RESEARCH PAPER

INTRODUCTION AND LITERATURE REVIEW

Topic: The Effect of COVID-19 Stringency Index measure on U.S. Housing Price Growth: Interactions with Unemployment, GDP Growth, Mortgage Rate, and Death Rate.

Introduction: The COVID-19 pandemic disrupted global economic systems and reshaped housing markets through a complex mix of public health policies, behavioral shifts, and macroeconomic responses. In the United States, these interventions included lockdowns, travel restrictions, and economic relief packages. The central question guiding this review is: How did the stringency of government responses to COVID-19 affect housing price growth in the U.S., especially with respect to unemployment rates, GDP growth, mortgage rates, and COVID-19 death rates? Understanding these relationships is crucial, as the housing market plays a major role in household wealth and macroeconomic stability. This review is organized by major themes emerging in the literature: (1) direct effects of stringency on housing prices, (2) the mediating roles of unemployment, GDP, and mortgage rates, (3) spatial and demographic disparities.

Direct Effects of Stringency on Housing Prices

Research consistently shows that the government stringency index measuring the restrictiveness of policies had a non-linear effect on U.S. housing prices during the pandemic. Several studies found that moderate restrictions tended to stabilize or even boost house price growth by reducing future uncertainty and quickly restoring market confidence(Gupta et al., 2021).However, when policies became extremely restrictive, housing price growth slowed or reversed, likely due to broader economic disruptions and reduced mobility(Gamber et al., 2023a).Some metropolitan regions experienced stronger effects than rural areas, reflecting differences in COVID-19 prevalence and local economic resilience(Yörük, 2022).Most studies use statistical methods that track changes over time across states or cities, often relying on data from Zillow or the Federal Housing Finance Agency (FHFA). However, they usually don't account for the fact that stricter lockdowns often happened in places with more COVID-19 deaths. This overlap makes it hard to tell whether higher policy stringency actually caused changes in housing prices, or if both were simply reactions to severe outbreaks, a gap this study aims to fix (Lee & Huang, 2022).

Mediating Roles: Unemployment, GDP Growth, and Mortgage Rates

The literature identifies unemployment and GDP contraction as primary mediators of the relationship between policy stringency and housing markets (Yilmazkuday, 2023).Despite sharp unemployment spikes at the pandemic's outset, house prices rose, an anomaly explained by rapid monetary policy easing, record-low mortgage rates, and widespread mortgage forbearance programs that sustained housing demand even as incomes declined (Gamber et al., 2023). Empirical evidence shows that forbearance and fiscal relief measures played a crucial role in mitigating the traditional negative link between unemployment and house prices by preventing forced sales and stabilizing household finances (Gupta et al., 2021). Simultaneously, GDP growth rebounded as restrictions eased, reinforcing house price growth by boosting consumer confidence and supporting housing investment. While this body of work convincingly links policy stringency to supply bottlenecks and cheap credit, most studies rely on national-level data, obscuring heterogeneity across states with varying supply elasticities and income distributions (Yilmazkuday, 2023). Moreover, few studies disentangle the interaction between local stringency and forbearance policies, leaving unresolved whether credit support or physical restrictions played the dominant role in shaping housing market resilience (Gupta et al., 2021).

Spatial and Demographic Mediation

A central theme in recent COVID-19 housing research concerns the spatial and demographic heterogeneity of policy stringency effects across U.S. housing markets. (Gupta et al., 2021) demonstrated that the impact of government restrictions varied widely across regions, with urban areas facing heightened volatility in house prices and mobility, especially where pandemic severity was greatest and policy enforcement strictest. (Gamber et al., 2023) find that migration patterns, facilitated by remote work trends, drove increased housing demand and rapid price appreciation in suburban and rural markets, reinforcing the role of perceived safety and lifestyle change (Gupta et al., 2021). Furthermore, studies highlight those demographic disparities deepened during the pandemic, with first-time buyers and lower-income households experiencing greater barriers to entry as prices outpaced incomes (Li & Zhang, 2021). Overall, the literature suggests that both spatial context and demographic vulnerability shaped the relationship between stringency and housing market outcomes, with policy impacts amplifying pre-existing inequities across regions and social groups (Wang, 2022).

DATA AND DATA ANALYSIS

DATA GENERATION

The dataset for this study was created by combining information from several public sources to understand how COVID-19 affected housing markets across U.S. states from 2020 to 2023. The Stringency Index, which measures how strict government policies like lockdowns and travel limits were, came from the *Oxford COVID-19 Government Response Tracker (OxCGRT)*. Data on COVID-19 deaths per million people were taken from *Our World in Data (OWID)* to show how severely each state was impacted. Housing price data were gathered from *Zillow Research* to capture quarterly changes in home values. Economic factors such as state GDP growth and mortgage interest rates were collected from the *Federal Reserve Economic Data (FRED)*. All data were aligned to a quarterly frequency, and merged by state and date, resulting in a unified panel dataset.

DATA DICTIONARY

The dataset contains 7344 rows and 8 columns.

- Dependent Variable: Housing Prices Growth Rate in USA

This is the Quarterly percentage growth in housing prices from 2020 to 2023.

- Independent Variable: Stringency Index.

This is the Government response index (lockdowns, travel limits, etc.) from 2020 to 2023.

- Control Variables:

Deaths per million- COVID-19 deaths per million residents.

