Survey On Open Source Software Contributions Code Contribution Vs Non-code Contributions

Swetha Jayapathy
Oregon State University

1. INTRODUCTION

The contributions in open source projects are typically characterized by the number of lines of code. This leads to two buckets of contribution types in OSS: code contributions (ex: code related commits) and non-code contributions (ex: nonrelated commits). The contributions live on the sidelines, but they are equally, if not more valuable than their code counterparts. Non-code contributions are given less recognition than code contributions in OSS communities though they play a significant role in the evolution of open source projects. The goal of this work is to gain an understanding of the impact of the non-code contribution to OSS by comparing them to code contributions.

As part of my research on contributions to OSS, I conducted an exploratory online survey to analyze the following research questions. I felt a survey would be appropriate for this research as it would help to reach out to the actual OSS contributors to get their thoughts and opinions. It is useful in analyzing the characteristics of a large population. Moreover, it helps to describe and explore variables and constructs of interest. Specifically, the questions I am trying to answer are

RQ1 What is the frequency of code and non-code contributions in OSS projects over the past 5 years?

RQ2 What are the components of non-code contributions and what are the key non-code contributions which impact OSS?

RQ3 How does gender have an effect on the type of contribution being made in OSS projects?

2. METHODOLOGY

The Survey was conducted online through Qualtrics software. The Survey was of 12 questions which consisted of demographic questions and specific questions needed to answer the research questions. The initial Survey design was evaluated by Rupika Dikkala, one of the students in the class and the Survey questions were reframed according to the feedback from her (Appendix A). The peer-evaluator mentioned that the survey introduction was in detail and that there are chances of the participant to feel biased with certain information. Therefore, the survey introduction was changed to a brief shorter one. She mentioned that the gender-related question did not have a text box for self-describe, which was fixed later. She stated that a participant might not be aware of what a non-code contribution might be and asked to include certain examples in the question. To fix this, examples such as Comments, issues and mentoring were provided in the question itself. Moreover, she helped in revising the wordings of a few questions so that the participant might get a clear understanding.

The survey questions were then corrected according to the feedback and the survey was deployed in Qualtrics. Refer to **Appendix A** for the survey URL. It was then distributed to people in the IT industry and to the students in the class.

Due to the short timeline, in order to get a quick response to the survey, the students were given points for participating in the survey. The students were asked to create a persona with a suitable background such as having experience in OSS, as a developer,

etc., using the hubspot website. They were asked to channel their respective persona while participating in the survey.

3. Survey Results

The survey was distributed to 27 people online and all of them participated in the survey leading to a 100% survey response. The majority of the participants were from class (60%). The participants were from 3 different countries - the US, China and India. The majority of the participants were from the US. Among the 27 participants, 7 replied that they had not worked with open source software and therefore, they were not eligible to participate. Among the rest of the 20 participants, 16 were Men and 10 were Women. 80% of them were of the age group 25 to 34. Most of them replied that they contributed to OSS because of self-interest to enhance their coding skills and to enhance their reputation.

For the first research question, the results of the survey imply that 78% of the participants majorly contributed to the code, whereas only 22% of people responded that their major contribution was non-code contribution (Figure 1).

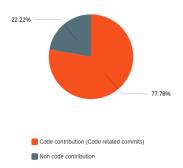


Figure 1. Major type of contribution of the survey participants

The survey included Likert scale questions to measure the frequency of the code and non-code contribution. This showed that 67% of the people contributed to code commits a

few times a month and 15% a few times a week. Whereas 56% of people contributed to non-code contribution a few times a month and 33% of people contributed a few times a week (Figure 2). The weekly contribution percentage is double for the non-code contribution when compared to code contributions.

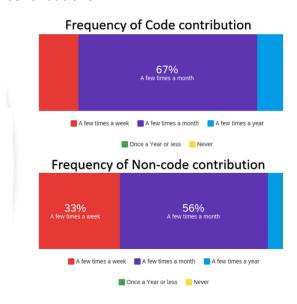


Figure 2 : Frequency of code contribution Vs non-code contribution

For the second research question regarding the types of non-code contributions, I have asked the participants to select the various non-code contributions which they contribute such as comments, reporting an issue, mentoring, code review, etc. 65% of the participants selected comments, reporting an issue and Testing as their most contributed one. 60% of the participants selected the code review (Figure 3).

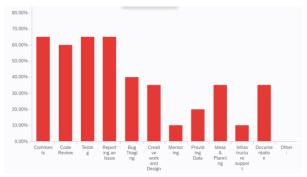


Figure 3: Most often contributed non-code contributions

Looking at Figure 3, we could see that Comments, reporting issues, code review and Testing are the most often contributed noncode contributions.

With respect to the third research question, the effect of gender on the type of contribution, I analyzed the gender of the participants and their respective survey results separately. 62% of the participants were men and 38% of them were women (Figure 4).

