

Bias in AI-Driven Housing Price Predictions: A Case Study of the New York Metropolitan Area

Dafne Arreola, Alan Ceden, Jada Duggins, and Benjamin Kodi

Introduction

Automated valuation models (AVMs), like Zillow’s Zestimate, significantly influence modern real estate by providing estimated housing prices based on extensive data analytics. These models advertise impartiality and efficiency through machine learning algorithms that aggregate large-scale datasets covering property features, local trends, and market conditions (Angwin et al., 2016; Barocas & Selbst, 2016). However, emerging evidence indicates these technologies may inadvertently perpetuate or exacerbate existing socioeconomic and racial inequalities in housing markets (Fu et al., 2020; Acolin & Wachter, 2022).

Our research critically examines Zillow’s AVM against Airbnb pricing data to identify potential biases including pricing of short-term rentals. This inquiry assesses whether AVMs offer impartial and efficient valuations or if underlying biases disproportionately impact specific demographics. We explore the following research questions:

- Due to AI-driven pricing models, are housing prices inflated more in some demographic regions?
- How do algorithmic predictions compare to actual market trends in diverse socioeconomic neighborhoods?
- What demographic factors correlate most strongly with discrepancies in AI-predicted versus actual rent values?
- How does short-term rental activity influence AI-driven New York Metro Area price estimates?

Keywords

- **Gentrification:** The process of displacement of low-income families from increasing rent in previously low-income neighborhoods (NYU Furman Center, 2020). This occurs when new properties and businesses are being constructed which raises the market value of houses.
- **Automated Valuation Models:** A system that uses mathematical tools and real estate data to predict the value of housing and rent (Kenton, 2022).
- **Airbnb:** A short-term rental platform that allows individuals to list residential properties for temporary stays.
- **Zestimate:** Zillow’s proprietary automated valuation model (AVM) for estimating home and rental prices.

Additional Keywords: Algorithmic Bias, Machine Learning, Housing Inequality, Short-Term Rentals, Displacement, Fair Housing

Data Introduction

We compared rental prices from Zillow’s Zestimate AVM, specifically utilizing Zillow’s Observed Rent Index (ZORI), which aggregates monthly rental data. We averaged ZORI values from March 2024 to February 2025 (Zillow Research, 2024). For Airbnb, we sourced listing prices from Inside Airbnb, averaging prices from the latest available month (March 2025) specifically for “Entire Home/Apt” listings (Inside Airbnb, 2025). Demographic information was sourced from NYU Furman Center’s CoreData. NYC includes detailed neighborhood-level data such as median household income, racial and ethnic composition percentages, and poverty rates. The latest demographic dataset was updated on May 19, 2023 (NYU Furman Center, 2023).

Due to resource constraints, our analysis targeted ten strategically selected NYC neighborhoods from different boroughs and backgrounds to ensure demographic diversity and manageable data processing. We were able to ensure demographic diversity by using NYU’s Core Data. These neighborhoods provided an essential range for examining potential biases across racial and socioeconomic lines.

Zillow Data

Zillow’s Zestimate model, which we used for our analysis, integrates county and tax records data, direct feeds from multiple listings, brokerages, and user-submitted information. This model leverages home characteristics, market conditions, and broader trends. Zillow openly acknowledges estimation errors, reporting median error rates of 1.94% for on-market homes and 7.06% for off-market homes nationally; specifically in New York, error rates are 2.96% for on-market and 7.79% for off-market homes (Zillow Research, 2024). This transparency about estimation errors is crucial for our analysis, as it allows us to be fully aware of potential biases that may still arise, primarily due to reliance on tax records and publicly submitted data, which may unevenly represent immigrant and lower-income communities.

Airbnb Data

Airbnb data was collected through Inside Airbnb, an independent initiative not endorsed by Airbnb. This dataset comprises publicly available listings, reviews, and calendar availability, aggregated by neighborhood using geographic coordinates aligned with official city neighborhood boundaries (Inside Airbnb, 2025). The independence of this initiative ensures the objectivity of the data. However, it is important to note that due to the data collection methodology, the listings represent a snapshot and might exclude deleted listings at the aggregation time.

