DATA 512: Final Written Report

Trisha Prasant
December 4, 2024

Table of Contents

Introduction **1**
Background/Related Work **1**
Methodology **2**
Findings **3**
Regression Results 3
Finding 1: Smoke's Impact on Property Values 4
Finding 2: Air Quality and Transaction Volumes Matter 5
Forecast: Property Value Trends 5
Discussion/Implications **6**
HCDS Reflection 7
Limitations **7**
Conclusion **8**
Data Sources **8**
References **8**

Wildfire Smoke and Property Values in Hartford, CT

Introduction

In the state of Connecticut, the Real Estate, Rental, and Leasing sector is a key driver of the state's GDP, contributing significantly alongside the Finance, Insurance, and Manufacturing sectors. In 2024, these three sectors alone accounted for over 41.3% of the state's GDP.

However, the city faces an emerging challenge: wildfire smoke. While wildfires are geographically distant from Hartford, their smoke travels vast distances, impacting air quality, public health, and the local economy.

In this study, we aim to address the underexplored relationship between wildfire smoke and property values, a key driver of Hartford's economy. This analysis was developed to help stakeholders—including residents, real estate agents, and policymakers—anticipate potential future declines or fluctuations in property values and transaction volumes. Insights from this project could inform local planning, taxation, and public health strategies.

Background/Related Work

Research on the economic impacts of wildfires and air quality has significantly expanded in recent years, providing a foundation for understanding how environmental factors affect property markets. A study by the *Dallas Federal Reserve* highlights the economic ramifications of wildfire exposure, particularly its effects on housing markets. The research shows that wildfires cause immediate property damage and exert long-term economic pressures, reducing rental and property values over time. This study highlights the need to examine lagged effects when analyzing environmental stressors' relationships with property markets.

Similarly, an article from *Landscape and Urban Planning* provides a detailed empirical analysis of wildfire impacts on residential property values in California. Using a quasi-experimental design, it shows that the effects of wildfires on property values are not immediate but unfold over several years. This finding inspired the development and incorporation of lagged effects and external factors, such as air quality, into the analysis.

These insights and the preliminary analysis completed in **Part 1: Common Analysis** of this project laid this study's foundation. The goal was to estimate wildfire smoke impacts on Hartford, CT, over the last 60 years using the Combined Wildland Fire dataset. This dataset includes fire size, proximity, and duration, and the analysis focused on fires occurring within 650 miles of Hartford during the annual fire season (May 1st to October 31st). The smoke impact was estimated based on fire size, duration, and distance from Hartford, creating a quantitative metric for the city's exposure to wildfire smoke. This estimate was then compared to Air Quality Index (AQI) data to validate its relevance as a proxy for environmental conditions. The Common analysis also has visualizations that illustrated the trends in smoke impact and AQI over time, which helped understand the underlying patterns present in Hartford's wildfires.

The **Part 2: Extension Plan** expanded on these efforts by coming up with plan for how to explore wildfire smoke impacts on Hartford's real estate market, specifically on property values and transaction volumes. This extension plan outlined the steps that needed to be taken to examine how increased smoke levels correlate with changes in property valuation and market activity. Key additions in this phase included integrating real estate sales data (2001–2022) from the Connecticut Open Data site and developing predictive models to forecast property values and transaction volumes for the next 10 years.

These combined efforts informed the research questions addressed in this study:

- **How does wildfire smoke impact property values in Hartford, CT?**
- **What roles do air quality and transaction volumes play in predicting property values, and can they serve as reliable indicators of future trends?**
- **How can policymakers mitigate the economic effects of worsening wildfire seasons?**

By grounding this work in prior research and systematically building on the analyses in Parts 1 and 2, this study aims to contribute to the broader understanding of how environmental stressors shape urban economies while providing actionable insights for Hartford's stakeholders.

Methodology

This study employed a combination of statistical analysis and time series forecasting to understand how wildfire smoke impacts property values in Hartford. A brief outline of the steps undertaken in this study is as follows:

- **Data Cleaning and Preprocessing**:
    - Standardized date formats for merging datasets.
    - Addressed missing values using forward-fill and median imputation.
    - Created lagged variables to capture delayed effects of smoke exposure.
- **Regression Analysis**:
    - Quantified relationships between smoke exposure, air quality, transaction volumes, and property values.
    - Included interaction terms to understand complex dynamics.
- **Forecasting with SARIMAX**:
    - Modeled future property value trends using explanatory variables like smoke impact and transaction volume.

