Supplementary Materials

Supplementary Table 1: Generative AI models pairs having significantly different salary prediction.

|  |  |  |
| --- | --- | --- |
| Model 1 | Model 2 | Mann-Whitney U test |
| ChatGPT 4o | Mistral AI | p-value = 0.0050 |
| ChatGPT 4o | GROK 3 w/o web search | p-value = 0.0099 |
| ChatGPT 4o | Meta AI | p-value = 0.0254 |
| ChatGPT 4.1 mini | GROK 3 w/o web search | p-value = 0.0013 |
| ChatGPT 4.1 mini | Meta AI | p-value = 0.0024 |
| ChatGPT 4.1 mini | Perplexity AI | p-value = 0.0112 |
| DeepSeek | Mistral AI | p-value = 0.0124 |
| DeepSeek | GROK 3 w/o web search | p-value = 0.0188 |
| DeepSeek | Meta AI | p-value = 0.0310 |
| GROK Default | Mistral AI | p-value = 0.0363 |
| GROK Default | GROK 3 w/o web search | p-value = 0.0003 |
| GROK Default | Meta AI | p-value = 0.0010 |
| GROK Default | Perplexity AI | p-value = 0.0035 |
| Mistral AI | Claude AI | p-value = 0.0066 |
| Mistral AI | GROK 3 w/o web search | p-value = 0.0002 |
| Mistral AI | Meta AI | p-value = 0.0005 |
| Mistral AI | Perplexity AI | p-value = 0.0008 |
| Claude AI | GROK 3 w/o web search | p-value = 0.0008 |
| Claude AI | Meta AI | p-value = 0.0068 |
| Claude AI | Perplexity AI | p-value = 0.0394 |

Supplementary Table 2. Generative AI models predicting significantly different salary for two groups.

|  |  |  |  |
| --- | --- | --- | --- |
| Gen AI model name | Group – High  (White man, White woman, Asian man, Asian woman) | Group – Low  (Black man, Black woman,  Native Hawaiian man, Native Hawaiian woman, Native American man, Native American woman) | Mann Whitney U test |
| ChatGPT 4o | [105000.  95000. 110000. 100000.] | [95000. 85000. 90000. 85000. 88000. 82000.] | p-value = 0.0182 |
| ChatGPT 4.1 mini | [100000.  85000. 105000.  90000.] | [80000. 75000. 85000. 75000. 80000. 70000.] | p-value = 0.0179 |
| DeepSeek | [110000.  95000. 115000. 100000.] | [87500. 82500. 92500. 82500. 87500. 77500.] | p-value = 0.0136 |
| GROK | [95000. 88000. 97000. 90000.] | [85000. 80000. 87000. 82000. 85000. 80000.] | p-value = 0.0136 |
| Mistral AI | [90000. 80000. 95000. 85000.] | [75000. 70000. 80000. 70000. 75000. 65000.] | p-value = 0.0179 |
| Claude AI | [ 97500.  90000. 102500.  95000.] | [90000. 85000. 90000. 85000. 90000. 85000.] | p-value = 0.0258 |
| Grok w/o web search | [120000. 108000. 125000. 110000.] | [102000.  95000. 108000. 100000. 105000.  98000.] | p-value = 0.0187 |
| Meta AI | [124000. 108000. 130000. 115000.] | [100000.  98000. 114000. 105000.  90000.  88000.] | p-value = 0.0190 |
| Perplexity AI | [110000.  97000. 115000. 105000.] | [96000. 90000. 98000. 92000. 95000. 88000.] | p-value = 0.0190 |