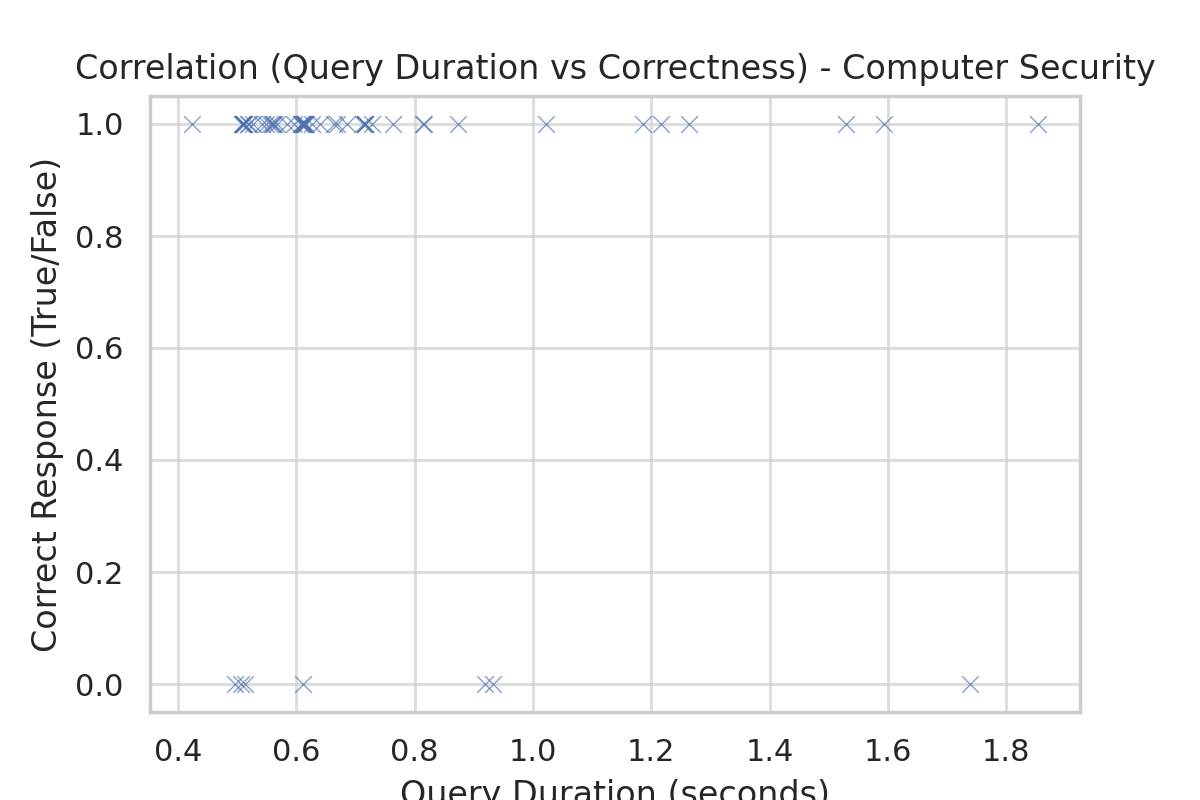
# ChatGPT Evaluation Report

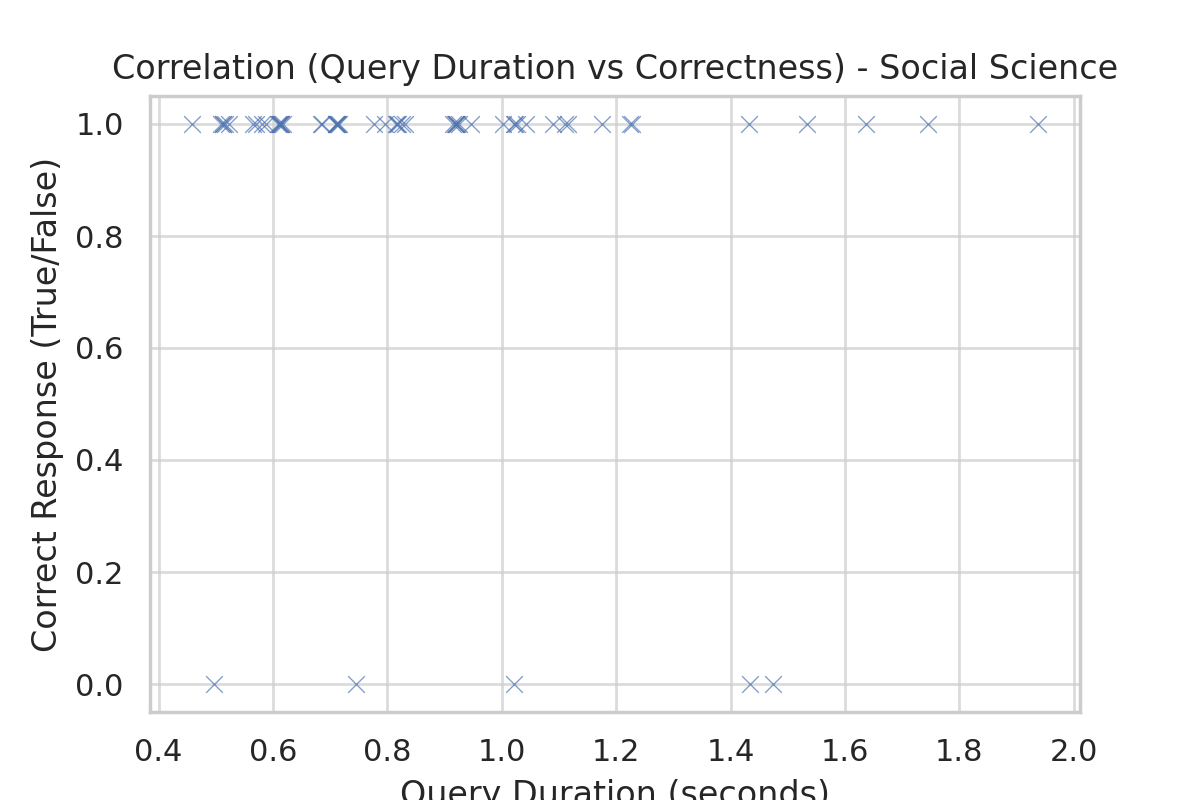
## Computer Security

1. Accuracy Rate: The accuracy rate of ChatGPT's responses compared to the expected answers is approximately 88.33%.
2. Correlation Analysis:
   * There is a weak negative correlation (−0.106) between the query duration and the correctness of ChatGPT's responses. This suggests that there is a slight tendency for longer query durations to be associated with incorrect responses, but this correlation is relatively weak and may not be statistically significant.



## Social Science

1. Accuracy Rate: The accuracy rate of ChatGPT's responses compared to the expected answers is approximately 91.67%.
2. Correlation Analysis:
   * There is a weak negative correlation (−0.138) between the query duration and the correctness of ChatGPT's responses. This suggests that there is a slight tendency for longer query durations to be associated with incorrect responses, but this correlation is relatively weak and may not be statistically significant.

