

Code vs. Non-code contributions

Rupika Dikkala
dikkalar@oregonstate.edu

Swetha Jayapathy
jayapats@oregonstate.edu

Amreeta Chatterjee
chattera@oregonstate.edu

1. INTRODUCTION

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: non-code related commits). The non-code contributions live on the sidelines, but they are equally, if not more, valuable than their code counterparts.

A non-code contribution includes documentation, reviewing pull requests, facilitating discussions, etc [12]. Ren et al.'s research states that a project's code contributions are more than the number of lines or commits; they delve into the importance non-code contributors have in fixing build errors, developing new features, and overall maintaining the source code hygiene [11]. For example, Github ranks their contributors by the number of code commits and doesn't factor in any non-code commits. And while there are external plugins [13] for Github that credits non-code contributors, very few projects leverage this service as it's not mandated to do so. In our project we want to show the discrepancy of importance given to code and non-code contributions in OSS projects through the following research questions:

1. How does the Github ranking differ when we include code and non-code contributions?
2. What is the frequency of code and non-code contributions in OSS projects over the past 5 years?
3. How does gender have an effect on the type of contribution being made in OSS projects in the last 5 years?
4. How do developers on Github perceive the importance of non-code contributions compared to code contributions?

2. BACKGROUND

Our SLR encompasses two processes: A query search and snowballing by doing a citation analysis. We established the following inclusion criteria: a full paper, written in the English language, and published in a peer-reviewed venue. The exclusion criteria consisted of excluding papers published earlier than 2010, since we only wanted to focus on the last ten years. Papers regarding code contributions only, which did not mention non-code contributions were not within the scope of the SLR.

In the initial process, we started with the following queries: ("non-code" AND "contributors") OR ("Non-code" AND "contributions") OR ("Non-code contributors" and "gender"). We then filtered the results by looking at the title and abstract. The title or abstract should mention non-code contributions and/or gender in OSS.

This resulted in seven papers. We further looked at the citations of these papers to find more relevant papers, and thus extracted 10 papers as shown in Figure 1.

Bosu et. al [1] gathered information about code review practices via surveys of open source software developers. They found that developers consider code review important and that code reviews work as a medium to mentor new teammates about the project

design, coding conventions, and available API or libraries. Kuechler et. al [2] studied eleven mailing lists and corresponding subscription logs from six FOSS projects. They found that only 1.5% women are code contributors and the participation of women declines with every step of the FOSS joining process.

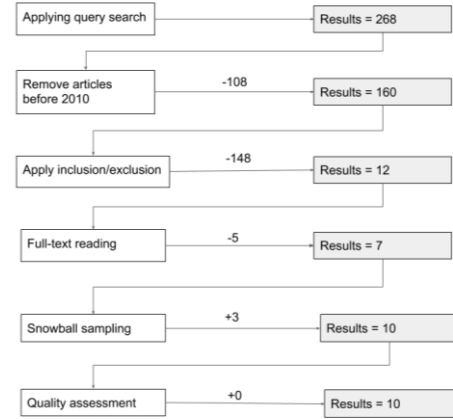


Fig 1. Number of papers included in the study selection process

Trikenreich et. al. [3] investigated the different roles performed by successful and recognized OSS contributors during their career in OSS communities. Through empirical evidence, they found that they found roles that are community-centric roles (e.g advocate, license manager, community founder) in addition to the well-known project-centric ones (e.g maintainer, core member). These community-centric roles remain hidden, since these roles may not leave traces in software repositories typically analyzed by researchers.

In [4], authors found that women perform majoritarily other types of contributions than coding, while men mostly contribute with code. More than half of the women are involved in leading, administering or coordinating FLOSS projects. The share of women leading projects is also lower than for men.

3. OVERALL DESIGN OF THE STUDY

For the metrics in this project, we first needed to define code contributions and non-code contributions. Since we are looking at the contributor ranking, we decided to base it off Github's contribution ranking. Thus for code contributions, we chose the number of code commits [14]. For non-code, we wanted to focus on issues and the merging of pull requests because they drive the debugging process [11]. Therefore we have a total of 4 non-code buckets: people who authored comments on both issues and pull requests, the issue author - the person who created the issue, and the merge author - the person who merged the pull request.

Our sample included 10 projects with a medium popularity (100 stars or more) and a maximum size of 50MB due to time constraints. They are as follows: **d3**, **EbookFoundation**, **sindresorhus**, **jwasham**, **getify**, **ohmyzsh**, **donnemartin**, **github**, **airbnb**, **public-apis**.

The main focus on this data is the contributor, and the aggregated count for all the 4 types of non-code contributions and the counts for the code contributions for each project. All data pulled from Github was further sanitized to remove invalid or "ghost" users, i.e. contributors who deleted their accounts. Furthermore, we filtered the data to only include contributions over the last 5 years (Jan 2015 - Dec 2019). There was some additional data handling done for individual research questions, which is delineated below.

4. METHODOLOGY

4.1 RQ1 - How does the Github ranking differ when we include code and non-code contributions?

In order to gather the number of code contributions, we collected the contributor list for each project and mined the 'login' and 'contribution' to extract the number of commits per person. In order to count the number of non-code contributions, we counted the number *instances* each contribution was made; where an *instance* is defined as a comment (issue or pull), an issue created, or a merged pull request.

For gathering comment contributors, we used the `repo/{}/issues/comments` API call (where {} is the project), which included comments for both issues and pull requests. To gather issue authors, we used the `repo/{}/issues` API call and filtered out pull requests from our query. In order to get the merge author, we used the same API call (this time filtering out all the issues and open pull requests), and then ran `repos/{}/pulls/pull_number`. Once we gathered the metadata for each individual PR, we extracted the merge-author if the PR was already merged. We also gleaned the created-by timestamp for each of these contributions in preparation for RQ2. This process resulted in 3 dataframes (Issue/PR comments, Issue Author, Merge Author) which we then concatenated and ran a group-by function in order to gather the counts for each contributor (as shown in Figure 2). Full source code is included in the Appendix.

In order to produce the final rank list that combined both code and non-code contributions, we had to determine the amount of contributions each contributor made. As aforementioned, we counted non-code contributions by number of instances, and since the value of one code commit and one non-code instance isn't necessarily equal, we normalized both metrics by taking a percentage of the total number of contributions. In order to account for users who made both code and non-code contributions, we summed the values to generate a total %. From there, we generated a rank including both code and non-code contributions (Total Rank column in Figure 3) while the "Github" rank column is the ordering from Github.

	login	contributions
0	mbostock	1425
1	curran	76
2	jasondavies	67
3	Fil	47
4	mgold	31
...
1242	r1m	1
1243	qian99	1
1244	mantri-govind	1
1245	dfranco	1
1246	LiamKarlMitchell	1

Fig 2. Non-code Contribution *instance* count per contributor for a sample project

Github Ranking	Total Ranking	login	noncode%	code%	total%
0	1	vhf	15.955	34.935	50.890
1	2	MHM5000	0.025	6.824	6.849
2	5	alexanderfefelev	0.000	3.038	3.038
3	4	MrS4w	0.225	2.947	3.172
4	8	esparta	0.000	2.403	2.403
5	7	kadhirash	0.937	1.950	2.887
6	9	ericguirbal	0.012	1.564	1.576
7	6	borgified	1.449	1.564	3.013
8	0	esthellman	51.049	1.383	52.432
9	10	HankiDesign	0.162	1.179	1.341
10	13	maieul	0.087	1.111	1.198
11	14	mciolella	0.062	0.998	1.060
12	16	jlightuse	0.000	0.771	0.771
13	19	luminousspice	0.000	0.680	0.680
14	20	danielkrupinski	0.087	0.589	0.676

Fig 3. From Left to Right: Github ranking is the ranking taken from Github, Total Ranking is Non-code+Code contributions, login is the contributor, noncode% and code% is normalized contribution count, total% is the sum of the two

4.2 RQ2 -What is the frequency of code and non-code contributions in OSS projects over the past 5 years?

We mined Git to get the code contributions and non-code contributions of the ten projects for the past 5 years. Code contributions included the code commits whereas non-code contributions included issues, pull requests, and comments. Comments consist of comments from issues and pull requests. The total number of Code contributions was 13420 and that of non-code contributions was 66943. The created date of this data was collected and grouped into years and then months. For each year, the frequency of the non-code contributions and the code-contributions were compared month-wise (Figure 2).

Descriptive statistics for both types of contribution was taken for each month per year. A T-test with a 95 % confidence interval for the below-mentioned hypothesis (Hypothesis 1) was performed. The difference between the number of non-code contributions and the number of code contributions for each month per year was taken and then a one-sample t-test was performed on the result.

Hypothesis 1- The mean population difference between the number of code contributions and the number of non-code contributions is zero.

4.3 RQ3: How does gender have an effect on the type of contribution being made in OSS projects in the last 5 years?

Using GithubAPI, we compiled the list of commits and the non-code contributions made in each of the ten projects, since 2015. From this data, we extract the login information of the contributor, which is then used to find the full names of unique authors and their locations, as mentioned on their Github profiles. We collected a total of 3589 code contributors and 6371 non-code contributors across ten projects.

The gender of these contributors were inferred using the genderComputer: male, female, or unisex, which has been used in several prior SE studies [5]. There were no results for 461 code contributors and for 910 non-code contributors, because of various reasons: some had their name and location details written in a language other than English, whereas some did not have any location information on their Github profiles.

4.4 RQ4 -How do developers on Github perceive the importance of non-code contributions compared to code contributions?

To answer RQ4, we thought that an online survey would be the best choice to get direct opinions from the contributors. The Survey was conducted online through Qualtrics software. 3 different surveys were clubbed together to analyze the solution. The survey consisted of an average of 12 questions each, which included demographic questions and specific questions needed to answer the research questions. The initial Survey design was evaluated by the peer evaluator for each survey. The peer evaluator was selected from students in the class. Then the Survey questions were reframed according to the feedback obtained from the peer evaluators (Appendix A).

Feedback provided was very much useful to direct the questions in such a way that the participants would have no difficulty in understanding and providing the answers. Feedback such as gender-related questions did not have a text box for self-describe, a participant might not be aware of what a non-code contribution might be and therefore it is better to provide examples for such questions, etc were provided.

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, 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.

5. RESULTS

5.1 RQ1 Results

In order to explore the differences in the contributor ranking, we wanted to compare the "Total Ranking" column we produced that included both code and non-code contributions with the "Github Ranking" column (recall Figure 3) containing only code contributions. To achieve this, we ran a Wilcoxon Signed Rank Test to evaluate for a significant difference between the two groups in each project. We chose this test because 1) our data was paired (two different ranks for a contributor), and 2) our data was non-parametric because it wasn't normally distributed (every contributor had a unique rank within each group).

	project	statistic	p-value	# contributors
0	d3	432120.0	0.429	1334
1	EbookFoundation	289623.0	0.560	1088
2	sindresorhus	395224.0	0.916	1262
3	jwasham	77099.0	0.990	561
4	getify	226160.0	0.459	966
5	ohmyzsh	10319569.5	0.851	6437
6	donnemartin	36522.0	0.910	387
7	github	409030.0	0.652	1295
8	airbnb	1132065.0	0.865	2139
9	public-apis	104735.0	0.005	694

Fig 4. From left to right: project name, Wilcoxon test Z-statistic, p-value, number of contributors in the project

As seen in Figure 4, only the public-api project indicated that the "Total Rank" was significantly different from the Github ranking $Z = 104735$, $p < 0.005$. The other 9 projects do not indicate a statistically significant difference between the Total Rank group and the Github Rank group as the p-values are higher than 0.05.

5.2 RQ2 Results

It can be seen from Figure 5 that the frequency of non-code contribution is consistently higher than the frequency of code contribution for each month of all the years. Moreover, the difference between them is huge.

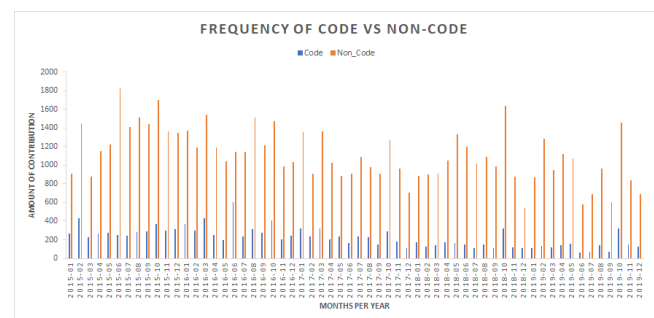


Fig 5. Comparison of number of code contributions and non-code contributions

Figure 6 shows the descriptive statistics for the number of code contributions and non-code contributions per year. We could see that the mean of the number of non-code contributions is far higher than the mean of the number of code contributions. For example, if we look at the year 2015, we see that the mean of code contribution is 292.5, while that of non-code contribution is 1350.75. Thus the difference is significant.

Statistics /Year	Mean		Median		Standard Deviation		Variance	
	Code	Non-Code	Code	Non-Code	Code	Non-Code	Code	Non-Code
2015	292.5	1350.75	279	1387.5	57.14	280.81	3266.09	78854.93
2016	318.5	1237.5	287	1190.5	117.04	191.6	13699.36	36711.36
2017	222.83	1030	229	972	65.78	203.61	4327.24	41460.72
2018	152.41	1034.5	146.5	1002.5	57.53	271.04	3309.9	73463.54
2019	132.08	925.83	129	908.5	66.81	272.06	4463.71	74020.69

Fig 6. Descriptive statistics for number of code contributions and non-code contributions per each year

We could also observe that there is a large standard deviation for both types of contributions. As we only look at issues, PR and comments for non-code contribution, and did not consider other types of contribution, it may have resulted in a large standard deviation.

From the one-sample t-test, there is convincing evidence that the population mean difference between the number of code contributions and the number of non-code contributions is not zero ($p\text{-value} = 2.2e-16$, $df = 59$, 95 % Confidence Interval from 833.88 to 950.21). The t-test result suggests that the non-code contributions are higher than the code contributions.

5.3 RQ3 Results

Figure 7 shows a bar graph of the percentage of male, female and unisex, as inferred by the genderComputer. To answer RQ3, we did not take into account the contributors whose gender was inferred as “unisex” by the genderComputer. This exclusion resulted in 3091 code contributors and 5380 non-code contributors.

A chi-square test of independence was performed to examine the relation between gender and the type of contributions made, that is, code or non-code. The relation between these two variables was not significant, $X^2(1, N = 8471) = 0.06$, $p = .80$.

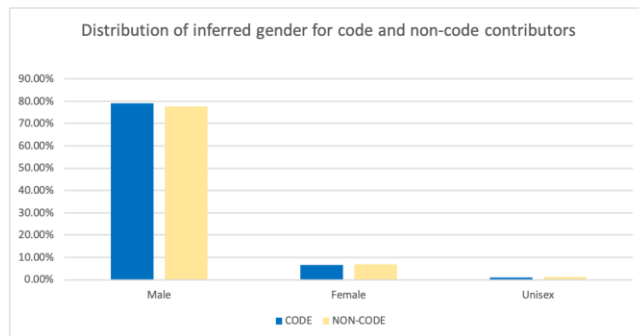


Fig 7. This bar chart shows the percentages of male and female contributors who contributed to the ten projects, either by committing code or doing non-code contributions.

5.4 RQ4 Results

Each of the 3 surveys was distributed to 20 people online and all of them participated in the survey leading to a 100% survey response. The majority of the participants were from class (80%). The participants were from 3 different countries – the US, China and India. A major number of the participants were from the US. Among the total participants, 7 replied that they had not worked with open source software and therefore, they were not eligible to participate. Most of them replied that they contributed to OSS because of self-interest to enhance their coding skills and to enhance their reputation.

59% of the participants have mentioned that they contribute most often toward code contributions while 23% toward non-code contributions and 18% of them contribute to both. The majority of participants replied that non-code contributions are equally important as code contributions. 29% of participants selected the strongly agree option and 34% of participants selected agree. With over 62% agreeing that non-code contributions are equally important as code contributions (Figure 8).

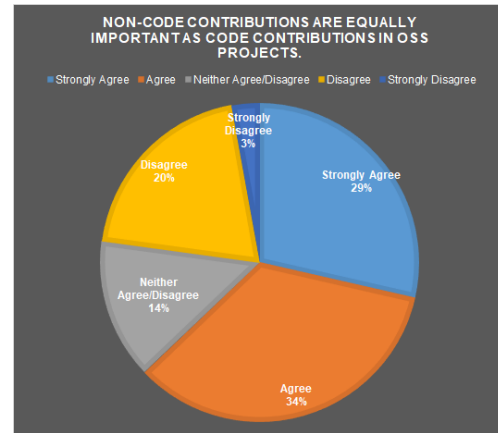


Fig 8. Percentage of responses for the question- Are non-code contributions are equally important as code contributions?