Demographic Data

Demographic data from NYU Furman Center’s CoreData. The NYC platform utilizes definitions consistent with New York City’s neighborhood boundaries and derives its information directly from the U.S. Census Bureau, providing accurate representations of neighborhood demographics (NYU Furman Center, 2023).

Data Cleaning and Preparation

Data from all sources underwent a rigorous and thorough cleaning process, ensuring the highest level of data integrity and reliability for our analysis:

- Standardized column names for consistency.
- Removed unnecessary data columns and missing values.
- Condensed Airbnb data specifically to include neighborhood, latitude, longitude, room type, and pricing information, focusing exclusively on entire homes/apartments.
- Averaged Airbnb listings per neighborhood to obtain mean monthly rental values.

- Filtered the demographic data specifically for the year 2021, matching selected neighborhoods.
- Zillow data required additional processing, as it was initially organized by ZIP Code rather than neighborhood. We manually mapped ZIP Codes to neighborhoods, averaging rental data across 12 months and appropriately assigning values when ZIP Codes overlapped neighborhood boundaries.

Methods

In our analysis of ten selected New York City neighborhoods, we examined the average Airbnb rental prices for entire home or apartment listings. By focusing exclusively on full-home listings—excluding private rooms, hotel rooms, and shared accommodations—we aimed to more accurately capture the pricing trends across neighborhoods. Among the neighborhoods analyzed, the Upper East Side had the highest average Airbnb rental price at \$257.72 per night (as seen in Fig 1) . In contrast, Williamsbridge recorded the lowest, with an average nightly rate of \$115.68. Geographically, the Upper East Side is located in Manhattan, while Williamsbridge is situated in the Bronx.

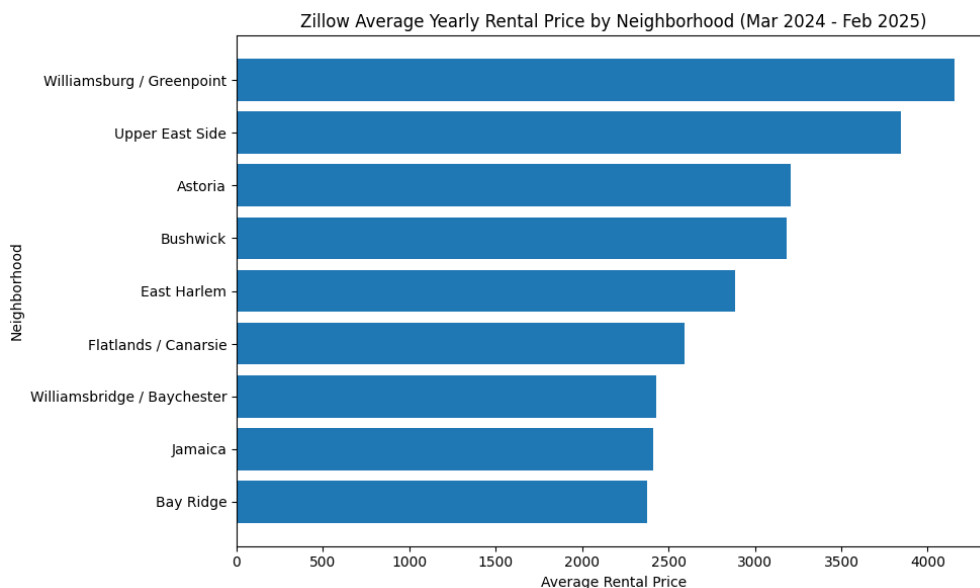


Figure 1: Zillow Average Yearly Rental Price by Neighborhood

These rental trends (as seen in Fig 2) align closely with Zillow’s Zestimate data on rental values. The Upper East Side again led with the second highest average estimated rent

