The 2020 Apache Software Foundation Community Survey

by Bitergia Analytics

Final Readout

The Research Team

Anita Sarma
Oregon State University

Daniel Izquierdo
Bitergia

Mariam Guizani
Oregon State University

Georg Link
Bitergia

Griselda Cuevas
The Apache Software Foundation

Special Thank You

Katia Rojas

Justin McLean

Kenn Knowles

Special Thank You

The ASF Community

Motivation

Survey design, data collection, and analysis aligned with the **ASF D&I strategy**.

2019 short term goals

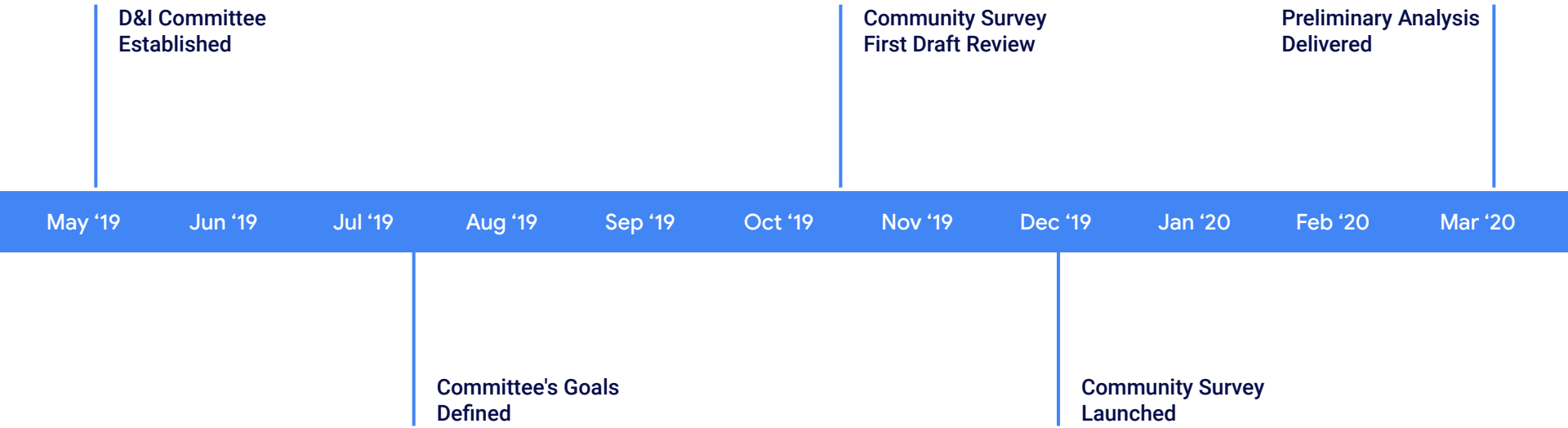
- **Gather scientific data** to study current status of Diversity and Inclusion at the ASF
- **Raise awareness** in our community about the importance of Diversity & Inclusion in the business, and in the open source industry
- **Find key indicators** to track over time

2020 medium term goals

- Instrument ASF's contributor funnel to **recommend a participation baseline** for underrepresented groups
- Create a toolkit to **address the top 3 entry barriers for new contributors** from under represented groups
- **Become a trusted thought partner** for PMCs when it comes to D&I

As stated in the 2019 and 2020 ASF EDI goals https://issues.apache.org/jira/browse/DI-25

The 2020 ASF Community Survey Timeline



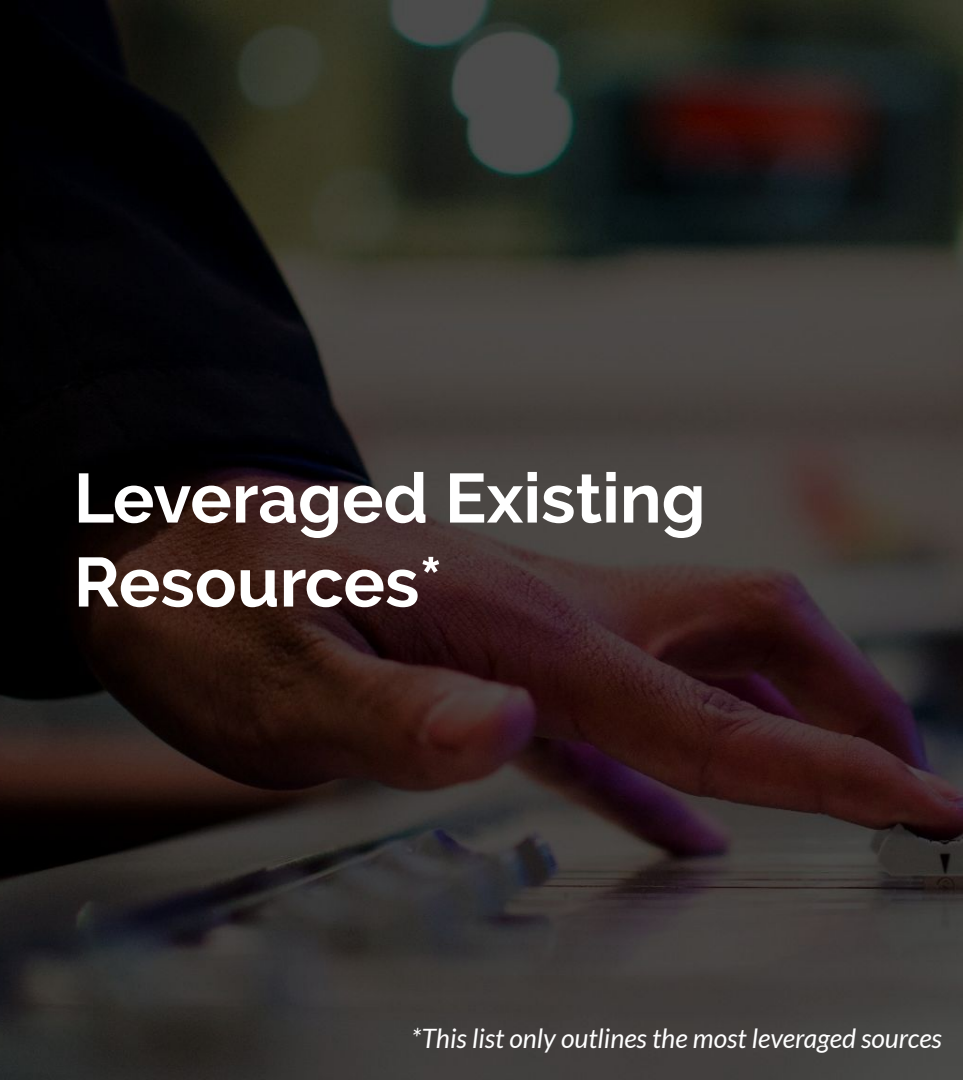
3 Efforts to gather data about the ASF community

The Community Survey

Project Metrics Analysis

New Contributor and Mentor Friction Logs

Survey Design



Leveraged Existing Resources*

**This list only outlines the most leveraged sources*

2016 ASF Committer Diversity Survey

Open Demographics

CHAOSS D&I Working Group

OpenStack Gender Diversity Report - Focus on mentorship

Stackoverflow Members Survey - Focus on Developer Role and Gender

Academic references omitted

Considerations were focused on topics covered, number of questions, time to complete, and extended to any participant in the community (past or present).

- Topics covered - extension of 2016 survey plus focus on understanding minorities challenges and barriers to contribution*.
- Number of questions limited to up to 25
- Time to complete the survey should not be more than 20 minutes
- Participants roles extended to anyone that has ever participated in the community

Reaching out to participants was based on:

- Emails to the committers, and D&I lists.
- Social networks
- Booth activity at FOSDEM

Limesurvey is the tool used to run the online survey. This is the professional hosting service on top of the open source project**.

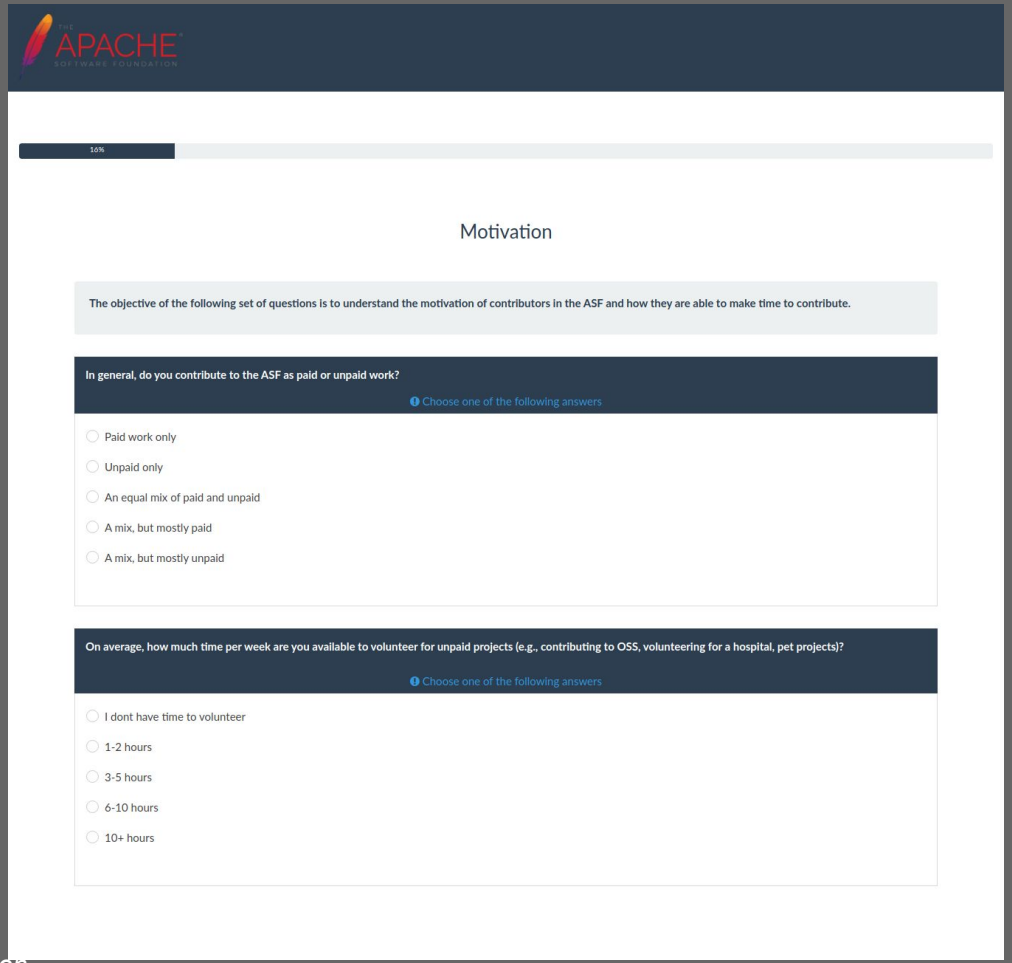
* As stated in the 2019 and 2020 ASF EDI goals https://issues.apache.org/jira/browse/DI-25
** https://www.limesurvey.org/index.php/community

Survey

Survey* is split into:

1. Contributor Role and Tenure (4 questions)
2. Motivation (2)
3. Availability of Protocols / Guidelines (2)
4. Support for Newcomers (4)
5. Diversity and Inclusion (11)
6. Wrap up (3)

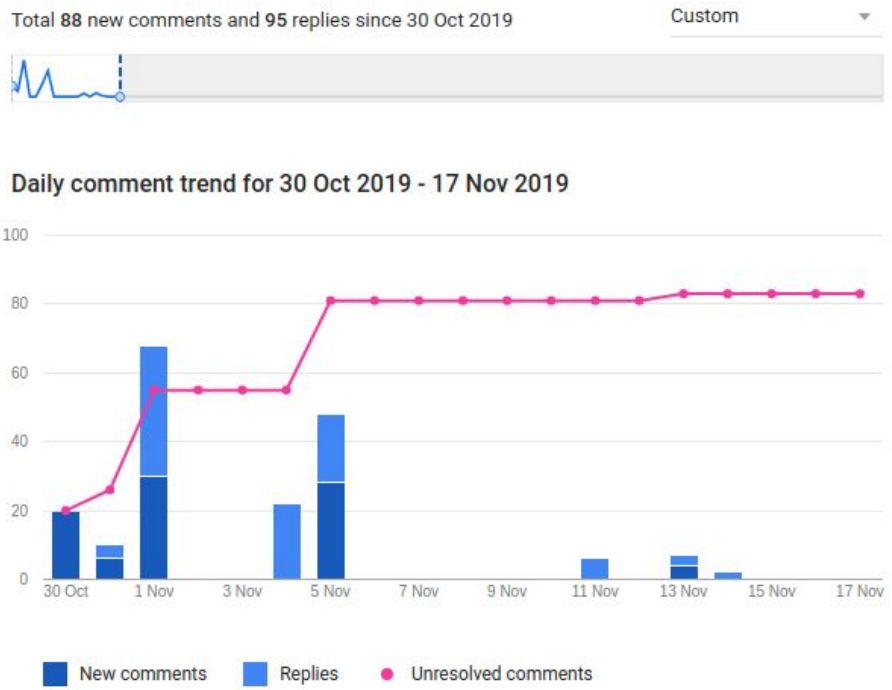
Each section contains a motivation introductory paragraph.



