

Sign the consent with your name ⓘ

Lydia Danas

Matthew Chen

Charlie Dobson

Sajed Jalil

Saikat Chakraborty

David Christopher Thames

Antonio Mastropaolo

Jake Richardson

Tyler M. Pringle

Henry Burke

Santiago Felipe Arteaga Martin

Manav Kapoor

Anshu Sharma

Conor O'Brien

Kobi Feldman

Christina Mauer

Steven Jia

Ryan Russell

Collin McMillan

Safwat Ali Khan

Stacey Lu

Kaushal Kafle

Grant Enderson

Jinhao Dong

Katelyn Beasley

Anna McKinney

Arpan Mahara

Noah Schulman

Alex Kolar

Sebastian Mendoza-Toledo

Anna Jordan

Kathryn Balint

Max Di Penta

Ting Zhang

Xin Xia

Nicolette Glut

Max Simon

2

4

4

9

27

7

~1.5

5

1

13

2

0.75

10

3

1

5

1.5

15

4

1.5

7 years

2

6

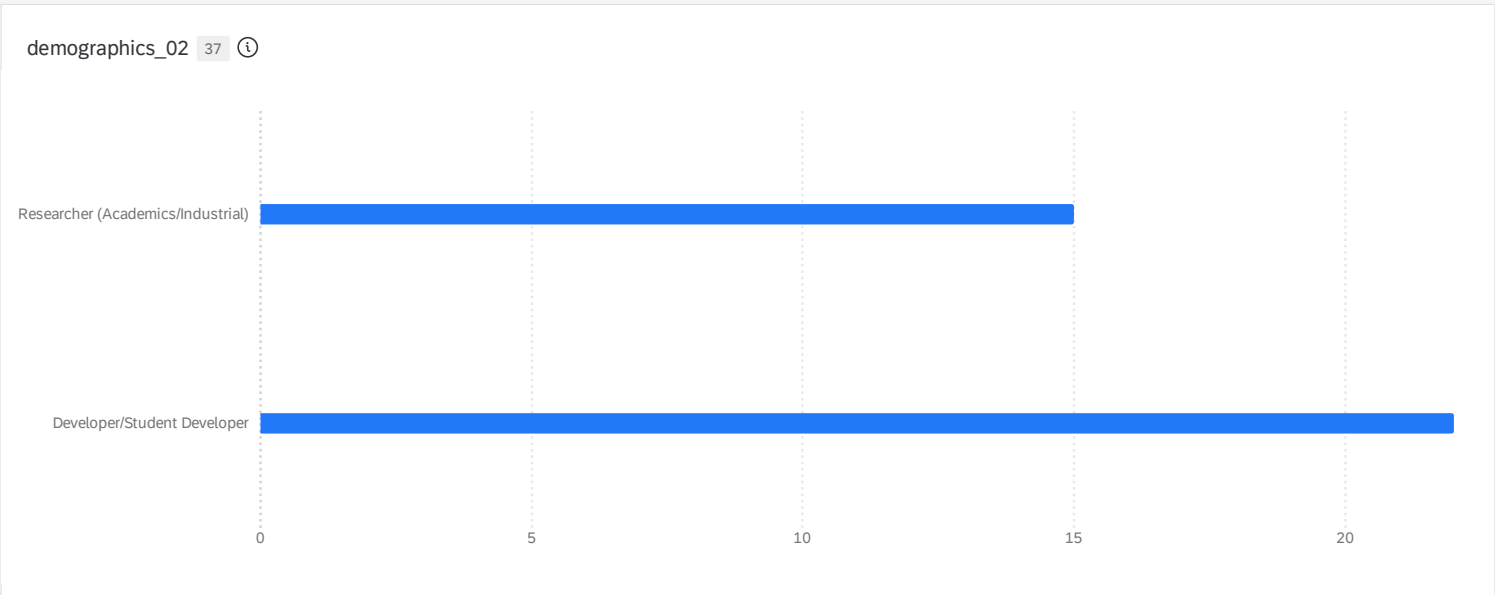
.5

3

2

2

2
2
1
5
35
8
10
3
2

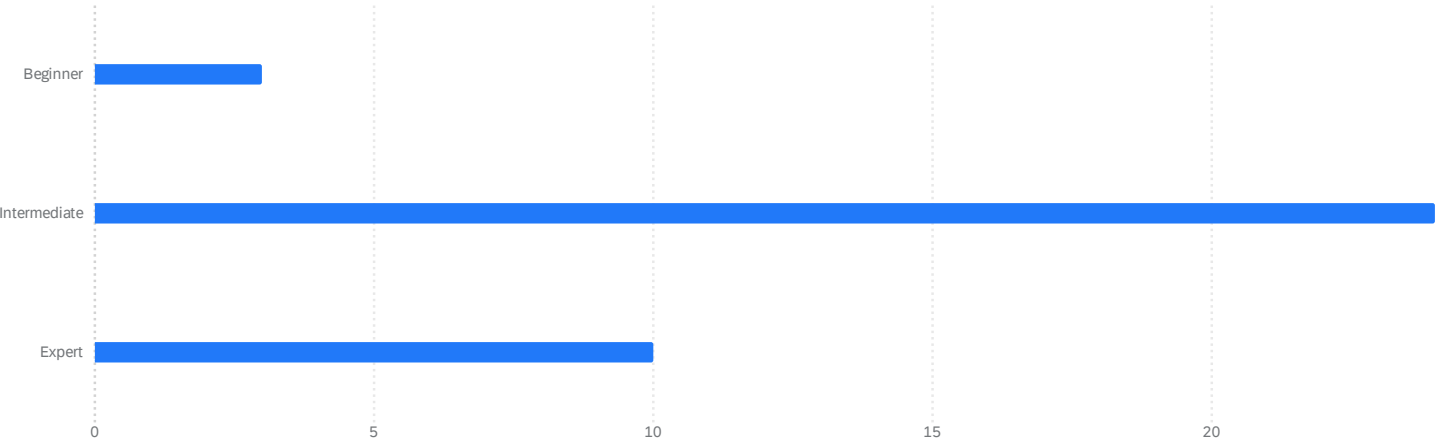


demographics_02 37 ⓘ		
Q2 - demographics_02	Percentage	Count
Researcher (Academics/Industrial)	41%	15
Developer/Student Developer	59%	22
Sum	100%	37

demographics\_02 37 ⓘ

demographics_02	Average	Minimum	Maximum	Count
Researcher (Academics/Industrial)	1.00	1.00	1.00	15
Developer/Student Developer	2.00	2.00	2.00	22

What is your level of expertise in python? 37 ⓘ



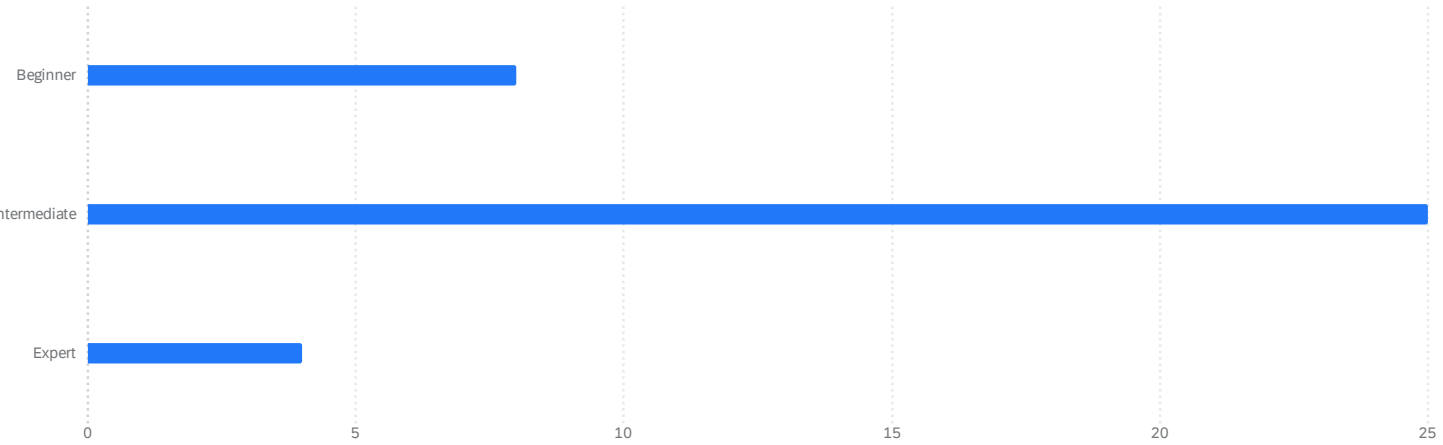
What is your level of expertise in python? 37 ⓘ

Q3 - What is your level of expertise in python?	Percentage	Count
Beginner	8%	3
Intermediate	65%	24
Expert	27%	10
Sum	100%	37

What is your level of expertise in python? 37 ⓘ

What is your level of expertise in python?	Average	Minimum	Maximum	Count
Beginner	1.00	1.00	1.00	3
Intermediate	2.00	2.00	2.00	24
Expert	3.00	3.00	3.00	10

What is your level of expertise in Java? 37 ⓘ



What is your level of expertise in Java? 37 ⓘ

Q213 - What is your level of expertise in Java?	Percentage	Count
Beginner	22%	8
Intermediate	68%	25
Expert	11%	4
Sum	100%	37

What is your level of expertise in Java? 37 ⓘ

What is your level of expertise in Java?	Average	Minimum	Maximum	Count
Beginner	1.00	1.00	1.00	8
Intermediate	2.00	2.00	2.00	25
Expert	3.00	3.00	3.00	4

- Machine learning
- Machine learning
- back-end development
- Machine learning, Script automation with notebooks
- machine learning
- front-end development

Scripting, Deep Learning

Back-End Development

Mainly back-end development

Machine Learning

machine learning, back-end development

Back end development

I rarely use python at my job and when I do it's running a container or script for data generation. I sometimes run python simulators of quantum networks for research.

Back-end development

back-end development, handling data

back-end development

Back-end development

back-end development

machine learning, prototyping

machine learning, back-end development, scripting

Machine learning

backend development, general scripting, research task automation

machine learning and back-end development

Deep learning

data analysis and visualization, back-end development

back-end development

Machine Learning

back-end development, school projects

Typically just scripting and occasional data visualization.

Scripts

---

I mostly use Python for doing practice problems (in the Competitive Programming class)

---

back-end development

---

Machine learning

---

machine learning

---

machine learning

---

Full-stack development

---

data analysis and leetcode

N/a

---

Full-stack development

---

back-end development

---

Full stack development

---

back-end, program analysis

---

classwork only

---

In House Prototype

---

Back-End Development

---

Full-stack development

---

Mobile Development

---

back-end development

---

back end development

---

I mostly used Java in coursework, sometimes in my job I'll come across a Java file that has some useful information.

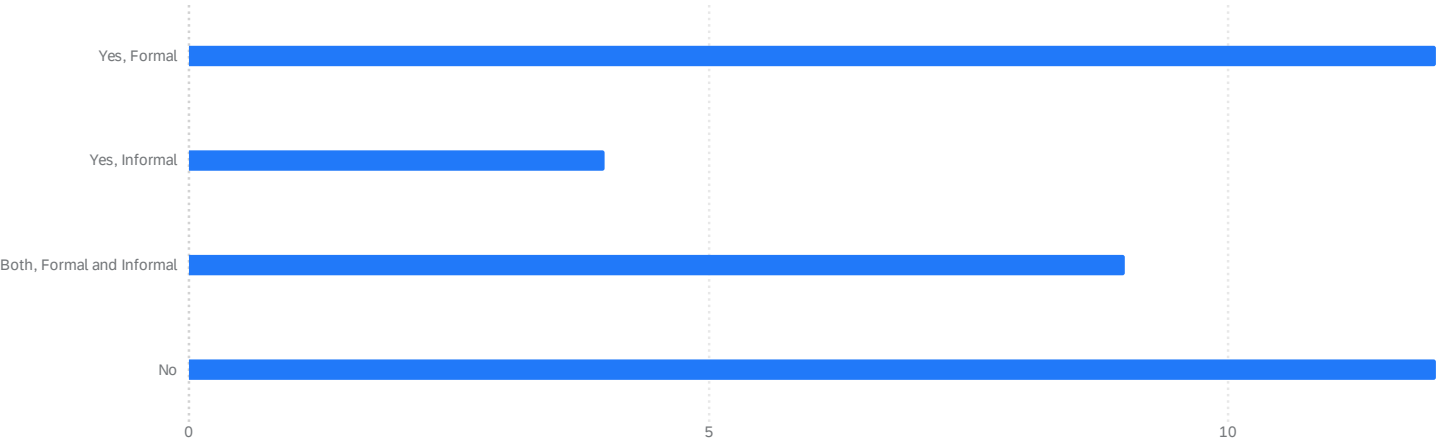


Back-end development
games, full-stack development
back-end development
Back-end development
back-end development
research, student project management
back-end development
Front-end development
code analysis
N/A, I do not work with Java currently
Back-end development
back-end development
back-end development
Back-End Development
school projects
Full stack development.
Back-End Development
Full-stack
back-end development
Back-end development
data structure and algorithms
back-end development

Full-stack development

full stack development

Do you have a background in machine learning? Formal (college classes, degree, certification) or informal (self-learning or other training)? 37 ⓘ



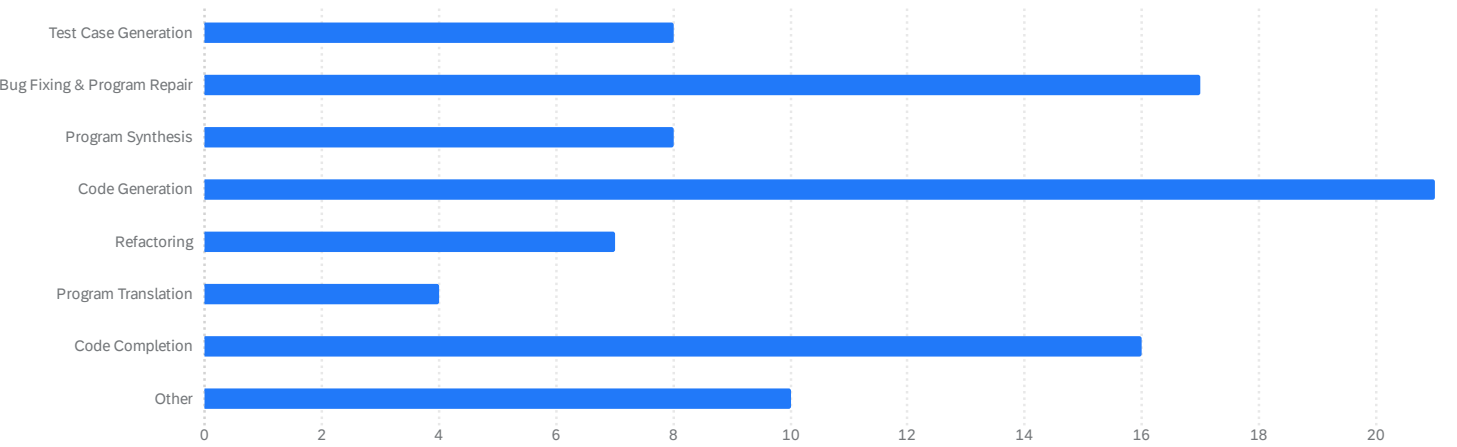
Do you have a background in machine learning? Formal (college classes, degree, certification) or informal (self-learning or other training)? 37 ⓘ

Q5 - Do you have a background in machine learning? Formal (college classes, degree, certification) or informal (self-learning or other training)?	Percentage	Count
Yes, Formal	32%	12
Yes, Informal	11%	4
Both, Formal and Informal	24%	9
No	32%	12
Sum	100%	37

Do you have a background in machine learning? Formal (college classes, degree, certification) or informal (self-learning or other training)? 37 ⓘ

Do you have a background in machine learning? Formal (college classes, deg...	Average	Minimum	Maximum	Count
Yes, Formal	1.00	1.00	1.00	12
Yes, Informal	2.00	2.00	2.00	4
Both, Formal and Informal	3.00	3.00	3.00	9
No	6.00	6.00	6.00	12