When asked about the frequency of their contributions, 11% responded that they contribute a few times a week and 57% responded they contribute a few times a month towards code contribution (Figure 9). While 31% responded that they contribute a few times a week and 37% responded they contribute a few times a month towards non-code contributions (Figure 10). Looking at both the graphs, we could observe that the overall frequency of non-code contribution is higher than the code contributions.

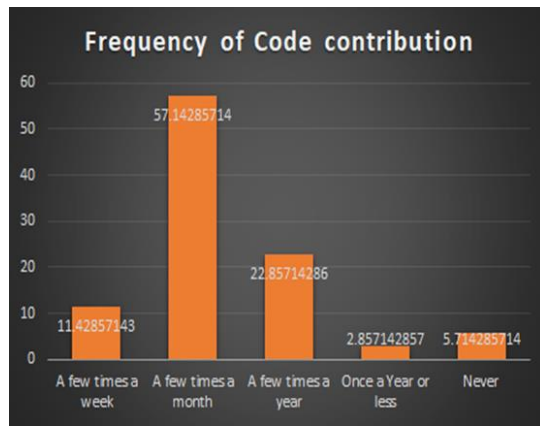


Fig 9: Frequency response for code contributions

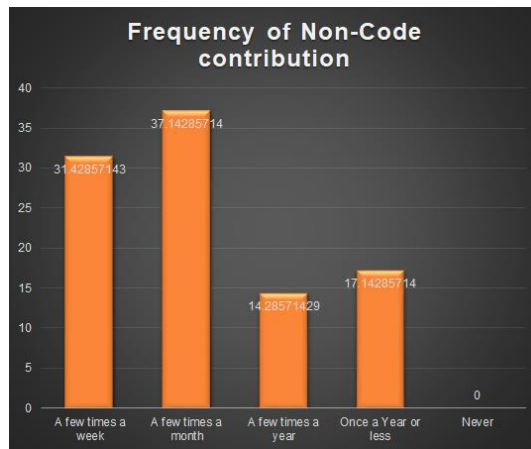


Fig 10: Frequency response for non-code contributions

The participants were asked to rank the contributions among code contributions, participation in OSS, posting on mailing lists, non-technical parts of contribution and understanding the goals of the project. The result is shown in Figure 11. We could see that code contribution occupies the first and second rank among the participants. Posting on the mailing list is among the 3rd rank and then comes non-technical contributions.

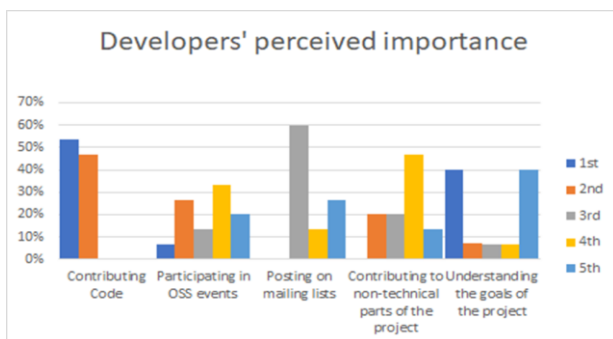


Figure 11: Ranking of the contributions

6. THREATS TO VALIDITY

A few threats to the validity of our analysis are as follows:

a. Internal

Inference of Gender using the genderComputer. On Github, disclosing name and location on user profiles is optional and might be absent altogether. In these cases, we could not infer the gender of these contributors. Furthermore, the genderComputer only accounts for Countries as locations, so we lost data that only contained a City as a location. As mentioned in RQ1 results, our definitions of a non-code instance may be flawed. We counted an *instance* as a comment, issue created or PR merged - however there is a difference between commenting on an issue and creating an issue (for example), so perhaps including a weighted score might have yielded a more accurate ranking.

b. Construct

The overarching definitions of a code and non-code contributions may be inaccurate. Furthermore, our definition of non-code contributions only pertained to issues and PRs, and completely excluded code reviews, README files, documentation, etc. On the same note, we overcompensated for code contributions by including all code commits instead of excluding README files or comments within the code.

c. External

Our results are only applicable to Github, not other OSS platforms such as Source forge, Bitbucket, etc. Due to our small sample size, our survey results might not be generalizable, as well. Additionally, the projects we selected for mining represents only a subset of all the repositories on Github, so it cannot be completely generalized across Github either. Furthermore the number of contributors is not the same across all projects which prevents generalizability.

7. DISCUSSION

RQ1: Our results indicated there's no significant difference between the Github Ranking and Total Ranking groups. The extremely high Z-statistic for these projects (refer to Figure 4) could be indicated by the lack of normality in our data; as aforementioned, each contributor has a unique rank within both groups. The large p-values could be attributed to our definition of a non-code *instance*, and could further relate to the way we defined a code and non-code contribution as delineated in the Threats section. In future studies, careful definitions for code and non-code contributions must be implemented in order to account for all contributions being made.

RQ2: We infer from the results that there is a significant difference between the frequency of non-code contributions and code contributions with a very less p-value. We could also say that the frequency of non-code contributions are higher than the frequency of code contributions from the test results and the bar graph (Fig 5). The non-code contributions used to analyze this result do not consist of all types of non-code contributions, we picked the most important ones and therefore, when all the contributions are included, we might get strong evidence.

RQ3: Our results showed that there was no significant difference in the number of men and women doing code and non-code

contribution. However, the participation of women in both code and non-code contribution was around 7%, which is more than the reported 1.5% in Robles [4]. This shows that although imbalance in gender remains, participation of women has increased in OSS. Future research should continue to act upon bridging this gap.

RQ4: We could see from the survey results and the mined data that the number of non-code contributions are being more frequent than the number of code contributions. Also, the amount of non-code contribution is comparatively higher than the code contribution. Whereas when the developers were asked to rank, we got a contrasting result that they perceive code contributions as the most important one. Therefore, further research needs to be done to analyze this.

Our research questions have changed substantially from the start of the term. Our original RQ1 was “*How do we compare non-code contributors with code contributors in the past 5 years?*”; we changed it to “*How does the Github ranking differ when we include code and non-code contributions?*” in order to make a more meaningful comparison. Furthermore, we added RQ4 in order to incorporate data from software practitioners and real OSS users. We also decided to shrink the size of our projects and narrow our definition of a non-code contribution due to time constraints. If we were to redo the study, we would add more concrete definitions for non-code/code contributions. Furthermore, in order to improve the genderComputer results, we would execute a script that outputs the Country based on the City in order to improve location accuracy and include more data in our analysis.

8. CONCLUSIONS

In this project we investigated the importance given to code and non-code contributions in OSS projects.

We compared the Github rankings (which only included code contributions) with a generated “Total Ranking” that included both code and non-code contributions. We did not find a significant difference between the two rankings, which could be attributed to our metric and data definitions.

We found that gender did not have a significant effect on the type of contributions being made. It can be concluded that the frequency of non-code contributions are much higher than the frequency of code contributions. Although it can be seen from the results that a lot of non-code contributions are being done, developers perceive that code contributions are the most important one.

9. REFERENCES

- [1] Bosu, A., Carver, J. C., Bird, C., Orbeck, J., & Chockley, C. (2016). Process aspects and social dynamics of contemporary code review: Insights from open source development and industrial practice at microsoft. *IEEE Transactions on Software Engineering*, 43(1), 56-75.
- [2] Kuechler, V., Gilbertson, C., & Jensen, C. (2012, September). Gender differences in early free and open source software joining process. In *IFIP International Conference on Open Source Systems* (pp. 78-93). Springer, Berlin, Heidelberg.
- [3] Trinkenreich, B., Guizani, M., Wiese, I., Sarma, A., & Steinmacher, I. (2020). Hidden Figures: Roles and Pathways of Successful OSS Contributors. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1-22.
- [4] Robles, G., Reina, L. A., González-Barahona, J. M., & Domínguez, S. D. (2016, May). Women in free/libre/open source software: The situation in the 2010s. In *IFIP International Conference on Open Source Systems* (pp. 163-173). Springer, Cham.
- [5] J. Terrell, A. Kofink, J. Middleton, C. Raine, E. Murphy-Hill, C. Parnin, and J. Stallings, “Gender differences and bias in open source: Pull request acceptance of women versus men,” *PeerJ Computer Science*, vol. 3, p. e111, 2017.
- [6] Rebouças, M., Santos, R. O., Pinto, G., & Castor, F. (2017, May). How does contributors' involvement influence the build status of an open-source software project?. In *2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR)* (pp. 475-478). IEEE.
- [7] Izquierdo, D., Huesman, N., Serebrenik, A., & Robles, G. (2018). Openstack gender diversity report. *IEEE Software*, 36(1), 28-33.
- [8] Nordin, D., & Meir, A. (2013). Development of a questionnaire to study motivations of Drupal contributors.
- [9] Chahal, K. K., & Saini, M. (2018, June). Developer Dynamics and Syntactic Quality of Commit Messages in OSS Projects. In *IFIP International Conference on Open Source Systems* (pp. 61-76). Springer, Cham.
- [10] Barcomb, A., Stol, K. J., Fitzgerald, B., & Riehle, D. (2020). Managing Episodic Volunteers in Free/Libre/Open Source Software Communities. *IEEE Transactions on Software Engineering*.
- [11] Ren, Jinglei, et al. "Towards quantifying the development value of code contributions." *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 2018.
- [12] GitHub. “How to Contribute to Open Source” Retrieved from <http://opensource.guide/how-to-contribute/>
- [13] “All Contributors” Retrieved from <https://allcontributors.org/>
- [14] GitHub Docs. “Viewing a project's contributors” Retrieved from <https://docs.github.com/en/free-pro-team@latest/github/visualizing-repository-data-with-graphs/viewing-a-projects-contributors>

APPENDIX A

Survey1 Link: https://oregonstate.qualtrics.com/jfe/form/SV_9zdqrnTaQOExfq_R

Survey2 Link: https://oregonstate.qualtrics.com/jfe/form/SV_8HxXTHrPl44w2CV

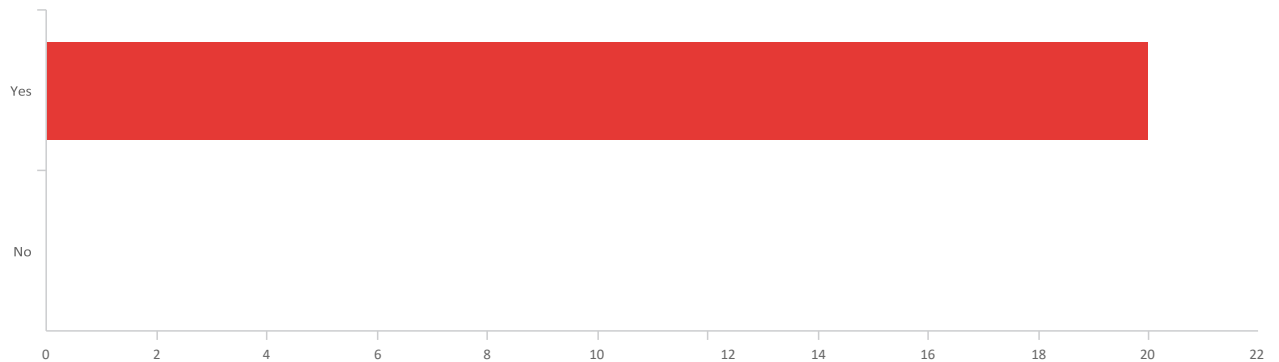
Survey3 Link: https://oregonstate.qualtrics.com/jfe/form/SV_9Nw8Vz8V2lpOnxX

Default Report

Survey Design - Code Contributors

November 30, 2020 4:03 PM MST

Q7 - Thanks for choosing to participate in the screening survey for the Code Contributor Study! This survey should only take 4-6 minutes of your time. We're student researchers at Oregon State University trying learn more about the differences between non-code and code contributions in OSS (Open Source Software) communities. In order to qualify for this survey, you must have experience with OSS projects. Your personal information will be kept strictly confidential and will not be sold, reused, rented, loaned or otherwise. Do you give your consent to participate in this survey?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Thanks for choosing to participate in the screening survey for the Code Contributor Study! This survey should only take 4-6 minutes of your time. We're student researchers at Oregon State University trying learn more about the differences between non-code and code contributions in OSS (Open Source Software) communities. In order to qualify for this survey, you must have experience with OSS projects. Your personal information will be kept strictly confidential and will not be sold, reused, rented, loaned or otherwise. Do you give your consent to participate in this survey?	1.00	7.00	6.71	1.28	1.63	21

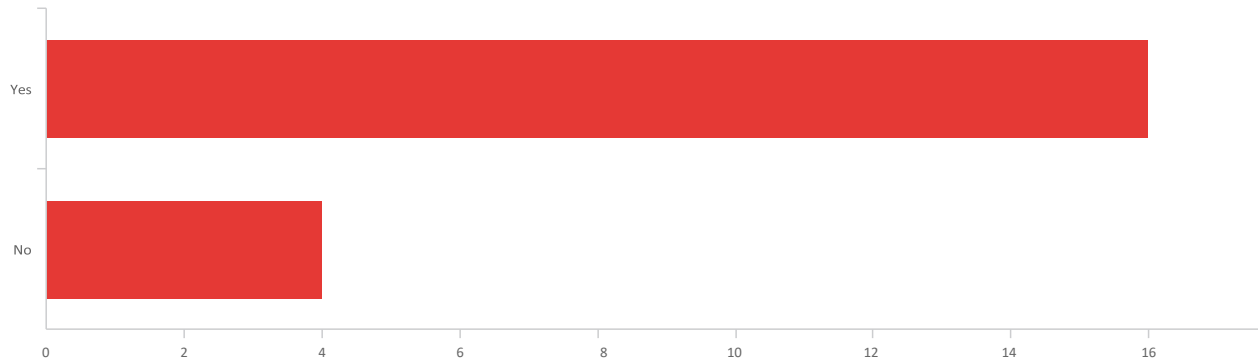
#	Field	Choice Count
---	-------	--------------

#	Field	Choice Count
---	-------	--------------

7	Yes	100.00%	20
8	No	0.00%	0
			20

Showing rows 1 - 3 of 3

Q17 - Do you have experience working with OSS (Open Source Software) projects in the last 3 years?