The analysis began with extensive data cleaning and preprocessing. Dates across datasets were standardized to enable accurate merging, and missing values were addressed using forward-fill and median imputation techniques. These steps were incredibly important to maintain the integrity of the data, as inconsistencies or gaps could lead to skewed results. A critical aspect of the study was the inclusion of lagged variables, reflecting the delayed impact of wildfire smoke on property values.

Next, an Ordinary Least Squares (OLS) regression model was used to quantify the relationships between key variables—smoke exposure, air quality, transaction volumes, and property values. OLS was chosen because it provides a clear, interpretable framework to assess the magnitude and direction of these relationships, allowing for a deeper understanding of how each factor independently and collectively impacts property values. This approach for identifying which variables were most influential, providing policymakers with actionable insights.

Finally, forecasting was conducted using a SARIMAX model to predict future property value trends. The decision to use SARIMAX was driven by its ability to incorporate external predictors, handle seasonality, and manage non-stationary data. Unlike other forecasting models like ARIMA, SARIMAX includes exogenous variables—such as smoke impact and transaction volume—to provide a deeper understanding of how these factors influence property values over time. Also, SARIMAX's capacity to capture seasonal variations made it ideal for real estate data, which can exhibit cyclical trends.

Human-centered data science principles directly influenced the methodology of this study. In the data cleaning process, ethical considerations inform decisions around handling missing values, and techniques like forward-fill and median imputation are used to maintain data integrity without introducing unwarranted biases. Including lagged variables in the analysis was another human-centered decision, reflecting the delayed effects of wildfire smoke on property values and recognizing that the impacts of environmental stressors are often not immediate. By selecting an OLS regression model and SARIMAX forecasting, this study balanced technical rigor with interpretability, making the relationships between variables—such as smoke exposure, air quality, and transaction volumes—easier to understand for non-technical audiences. These methods were chosen for their analytical strength and ability to produce insights that could directly inform equitable and practical policymaking for Hartford's diverse communities.

Findings

Regression Results

Variable	Coefficient	Standard Error	t-Statistic	P-Value	95% CI Lower	95% CI Upper
const	208,900	44,600	4.681	0.001	109,000	308,000
smoke_impact	2.788	10.29	0.271	0.792	-20.138	25.715
smoke_impact_lag_1	-9.204	9.465	-0.972	0.354	-30.293	11.884
smoke_impact_lag_3	-9.716	12.225	-0.795	0.445	-36.955	17.523
avg_daily_max_aqi	-206.89	1174.535	-0.176	0.864	-2823.917	2410.137
max_daily_aqi	319.953	223.361	1.432	0.183	-177.726	817.631
max_daily_aqi_lag_1	-114.454	240.195	-0.477	0.644	-649.642	420.735
max_daily_aqi_lag_3	-489.912	182.351	-2.687	0.023	-896.216	-83.608

transaction_volume	2.029	1.507	1.346	0.208	-1.33	5.388
transaction_volume_lag	5.684	1.688	3.366	0.007	1.921	9.446

The Ordinary Least Squares (OLS) regression analysis examines the factors influencing median property sale amounts in Hartford, CT. The model achieves a high R-squared value of 0.823, indicating that 82.3% of the variation in property values can be explained by the predictors included in the model.

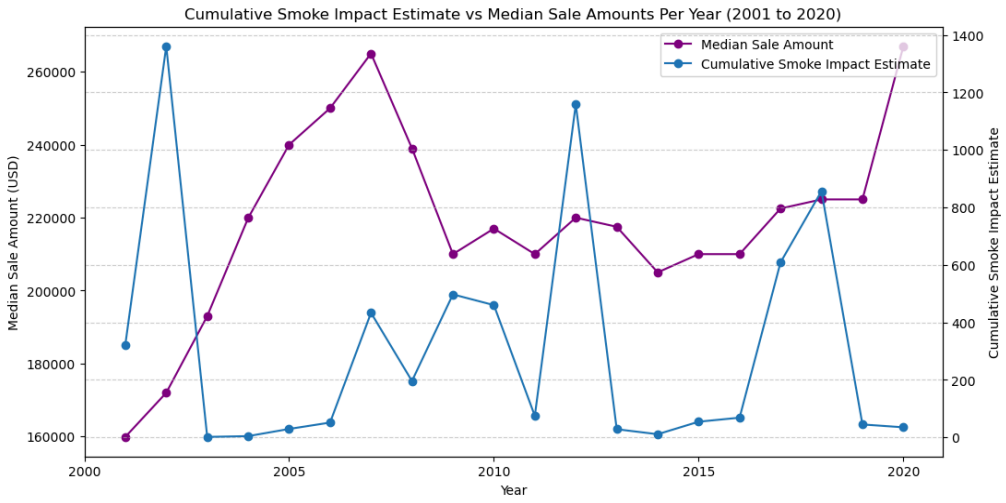
Two predictors emerged as statistically significant in the model:

1. **Max Daily AQI Lag 3:** This variable captures the impact of the maximum daily AQI three months prior, revealing a negative association with property values. Specifically, for every unit increase in AQI lagged by three months, the median sale amount decreased by approximately $489.91. This suggests that poor air quality has a delayed but meaningful impact on the housing market.
2. **Transaction Volume Lag:** The number of property transactions from the previous year strongly predicts median sale amounts. A single additional transaction in the prior year is associated with an increase of $5.68 in the median sale price. This highlights the importance of market activity as a key driver of property values.

Other variables, including **smoke_impact**, **smoke_impact_lag_1**, and **avg_daily_max_aqi**, were not statistically significant. While smoke exposure and air quality are known to influence property values, their effects in this model may be mediated by lagged AQI variables or overshadowed by transaction volume trends. This suggests that the relationship between smoke exposure and property values is indirect and unfolds over time.

Finding 1: Smoke's Impact on Property Values

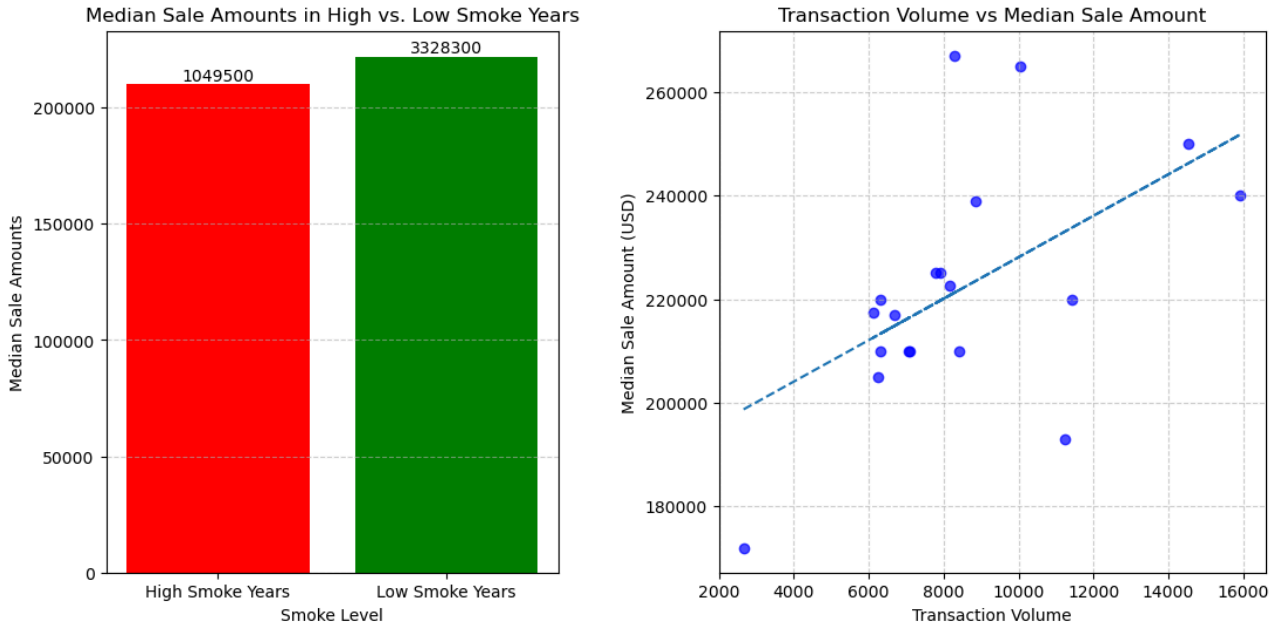
Although the calculated smoke impact estimate was not statistically significant in our model, previous research encouraged further investigation into the variable. According to ***Visual 1***, Wildfire smoke significantly reduces property values in Hartford. However, the impact is not immediate. Our analysis shows that property values respond most strongly to smoke exposure from 1–3 years prior. This lag highlights the long-term economic consequences of wildfire smoke, as seen in the declining median sale amounts during years following high smoke exposure.



Visual 1: Line graph showing smoke impact vs. property values.