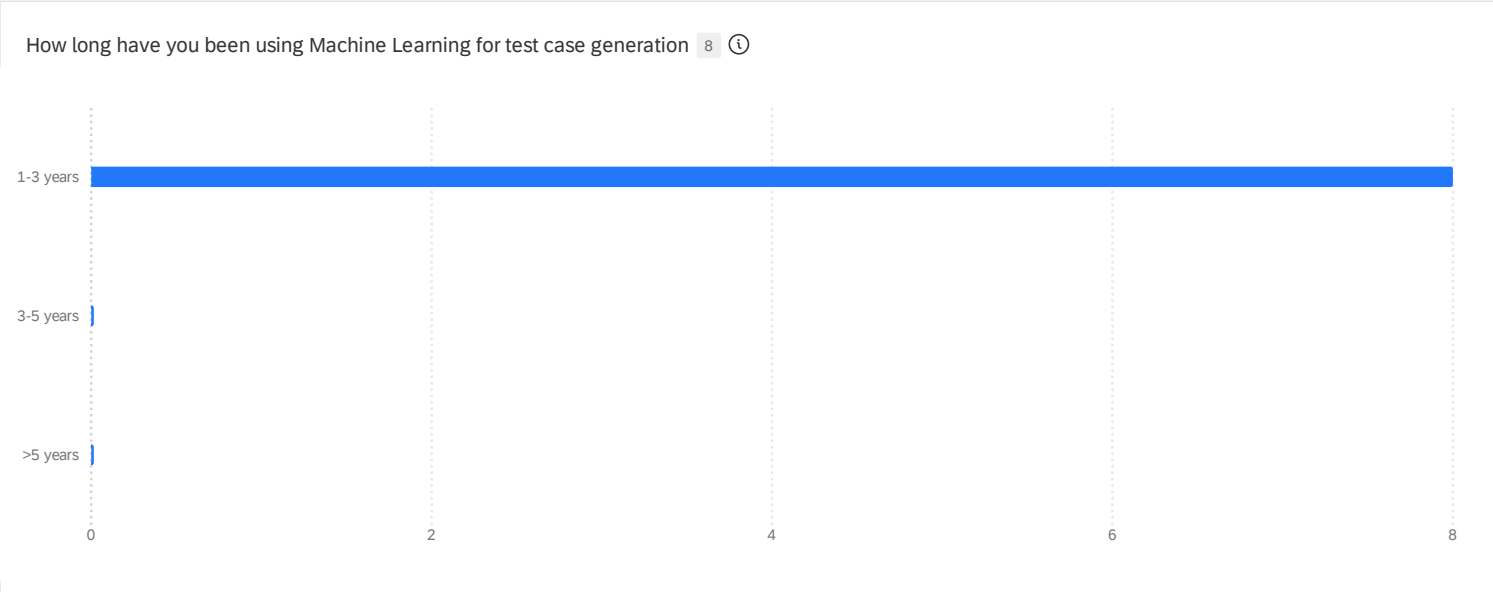
Please indicate for which programming tasks you have used an AI-assisted tool or Large Language Model. 33 ⓘ



Please indicate for which programming tasks you have used an AI-assisted tool or Large Language Model. 33 ⓘ

Q8 - Please indicate for which programming tasks you have used an AI-assisted tool or Large Language Model.	Percentage	Count
Test Case Generation	24%	8
Bug Fixing & Program Repair	52%	17
Program Synthesis	24%	8

Q8 - Please indicate for which programming tasks you have used an AI-assisted tool or Large Language Model.	Percentage	Count
Code Generation	64%	21
Refactoring	21%	7
Program Translation	12%	4
Code Completion	48%	16
Other	30%	10
Sum	276%	91



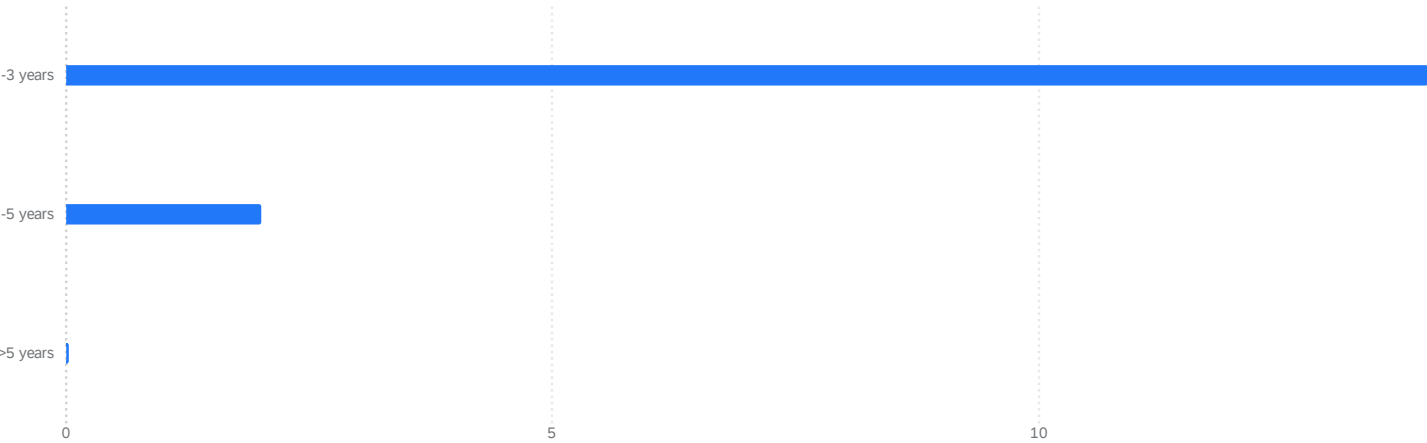
How long have you been using Machine Learning for test case generation 8 ⓘ

Q108 - How long have you been using Machine Learning for test case generation	Percentage	Count
1-3 years	100%	8
3-5 years	0%	0
>5 years	0%	0
Sum	100%	8

How long have you been using Machine Learning for test case generation 8 ⓘ

How long have you been using Machine Learning for test case generation	Average	Minimum	Maximum	Count
1-3 years	1.00	1.00	1.00	8
3-5 years	-	-	-	0
>5 years	-	-	-	0

How long have you been using Machine Learning for Code Completion 16 ⓘ



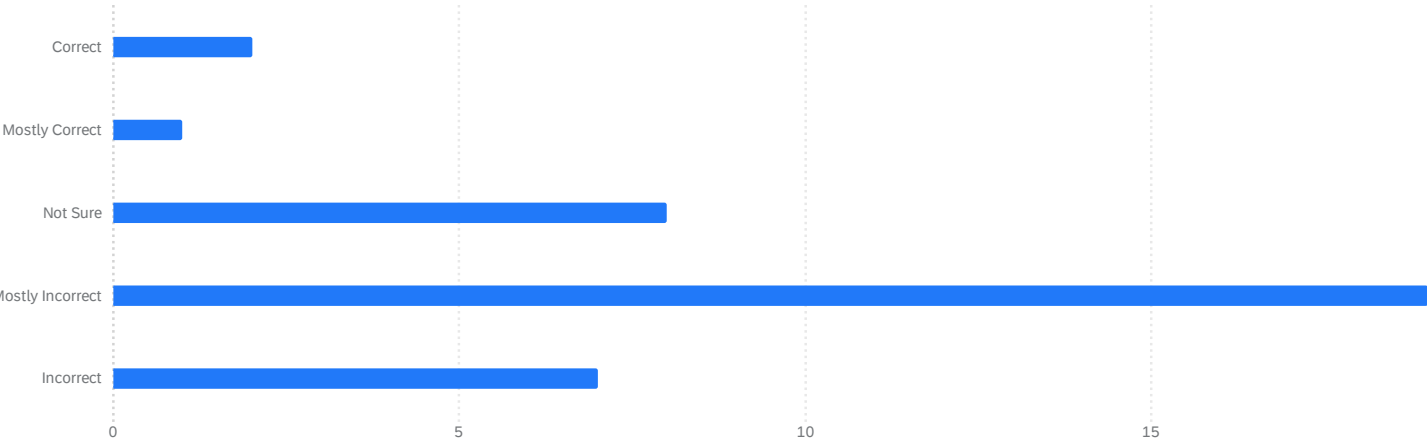
How long have you been using Machine Learning for Code Completion 16 ⓘ

Q117 - How long have you been using Machine Learning for Code Completion	Percentage	Count
1-3 years	88%	14
3-5 years	13%	2
>5 years	0%	0
Sum	100%	16

How long have you been using Machine Learning for Code Completion 16 ⓘ

How long have you been using Machine Learning for Code Completion	Average	Minimum	Maximum	Count
1-3 years	1.00	1.00	1.00	14
3-5 years	2.00	2.00	2.00	2
>5 years	-	-	-	0

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q148 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	5%	2
Mostly Correct	3%	1
Not Sure	22%	8
Mostly Incorrect	51%	19
Incorrect	19%	7
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	2
Mostly Correct	2.00	2.00	2.00	1
Not Sure	6.00	6.00	6.00	8
Mostly Incorrect	7.00	7.00	7.00	19
Incorrect	8.00	8.00	8.00	7

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ def ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 35 ⓘ

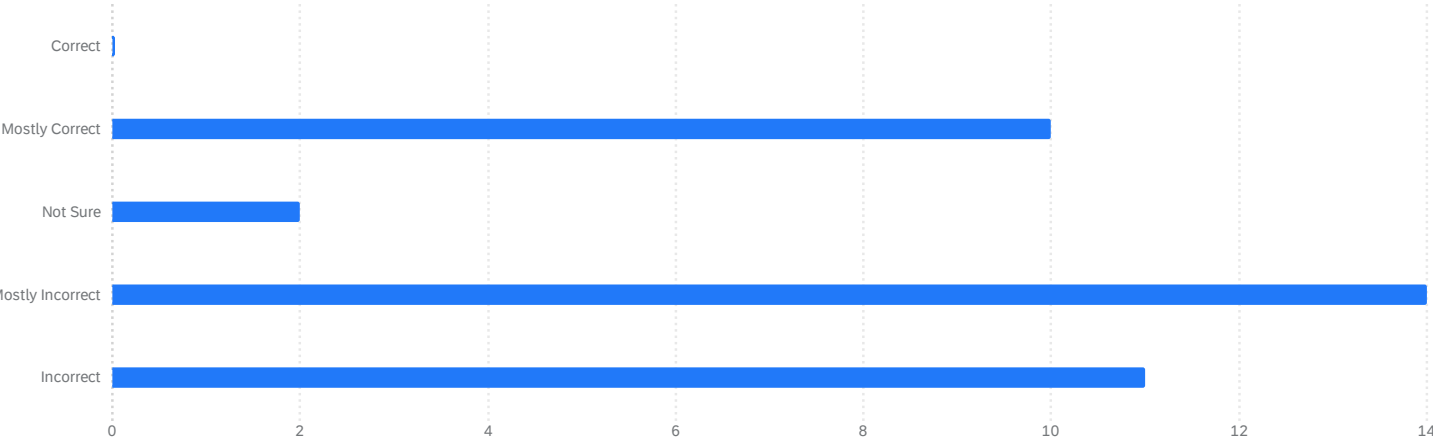
Q150_4 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ def ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply	Percentage	Count
1: def	43%	15
2: data	9%	3
3: (	0%	0
4: self	20%	7
5: )	0%	0
6: -	0%	0
7: &gt;	0%	0
8: Tensor	6%	2
9: Image	11%	4
10: :	0%	0

Q150\_4 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ def ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

PercentageCount

11: "	0%	0
12: Return	29%	10
13: this	11%	4
14: images	20%	7
15: pixels	14%	5
16: as	3%	1
17: a	3%	1
18: tensor	14%	5
19: .	0%	0
20: "	0%	0
21: return	57%	20
22: self	31%	11
23: .	26%	9
24: px	29%	10
Sum	326%	114

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q158 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	0%	0
Mostly Correct	27%	10
Not Sure	5%	2
Mostly Incorrect	38%	14
Incorrect	30%	11
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	-	-	-	0
Mostly Correct	2.00	2.00	2.00	10
Not Sure	7.00	7.00	7.00	2
Mostly Incorrect	8.00	8.00	8.00	14
Incorrect	9.00	9.00	9.00	11

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ return ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 34 ⓘ

Q159_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ return ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply	Percentage	Count
1: def	38%	13
2: Set	24%	8
3: Bit	15%	5
4: map	15%	5
5: Disabled	29%	10
6: (	3%	1
7: *	0%	0
8: args	6%	2
9: ,	0%	0
10: **	0%	0



Q159\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ return ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

PercentageCount

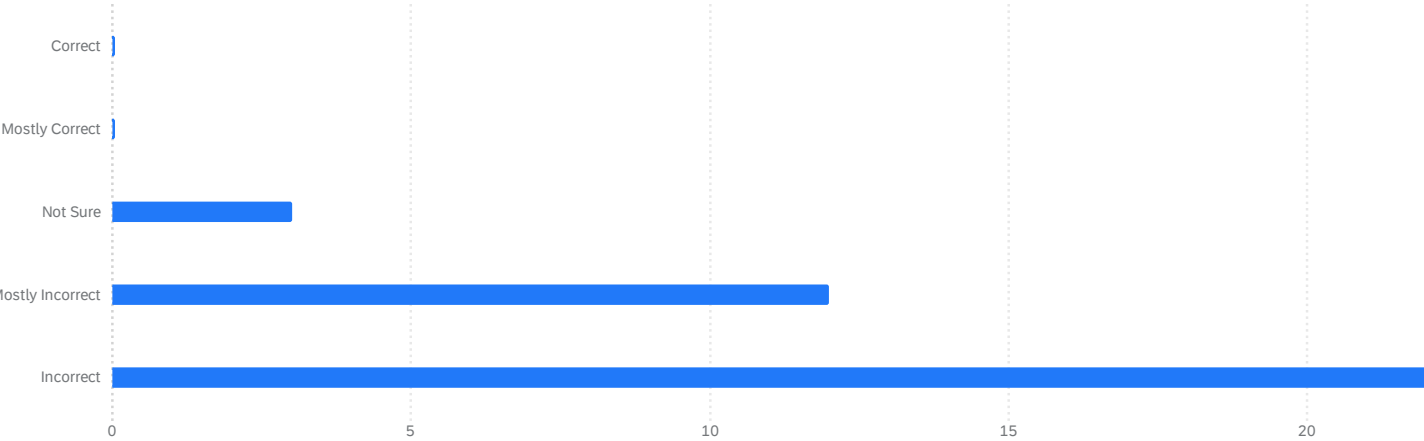
11: kwargs	3%	1
12: )	3%	1
13: :	3%	1
14: ""	3%	1
15: Set	24%	8
16: Bit	12%	4
17: map	12%	4
18: Disabled	26%	9
19: (	9%	3
20: self	9%	3
21: ,	9%	3
22: Bit	9%	3
23: map	9%	3
24: bit	9%	3
25: map	9%	3
26: )	9%	3
27: ""	3%	1
28: return	29%	10
29: _	0%	0
30: controls	3%	1
31: _	0%	0
32: .Any	3%	1
33: Button	0%	0
34: _	0%	0
35: Set	12%	4
36: Bit	12%	4
37: map	12%	4
38: Disabled	21%	7

Q159\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ return ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

PercentageCount

39: (	15%	5
40: *	12%	4
41: args	15%	5
42: ,	12%	4
43: **	12%	4
44: kwargs	18%	6
45: )	18%	6
Sum	479%	163

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q252 - Q.1- How correct do you think the generated code is?

PercentageCount

Correct	0%	0
Mostly Correct	0%	0
Not Sure	8%	3
Mostly Incorrect	32%	12
Incorrect	59%	22
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	-	-	-	0
Mostly Correct	-	-	-	0
Not Sure	7.00	7.00	7.00	3
Mostly Incorrect	8.00	8.00	8.00	12
Incorrect	9.00	9.00	9.00	22

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ else ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 34 ⓘ

Q253_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ else ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply	Percentage	Count
1: ""	0%	0
2: Generate	3%	1
3: Python	9%	3
4: code	3%	1
5: that	3%	1
6: True	29%	10
7: if	47%	16
8: this	6%	2
9: Entry	9%	3
10: has	9%	3
11: references	9%	3
12: from	9%	3
13: any	12%	4
14: App	12%	4
15: Session	12%	4
16: .	6%	2
17: If	79%	27
18: not	71%	24
19: ,	12%	4

Q253\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ else ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

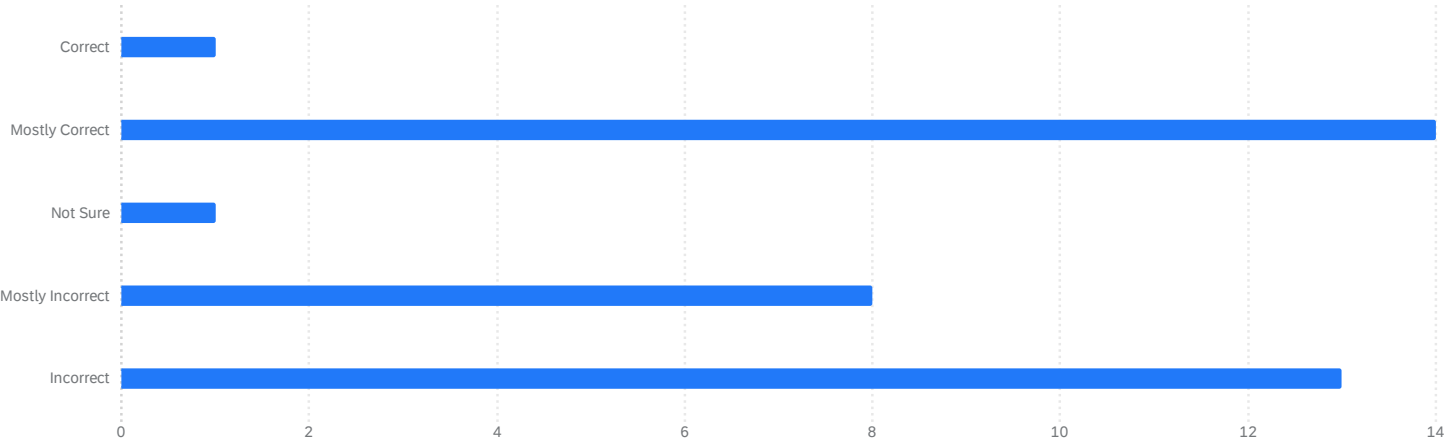
PercentageCount

20: it	12%	4
21: can	12%	4
22: be	9%	3
23: removed	12%	4
24: from	9%	3
25: the	9%	3
26: cache	9%	3
27: .	6%	2
28: and	0%	0
29: signature	0%	0
30: is	0%	0
31: ""	0%	0
32: def	6%	2
33: has	3%	1
34: _	0%	0
35: refs	0%	0
36: (	3%	1
37: self	3%	1
38: )	3%	1
39: -	3%	1
40: &gt;	3%	1
41: bool	21%	7
42: :	9%	3
43: self	3%	1
44: .	0%	0
45: ref	0%	0
46: ,	0%	0
47: self	3%	1

Q253\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ else ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

	Percentage	Count
48: .	0%	0
49: context	3%	1
50: =	3%	1
51: None	12%	4
Sum	491%	167

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q260 - Q.1- How correct do you think the generated code is?