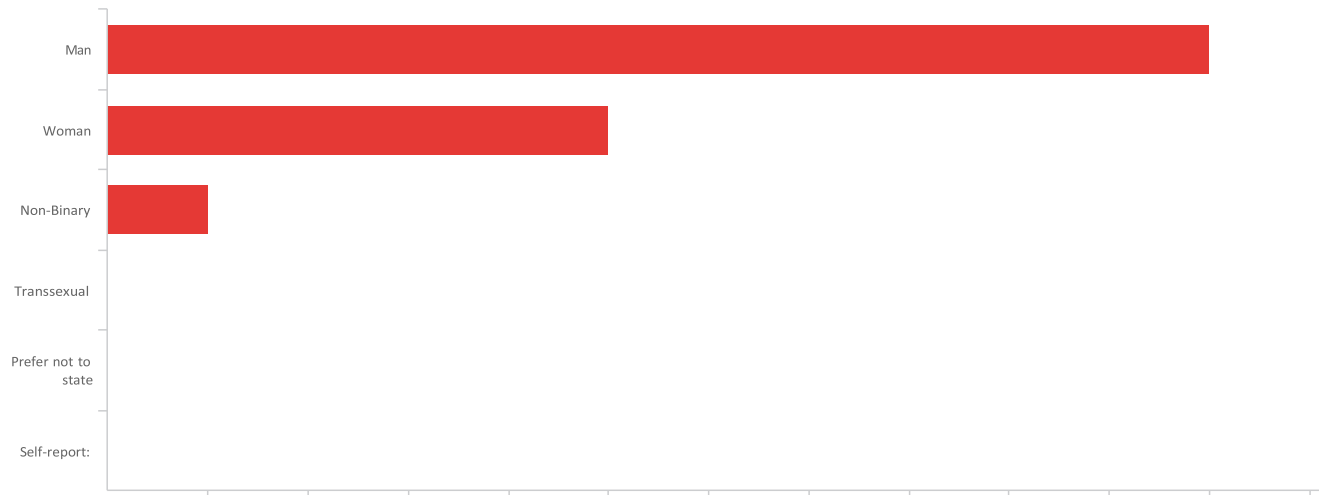
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Do you have experience working with OSS (Open Source Software) projects in the last 3 years?	1.00	2.00	1.20	0.40	0.16	20

#	Field	Choice Count
1	Yes	80.00% 16
2	No	20.00% 4

20

Showing rows 1 - 3 of 3

Q3 - What gender (if any) do you identify yourself by? (check all that apply)



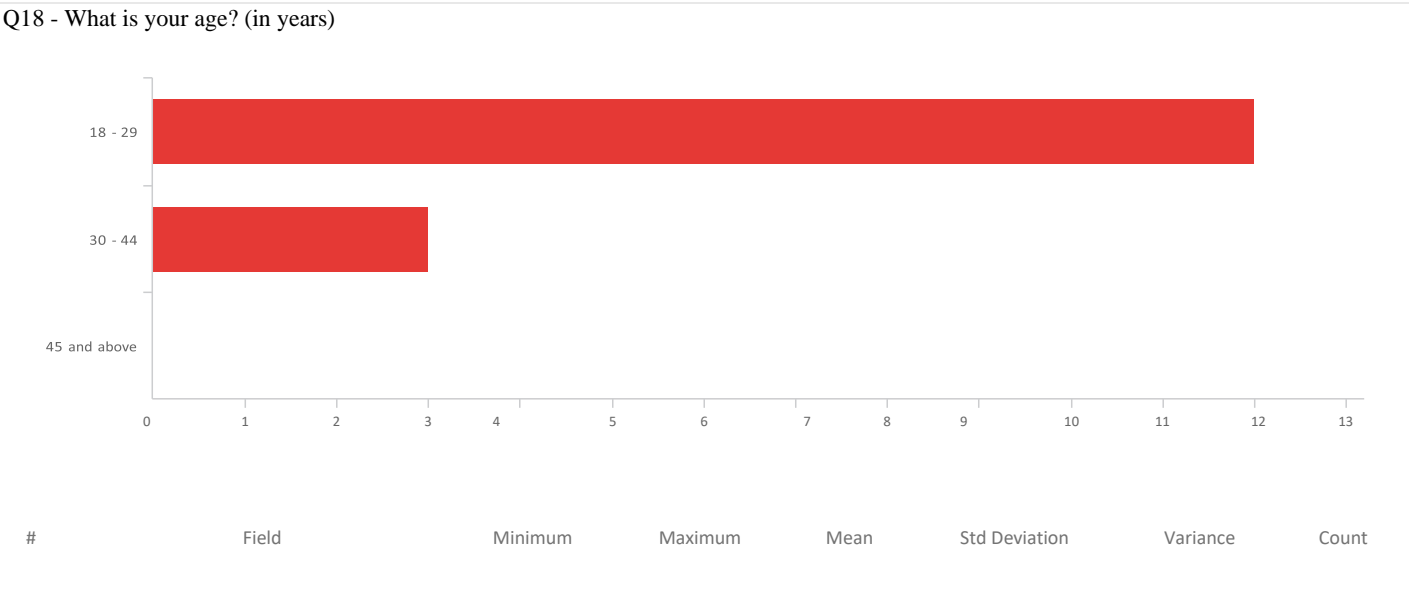
0123456789101112

#	Field	Choice Count
1	Man	64.71% 11
2	Woman	29.41% 5
3	Non-Binary	5.88% 1
4	Transsexual	0.00% 0
5	Prefer not to state	0.00% 0
6	Self-report:	0.00% 0
		17

Showing rows 1 - 7 of 7

Q3_6_TEXT - Self-report:

Self-report:

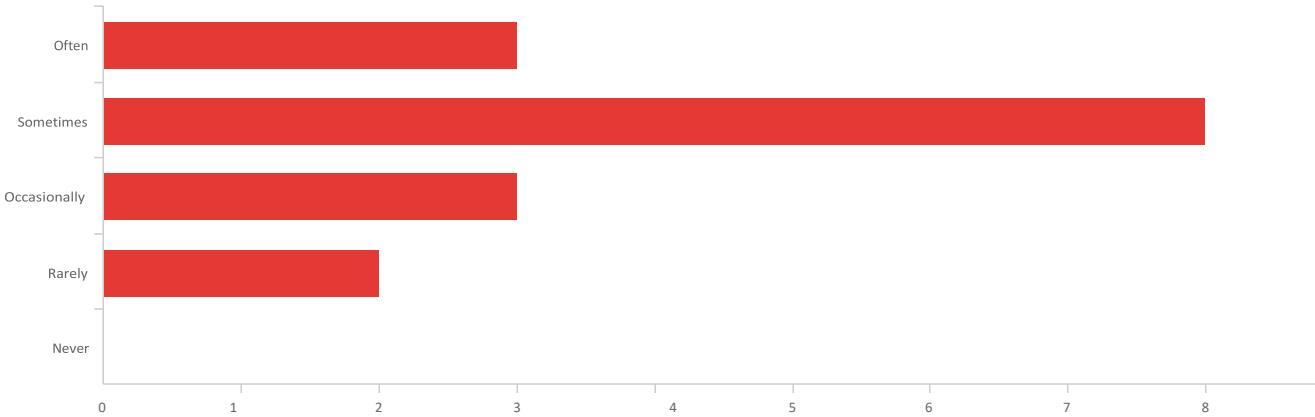


1	What is your age? (in years)	1.00	4.00	1.60	1.20	1.44	15
---	------------------------------	------	------	------	------	------	----

#	Field	Choice Count
1	18 - 29	80.00% 12
4	30 - 44	20.00% 3
5	45 and above	0.00% 0
		15

Showing rows 1 - 4 of 4

Q8 - How often have you contributed to OSS (Github, Bitbucket, Gitlab, etc) in the past 3 years?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How often have you contributed to OSS (Github, Bitbucket, Gitlab, etc) in the past 3 years?	1.00	4.00	2.25	0.90	0.81	16

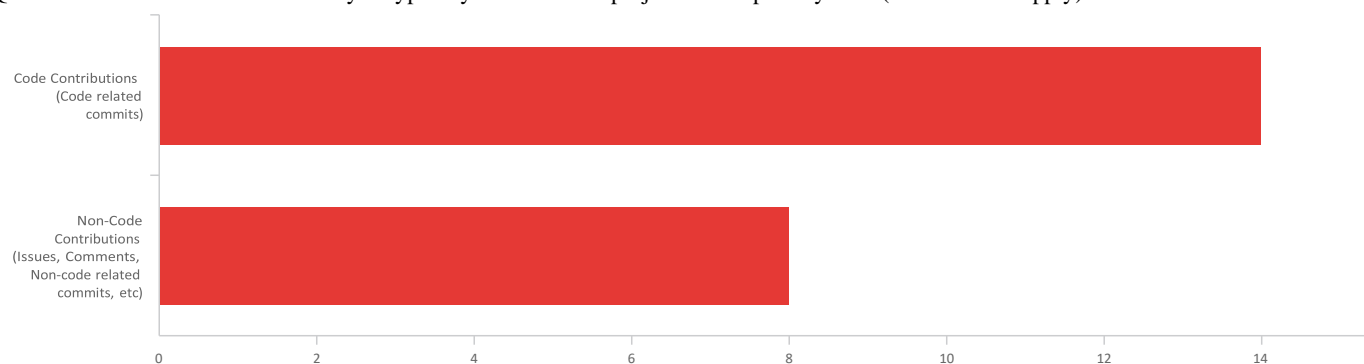
#	Field	Choice Count
---	-------	--------------

1	Often	18.75%	3
2	Sometimes	50.00%	8
3	Occasionally	18.75%	3
4	Rarely	12.50%	2
5	Never	0.00%	0

			16
--	--	--	----

Showing rows 1 - 6 of 6

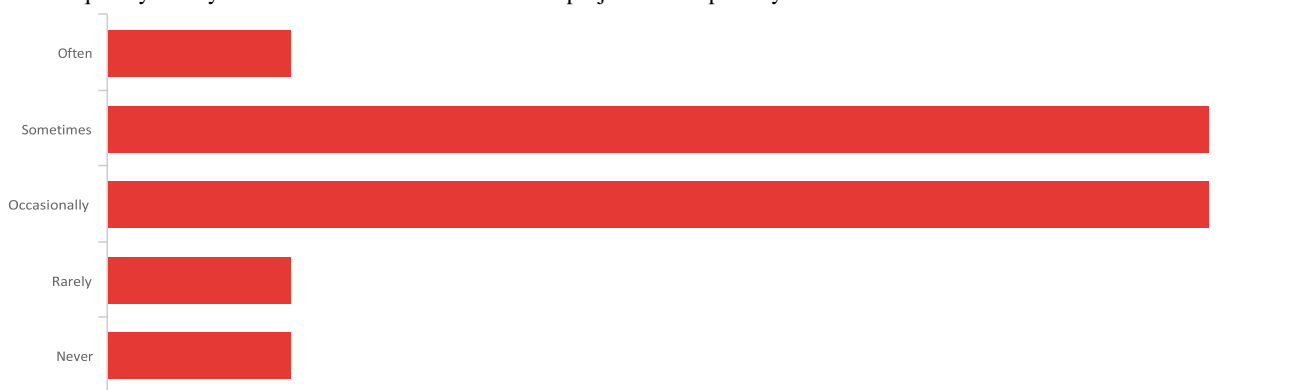
Q12 - What kind of contributions do you typically make to OSS projects in the past 3 years? (select all that apply)



#	Field	Choice Count
1	Code Contributions (Code related commits)	63.64% 14
2	Non-Code Contributions (Issues, Comments, Non-code related commits, etc)	36.36% 8
		22

Showing rows 1 - 3 of 3

Q11 - How frequently have you made code contributions to OSS projects in the past 3 years?



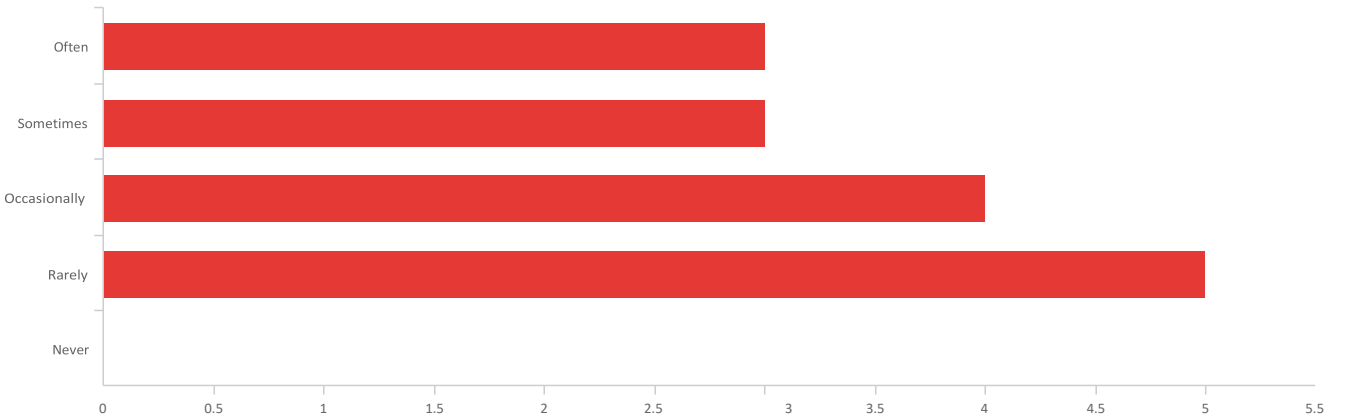
00.511.522.533.544.555.566.5

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How frequently have you made code contributions to OSS projects in the past 3 years?	1.00	5.00	2.67	0.94	0.89	15

#	Field	Choice Count
1	Often	6.67% 1
2	Sometimes	40.00% 6
3	Occasionally	40.00% 6
4	Rarely	6.67% 1
5	Never	6.67% 1
		15

Showing rows 1 - 6 of 6

Q19 - How frequently have you made non-code contributions to OSS projects in the past 3 years?

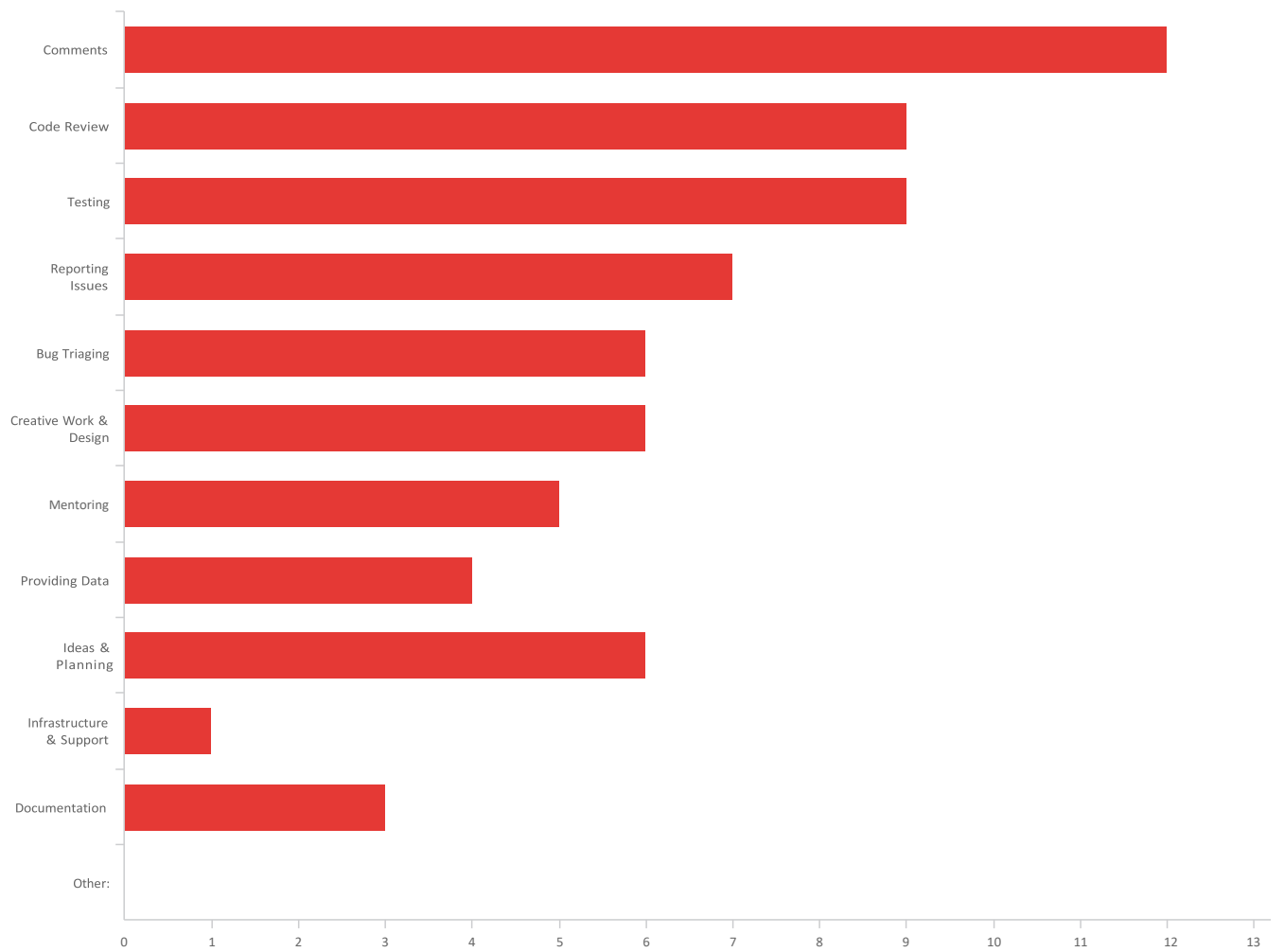


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
---	-------	---------	---------	------	---------------	----------	-------

1	ave you made non-code contributions to OSS projects in the past 3 years?		1.00	4.00	2.73	1.12	1.26	15
#	Field							Choice Count
1	Often						20.00%	3
2	Sometimes						20.00%	3
3	Occasionally						26.67%	4
4	Rarely						33.33%	5
5	Never						0.00%	0
								15

Showing rows 1 - 6 of 6

Q20 - If applicable, what kind of non-code contributions did you make to OSS projects in the last 3 years? (select all that apply)



