**1. Overview**This project investigates potential biases in LLM-generated data narratives using the same dataset under varied prompt framings. The goal is to detect and measure whether phrasing, demographic context, or hypothesis priming influences LLM interpretations. The experiment builds on previous SU research tasks in data storytelling and introduces a controlled prompt-based methodology to quantify and visualize model bias.

**2. Objectives**Identify framing, demographic, confirmation, and selection biases in LLM narratives. Evaluate consistency and fairness across multiple LLMs (GPT-4, Claude.ai, Gemini). Develop quantitative and qualitative metrics for measuring bias impact.

**3. Dataset**Dataset: Syracuse Performance Data (team statistics across game periods)Fields: Period, Goals, Saves, Shots, and Shots on Goal for Syracuse and Opponents.

This dataset provides neutral numeric data ideal for testing how prompts affect LLM interpretation of identical facts.

**4. Hypotheses**

|  |  |  |
| --- | --- | --- |
| **ID** | **Hypothesis** | **Bias Type** |
| H1 | Positive vs. negative wording changes recommendations. | Framing |
| H2 | Adding demographics alters suggested players. | Demographic |
| H3 | Primed hypotheses increase model agreement. | Confirmation |
| H4 | LLMs emphasize different stats based on phrasing. | Selection |

**5. Methodology**

**Design:**Create 3–5 prompt pairs differing in framing or context.  
Define unbiased “ground truth” using dataset statistics.

**Execution:**Query GPT-4, Claude.ai, and Gemini (3–5 runs each).  
Log all responses (JSON/CSV) with timestamps and model details.

**Analysis:**Perform sentiment, keyword, and frequency analysis using TextBlob, VADER, and scipy.stats.

Validate claims against numerical data to detect exaggerations or contradictions.

**6. Tools & Deliverables**

Category: Tools / Outputs

LLMs: GPT-4, Claude.ai, Gemini

Analysis: Python (Pandas, SciPy, TextBlob, VADER)

Visualization: Matplotlib, Seaborn

Deliverables experiment\_design.py, run\_experiment.py, analyze\_bias.py, REPORT.md, .gitignore

**7. Ethics & Compliance**

All player names and identifiers will be anonymized.

No raw datasets will be uploaded to GitHub.

.gitignore will exclude sensitive or intermediate files.

Work aligns with Syracuse University’s research ethics and OPT compliance