	Percentage	Count
Correct	3%	1
Mostly Correct	38%	14
Not Sure	3%	1
Mostly Incorrect	22%	8
Incorrect	35%	13
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

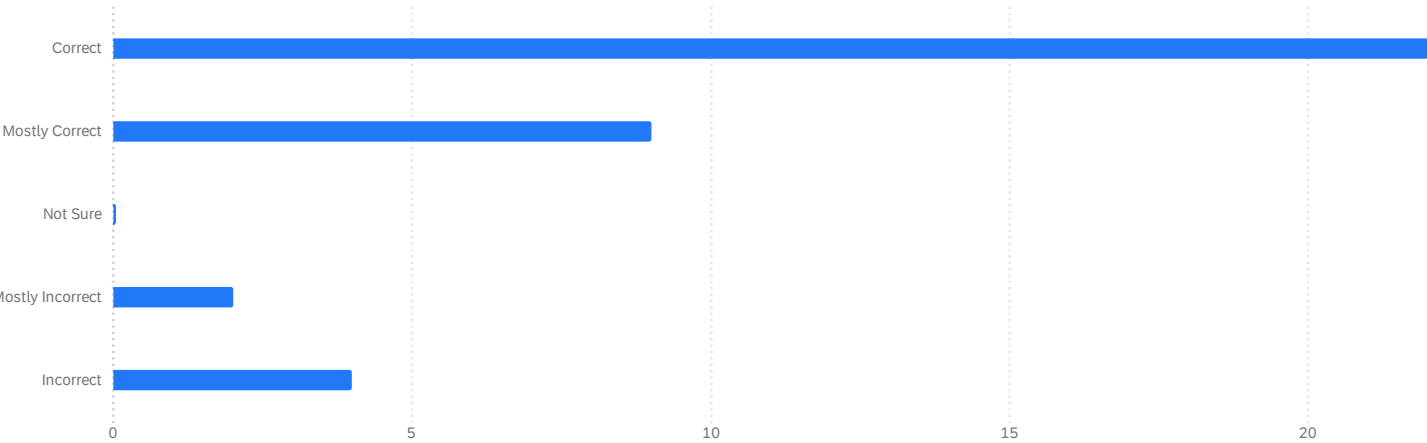
Q.1- How correct do you think the generated code is?

	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	1

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Mostly Correct	2.00	2.00	2.00	14
Not Sure	7.00	7.00	7.00	1
Mostly Incorrect	8.00	8.00	8.00	8
Incorrect	9.00	9.00	9.00	13

<p>Q.2- What words/tokens do you think are responsible for the generation of the word/token [ close ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply):</p> <div> <div>35</div> <div></div> </div>				
<p>Q261_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ close ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply):</p> <div> <div>- Apply</div> </div>			Percentage	Count
1: def			6%	2
2: read			23%	8
3: lines			20%	7
4: (			0%	0
5: self			3%	1
6: )			0%	0
7: :			0%	0
8: #			3%	1
9: remember			31%	11
10: to			29%	10
11: close			89%	31
12: after			17%	6
13: open			31%	11
14: f			14%	5
15: =			6%	2
Sum			271%	95

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q264 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	59%	22
Mostly Correct	24%	9
Not Sure	0%	0
Mostly Incorrect	5%	2
Incorrect	11%	4
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	22
Mostly Correct	2.00	2.00	2.00	9
Not Sure	-	-	-	0
Mostly Incorrect	8.00	8.00	8.00	2
Incorrect	9.00	9.00	9.00	4

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 36 ⓘ

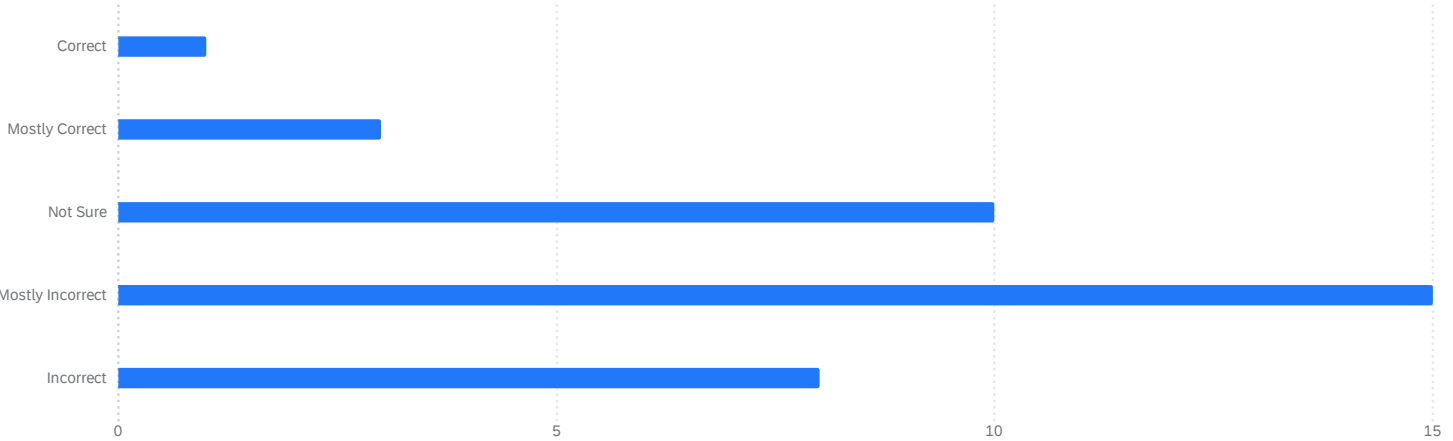
Q265_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply	Percentage	Count
1: ""	3%	1
2: Generate	3%	1
3: a	3%	1
4: python	8%	3
5: code	6%	2
6: that	3%	1
7: calculates	22%	8
8: summation	78%	28
9: of	31%	11
10: list	72%	26
11: elements	47%	17
12: ""	3%	1
13: def	8%	3
14: list	33%	12
15: _	14%	5
16: sum	36%	13
17: (	3%	1
18: *	8%	3
19: inp	22%	8



Q265\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - Apply

	Percentage	Count
20: _	22%	8
21: lst	28%	10
22: )	3%	1
23: :	3%	1
24: s	17%	6
25: =	11%	4
26: inp	22%	8
27: _	22%	8
28: lst	25%	9
29: [	22%	8
30: 0	22%	8
31: ]	22%	8
Sum	622%	224

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q183 - Q.1- How correct do you think the generated code is?

	Percentage	Count
Correct	3%	1
Mostly Correct	8%	3
Not Sure	27%	10

Q183 - Q.1- How correct do you think the generated code is?	Percentage	Count
Mostly Incorrect	41%	15
Incorrect	22%	8
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	1
Mostly Correct	6.00	6.00	6.00	3
Not Sure	7.00	7.00	7.00	10
Mostly Incorrect	8.00	8.00	8.00	15
Incorrect	2.00	2.00	2.00	8

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 35 ⓘ

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

	Percentage	Count
5: Balance	3%	1
6: Books	3%	1
7: implements	0%	0
8: Closeable	0%	0
9: {	0%	0
10: public	0%	0
11: Balance	0%	0
12: Books	3%	1
13: (	0%	0
14: int	9%	3
15: total	11%	4
16: Clients	6%	2
17: ,	3%	1
18: int	23%	8
19: iterations)	54%	19

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

	Percentage	Count
20: {	0%	0
21: this	0%	0
22: (	0%	0
23: total	3%	1
24: Clients	0%	0
25: ,	0%	0
26: iterations	34%	12
27: ,	0%	0
28: new	0%	0
29: Configuration	0%	0
30: Factory	0%	0
31: (	0%	0
32: )	0%	0
33: .get	0%	0
34: (	0%	0
35: )	0%	0
36: )	0%	0
37: ;	0%	0
38: }	0%	0
39: Balance	9%	3
40: Books	9%	3
41: (	0%	0
42: int	14%	5
43: total	14%	5
44: Clients	9%	3
45: ,	0%	0
46: int	20%	7
47: iterations	49%	17

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

	Percentage	Count
48: )	0%	0
49: ;	0%	0
50: Balance	14%	5
51: Books	14%	5
52: (	6%	2
53: int	17%	6
54: total	20%	7
55: Clients	14%	5
56: ,	6%	2
57: int	23%	8
58: iterations	51%	18
59: ,	6%	2
60: Configuration	11%	4
61: conf	11%	4
62: )	6%	2
63: ;	6%	2
64: void	0%	0
65: init	0%	0
66: (	0%	0
67: )	0%	0
68: ;	0%	0
69: void	0%	0
70: run	3%	1
71: (	0%	0
72: )	0%	0
73: ;	0%	0
74: boolean	0%	0
75: verify	0%	0

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

76: (	0%	0
77: )	0%	0
78: ;	0%	0
79: void	0%	0
80: close	0%	0
81: (	0%	0
82: )	0%	0
83: ;	0%	0
84: static	0%	0
85: void	0%	0
86: main	0%	0
87: (	0%	0
88: String	0%	0
89: [	0%	0
90: ]	0%	0
91: args	0%	0
92: )	0%	0
93: ;	0%	0
94: }	0%	0
99: @	0%	0
100: Test	3%	1
101: public	3%	1
102: void	0%	0
103: test	6%	2
104: Balance	3%	1
105: Books	6%	2
106: (	0%	0
107: )	0%	0

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

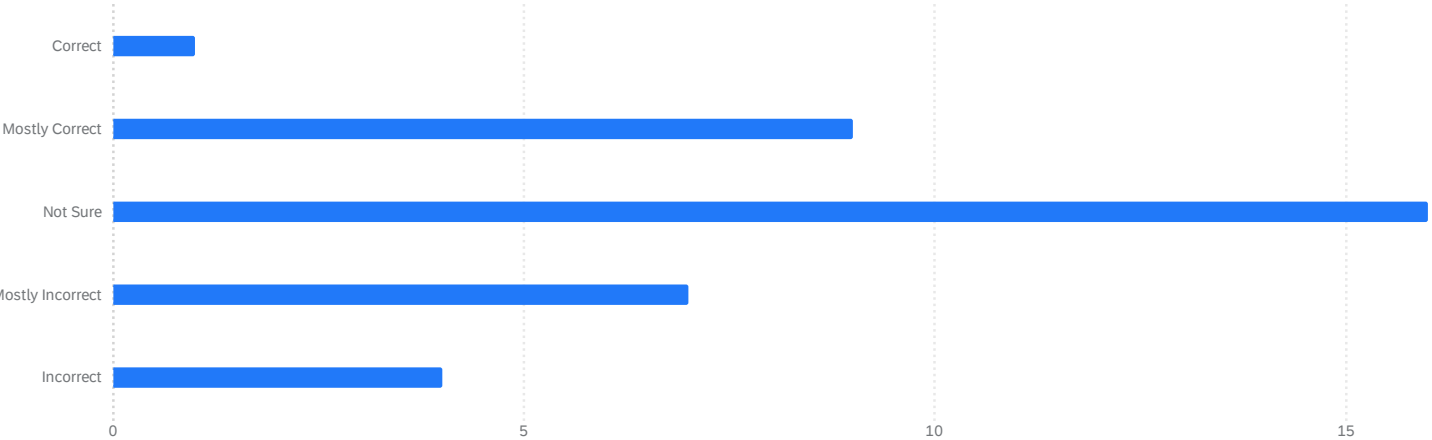
108: throws	0%	0
109: Exception	0%	0
110: {	3%	1
111: for	49%	17
112: (	11%	4
113: int	20%	7
114: i	20%	7
115: =	14%	5
116: 0	14%	5
117: ;	14%	5
118: i	20%	7
119: &lt;	14%	5
120: 1000	17%	6
121: ;	11%	4
122: i	17%	6
123: +	14%	5
124: +)	14%	5
125: {	3%	1
126: Balance	6%	2
127: Books	11%	4
128: book	3%	1
129: =	3%	1
130: new	9%	3
131: Balance	9%	3
132: Books	9%	3
133: (	9%	3
134: i	17%	6
135: )	9%	3

Q185\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ for ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

136: ;	9%	3
Sum	851%	298

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q272 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	3%	1
Mostly Correct	24%	9
Not Sure	43%	16
Mostly Incorrect	19%	7
Incorrect	11%	4
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	1
Mostly Correct	6.00	6.00	6.00	9
Not Sure	7.00	7.00	7.00	16
Mostly Incorrect	8.00	8.00	8.00	7

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Incorrect	2.00	2.00	2.00	4

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ cleanUpDescriptors ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 36 ⓘ

Q273_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ cleanUpDescriptors ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply	Percentage	Count
4: Sequence	8%	3
5: Cleaner	22%	8
6: Task	6%	2
7: implements	0%	0
8: Runnable	3%	1
9: {	0%	0
10: @	0%	0
11: Override	0%	0
12: public	3%	1
13: void	3%	1
14: run	8%	3
15: (	3%	1
16: )	3%	1
17: {	0%	0
18: manager	14%	5
19: .	14%	5
20: clean	97%	35
21: Up	94%	34
22: Descriptors	94%	34
23: (	22%	8
24: )	22%	8
25: ;	8%	3
26: logger	3%	1
27: .	0%	0
28: debug	6%	2



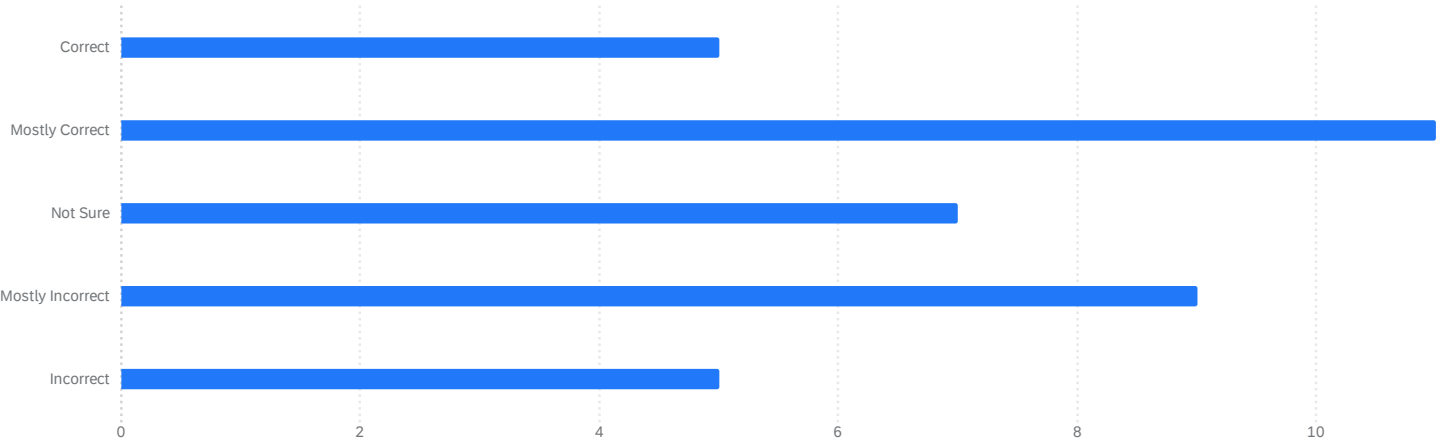
Q273\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ cleanUpDescriptors ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

	Percentage	Count
29: (	0%	0
30: "	0%	0
31: Clean	31%	11
32: process	11%	4
33: has	8%	3
34: ran	8%	3
35: .	0%	0
36: "	0%	0
37: )	0%	0
38: ;	0%	0
39: }	0%	0
40: @	0%	0
41: Override	0%	0
42: void	0%	0
43: run	0%	0
44: (	0%	0
45: )	0%	0
46: ;	0%	0
47: }	0%	0
51: @	0%	0
52: Test	3%	1
53: public	3%	1
54: void	3%	1
55: test	6%	2
56: Run	8%	3
57: (	0%	0
58: )	0%	0
59: throws	0%	0

Q273\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ cleanUpDescriptors ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

	Percentage	Count
60: Exception	0%	0
61: {	0%	0
62: task	6%	2
63: .	3%	1
64: run	6%	2
65: (	3%	1
66: )	3%	1
67: ;	0%	0
68: verify	6%	2
69: (	3%	1
Sum	542%	195

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q276 - Q.1- How correct do you think the generated code is?