*Questions are listed at the end of this presentation

Process

- 2 Internal iteration followed by feedback process. Meeting notes available*.
- Public feedback and discussions followed using Google Docs**.
    - Comments were left “unresolved” to leave communication trace
- Voting and veto process followed as usual at the ASF under the Equity, Diversity, and Inclusion PMC.



Referenced Google Docs Activity

* E.g., https://cwiki.apache.org/confluence/display/EDI/2019-11-13+Meeting+notes+survey
** E.g., https://docs.google.com/document/d/1dfyMZ58ZENOO21sZvj__g1fjIP-BJfDuIECt066Dsuc/edit

Challenges

- Nomenclature differences across countries
- Feedback retrieval process (initially done in Google Docs)
- Several iterations were needed to address all requirements and needs
- Highly motivated community with good ideas, but hard to get all of them reflected in the document.
- Balance between ideas, number of questions, and total time to answer the survey
- Providing incentives to respond was an issue because of legal reasons
- Delay due to open discussions
- Many incomplete surveys (ignoring non-complete surveys)

Individual Identification Challenges

- Some of the questions were optional to avoid identity recognition
- Data should be aggregated when having a smaller amount of data points
- The focus of the analysis is now extended to all of the participants and not only committers

New Challenges based on Feedback

- Not Applicable option should be allowed, but leads to data loss

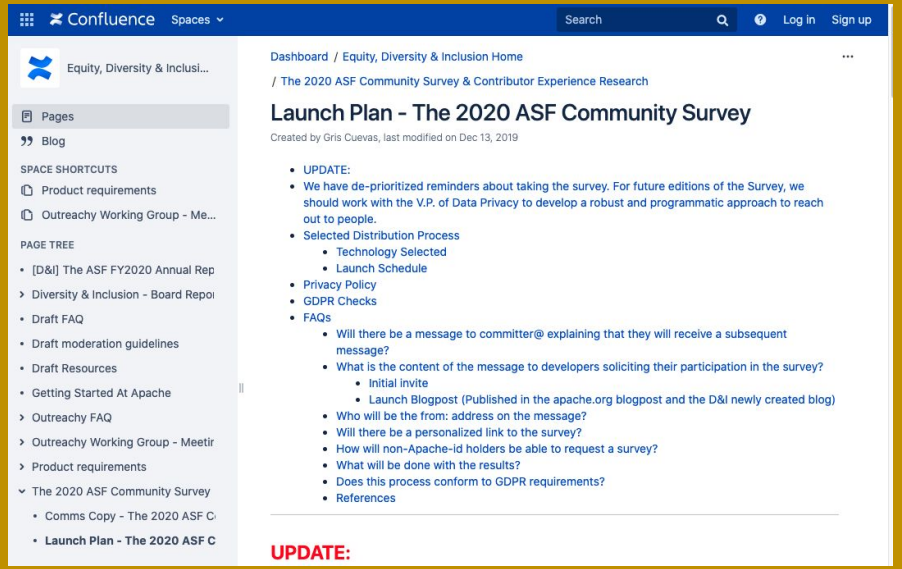
Reproducibility

Survey methodology and questions will be published in the D&I wiki so can be leveraged in the future.

Our recommendation is to apply this survey every other year.

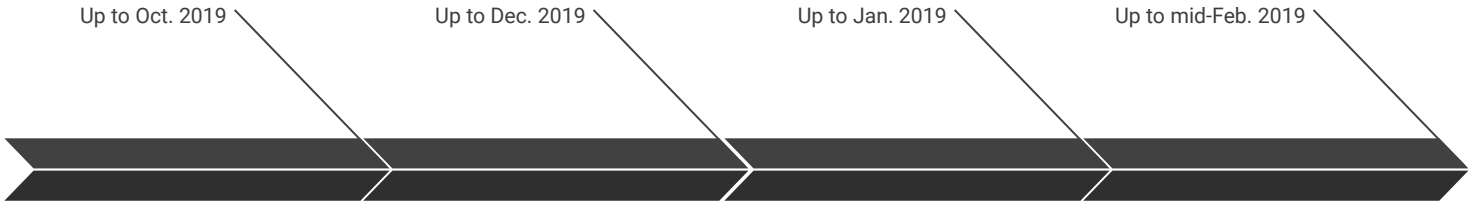
D&I Wiki page

https://cwiki.apache.org/confluence/display/EDI



Data Collection

Data Collection Timeline



Survey Design

First Survey Design available

Community Feedback

Open discussion and addition of ideas, comments, and concerns

Data Collection

Data collection starts

7010 emails sent

Deadline Extension

Booth FOSDEM opportunity

624
Total Responses

8.5%
Response Rate*

** Based on a considered total community size of 7,500*

With token (committers)

Average Time	18 min. 9 sec.
Median Time	10 min. 19 sec.
Partially filled	166
Fully filled	466

Without token (public)

Average Time	18 min. 16 sec.
Median Time	9 min. 22 sec.
Partially filled	352
Fully filled	158

Data Analysis

Scope

Survey responses are analyzed.

- Committers (7010 emails sent: 466 Full; 166 Partial)
- Open survey (158 Full; 352 Partial)

Initial results are shared with ASF as a presentation in a 1-hour session. Slides are shared on dev@diversity.apache.org.

Initial results inform the design of interviews.

ASF Goals: Instrument the Apache contributor funnel to establish a participation baseline for underrepresented groups

ASF Questions:

- What is the state of diversity in ASF?
- Are contributors facing challenges when participating at the ASF?
  - Any association with gender?
  - Any association with English fluency?
- Look for interview participants?

Analysis Dimensions

- Demographics:
  - Age, Gender, English fluency, Background culture
- Socio-economic aspects:
  - Education, Compensation, Time to volunteer
- Experiences in the ASF:
  - Tenure, mentorship, challenges

Method

Demographics Analysis

Population analysis that help draw a first approach to the ASF ecosystem. This section mixes several attributes to understand specific areas, including education,gender, migrants, newcomers and challenges.

Profile Analysis

Identification of key profiles to be part of the in *depth interview process*. These are dependent of significant attributes important for ASF EDI working group such as gender, English speaking confidence, and others.

Survey Results

Average ASF contributor

Demographics

40 Years old

Man

Confident in English

Born/lives in the USA

Socio-economic Aspects

Bachelor's Degree

No compensation

1 or 2 hours for volunteering

Experience in the ASF

5 Years in the community

Didn't have a mentor

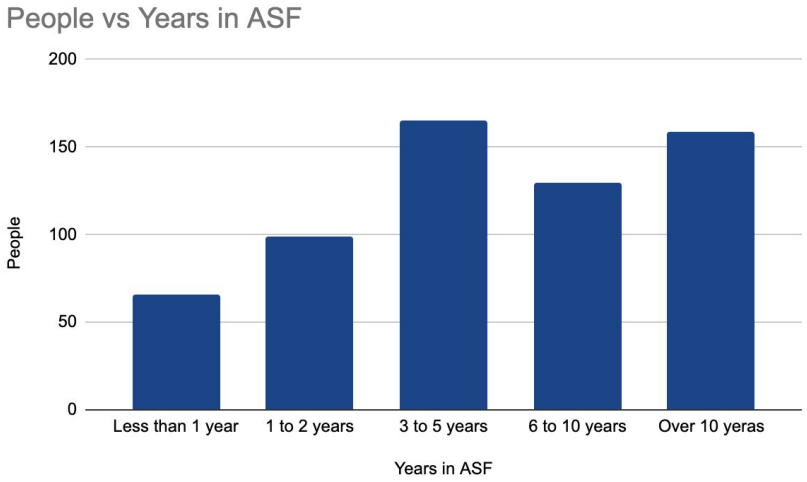
Faced no challenges

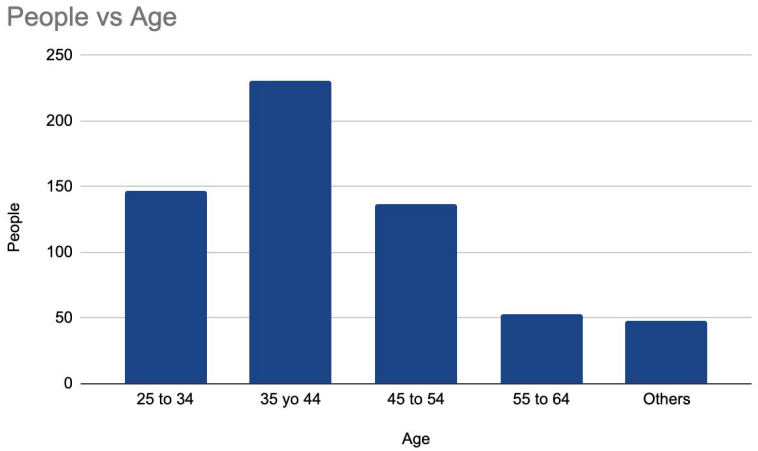
Getting Experienced

People with less than 1 year of experience represents **10%** of the answers

Participants under 24 represents less than a **5 %**

Are there barriers for newcomers or young adults?



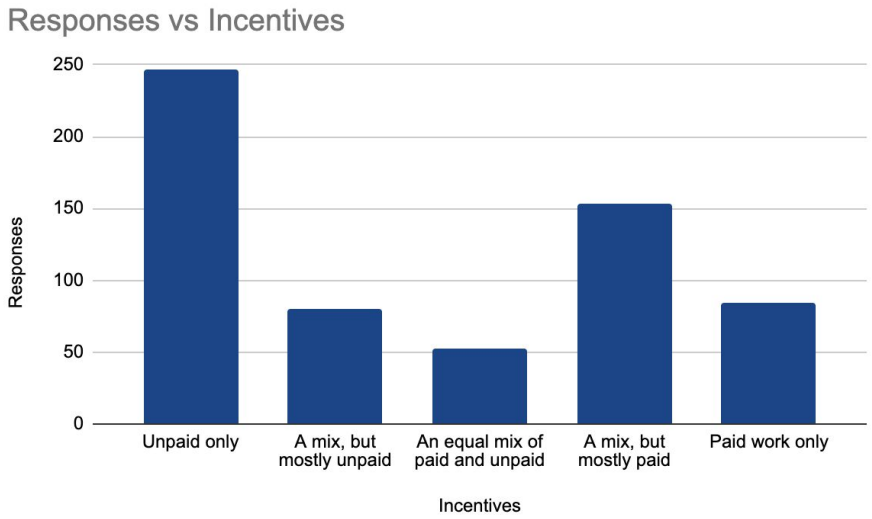


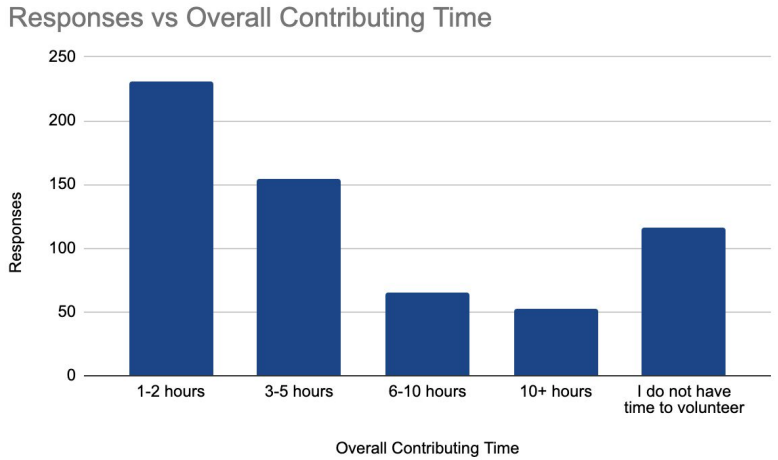
Mostly Volunteers

40% work as volunteers

81% have less than 5 hrs/week for volunteering

Are there ways to be more efficient? To help those unpaid volunteers to contribute?