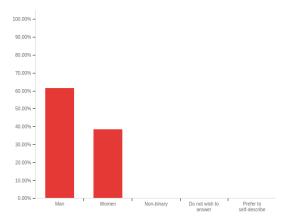


Figure 4: Percentage of Men and Women participants

When the major type of contribution is analyzed for men and women separately, we see that women contribute more toward noncode contribution from Figure 5. Only 10% of men have mentioned their major contribution as non-code contribution, whereas 38% of women have mentioned non-code contribution as their major type.

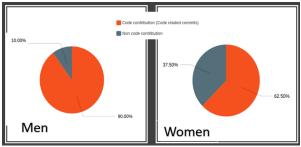


Figure 5 – Major type of contribution with respect to gender

The distribution of code contribution (Figure 6) among men and women has a slight difference. 80% of men has responded that they contribute a few times a month, whereas 50% of women has responded they contribute a few times a month and 38% of women responded a few times a week. Overall, the frequency of code contribution among men and women looks almost similar.

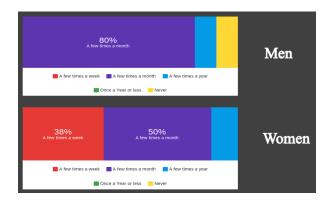


Figure 6: Distribution of code contribution among men and women

Looking at the distribution of non-code contribution among men and women, 60% of men responded they contribute a few times a month and 20% of men responded they contribute a few times a week. Comparatively, we could see from Figure 7 that the frequency of non-code contribution among women is higher. 50% of women responded that they

contribute a few times a month and 50% responded that they contribute a few times a week.



Figure 7: Distribution of non-code contribution among men and women

4. Conclusion

The goal of this work is to analyze the different types of contributions to OSS and show that non-code contribution is as important as code contribution. The survey results have helped to analyze this data and we could infer the answers to our research questions from it. For the first research regarding the comparison of frequency distribution among code and non-code contributions, we could see that the overall frequency of non-code contributions is higher than the code contribution. Thus we could conclude that although code contribution is the major type of contribution to OSS, the non-code contribution is the most frequent type of contribution.

The answer to the second research question is quite clear from the survey response that is the most frequent types of non-code contributions are comments, code review, reporting issues, mentoring, ideas & planning, bug triaging, infrastructure and support and documentation. The major type of non-code contribution resulted in comments, reporting issues, code review and testing. We could see from the survey results

that the answer to the third research question is women tend to contribute more frequently towards non-code contributions compared to men and that the percentage of women for the code-contribution per week is slightly higher than the percentage of men.

Threats to Validity

My study reveals many threats to validity. Internal validity: The survey had a relatively low number of participants from a self-selected population. Since the majority of the participants were from class, we could not completely rely on the results. Therefore, it seems to be biased. The survey needs to be circulated to a wide range of populations and to people from different backgrounds to infer the results. Moreover, the result of the survey alone cannot conclude the answers to the research questions. They mostly reflect the opinions of the population which might be biased, we need to have mixed analysis to come to a concrete conclusion. The study design needs to have multiple analysis. External validity: As the sample population is very less and biased, the survey cannot represent the actual population.

Suggestions:

The research could include mining of OSS data and analyzing it for good results, as the history of data tells us the actual story. It can analyze 3 to 4 projects which are highly active currently and mine the data related to code and non-code contributions from it. Correlation tests can be performed to find whether there is a relationship between the data. This would give a strong conclusion.

The study can also include a literature review which could clearly show some valid ground truth that the study cannot alone come up with. This minimizes the biases and errors inherent to traditional and narrative reviews.

Appendix A

- Survey Link –
 https://oregonstate.qualtrics.com/jfe/form/SV_9zdqrnTaQOExfqR
- 2. See the attached file for Survey Critique

Survey Designer Name: Swetha Jayapathy

Survey Link: https://oregonstate.qualtrics.com/jfe/form/SV_9zdqrnTaQOExfq R Google Doc Link

to RQ & Data:

https://docs.google.com/document/d/1N9rhDR-dAoKrMZmdNBPat17y5DA-59ZDCMuOrJc WpUU/edit?usp=sharing

Link to Persona (if already created): Survey Critique by: Rupika Dikkala

Instructions:

- 1. By the end of the day upload your critique to Canvas (Survey Design Critique)
- 2. Time breakdown: 30 min to review; 30 min each giving feedback
- 3. Please pay attention to **best practices** when critiquing the survey:
 - a. Survey results answers research questions
 - b. Front matter (goal of work, how long it takes to complete, confidentiality etc)
 - c. Clear instructions for filling out the survey
 - d. Wording of the questions (unambiguous, unbiased, concrete/past tense, avoids negative)
 - e. Framing of the answer options (avoid overlap, ranges, NA/dont know options etc.)
 - f. Ordering of questions (blocks, where are sensitive questions placed)

Critique

- In the description part: try not to give too much away about the study or it might bias the user who is taking it
- Consent looks great, maybe include more info Gender question
 - "Woman" instead of "women"
 - Put non-binary after "woman" o
 Include a text box for "self-describe"
- Country question have them fill out the country?
- Age question
 - Say "24 or under" instead of "less than"