	Percentage	Count
Correct	14%	5
Mostly Correct	30%	11
Not Sure	19%	7
Mostly Incorrect	24%	9
Incorrect	14%	5
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	5
Mostly Correct	6.00	6.00	6.00	11
Not Sure	7.00	7.00	7.00	7
Mostly Incorrect	8.00	8.00	8.00	9
Incorrect	2.00	2.00	2.00	5

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ finditerative ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 36 ⓘ

Q277_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ finditerative ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply	Percentage	Count
4: Find	6%	2
5: Longest	11%	4
6: Consecutive	11%	4
7: Sequence	14%	5
8: {	0%	0
9: public	0%	0
10: int	3%	1
11: find	17%	6
12: Recursive	8%	3
13: (	0%	0
14: int	3%	1
15: [	3%	1
16: ]	3%	1
17: array	6%	2
18: )	0%	0
19: {	0%	0
20: validate	0%	0
21: Input	0%	0
22: (	0%	0
23: array	0%	0
24: )	0%	0
25: ;	0%	0
26: return	0%	0
27: find	17%	6
28: Recursive	6%	2
29: Inner	0%	0
30: (	0%	0
31: array	3%	1

Q277\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ findIterative ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

32: ,	0%	0
33: 1	0%	0
34: ,	0%	0
35: 0	0%	0
36: ,	0%	0
37: 0	0%	0
38: )	0%	0
39: ;	0%	0
40: }	0%	0
41: int	22%	8
42: findIterative	97%	35
43: (	17%	6
44: int	31%	11
45: [	28%	10
46: ]	28%	10
47: numbers	31%	11
48: )	17%	6
49: ;	8%	3
50: int	0%	0
51: find	8%	3
52: Recursive	3%	1
53: (	0%	0
54: int	3%	1
55: [	3%	1
56: ]	3%	1
57: array	6%	2
58: )	0%	0
59: ;	0%	0

Q277\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ findIterative ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

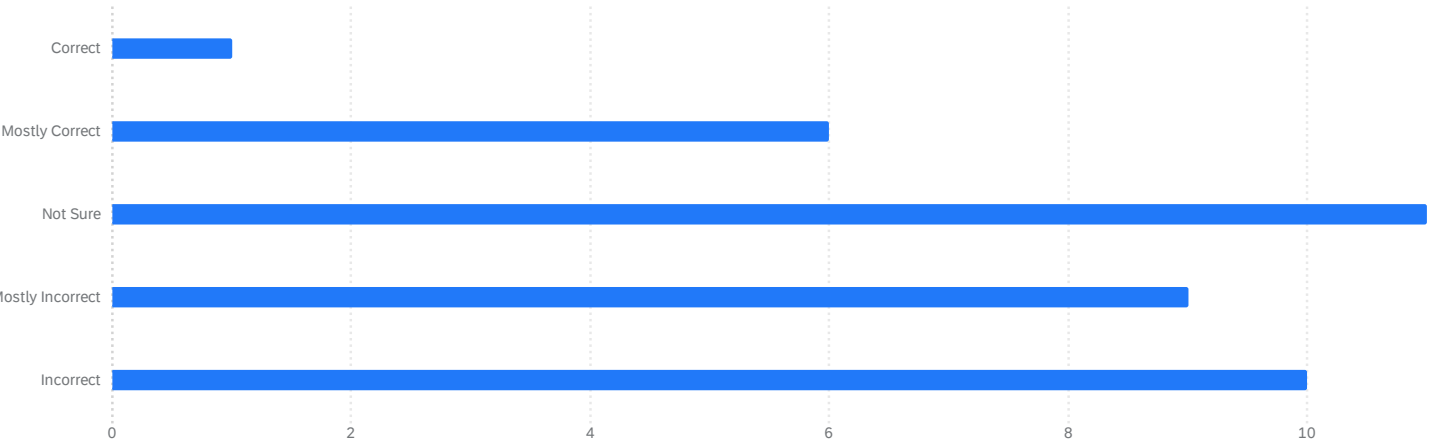
60: }	0%	0
64: @	0%	0
65: Test	0%	0
66: public	0%	0
67: void	0%	0
68: should	8%	3
69: Find	17%	6
70: Longest	8%	3
71: Consecutive	11%	4
72: Sequence	14%	5
73: Recursive	0%	0
74: (	0%	0
75: )	0%	0
76: {	0%	0
77: int	14%	5
78: [	11%	4
79: ]	11%	4
80: array	19%	7
81: =	0%	0
82: {	0%	0
83: 1	3%	1
84: ,	0%	0
85: 3	3%	1
86: ,	0%	0
87: 4	3%	1
88: ,	0%	0
89: 5	3%	1
90: ,	0%	0

Q277\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ findIterative ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

91: 64	3%	1
92: ,	0%	0
93: 4	3%	1
94: ,	0%	0
95: 5	3%	1
96: ,	0%	0
97: 6	3%	1
98: ,	0%	0
99: 7	3%	1
100: ,	0%	0
101: 8	3%	1
102: ,	0%	0
103: 9	3%	1
104: ,	0%	0
105: 98	3%	1
106: ,	0%	0
107: -1	3%	1
108: ,	0%	0
109: -2	3%	1
110: }	0%	0
111: ;	0%	0
112: int	6%	2
113: sequence	17%	6
114: Length	3%	1
115: =	3%	1
116: lcs	14%	5
117: .	6%	2
Sum	611%	220

Q.1- How correct do you think the generated code is? 37 ⓘ



Q.1- How correct do you think the generated code is? 37 ⓘ

Q280 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	3%	1
Mostly Correct	16%	6
Not Sure	30%	11
Mostly Incorrect	24%	9
Incorrect	27%	10
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ

Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	1
Mostly Correct	6.00	6.00	6.00	6
Not Sure	7.00	7.00	7.00	11
Mostly Incorrect	8.00	8.00	8.00	9
Incorrect	2.00	2.00	2.00	10



Q.2- What words/tokens do you think are responsible for the generation of the word/token [ latch ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 34 ⓘ

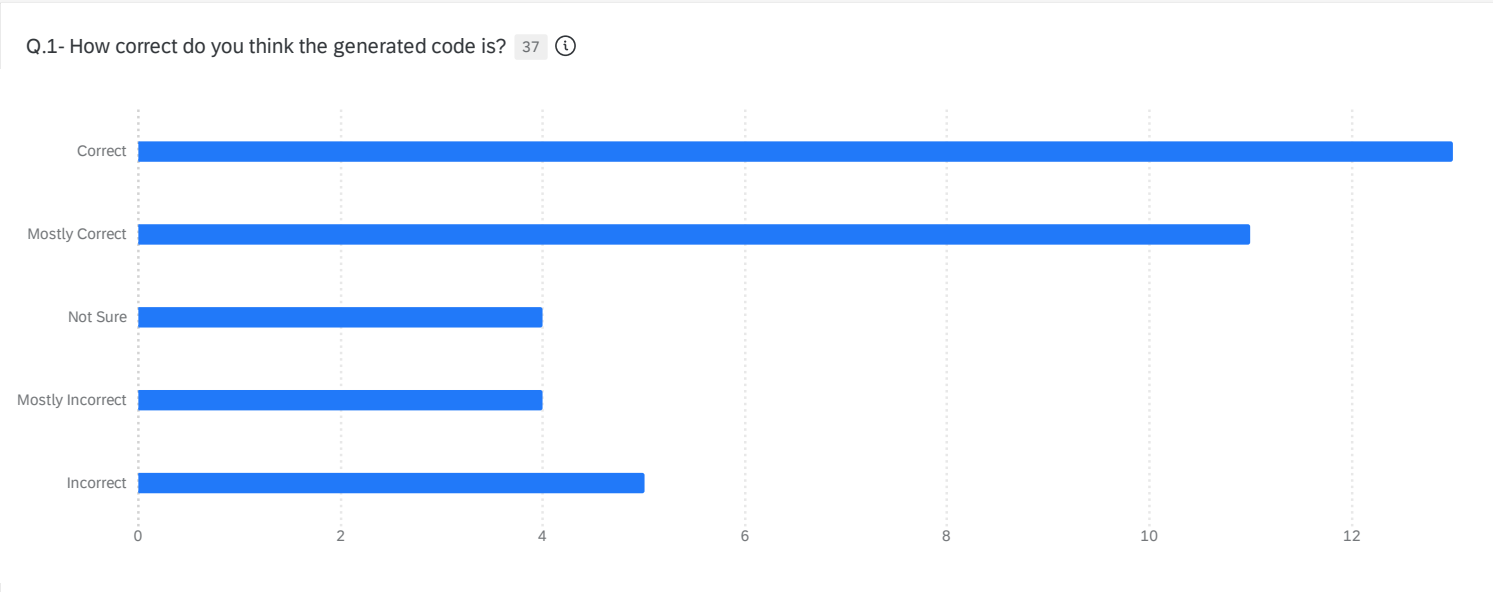
Q281_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ latch ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply	Percentage	Count
4: Cached	3%	1
5: Thread	6%	2
6: Scheduler	3%	1
7: extends	0%	0
8: Scheduler	6%	2
9: {	0%	0
10: @	0%	0
11: Override	0%	0
12: public	3%	1
13: Worker	9%	3
14: create	6%	2
15: Worker	12%	4
16: (	0%	0
17: )	0%	0
18: {	0%	0
19: return	3%	1
20: new	6%	2
21: Event	9%	3
22: Loop	12%	4
23: Worker	12%	4
24: (	0%	0
25: Cached	6%	2
26: Worker	6%	2
27: Pool	3%	1
28: .	3%	1
29: INSTANCE	3%	1
30: .	3%	1
31: get	3%	1

Q281\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ latch ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply

PercentageCount

32: (	3%	1
33: )	3%	1
34: )	0%	0
35: ;	0%	0
36: }	0%	0
37: @	0%	0
38: Override	0%	0
39: Worker	3%	1
40: create	3%	1
41: Worker	6%	2
42: (	0%	0
43: )	0%	0
44: ;	0%	0
45: }	0%	0
49: @	0%	0
50: Test	0%	0
51: public	0%	0
52: void	0%	0
53: test	3%	1
54: Create	3%	1
55: Worker	3%	1
56: (	0%	0
57: )	0%	0
58: throws	0%	0
59: Exception	0%	0
60: {	0%	0
61: final	9%	3
62: Count	21%	7

Q281_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ latch ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply	Percentage	Count
63: Down	21%	7
64: Latch	59%	20
65: latch	65%	22
66: =	6%	2
67: new	9%	3
68: Count	12%	4
69: Down	12%	4
70: Latch	24%	8
71: ;	3%	1
72: final	9%	3
73: Count	21%	7
74: Down	21%	7
75: Latch	59%	20
Sum	488%	166



Q292 - Q.1- How correct do you think the generated code is?	Percentage	Count
Correct	35%	13
Mostly Correct	30%	11
Not Sure	11%	4

Q292 - Q.1- How correct do you think the generated code is?	Percentage	Count
Mostly Incorrect	11%	4
Incorrect	14%	5
Sum	100%	37

Q.1- How correct do you think the generated code is? 37 ⓘ				
Q.1- How correct do you think the generated code is?	Average	Minimum	Maximum	Count
Correct	1.00	1.00	1.00	13
Mostly Correct	6.00	6.00	6.00	11
Not Sure	7.00	7.00	7.00	4
Mostly Incorrect	8.00	8.00	8.00	4
Incorrect	2.00	2.00	2.00	5

Q.2- What words/tokens do you think are responsible for the generation of the word/token [ NoSuchElementException ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): 36 ⓘ				
Q293_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ NoSuchElementException ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply): - apply			Percentage	Count
4: Jcr			0%	0
5: Empty			3%	1
6: Node			3%	1
7: Iterator			3%	1
8: implements			0%	0
9: Node			0%	0
10: Iterator			0%	0
11: {			0%	0
12: @			0%	0
13: Override			0%	0
14: public			11%	4
15: Object			11%	4
16: next			22%	8
17: (			6%	2
18: )			6%	2

Q293\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ NoSuchElementException ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply):  
- apply

	Percentage	Count
19: {	6%	2
20: throw	61%	22
21: new	56%	20
22: No	97%	35
23: Such	97%	35
24: Element	97%	35
25: Exception	100%	36
26: (	22%	8
27: )	22%	8
28: ;	17%	6
29: }	8%	3
30: private	0%	0
31: Jcr	0%	0
32: Empty	3%	1
33: Node	3%	1
34: Iterator	0%	0
35: (	0%	0
36: )	0%	0
37: ;	0%	0
38: @	0%	0
39: Override	0%	0
40: Node	0%	0
41: next	0%	0
42: Node	0%	0
43: (	0%	0
44: )	0%	0
45: ;	0%	0
46: @	0%	0

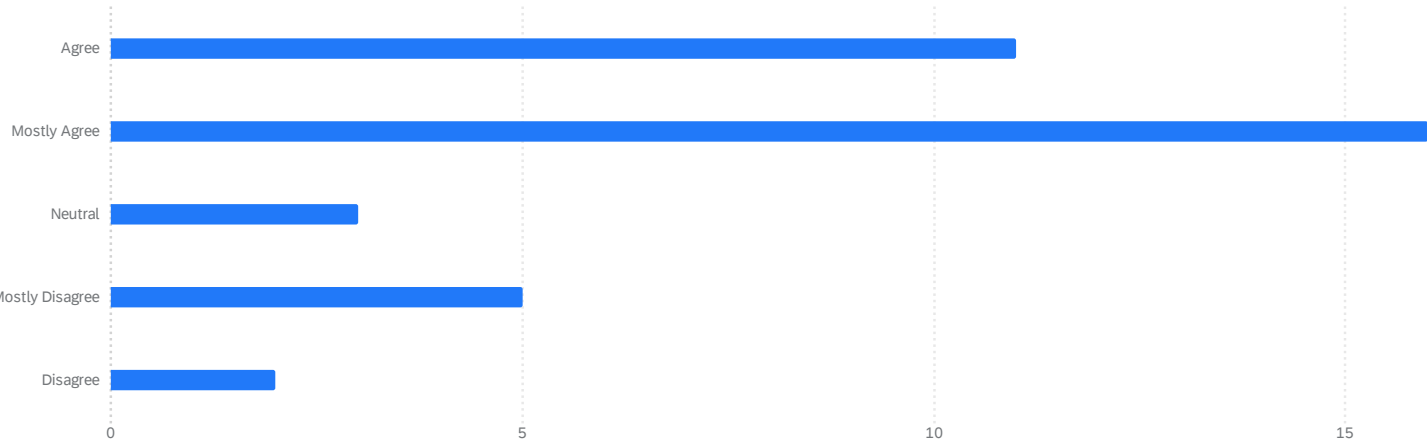
Q293\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ NoSuchElementException ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply):  
- apply

	Percentage	Count
47: Override	0%	0
48: long	0%	0
49: get	0%	0
50: Position	0%	0
51: (	0%	0
52: )	0%	0
53: ;	0%	0
54: @	0%	0
55: Override	0%	0
56: long	0%	0
57: get	0%	0
58: Size	0%	0
59: (	0%	0
60: )	0%	0
61: ;	0%	0
62: @	0%	0
63: Override	0%	0
64: void	0%	0
65: skip	0%	0
66: (	0%	0
67: long	0%	0
68: skip	0%	0
69: Num	0%	0
70: )	0%	0
71: ;	0%	0
72: @	0%	0
73: Override	0%	0
74: boolean	0%	0

Q293\_1 - Q.2- What words/tokens do you think are responsible for the generation of the word/token [ NoSuchElementException ], highlighted inside a red box, in the previous sample? A sequence of words is shown below for your convenience, please select all that apply (click on the word/token then apply):  
- apply

	Percentage	Count
75: has	0%	0
76: Next	0%	0
77: (	0%	0
78: )	0%	0
79: ;	0%	0
80: @	0%	0
81: Override	0%	0
82: Object	0%	0
83: next	3%	1
84: (	0%	0
85: )	0%	0
86: ;	0%	0
87: @	0%	0
88: Override	0%	0
89: void	0%	0
90: remove	0%	0
91: (	0%	0
92: )	0%	0
93: ;	0%	0
94: }	0%	0
98: @	3%	1
99: Test	8%	3
100: (	3%	1
101: expected	28%	10
102: =	17%	6
Sum	714%	257

Q.1- Do you agree with the code rationales generated for token [def] highlighted inside the red box? 37 ⓘ



Q.1- Do you agree with the code rationales generated for token [def] highlighted inside the red box? 37 ⓘ

