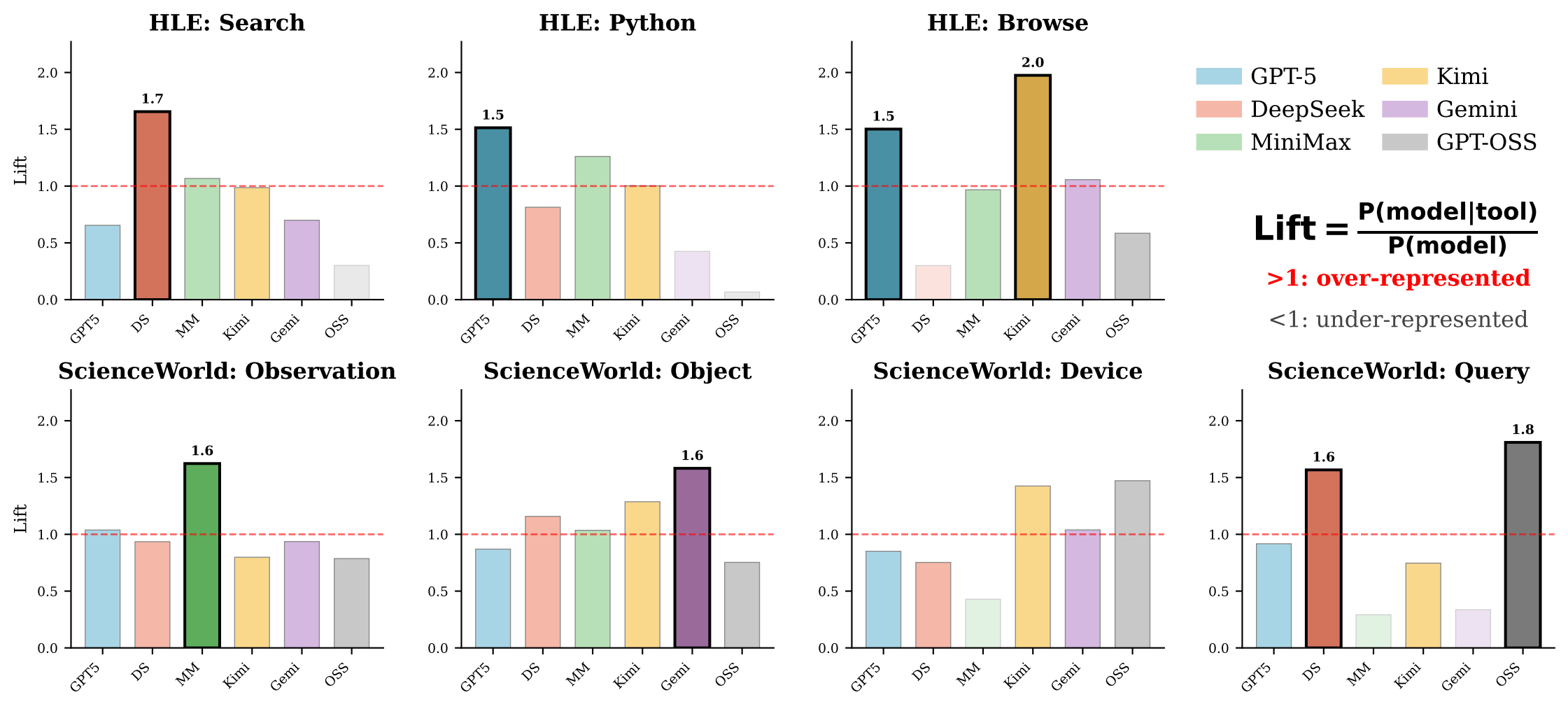
HLE: Search
HLE: Python
HLE: Browse
ScienceWorld: Observation
ScienceWorld: Object
ScienceWorld: Device
ScienceWorld: Query
Lift
GPT5
DS
MM
Kimi
Gemi
OSS
1.7
1.5
1.5
2.0
1.6
1.6
1.8
GPT-5
DeepSeek
MiniMax
Kimi
Gemini
GPT-OSS
Lift = P(model|tool) / P(model)
>1: over-represented
<1: under-represented