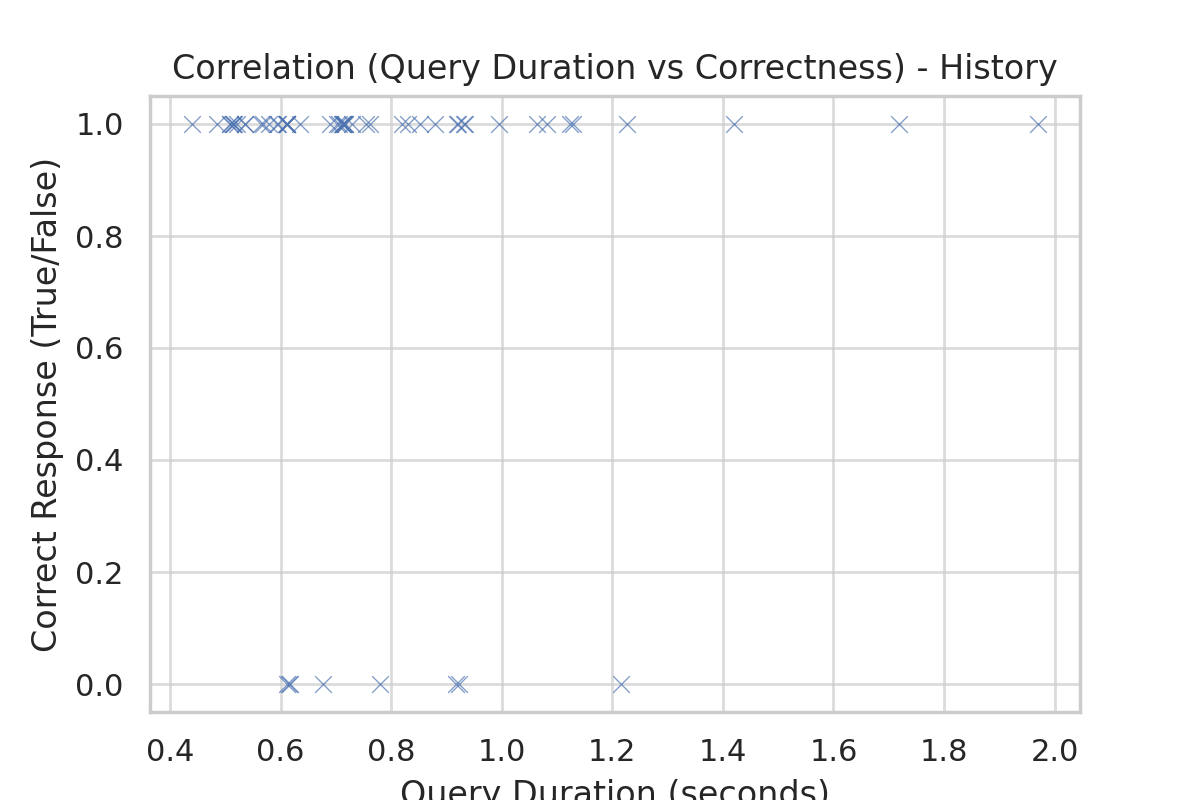


## History

1. Accuracy Rate: The accuracy rate of ChatGPT's responses compared to the expected answers is approximately 86.67%.
2. The average query durations for each domain:

* History: 0.78 seconds.
* Social Science: 0.88 seconds.
* Computer Security: The average query duration is approximately 0.73 seconds.

1. Correlation Analysis:
   * There is a weak negative correlation (−0.019) between the query duration and the correctness of ChatGPT's responses. This suggests that there is a slight tendency for longer query durations to be associated with incorrect responses, but this correlation is relatively weak and may not be statistically significant.



**Conclusion**

The analysis of ChatGPT's performance on the History, Social Science, and Computer Security datasets provides valuable insights:

1. **Accuracy Rates:**
   * History: 86.67%
   * Social Science: 91.67%
   * Computer Security: 88.33%

These figures indicate a high level of accuracy across all three domains, with Social Science achieving the highest accuracy rate. This suggests that ChatGPT has a strong ability to provide correct answers in these varied fields, demonstrating versatility and reliability in understanding and responding to multiple-choice questions.

1. **Correlation Between Query Duration and Correctness:**
   * All three datasets showed a very weak negative correlation between the query duration and the correctness of responses.
   * This implies that longer query durations do not significantly impact the likelihood of an incorrect response. In other words, the time taken to generate a response is not a strong indicator of its correctness.
2. **Implications:**
   * The high accuracy rates across different subject’s underline ChatGPT's effectiveness in handling a wide range of topics, making it a useful tool for educational and informational purposes.
   * The minimal impact of query duration on correctness suggests that response time should not be a major concern when using ChatGPT for answering questions in these areas.

In conclusion, ChatGPT demonstrates high effectiveness in providing accurate responses across diverse subjects with no significant correlation between the response time and the accuracy of its answers. This makes it a reliable tool for educational and informational queries in these domains.

Top of Form

Bottom of Form