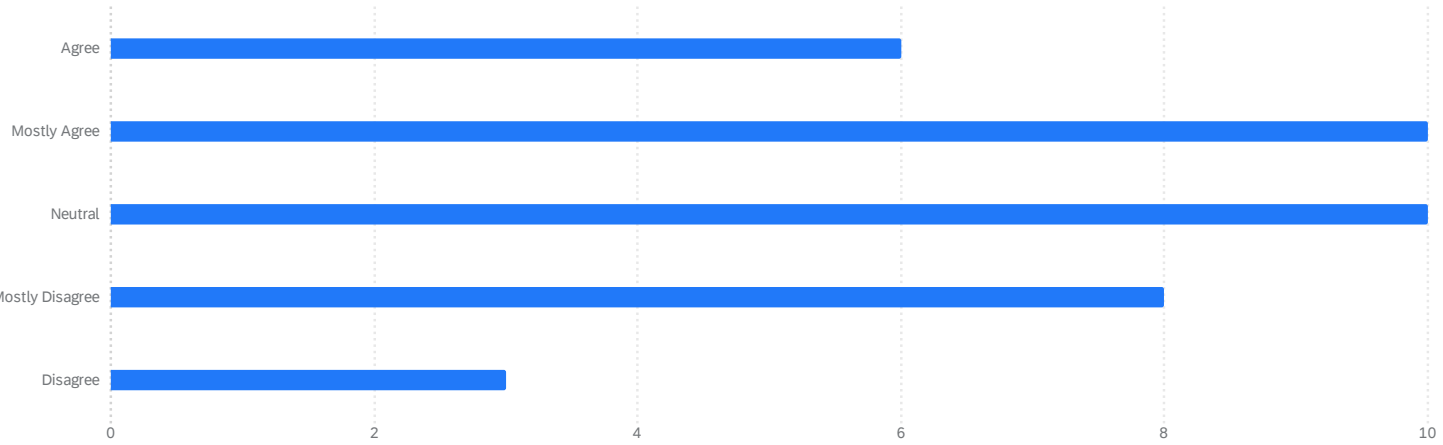
Q304 - Q.1- Do you agree with the code rationales generated for token [def] highlighted inside the red box?	Percentage	Count
Agree	30%	11
Mostly Agree	43%	16
Neutral	8%	3
Mostly Disagree	14%	5
Disagree	5%	2
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [def] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [def] highli...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	11
Mostly Agree	2.00	2.00	2.00	16
Neutral	6.00	6.00	6.00	3
Mostly Disagree	7.00	7.00	7.00	5
Disagree	8.00	8.00	8.00	2



Q.1- Do you agree with the code rationales generated for token [return] highlighted inside the red box? 37 ⓘ



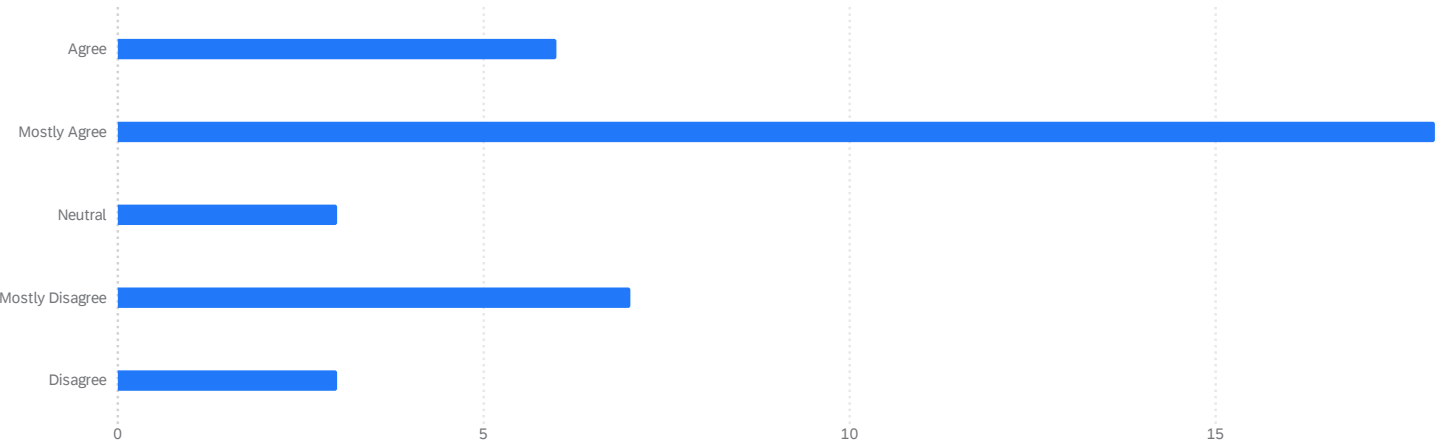
Q.1- Do you agree with the code rationales generated for token [return] highlighted inside the red box? 37 ⓘ

Q312 - Q.1- Do you agree with the code rationales generated for token [return] highlighted inside the red box?	Percentage	Count
Agree	16%	6
Mostly Agree	27%	10
Neutral	27%	10
Mostly Disagree	22%	8
Disagree	8%	3
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [return] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [return] hig...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	6
Mostly Agree	2.00	2.00	2.00	10
Neutral	7.00	7.00	7.00	10
Mostly Disagree	8.00	8.00	8.00	8
Disagree	9.00	9.00	9.00	3

Q.1- Do you agree with the code rationales generated for token [else] highlighted inside the red box? 37 ⓘ



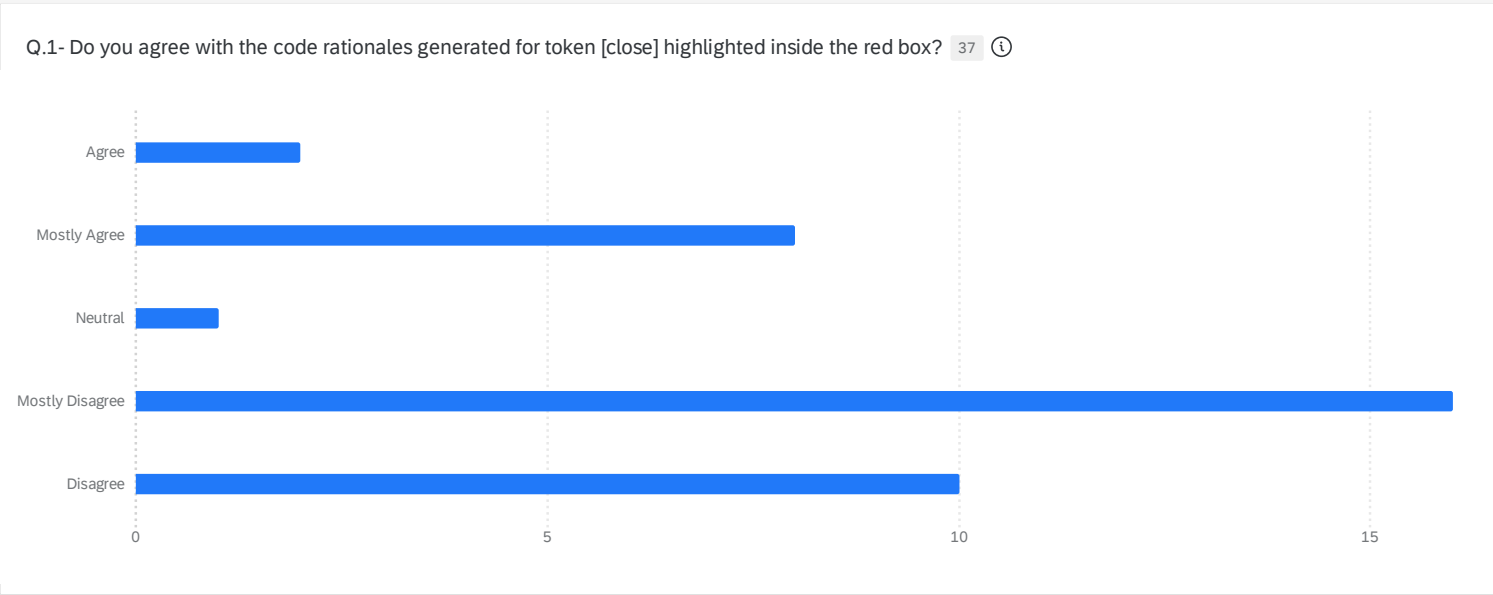
Q.1- Do you agree with the code rationales generated for token [else] highlighted inside the red box? 37 ⓘ

Q320 - Q.1- Do you agree with the code rationales generated for token [else] highlighted inside the red box?	Percentage	Count
Agree	16%	6
Mostly Agree	49%	18
Neutral	8%	3
Mostly Disagree	19%	7
Disagree	8%	3
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [else] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [else] highl...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	6

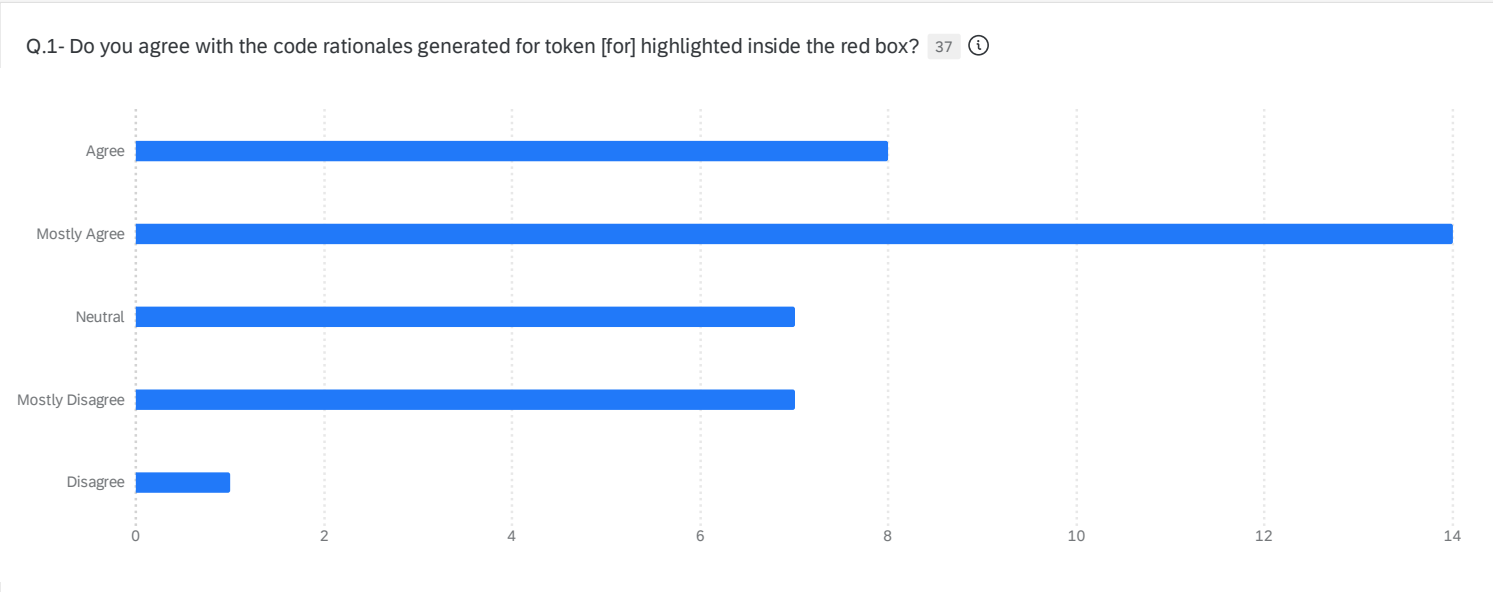
Q.1- Do you agree with the code rationales generated for token [else] highl...	Average	Minimum	Maximum	Count
Mostly Agree	2.00	2.00	2.00	18
Neutral	7.00	7.00	7.00	3
Mostly Disagree	8.00	8.00	8.00	7
Disagree	9.00	9.00	9.00	3



Q.1- Do you agree with the code rationales generated for token [close] highlighted inside the red box? 37 ⓘ				
Q328 - Q.1- Do you agree with the code rationales generated for token [close] highlighted inside the red box?		Percentage	Count	
Agree		5%	2	
Mostly Agree		22%	8	
Neutral		3%	1	
Mostly Disagree		43%	16	
Disagree		27%	10	
Sum		100%	37	

Q.1- Do you agree with the code rationales generated for token [close] highlighted inside the red box? 37 ⓘ				
Q.1- Do you agree with the code rationales generated for token [close] high...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	2
Mostly Agree	2.00	2.00	2.00	8

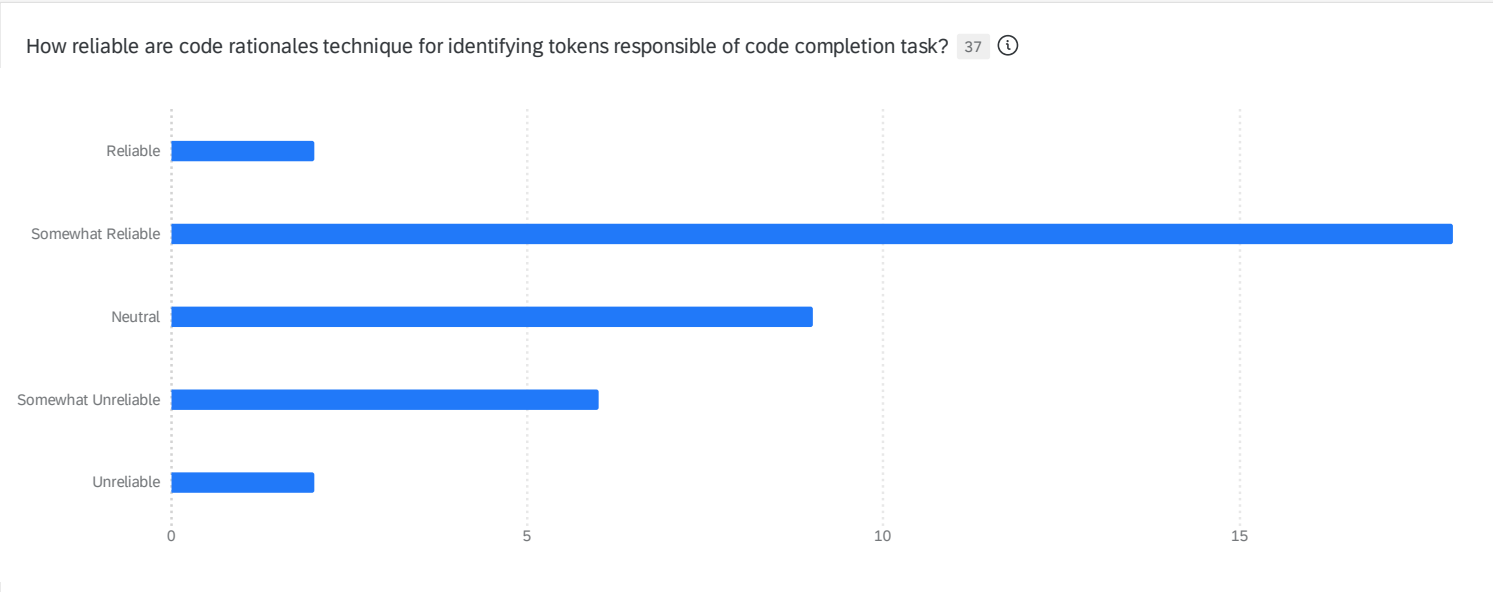
Q.1- Do you agree with the code rationales generated for token [close] high...	Average	Minimum	Maximum	Count
Neutral	7.00	7.00	7.00	1
Mostly Disagree	8.00	8.00	8.00	16
Disagree	9.00	9.00	9.00	10



Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box? 37 ⓘ			
Q332 - Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box?	Percentage	Count	
Agree	22%	8	
Mostly Agree	38%	14	
Neutral	19%	7	
Mostly Disagree	19%	7	
Disagree	3%	1	
Sum	100%	37	

Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box? 37 ⓘ				
Q.1- Do you agree with the code rationales generated for token [for] highli...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	8
Mostly Agree	2.00	2.00	2.00	14
Neutral	7.00	7.00	7.00	7
Mostly Disagree	8.00	8.00	8.00	7

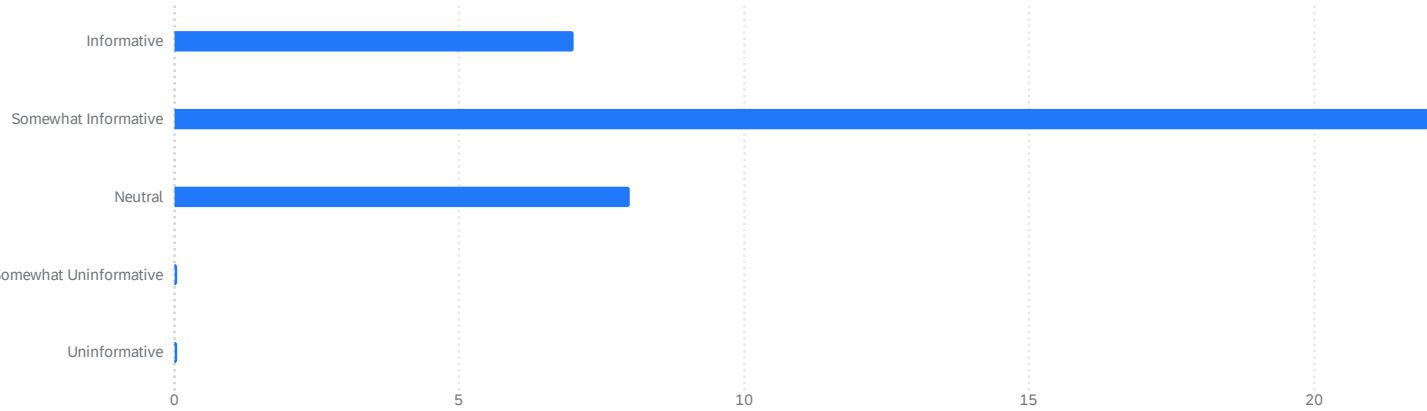
Q.1- Do you agree with the code rationales generated for token [for] highli...	Average	Minimum	Maximum	Count
Disagree	9.00	9.00	9.00	1