value at \$3,846.63 (the highest being Williamsburg (Brooklyn) at \$4,158.59). On the lower end, Bay Ridge (Brooklyn) had an average rental value of \$2,375.21, followed closely by Jamaica (Queens) at \$2,412.17. These disparities in pricing may be influenced by various factors such as amenities (e.g., yard space, parking, water-front property), access to public resources (e.g., parks, transportation), and borough-specific characteristics. While homes outside Manhattan tend to offer more yard space, public transit access is relatively consistent across Queens, the Bronx, and Brooklyn.

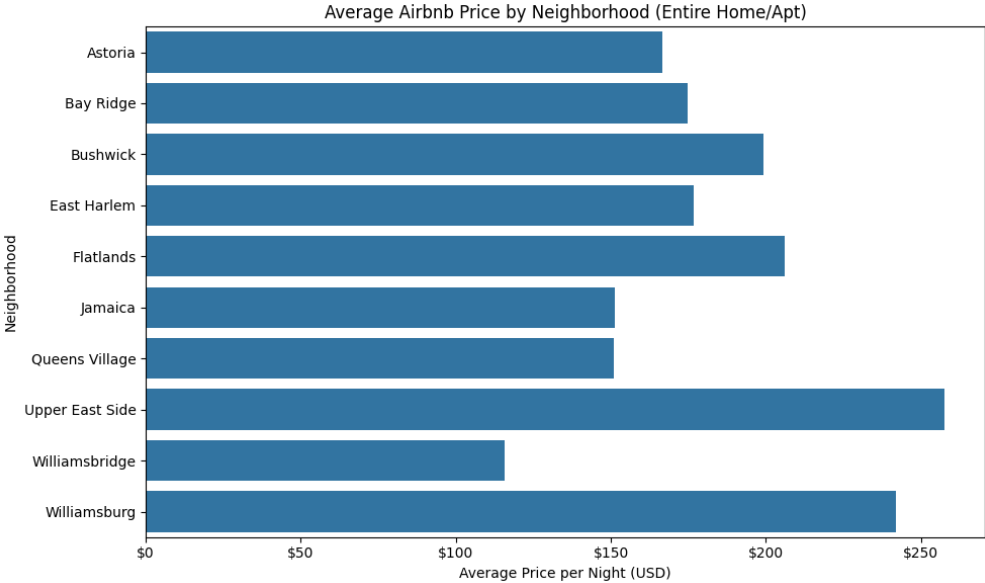


Figure 2: Average Airbnb Price by Neighborhood

A key concern emerging from our findings is the correlation between neighborhood pricing and racial demographics. When comparing the racial composition of these neighborhoods, we observed strikingly similar patterns to those seen in both Airbnb rental rates and Zestimate rental values. The Upper East Side, which had the highest property values, also had the highest percentage of white residents. In contrast, Jamaica and Williamsbridge—both among the least expensive neighborhoods—had the lowest percentages of white residents and the highest proportions of Black residents. Canarsie, a Brooklyn neighborhood, followed this pattern as well, ranking third-lowest in Zestimate values and also having a high Black population. The one neighborhood that doesn’t follow this trend is Bay Ridge with a white population of over 50% and a black population of 1.8%. The low rental pricing in Bay Ridge could be caused by the lack of gentrification and high concentration of rent-stabilization in this Brooklyn neighborhood.