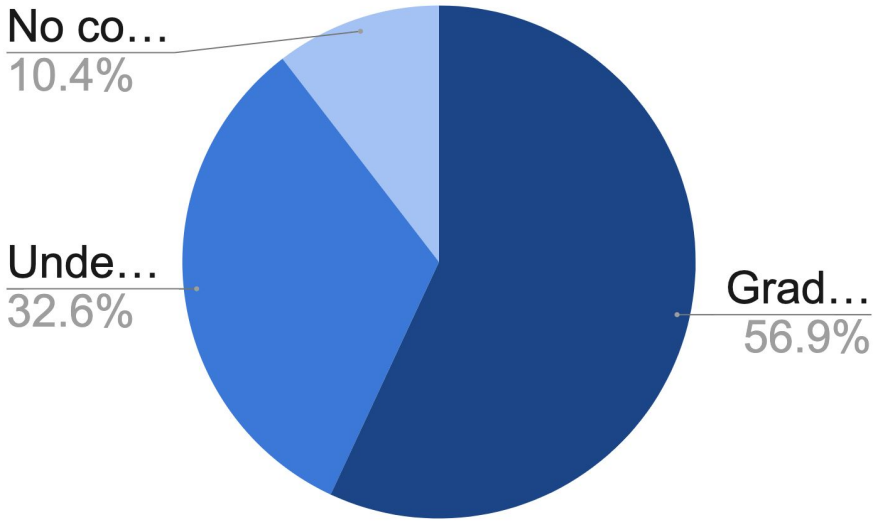


Highly Educated

Around 90% of ASF contributors are college educated or higher

Do these contributors know about OSS at their schools?

Are the “no college” group aware of OSS at the same level?

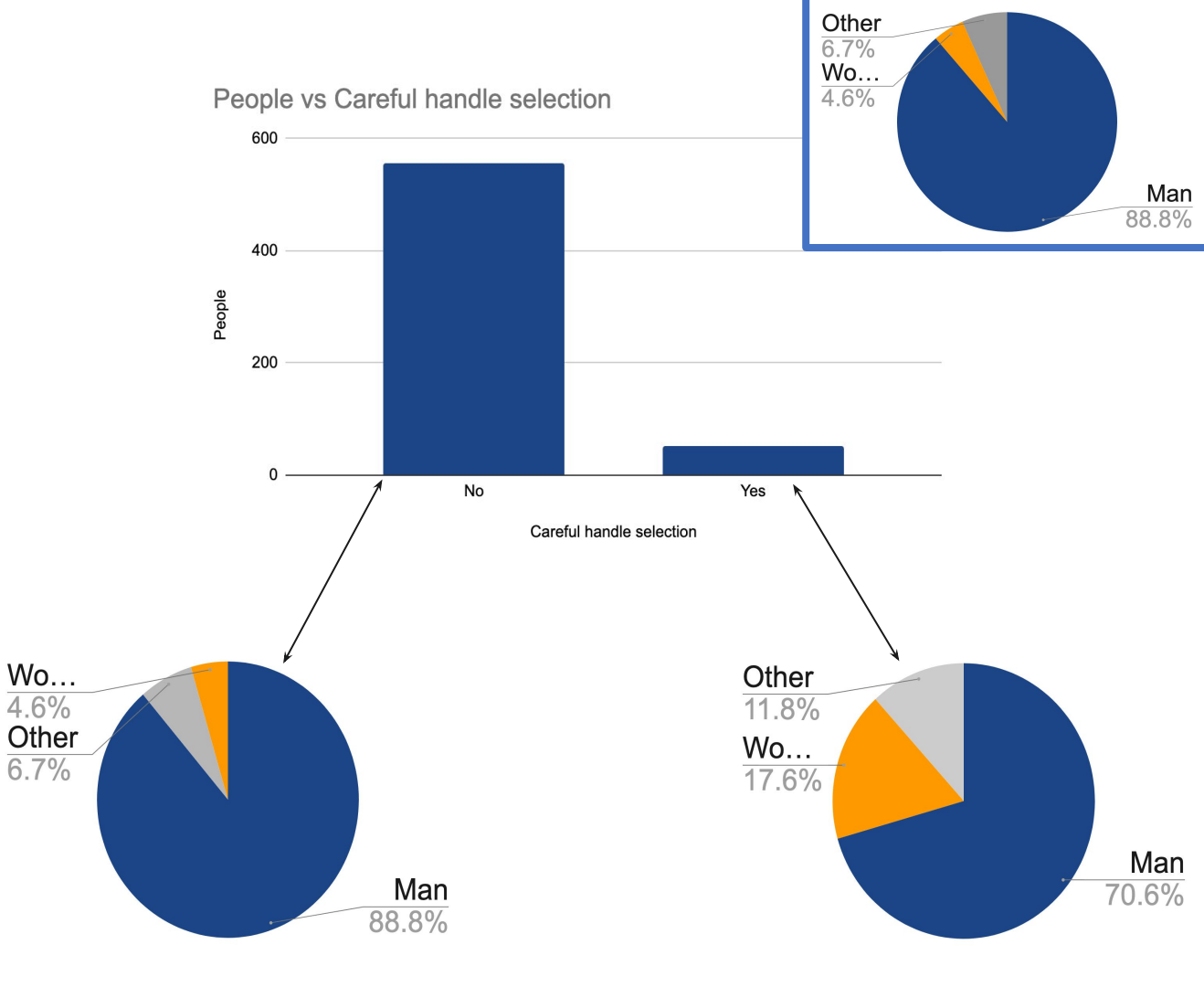


Mostly Men

11% have self-defined as women or with another gender

8.4% have carefully think what gender their handles represent

Are there issues related to non-gender neutral handle? How can we improve the percentage of people coming from the Women and Others group?

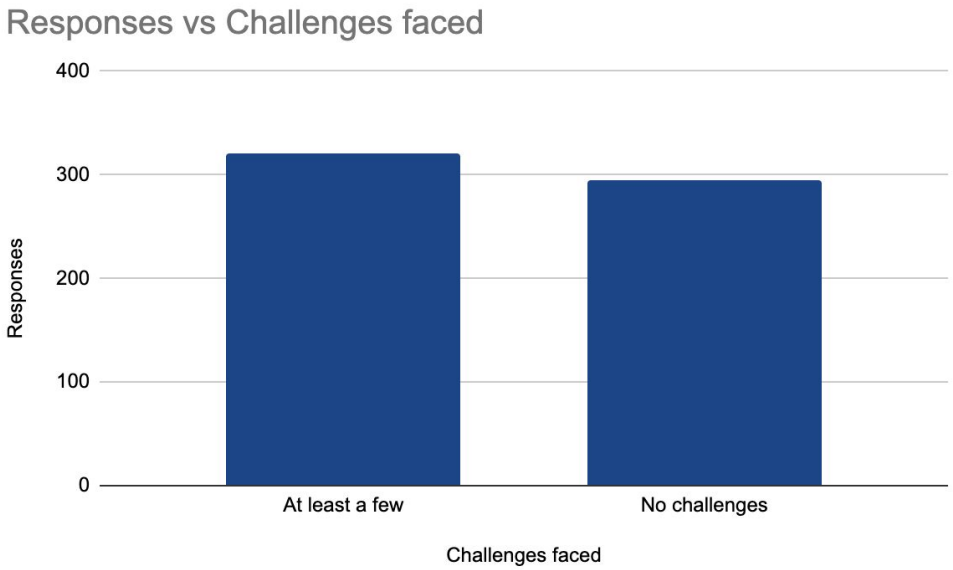


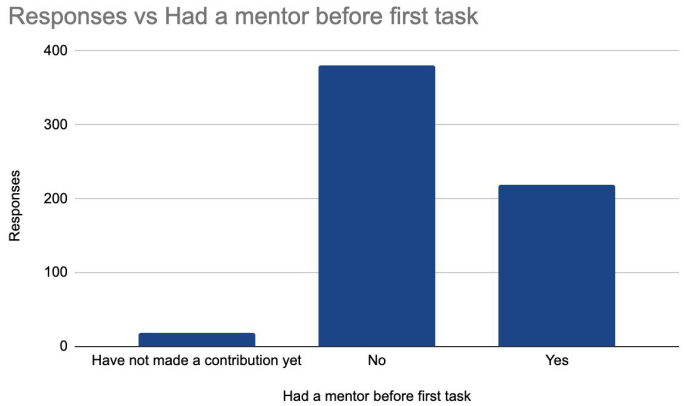
There are Challenges

About 52% of ASF contributors have faced challenges

62% of contributors didn't have a mentor

How can we reduce these challenges to make the contributors path easier?

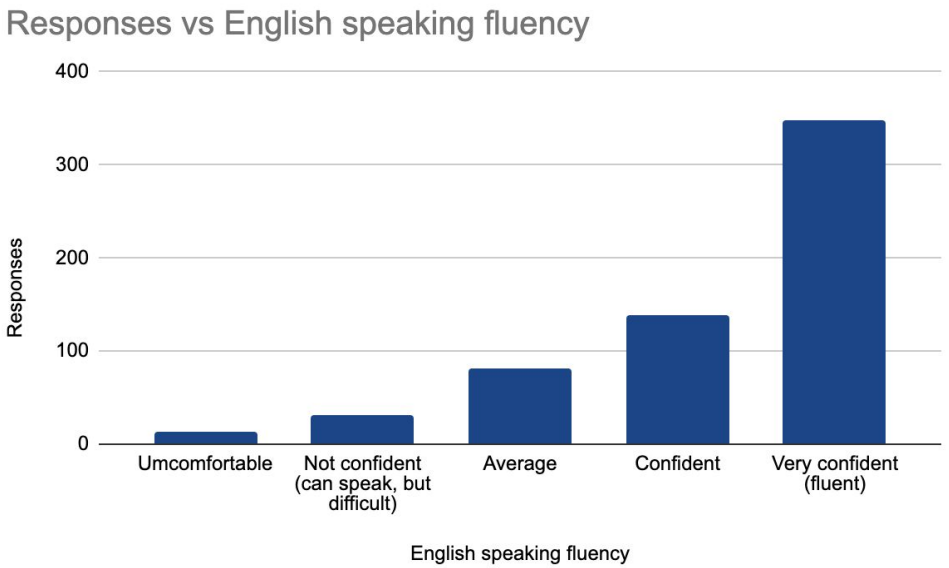




7% not English confident

Around 20% of contributors feel they have an average English level or less

Language seems to be a barrier for communication



Culture	# Responses	% Responses
United States	171	28.84%
Germany	69	11.64%
China	49	8.26%
France	32	5.40%
India	28	4.72%
United Kingdom	25	4.22%
Russian Federation	23	3.88%
Canada	18	3.04%
Spain	14	2.36%
Japan	11	1.85%

Top 10 countries by Culture

Residence	# Responses	% Responses
United States	230	38.27%
Germany	76	12.65%
China	37	6.16%
United Kingdom	28	4.66%
France	27	4.49%
Canada	18	3.00%
Netherlands	15	2.50%
Russian Federation	14	2.33%
Switzerland	12	2.00%
Japan	11	1.83%