How reliable are code rationales technique for identifying tokens responsible of code completion task? 37 ⓘ		
Q123 - How reliable are code rationales technique for identifying tokens responsible of code completion task?	Percentage	Count
Reliable	5%	2
Somewhat Reliable	49%	18
Neutral	24%	9
Somewhat Unreliable	16%	6
Unreliable	5%	2
Sum	100%	37

How reliable are code rationales technique for identifying tokens responsible of code completion task? 37 ⓘ				
How reliable are code rationales technique for identifying tokens responsib...	Average	Minimum	Maximum	Count
Reliable	1.00	1.00	1.00	2
Somewhat Reliable	2.00	2.00	2.00	18
Neutral	5.00	5.00	5.00	9
Somewhat Unreliable	6.00	6.00	6.00	6
Unreliable	3.00	3.00	3.00	2

We proposed a new taxonomy based on AST elements (e.g., punctuation, oop, strings) for grouping tokens, how informative do you find these for code completion? 37 ⓘ



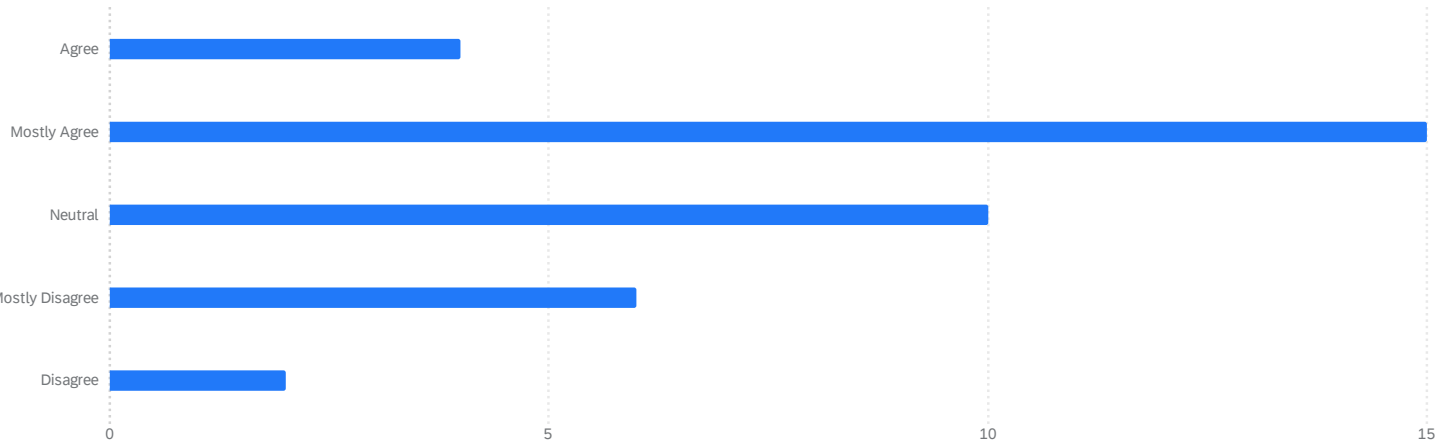
We proposed a new taxonomy based on AST elements (e.g., punctuation, oop, strings) for grouping tokens, how informative do you find these for code completion? 37 ⓘ

Q127 - We proposed a new taxonomy based on AST elements (e.g., punctuation, oop, strings) for grouping tokens, how informative do you find these for code completion?	Percentage	Count
Informative	19%	7
Somewhat Informative	59%	22
Neutral	22%	8
Somewhat Uninformative	0%	0
Uninformative	0%	0
Sum	100%	37

We proposed a new taxonomy based on AST elements (e.g., punctuation, oop, strings) for grouping tokens, how informative do you find these for code completion? 37 ⓘ

We proposed a new taxonomy based on AST elements (e.g., punctuation, oop, s...	Average	Minimum	Maximum	Count
Informative	1.00	1.00	1.00	7
Somewhat Informative	2.00	2.00	2.00	22
Neutral	6.00	6.00	6.00	8
Somewhat Uninformative	-	-	-	0
Uninformative	-	-	-	0

I found this Rational Representation easy to read 37 ⓘ



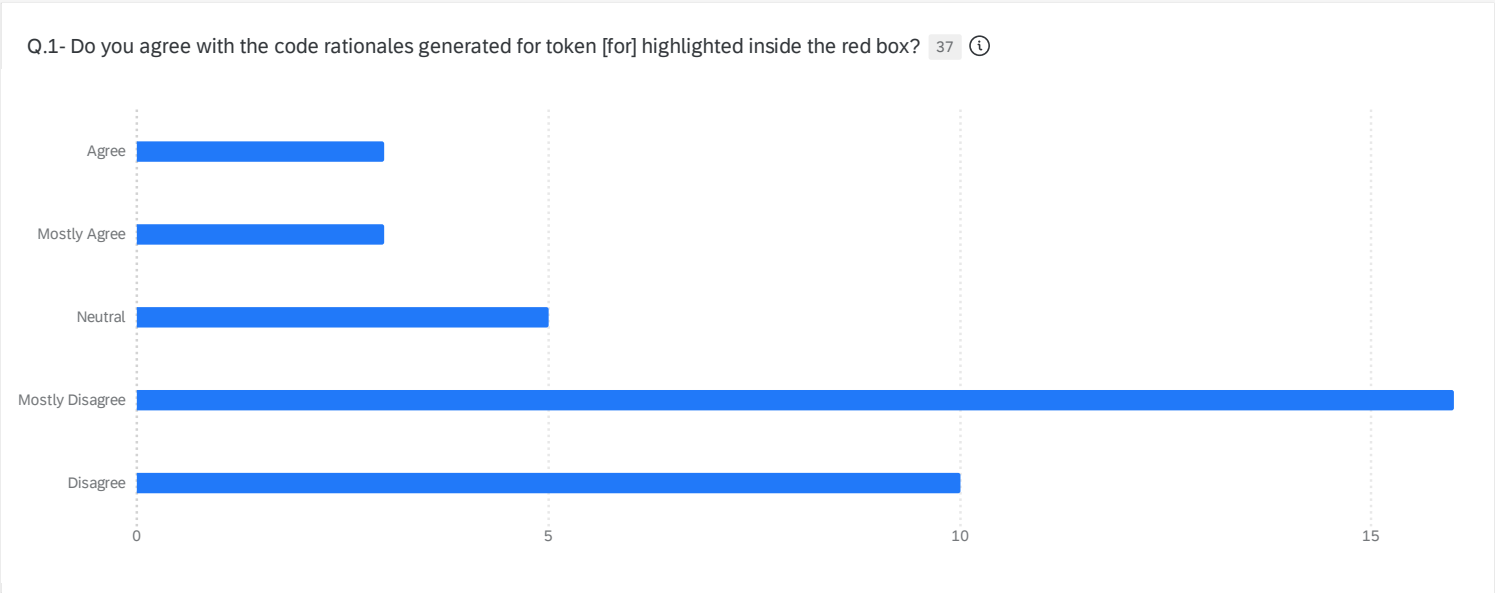
I found this Rational Representation easy to read 37 ⓘ

Q170 - I found this Rational Representation easy to read	Percentage	Count
Agree	11%	4
Mostly Agree	41%	15
Neutral	27%	10
Mostly Disagree	16%	6
Disagree	5%	2
Sum	100%	37

I found this Rational Representation easy to read 37 ⓘ

I found this Rational Representation easy to read	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	4
Mostly Agree	9.00	9.00	9.00	15

I found this Rational Representation easy to read	Average	Minimum	Maximum	Count
Neutral	10.00	10.00	10.00	10
Mostly Disagree	11.00	11.00	11.00	6
Disagree	12.00	12.00	12.00	2



Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box? 37 ⓘ

Q344 - Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box?

	Percentage	Count
Agree	8%	3
Mostly Agree	8%	3
Neutral	14%	5
Mostly Disagree	43%	16
Disagree	27%	10
Sum	100%	37

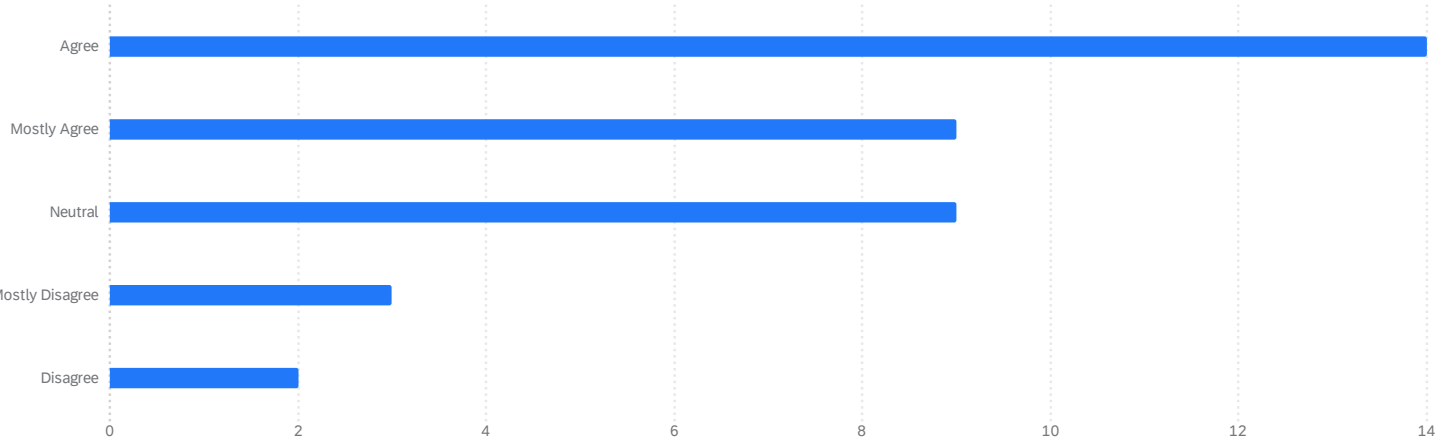
Q.1- Do you agree with the code rationales generated for token [for] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [for] highli...

	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	3
Mostly Agree	6.00	6.00	6.00	3
Neutral	7.00	7.00	7.00	5
Mostly Disagree	8.00	8.00	8.00	16
Disagree	2.00	2.00	2.00	10



Q.1- Do you agree with the code rationales generated for token [cleanUpDescriptors] highlighted inside the red box? 37 ⓘ



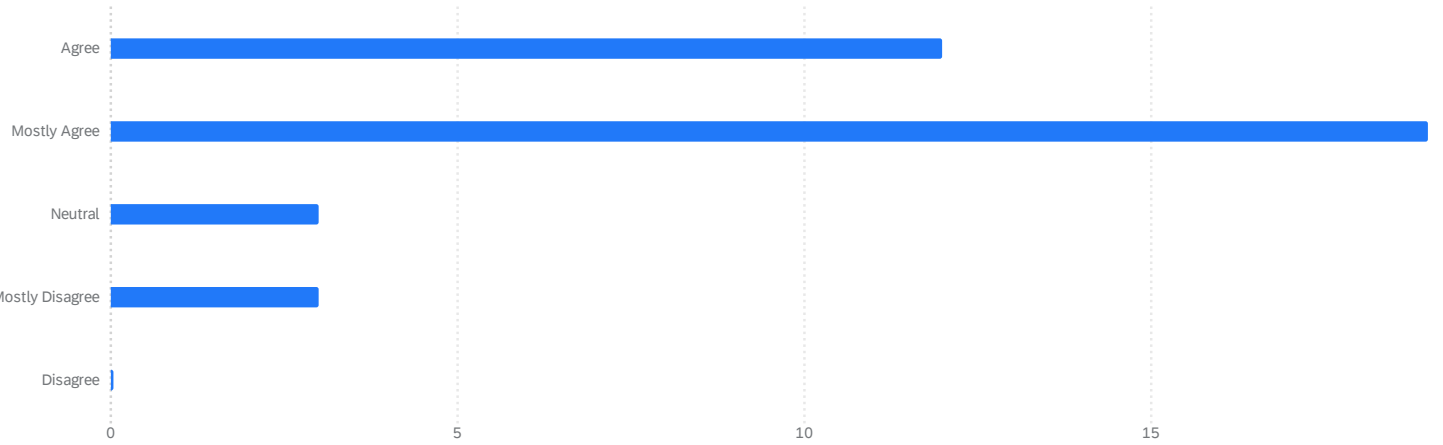
Q.1- Do you agree with the code rationales generated for token [cleanUpDescriptors] highlighted inside the red box? 37 ⓘ

Q352 - Q.1- Do you agree with the code rationales generated for token [cleanUpDescriptors] highlighted inside the red box?	Percentage	Count
Agree	38%	14
Mostly Agree	24%	9
Neutral	24%	9
Mostly Disagree	8%	3
Disagree	5%	2
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [cleanUpDescriptors] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [cleanUpDesc...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	14
Mostly Agree	6.00	6.00	6.00	9
Neutral	7.00	7.00	7.00	9
Mostly Disagree	8.00	8.00	8.00	3
Disagree	2.00	2.00	2.00	2

Q.1- Do you agree with the code rationales generated for token [findlterative] highlighted inside the red box? 37 ⓘ



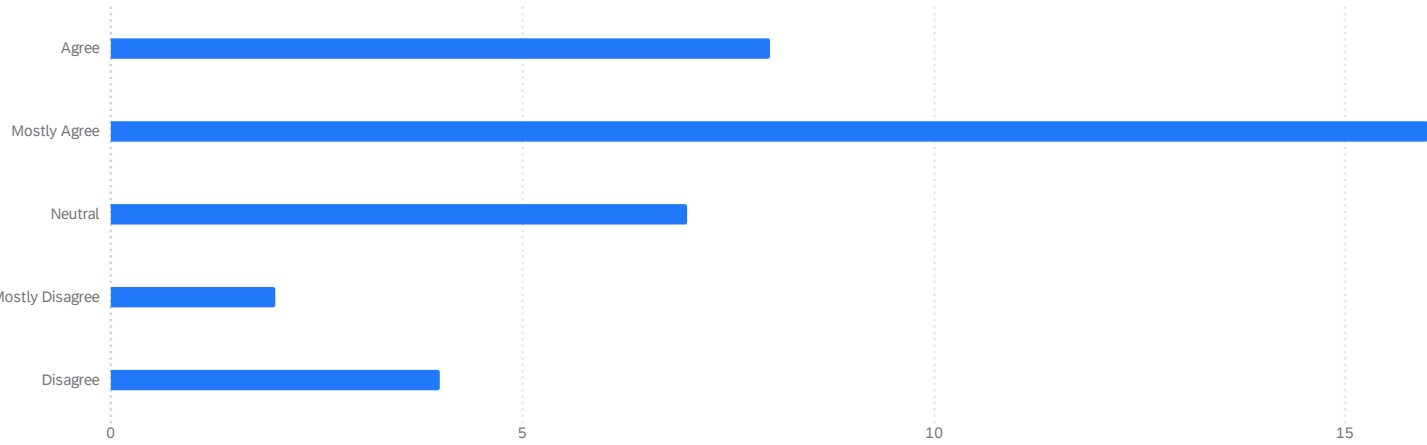
Q.1- Do you agree with the code rationales generated for token [findlterative] highlighted inside the red box? 37 ⓘ

Q356 - Q.1- Do you agree with the code rationales generated for token [findlterative] highlighted inside the red box?	Percentage	Count
Agree	32%	12
Mostly Agree	51%	19
Neutral	8%	3
Mostly Disagree	8%	3
Disagree	0%	0
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [findlterative] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [findlterati...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	12
Mostly Agree	6.00	6.00	6.00	19
Neutral	7.00	7.00	7.00	3
Mostly Disagree	8.00	8.00	8.00	3
Disagree	-	-	-	0

Q.1- Do you agree with the code rationales generated for token [latch] highlighted inside the red box? 37 ⓘ



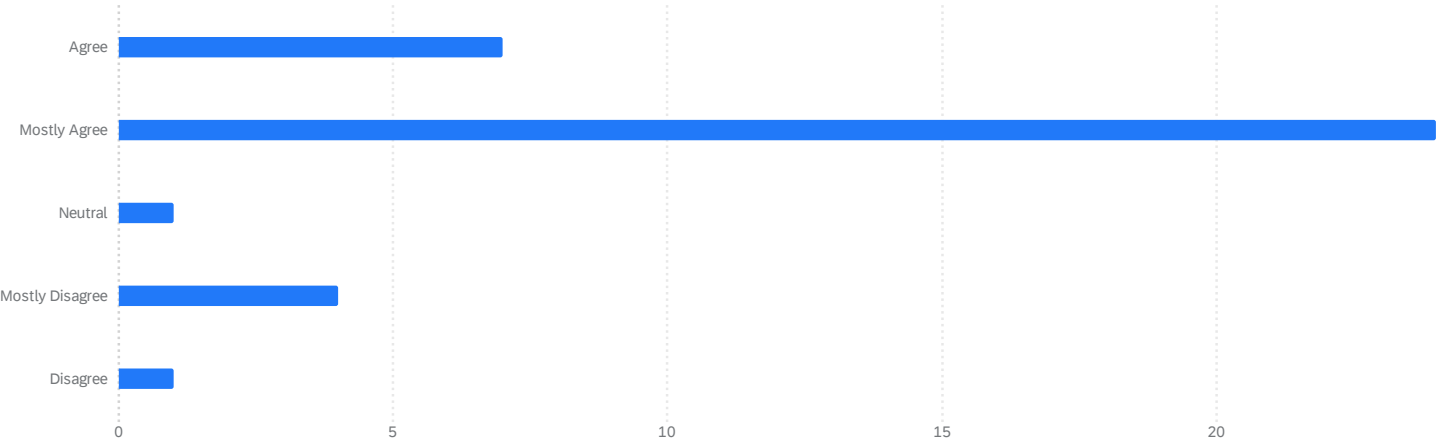
Q.1- Do you agree with the code rationales generated for token [latch] highlighted inside the red box? 37 ⓘ