- O Spacing for 25 to 34
- o Make it 35 to 44
- Then make it "45 or above"
- Give example for code/non code contribution Have you contributed to OSS question
 - If they say "no", end the survey
- What kind of projects do you spend most of your time on?
 - Keep OSS consistent with open source software o Give examples of closed source projects
 - Maybe change the wording to individual project For "N/A" option
 - Can direct the next few questions towards people who have an idea of how OSS works even if they don't use it
- Which is your majority of contribution towards? -- change the wording of this Select the reason for your contribution to OSS
 - Say "other reasons" instead of "any other reasons"
- Rank your amount of code contribution from 1 being the lowest to 5 being the highest?
 - Rank the "frequency"
 - Use the frequency of use scale
- Select the non-code contribution which you contribute mostly, select all that apply. O Change mostly to "most often"
- Do you think non-code contribution is equally important to code-contribution?
 - Reorder the options ("strongly agree" comes before "agree") How often do you contribute towards code commits?
 - O Change "code commits" to code contributions

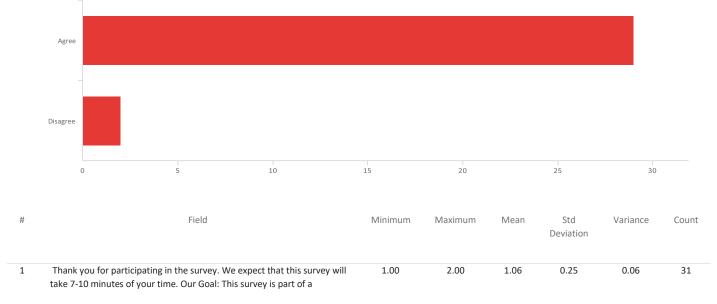
Appendix B

1.See the attached file for the raw data of the Survey

Default Report

Survey on Code Vs Non-code contribution to Open Source Software November 19, 2020 12:08 AM MST

Thank you for participating in the survey. We expect that this survey will take 7-10 minutes of your time. Our Goal: This survey is part of a research study to examine the contributions to Open Source Software (OSS). Our goal is to understand the importance of various kinds of contributions and their impact on OSS. Your personal information will kept as confidential and no individually identifying data will ever be released. If you wish to proceed with the Survey, please agree to begin.



research study to examine the contributions to Open Source Software (OSS). Our goal is to understand the importance of various kinds of contributions and their impact on OSS. Your personal information will

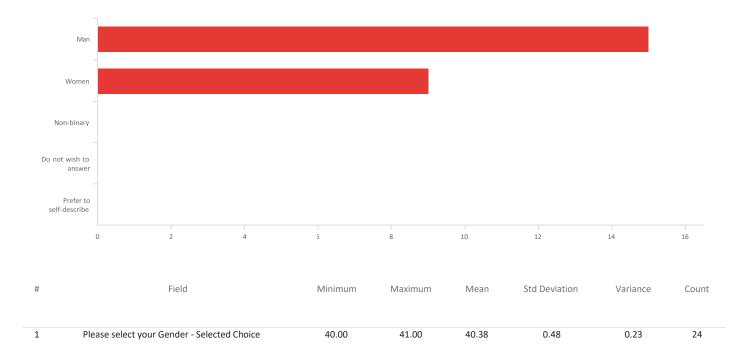
kept as confidential and no individually identifying data will ever be released. If you wish to proceed with the Survey, please agree to begin.

#	Field	Ch Cr	noice ount	
1	Agree	93.55	5%	29
2	Disagree	6.4	.5%	2
		Choice		

Field

Showing rows 1 - 3 of 3

Q1 - Please select your Gender



#	Field	Choice Coun	
40	Man	62.50%	15
41	Women	37.50%	9
42	Non-binary	0.00%	0
43	Do not wish to answer	0.00%	0
44	Prefer to self-describe	0.00%	0
			24

Q1_44_TEXT - Prefer to self-describe

Prefer to self-describe

Q2 - List of Countries

Bhutan Bolivia Bosnia and Herzegovina Botswana Brazil Brunei Darussalam Bulgaria Burkina Faso Burundi Cambodia Cameroon Canada Cape Verde Central African Republic Chad Chile China Colombia Comoros