GDP growth - Quarterly percentage change in state GDP.

Average interest rate - Average mortgage interest rate for the quarter.

Unemployment rate - Unemployment rate for the state and quarter.

State- Name of USA state.

Quarter- Quarter and year of observation.

This dataset is a great fit for studying how COVID-19 restrictions affected U.S. housing prices. The data is state-level and quarterly from 2020 to 2023 and helps us track how housing markets changed over time and how different factors interacted during the pandemic.

EXPLORATORY DATA ANALYSIS

EDA was performed on the unified dataset to deal with missing values and outlier detection to understand the underlying structure of the dataset. This included examining the distribution of each variable, assessing correlations among predictors, and identifying patterns or anomalies that could bias the analysis.

Missing Values: There were 1944 missing values in Stringency index column and 1494 in deaths per million columns. These were handled by filling them with zeroes because in early 2020 there were no stringency measures imposed and no deaths from covid.

Normalization: The dataset was normalized so that all the numeric predictor columns were scaled between 0 and 1 using MinMaxScaler. This helped keep everything on the same level and avoid one large-valued feature overpowering the others in the model. Columns like state and quarter were not changed because they are not numeric. After scaling, the dataset became easier for the regression model to understand and compare across features.

Descriptive Statistics: This was computed using the. describe() method.

variable	count	mean %	median	std dev	min	max
housing price growth rate	7344	0.701301	0.628134	0.735119	-1.35733	2.779709
stringency index	7344	35.78153	29.93696	16.17362	13.56198	78.42937
GDP growth rate	7344	0.672327	0.691223	1.501945	-2.39912	3.750001
average interest rate	7344	0.922277	0.85	0.347637	0.35	1.778125
deaths per million	7344	204.020366	119.447209	238.988262	-23.89	1589.83
unemployment rate	7344	4.467688	3.9	1.859337	1.7	8.6

The descriptive statistics show that the dataset is large and consistent, with 7,344 observations across all variables. Housing price growth varies widely across states and years, ranging from negative growth to strong positive growth. Stringency index, GDP growth, interest rates, unemployment, and deaths per million also show substantial variation, indicating meaningful differences in economic and health conditions across states during the COVID period.

Outlier Detection: The outliers were detected by computing Z scores for each numeric column. Data points with Z>3 was marked as outliers.

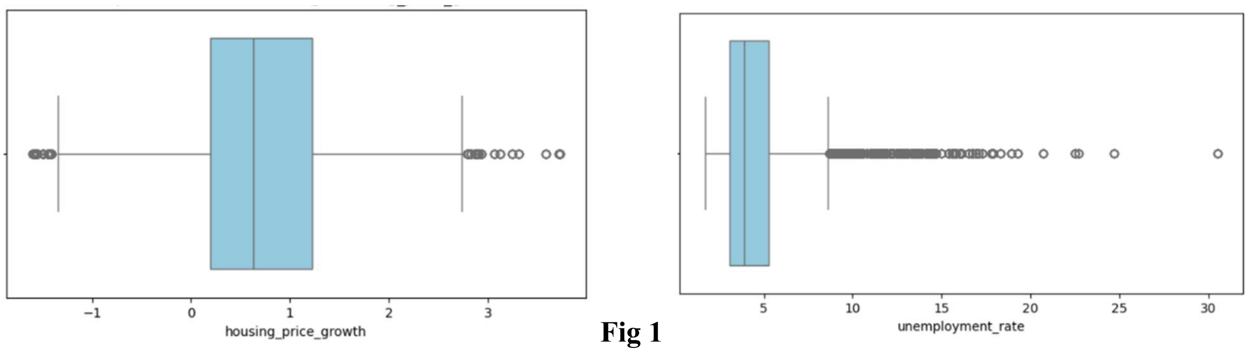

Fig 1

Fig 1 shows the boxplot of outliers which were handled using the IQR method.

Correlation Matrix:

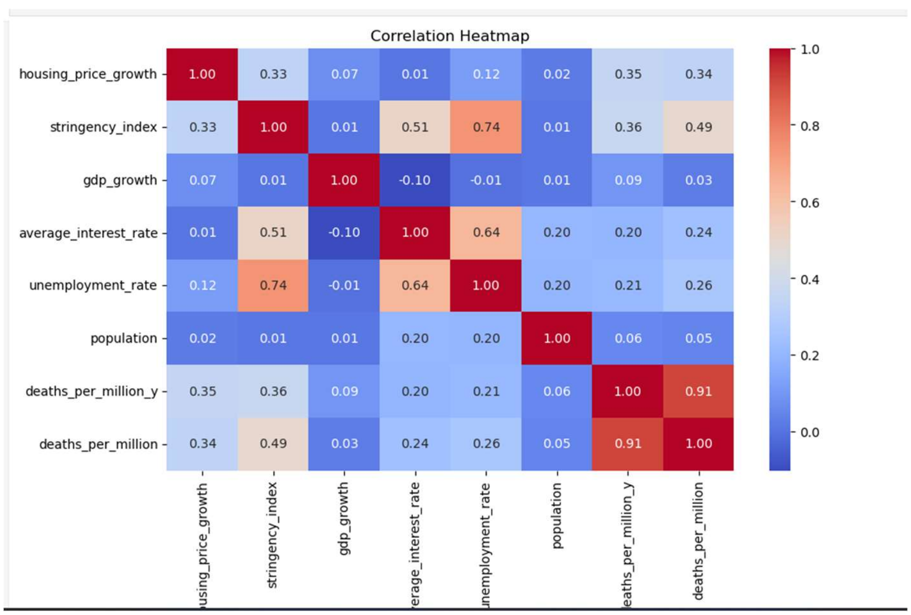

Fig 2

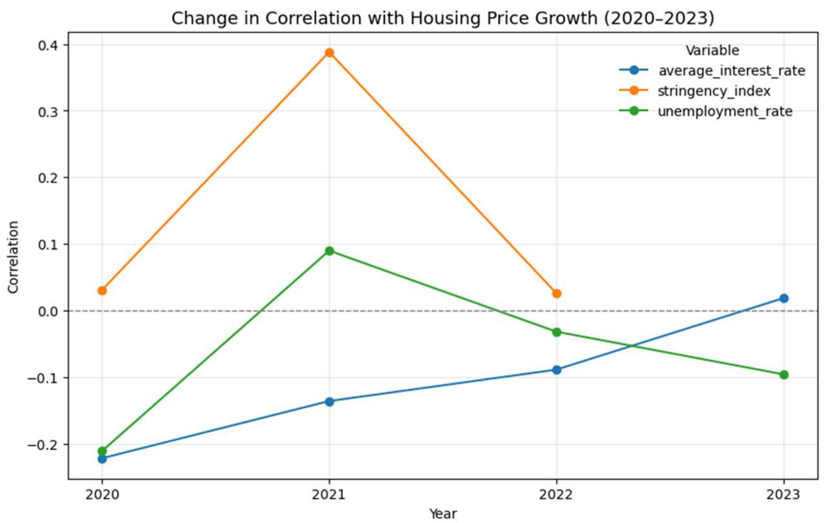

Fig 3

Fig 2 shows the heatmap of how housing price growth related to other factors during the pandemic. Overall, stricter lockdowns were linked with higher unemployment, while housing prices had only weak direct links to most factors. Fig 3 shows that in 2021, housing prices rose the most when interest rates were low and stringency was high, but as rates increased in 2022–2023, housing growth slowed down.

Pair Plots: Fig 4 shows the pair plots were plotted to analyze the numeric relationships.

