Predicting Place, Revealing Bias: GPT-4.0's Geographic Inferences from Demographics

Mary Ganey, Xinyao Ma, Haixu Tang *Luddy School of Informatics, Computing and Engineering*

Abstract

This project explores how large language models (LLMs) like GPT-4.0 infer geographic location from demographic inputs. Using approximately 100 real participant profiles drawn from prior research[1], each of which vary in race, gender, age, education, and state, GPT was prompted to guess where each person might live. Then, we compared its ZIP code predictions to actual ZIPs using U.S. Census data on race, income, and rurality. The trends found mirror prior research on LLM geographic bias [2] and suggest that models may reinforce demographic clustering, raising equity concerns for real-world applications.

These patterns indicate that GPT's location predictions may reinforce demographic clustering and reflect stereotypical associations, raising concerns about fairness and representation in AI systems. As such models are integrated into decision-making tools in healthcare, policy, and consumer platforms, understanding and mitigating geographic bias is essential for building equitable AI.

Methodology

We prompted GPT-4.0 with demographic profiles from approximately 100 real participants, including race, gender, age, education, employment, and healthcare access. Each profile followed a natural-language prompt format such as:

*"You are a **[ethnicity]** person whose gender is **[gender]** who has **[educational background]** educational background. Your current job is **[employment]**. You are between the ages of **[age]**. Your data was measured in the state of **[state]**. You visit the doctor **[doctor visits]**, and the ER **[ER visits]**. Based on your background, respond with the top three zip code locations where you might live within the **[state]**. Only respond with the top three zip code locations ranked 1 to 3 with the best guess as number 1."*

GPT returned three ranked ZIP code guesses per participant. These predicted ZIP codes were compared to each participant's actual ZIP using U.S. Census data on race/ethnicity, income, education, and rurality. Differences were measured as absolute and directional biases (e.g., predicted Black population % – actual %). To measure demographic bias, during some experiments, some demographic information (e.g. [gender], [ethnicity]) was not fed into the prompt to GPT.

Our final dataset included participants from 28 U.S. states, balanced across age, race, and education levels. Four states (CO, KS, NH, NJ) were excluded due to missing data.

Conclusion

GPT-4.0's ZIP code predictions exhibit clear biases tied to race, income, and rurality. Rather than producing neutral or demographically balanced outputs, the model often reinforced social and geographic stereotypes.

Rurality was consistently underpredicted. In states like South Dakota (-29.4%), Missouri (-19.6%), Utah (-17.5%), and New York (-15.4%), GPT showed a strong tendency to assign ZIPs that were more urban than participants' actual locations. This bias was especially pronounced for racially marginalized groups, suggesting a pattern of urban clustering in the model's outputs.

While some findings, particularly those involving Asian and Black participants, are based on smaller sample sizes and should be interpreted with care, the broader trend is clear: demographic-based ZIP code predictions by large language models can reproduce and even amplify real-world inequities.

As LLMs are adopted in critical fields like healthcare, education, and housing, geographic fairness must be a central consideration. Without transparent safeguards and demographic accountability, AI tools risk reinforcing the very disparities they claim to transcend.

Future work should explore model behavior under varied prompts, larger and more diverse samples, and mitigation strategies to reduce bias in location-based predictions.

Data and Results

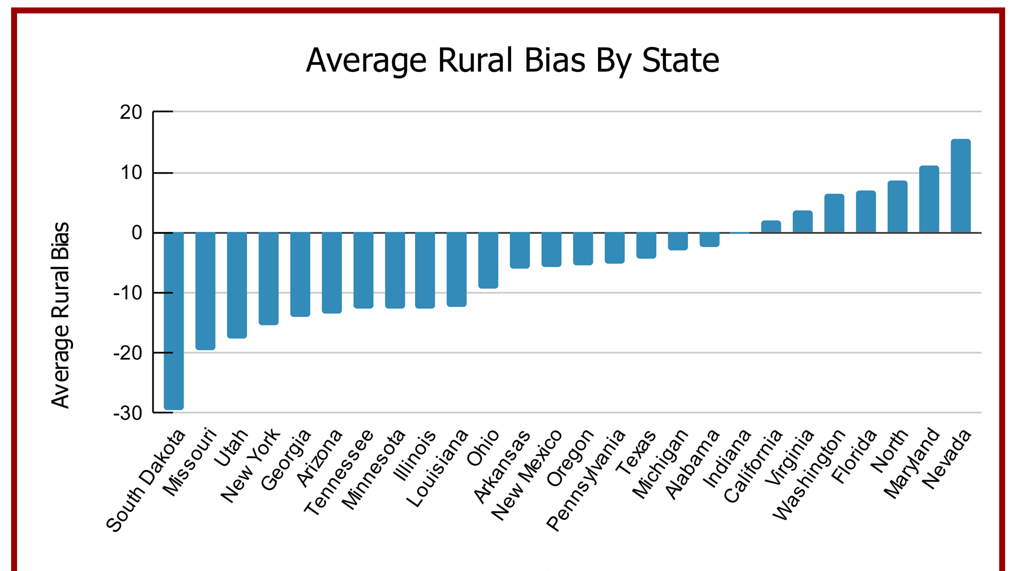

Figure 1

This chart shows average rurality difference between GPT's predicted ZIP and actual ZIP.