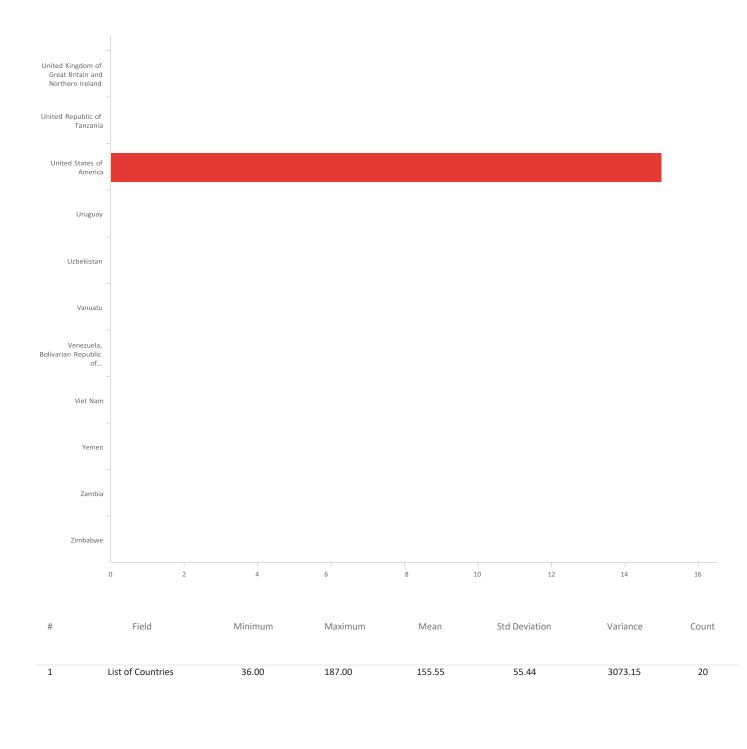
Congo, Republic of the...

France Gabon Gambia Georgia Germany Ghana Greece Grenada Guatemala Guinea Guinea-Bissau Guyana Haiti Honduras Hong Kong (S.A.R.) Hungary Iceland India Indonesia

Iran, Islamic Republic of...

Iraq

1 -	



Field Choice Count

1 Afghanistan 0.00% 0

2 Albania	0.00% 0
3 Algeria	0.00% 0
4 Andorra	0.00% 0
5 Angola	0.00% 0
6 Antigua and Barbuda	0.00% 0

#	Field	Choic Coun	
7	Argentina	0.00%	0
8	Armenia	0.00%	0
9	Australia	0.00%	0
10	Austria	0.00%	0
11	Azerbaijan	0.00%	0
12	Bahamas	0.00%	0
13	Bahrain	0.00%	0
14	Bangladesh	0.00%	0
15	Barbados	0.00%	0
16	Belarus	0.00%	0
17	Belgium	0.00%	0
18	Belize	0.00%	0
19	Benin	0.00%	0
20	Bhutan	0.00%	0
21	Bolivia	0.00%	0
22	Bosnia and Herzegovina	0.00%	0
23	Botswana	0.00%	0
24	Brazil	0.00%	0
25	Brunei Darussalam	0.00%	0
26	Bulgaria	0.00%	0
27	Burkina Faso	0.00%	0
28	Burundi	0.00%	0
29	Cambodia	0.00%	0
30	Cameroon	0.00%	0
31	Canada	0.00%	0
32	Cape Verde	0.00%	0
33	Central African Republic	0.00%	0

34	Chad	0.00%	0
35	Chile		
#	Field	Choic Coun	
36	China	10.00%	2
37	Colombia	0.00%	0
38	Comoros	0.00%	0
39	Congo, Republic of the	0.00%	0
40	Costa Rica	0.00%	0
41	Côte d'Ivoire	0.00%	0
42	Croatia	0.00%	0
43	Cuba	0.00%	0
44	Cyprus	0.00%	0
45	Czech Republic	0.00%	0
46	Democratic People's Republic of Korea	0.00%	0
47	Democratic Republic of the Congo	0.00%	0
48	Denmark	0.00%	0
49	Djibouti	0.00%	0
50	Dominica	0.00%	0
51	Dominican Republic	0.00%	0
52	Ecuador	0.00%	0
53	Egypt	0.00%	0
54	El Salvador	0.00%	0
55	Equatorial Guinea	0.00%	0
56	Eritrea	0.00%	0
57	Estonia	0.00%	0
58	Ethiopia	0.00%	0
59	Fiji	0.00%	0
60	Finland	0.00%	0

61	France	0.00%	0
62	Gabon	0.00%	0
63	Gambia	0.00%	0
64	Georgia	0.00%	0
#	Field	Choic Coun	
65	Germany	0.00%	0
66	Ghana	0.00%	0
67	Greece	0.00%	0
68	Grenada	0.00%	0
69	Guatemala	0.00%	0
70	Guinea	0.00%	0
71	Guinea-Bissau	0.00%	0
72	Guyana	0.00%	0
73	Haiti	0.00%	0
74	Honduras	0.00%	0
75	Hong Kong (S.A.R.)	0.00%	0
76	Hungary	0.00%	0
77	Iceland	0.00%	0
78	India	15.00%	3
79	Indonesia	0.00%	0
80	Iran, Islamic Republic of	0.00%	0
81	Iraq	0.00%	0
82	Ireland	0.00%	0
83	Israel	0.00%	0
84	Italy	0.00%	0
85	Jamaica	0.00%	0
86	Japan	0.00%	0
87	Jordan	0.00%	0

