Table of contents

1 Introduction

2 Data

Data used in this paper was cleaned, processed and tested with the programming language R (R Core Team 2022). Also with support of additional packages in R: `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `janitor` (Firke 2023), `readr` (Wickham, Hester, and Bryan 2023), `knitr` (Xie 2014), `rstanarm` (Goodrich et al. 2023), `modelsummary` (Arel-Bundock 2023), `tidybayes` (Kay 2023), `loo` (Vehtari et al. 2023), `testthat` (Wickham Year of publication), `KableExtra` (Zhu 2023).

Table 1: Summary statistics of the California housing dataset

	longitude
	Min. :-124.3
	1st Qu.:-121.8
	Median :-118.5
	Mean :-119.6
	3rd Qu.:-118.0
	Max. :-114.3

	latitude
	Min. :32.54
	1st Qu.:33.93
	Median :34.26
	Mean :35.63
	3rd Qu.:37.71
	Max. :41.95

	housing_median_age
	Min. : 1.00
	1st Qu.:18.00
	Median :29.00
	Mean :28.64
	3rd Qu.:37.00
	Max. :52.00

	total_rooms
	Min. : 2
	1st Qu.: 1448
	Median : 2127
	Mean : 2636
	3rd Qu.: 3148
	Max. :39320

	total_bedrooms
	Min. : 1.0
	1st Qu.: 296.0
	Median : 435.0
	Mean : 537.9
	3rd Qu.: 647.0
	Max. :6445.0
	NA's :207

	population
	Min. : 3
	1st Qu.: 787
	Median : 1166
	Mean : 1425
	3rd Qu.: 1725
	Max. :35682

	households
	Min. : 1.0
	1st Qu.: 280.0
	Median : 409.0
	Mean : 499.5
	3rd Qu.: 605.0
	Max. :6082.0

	median_income
	Min. : 0.4999
	1st Qu.: 2.5634
	Median : 3.5348
	Mean : 3.8707
	3rd Qu.: 4.7432
	Max. :15.0001

	median_house_value
	Min. : 14999
	1st Qu.:119600
	Median :179700
	Mean :206856
	3rd Qu.:264725
	Max. :500001

Table 2: Count of missing values for each variable

Variables	MissingValues
longitude	0
latitude	0
housing_median_age	0
total_rooms	0
total_bedrooms	207
population	0
households	0
median_income	0
median_house_value	0
ocean_proximity	0

Table 3: Count of missing values for each variable after cleaning

Variables	MissingValues
longitude	0
latitude	0
housing_median_age	0
total_rooms	0
total_bedrooms	0
population	0
households	0
median_income	0
median_house_value	0
ocean_proximity	0

Table 4: summary of modeling

2.1 Source

2.2 Method

3 Results

3.1 Data Trend

3.2 Modeling

Model Results

Table 5: cross validation of modeling

4 Discussion

4.1 Demographic Shifts

4.2 Health-related Behaviors

4.3 Government Policies

4.4 Environmental Changes

4.5 Possible Improvements

5 Conclusion

References

Arel-Bundock, Vincent. 2023. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. https://vincentarelbundock.github.io/modelsummary/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Goodrich, Ben, Jonah Gabry, Imad Ali, Sam Brilleman, and and others. 2023. *Rstanarm: Bayesian Applied Regression Modeling via Stan*. https://mc-stan.org/rstanarm.

Kay, Matthew. 2023. *Tidybayes: Tidy Data and Geoms for Bayesian Models*. https://mjskay.github.io/tidybayes/.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Vehtari, Aki, Jonah Gabry, Yuling Yao, and Andrew Gelman. 2023. *Loo: Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models*. https://mc-stan.org/loo.

Wickham, Hadley. Year of publication. *Testthat: Get Started with Testing*. https://CRAN.R-project.org/package=testthat.

———. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Xie, Yihui. 2014. *Knitr: A Comprehensive Tool for Reproducible Research in R*. Edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

Zhu, Hao. 2023. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.