

Does the inclusion of code in a StackOverflow question increase the solvability the post?

Sarah Telford

Science, Engineering and
Technology (SET)

The University of Adelaide

Adelaide, SA, Australia

a1810750@student.adelaide.edu.au

CCS CONCEPTS

- StackOverflow → Format
- StackOverflow → Solution
- StackOverflow → Questions

KEYWORDS

StackOverflow, code segments, question format, Python, response time

ACM Reference format:

Sarah Telford. 2022. Does the inclusion of code in a Stack Overflow question increase the solvability of a post?. In *Proceedings of ACM Woodstock conference (WOODSTOCK'18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/1234567890>

Abstract

Given the popularity and importance of online platforms, such as StackOverflow, within the programming community users need to ensure that their questions are being seen and resolved by other users. One common feature in countless StackOverflow questions posted is the inclusion of code segments in a question. However, I want to know if the inclusion of code in a StackOverflow

question increases the solvability of the post? In this paper, I am determined to find if there is a causal relationship between solvability and questions containing code. Data has been gathered from Python-related questions to generate percentage calculations and graphs. Other research papers were analysed to consolidate any findings and provided stronger evidence for results. The key finding of this study is that including code segments in a question is beneficial to users as it will increase the likelihood of questions receiving a solution. Furthermore, the response time of answers is increased on questions that have code.

1 Introduction

StackOverflow is “a public platform building the definitive collection of coding questions & answers”¹. According to StackOverflows insights, “about 100+ million people visit” the site each month, most of whom are professional or recreational programmers². Programmers and developers use StackOverflow to ask questions and find solutions to issues and errors encountered while developing programs and code. Currently, StackOverflow reports “6,826,036 questions with no upvoted or accepted answers” posted to the site, with a new question being asked every 13.6 seconds¹.

With such a saturated platform, users need to ensure that their posts are concise and easily understood. By looking at the impact of including code segments in a question on the likelihood of receiving a useful response, users will be able to discern whether they should include code in their own questions.

2 Motivation

StackOverflow provides users with some “basic tips” to “improve your chances of getting an answer”³. Included in these tips are suggestions for introducing the issue before posting any code, as well as suggested formats for posting code. The page mentions that the inclusion of code segments in a question improves the likeliness of getting an answer, however, it does provide any data about the success rates.

Most of the users of StackOverflow are new to programming. According to the 2021 developer survey conducted by StackOverflow, 50% of users have less than a decade of experience, with 35% reporting less than five years⁴. Since the majority of users are beginner programmers, it is important that they are utilising these online platforms to receive useful advice and solutions. These insights also report that “almost 60% of respondents learned how to code from online resources” and that “younger respondents tend to learn from online courses, forums, and other online resources”⁴.

3 Background

Research suggests that the usefulness of posted code segments gravely impacts the success of questions. A 2021 study by Mondal, et al. from the University of Cornell states that “it is not always possible to reproduce the issues from the code segments that may impede questions from receiving prompt and appropriate solutions”⁵. Furthermore, “a

reproducible question has at least two times higher chance of receiving an accepted answer than an irreproducible question”⁵. Therefore, it is vital that included code segments are easily reproducible for questions to be successfully answered.

Further research suggests the explanations given by users providing solutions is another issue. According to Silva, et al., solutions “containing both code examples and their explanations”⁶ are harder to find. Additionally, solutions containing both explanations and example code are sometimes unclear and hard to follow.

Asaduzzaman, et al. conducted a study in 2012 into “why questions remain unanswered”. The study aims to understand why questions remain unanswered. In turn, users will be able to improve the quality of their questions, thus increasing their chances of getting answers and their ability to use StackOverflow's services effectively. This study mined data from StackOverflow about unanswered questions. To find out the most frequent cause of questions remaining unanswered they took 400 randomly selected unanswered questions from Stack Overflow. These were then analysed and categorized by characteristics. Table 1 depicts the percentage of frequency for the top five characteristics of unanswered questions⁷.

Percent	Characteristic
21.75	Fails to attract an expert member
17.0	Too short, unclear, vague or hard to follow
12.0	A duplicate question
11.75	Impatient, irregular or inconsiderate members
9.0	Too hard, too specific or too time consuming

Table 1: Top five characteristics of unanswered questions⁷

4 Research Method

4a Research Questions

This study aims to determine if the inclusion of code in a StackOverflow question increases the solvability of a post. To find a solution to this question, the following sub-questions will need to be answered:

What percentage of StackOverflow questions receive an answer?

What percentage of approved solutions have code included in the question?