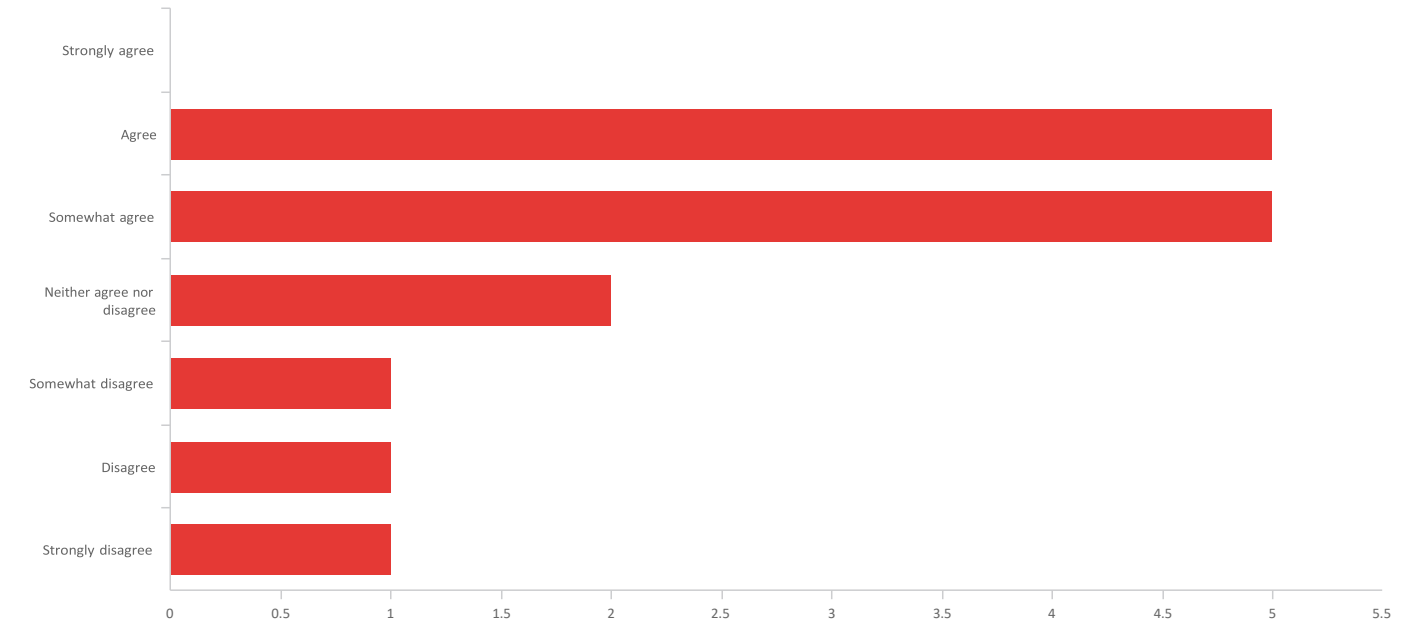
#	Field	Choice Count
1	Comments	17.65% 12
4	Code Review	13.24% 9
5	Testing	13.24% 9
6	Reporting Issues	10.29% 7
7	Bug Triaging	8.82% 6
8	Creative Work & Design	8.82% 6
9	Mentoring	7.35% 5
10	Providing Data	5.88% 4

#	Field	Choice Count
11	Ideas & Planning	8.82% 6
12	Infrastructure & Support	1.47% 1
13	Documentation	4.41% 3
14	Other:	0.00% 0
Showing rows 1 - 13 of 13		68

Q20_14_TEXT - Other:

Other:

Q9 - I feel satisfied with my effort after submitting a non-code contribution to an OSS project.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	I feel satisfied with my effort after submitting a non-code contribution to an OSS project.	2.00	7.00	3.40	1.50	2.24	15

#	Field	Choice Count
1	Strongly agree	0.00% 0
2	Agree	33.33% 5
3	Somewhat agree	33.33% 5
4	Neither agree nor disagree	13.33% 2
5	Somewhat disagree	6.67% 1
6	Disagree	6.67% 1

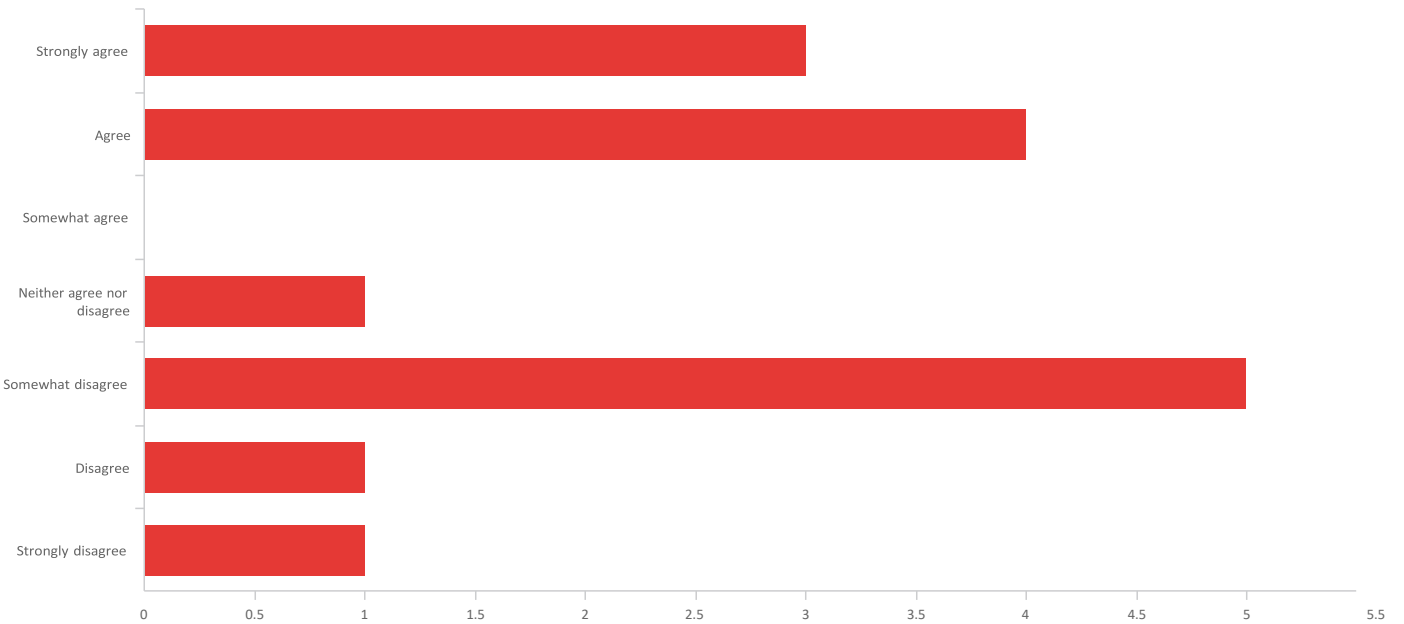
7 Strongly disagree

6.67% 1

15

Showing rows 1 - 8 of 8

Q15 - I believe non-code contributions are equally as important as code contributions in OSS projects.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	I believe non-code contributions are equally as important as code contributions in OSS projects.	1.00	7.00	3.53	1.96	3.85	15

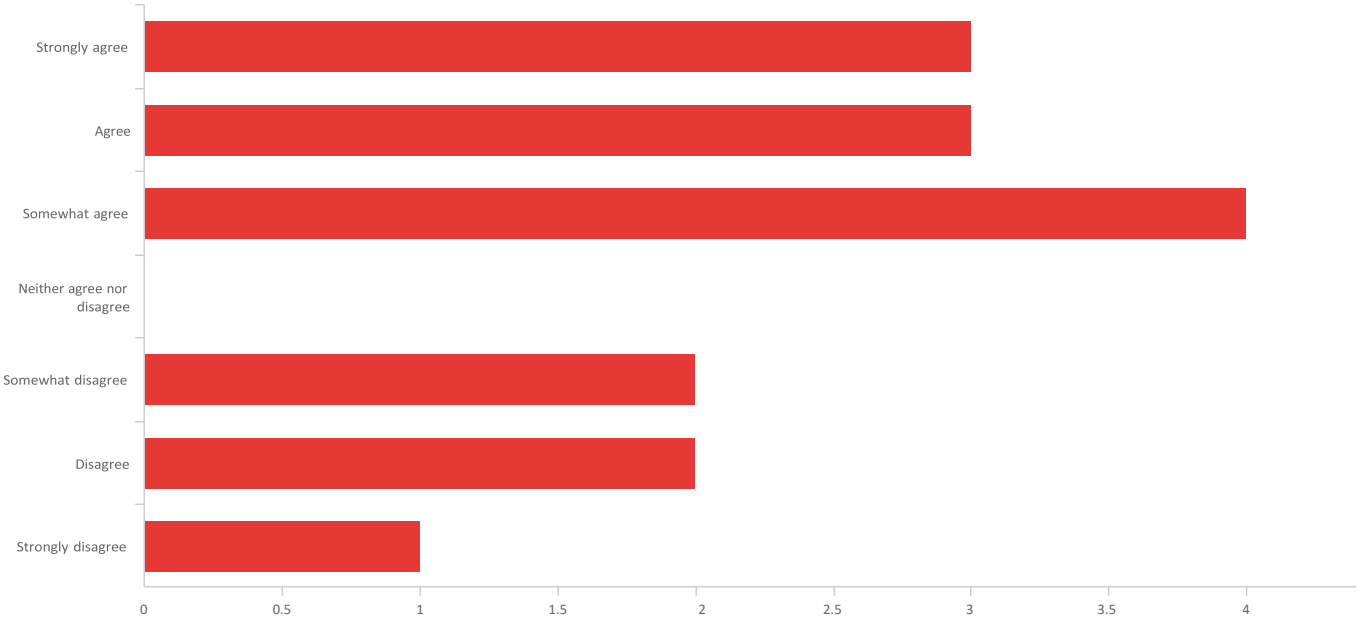
#	Field	Choice Count
1	Strongly agree	20.00% 3
2	Agree	26.67% 4
3	Somewhat agree	0.00% 0
4	Neither agree nor disagree	6.67% 1

5	Somewhat disagree	33.33%	5
6	Disagree	6.67%	1
7	Strongly disagree	6.67%	1

15

Showing rows 1 - 8 of 8

Q16 - I believe non-code contributors in OSS projects should be given the same amount of recognition as code contributors.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	I believe non-code contributors in OSS projects should be given the same amount of recognition as code contributors.	1.00	7.00	3.33	1.92	3.69	15

#	Field	Choice Count
1	Strongly agree	20.00% 3
2	Agree	20.00% 3

3	Somewhat agree	26.67%	4
4	Neither agree nor disagree	0.00%	0
5	Somewhat disagree	13.33%	2
6	Disagree	13.33%	2
7	Strongly disagree	6.67%	1
			15

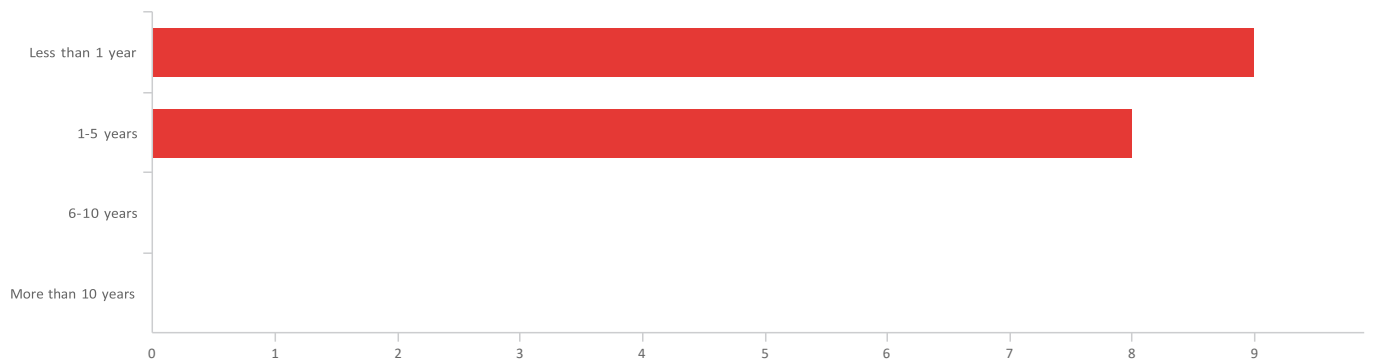
Showing rows 1 - 8 of 8

5. End of Report

Default Report

CS 560: Survey Design - Amreeta Chatterjee
November 30, 2020 3:29 PM MST

Q2 - For how long have you been contributing to OSS?



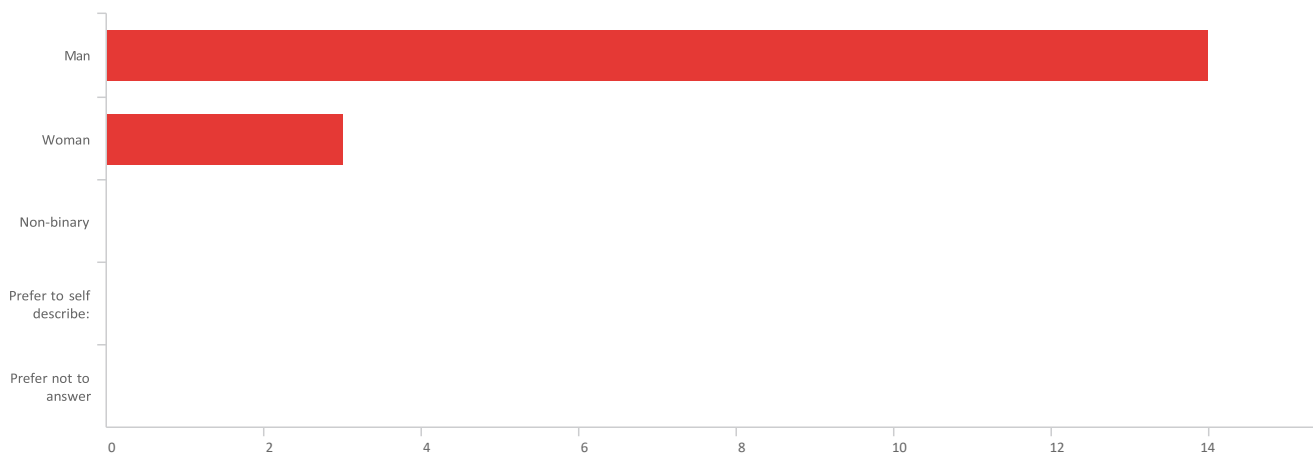
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	For how long have you been contributing to OSS?	1.00	2.00	1.47	0.50	0.25	17

#	Field	Choice Count
1	Less than 1 year	52.94% 9
2	1-5 years	47.06% 8
3	6-10 years	0.00% 0
4	More than 10 years	0.00% 0

17

Showing rows 1 - 5 of 5

Q3 - How do you identify yourself?