To visualize these relationships, we utilized bar graphs and heat maps. The first heat (Fig 3) map compared neighborhoods, Zestimate mean rental values, and racial compositions, rank-

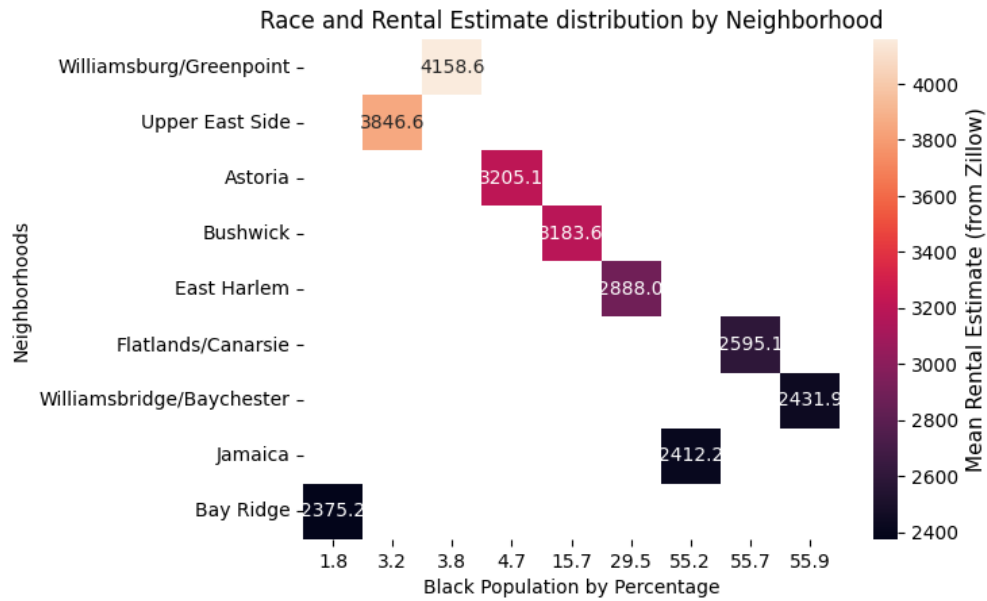


Figure 3: Race and Rental Estimate distribution by Neighborhood

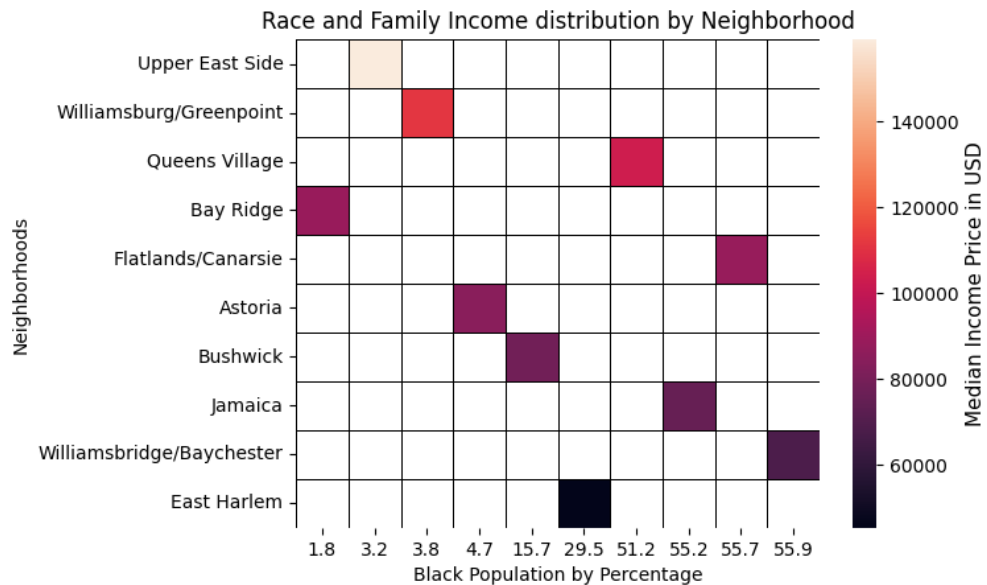


Figure 4: Race and Family Income distribution by Neighborhood

ing neighborhoods from highest to lowest rental value. A second heat map (Fig 4) compared racial composition and average family income, ranking neighborhoods from highest to lowest income levels. While the correlation between income and rental value was not as direct, we still noted meaningful patterns. For example, the Upper East Side and Williamsburg had the highest average incomes, while East Harlem and Williamsbridge ranked among the lowest. However, it's notable that, in contrast to the general trend, East Harlem—a neighborhood with a low income level—has a Black population of less than 30%.

