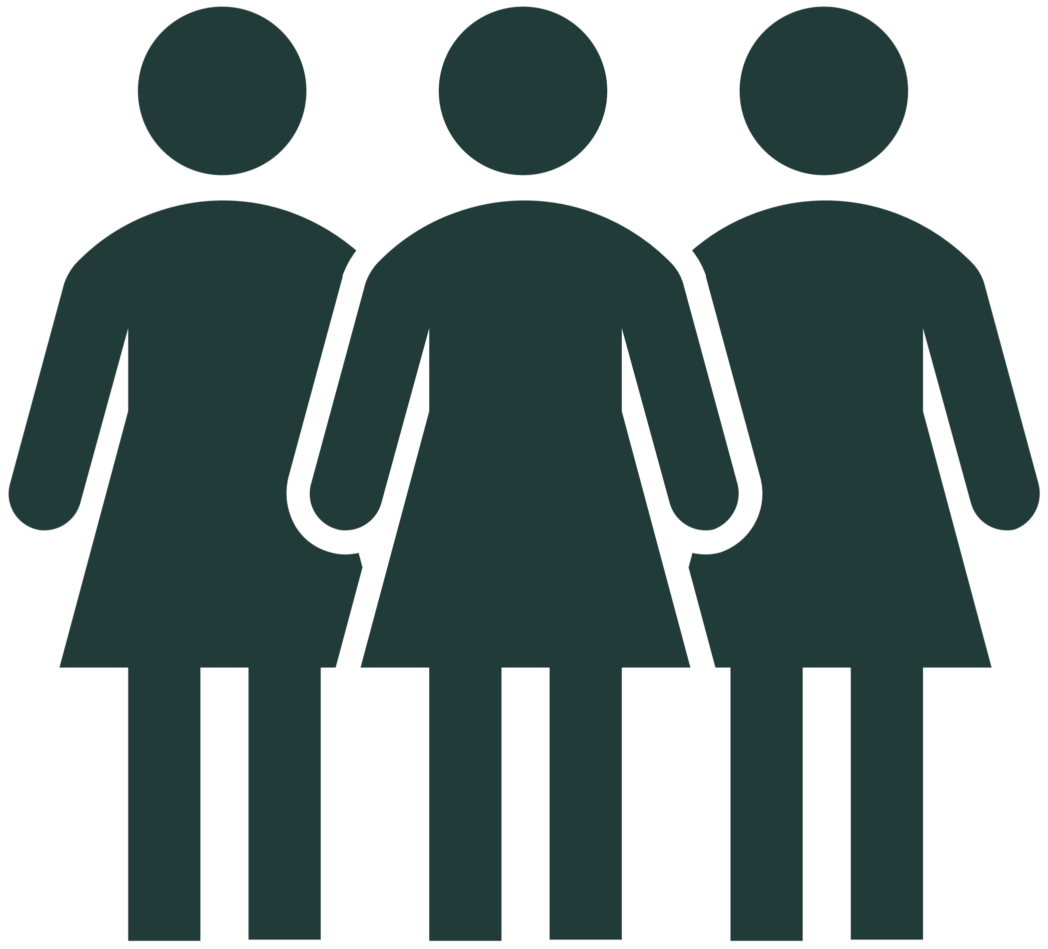
## strongest model:

## ARIMA1 (SARIMAX)

```
                               SARIMAX Results
===============================================================================
Dep. Variable:     Principal Component   No. Observations:                  51
Model:                  ARIMA(1, 0, 0)   Log Likelihood                -142.871
Date:                 Sat, 12 Oct 2024   AIC                            291.742
Time:                         09:25:13   BIC                            297.538
Sample:                              0   HQIC                           293.957
                                  - 51
Covariance Type:                   opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0006      4.724      0.000      1.000      -9.258       9.259
ar.L1         -0.0333      4.870     -0.007      0.995      -9.579       9.512
sigma2        15.8771      4.667      3.402      0.001       6.731      25.023
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):              4416.83
Prob(Q):                              0.99   Prob(JB):                         0.00
Heteroskedasticity (H):               0.03   Skew:                             6.68
Prob(H) (two-sided):                  0.00   Kurtosis:                        46.59
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

# forecasting & results



# bibliography & datasets

*Bibliography*

- Goines, David Lance. “Anatole France.” Poster accessed at [https://www.goines.net/Gallery/gal_xtra/007_anatole_france.gif] in September 2024.
- Ludden, Jennifer. 06/28/2024. “The Supreme Court says cities can punish people for sleeping in public places.” *Morning Edition.* National Public Radio Broadcast, accessed at [https://www.npr.org/2024/06/28/nx-s1-4992010/supreme-court-homeless-punish-sleeping-encampments] in September 2024.
- Wikipedia article: [https://en.wikipedia.org/wiki/Affordable_Care_Act], accessed in 2024.

*Databases*

- Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group. “Table 1: Total All Payers State Estimates by State of Residence (1991 - 2020) - Personal Health Care (Millions of Dollars).” [https://www.chcf.org/publication/2023-edition-california-health-care-spending/#related-links-and-downloads]
- California Government dataset, https://www.dir.ca.gov/OPRL/CPI/CPICalculator/CpiCalculator.aspx, accessed 2024.
- Department of Housing and Urban Development [https://www.hudexchange.info/resource/3031/pit-and-hic-data-since-2007/], accessed 2024.
- U.S. Bureau of Labor Statistics, Unemployment Rate in California [CAUR], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CAUR, June 2, 2024.
- U.S. Federal Housing Finance Agency, All-Transactions House Price Index for California [CASTHPI]. Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CASTHPI, June 2, 2024.
- The World Bank, World Development Indicators (2024). *US Health Indicators; Poverty and health indicators*. Retrieved from [https://databank.worldbank.org/source/world-development-indicators] in 2024.