88	Kazakhstan	0.00%	0
89	Kenya	0.00%	0
90	Kiribati	0.00%	0
91	Kuwait	0.00%	0
92	Kyrgyzstan	0.00%	0
93	Lao People's Democratic Republic		
#	Field	Choice	
94	Latvia	0.00%	0
95	Lebanon	0.00%	0
96	Lesotho	0.00%	0
97	Liberia	0.00%	0
98	Libyan Arab Jamahiriya	0.00%	0
99	Liechtenstein	0.00%	0
100	Lithuania	0.00%	0
101	Luxembourg	0.00%	0
102	Madagascar	0.00%	0
103	Malawi	0.00%	0
104	Malaysia	0.00%	0
105	Maldives	0.00%	0
106	Mali	0.00%	0
107	Malta	0.00%	0
108	Marshall Islands	0.00%	0
109	Mauritania	0.00%	0
110	Mauritius	0.00%	0
111	Mexico	0.00%	0
112	Micronesia, Federated States of	0.00%	0
113	Monaco	0.00%	0
114	Mongolia	0.00%	0

115	Montenegro	0.00%	0
116	Morocco	0.00%	0
117	Mozambique	0.00%	0
118	Myanmar	0.00%	0
119	Namibia	0.00%	0
120	Nauru	0.00%	0
121	Nepal	0.00%	0
122	Netherlands	0.00%	0
#	Field	Choic Coun	
123	New Zealand	0.00%	0
124	Nicaragua	0.00%	0
125	Niger	0.00%	0
126	Nigeria	0.00%	0
127	North Korea	0.00%	0
128	Norway	0.00%	0
129	Oman	0.00%	0
130	Pakistan	0.00%	0
131	Palau	0.00%	0
132	Panama	0.00%	0
133	Papua New Guinea	0.00%	0
134	Paraguay	0.00%	0
135	Peru	0.00%	0
136	Philippines	0.00%	0
137	Poland	0.00%	0
138	Portugal	0.00%	0
139	Qatar	0.00%	0
140	Republic of Korea	0.00%	0
141	Republic of Moldova	0.00%	0

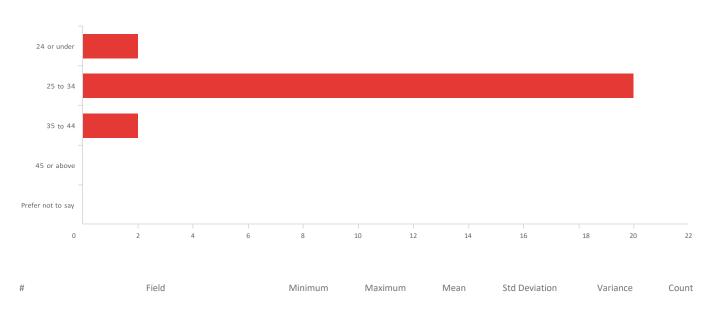
142	Romania	0.00%	0
143	Russian Federation	0.00%	0
144	Rwanda	0.00%	
145	Saint Kitts and Nevis	0.00%	
146	Saint Lucia	0.00%	0
147	Saint Vincent and the Grenadines	0.00%	0
148	Samoa	0.00%	0
149	San Marino	0.00%	0
150	Sao Tome and Principe	0.00%	0
151	Saudi Arabia		
#	Field	Choice	
152	Senegal	0.00%	0
153	Serbia	0.00%	0
154	Seychelles	0.00%	0
155	Sierra Leone	0.00%	0
156	Singapore	0.00%	0
157	Slovakia	0.00%	0
158	Slovenia	0.00%	0
159	Solomon Islands	0.00%	0
160	Somalia	0.00%	0
161	South Africa	0.00%	0
162	South Korea	0.00%	0
163	Spain	0.00%	0
164	Sri Lanka	0.00%	0
165	Sudan	0.00%	0
166	Suriname	0.00%	0
167	Swaziland	0.00%	0
168	Sweden	0.00%	0

169 Switzerland		0.00%	0
170 Syrian Arab Republic		0.00%	0
171 Tajikistan		0.00%	0
172 Thailand		0.00%	0
173 The former Yugoslav Repub	lic of Macedonia	0.00%	0
174 Timor-Leste		0.00%	0
175 Togo		0.00%	0
176 Tonga		0.00%	0
177 Trinidad and Tobago		0.00%	0
178 Tunisia		0.00%	0
179 Turkey		0.00%	0
180 Turkmenistan		0.00%	0

#	Field	Choice Coun	
181	Tuvalu	0.00%	0
182	Uganda	0.00%	0
183	Ukraine	0.00%	0
184	United Arab Emirates	0.00%	0
185	United Kingdom of Great Britain and Northern Ireland	0.00%	0
186	United Republic of Tanzania	0.00%	0
187	United States of America	75.00%	15
188	Uruguay	0.00%	0
189	Uzbekistan	0.00%	0
190	Vanuatu	0.00%	0
191	Venezuela, Bolivarian Republic of	0.00%	0
192	Viet Nam	0.00%	0
193	Yemen	0.00%	0
580	Zambia	0.00%	0
1357	Zimbabwe	0.00%	0
			20