- Positive = assigned to more rural ZIP
- Negative = more urban ZIP than actual

Key patterns:

- Rural-heavy states like South Dakota, Missouri, Utah show strong urban skew in GPT predictions.

This suggests GPT defaults toward urban locations even for rural participants.

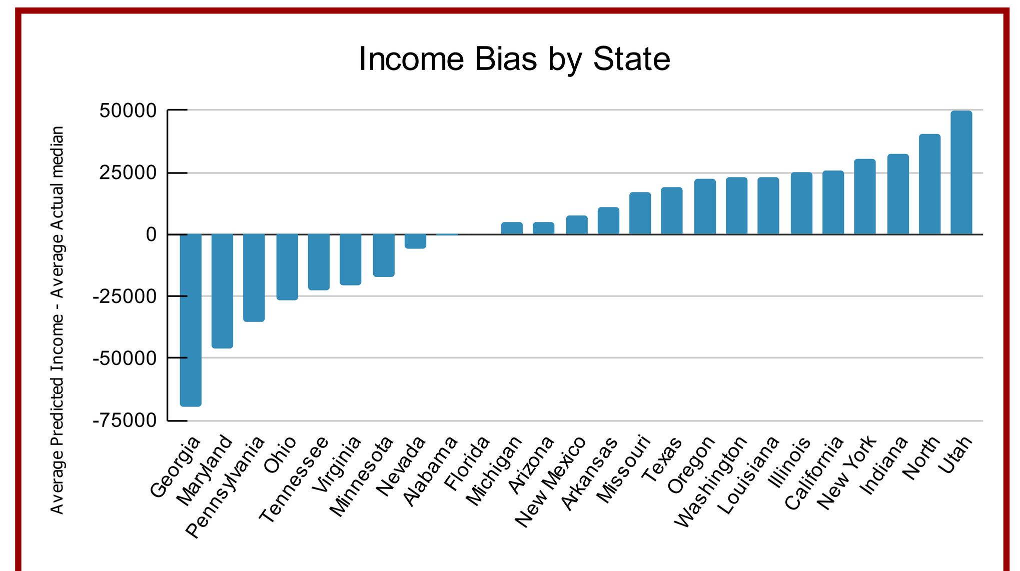

Figure 2

This chart shows the average income bias across participants in each state.

- Positive = GPT predicted ZIPs with higher income than actual
- Negative = GPT predicted ZIPs with lower income than actual

Notable patterns:

- Strong underestimation in Georgia and Maryland (negative bias).
- Overestimation in Utah, North Carolina, Indiana (positive bias).

This suggests that GPT's predictions can amplify or obscure socioeconomic reality at the state level.

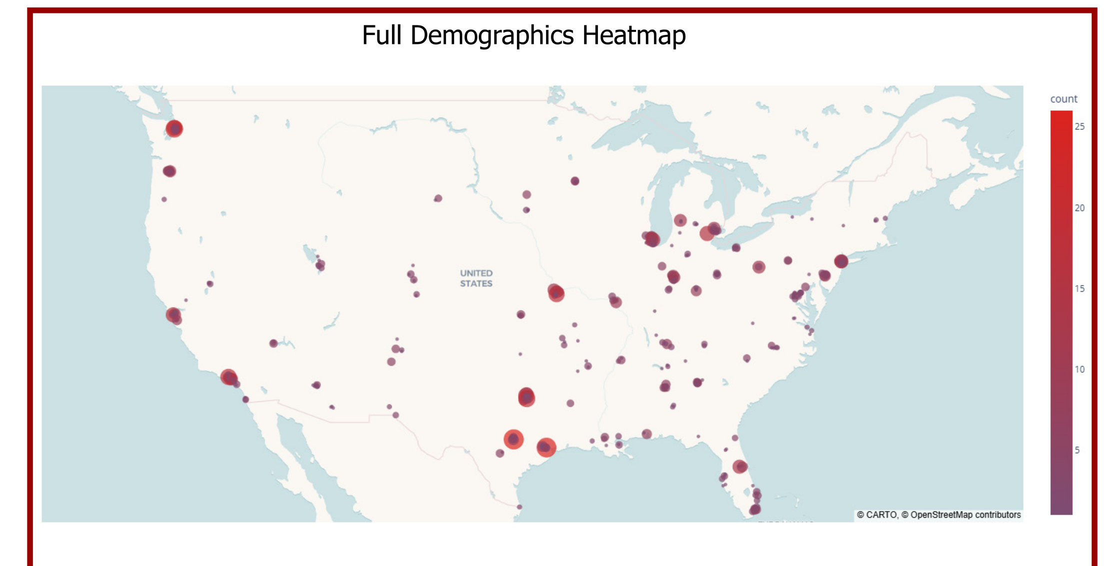

Figure 3

This heatmap groups GPT-predicted ZIPs by participants' actual state.