Top 10 countries by Residence

Hypotheses

Hypotheses

Education impacts compensation

Those in minority face challenges

Men who face challenges have different demographics

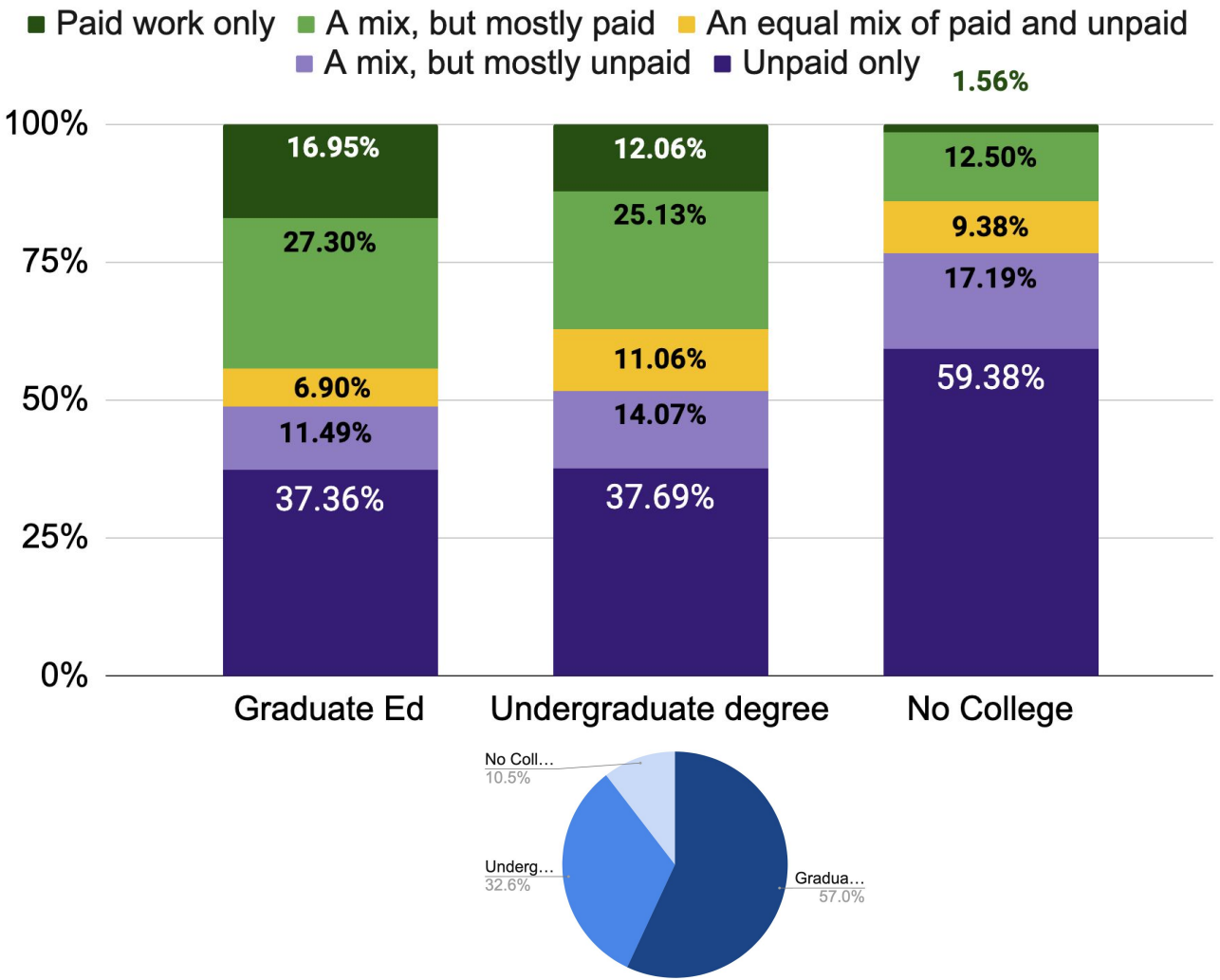
The culture where you grew up makes a difference

Education impacts compensation

Education Compensation

Increase of Unpaid population for those with no college education

Willing to work as volunteer to gain experience?



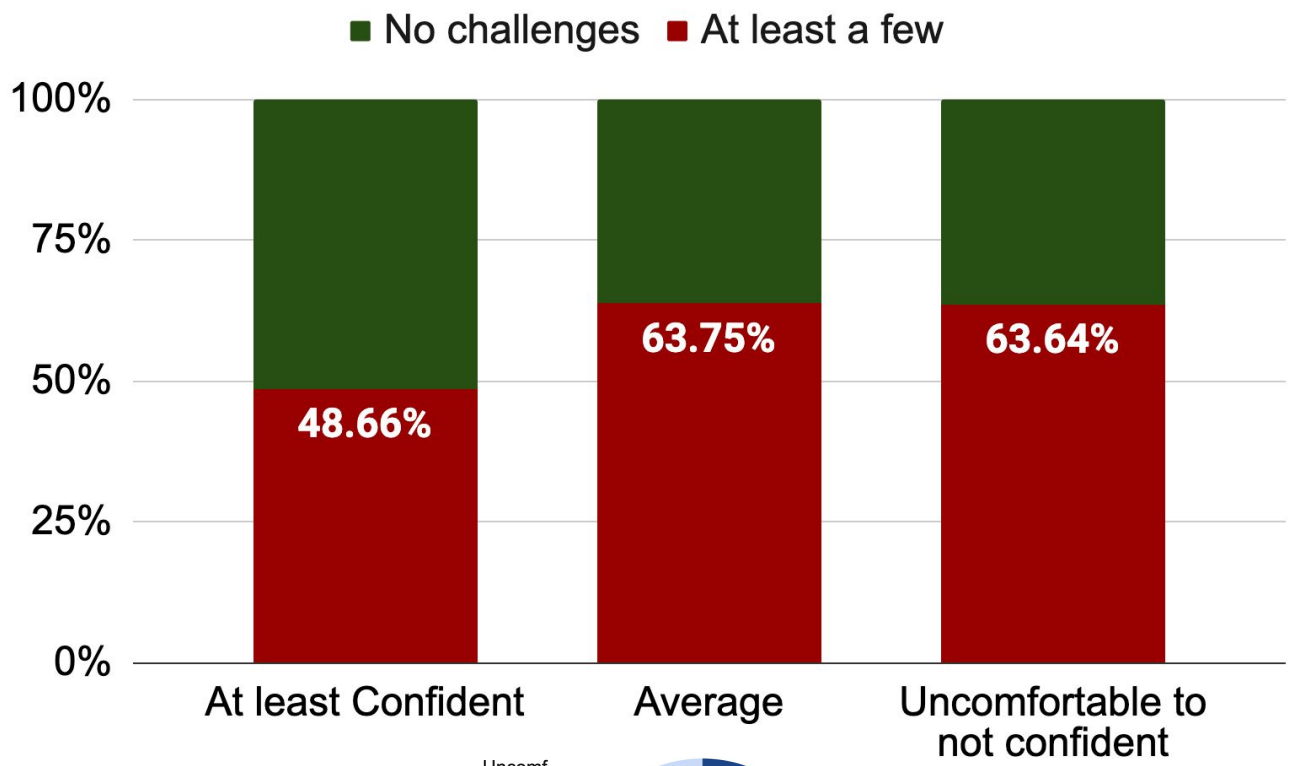
Those in minority face challenges

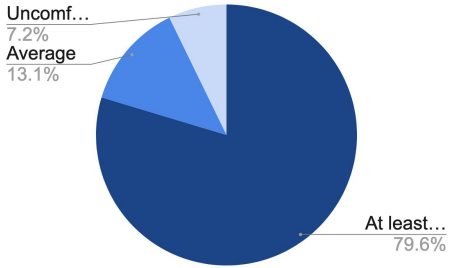
- English fluency
- Gender
- Having a mentor

Language

Those with an average English level or less faced more challenges

How can we lower the language barrier?

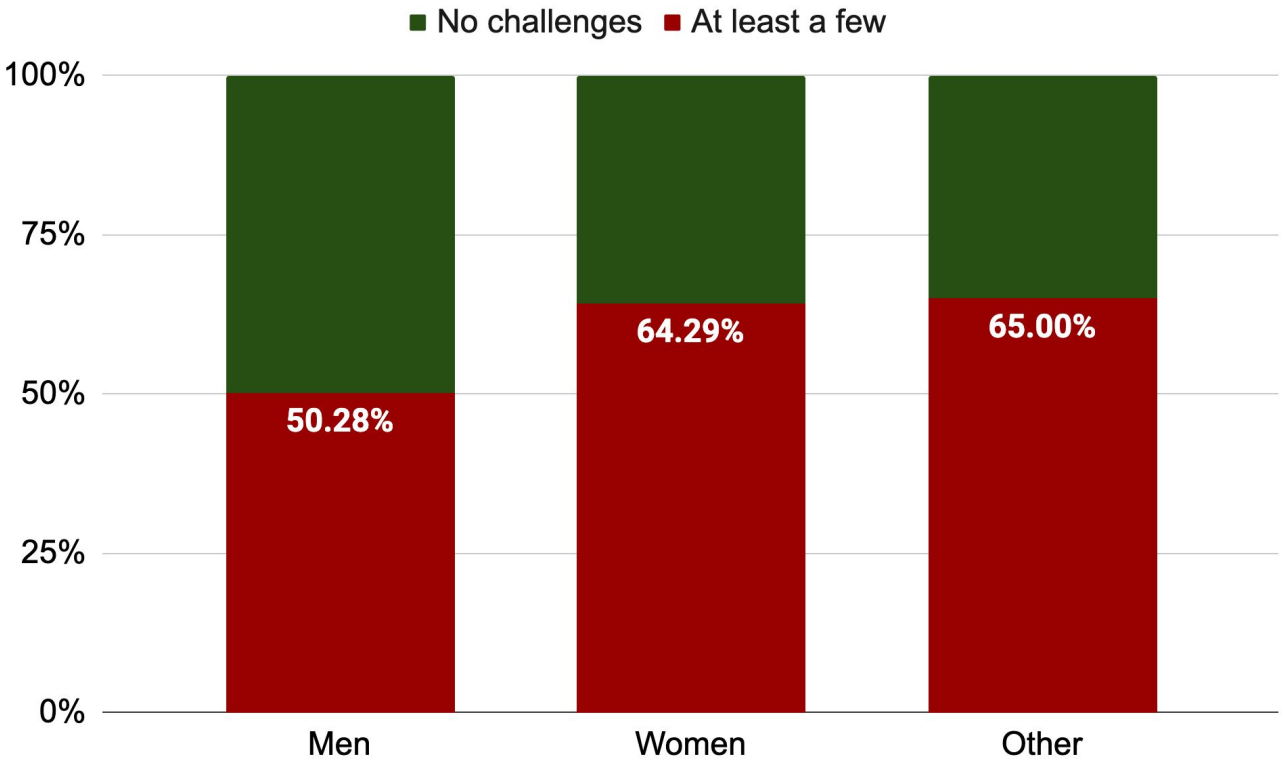


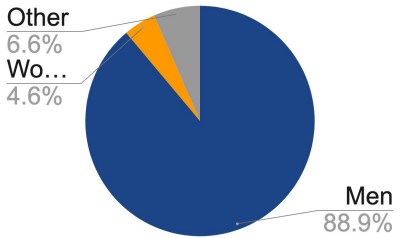


Gender

Those that self-assigned as Women or other groups (not men) faced more challenges

How can we reduce this barrier?

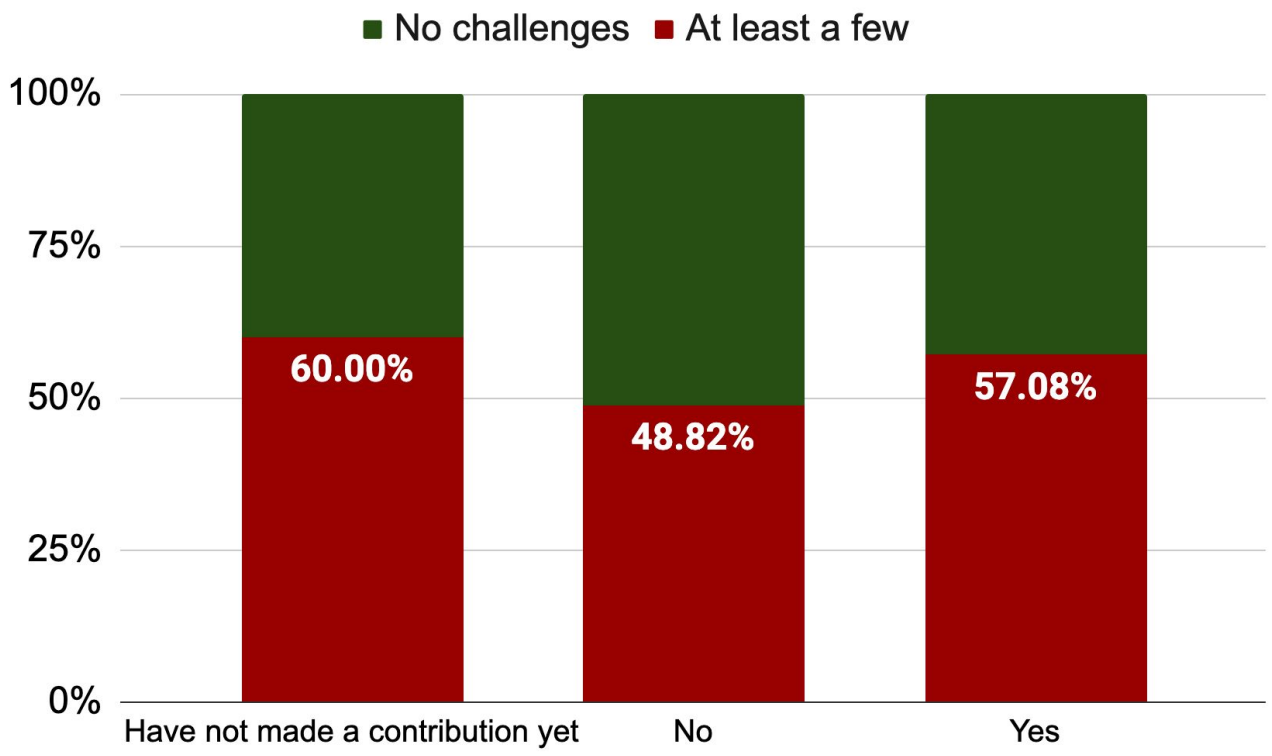


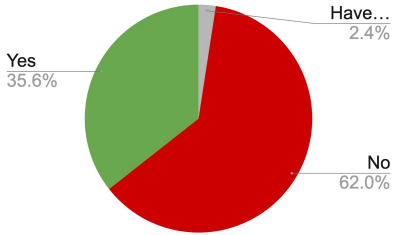


Having a Mentor

Those with a mentor faced more challenges

How can we reduce this barrier?