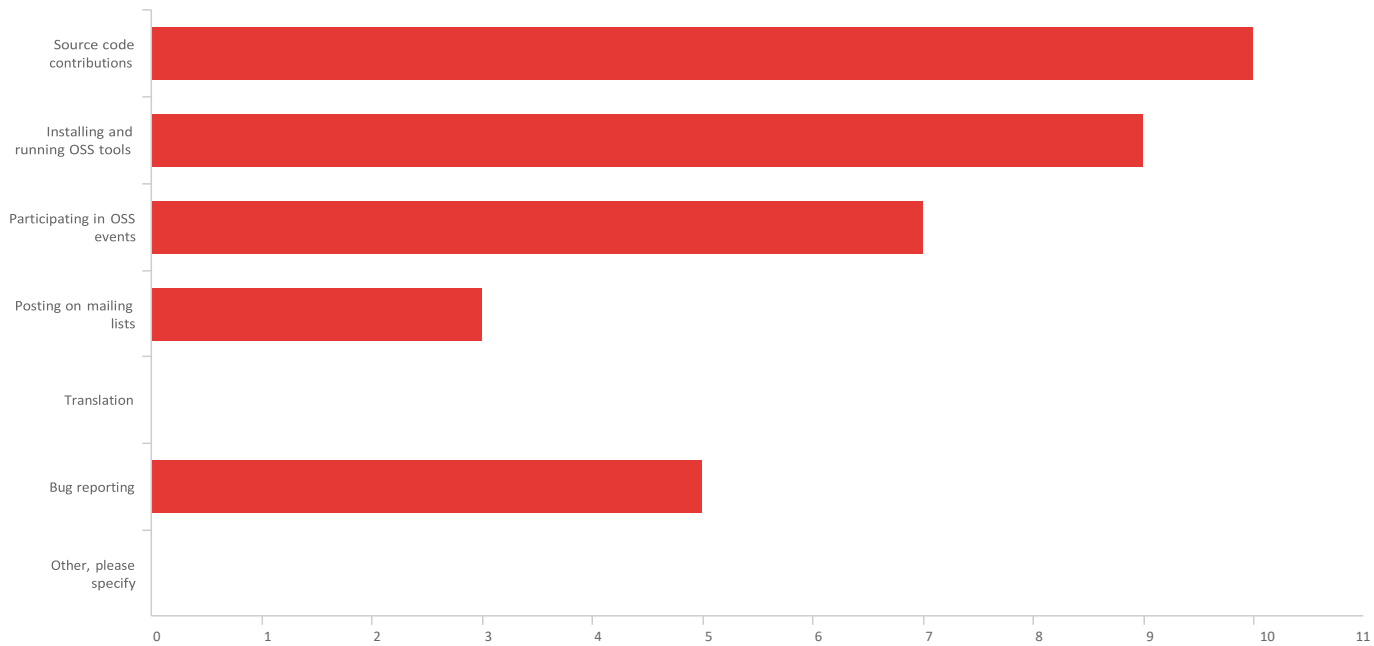
#	Field	Choice Count
1	Man	82.35% 14
2	Woman	17.65% 3
3	Non-binary	0.00% 0
4	Prefer to self describe:	0.00% 0
5	Prefer not to answer	0.00% 0

Showing rows 1 - 6 of 6

Q3_4_TEXT - Prefer to self describe:

Prefer to self describe:

Q4 - How would you describe your initial interactions in OSS?



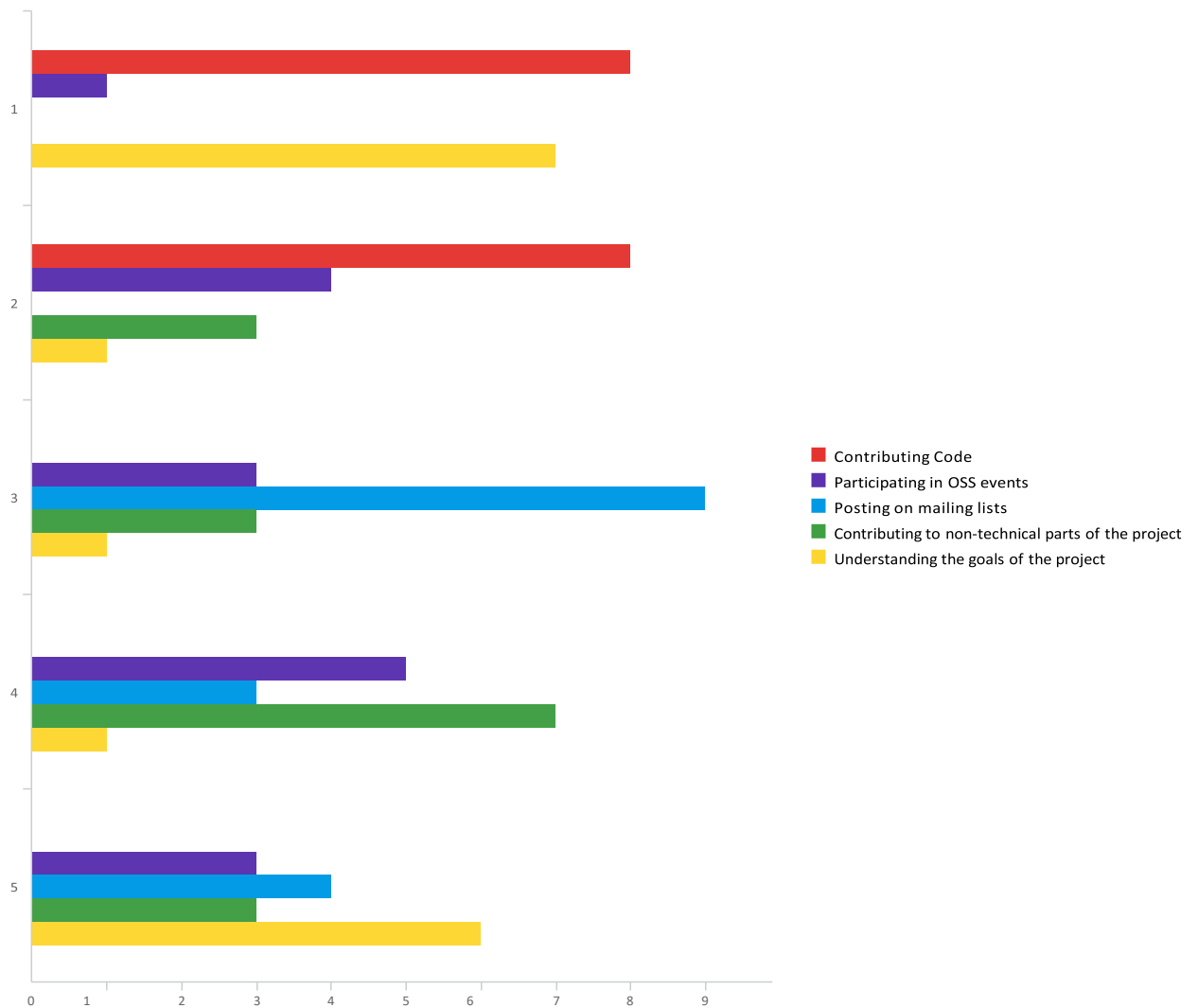
#	Field	Choice Count
1	Source code contributions	29.41% 10
2	Installing and running OSS tools	26.47% 9
3	Participating in OSS events	20.59% 7
4	Posting on mailing lists	8.82% 3
5	Translation	0.00% 0
6	Bug reporting	14.71% 5

Showing rows 1 - 8 of 8

Q4_7 TEXT - Other, please specify

r, please specify
Oth

Q5 - Rank the following interactions in order of their importance to becoming a core contributor in OSS, according to you.

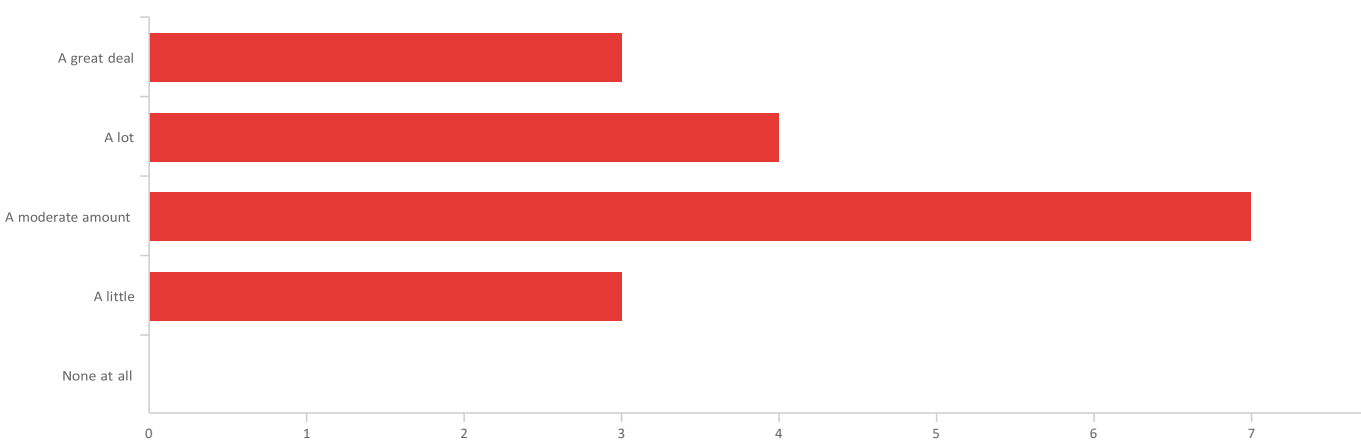


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Contributing Code	1.00	2.00	1.50	0.50	0.25	16
2	Participating in OSS events	1.00	5.00	3.31	1.21	1.46	16
3	Posting on mailing lists	3.00	5.00	3.69	0.85	0.71	16
4	Contributing to non-technical parts of the project	2.00	5.00	3.63	0.99	0.98	16
5	Understanding the goals of the project	1.00	5.00	2.88	1.83	3.36	16

#	Field	1	2	3	4	5	Total
1	Contributing Code	50.00% 8	50.00% 8	0.00% 0	0.00% 0	0.00% 0	16
2	Participating in OSS events	6.25% 1	25.00% 4	18.75% 3	31.25% 5	18.75% 3	16
3	Posting on mailing lists	0.00% 0	0.00% 0	56.25% 9	18.75% 3	25.00% 4	16
4	Contributing to non-technical parts of the project	0.00% 0	18.75% 3	18.75% 3	43.75% 7	18.75% 3	16
5	Understanding the goals of the project	43.75% 7	6.25% 1	6.25% 1	6.25% 1	37.50% 6	16

Showing rows 1 - 5 of 5

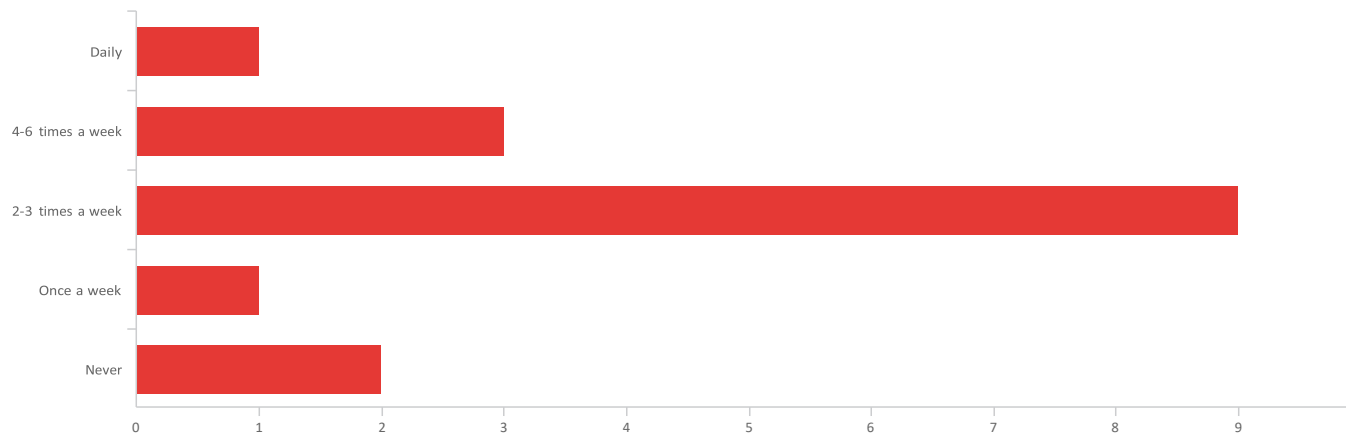
Q6 - How aware were you about the functionality of the project when you started contributing?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	ut the functionality of the project when you started contributing?	1.00	4.00	2.59	0.97	0.95	17

#	Field	Choice Count
1	A great deal	17.65% 3
2	A lot	23.53% 4
3	A moderate amount	41.18% 7
4	A little	17.65% 3
5	None at all	0.00% 0
		17

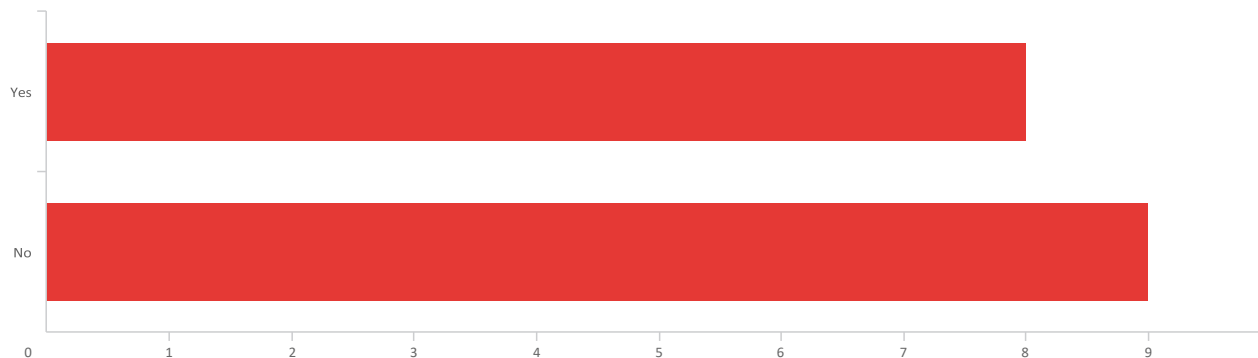
Q7 - During your first year of contributing to OSS, how much time did you spend on contributions per week?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	During your first year of contributing to OSS, how much time did you spend on contributions per week?	1.00	5.00	3.00	1.00	1.00	16

#	Field	Choice Count
1	Daily	6.25% 1
2	4-6 times a week	18.75% 3
3	2-3 times a week	56.25% 9
4	Once a week	6.25% 1
5	Never	12.50% 2
		16

Q8 - During your first year of contributing to OSS, did you make contributions to different OSS projects?

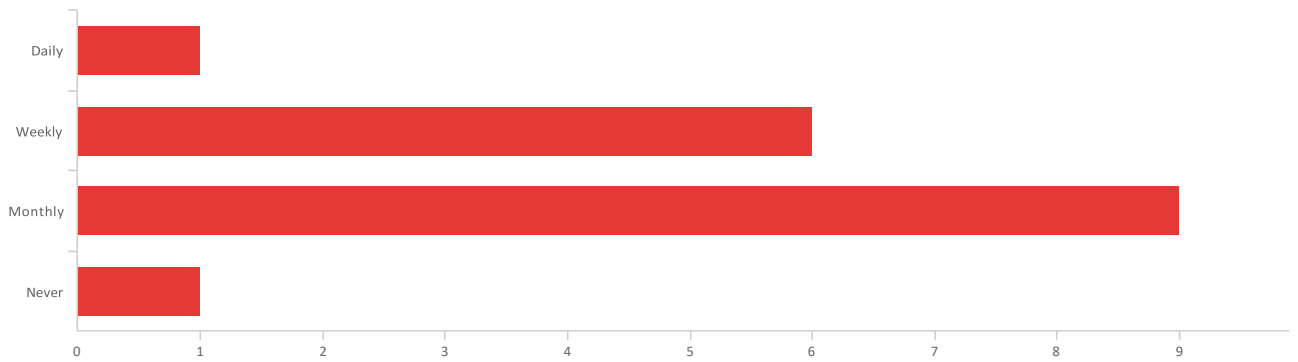


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	During your first year of contributing to OSS, did you make contributions to different OSS projects?	1.00	2.00	1.53	0.50	0.25	17

#	Field	Choice Count
1	Yes	47.06% 8
2	No	52.94% 9
		17

Showing rows 1 - 3 of 3

Q9 - How frequently did you submit a pull-request to different projects in your first year of contributions?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	I you submit a pull-request to different projects in your first year of contributions?	1.00	4.00	2.59	0.69	0.48	17

#	Field	Choice Count
1	Daily	5.88% 1
2	Weekly	35.29% 6
3	Monthly	52.94% 9
4	Never	5.88% 1
		17

Showing rows 1 - 5 of 5

Q10 - In your opinion, what initial interactions pave the path towards becoming a core developer in OSS?

In your opinion, what initial interactions pave the path towards becoming a...

Interactions with people who actively contribute in OSS.

I don't quite get what do you mean by initial interactions. But I think effective communication and be aware of the goal of the whole project is very important for a core developer.