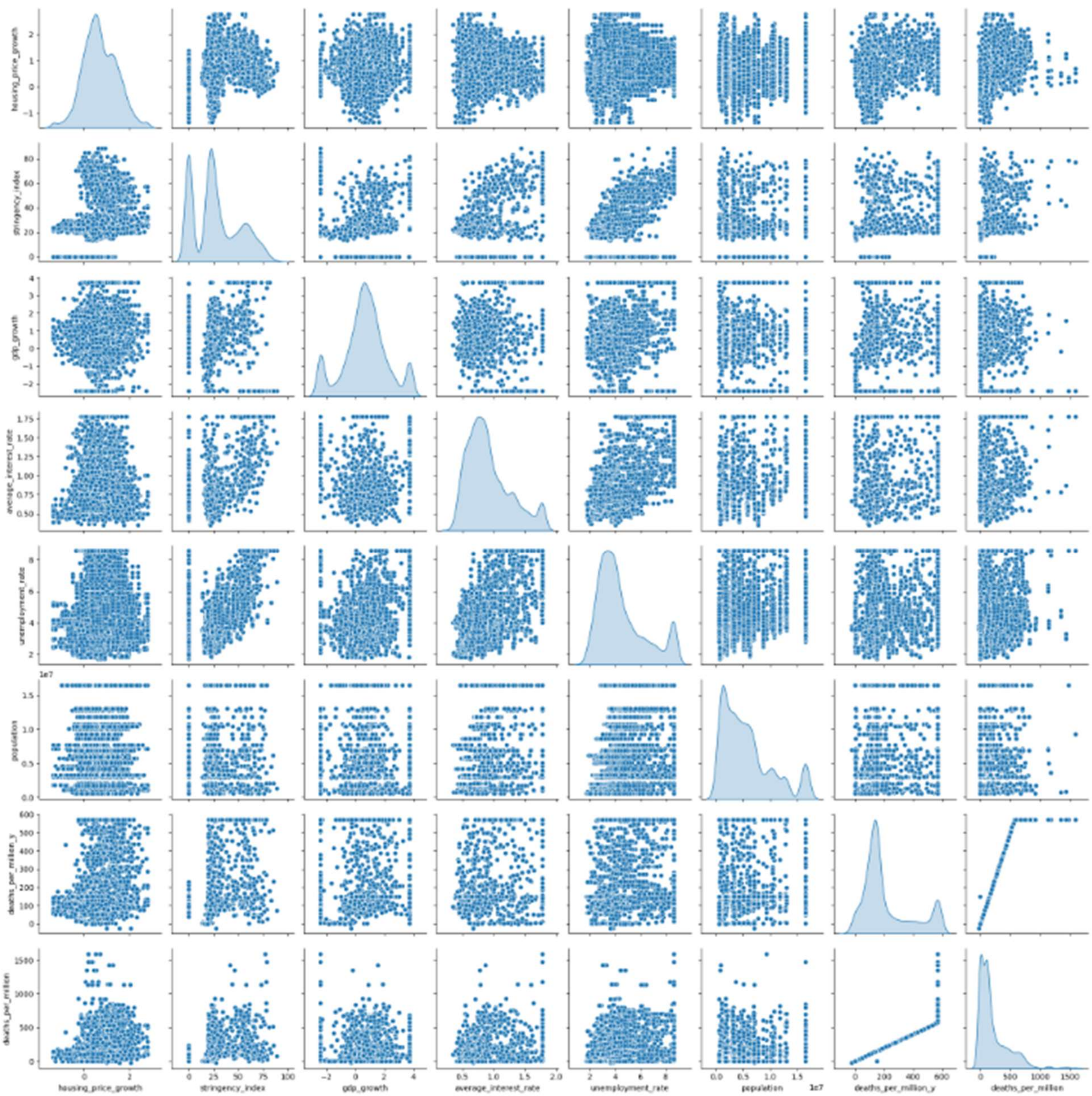

Fig 4

Housing Price Growth- This shows no strong linear pattern with any of the variables.

Stringency Index- This shows a negative trend with deaths per million and positive trend with unemployment rate**.**

Deaths per million- Negatively related with stringency and unemployment. Slightly right-skewed distribution means a few states had much higher death counts.

Average Interest Rate and Unemployment: The scatter plot between them shows a positive upward trend indicating that they increased together during tighter pandemic or recovery phases.

The bimodal distributions suggest phase-wise effects of COVID-19 restrictions across states.

Distribution Plots:

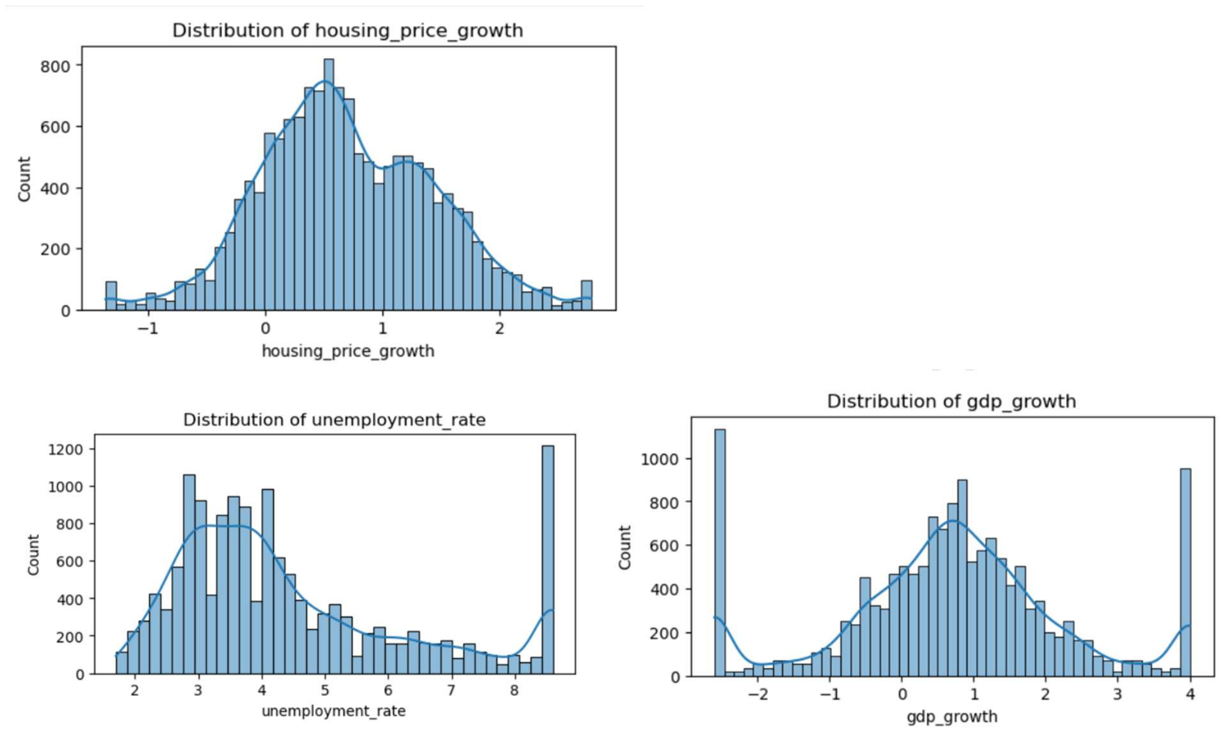

Fig 5

- The housing price growth curves looks bell shaped but slightly right skewed. This means that most states had **moderate, positive housing price growth** during the 2020–2023 period. Few states experienced **very high growth** (the right tail) whereas the 0 indicates that some states even saw price declines.

- GDP growth varied dramatically by quarter, some states plunged while others rebounded faster, showing a cyclical COVID impact.

- Unemployment rate distribution is very **right-skewed**, meaning most states had moderate unemployment but a few faced extreme spikes around 10–30%.