Men who face challenges have different demographics

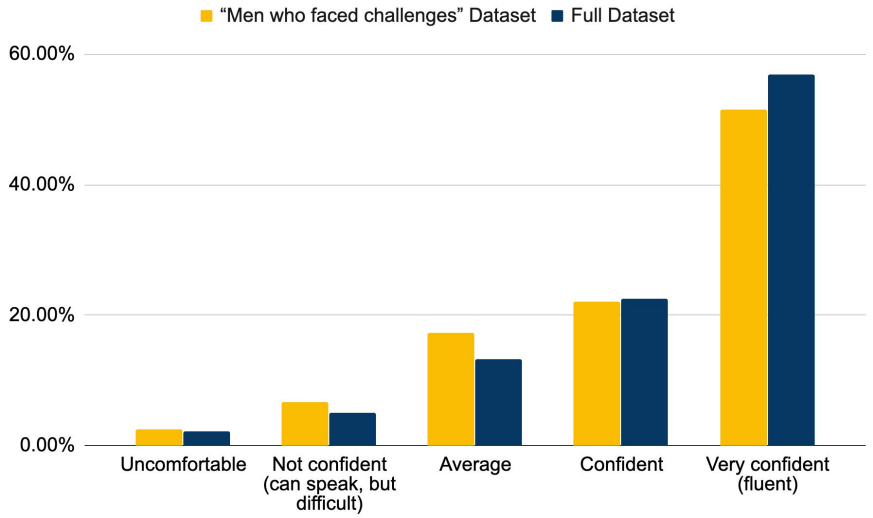
* Population of 272 people

No main differences

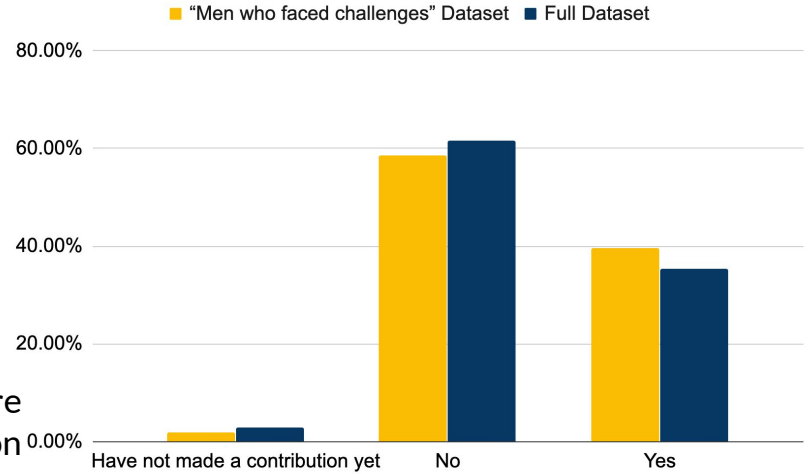
Men with certain challenges do not show a different distribution if compared to the full dataset

Less confident in English

Slightly higher distribution of those that had a mentor before their first contribution



Men that had a mentor before their first contribution



Those who move from the country where they grew up have...

... different demographics

... challenges to contributing

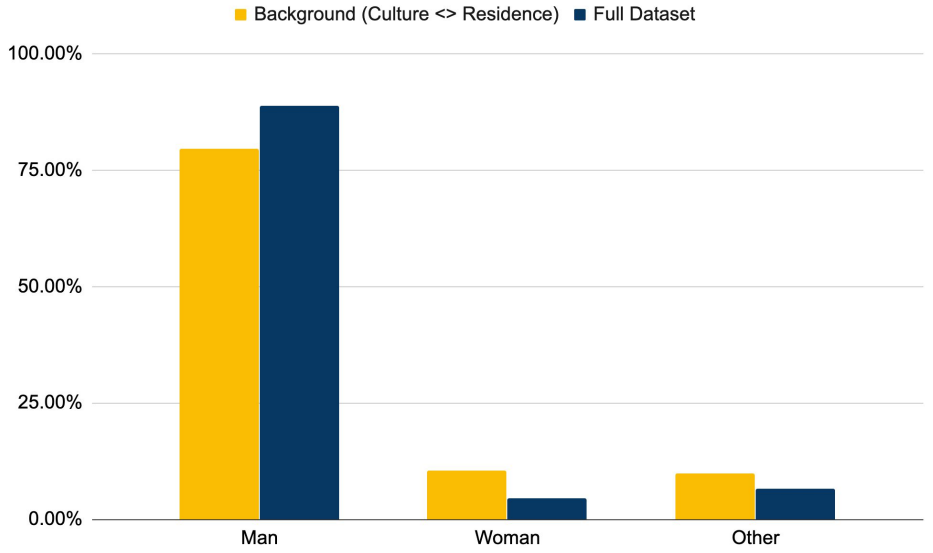
* Population of 134 people

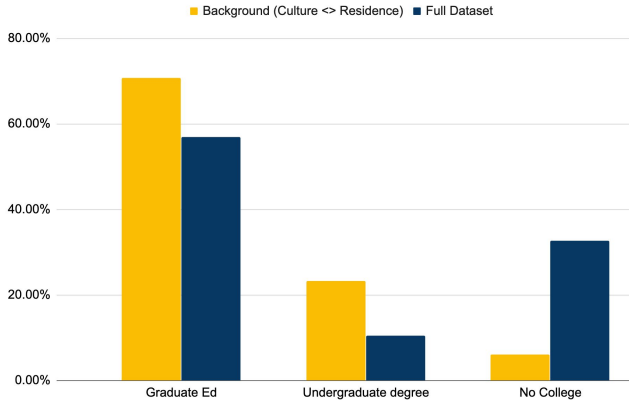
Cultural Background

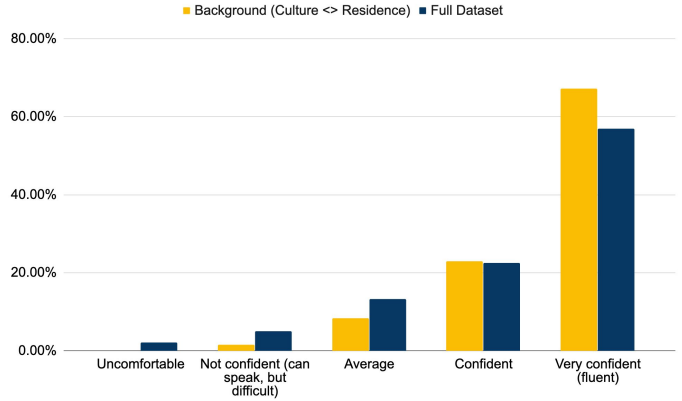
There are more women with a different cultural background and current residence

In general, those that moved to another country have a higher education and feel more English confident

And more are compensated for participating







Recap

1. Education impacts compensation: yes ($p<0.05$)*
2. Those in minority face challenges
   a. English fluency: Yes ($p<0.05$)
   b. Gender: Yes ($p < 0.05$)
   c. Having a mentor: No ($p = 0.05$)
3. Men who face challenges have different demographics: No
4. Those who move country where they grew up...
   a. ... different demographics: No
   b. ... challenges to contributing: No ($p>0.05$)

*Chi^2 test for significance

Question: what barriers newcomers face?

* Population of 18 people

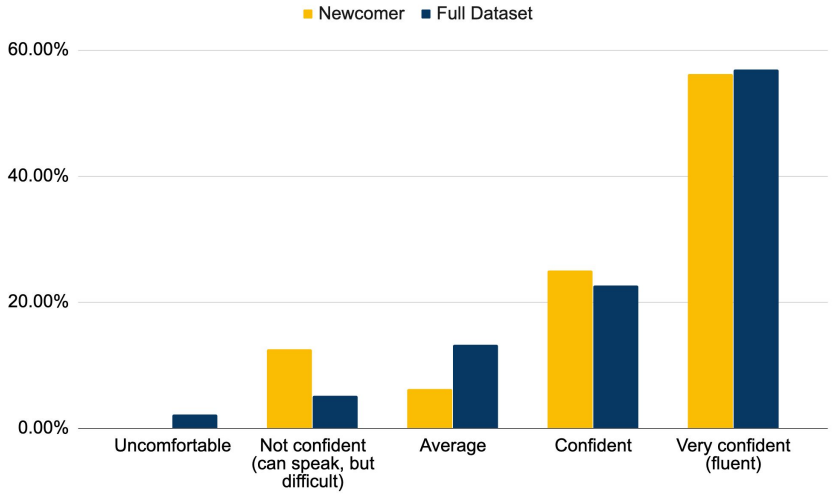
Newcomers have barriers

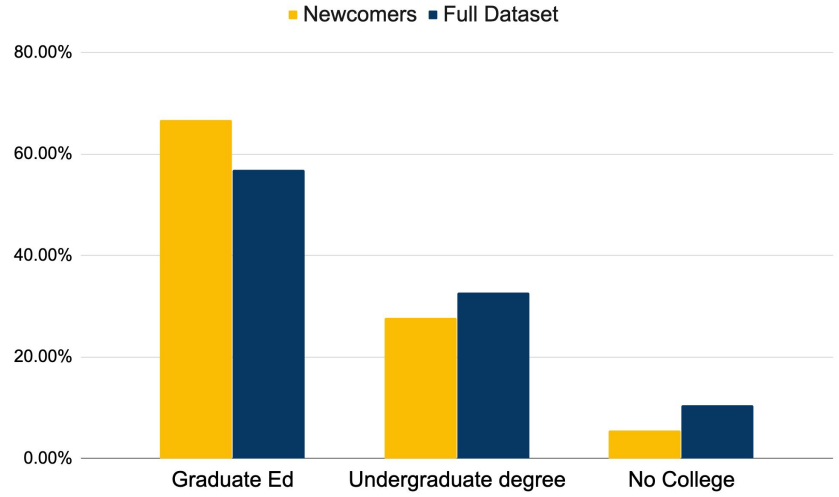
A higher percentage of them face challenges

There are more women in percentage. Is gender a factor for this?

The language might be another barrier

And most of them are volunteering with no compensation





Next Steps

Next steps

Interviews to validate hypothesis

- Interviewee identification distributed across studied dimensions
  - Minorities
  - Men with challenges
  - Average profile
- Additional interviews with groups of interest
  - Interns
  - Those who left

Quantitative analysis

- Selection of projects to analyze
- Dashboard generation with gender info

Questions?

Recap

1. Education impacts compensation: yes ($p<0.05$)*
2. Those in minority face challenges
   a. English fluency: Yes ($p<0.05$)
   b. Gender: Yes ($p < 0.05$)
   c. Having a mentor: No ($p = 0.05$)
3. Men who face challenges have different demographics: No
4. Those who move country where they grew up...
   a. ... different demographics: No
   b. ... challenges to contributing: No ($p>0.05$)

*Chi^2 test for significance

Appendix

Survey Scope

Initial Goals and Scope

Include relevant CHAOSS D&I WG metrics and survey questions.

Feedback collected in a 1-hour session with ASF members. One week before, draft survey submitted to dev@diversity.apache.org.

ASF pilots the survey with ComDev and incubator for one week.