Showing rows 1 - 196 of 196

Q3 - Select the range in which your age falls



#	Field	Choice Coun	
1	24 or under	8.33%	2
2	25 to 34	83.33%	20
3	35 to 44	8.33%	2
4	45 or above	0.00%	0

3.00

2.00

0.41

0.17

24

24

1.00

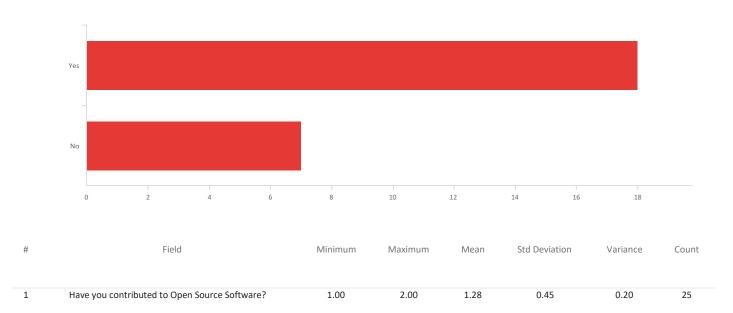
Prefer not to say 0.00% 0

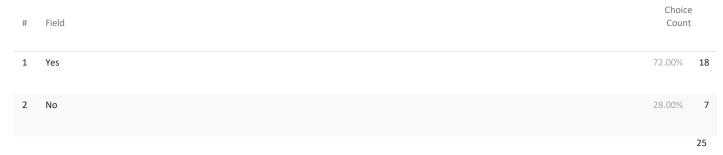
Showing rows 1 - 6 of 6

Q4 - Have you contributed to Open Source Software?

1

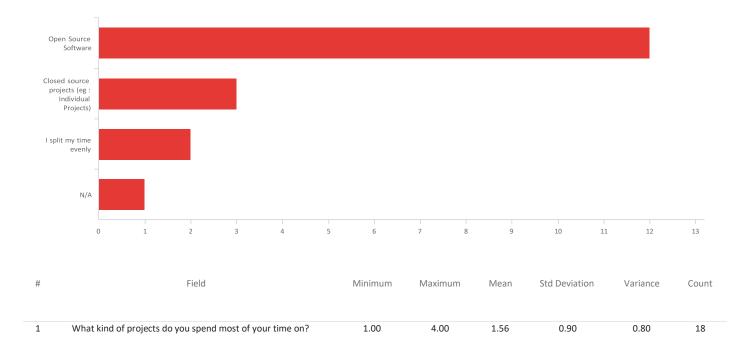
Select the range in which your age falls





Showing rows 1 - 3 of 3

Q5 - What kind of projects do you spend most of your time on?



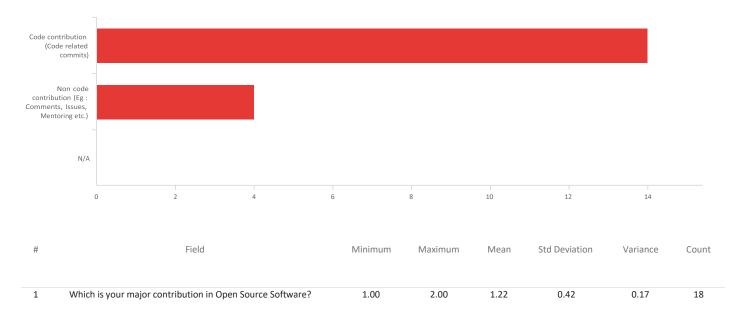
#	Field	Choice Count	
1	Open Source Software	66.67%	12
2	Closed source projects (eg : Individual Projects)	16.67%	3
3	I split my time evenly	11.11%	2

4 N/A 5.56% 1

Showing rows 1 - 5 of 5

18

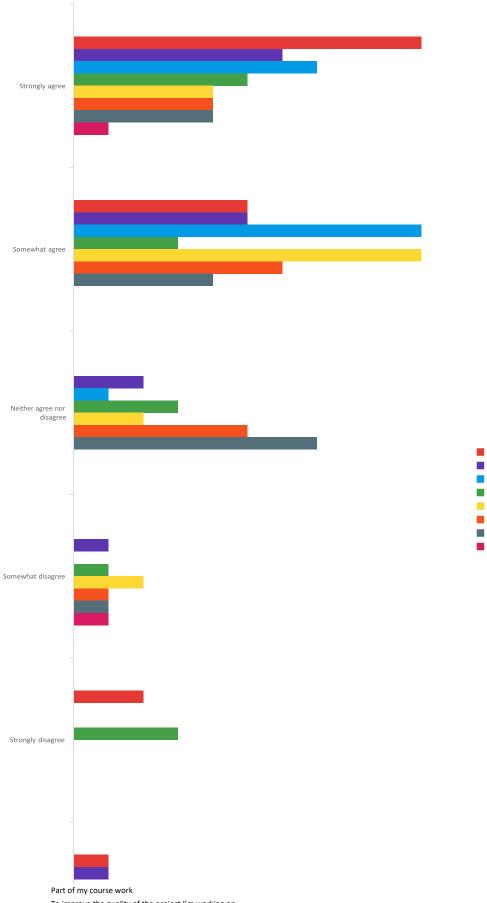
Q6 - Which is your major contribution in Open Source Software?