Statewise Analysis

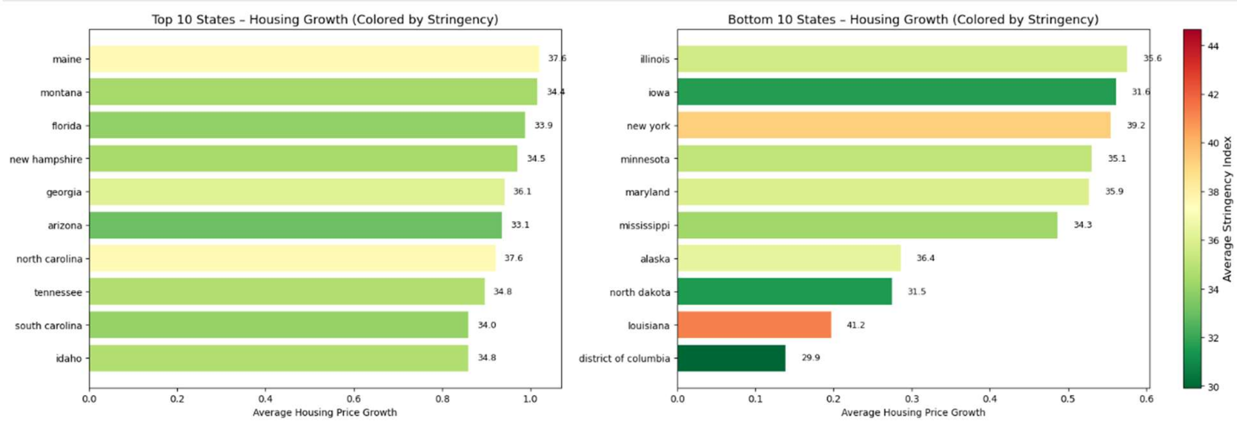

Fig 6

Fig 6 compares the top and bottom ten states by housing price growth, with color intensity representing average COVID Stringency Index. States with lower restrictions (lighter green) generally experienced stronger housing growth, whereas states with stricter lockdown measures (darker/red tones) showed weaker price appreciation. This pattern suggests that stricter containment policies may have dampened housing market momentum during the pandemic period.

Fig 7

In 2020–2021, interest rates dropped to very low levels, and most states saw strong growth in housing prices. But when rates started rising again in 2022–2023, housing price growth decreased everywhere. This suggests that low borrowing costs during the pandemic made it easier for people to buy homes, which pushed prices higher.

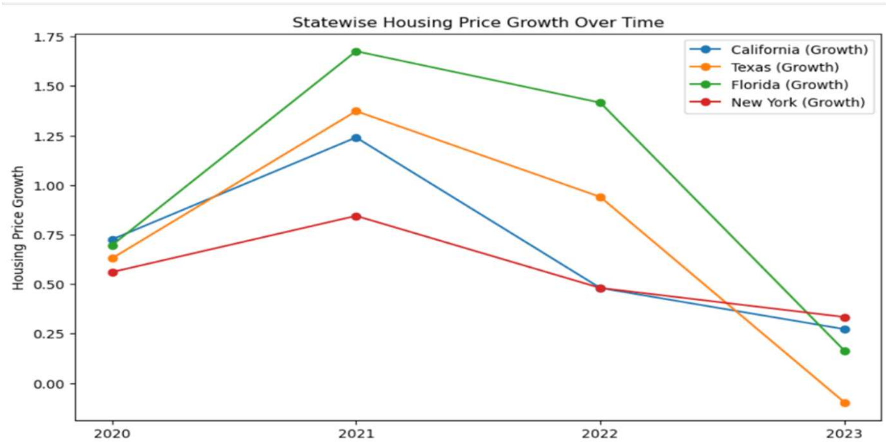

Fig 8

Fig 8 is the time-series chart which shows that after major interest rate cuts in 2021, states with fewer restrictions like Florida and Texas experienced the fastest housing price growth, driven by migration and easier borrowing. When interest rates increased in 2022 and 2023, growth rates fell sharply across all states with New York's line lower, reflecting slower recovery due to stricter lockdowns and limited movement.

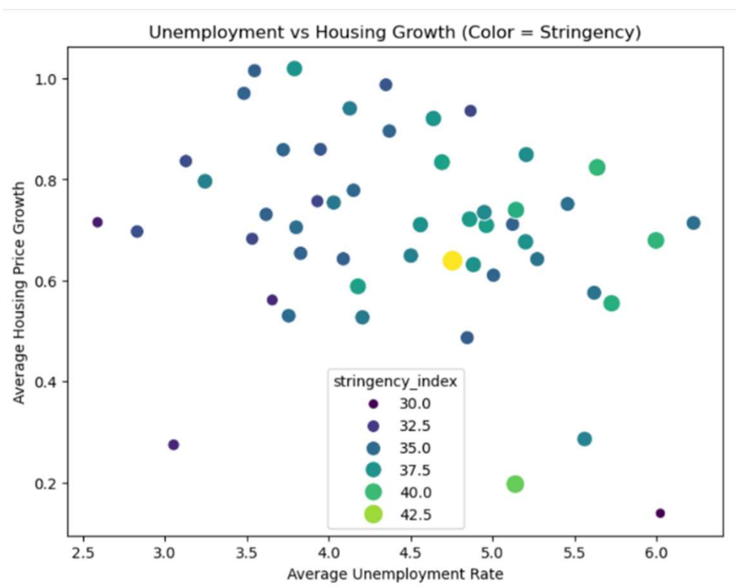

Fig 9

Fig 9 shows that states with stricter COVID measures faced higher unemployment and slower housing price growth. In contrast, states with looser restrictions had lower unemployment and stronger housing markets. This suggests that strict lockdowns indirectly slowed housing growth by increasing unemployment and reducing economic activity.