What percentage of unanswered questions on StackOverflow contain code?

Are questions that contain code segments answered faster than those that don't?

What causes StackOverflow questions to remain unanswered?

Does the length of code segments impact the success rate of questions receiving responses?

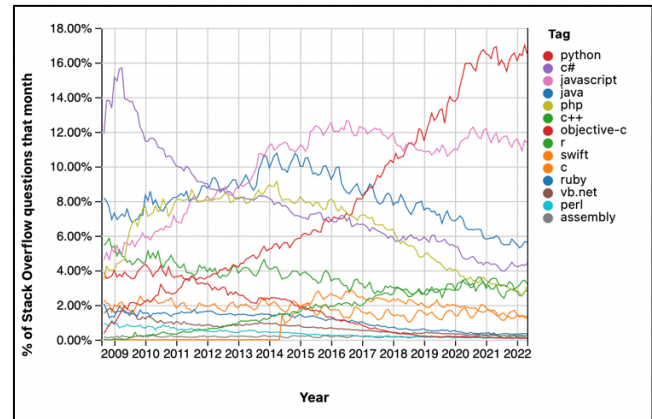
4b Data Collection

StackOverflow's annual developer surveys were used as the main source for data collection. Annually, the developer surveys "examine all aspects of the developer experience ... on open-source software"⁸. The survey receives "nearly 80,000 responses fielded from over 180 countries and dependent territories"⁸.

Due to the large number of posts made to StackOverflow, the sample pool was restricted to Python programming posts. Python has been selected as the focus language because it is one of the

most popular programming languages used by StackOverflow developers. According to a 2020 survey conducted by StackOverflow, 44.1% of users program using Python².

StackOverflow trends depict the frequency of questions relating to different programming languages.



Graph 1: StackOverflow Trends⁹

This trend graph depicts Python as the most common language asked about on the site, with the frequency of questions continually growing each year.

StackOverflow insights provide detailed data about users, posts, programming languages, and formats. These insights were used to gather data about Python-related posts and the rate of success in receiving an acceptable answer on questions containing and not containing code segments. Data was also gathered regarding what causes questions to be left unanswered and what makes a successful question.

Data regarding whether a submitted question contains code segments were gathered from both unanswered and answered Python questions. Each category had 50 random samples selected from the question pools, totaling 100 samples in total. These samples have been compared to determine if there is

a correlation between questions containing code and a solution.

The sample pool was taken from the following filters under that Python tag.

- Newest
- Recent activity
- Highest score
- Most frequent

Data has been taken from each filter to ensure that the sample accurately represents the sample.

Due to limited time and resources, other studies' data was used in the generation of findings. These studies were selected because they used large sample sizes. Both studies analysed data collected by StackOverflow between July 31, 2008, and July 31. "Answering Questions about Unanswered Questions of StackOverflow" by Asaduzzaman, et al. randomly sampled 400 unanswered questions, selecting 100 questions from each of the four years for which they collected data⁷. Whereas Min(e)d Your Tags: Analysis of Question Response Time in StackOverflow by Bhat, et al. analysed the whole dataset.

4c Data Analysis

The data collected was used in a range of different statistical comparisons. A random sample of 100 questions was taken from the StackOverflow Python tags and analysed to generate statistics about the population. The sample was data categorised into two groups:

- Answered question
- Unanswered questions

With each category containing 50 questions.

The two categories were analysed, and questions were sorted into the subcategories:

- Contains code segments
- Does not contain code segments

Next, percentages were generated for each of the subcategories:

- answered and contained code segments (Appendix A, Table 1)
- answered and not contained code segments (Appendix A, Table 1)
- unanswered and contained code segments (Appendix A, Table 2)
- unanswered and not contained code segments (Appendix A, Table 2)

A comparison between the percentages for each category was conducted to determine if there is a causal relationship between questions solved and questions containing code segments. Using the percentages generated in appendix A, table 1, and appendix A, table 2, pie charts were created to illustrate the difference between groups.

These percentages were applied to the population to find the total number of questions that would fall in each category. This was achieved through a simple percentage calculation, where the total number of questions was multiplied by the percentage.

When calculating timeframes for questions being answered, 10 samples were taken from both the questions containing code and the questions not containing code (Appendix B). The timeline of each of the 20 samples were analysed to find when each question was asked and answered. The difference in time was calculated by subtracting the asked time from the response time. During these calculations, caution was taken to account for time measurements. A dot plot was created to display this data and the trendline for each category was generated to illustrate the averages.