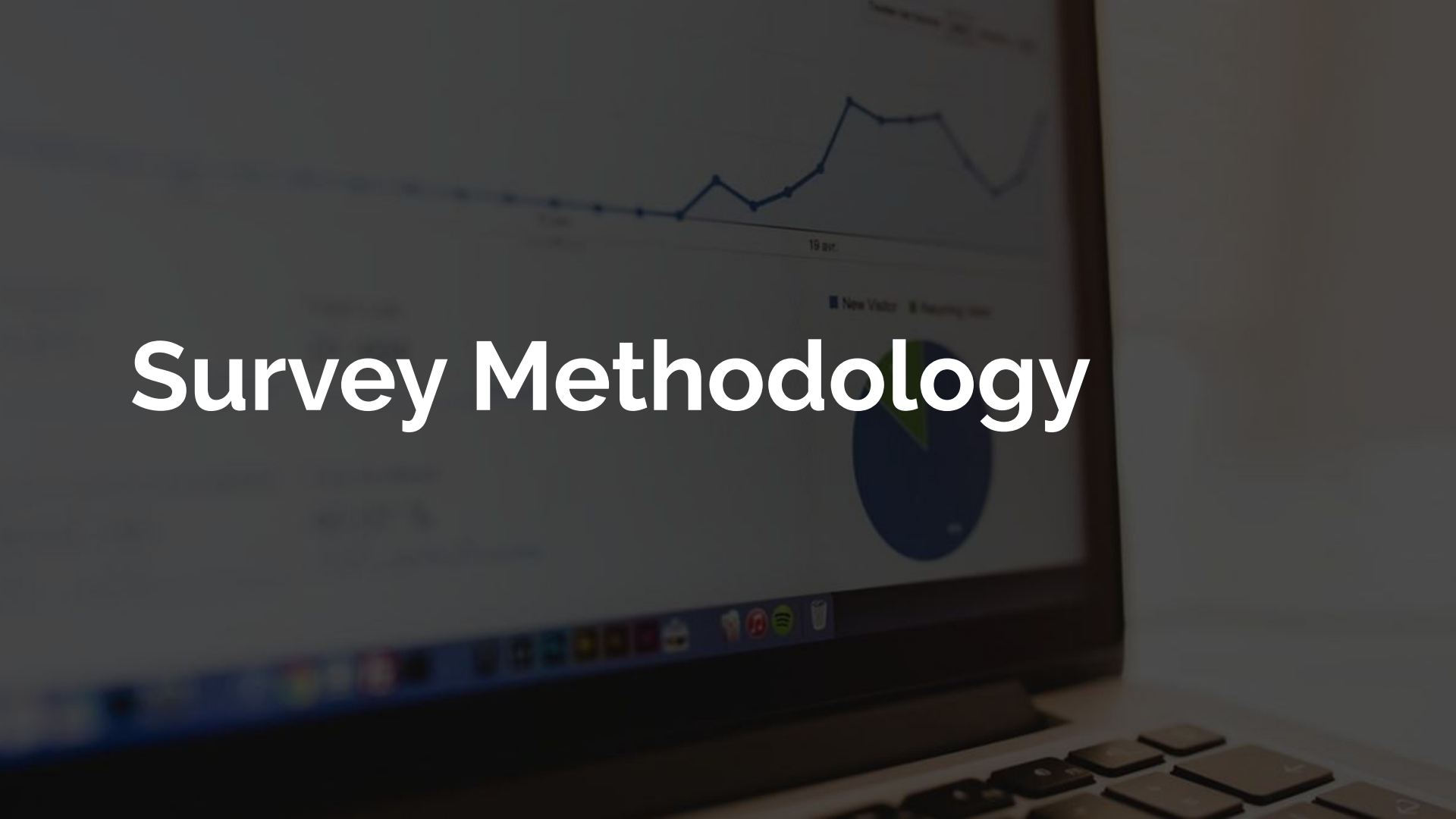
Final survey is provided to ASF, ready to begin data collection.

Any participant in the community can fill the survey.

Committers were sent tokenized invitation via email so we could compute response rates. This is anonymous, and this is done through a randomly generated token by the tool. We also sent (non-tokenized) links via social media

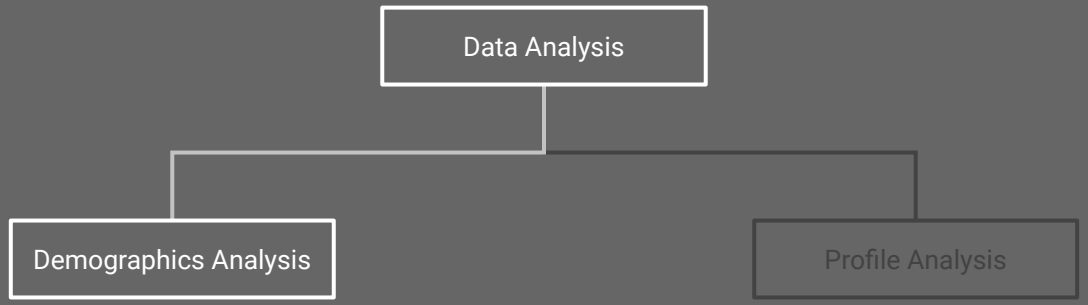
The goal is to have as many people as possible answering the survey and reach a 10% of the total number of committers.

The survey extends the 2016 survey demographics and Contributor Role section and adds other new categories. These focus on understanding the challenges of contributing to the ASF.



Survey Methodology

Method



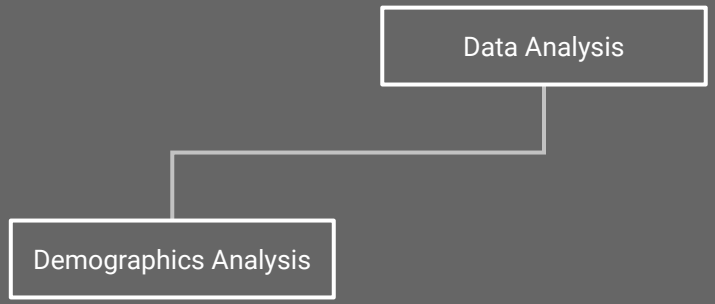
Demographics Analysis

Population analysis that help draw a first approach to the ASF ecosystem. This section mix several attributes to understand specific areas. This focuses on education, challenges, migrants and newcomers.

Profile Analysis

Identification of key profiles to be part of the in depth interview process. These are dependent of significant attributes important for ASF EDI working group such as gender, English speaking confidence, and others.

Method



Demographics Analysis

Population analysis that help draw a first approach to the ASF ecosystem.

This section mix several attributes to understand specific areas. This focuses on education, challenges, migrants and newcomers.

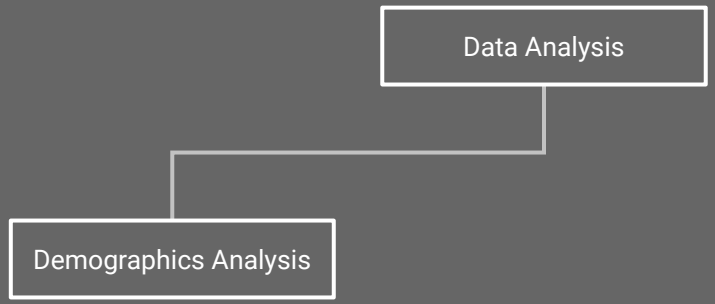
Details

We **aggregated** both the closed (158) and open (466) survey data

We only analyzed **complete** survey data

We **combine** data when <5% and when it makes logical sense

Method



Attributes

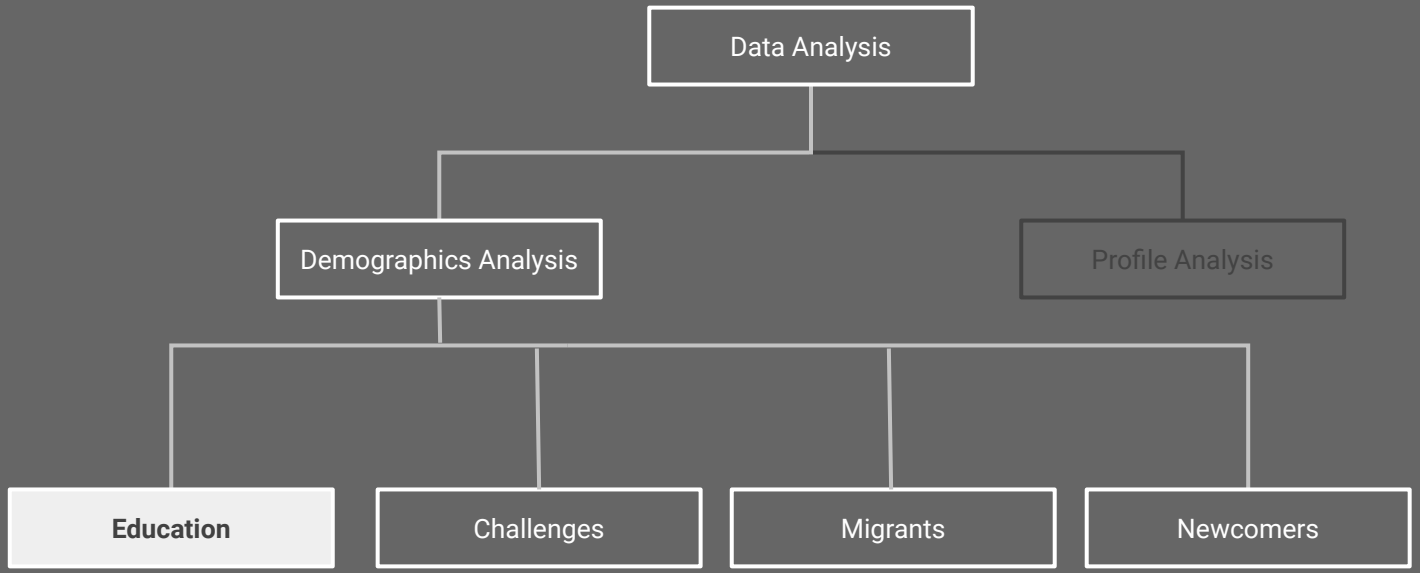
Years in ASF
Age
Overall Contributing Time
Education
Gender
Careful Handle
Have Mentor
Faced Challenges
English Speaking Fluency
Background Culture
Current Residence

Paired Attributes

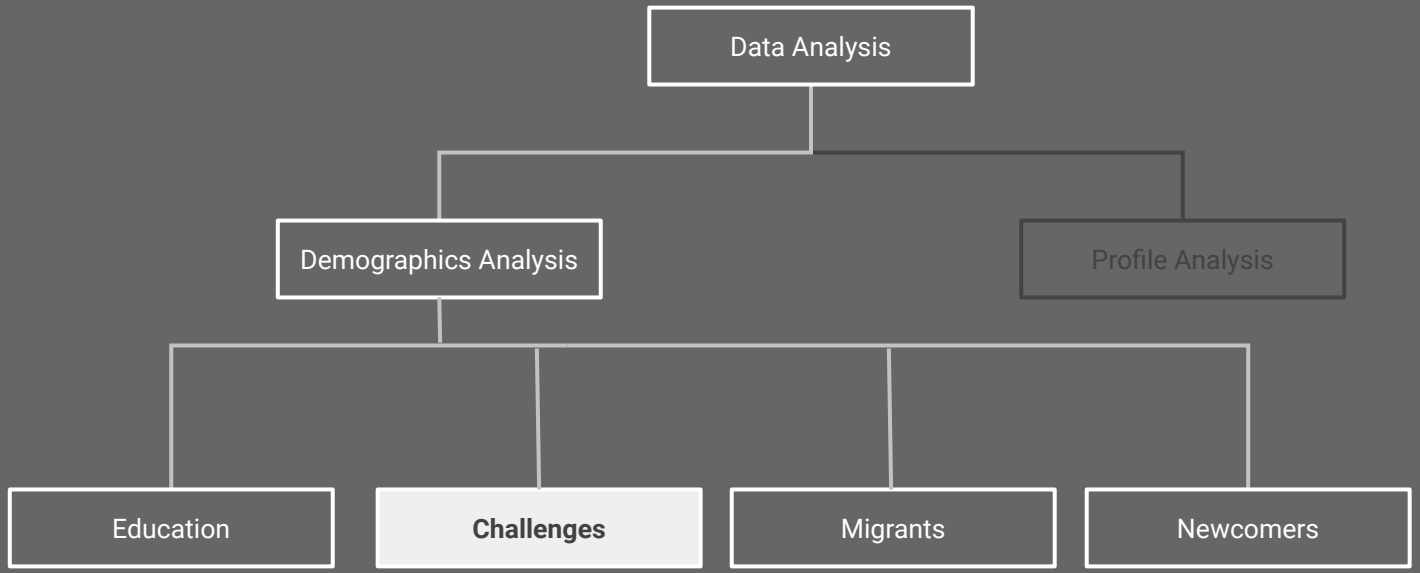
Education
Faced Challenges*
Cultural Background
Newcomers
English Fluency
Gender / Careful Handle
Have Mentor
Paid / Volunteering work