METHODOLOGY:

This study uses a quantitative, regression-based methodology to examine how the COVID-19 Stringency Index influenced housing price growth across U.S. states from 2020 to 2023. The dataset combines both state-level and quarterly observations, allowing the model to capture how economic conditions changed over time within each state during the pandemic. Before modeling, extensive data preprocessing was performed, including checking for missing values, detecting outliers, and normalizing all numerical predictor variables using MinMaxScaler to ensure they were on a comparable scale. Non-numeric variables such as state and quarter were left unchanged. The dependent variable in the regression is Housing Price Growth, which measures quarterly changes in state-level home prices. The key explanatory variable is the COVID-19 Stringency Index, which reflects the strictness of government policies such as lockdowns, travel restrictions, business closures, and school shutdowns. Additional control variables incorporated into the model include Unemployment Rate**,** GDP Growth**,** Average Interest Rate**,** Deaths **per** Million, and Population**,** each capturing different dimensions of economic and pandemic-related impacts. Before modeling, the data underwent preprocessing including handling missing values, outlier detection, and Minmax normalization of numerical variables to ensure consistent scaling. To refine the model and identify the most influential predictors, a forward selection procedure based on AIC was applied. This iterative process added variables only when they improved the model's AIC score, resulting in a final set of six predictors used for the OLS regression. The statistical technique applied is **Multiple Linear Regression using Ordinary Least Squares (OLS)**, implemented through Statsmodels' `smf.ols()` function. This approach is appropriate because the target variable is continuous and all predictors are numerical, enabling the model to estimate the marginal effect of each factor on housing price growth while holding the others constant. The final regression model follows the form:

$$\text{HPG} = \beta_0 + \beta_1 St + \beta_2 Ur + \beta_3 Gr + \beta_4 Ir + \beta_5 Dm + \epsilon$$

St → Stringency Index

Ur → Unemployment Rate

Gr → GDP Growth Rate

Ir → Interest Rate

Dm → Deaths per Million

Overall, this methodology provides a structured and statistically sound framework for identifying the extent to which COVID-19 restrictions and economic conditions shaped state-level housing market performance during the pandemic.

RESULTS AND DISCUSSION

OLS Regression Results

Dep. Variable:	housing_price_growth	R-squared:	0.226			
Model:	OLS	Adj. R-squared:	0.225			
Method:	Least Squares	F-statistic:	214.2			
No. Observations:	4406	Prob (F-statistic):	1.76E-240			
Df Residuals:	4399	Log-Likelihood:	-4331.5			
Df Model:	6	AIC:	8677			
Covariance Type:	nonrobust	BIC:	8722			
	Co efficient	**std error**	**t**	**p> |t|**	**0.025**	**0.975**
Intercept	0.3348	0.031	10.947	0	0.275	0.395
Stringency Index	1.3596	0.06	22.64	0	1.242	1.477
GDP Growth	0.0856	0.04	2.139	0.033	0.007	0.164
Average interest rate	-0.5219	0.053	-9.852	0	-0.626	-0.418
Unemployment rate	-0.5291	0.062	-8.521	0	-0.651	-0.407
Population	0.21	0.036	5.861	0	0.14	0.28
Deaths per million	0.6179	0.037	16.787	0	0.546	0.69

Evaluation Metrics:

$R^2 = 0.226$

The model explains 22.6% of the variation in housing price growth. This is because housing markets depend on many unobserved factors.

Adjusted $R^2 = 0.225$

This adjusts for the number of predictors. After adjusting for the number of variables, the model still explains about 22.5% of the variability.

AIC = 8677

Used mainly for model selection such as forward selection. Lower AIC = better model fit when comparing models.

Regression Statistics

Mean Error (ME)	-0.0037
Root Mean Squared Error (RMSE)	0.6491
Mean Absolute Error (MAE)	0.5101
Mean Percentage Error (MPE)	-22.9936
Mean Absolute Percentage Error (MAPE)	251.8969

The RMSE (0.6491) and MAE (0.5101) suggest that the model predicts housing price growth within about half a percentage point on average, which is reasonable given the volatility of the 2020–2023 period. The negative Mean Percentage Error indicates a slight tendency toward underprediction, while the extremely high MAPE occurs because percentage errors become unstable when actual growth values are close to zero. Therefore, MAE and RMSE provide the most reliable indicators of model accuracy in this case.

Overall, the results indicate that the regression model is statistically significant, and all predictors included in the analysis show meaningful contributions to explaining housing price growth. The residuals exhibit no autocorrelation, indicating that the model's errors behave as expected and do not follow any systematic pattern. The R^2 value is reasonable and the AIC and BIC values indicate that the model is well-suited for comparison and selection among alternative specifications.

INTERPRETATIONS

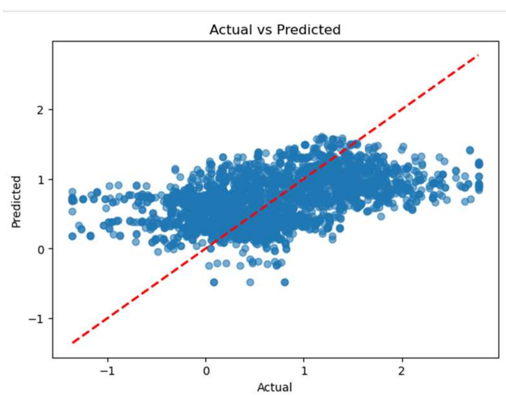

Fig 10

Fig 10 is the scatter plot that compares the true housing price growth values with the model's predicted values. The red dashed line represents the ideal situation where predictions perfectly match actual values. The points cluster around the general upward trend, which means the model is capturing the direction of housing price changes reasonably well. However, the scatter is wide, especially at higher values, showing that the model has moderate prediction error and tends to underpredict when actual values are high. This confirms the earlier error metrics (RMSE and MAE) showing that the model is accurate enough but not perfect.