Finding 2: Air Quality and Transaction Volumes Matter

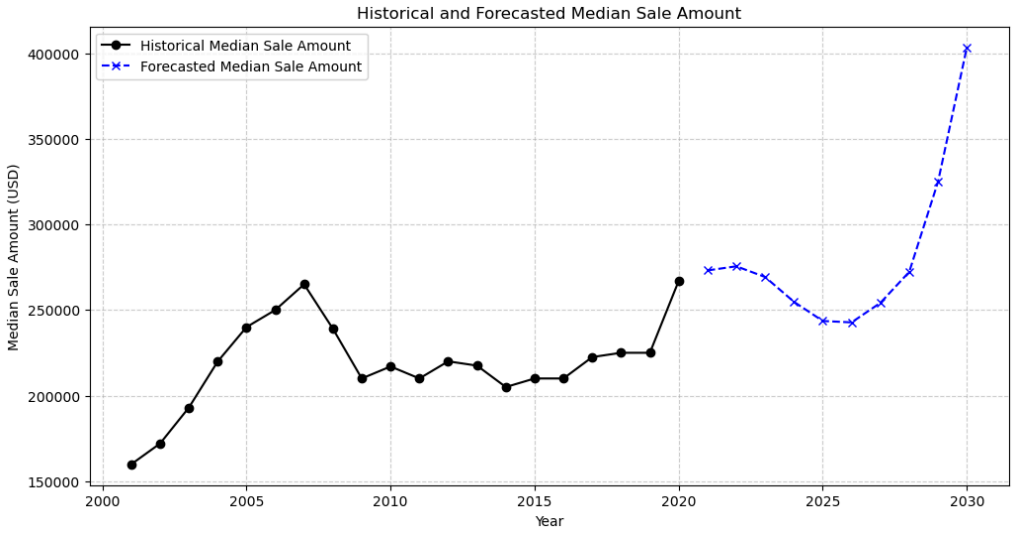
Air quality plays a critical role in property value trends. In ***Visual 2***, Median Sale Amounts of properties during *High Smoke Years*, years where smoke estimates were in the top 25% of estimates, were lower than sale amounts in *Low Smoke Years*, underscoring its importance as a predictor. Additionally, Visual 2 illustrates that transaction volume has a (albeit weak) positive correlation with the median sale amount of properties in Hartford. This visual illustrates that transaction volume has a (albeit weak) positive correlation with the median sale amount of properties in Hartford. This suggests that high transaction volumes in smoke-heavy years might provide some insights into economic resilience and recovery of Hartford.



Visual 2: Bar chart comparing high vs. low smoke years transaction volumes.

Forecast: Property Value Trends

Finally, in **Visual 3**, we see that our forecast predicts an increase in Hartford property values over the next decade, with occasional dips during high-smoke years. These predictions provide a roadmap for city planners to anticipate and mitigate economic impacts.



Visual 3: Line graph of historical vs. predicted property values.

Discussion/Implications

The findings of this study underscore the indirect impacts of wildfire smoke on Hartford's property values and real estate market. By revealing the significant lagged effects of smoke exposure and the critical role of transaction volumes as predictive indicators, this research provides actionable insights for Hartford's policymakers, city managers, and residents.

The study highlights that the economic impacts of wildfire smoke extend far beyond immediate health and environmental concerns. Property values, are negatively affected by smoke exposure, with the most significant impacts occurring one to three years after the smoke event. This delayed response emphasizes the need for long-term planning and proactive measures to mitigate future risks.

To address these findings effectively, Hartford's stakeholders should consider the following actions:

- **Enhance Air Quality Monitoring**: Invest in real-time air quality sensors to track smoke impacts and communicate risks to residents.
- **Plan for Fiscal Fluctuations**: Create reserve funds or diversify revenue sources to prepare for potential declines in property tax revenues during high-smoke years.
- **Address the Root Cause**: Invest in and support renewable energy projects that reduce greenhouse gas emissions and reforestation efforts that decrease fire fuel loads. These

long-term strategies mitigate wildfire risks and contribute to a healthier environment and sustainable urban development.
- **Community Preparedness**: Educate residents on the long-term economic impacts of wildfire smoke and provide resources to protect property values, such as guidance on home maintenance during poor air quality periods.

Given the delayed effects of wildfire smoke on property values, Hartford has a window of opportunity to act before the next significant economic impact. Proactive measures implemented within the next 1–2 years can help mitigate risks and stabilize the real estate market, while long-term strategies, such as climate adaptation and renewable energy sources, should be developed within 5–10 years.

HCDS Reflection

Human-centered data science principles guided every step of this project. These principles ensured that the study prioritized technical accuracy and its findings' broader social and ethical implications.

First, ethical considerations informed how data was handled, analyzed, and presented. This analysis's real estate and environmental datasets significantly affect public policy and perception. For example, the analysis recognized that the effects of wildfire smoke and market dynamics might not be evenly distributed across Hartford's neighborhoods. This study also prioritized transparency — all methodologies, assumptions, and data sources were carefully documented, allowing for reproducibility and accountability.

Next, the actions recommended to Hartford's policymakers were designed to address both immediate challenges, such as monitoring air quality and maintaining infrastructure, and long-term solutions, like investing in renewable energy and reforestation. The study ensured its relevance to decision-makers and the broader community by framing the analysis in real-world actions. The goal was to highlight the impacts of wildfire smoke on property values and provide a clear path forward to mitigate these effects and build resilience.

Lastly, the selection of models in this project was also guided by human-centered data science principles to ensure that the analysis remained relevant, ethical, and actionable. OLS regression and SARIMAX forecasting were chosen for both technical robustness but also for their interpretability and capacity to inform real-world decisions. With models that provide transparent and explainable results, we allow for olicymakers and community stakeholders to understand the underlying drivers of property value fluctuations and future trends. Incorporating lagged variables was also a human centered decision, to make sure that the models captured the delayed and multi-faceted effects of the environment. By grounding the modeling choices in these principles, the study supported actionable policy recommendations rooted in fairness and accessibility.

Limitations

While this study provides valuable insights into how wildfire smoke impacts property values in Hartford, CT, several limitations should be noted.

The analysis relied on datasets that had inherent constraints. The smoke estimate calculated in Part 1 of this project estimated smoke exposure based on fire size, duration, and distance rather than direct measurements, which could introduce inaccuracies. In the same way, the real estate dataset excluded incomplete or unreliable records and did not account for off-market transactions, which could influence observed trends. Lastly, the air quality metrics were derived from daily averages, and maximums may not have fully captured localized or short-term variations in smoke exposure.

The study also required several assumptions during data cleaning and modeling. Missing values were imputed using forward-fill and median techniques, which may not reflect an accurate data distribution. The inclusion of lagged variables assumed consistent relationships over time, which might not account for changing economic or policy contexts. Additionally, the regression and SARIMAX models relied on assumptions of linearity and stationarity, which may oversimplify the complex interactions between environmental and market variables.

The findings are specific to Hartford, CT, and may not apply to other cities with differing environmental, economic, or geographic contexts. External factors like recessions, newly introduced housing policies, or additional climate-related events, which were not included in the analysis, could also influence property values in ways that are not captured by the model.

Future research could improve accuracy by incorporating higher-resolution smoke exposure data, exploring non-linear relationships, and including additional external variables to account for broader economic trends.

Conclusion

This study explored the complex relationship between wildfire smoke and property values in Hartford, CT, revealing major insights with long-term implications. We demonstrated that smoke exposure negatively impacts property values, with effects that manifest over several years, highlighting the lasting economic consequences of environmental stressors. Key predictors like air quality and transaction volumes provides a new perspective on Hartford's market dynamics, highlighting their importance in future trends. By leveraging these findings, Hartford has an opportunity to address these challenges early, either through investments in air quality monitorin or sustainable urban planning. These forward-thinking strategiess can ensure economic stability and enhance the quality of life for Hartford's residents amidst growing wildfires.

Data Sources

1. **Wildfire datasets (1961–2021):** Information on wildfire size, duration, and proximity to Hartford.
2. **Hartford real estate sales records (2000 - 2020) :** Included property values, transaction volumes, and assessed values.
3. **US EPA's Air Quality System (AQS) API:** Captured daily air quality levels to contextualize the environmental impact.

References

BEA. (2024, September). *Connecticut real GDP by industry U.S. 2023*. Statista. https://www.statista.com/statistics/1064795/connecticut-real-gdp-by-industry/

Dong, H. (2024). Climate change and real estate markets: An empirical study of the impacts of wildfires on home values in California. *Landscape and Urban Planning*, *247*, 105062. https://doi.org/10.1016/j.landurbplan.2024.105062

IBISWorld. (2024). *IBISWorld - Industry market research, reports, and statistics*. IBISWorld. https://www.ibisworld.com/united-states/economic-profiles/connecticut/

Lopez, L. A., Tzur-Ilan, N., & Owen, J. (2024, August 27). *Rents, home values depressed in air pollution hotspots*. Dallasfed.Org. https://www.dallasfed.org/research/economics/2024/0827?