Q360 - Q.1- Do you agree with the code rationales generated for token [latch] highlighted inside the red box?	Percentage	Count
Agree	22%	8
Mostly Agree	43%	16
Neutral	19%	7
Mostly Disagree	5%	2
Disagree	11%	4
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [latch] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [latch] high...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	8
Mostly Agree	6.00	6.00	6.00	16
Neutral	7.00	7.00	7.00	7
Mostly Disagree	8.00	8.00	8.00	2
Disagree	2.00	2.00	2.00	4

Q.1- Do you agree with the code rationales generated for token [NoSuchElementException] highlighted inside the red box? 37 ⓘ



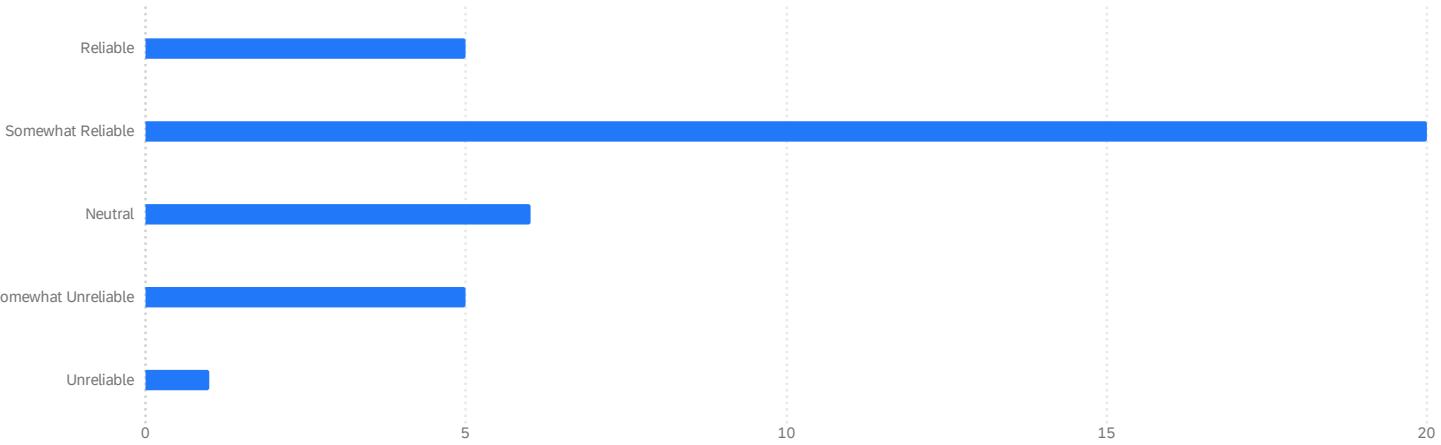
Q.1- Do you agree with the code rationales generated for token [NoSuchElementException] highlighted inside the red box? 37 ⓘ

Q372 - Q.1- Do you agree with the code rationales generated for token [NoSuchElementException] highlighted inside the red box?	Percentage	Count
Agree	19%	7
Mostly Agree	65%	24
Neutral	3%	1
Mostly Disagree	11%	4
Disagree	3%	1
Sum	100%	37

Q.1- Do you agree with the code rationales generated for token [NoSuchElementException] highlighted inside the red box? 37 ⓘ

Q.1- Do you agree with the code rationales generated for token [NoSuchEleme...	Average	Minimum	Maximum	Count
Agree	1.00	1.00	1.00	7
Mostly Agree	6.00	6.00	6.00	24
Neutral	7.00	7.00	7.00	1
Mostly Disagree	8.00	8.00	8.00	4
Disagree	2.00	2.00	2.00	1

How reliable are code rationales technique for identifying tokens responsible of test generation task? 37 ⓘ



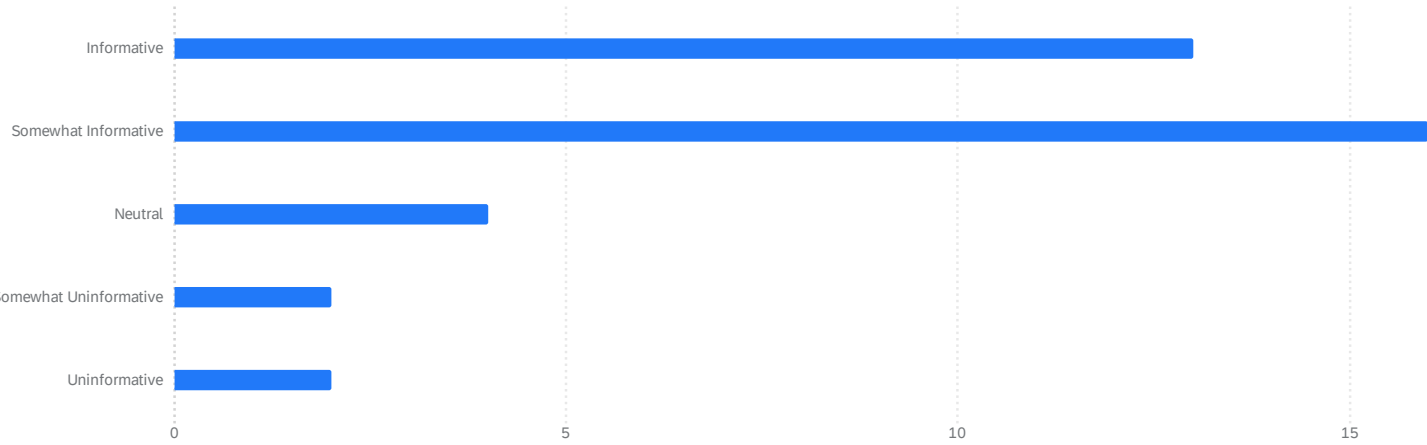
How reliable are code rationales technique for identifying tokens responsible of test generation task? 37 ⓘ

Q220 - How reliable are code rationales technique for identifying tokens responsible of test generation task?	Percentage	Count
Reliable	14%	5
Somewhat Reliable	54%	20
Neutral	16%	6
Somewhat Unreliable	14%	5
Unreliable	3%	1
Sum	100%	37

How reliable are code rationales technique for identifying tokens responsible of test generation task? 37 ⓘ

How reliable are code rationales technique for identifying tokens responsib...	Average	Minimum	Maximum	Count
Reliable	1.00	1.00	1.00	5
Somewhat Reliable	2.00	2.00	2.00	20
Neutral	5.00	5.00	5.00	6
Somewhat Unreliable	6.00	6.00	6.00	5
Unreliable	3.00	3.00	3.00	1

We use context levels (e.g., focal method, signature, fields) for test case generation, how informative do you find these? 37 ⓘ



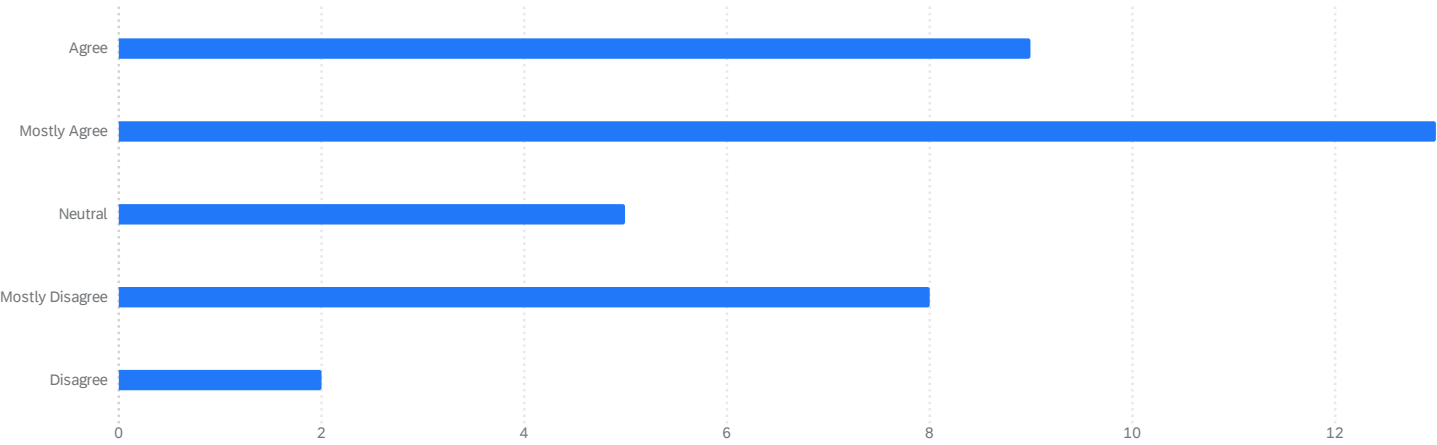
We use context levels (e.g., focal method, signature, fields) for test case generation, how informative do you find these? 37 ⓘ

Q224 - We use context levels (e.g., focal method, signature, fields) for test case generation, how informative do you find these?	Percentage	Count
Informative	35%	13
Somewhat Informative	43%	16
Neutral	11%	4
Somewhat Uninformative	5%	2
Uninformative	5%	2
Sum	100%	37

We use context levels (e.g., focal method, signature, fields) for test case generation, how informative do you find these? 37 ⓘ

We use context levels (e.g., focal method, signature, fields) for test case...	Average	Minimum	Maximum	Count
Informative	1.00	1.00	1.00	13
Somewhat Informative	2.00	2.00	2.00	16
Neutral	6.00	6.00	6.00	4
Somewhat Uninformative	3.00	3.00	3.00	2
Uninformative	5.00	5.00	5.00	2

I found Context level Dependency map of rationals representation easy to read 37 ⓘ



I found Context level Dependency map of rationals representation easy to read 37 ⓘ

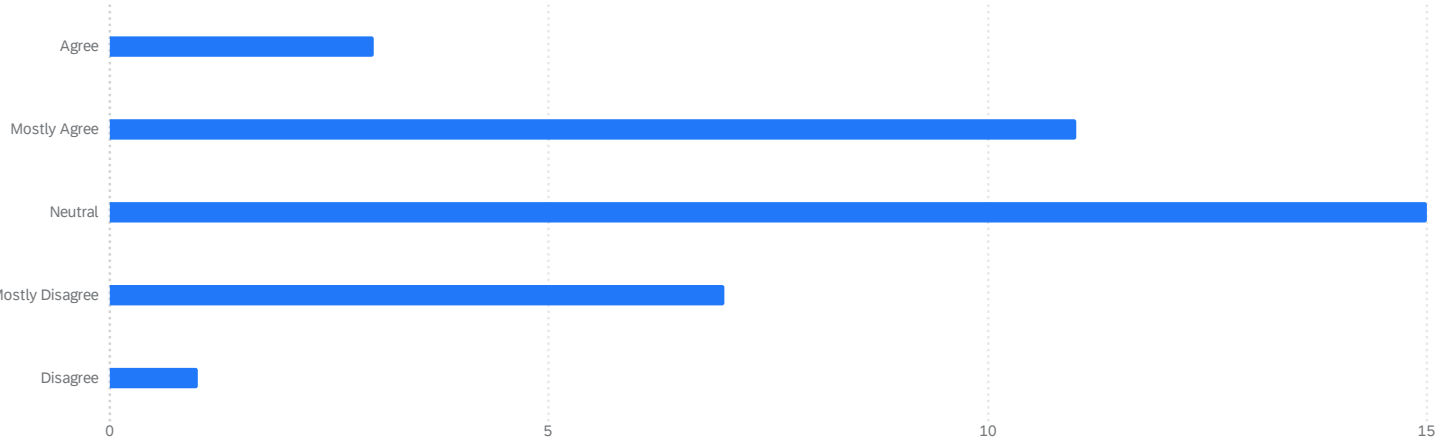
Q222 - I found Context level Dependency map of rationals representation easy to read	Percentage	Count
Agree	24%	9
Mostly Agree	35%	13
Neutral	14%	5
Mostly Disagree	22%	8
Disagree	5%	2
Sum	100%	37

I found Context level Dependency map of rationals representation easy to read 37 ⓘ

I found Context level Dependency map of rationals representation easy to re...	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	9
Mostly Agree	9.00	9.00	9.00	13
Neutral	10.00	10.00	10.00	5
Mostly Disagree	11.00	11.00	11.00	8
Disagree	12.00	12.00	12.00	2



I would rely on concept of code rational to describe prediction of any token for code generation tasks 37 ⓘ

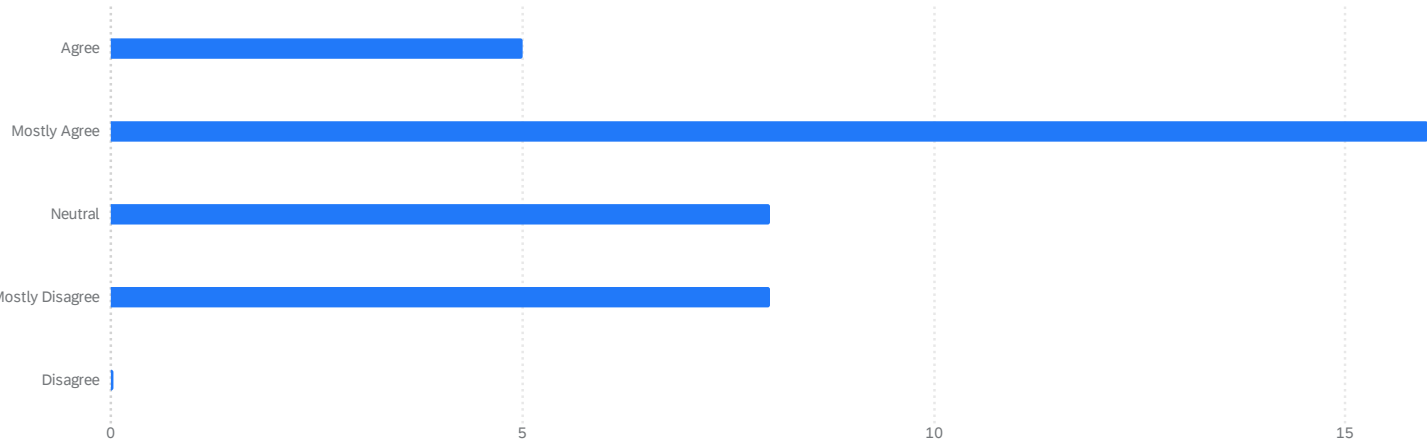


I would rely on concept of code rational to describe prediction of any token for code generation tasks 37 ⓘ

Q238 - I would rely on concept of code rational to describe prediction of any token for code generation tasks	Percentage	Count
Agree	8%	3
Mostly Agree	30%	11
Neutral	41%	15
Mostly Disagree	19%	7
Disagree	3%	1
Sum	100%	37

I would rely on concept of code rational to describe prediction of any token for code generation tasks 37 ⓘ

I would rely on concept of code rational to describe prediction of any token for code generation tasks	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	3
Mostly Agree	9.00	9.00	9.00	11
Neutral	10.00	10.00	10.00	15
Mostly Disagree	11.00	11.00	11.00	7
Disagree	12.00	12.00	12.00	1



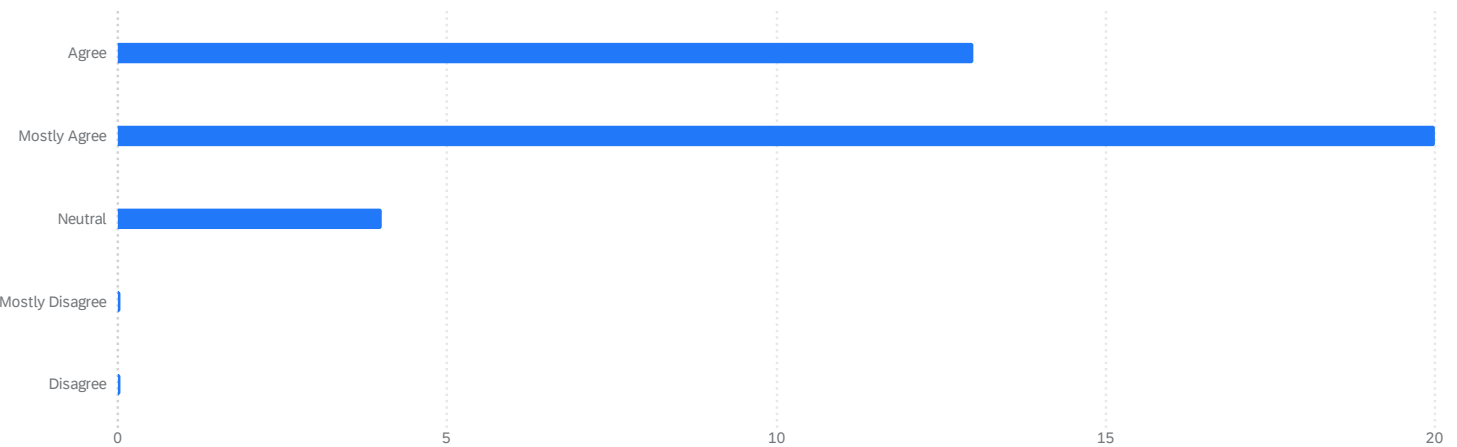
Code Rationales Technique can be used to debug a Neural Code Model or Deep Code Generator 37 ⓘ