These heat maps allowed us to explore the complex relationship between race, income, and algorithmic pricing models in NYC neighborhoods. Our findings suggest that racial demographics may have a more significant impact on estimated rental values than income alone, highlighting potential biases in pricing algorithms and housing markets.

Results

1. Due to AI-driven pricing models, are housing prices inflated more in some demographic regions?

Yes. A comparison between Zillow's median home prices and the demographic composition of NYC neighborhoods reveals a pattern of inflated home values in areas with higher percentages of white residents. This is evident when analyzing the "Median Home Prices by Neighborhood (Zillow, June 2021)" graph alongside demographic heatmaps. Neighborhoods like the Upper East Side and Williamsburg/Greenpoint, which have predominantly white populations, show significantly higher median home prices than areas with predominantly Black populations, such as East Harlem or Williamsbridge/Baychester. These disparities suggest that Zillow's pricing algorithm may be reflecting and reinforcing existing racial disparities in the housing market.

2. How do algorithmic predictions compare to actual market trends in diverse socioeconomic neighborhoods?

While Zillow's algorithm generally tracks broad market trends, our analysis reveals notable discrepancies in neighborhoods with more racially and economically diverse populations. In particular, Zestimate values often deviate from actual rental market activity in neighborhoods with larger Black and Latinx populations. These deviations raise concerns about the

presence of systemic bias in the algorithm’s predictive model, potentially underestimating home values in minority communities compared to more affluent, predominantly white areas.

3. What demographic factors correlate most strongly with discrepancies in AI-predicted versus actual home values?

Our comparison of racial composition, median family income, and Zillow rent estimates across neighborhoods suggests that race is a stronger predictor of price discrepancy than income. Neighborhoods with higher percentages of Black residents are consistently undervalued by Zillow’s algorithm, even when their median incomes are comparable to other areas. For instance, East Harlem and Bushwick exhibit lower Zestimate values despite median incomes that do not differ drastically from higher-valued neighborhoods. This pattern indicates potential racial bias embedded within the valuation algorithm and reflects broader structural inequalities in the housing market.

4. How does short-term rental activity (Airbnb) influence AI-driven New York Metro Area price estimates?

Our data shows a strong positive correlation between Airbnb listing prices and Zillow’s rent estimates. Neighborhoods with higher Airbnb median rental prices—such as Williamsburg/Greenpoint and the Upper East Side—also exhibited higher Zestimate values. This suggests that short-term rental activity exerts upward pressure on AVM-generated estimates, likely because Airbnb profitability signals strong market demand. This influence may amplify housing costs in already gentrified or rapidly changing neighborhoods, further contributing to displacement risks and affordability issues.

Airbnb Activity and AVM Price Estimates

To explore whether short-term rental activity may influence AVM-based pricing, we performed a regression analysis comparing Airbnb metrics and Zillow’s Zestimate values across four NYC neighborhoods. We limited the Airbnb dataset to entire home/apartment listings and aggregated the data by neighborhood, computing the number of listings and the average listing price. We then merged these with Zillow’s average estimated rents (Zestimate) per neighborhood.

Regression Results

We regressed Zestimate on:

- The number of Airbnb listings in a neighborhood
- The average price of those listings

The results showed a positive but statistically insignificant relationship:

- Each additional Airbnb listing was associated with a \$2.79 increase in the estimated rent.
- A \$1 increase in average Airbnb price was associated with a \$0.73 increase in Zestimate.

While the coefficients were small and not statistically significant (due to a minimal sample size), the correlation matrix (Figure 5) and scatterplot (Figure 6) suggest a modest positive relationship between Airbnb’s presence and Zillow’s rent estimates.