Understanding the project

communication between teammates

Collabrations

Hard to answer

Quality of contributions and supporting the project's goal

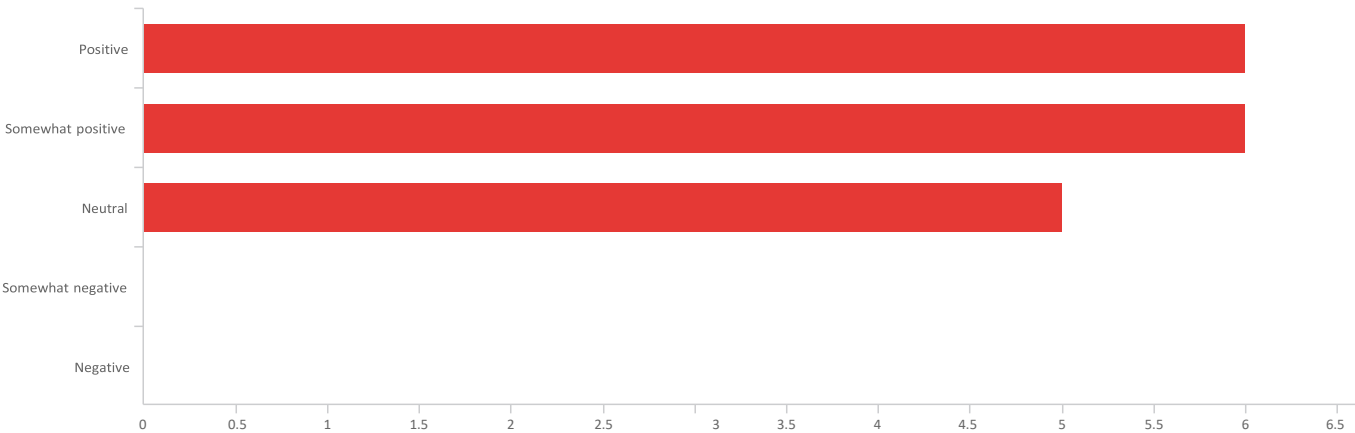
Weekly meeting

Fixing Bugs

Trying to be as involved in the community as possible and contributing in every way you can

Submit high quality pull requests frequently.

Q11 - How would you describe your initial communications with established members of the OSS project?

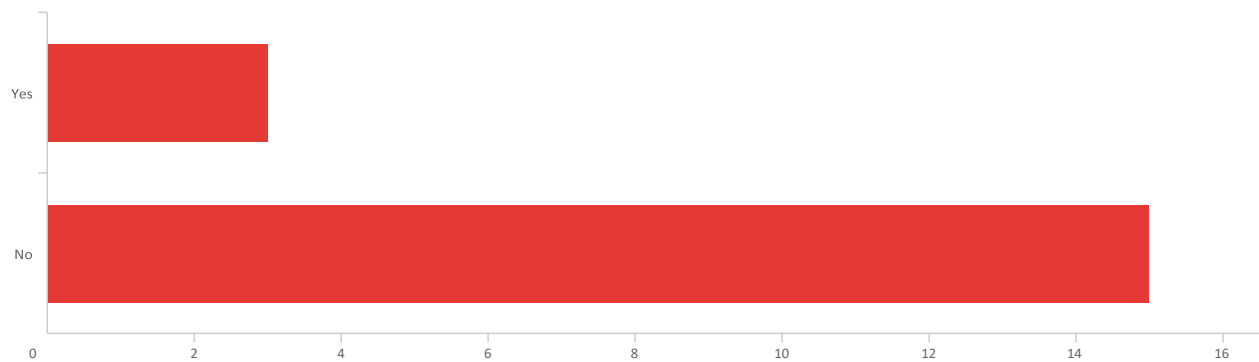


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How would you describe your initial communications with established members of the OSS project?	1.00	3.00	1.94	0.80	0.64	17

#	Field	Choice Count
1	Positive	35.29% 6
2	Somewhat positive	35.29% 6
3	Neutral	29.41% 5
4	Somewhat negative	0.00% 0
5	Negative	0.00% 0
		17

Showing rows 1 - 6 of 6

Q12 - Did you have a mentor work with you prior to your first contribution?



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Did you have a mentor work with you prior to your first contribution?	1.00	2.00	1.83	0.37	0.14	18

#	Field	Choice Count
1	Yes	16.67% 3

2 No

83.33% 15

18

Showing rows 1 - 3 of 3

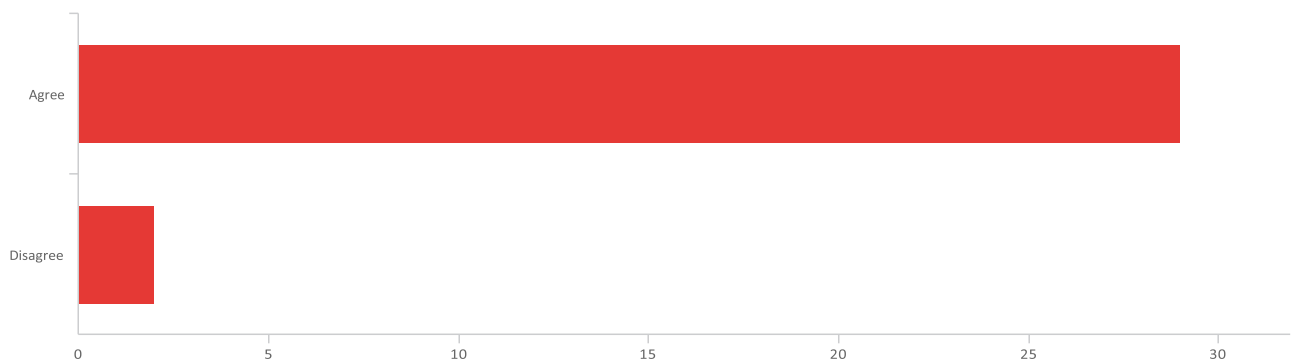
6. End of Report

Default Report

Survey on Code Vs Non-code contribution to Open Source Software

November 19, 2020 12:08 AM MST

Q4 - 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.



#

Field

Minimum

Maximum

Mean

Std
Deviation

Variance

Count

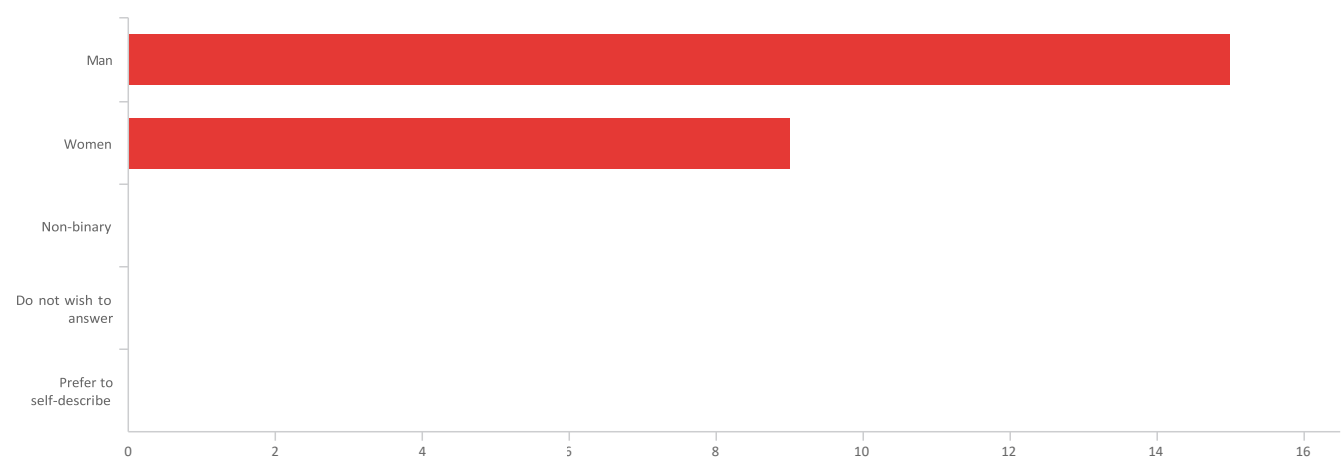
1	Thank you for participating in the survey. We expect that this survey will take 7-10 minutes of your time. Our Goal: This survey is part of a research study to examine the contributions to Open Source Software (OSS). Our goal is to understand the importance of various kinds of contributions and their 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.	1.00	2.00	1.06	0.25	0.06	31
---	---	------	------	------	------	------	----

#	Field	Choice Count
1	Agree	93.55% 29
2	Disagree	6.45% 2
#	Field	Choice Count

31

Showing rows 1 - 3 of 3

Q1 - Please select your Gender



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Please select your Gender - Selected Choice	40.00	41.00	40.38	0.48	0.23	24

#	Field	Choice Count
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

Showing rows 1 - 6 of 6

Q1_44_TEXT - Prefer to self-describe

Prefer to self-describe

Q2 - List of Countries

f
g
h
a
n
i
s
t
a
n

A
l
b
a
n
i
a

A
l
g
e
r
i
a

A
n
d
o
r
r
a

A
n
g
o
l
a

Antigua
and
Barbuda

A
r
g
e
n
t
i
n
a

A
r
m
e
n
i
a

A
u
s
t

r
a
l
i
a

A
u
s
t
r
i
a

A
z
e
r
b
a
i
j
a
n

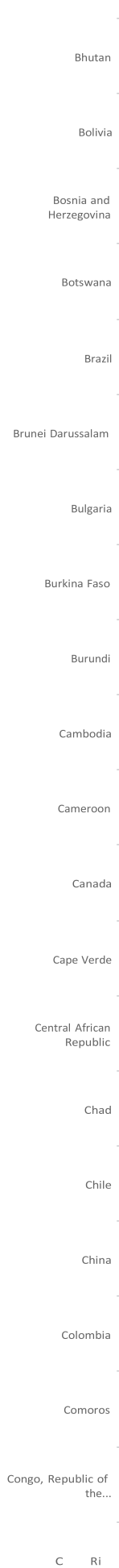
B
a
h
a
m
a
s

B
a
h
r
a
i
n

B
a
n
g
l
a
d
e
s
h

B
a
r
b
a
d
o
s

B
e
l
a
r
u
s



Côte
d'Ivoire

Croatia

Cuba

C
y
p
r
u
s

Czech
Republic

Democra
tic
People's
Republic
of Korea

Democrat

ic

Republic

of the

Congo

Denmark

D
j
i
b
o
u
t
i

Dominica

Dominica
n
Republic

Ecuador

Egypt

El
Salvador

Equatorial
Guinea

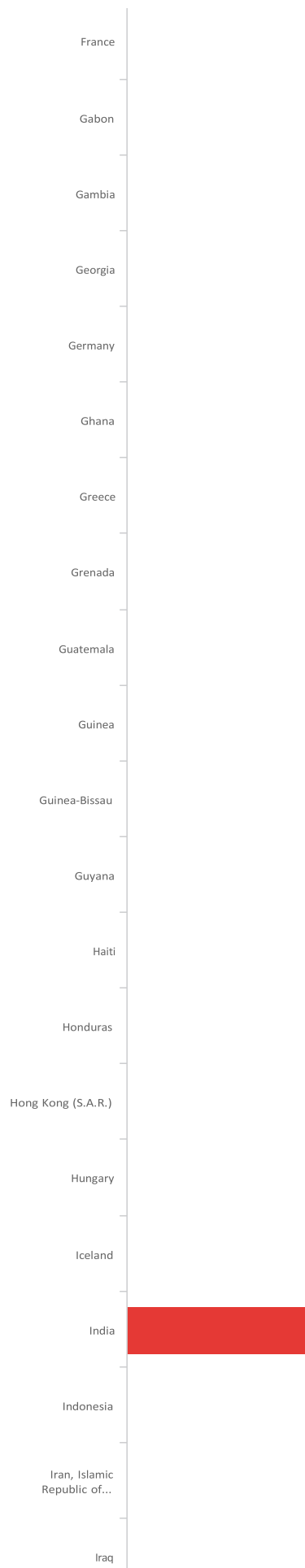
Eritrea

E
s
t
o
n
i
a

Ethiopia

Fiji

F
i
n
l
a
n
d



Israel

Italy

Jamaica

J
a
p
a
n

Jordan

Kazakhstan

K
e
n
y
a

Kiribati

K
u
w
a
i
t

Kyrgyzstan

Lao
People's
Democratic
Republic

Latvia

Lebanon

Lesotho

L
i
b
e
r

i
a

Libyan
Arab
Jamahiriya

Liechtenstein

Lithuania

urg
Madagas
car

Malawi

M
a
l
a
y
s
i
a

M
a
l
d
i
v
e
s

Mali

Malta

Mars
hall
Islan
ds

Mauritani
a

M
a
u
r
i
t
i
u
s

Mexico

Micronesi
a,
Federated
States
of...

Monaco

M
o
n

g
o
l
i
a

Montene
gro

M
o
r
o
c
c
o

Mozambi
que

M
y
a
n
m
a
r

Namibia

Nauru

Nepal

Netherlan
ds

Zealand

Nicaragua

Niger

Nigeria

North
Korea

Norway

Oman

Pakistan

Palau

P
a
n
a
m
a

Papua
New
Guinea

Paraguay

Peru

Philippine
s

Poland

Portugal

Qatar

Republic
of Korea

Republic
of
Moldova

Romania

i

F d

ti
Russian
Federat
ion

R
w
a
n
d
a

Saint Kitts
and
Nevis

Saint
Lucia

Saint

Vincent

and the

Grenadin

es Samoa

San
Marino

Sao Tome
and

P
r
i
n
c
i
p
e

Saudi
Arabia

S
e
n
e
g
a
l

Serbia

Seychelle
s

Sierra
Leone

Singapore

S
l
o
v
a
k
i
a

Slovenia

Solom
on
Island
s

S
o
m
a
l
i
a

South
Africa

South
Korea

Spain

S
u
d
a
n

Suriname

Swaziland

Sweden

Switzerla
nd

Syrian
Arab
Republic

Tajikistan

Thailand

The
former
Yugoslav
Republic
of
Macedoni
a

Timor-
Leste

Togo

Tonga

Trinidad
and
Tobago

T
u
n
i
s
i
a

T
u
r
k

e
y

Turkmeni
stan

Tuvalu

Uganda

Ukraine

United
Arab
Emirates

United Kingdom of
Great Britain and
Northern Ireland

United Republic of
Tanzania

United States of
America

Uruguay

Uzbekistan

Vanuatu

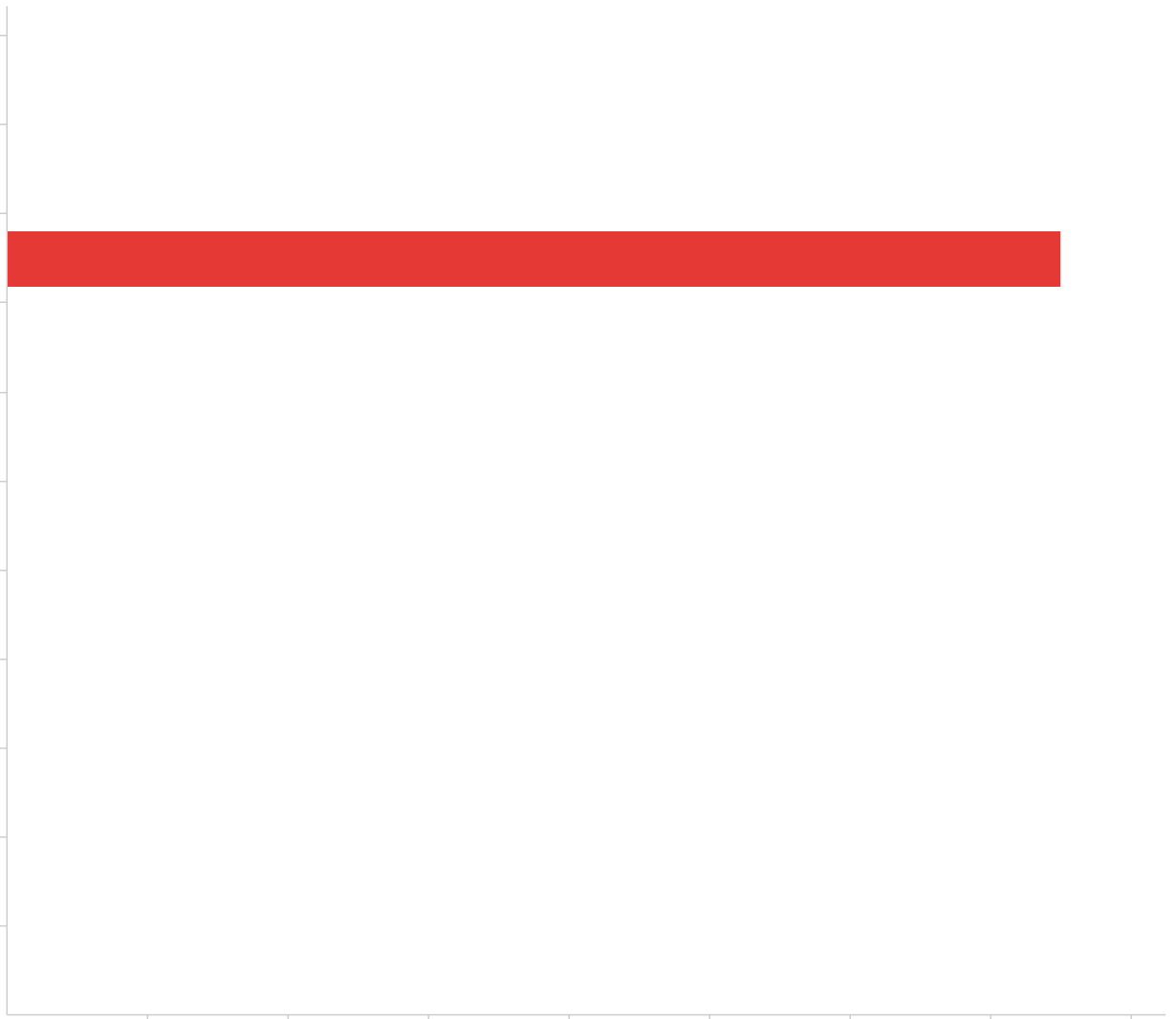
Venezuela,
Bolivarian Republic
of...