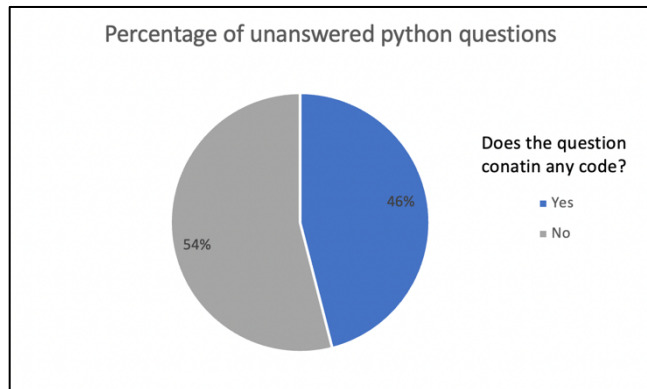
5 Findings

5a What percentage of StackOverflow questions receive an answer?

According to Stack Exchange, ~70% of questions posted receive an answer. It is assumed that all received answers have been approved because if a response is not useful, users can dispute the answer and mark the question as unsolved¹⁰.

5b What percentage of approved solutions have code included in the question?

The Python tag currently has 181,103 questions that have received answers. Graph 2 depicts the percentages of questions that do and do not contain code segments.



Graph 2: Percentage of unanswered Python questions

Assuming the sample is representative of the population, we used the percentages in appendix A, table 1 to calculate the total number of answered questions that do and do not contain code segments.

The calculation for questions containing code:

$$0.74 * 181\,103 = 134\,016$$

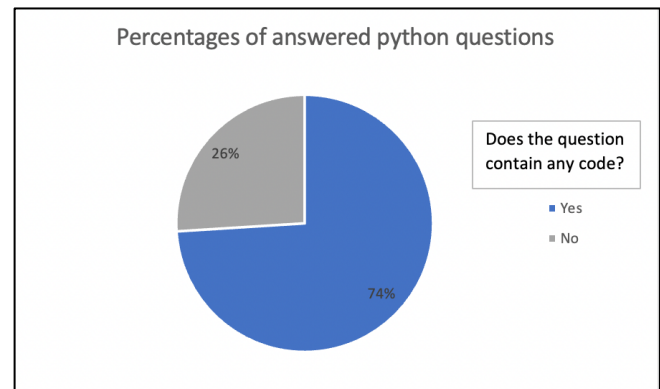
The calculation for questions not containing code:

$$0.26 * 181\,103 = 47\,086$$

This calculation shows that 134 016 of the answered Python questions contain code segments. Whilst only 47 086 of the answered questions do not have code in the question.

5c What percentage of unanswered questions on StackOverflow contain code?

Under the Python tag, there are currently 278,706 questions that have received no answers. The below pie chart illustrates the percentages for questions that do and do not contain code segments.



Graph 3: Percentage of unanswered Python questions

Assuming the population has a similar divide, using the percentages calculated from the sample, appendix A table 2, we can calculate the total number of unanswered questions that do and do not contain code segments.

The calculation for posts containing code:

$$0.46 * 278\,706 = 128\,204$$

The calculation for posts not containing code:

$$0.54 * 278\,706 = 150\,501$$

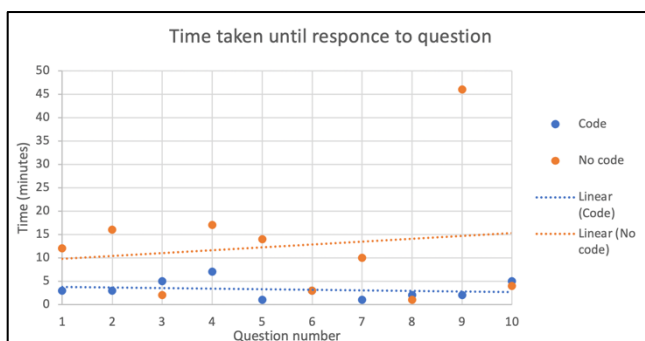
These calculations inform us that 128 204 of the unanswered Python questions contain code segments, while 150 501 of them do not.

5d Comparison of 5b and 5c

Comparing the percentages illustrated in graph 2 and graph 3 we can discern that there is a causal relationship between the inclusion of code segments and receiving a solution. This is evident by the fact that there is a smaller percentage of unanswered (46%) and a larger percentage of answered (74%) questions containing code. This infers that if users included code in their questions, they are more likely to receive a response.

5e Are questions that contain code segments answered faster than those that don't?

The data collected shows that the average time taken to respond to questions containing code is shorter than those that do not. This is clearly illustrated in the below graph.