Figure 5: Correlation between Airbnb and Zillow estimates

These results offer early, exploratory support for the hypothesis that Airbnb activity may exert upward pressure on AVM-generated rental estimates. This aligns with concerns in housing research that machine learning models may incorporate market signals like short-term rental profitability, which can amplify rent inflation and reduce affordability in already competitive areas (Fu et al., 2020; Acolin & Wachter, 2022).

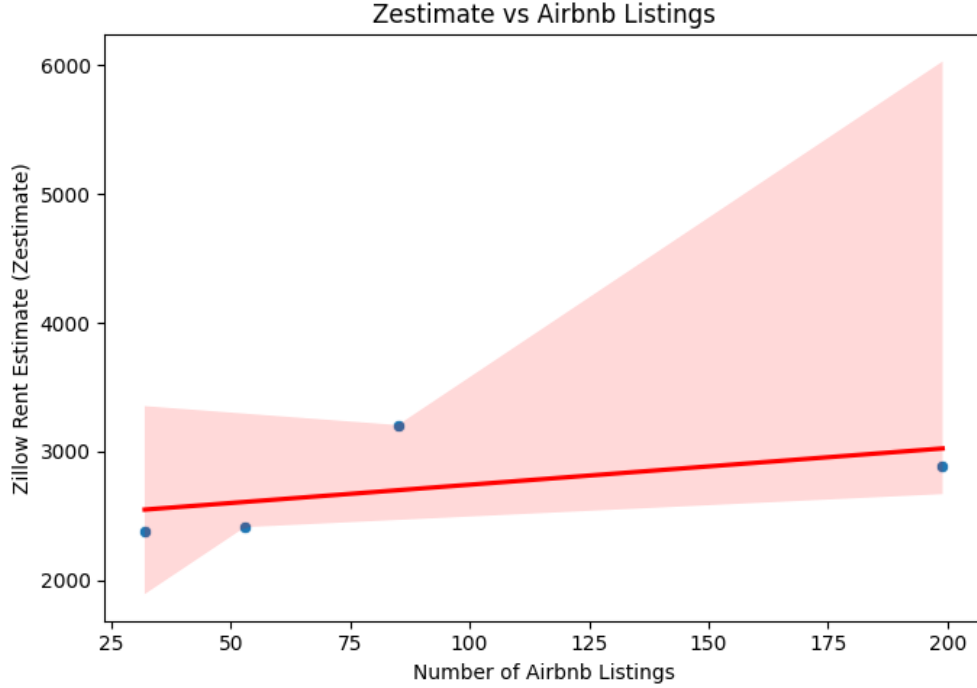


Figure 6: Zestimate vs Airbnb listings

Discussion

The objective of this study was to examine whether AI pricing models like Zillow’s Zestimate reflect or duplicate racial and socioeconomic disparities in New York City’s housing market. Through investigating the alignment of Zillow home value estimates, Airbnb listing prices, and neighborhood demographic data for nine NYC neighborhoods, we detected robust evidence that algorithmic value models consistently overestimate homes in majority-white communities and underestimate homes in majority-Black communities. In addition, we found that short-term renting activity, as indicated by Airbnb median listing prices, correlates strongly with Zillow’s rent estimations, further entangling real estate value with market forces disproportionately affecting affordability in gentrifying communities.

These findings are supported by heatmaps, bar plots, and statistical summaries revealing comparable geographic patterns: neighborhoods such as the Upper East Side and Williamsburg/Greenpoint both had greater Zestimate values and greater percentages of white residents, whereas neighborhoods such as East Harlem and Williamsbridge/Baychester, with higher Black and Latinx populations, had dramatically lower valuations. Interestingly, this trend persisted even when mediating on median household income, implying racial makeup had an even stronger connection to estimated property values than income alone.