Viet Nam

Yemen

Zambia

Zimbabwe



0246810121416

#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	List of Countries	36.00	187.00	155.55	55.44	3073.15	20

#	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	Choice Count
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		Choice Count	
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	Choice Count	
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 Count
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	Choice	Count
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%	0
145	Saint Kitts and Nevis	0.00%	0
146	Saint Lucia	0.00%	0

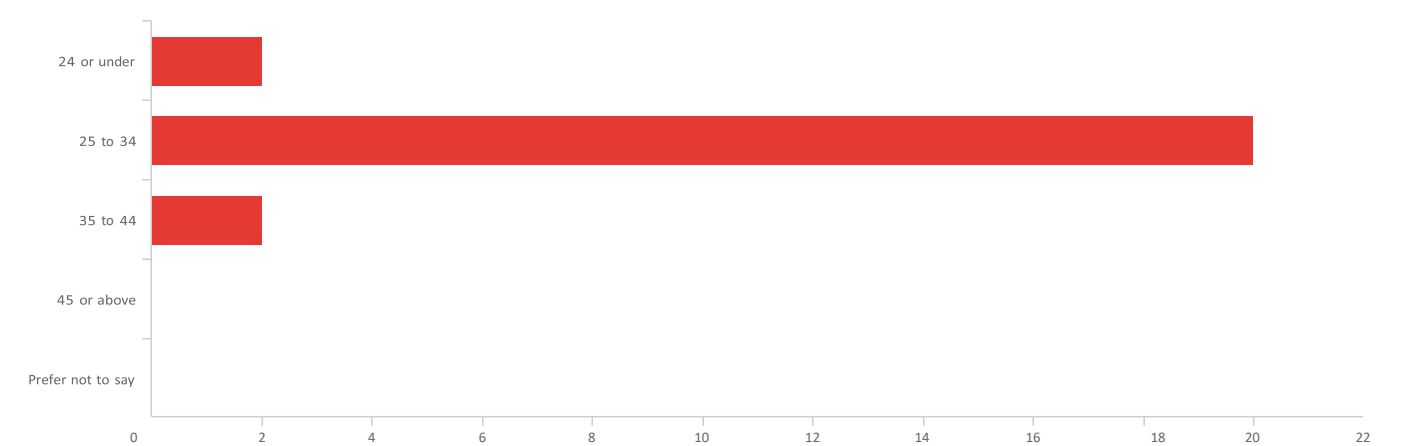
147	Saint Vincent and the Grenadines	0.00%	0
148	Samoa	0.00%	0
149	San Marino	0.00%	0
150	Sao Tome and Principe	0.00%	0
151	Saudi Arabia		
#	Field		Choice Count
152	Senegal	0.00%	0
153	Serbia	0.00%	0
154	Seychelles	0.00%	0
155	Sierra Leone	0.00%	0
156	Singapore	0.00%	0
157	Slovakia	0.00%	0
158	Slovenia	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 Republic 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 Count
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



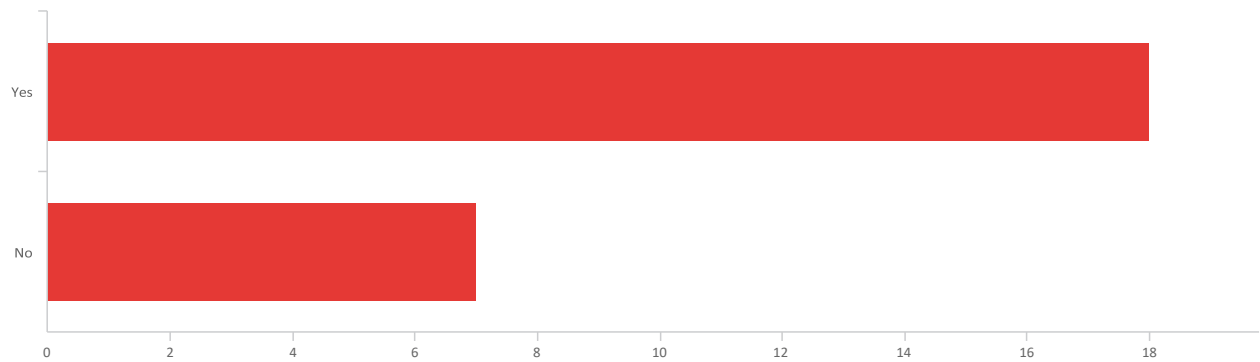
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
---	-------	---------	---------	------	---------------	----------	-------

1	Select the range in which your age falls	1.00	3.00	2.00	0.41	0.17	24
---	--	------	------	------	------	------	----

#	Field	Choice Count
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
5	Prefer not to say	0.00% 0
		24

Showing rows 1 - 6 of 6

Q4 - Have you contributed to Open Source Software?

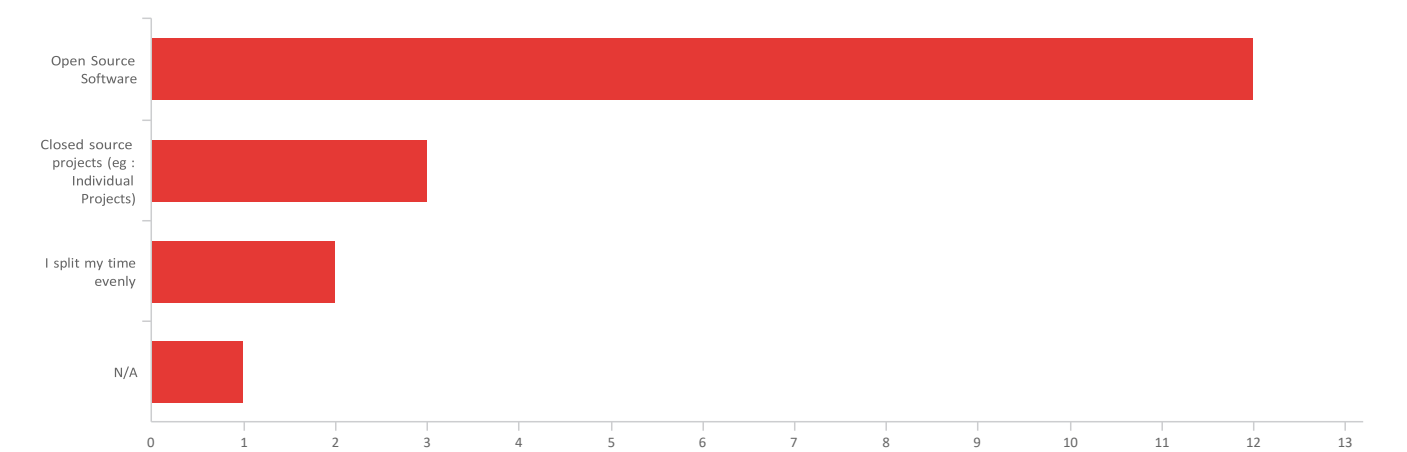


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Have you contributed to Open Source Software?	1.00	2.00	1.28	0.45	0.20	25

#	Field	Choice Count
1	Yes	72.00% 18
2	No	28.00% 7
		25

Showing rows 1 - 3 of 3

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



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	What kind of projects do you spend most of your time on?	1.00	4.00	1.56	0.90	0.80	18

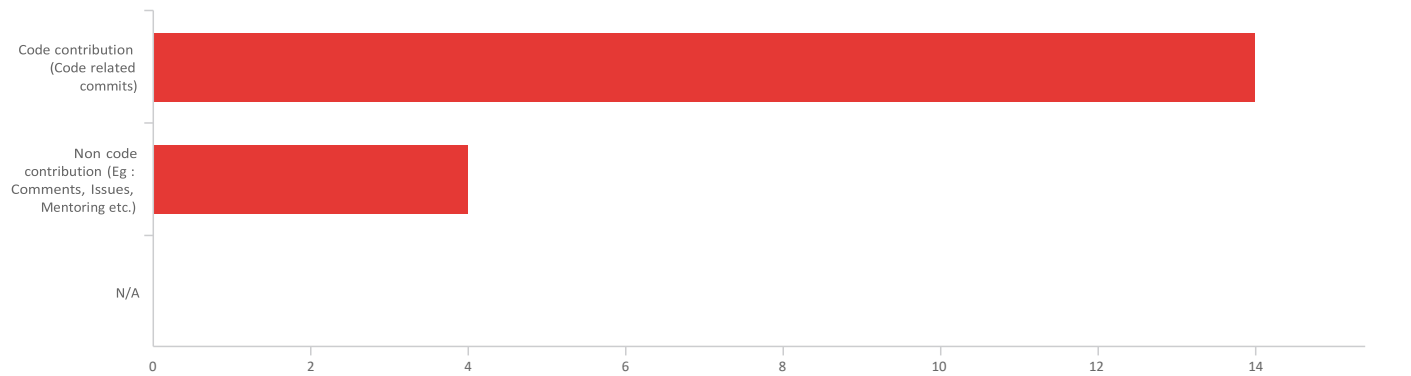
#	Field	Choice Count
1	Open Source Software	66.67% 12
2	Closed source projects (eg : Individual Projects)	16.67% 3
3	I split my time evenly	11.11% 2

4	N/A	5.56%	1
---	-----	-------	---

18

Showing rows 1 - 5 of 5

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



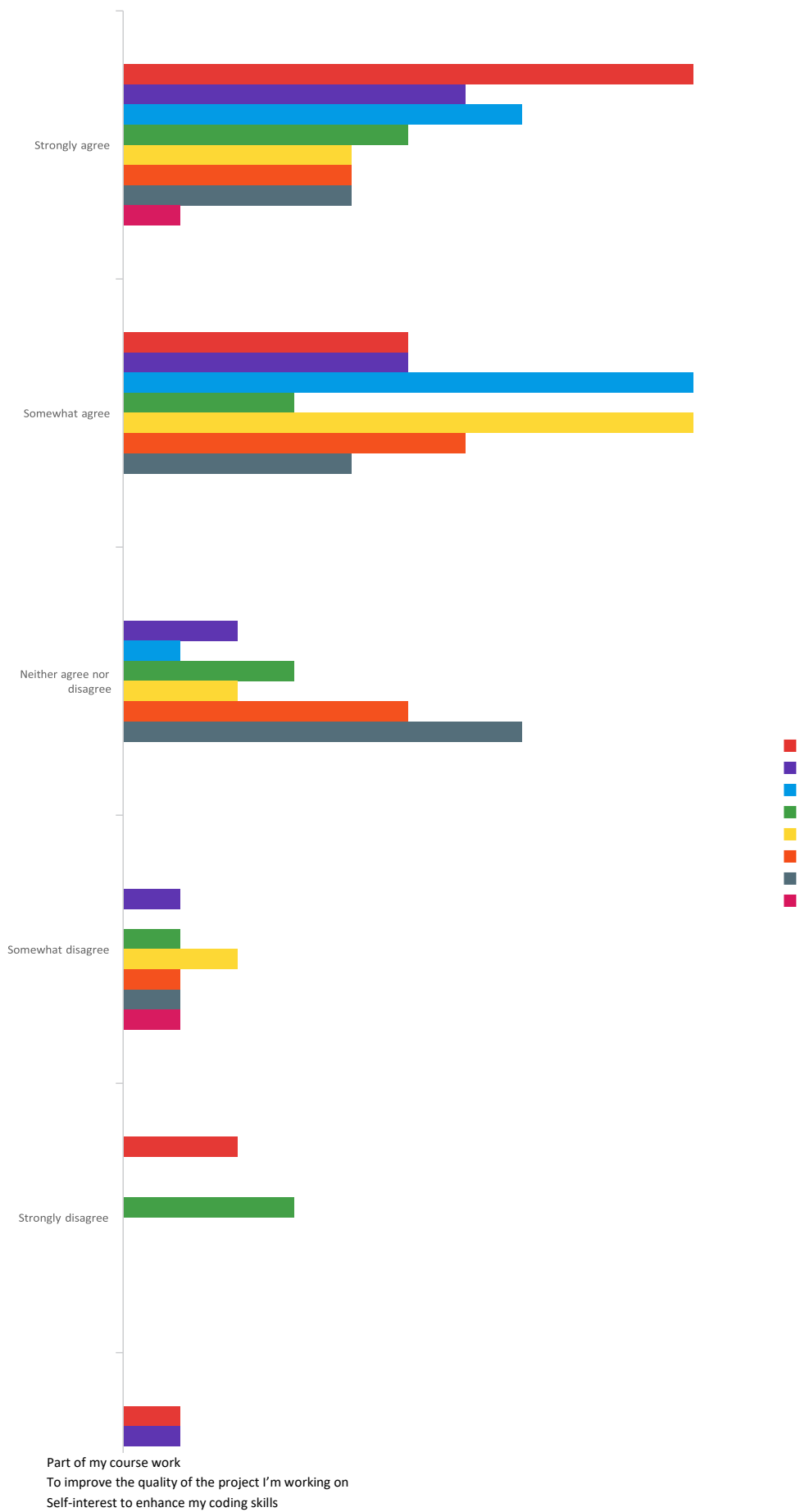
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Which is your major contribution in Open Source Software?	1.00	2.00	1.22	0.42	0.17	18

#	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

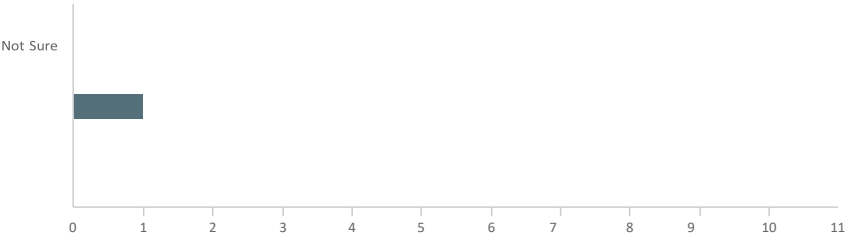
18

Showing rows 1 - 4 of 4

Q7 - Select the reason for your contribution to OSS



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



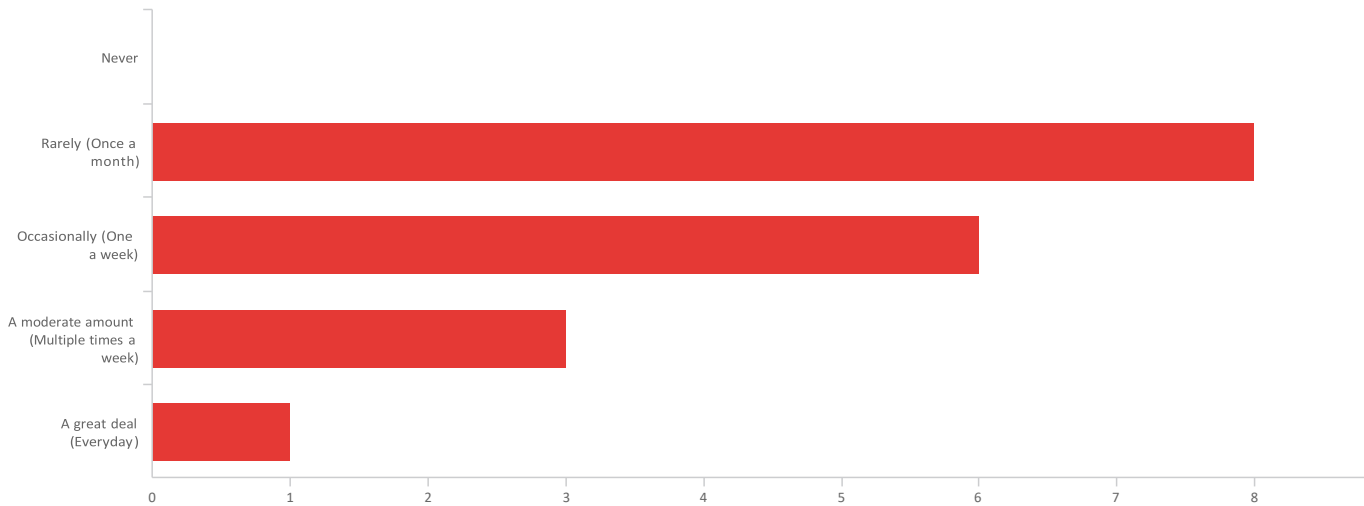
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	Part of my course work	1.00	6.00	2.00	1.56	2.44	18
2	To improve the quality of the project I'm working on	1.00	6.00	2.13	1.36	1.85	15
3	Self-interest to enhance my coding skills	1.00	3.00	1.67	0.58	0.33	18
4	I am paid for my contribution	1.00	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	I 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		Neither agree		Somewhat nor		Strongly		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?