Graph 4: Time taken until response to a question is received

Other studies found that one of the major contributing factors to a question being answered was the response time received on a question. Bhat, et al. found that 50% of the questions receive a response within 15 minutes of being posted¹¹. When comparing the results in graph 4, we can see solutions were found in this timeframe for all the questions that contain code. However, 30% of those that did have code not were outside of this range. The timeframe for receiving a response is important because with such a saturated platform new questions will quickly overshadow existing ones.

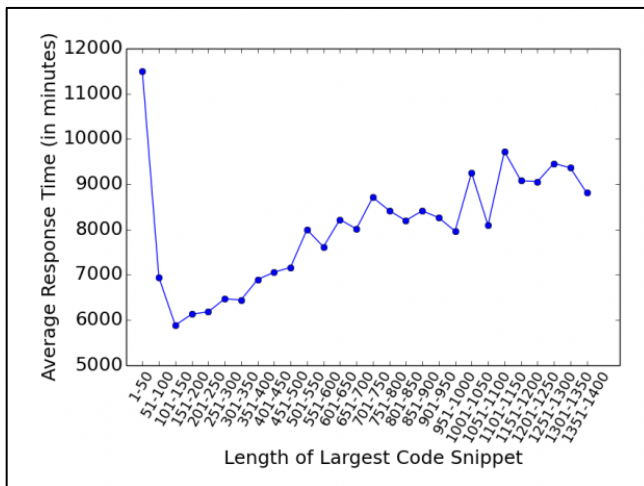
5f What causes StackOverflow questions to remain unanswered?

Several factors that can contribute to a StackOverflow question remaining unanswered. Asaduzzaman, et al. found that there were 13 main reasons for a question remaining unanswered.

1. The question is too short, unclear, vague, or hard to follow.
2. The question is program-specific without a program snippet or proper explanation.
3. The question is too hard, specific, or time-consuming.
4. The question is related to proprietary technology
5. Impatient, irregular, or inconsiderate users.
6. Not a question for StackOverflow
7. User answered their own question
8. The question does not have any answer
9. The question is a duplicate
10. Answer no longer relevant or needed.
11. The question fails to attract an expert member
12. The bug is not reproducible
13. The question is related to course projects or homework.

5g Does the length of code segments impact the success rate of questions receiving responses?

According to Bhat, et al. questions with code snippets are likely to receive a faster response than those without. However, this is conditional to the length of the code. Bhat found that the ideal length for code segments should be between 50 and 150 characters. If the code is too short or increasingly long, then the response time tends to increase. This change in response time is illustrated in the below graph.¹¹



Graph 5: The size of code snippets in questions vs response time¹¹

6 Discussion

It is evident from the findings that there is a positive causal relationship between questions containing code and receiving responses. Regardless, other factors such as tags, the length of the question, programming language, and upvotes contribute to the probability of a response. Furthermore, users that included code segments in questions need to ensure that they are concise and easy to follow, while still containing all the necessary code. As complicated and long code usually results in a question remaining unsolved. Overall, one can

conclude that including code segments in your question will increase the likelihood of receiving a response, however, it does not guarantee one.

7 Threats to Validity

Due to limited time and resources, the sample size of this study was 100 questions, which is a small sample size. Consequently, the data may not accurately represent the population, in turn, impacting the validity of the results since assumptions have been made based on the results from the sample.

Furthermore, the small sample size allowed outlier data to significantly impact the results. This is evident in graph 4, where one of the data points for “no-code” was higher than the others in this category. Since the questions were randomly selected the outlier could not be omitted from the dataset. Furthermore, purposefully omitting data would decrease the validity of the study.

To limit bias when selecting data there were procedures undertaken. These included random questions selection and questions selected from different filters. However, this does not prevent subconscious bias in the selection process.

Additionally, by limiting the sample pool to Python-related questions the sample cannot be generalized to the whole population of StackOverflow. Instead, the data can only be applied to other Python questions. Additionally, there are discrepancies with the findings data. Whilst the data collected in this study was limited to Python questions the data in the other research papers did not have this restriction. Therefore, their data can be generalized to the whole population. These discrepancies caused assumptions to be applied to populations that are not being correctly represented.

Overall, when making assumptions based on the sample, caution must be taken to ensure that the data is being applied correctly to the population.

8 Related Works

The research conducted in this paper supports the various studies mentioned throughout the report. The response times calculated in section 5e corroborate the findings of Bhat, et al. since over 50% of questions received a response in less than 15 minutes.

Additionally, the percentages of answered Python questions also validates StackOverflows claim that ~70% of questions receive a response since 74% of the sample had been successfully resolved.