*With detail in men who faced challenges subset

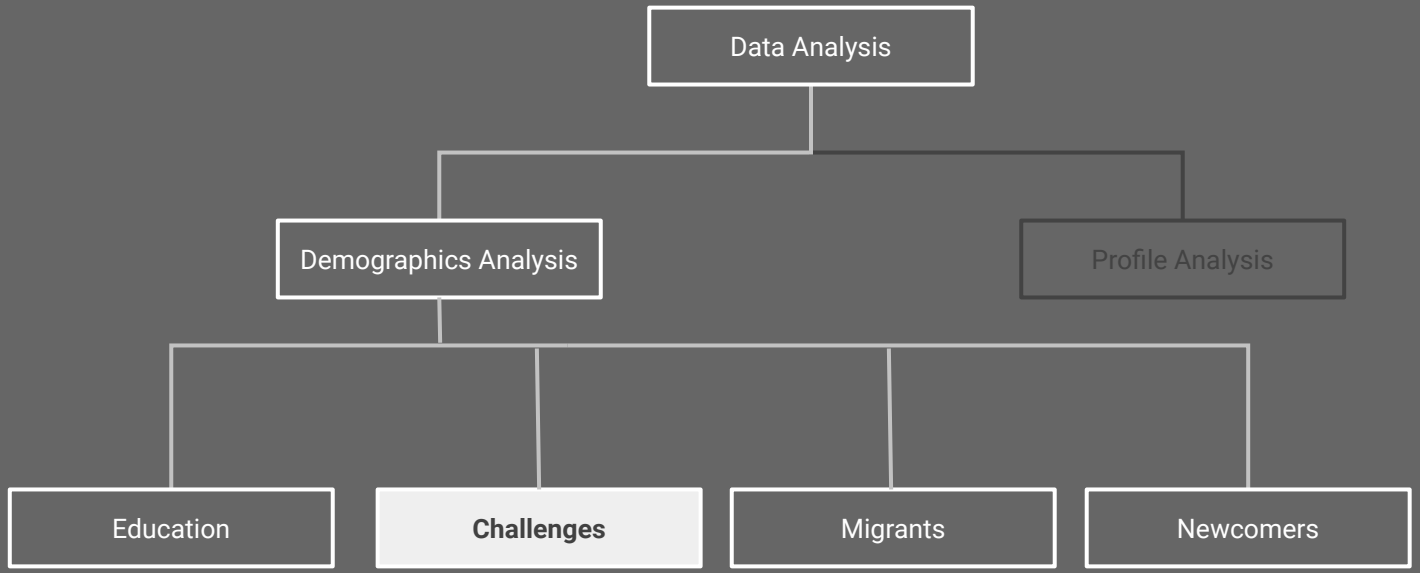
Method


Data Analysis
Demographics Analysis
Profile Analysis
Education
Challenges
Migrants
Newcomers

Method

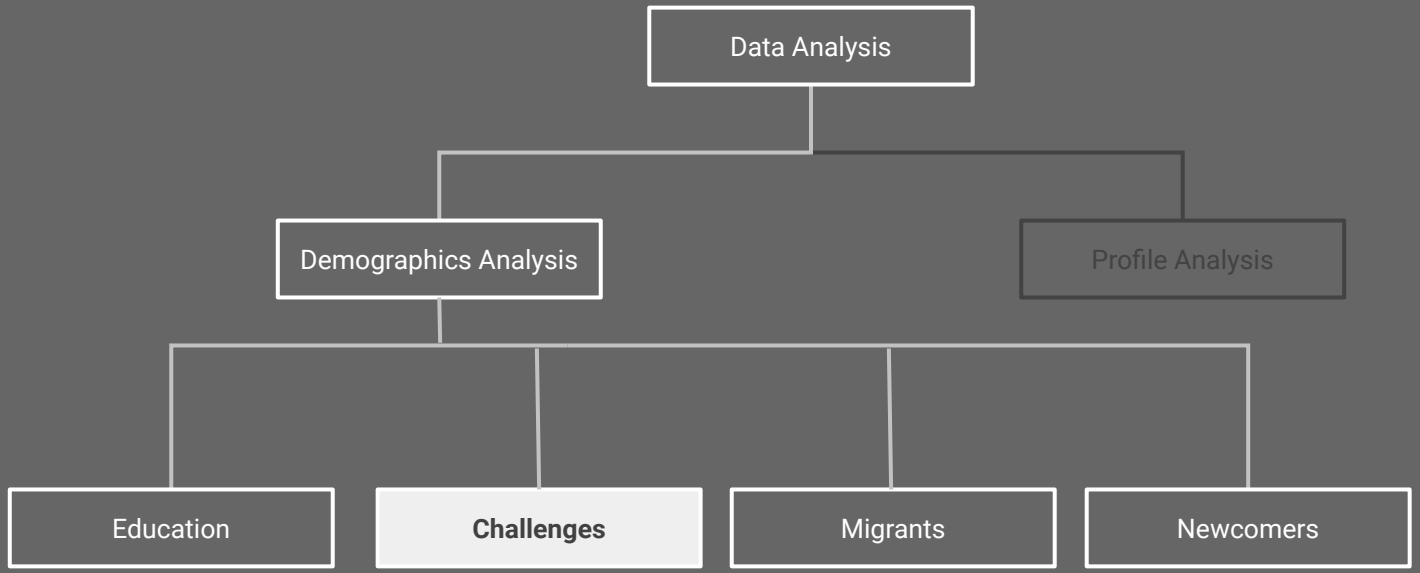

Data Analysis
Demographics Analysis
Profile Analysis
Education
Challenges
Migrants
Newcomers

Method



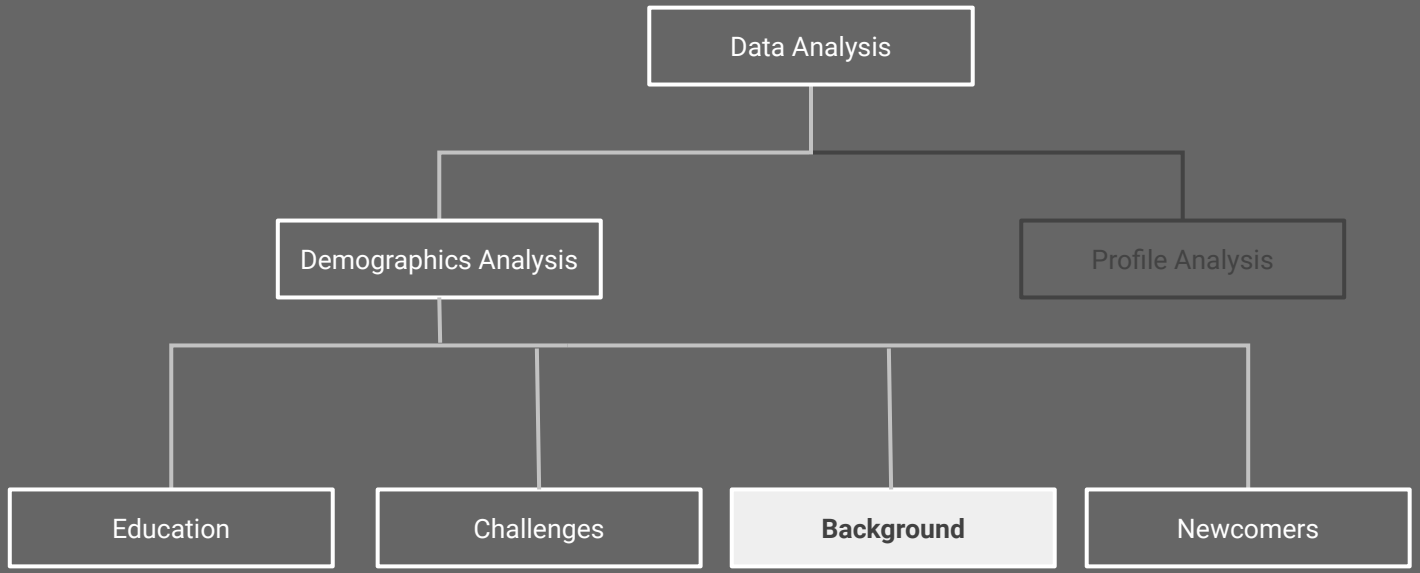
Focus on people that carefully selected their handle
Population of 51 people

Method



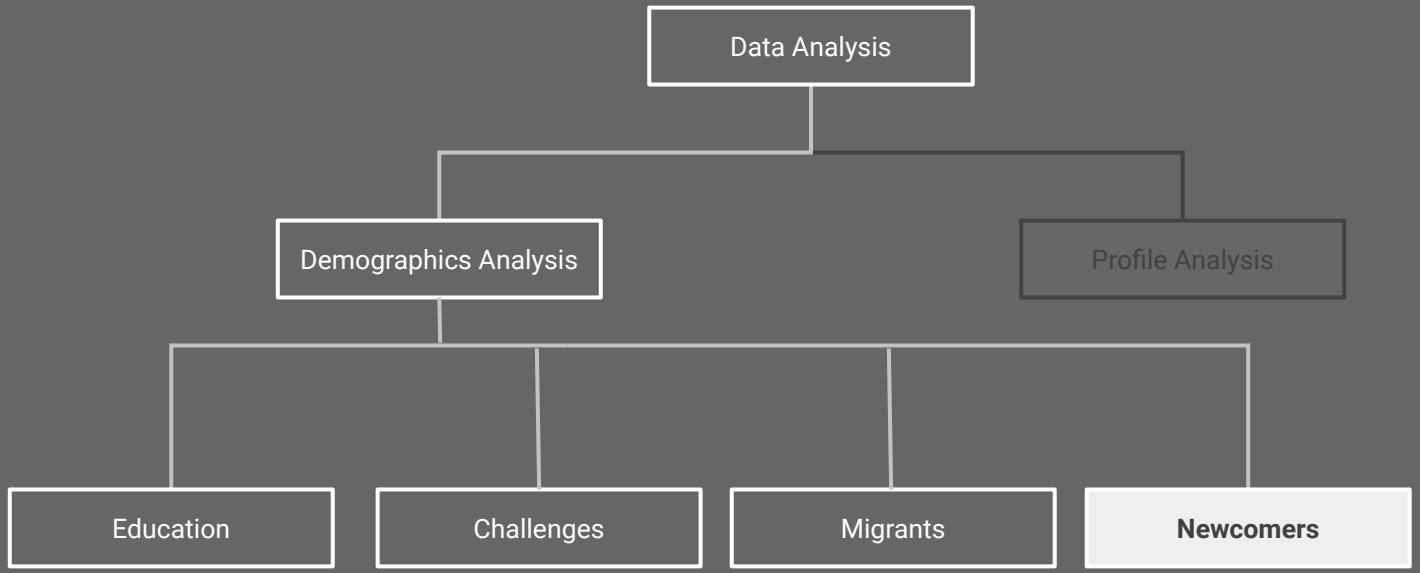
Focus on Men who faced challenges dataset
Population of 272 people

Method



Focus on people whose residence is different from where they currently live
Population of 134 people

Method



Focus on newcomers (those without a contribution yet)
Population of 18 people

Survey design based on previous experiences:

- 2016 ASF Committer Diversity Survey
- To be completed

There were two surveys to keep track of the number of committers (anonymous way) that answered.
Open to anyone in the community.