- Urban clustering is prominent, there was repeated neglect of rural areas in GPT predictions.

To explore variations by input trial, scan the QR code for additional heatmaps.

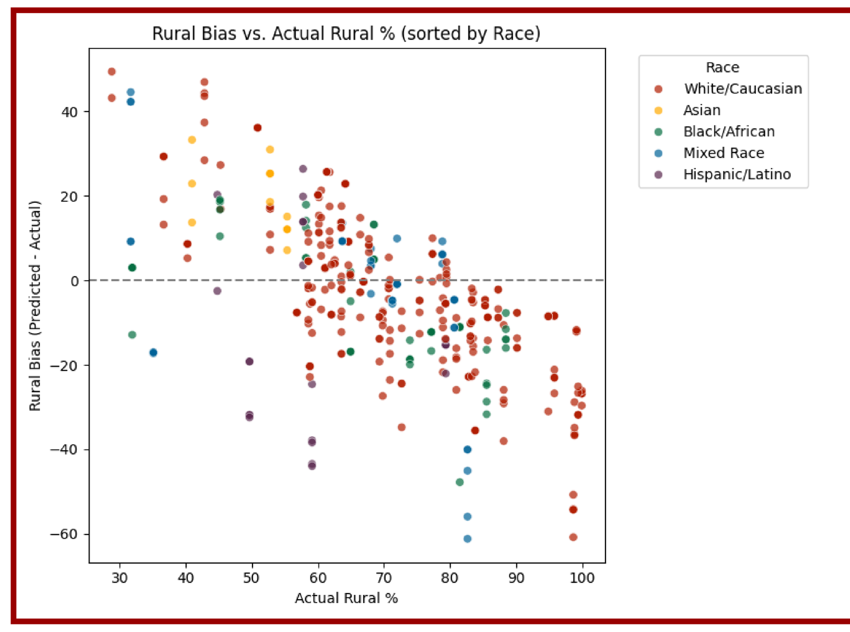

Figure 4

This rural bias scatter plot compares rurality prediction error across ethnic groups.

- GPT underpredicts rurality in high-rural ZIPs, especially for Black and White participants.

Suggests potential bias toward assigning more urban ZIPs across the board, disproportionately for certain races.

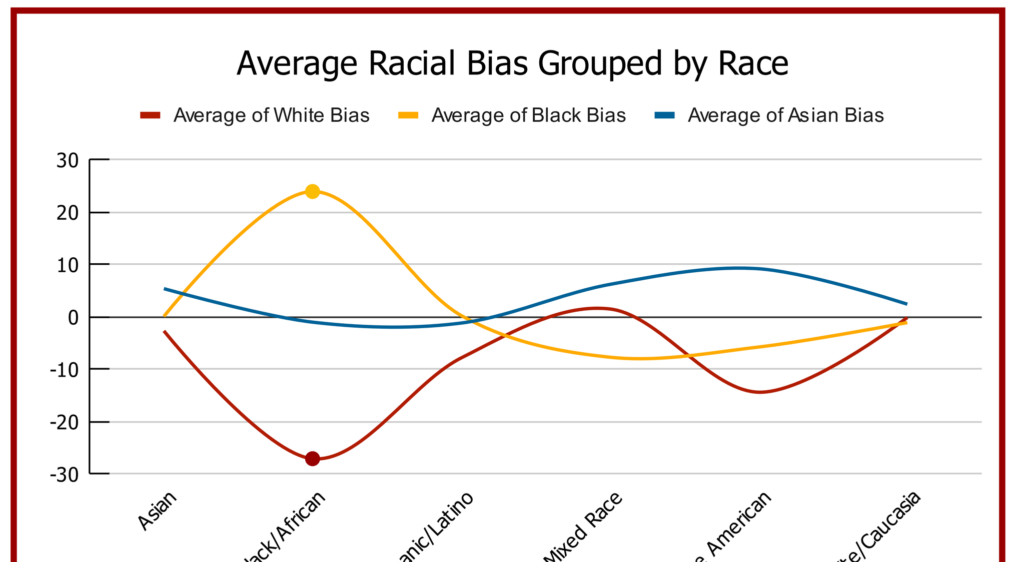

Figure 5

This chart shows average racial demographic bias in ZIP predictions.