Q230 - Code Rationales Technique can be used to debug a Neural Code Model or Deep Code Generator	Percentage	Count
Agree	14%	5
Mostly Agree	43%	16
Neutral	22%	8
Mostly Disagree	22%	8
Disagree	0%	0
Sum	100%	37

Code Rationales Technique can be used to debug a Neural Code Model or Deep Code Generator 37 ⓘ

Code Rationales Technique can be used to debug a Neural Code Model or Deep...	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	5
Mostly Agree	9.00	9.00	9.00	16
Neutral	10.00	10.00	10.00	8
Mostly Disagree	11.00	11.00	11.00	8
Disagree	-	-	-	0

Code Rationales Technique could be helpful in fine-tuning models for specific downstream task 37 ⓘ



Code Rationales Technique could be helpful in fine-tuning models for specific downstream task 37 ⓘ

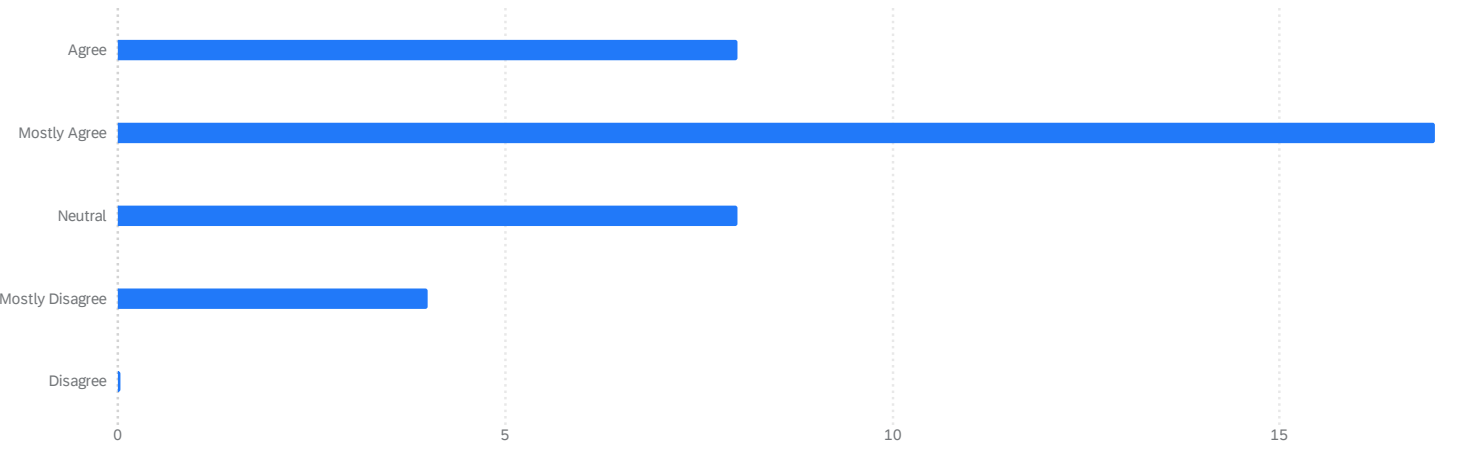
Q231 - Code Rationales Technique could be helpful in fine-tuning models for specific downstream task	Percentage	Count
Agree	35%	13

Q231 - Code Rationales Technique could be helpful in fine-tuning models for specific downstream task	Percentage	Count
Mostly Agree	54%	20
Neutral	11%	4
Mostly Disagree	0%	0
Disagree	0%	0
Sum	100%	37

Code Rationales Technique could be helpful in fine-tuning models for specific downstream task 37 ⓘ

Code Rationales Technique could be helpful in fine-tuning models for specif...	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	13
Mostly Agree	9.00	9.00	9.00	20
Neutral	10.00	10.00	10.00	4
Mostly Disagree	-	-	-	0
Disagree	-	-	-	0

Code Rationales Technique could be helpful in curating training/validation/test sets 37 ⓘ



Code Rationales Technique could be helpful in curating training/validation/test sets 37 ⓘ

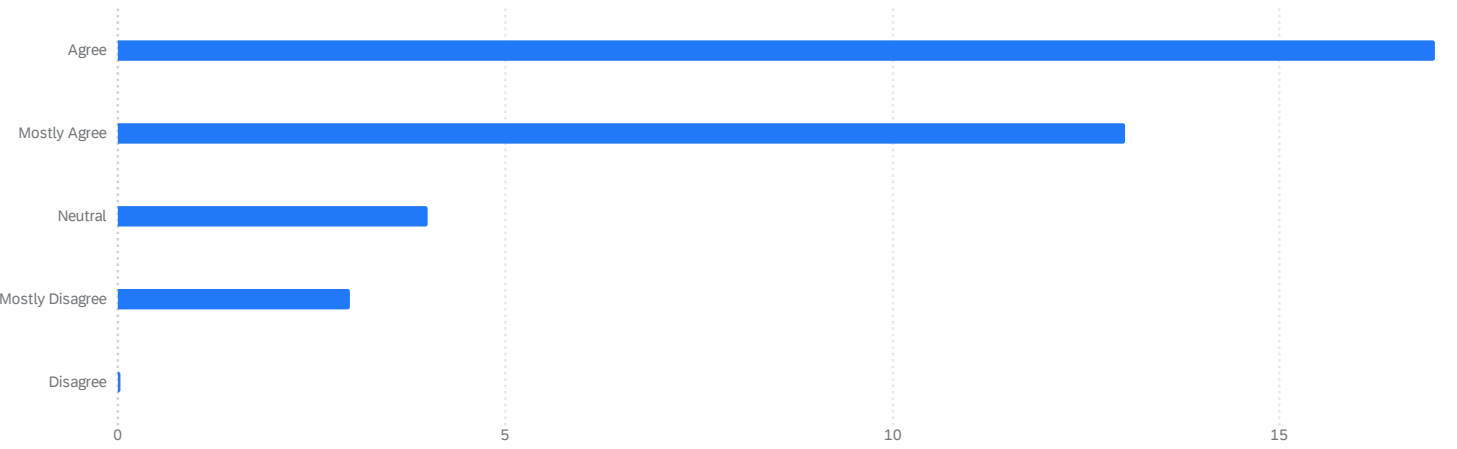
Q232 - Code Rationales Technique could be helpful in curating training/validation/test sets	Percentage	Count
Agree	22%	8
Mostly Agree	46%	17
Neutral	22%	8
Mostly Disagree	11%	4

Q232 - Code Rationales Technique could be helpful in curating training/validation/test sets	Percentage	Count
Disagree	0%	0
Sum	100%	37

Code Rationales Technique could be helpful in curating training/validation/test sets 37 ⓘ

Code Rationales Technique could be helpful in curating training/validation/...	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	8
Mostly Agree	7.00	7.00	7.00	17
Neutral	9.00	9.00	9.00	8
Mostly Disagree	10.00	10.00	10.00	4
Disagree	-	-	-	0

Code Rationales Technique helps infer the causal relationships between inputs and outputs of the model 37 ⓘ



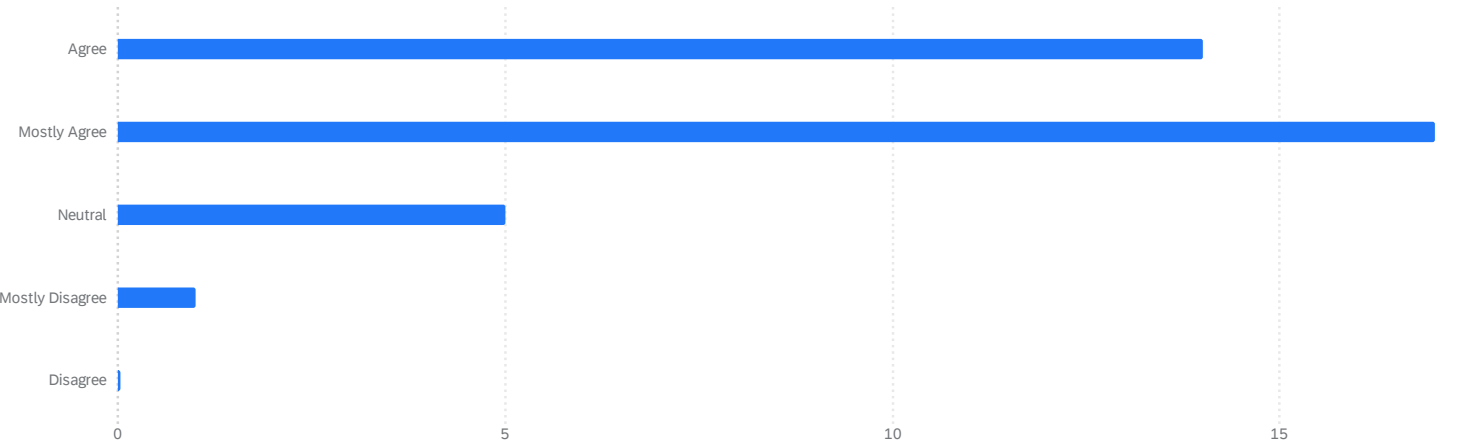
Code Rationales Technique helps infer the causal relationships between inputs and outputs of the model 37 ⓘ

Q233 - Code Rationales Technique helps infer the causal relationships between inputs and outputs of the model	Percentage	Count
Agree	46%	17
Mostly Agree	35%	13
Neutral	11%	4
Mostly Disagree	8%	3
Disagree	0%	0
Sum	100%	37

Code Rationales Technique helps infer the causal relationships between inputs and outputs of the model 37 ⓘ

Code Rationales Technique helps infer the causal relationships between inputs and outputs of the model	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	17
Mostly Agree	9.00	9.00	9.00	13
Neutral	10.00	10.00	10.00	4
Mostly Disagree	11.00	11.00	11.00	3
Disagree	-	-	-	0

Code Rationales Technique is useful to interpret the outcomes of Neural Code Models 37 ⓘ



Code Rationales Technique is useful to interpret the outcomes of Neural Code Models 37 ⓘ

Q239 - Code Rationales Technique is useful to interpret the outcomes of Neural Code Models	Percentage	Count
Agree	38%	14
Mostly Agree	46%	17
Neutral	14%	5
Mostly Disagree	3%	1
Disagree	0%	0
Sum	100%	37

Code Rationales Technique is useful to interpret the outcomes of Neural Code Models 37 ⓘ

Code Rationales Technique is useful to interpret the outcomes of Neural Code Models	Average	Minimum	Maximum	Count
Agree	6.00	6.00	6.00	14

Code Rationales Technique is useful to interpret the outcomes of Neural Cod...	Average	Minimum	Maximum	Count
Mostly Agree	9.00	9.00	9.00	17
Neutral	10.00	10.00	10.00	5
Mostly Disagree	11.00	11.00	11.00	1
Disagree	-	-	-	0

How useful did you find the most about the Code Rationales Technique? ⓘ

Useful for identifying what input data leads to generated code

I found it quite useful as it showed where potential errors and misunderstandings stemmed from.

I thought that it was fairly useful, but somewhat unreliable, at least compared to my interpretation of what input tokens caused certain output tokens. I thought the visualization was very useful for breaking up the different tokens into easy to understand categories.

Not appealing visually for human user but useful for the LLMs

Somewhat, specially given sufficient context

Not perfect, but helpful

Somewhat Useful I would say

I think it is useful in that it helps lay a framework, but can not be totally relied on.

It was useful in the sense that now, if I use an LLM to generate code, I know that I have to think carefully about the tokens I use in my prompt.

I found it useful and informative

its useful to infer or analyze generated codes and get insights about how it was generated

It was useful to see which inputs the llm was picking up on, so very.

Somewhat useful, although I'd like more information about why some rationals were used, although I don't know if that can be provided

Somewhat useful. Shows promise.

I found it helpful in understanding what pieces of the code snippet were being looked at to inform the models decisions

Quite useful

I think it is useful to get a more detailed look at what the model is doing. It would be useful for debugging and understanding the model better.

Slightly useful

Very useful for understanding what the model was "thinking." But it often did not align with what I thought.

somewhat useful

It was pretty useful to know the tokens that caused the output token.

It was useful to infer the potential causality between inputs and outputs of the models.



It appears to be moderately useful at this stage, but there do seem to be some issues with the generated code

---

It is useful to know how the neural network works and whether it is reliable

---

I found it fairly useful, sometimes I think it adds extraneous features to try to fill out the chart where it may only be 1 or 2 code features that are relevant.

---

It was useful to understand the connection between input and output and see areas that could be improved upon.

---

Somewhat useful; During code reviews, rationales can provide valuable context, making the process more efficient.

---

I don't understand the wording of this question ??

---

I think the technique is good for inferring simple decisions that a LLM makes, like how to name a variable or conditional structure. But it seems to fall apart when the request become more complex.

---

The arrows to the boxes of categories.

---

Having an organized, visual representation of what the language model uses to generate its output is helpful for understanding the model & maybe provides insight on how to format your code to maximize useful outputs

---

I found it helpful in understanding how and why neural code models come up with specific outputs based on prompts

---

Mostly useful, yet in some cases its reliability could be improved

---

we can understand why the code has been generated

---

Not quite useful

---

I found the Code Rationale Technique to be useful, and to help fill in the gaps when I was unsure how/why the model would predict what it did.

---

Most function calls, exceptions, and conditionals within the created code were decided by the syntax of python and java, but also natural language in the prompt. I found code rationales to be fairly useful.

Larger text size for larger rationales

---

Maybe show why one symbol is associated with another. E.g., why is ) and ; associated with return or def?

---

Higher accuracy, perhaps an indication of confidence in the accuracy of the prediction.

---

N/A

---

perhaps looking from a higher level, instead of one token, probably collection of tokens jointly influenced a prediction.

---

Better use of colors to assist in correlating result to snippets.

Well what if the model pays attention to specific sections of the AST rather than just tokens?

I am not 100% sure how to improve it, but I would think surveys like this one would help to give the Neural Network more information and data to improve on.

As of now, I think the technique is simple enough to be understood quickly, while still providing a good amount of information on why an LLM made a certain decision.

If it was more accurate it could definitely be applied to more applications

is difficult to ascertain the causality of the generated tokens with just code snippets a large language model intakes large ammount of data to infer or predict code, just a litte piece ins not enoguh, maybe use a heatmap to relate the analized tokens and the generated ones, it would be easier to understand

It should show more of what is using as inputs, as I think it may be leaving a few of them out.

If it possible, provide more info about why a specific set of tokens constitutes a rational and how rationals were used to produce the next token (although this may be more me needing to learn about how LLMs and NLP work)

Incorporating more Tokens as justification, perhaps with weights corresponding to how strong particular tokens influence the output over others

An explanation on why certain tokens might appear under multiple categories or some way of understanding why the model is choosing to look at certain tokens specifically.

Indicate how significant each previous token was to the generated one compared to other tokens.

Perhaps more info can be learned by showing exactly out how much influence each rationale had in the generation.

A better model to make sure the input tokens get mapped to the correct type of component.

Perhaps an interactive interface where when you scroll over any given word, it automatically highlights the most important token or two in the prediction of the scrolled-over token.

use natural language to provide human readable rationales

I can't think of anything right now.

I think teaching developers to reason or link different code/NL snippets to code rationale components will be useful

Perhaps it needs to take into account more of the input/output data before making a prediction

I think not all related tokens are predicted, and the results are not complete. May I ask whether you leverage attention to find the code rationales and which layer do you leverage?

I think it sometimes picks syntax features/keywords that are unrelated

More divisions and focus for the more ambiguous cases.

It seems they are logical. May be there should be a way to receive user feedback and then update with iterative improvement.

Not really sure, honestly

More "levels" need to be considered. The LLM is too intricate to be broken into 5, in my opinion.

---

Slightly better UI perhaps, but no issue with the technique itself

---

I don't know

---

I would have liked brief explanations of why certain things were not included as rationals

---

It does not seem comment/Javadoc is taken into account. Unless it is pruned when training the generative model, it also determines the next tokens IMHO.

---

now the token-level is too fine-grained, a more coarser-grain level may be easier to understand

---

NA

---

I am unsure how it could be improved. Perhaps if the technique required slightly less proficiency in the specific language that it is being used on it could be improved (for example, for a non-fluent English speaker it may be hard to parse what words are officially nouns).

---

The code rationales technique can be improved with better prompting. The more that the AI can go off of, the more accurate the code could be, and more prompting means more chances for code rationales to be corrected.