9 Conclusion

The inclusion of code in a StackOverflow question increases its solvability. Data collected from StackOverflow was analysed to determine the extent of this correlation and to identify any contributing evidential factors. I conjecture that if users want to maximise the utility of StackOverflow and receive insightful responses they need to include code segments in their questions.

10 Future Works

Future work can easily be conducted to expand the findings of this paper. Researchers who wish to continue this study should:

- *Increase the sample pool size:* The sample pool used in this study is too small and is easily subject to outlier data influence. Therefore, future work should increase the sample size.
- *Expand on the sample pool:* The sample pool used to generate percentages and timeframes was limited to Python-related questions.

However, the study is intended to be applied to all StackOverflow questions. Therefore, any future work should expand on the question pool to ensure questions are taken from all tags.

- *Conduct a survey:* A survey is an easy way to gather insights from a community. Participants can be asked a series of questions relating to their use of StackOverflow and the format of posts on the platform. This survey can be distributed online so that participants can be from anywhere.

Appendices

Appendix A: Percentages Tables

Answered Python Questions	
Does the question contain code?	Percentage
Yes	74%
No	26%

Table 1: Percentages (%) calculated for answered Python questions

Unanswered Python Questions	
Does the question contain code?	Percentage
Yes	46%
No	54%

Table 2: Percentages (%) calculated for unanswered Python questions

Appendix A:

Title	Code?	asked	answered	Time (mins)
Does Python have a ternary conditional operator?	no	8:32	8:44	12
What are metaclasses in Python?	no	6:10	6:26	16
How do I execute a program or call a system command?	no	1:35	1:37	2
I need a good explanation (references are a plus) on Python slicing.	no	22:31	22:48	17
Using global variables in a function	no	5:45	5:59	14
How do I get the current time?	no	4:54	4:57	3
How do I sort a dictionary by value?	no	12:49	12:59	10
What is the difference between <code>__str__</code> and <code>__repr__</code> ?	no	4:27	4:28	1
Compute a confidence interval from sample data	no	21:29	22:15	46
Regex: get all numeric and special	no	4:17	4:21	4

characters starting with specific letters, stop when space occurs after number				
What does the "yield" keyword do?	yes	22:21	22:24	3
What does if <code>__name__ == "__main__":</code> do?	yes	4:11	4:14	3
How do I merge two dictionaries in a single expression (take union of dictionaries)?	yes	7:44	7:49	5
Does Python have a string 'contains' substring method?	yes	2:52	2:59	7
Catch multiple exceptions in one line (except block)	yes	2:55	2:56	1
How to iterate over rows in a DataFrame in Pandas	yes	7:04	7:07	3
Finding the index of an item in a list	yes	1:39	1:40	1
Iterating over dictionaries using 'for' loops	yes	22:27	22:29	2
Accessing the index in 'for' loops	yes	22:47	22:49	2
How do I make a flat list out of a list of lists?	yes	20:30	20:35	5

Table 1: Timeframe (min) calculations for answered Python questions

REFERENCES

- [1] (2008 Accessed: 2022-4-18). StackOverflow.
- [2] Stack Overflow Developer Survey 2020. (2022). Retrieved 5 April 2022, from <https://insights.stackoverflow.com/survey/2020>
- [3] How do I ask a good question? - Help Center. (2022). Retrieved 5 April 2022, from <https://stackoverflow.com/help/how-to-ask>
- [4] Stack Overflow Developer Survey 2021. (2022). Retrieved 5 April 2022, from <https://insights.stackoverflow.com/survey/2021>
- [5] Mondal, S., Rahman, M., Roy, C., & Schneider, K.. (2021). The Reproducibility of Programming-Related Issues in Stack Overflow Questions.
- [6] Silva, R., Rahman, M., Dantas, C., Roy, C., Khomh, F., & Maia, M.. (2021). Improved Retrieval of Programming Solutions With Code Examples Using a Multi-featured Score.
- [7] Asaduzzaman, M., Mashiyat, A., Roy, C. & Schneider, K. (2013) Answering Questions about Unanswered Questions of Stack Overflow
- [8] Stack Overflow Annual Developer Survey. (2022). Retrieved 10 June 2022, from <https://insights.stackoverflow.com/survey>
- [9] (2022 Accessed: 2022-5-14). StackOverflow Trends.
- [10] (2022 Accessed: 2022-5-16). Stack Exchange.
- [11] Bhat, V., Gokhale, A., Jadhav, R., Pudipeddi, J., & Akoglu, L. (ND). Min(e)d Your Tags: Analysis of Question Response Time in StackOverflow