- For Black participants, GPT overestimates Black population by nearly +24% and underestimates White population by as much as -27%.

This indicates potential racial clustering, where GPT aligns race with heavily race-skewed ZIPs.

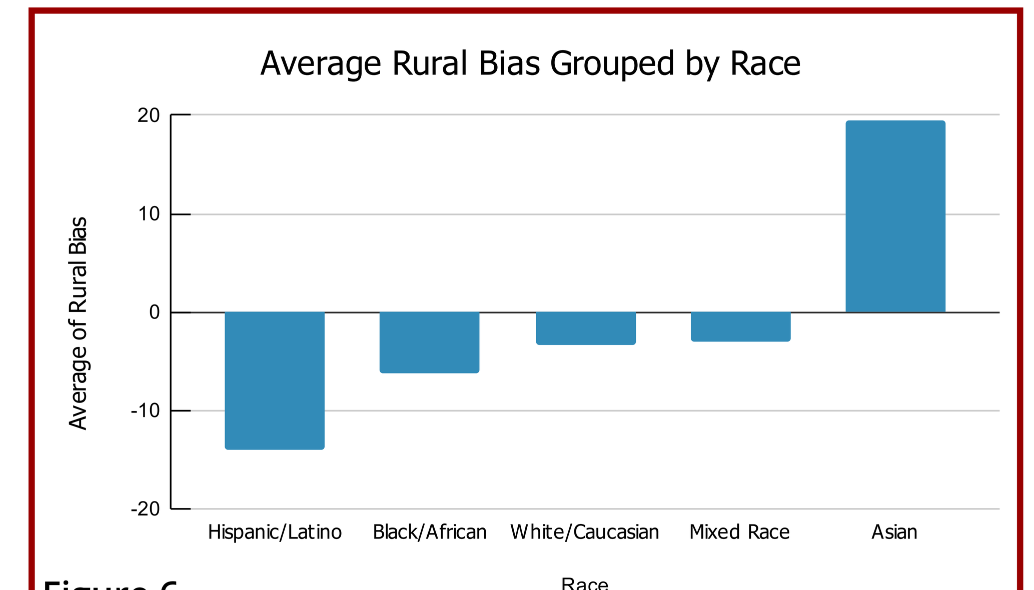

Figure 6

This chart shows rural bias grouped by ethnicity.

- Hispanic/Latino participants had the most urban-biased predictions.
- Asian participants had the most rural-biased predictions.

This may suggest racialized assumptions about rurality in model outputs.

Key Takeaways

Demographic-based ZIP predictions by GPT are not neutral.

- GPT's Predictions reflect systematic biases in socioeconomic status, race, and rurality.

Rural ZIPs are underrepresented in GPT outputs

- This, especially for participants from rural states and for Black and White individuals living in rural ZIPs.

GPT over-associates race with demographic clustering.

- Black participants were predicted to live in ZIPs with a +24% higher Black population than their actual ZIPs.

State-level income biases show large errors

- GPT underestimates income for participants in some high-income states (e.g., Georgia, Maryland)
- GPT also overestimates for participants (e.g. Indiana, Utah)

Significant differences exist in GPT predictions across race and rurality, showing bias isn't random, it's patterned.

T-Test Results

White vs Black bias: $p = 0.0012$.

- GPT's ZIPs differ significantly in racial composition when predicting for Black vs White participants.

White vs Asian bias: $p = 0.000001$

- Even stronger difference in predicted racial makeup between White and Asian participants.

Black bias vs Rural bias: $p = 0.000117$

- Significant relationship; GPT predicts ZIPs that are both more Black and more urban for Black participants.

Asian bias vs Rural bias: $p < 0.00000001$

- Extremely strong correlation; GPT tends to assign Asian participants to more rural and less racially accurate ZIPs.

***Asian and Black participant groups were underrepresented in this dataset; results may reflect sample-specific patterns.*

Acknowledgements
This research was supported in part by the National Science Foundation under CNS-2207231. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the the US Government, the National Science Foundation, nor Indiana University.

[1] X. Ma et al., "Enhancing Patient-Centric Communication: Leveraging LLMs to Simulate Patient Perspectives," Jan. 12, 2025, arXiv: arXiv:2501.06964. doi: 10.48550/arXiv.2501.06964.
[2] R. Manvi, S. Khanna, M. Burke, D. Lobell, and S. Ermon, "Large Language Models are Geographically Biased," Oct. 05, 2024, arXiv: arXiv:2402.02680. doi: 10.48550/arXiv.2402.02680.

Contact Information: Marganey@iu.edu