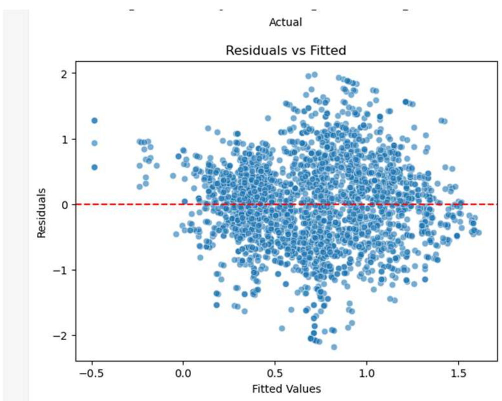

Fig 11

Fig 11 shows residuals (prediction errors) on the y-axis against the predicted values on the x-axis, with the red dashed line at zero showing the ideal "no error" baseline. The residuals are randomly scattered around zero without forming any clear pattern which means that the model errors are not dependent on the fitted values, and the assumption of linearity and homoscedasticity is reasonably satisfied.

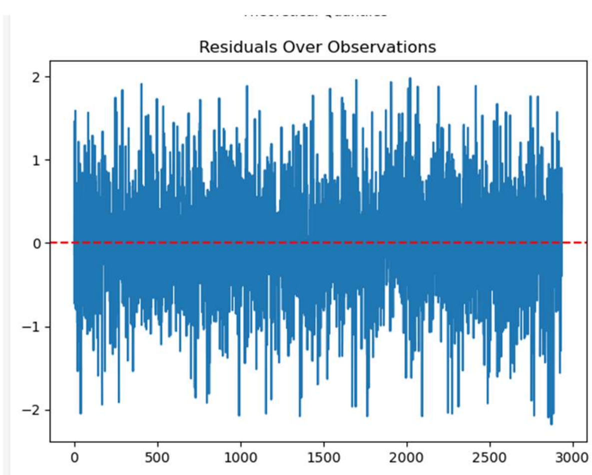

Fig 12

Fig 12 shows that the residuals fluctuate randomly around zero without any noticeable pattern or trend. This supports the Durbin-Watson statistic (~2.0) and confirms there is no autocorrelation, meaning the model errors are independent and do not follow a time-based or sequential pattern. This strengthens the reliability of the regression results.

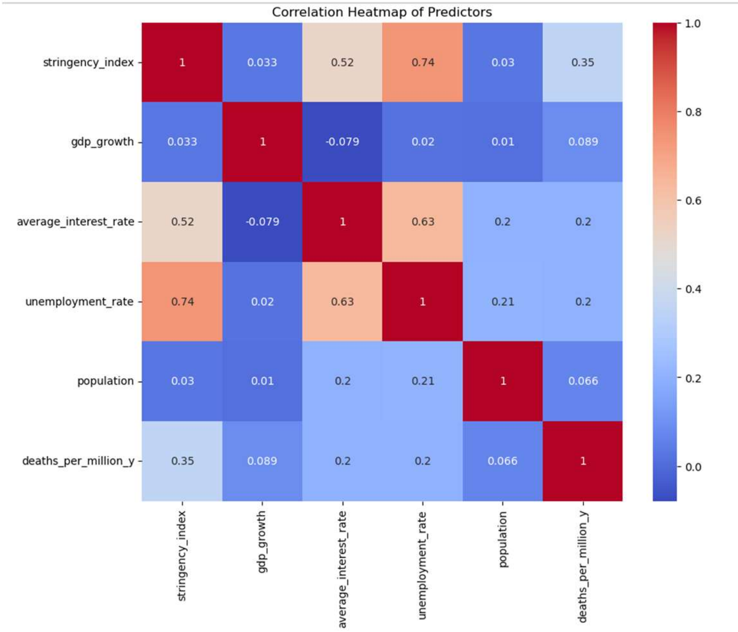

Fig 13

Fig 13 shows pairwise correlations between all predictor variables. The only relatively strong correlations are Stringency Index and Unemployment rates. No correlations are so high that they indicate severe multicollinearity. This is good because it means the predictors provide unique information to the model, and coefficient estimates remain stable and interpretable.

Overall, the visualizations show that the model fits the data reasonably well, with predictions tracking actual values despite some noise. The residuals are random, unbiased, and approximately normal, and they show no autocorrelation. The predictors also do not exhibit strong multicollinearity. Together, these checks confirm that the main OLS assumptions are largely satisfied.