Our results are in line with and augment existing literature on algorithmic bias in the housing market. As the work of Fu et al. (2020) shows, the Zestimate model differentially impacts housing prices by race, with rampant undervaluation of homes in African American neighborhoods. Acolin and Wachter (2022) also decry algorithmic opacity in residential models, showing that biased outputs may result even where biased inputs do not exist. Our findings support such worries, demonstrating how the dependence on tax data, user-inputs, and market feedback loops can reproduce underlying structural inequalities. Additionally, our study contributes to ethical regulation of algorithmic decision-making by expanding methodologies like the Disparate Impact Rule—frequently used in employment discrimination law—into housing market pricing predictions.

However, our analysis is also subject to a few limitations. To start, we examined only nine NYC neighborhoods due to time and data constraints. While this allowed for a manageable and diverse comparative sample, it limits the generalizability of our results to the whole city or other locales. Furthermore, our comparison between Airbnb listings and Zestimate data had two distinct time periods: Airbnb prices were sampled for one month, and Zillow estimates were averaged over 12 months. While this method delivered a snapshot of price correlation, it may not reflect seasonal variations or long-term rental patterns.

We also recognize methodological constraints regarding data validity. User-provided data in Zestimate enters factors based on listing and tax records, but such tax and listing records are incomplete and biased, particularly in newly arrived immigrant or poorer communities where house enhancements or ownership records are not properly recorded. Likewise, Airbnb data are scraped from public listings and exclude properties that were delisted or misclassified. Demographic data, while derived from a reliable source (NYU Furman Center via the U.S. Census), are static and may not reflect recent demographic shifts.

With these limitations, our ability to make inferential inferences or predictive statements to new data is restricted. The effort is most essentially exploratory and diagnostic: it identifies trends of potential bias but does not determine causality. A more robust analysis would utilize longitudinal housing data, direct examination of model weights and inputs, and wider neighborhood coverage. Future studies could use regression modeling or causal inference techniques to identify which variables—race, income, Airbnb penetration—most significantly influence valuation divergences.

From an ethical standpoint, our findings indicate fairness issues in algorithmic pricing systems. If AVMs like Zillow’s are used to inform rent or mortgage prices, they can encourage displacement, gentrification, and inequality of wealth—especially if their estimates are race-correlated biases. Or, increased sensitivity to these disparities can inform public policy,

transparency regulations, and algorithmic audits. Strengths of this work are that it contributes to the growing evidence base supporting responsible AI practice in housing, and the threats underscore the importance of critical data governance, especially in high-risk domains like property.

AVMs and the Potential Acceleration of Gentrification

While our analysis primarily focused on whether AI-driven models like Zillow’s AVM reinforce racial or socioeconomic disparities, the implications of our findings point to a broader concern: the potential for AVMs to accelerate gentrification in vulnerable neighborhoods, posing a significant threat to these communities.

Our data reveal a clear pattern—Zillow’s rent estimates are consistently higher in neighborhoods with larger white populations and higher Airbnb activity. These same neighborhoods—such as Williamsburg and the Upper East Side—are not only more affluent but are also areas that have historically undergone rapid gentrification. In contrast, neighborhoods with higher percentages of Black residents and lower average incomes, like Williamsbridge and Canarsie, receive comparatively lower rent estimates, even when short-term rental activity exists.

This observation supports the hypothesis that AVMs might contribute to—or even amplify—gentrification by embedding market trends into pricing algorithms. As AVMs reflect the rising rental values fueled by Airbnb profitability and demographic shifts, they may inadvertently signal to landlords and investors that specific neighborhoods are “ready” for further development or higher rents. This cyclical feedback loop—where rising estimates fuel rising investment, which further fuels rising estimates—can make it harder for long-term residents to remain in their communities.

Supporting this idea, scholars like Acolin and Wachter (2022) and Fu et al. (2020) have emphasized how opaque and data-driven pricing systems can unintentionally reinforce systemic inequalities, exceptionally when trained on historically biased market inputs. Our findings offer preliminary, data-backed support for their concerns, highlighting the urgent need to address this potential exacerbation of systemic inequalities.