#	Field	Choice Count	
1	Code contribution (Code related commits)	77.78%	14
2	Non code contribution (Eg : Comments, Issues, Mentoring etc.)	22.22%	4
3	N/A	0.00%	0
			18

Showing rows 1 - 4 of 4

Q7 - Select the reason for your contribution to OSS



To improve the quality of the project I'm working on Self-interest to enhance my coding skills

I am paid for my contribution It would enhance my reputation It is fun to code I like to help others Other reasons



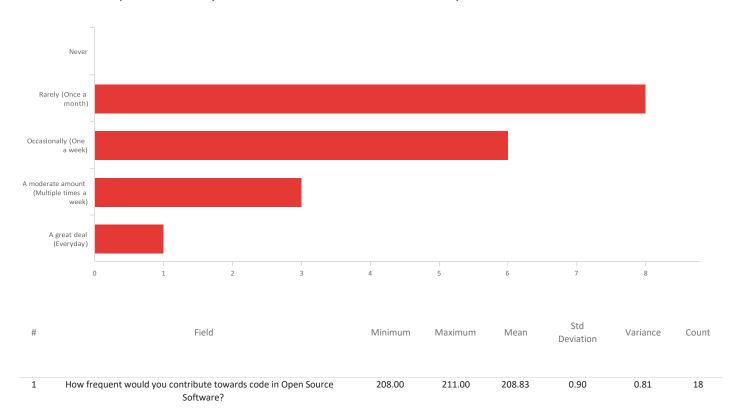
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Part of my course work	1.00	6.00	2.00	1.56	2.44	18
2	To improve the quality of the project I'm working on	1.00	6.00	2.13	1.36	1.85	15
3	Self-interest to enhance my coding skills	1.00	3.00	1.67	0.58	0.33	18
4	I am paid for my contribution	1.00	5.00	2.60	1.50	2.24	15
5	It would enhance my reputation	1.00	4.00	2.11	0.87	0.77	18
6	It is fun to code	1.00	4.00	2.19	0.88	0.78	16
7	l like to help others	1.00	6.00	2.53	1.24	1.54	17
8	Other reasons	1.00	4.00	2.50	1.50	2.25	2

#	Field agree	Strongly	Somewhat agree	Neither agree Somewhat nor disagree disagree	Strongly disagree	Not Sure	Total
1	Part of my course work	55.56% 10	27.78% 5	0.00% 0 0.00% 0	11.11% 2	5.56% 1	18
2	To improve the quality of the project I'm working on	40.00% 6	33.33% 5	13.33% 2 6.67% 1	0.00% 0	6.67% 1	15

3	Self-interest to enhance my coding skills	38.89%	7	55.56%	10	5.56%	1	0.00%	0	0.00%	0	0.00%	0	18
4	I am paid for my contribution	33.33%	5	20.00%	3	20.00%	3	6.67%	1	20.00%	3	0.00%	0	15
5	It would enhance my reputation	22.22%	4	55.56%	10	11.11%	2	11.11%	2	0.00%	0	0.00%	0	18
6	It is fun to code	25.00%	4	37.50%	6	31.25%	5	6.25%	1	0.00%	0	0.00%	0	16
7	I like to help others	23.53%	4	23.53%	4	41.18%	7	5.88%	1	0.00%	0	5.88%	1	17
8	Other reasons	50.00%	1	0.00%	0	0.00%	0	50.00%	1	0.00%	0	0.00%	0	2

Showing rows 1 - 8 of 8

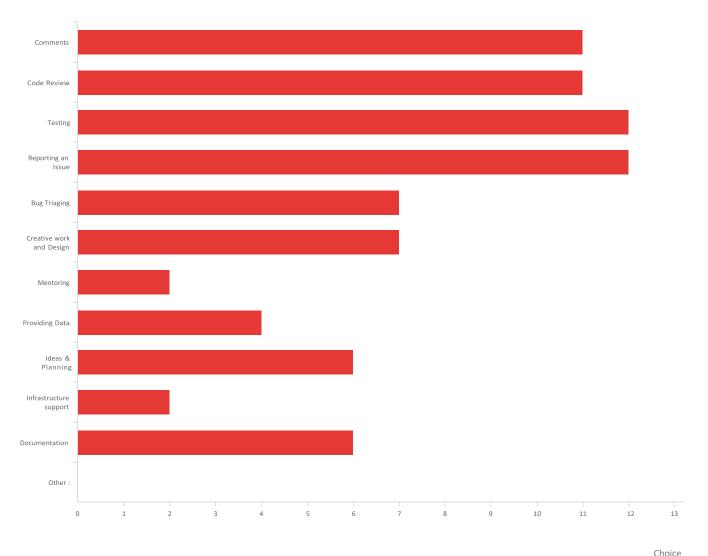
Q8 - How frequent would you contribute towards code in Open Source Software?