REGRESSION EQUATION

$$HPG = 0.3348 + 1.3596St + 0.0856Gr - 0.5219Ir - 0.5291Ur + 0.6179Dm$$

Overall, the study successfully answers the initial research question by demonstrating how COVID-19 policy stringency influenced U.S. housing price growth and how economic conditions mediated this relationship, aligning closely with the themes identified in the literature. Consistent with studies such as Gupta et al., 2021 and Gamber et al., 2023, the results show that stricter government restrictions were strongly associated with higher housing price growth, reflecting the combined effects of remote-work migration, limited housing supply, and reduced mobility that intensified demand even in highly regulated states. The expected economic patterns also held true, unemployment and mortgage interest rates significantly reduced housing price growth, while GDP growth produced a modest positive effect, supporting earlier findings that monetary and fiscal interventions helped sustain market activity despite economic strain. Pandemic severity, measured through deaths per million, showed a positive relationship with housing price growth, echoing the literature's discussion of spatial disparities and migration away from high-risk urban centers toward safer suburban and rural areas (Yörük, 2022; Li &

Zhang, 2021). These findings validate the assumptions outlined in the introduction and literature review particularly that stringency, economic shocks, and demographic factors jointly shaped pandemic-era housing dynamics and they clarify the mixed expectations around whether strict policies would suppress or stimulate housing markets. Ultimately, the study confirms that a combination of strong restrictions, economic disruption, and migration behavior produced the unusual surge in housing price growth observed from 2020 to 2023.

CONCLUSION

This study shows that COVID-19 restrictions had a real impact on how housing prices changed across the United States. Surprisingly, states with stricter lockdowns often saw bigger increases in housing prices. This happened because remote work, limited housing supply, and people moving to new areas created strong demand, even during strict government rules. The results also showed that the economy still mattered: higher unemployment and rising mortgage rates slowed price growth, while stronger GDP and larger populations helped push prices up. States with more COVID-19 deaths also saw unusual housing patterns, likely because people moved away from high-risk cities into safer or less crowded areas. Overall, the findings make it clear that the housing market during COVID-19 was shaped by a mix of government policies, economic conditions, and people's movement between states. By looking at all these factors together, this study helps explain why the housing market behaved so differently during the pandemic and provides useful insight for leaders planning how to handle future crises.

GROUP MEMBERS

Sana Ambreen- Filled the dataset with missing GDP values, evaluated the suspicious values for deaths per million and fixed them, EDA visualizations, Methodology, Performed Linear Regression, Interpretations, Results and discussions, Conclusion

Azmath Noorain

Yitian Liu

REFERENCES

Gamber, W., Graham, J., & Yadav, A. (2023). Stuck at home: Housing demand during the COVID-19 pandemic. *Journal of Housing Economics*, *59*, 101908. https://doi.org/10.1016/j.jhe.2022.101908

Gupta, A., Mittal, V., Peeters, J., & Van Nieuwerburgh, S. (2021). *Flattening the Curve: Pandemic-Induced Revaluation of Urban Real Estate*. https://doi.org/10.3386/w28675

Lee, J., & Huang, Y. (2022). Covid-19 impact on US housing markets: evidence from spatial regression models. *Spatial Economic Analysis*, *17*(3), 395–415. https://doi.org/10.1080/17421772.2021.2018028

Li, X., & Zhang, C. (2021). Did the COVID-19 Pandemic Crisis Affect Housing Prices Evenly in the U.S.? *Sustainability*, *13*(21), 12277. https://doi.org/10.3390/su132112277

Wang, B. (2022). Housing market volatility under COVID-19: Diverging response of demand in luxury and low-end housing markets. *Land Use Policy*, *119*, 106191. https://doi.org/10.1016/j.landusepol.2022.106191

Yilmazkuday, H. (2023). COVID-19 and housing prices: evidence from U.S. county-level data. *Review of Regional Research*, *43*(2), 241–263. https://doi.org/10.1007/s10037-023-00187-4

Yörük, B. K. (2022). Early effects of COVID-19 pandemic-related state policies on housing market activity in the United States. *Journal of Housing Economics*, *57*, 101857. https://doi.org/10.1016/j.jhe.2022.101857

Chatgpt:

yes give me a polished methodology paragraph combining this one and my actual followed method

Here is a polished, academic-style **Methodology** paragraph that blends:

- ✔ the paragraph you wrote
- ✔ the regression procedure you actually followed
- ✔ the OLS model, variable descriptions, and preprocessing steps you implemented

It is clean, formal, and ready for your report:

📌 Final Polished Methodology Paragraph

This study uses a quantitative, regression-based methodology to examine how the COVID-19 Stringency Index influenced housing price growth across U.S. states from 2020 to 2023. The dataset combines both state-level and quarterly observations, allowing the model to capture how economic conditions changed over time within each state during the pandemic. Before modeling, extensive data preprocessing was performed, including checking for missing values, detecting outliers, and normalizing all numerical predictor variables using MinMaxScaler to ensure they were on a comparable scale. Non-numeric variables such as state and quarter were left unchanged. The dependent variable in the regression is **Housing Price Growth**, which measures quarterly changes in state-level home prices. The key explanatory variable is the **COVID-19 Stringency Index**, which reflects the strictness of government policies such as lockdowns, travel restrictions, business closures, and school shutdowns. Additional control variables incorporated into the model include **Unemployment Rate**, **GDP Growth**, **Average Interest Rate**, **Deaths per Million**, and **Population**, each capturing different dimensions of economic and pandemic-related impacts. The statistical technique applied is **Multiple Linear Regression using Ordinary Least Squares (OLS)**, implemented through Statsmodels' `smf.ols()` function. This approach is appropriate because the target variable is continuous and all predictors are numerical, enabling the model to estimate the marginal effect of each factor on housing price growth while holding the others constant. The final regression model follows the form: