

2020 Post-Enumeration Survey Estimation Information Retrieval Using AI Agents

Guorong Chen PhD, Mathematical Statistician Estimation Branch, Decennial Statistical Studies Division, U.S. Census Bureau



INTRODUCTION

This poster explores using artificial intelligence (AI) to retrieve information from the U.S. Census Bureau's **2020 Post-Enumeration Survey (PES) Reports**. The project leverages multiple AI agents to extract both qualitative and quantitative information from these published reports.

A range of tailored prompts was developed to target different types of content, enabling structured information extraction and synthesis. Selected insights and observations from this exploration are highlighted in this poster.

Disclaimer

This poster does not aim to evaluate, rank, or compare the performance of the AI agents used in this demonstration.

Any views or opinions expressed on this poster are those of the author and not necessarily of the Census Bureau.

METHODS

Data: 2020 Post-Enumeration Survey Reports

Design:

- **Qualitative prompts:** Define key concepts (e.g., omissions) and retrieve insights (below).
- **Quantitative prompts:** Extract estimates and results (below).
- In addition, some open-ended prompts will be used to further explore the consistency of AI agent performance by reviewing the same prompts on two consecutive days using the same set of AI agents with memory usage turned off.

Prompts Type	Prompts
Qualitative	Define omissions for 2020 PES.
	Define the net coverage error rate for 2020 Census.
	Define whole person census imputation.
	What caused erroneous enumerations for population coverage in the 2020 Census?
	What caused erroneous enumerations for housing unit coverage in the 2020 Census?
Qualitative	How to reduce erroneous enumerations for 2030 Census.
	What was the percentage of correct enumeration in population on 2020 Census?
	What was the net coverage error rate in population on 2020 Census?
	What was the 2020 Census count?
	Is the 2020 Census net coverage error rate in person population significantly different from the 2010 one?
	How to reduce erroneous enumerations for 2030 Census.

Prompt tuning:

- Prompts ran without and with references¹⁻⁶
- Revise the prompt using the benchmark answers for the initially disagreeing prompts. For example, we first asked AI to define the net coverage error rate for the 2020 Census. We then asked AI to define the net coverage error rate for the 2020 Census in terms of Dual-System Estimation (DSE) DSE.

Analysis⁷

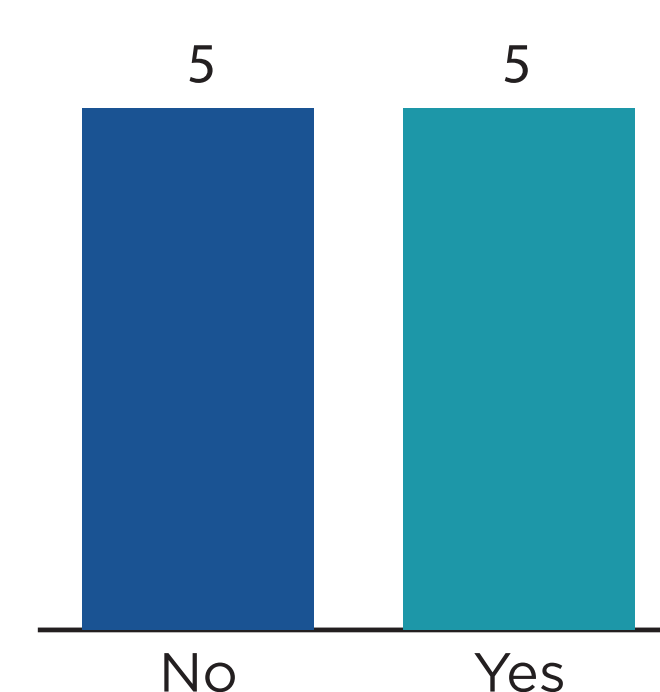
Reference-based definitions and statistics serve as benchmarks. If the AI extracted answers are the same as benchmark, AI answers are categorized as “Agree.” Otherwise, AI answers are categorized as “Disagree.”

CONCLUSION

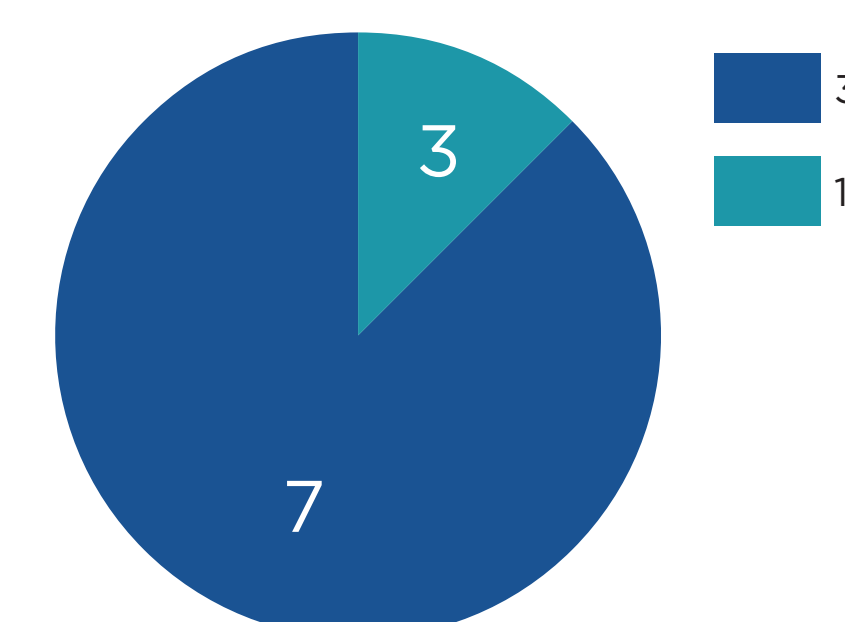
- Overall, it performed very well, but it still often required a “human in the loop,” such as prompt tuning or the use of multiple AI agents.
- It had a tendency to “hallucinate,” meaning it may generate or logically infer data that isn’t in the report, instead of clearly stating that it cannot find the information or that it does not exist in the document.
- It frequently generalized responses from other data sources that were not necessarily census data and brought into question the validity of expanding and combining disjoint conclusions as casual and comparative.

RESULTS

Number of Prompts by Prompt Tuning



Number of Prompts by Number of AI Agents Used



- After using multiple AI agents and prompt tuning, we were able to extract agreeing information for both qualitative and quantitative prompts (A,B).

- To obtain agreeing answers, 70 percent of cases required multiple AI agents (A), and 50 percent required prompt tuning (B).
- Another example of prompt tuning is that when asked to define omissions, AI did not get it right initially. When asked to define omissions in terms of DSE, it produced the correct definition. Results for this prompt, along with other prompts that yielded disagreements, are included below.

Prompts type	Qualitative prompts	Benchmark answers	AI agent's final answers	# AI agents used	Prompt tuned?	Result type
Qualitative	Define omissions for 2020 PES.	Omissions are people or housing units that should have been correctly counted in the census but were not.	It represents the portion of the P-Sample (the independent Post-Enumeration Survey sample) that does not match any correct enumeration in the census.	3	Yes	Agree
Qualitative	Define whole person census imputation.	A record is considered a whole-person census imputation if relationship, sex, age/date of birth, Hispanic origin, and race are all imputed.	A Whole Person Census Imputation is a census enumeration with no direct person-level data, created entirely through imputation, and treated in PES as a non-matchable census-only component handled outside the core matching process.	3	Yes	Agree
Quantitative	What was the 2020 Census count?	331,449,281	331,449,281 people including group quarters and remote Alaska.	3	Yes	Agree

- When asked, “What do you think the report is missing?”, the AI agent’s response sparked interesting internal discussions and introduced new perspectives.

REFERENCES

1. Hill, Courtney; Heim, Krista; Hong, Jinhee; and Phan, Nam. *2020 Post-Enumeration Survey Estimation Report: Census Coverage Estimates for People in the United States by State and Census Operations*. PES20-G-02RV. Washington, DC: U.S. Census Bureau, U.S. Government Publishing Office, June 2022.
2. James B. Lawrence and Jinhee Hong U.S. Census Bureau, *2020 Post-Enumeration Survey Estimation Report*, PES20-G-03, Census Coverage Estimates for Housing Units in the United States, U.S. Government Publishing Office, Washington, DC, August 2022.
3. Timothy Kennel, *2020 Post-Enumeration Survey Estimation: Defining Whole-Person Census Imputations*, July 21, 2023.
4. Julianne Zamora, *2020 Post-Enumeration Survey Estimation Design*, November 29, 2022.
5. Richard N. Turner, Michael Beaghen, and Courtney Hill, *2020 Post-Enumeration Survey Estimation Methods: Components of Census Coverage Methodology*, September 28, 2022.
6. Krista Heim, *2020 Post-Enumeration Survey Estimation Methods: Net Coverage Estimation*, September 21, 2022.
7. Batavia, Sakshi. *AI Agent Evaluation Methods and Best Practices*, Nurix, www.nurix.ai/blogs/ai-agent-evaluation-methods-best-practices. Accessed 22 Apr. 2026.