207	Never	0.00%	0
208	Rarely (Once a month)	44.44%	8
209	Occasionally (One a week)	33.33%	6
210	A moderate amount (Multiple times a week)	16.67%	3
211	A great deal (Everyday)	5.56%	1
			18

Showing rows 1 - 6 of 6

Q9 - Select the non-code contribution which you contribute most often, select all that apply.

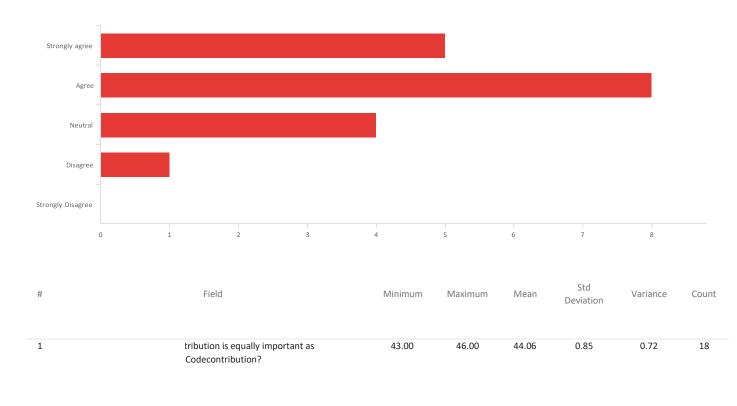


#	Field	Choice Coun	
1	Comments	13.75%	11
2	Code Review	13.75%	11
3	Testing	15.00%	12
4	Reporting an Issue	15.00%	12
5	Bug Triaging	8.75%	7
6	Creative work and Design	8.75%	7
7	Mentoring	2.50%	2
8	Providing Data	5.00%	4

#	Field		Choice Count	
9	Ideas & Planning	7	7.50%	6
10	Infrastructure support	2	2.50%	2
11	Documentation	7	7.50%	6
12	Other:		0.00%	0
	•	Showing rows 1 - 13 of 13		80
Q9_12	_TEXT - Other :			

Other:

Q10 - Do you think Non-code contribution is equally important as Codecontribution?

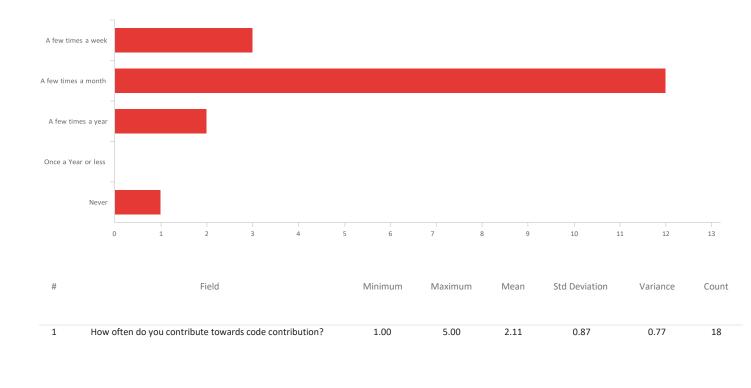


#	Field	Choic Coun	
43	Strongly agree	27.78%	5
44	Agree	44.44%	8
45	Neutral	22.22%	4
46	Disagree	5.56%	1
47	Strongly Disagree	0.00%	0

18

Showing rows 1 - 6 of 6

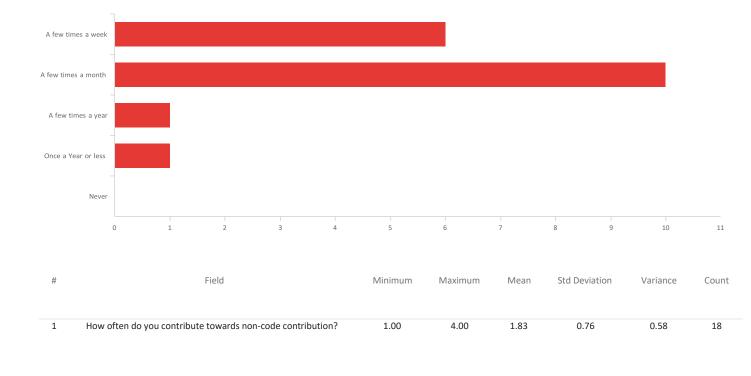
Q11 - How often do you contribute towards code contribution?



#	Field	Choice	
1	A few times a week	16.67%	3
2	A few times a month	66.67%	12
3	A few times a year	11.11%	2
4	Once a Year or less	0.00%	0
5	Never	5.56%	1
			18

Showing rows 1 - 6 of 6

Q12 - How often do you contribute towards non-code contribution?



#	Field	Choice Count	
1	A few times a week	33.33%	6
2	A few times a month	55.56%	10
3	A few times a year	5.56%	1
4	Once a Year or less	5.56%	1
5	Never	0.00%	0

Showing rows 1 - 6 of 6

End of Report