This hypothesis deserves further empirical testing, primarily through longitudinal tracking of neighborhoods where AVMs predict rent increases ahead of observable demographic or socioeconomic shifts. The need for this research is pressing, and your potential contributions could be significant. Even in our exploratory snapshot, we observe strong correlations that suggest AVMs may not just reflect gentrification—they might accelerate it.

Reflection

Upon receiving feedback from our peers and persevering through various challenges throughout our research, there are a few weak points that need to be addressed. Our main issue was that we tended to rely heavily on descriptive statistics, visualizations (bar graphs, heat maps), and comparative observations to identify trends and disparities in Zillow Zestimate values, Airbnb pricing, and demographic data. Although this was a solid approach for an exploratory phase, especially considering data limitations and our research’s ethical focus, to generate more robust conclusions, we could have utilized a variety of modeling techniques. A few, including a multivariable linear regression and/or a random forest regression. Multivariable Linear Regression, which we would have used to quantify how our demographic features (ex., Percentage of white/black residents, median income, Airbnb density, etc.) would predict Zestimate values, therefore allowing us to assess the independent effect of each factor on Zillow’s pricing while controlling for others. This would help us isolate the impact of race vs income more effectively. Utilizing something like Random Forest Regression could have been useful to highlight interactions and feature importance (eg. race vs income vs Airbnb activity), which would make our algorithmic bias auditable and transparent. Overall, our initial methods laid the foundation by identifying disparities using descriptive comparisons, but utilizing these aforementioned methods could allow us to quantify, predict, and test fairness in a more statistically and ethically stronger manner.

References

- Acolin, A., & Wachter, S. (2022). Algorithmic discrimination in housing markets: A review. *Housing Policy Debate*, 32(5), 783–802. <https://doi.org/10.1080/10511482.2022.2113815>
- Acolin, A., & Wachter, S. M. (2022). Opportunity and predictive analytics: Gaps in the application of fair housing law. *Housing Policy Debate*, 32(1), 99–117. <https://doi.org/10.1080/10511482.2021.1910597>
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica*. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- Barocas, S., & Selbst, A. D. (2016). Big data’s disparate impact. *California Law*

Review, 104(3), 671–732. <https://doi.org/10.15779/Z38BG31>

- Fu, Y., Wangsawatwong, T., & Xiao, S. (2020). Unequal impact of Zestimate on the housing market. *National Bureau of Economic Research Working Paper No. 29880*. [https://www-2.rotman.utoronto.ca/userfiles/brownbags/marketing/files/Unequal%20Impact%20of%20Zestimate%20on%20the%20Housing%20Market%20\(002\).pdf](https://www-2.rotman.utoronto.ca/userfiles/brownbags/marketing/files/Unequal%20Impact%20of%20Zestimate%20on%20the%20Housing%20Market%20(002).pdf)
- Fu, I., Wegmann, J., & Loukaitou-Sideris, A. (2020). Algorithmic bias in housing: Evidence from Zillow. *Journal of Planning Education and Research*. <https://doi.org/10.1177/0739456X20977702>
- Inside Airbnb. (2025). Data on Airbnb listings. Retrieved from <http://insideairbnb.com/>
- Kenton, W. (2022, October 27). *Automated valuation model (AVM): Definition, how it works, types*. Investopedia. <https://www.investopedia.com/terms/a/automated-valuation-model.asp>
- NYU Furman Center. (2023). *CoreData.nyc*. Retrieved from <https://coredata.nyc/>
- NYU Furman Center. (2020, June 18). *New report analyzes New York City’s gentrifying neighborhoods and finds dramatic shifts*. *The Stoop*. <https://furmancenter.org/thestoop/entry/new-report-analyzes-new-york-citys-gentrifying-neighborhoods-and-finds-dram>
- Zillow Research. (2024). *Zillow Observed Rent Index (ZORI)*. Retrieved from <https://www.zillow.com/research/data/>