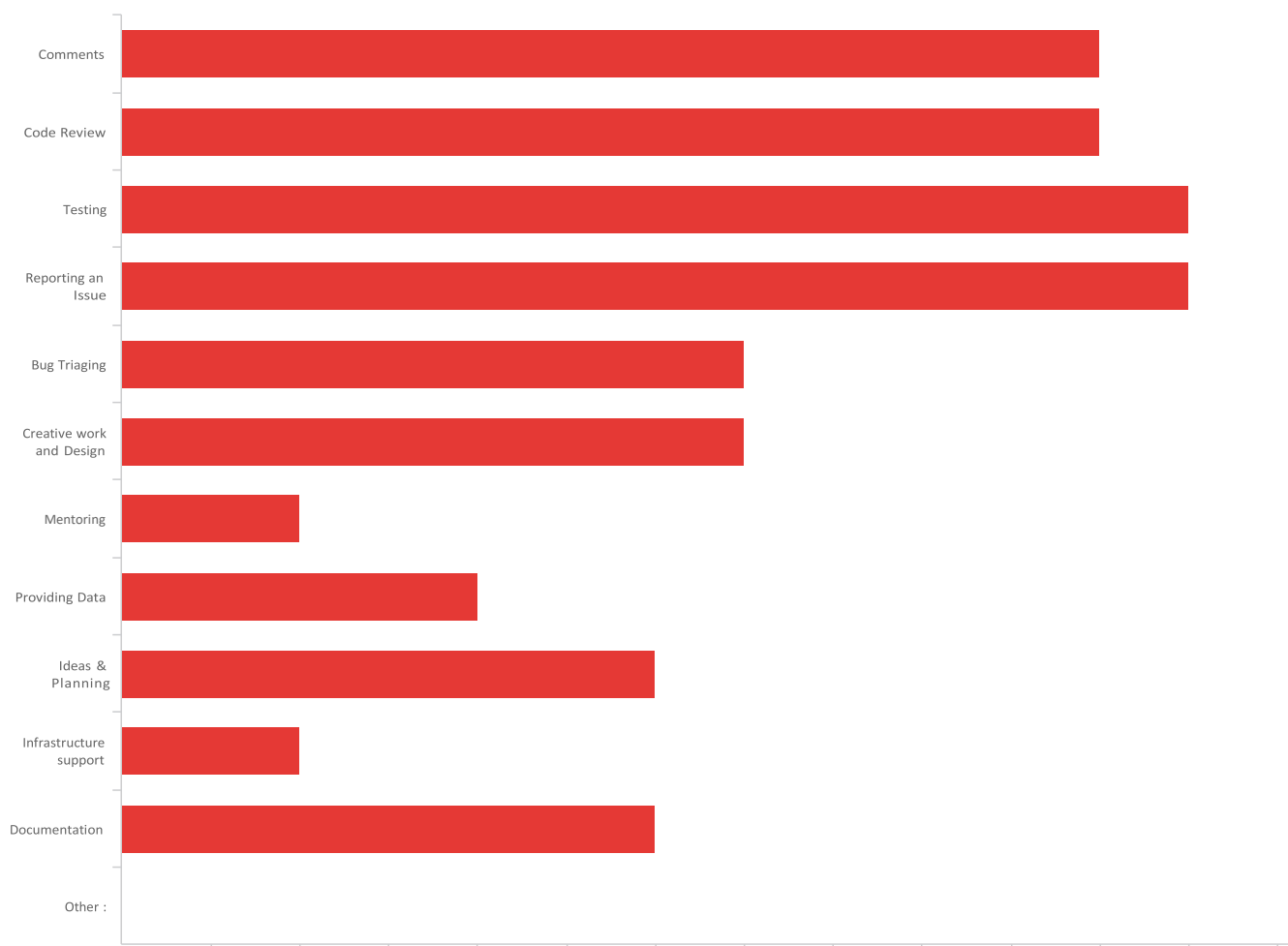
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How frequent would you contribute towards code in Open Source Software?	208.00	211.00	208.83	0.90	0.81	18

#	Field	Choice Count
---	-------	--------------

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.



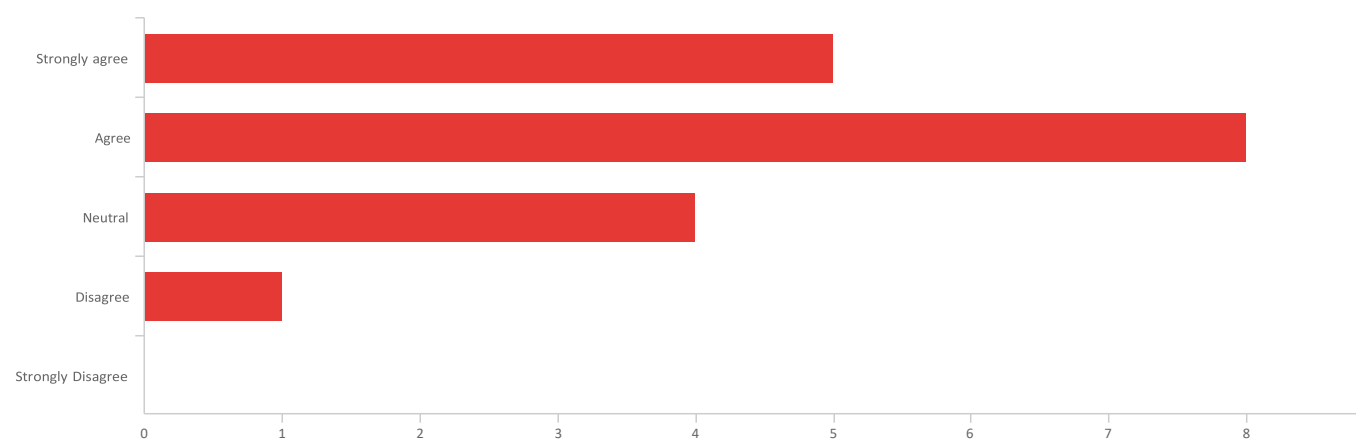
0 1 2 3 4 5 6 7 8 9 10 11 12 13

#	Field	Choice Count
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	7.50% 6
10	Infrastructure support	2.50% 2
11	Documentation	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 Code-contribution?

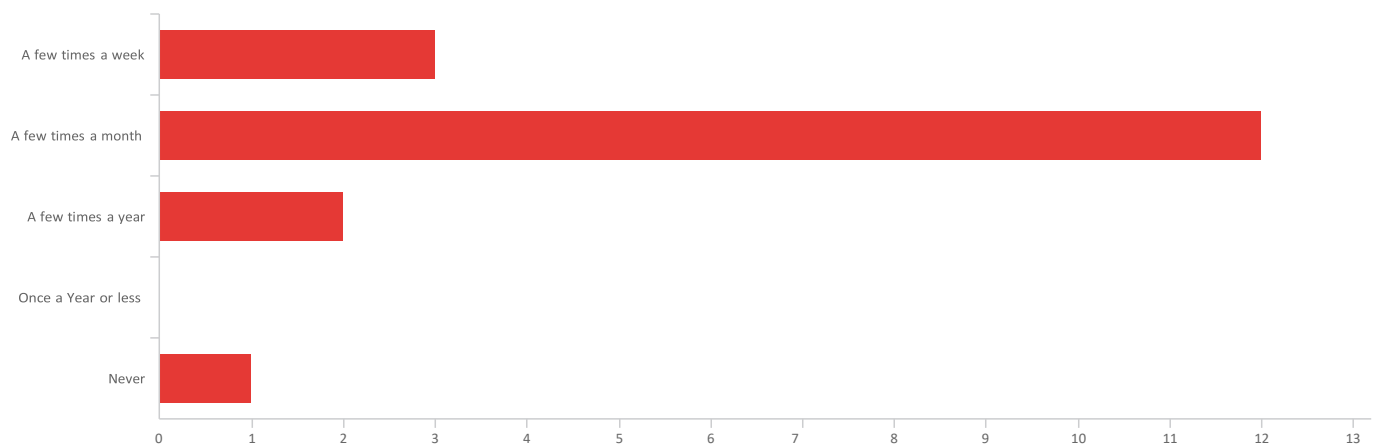


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	tribution is equally important as Codecontribution?	43.00	46.00	44.06	0.85	0.72	18

#	Field	Choice Count
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?

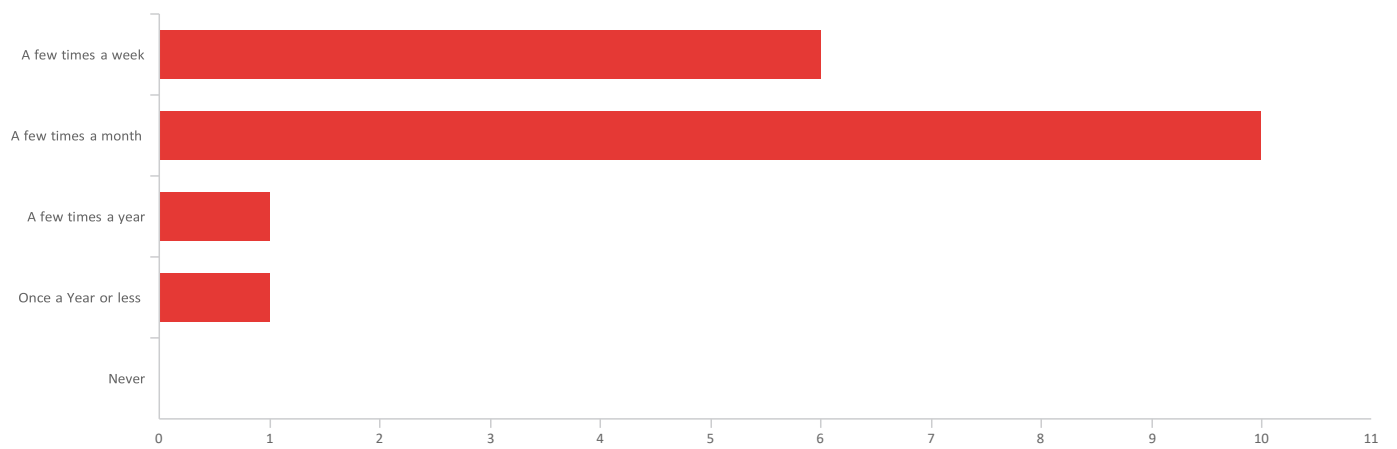


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How often do you contribute towards code contribution?	1.00	5.00	2.11	0.87	0.77	18

#	Field	Choice Count
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	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	How often do you contribute towards non-code contribution?	1.00	4.00	1.83	0.76	0.58	18

#	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
		18

Showing rows 1 - 6 of 6

7. End of Report

Appendix B

560_proj

Group 1 12/4/2020

```
library(ggplot2)
```

```
Data <- read.csv("D:/D drive contents/Fall 2020/CS560_Data Driven/MSR2/Final/RQ2test_data.csv")
head(Data)
```

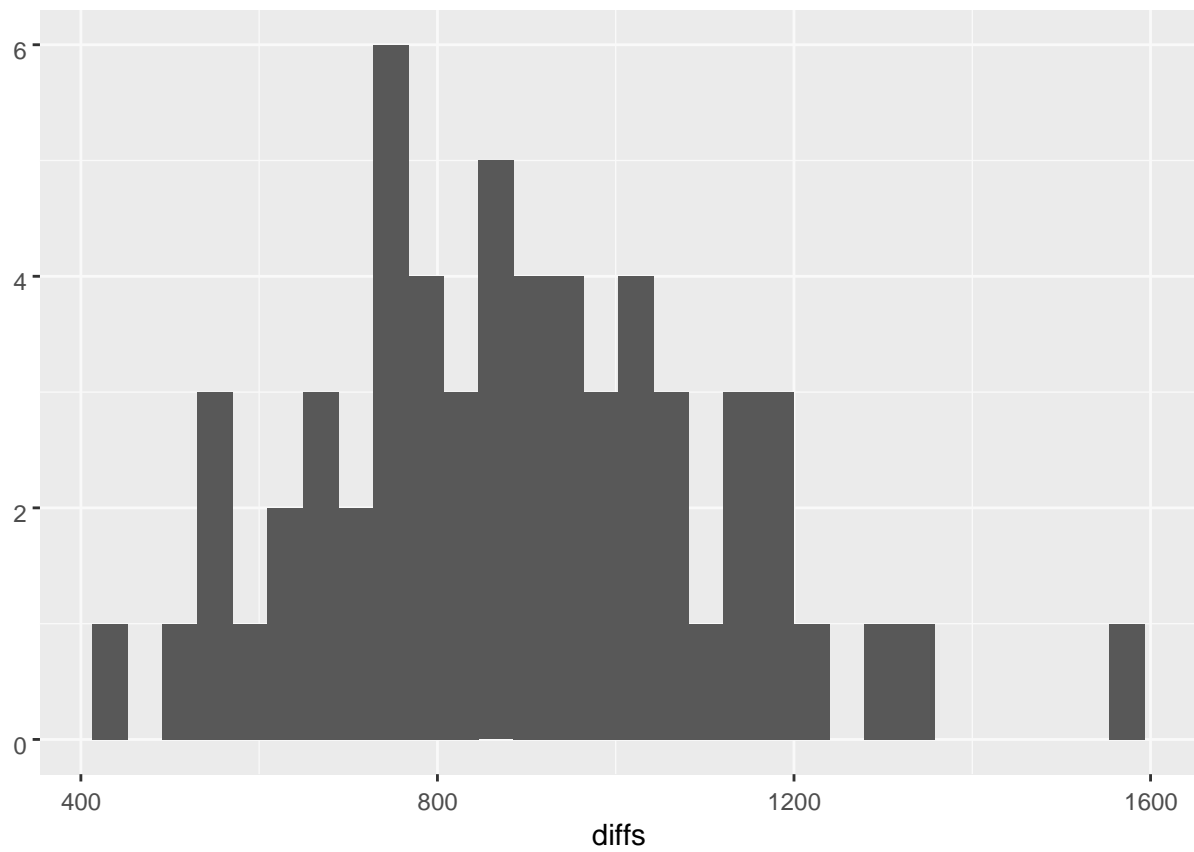
```
##      i..Date Code Non_Code
## 1 2015-01 268      911
## 2 2015-02 433     1447
## 3 2015-03 228      878
## 4 2015-04 265     1150
## 5 2015-05 273     1219
## 6 2015-06 248     1822
```

```
diffs <- with(Data, Non_Code-Code) diffs
```

```
## [1] 643 1014 650 885 946 1574 1167 1224 1152 1337 1068 1039 1001 894 1109
## [16] 938 847 543 913 1200 937 1068 783 795 1034 673 1041 817 653 741
## [31] 857 758 759 981 778 594 710 779 764 882 1173 1044 914 936 871 ## [46] 1317 762 433 766 1148 829 978 921 513
## [61] 618 823 538 1137 690 564
```

```
qplot(diffs, geom="histogram")
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
t.test(diffs)
```

```
##
## One Sample t-test
##
## data: diffs
## t = 30.689, df = 59, p-value < 2.2e-16 ## alternative hypothesis: true mean is
not equal to 0 ## 95 percent confidence interval: ## 833.8871 950.2129 ## sample
estimates:
## mean of x
##      892.05
```