Data collection took place in a two months period (Christmas in between) and did some advertisement using:

- Committers mailing list
- Social media
-

Analysis

- Data downloaded on: Mar 21
- Data analyzed: 466 + 158
- 7010 emails sent

Challenges

- Motivate committers to fill the survey
- Motivate community participants to fill the survey
- Most of the answers (>50%) came during the first call (committers mailing list)
- Difference between completed surveys for those with token or without token

LimeSurvey Challenges

- Two surveys are provided, data is merged

FOSDEM Booth Challenges

- Only few people were interested in the survey
- Some were willing to fill the survey at the booth (this may lead to a lack of privacy), we asked them not to.
- Lack of resources, perhaps having a QR or some flyer may have helped.
- Committers were aware of this
- Other people were not aware of this

Profile Analysis for interviewee selection

Method

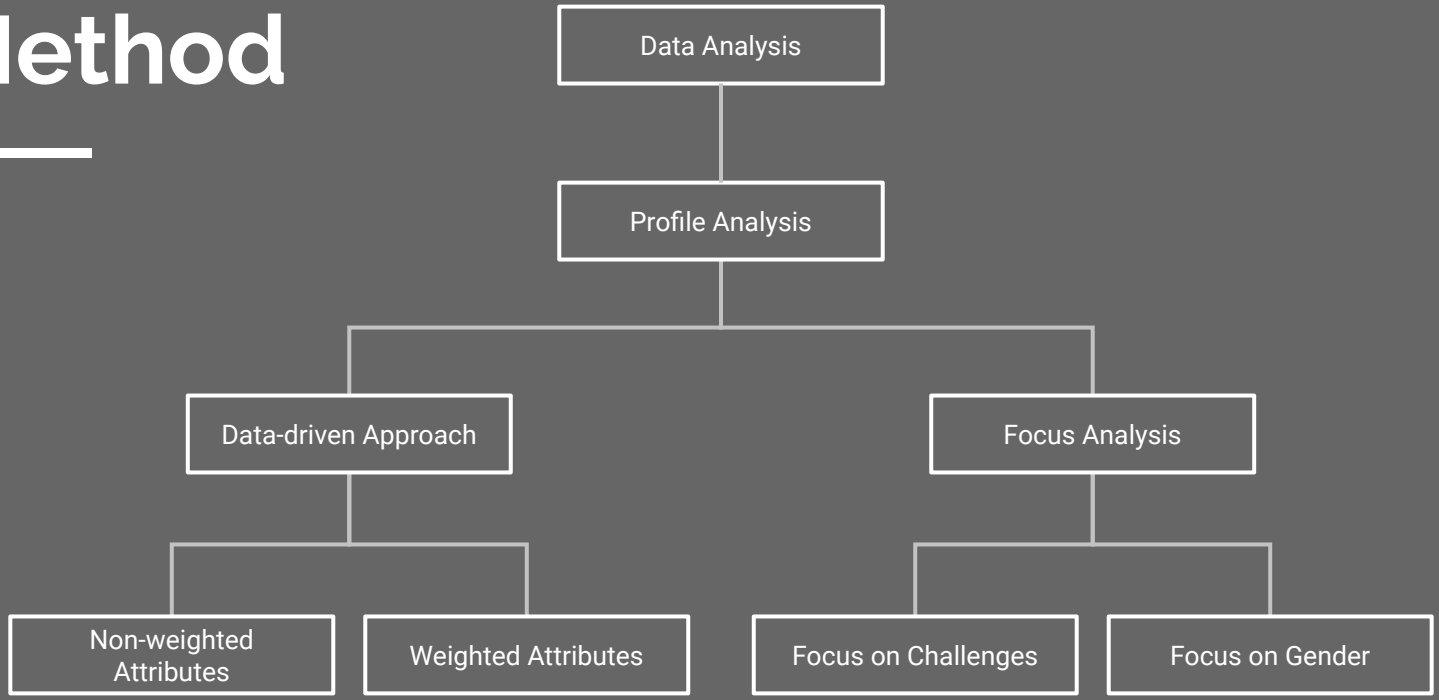
Demographics Analysis

Population analysis that help draw a first approach to the ASF ecosystem. This section mix several attributes to understand specific areas. This focuses on education, challenges, migrants and newcomers.

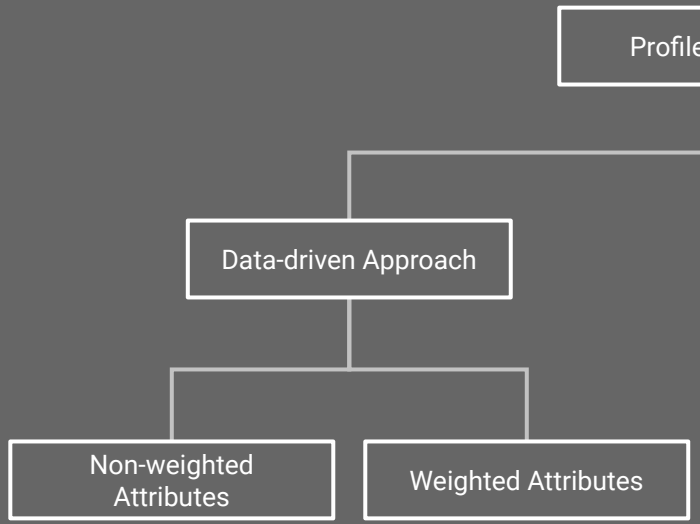
Profile Analysis

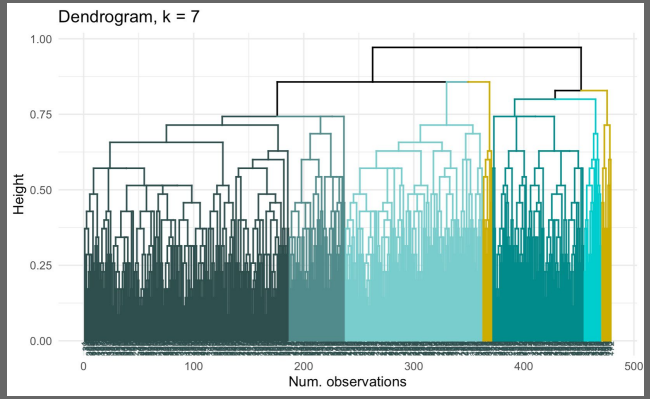
Identification of key profiles to be part of the in depth interview process. These are dependent of significant attributes important for ASF EDI working group such as gender, English speaking confidence, and others.

Method


Data Analysis
Profile Analysis
Data-driven Approach
Focus Analysis
Non-weighted Attributes
Weighted Attributes
Focus on Challenges
Focus on Gender

Method





Data analysis focused on creating clusters to identify majority profiles and outliers (all dataset included)

Two approaches* for clustering:

- Partitioning around Medoids
- Agglomerative Hierarchical Clustering adopted
  - Retain the structure even when we increase the number of clusters.
  - Starts by putting each participant in it's own cluster, examines distance and pairs clusters together.
  - Visual dendrogram (visually understand how cluster are divided).

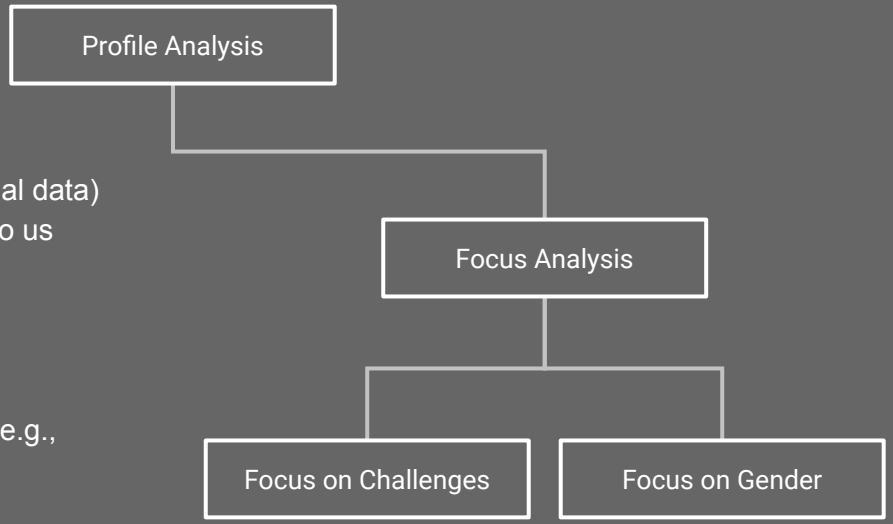
Silhouette analysis** to determine the number of clusters.

Weighted attributes to give focus to attributes of Diversity interest: challenges faced, gender, culture, english proficiency

- Agglomerative Hierarchical Clustering adopted

* https://www.stat.berkeley.edu/~s133/Cluster2a.html
** https://medium.com/codesmart/r-series-k-means-clustering-silhouette-794774b46586

Method



The focused analysis using attributes of interest (partial data)
This allows direct focus on the questions that matter to us according to the general ASF goals:

- Challenges faced
- Gender identity

Analysis between multiple pairs of factors of interest (e.g., challenges faced and gender)

Two partitioned dataset:

- “challenged” dataset: responses from individuals who at least faced some challenges (<> no challenges)
- “gender” dataset: <> men

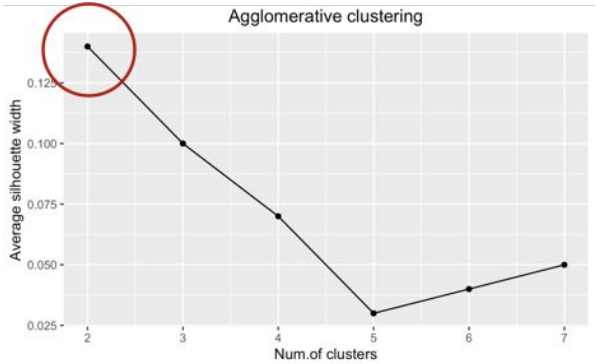
Clustering for majority profiles

- Non-Weighted
- Weighted (challenges faced, gender, culture, english proficiency)

Method Summary

Non-weighted Attributes	Weighted Attributes	Focus on Challenges	Focus on Gender
Clustering analysis on dataset with all factors given equal weights	This method provides weight to certain attributes. This analysis gives weight to: - Challenges faced - Gender - Culture - English proficiency	Partitioned data to focus on those who faced challenges - non weighted - weighted	Partitioned data to focus on “non men” participants - non weighted - weighted

Non-Weighted	Cluster 1	Cluster 2
Years at ASF	3 to 5	3 to 5
Compensation	Unpaid only	Unpaid only
Volunteering	1 to 2 hours	1 to 2 hours
Challenges	No	Few
Residence/Culture	US /US	Germany/ Germany
English Speaking	Very Confident	Confident
Age	35 to 44	35 to 44
Gender	Man	Man
Education	Undergrad	Masters
Mentor	No	No
Careful handle	No	No



Weighted	Cluster 1	Cluster 2
Years at ASF	3 to 5	3 to 5
Compensation	Unpaid only	Unpaid only
Volunteering	1 to 2 hours	1 to 2 hours
Challenges	No	Few
Residence/Culture	US /US	Germany/ Germany
English Speaking	Very Confident	Confident
Age	35 to 44	35 to 44
Gender	Man	Man
Education	Undergrad	Masters
Mentor	No	No

Weighted and non-weighted give exactly the same majority profiles

Adopt non-weighted dataset

Results: Outlier Profiles Non Weighted

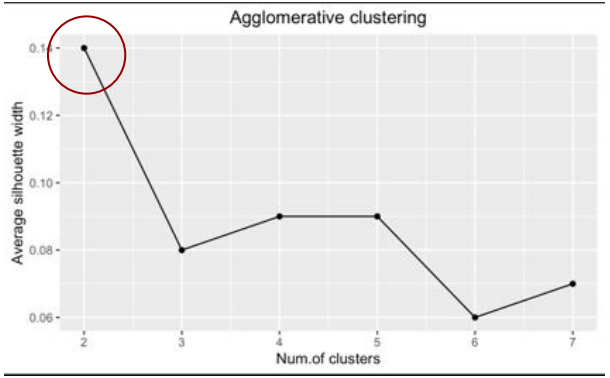
Non-Weighted	Outlier 1	Outlier 2	Outlier 3
Years at ASF	3 to 5	3 to 5	1 to 2
Compensation	Unpaid only	Equal Mix, paid and unpaid	A mix, but mostly paid
Volunteering	1 to 2 hours	1 to 2 hours	1 to 2 hours
Challenges	Few	No	Few
Residence/Culture	India /India	Canada/Canada	US/US
English Speaking	Very Confident	Confident	Very confident
Age	35 to 44	35 to 44	24 or younger
Gender	Man	Man	Gender Var/non-conform/
Careful Handle	No	No	Yes
Education	Undergrad	Technical training	Undergrad
Mentor	No	Yes	No

Challenged Partition	Cluster 1	Cluster 2
Years at ASF	3 to 5	Over 10 years
Compensation	Unpaid only	Mix but mostly paid
Volunteering	1 to 2 hours	3 to 5 hours
Challenges	Few	Few
Residence/Culture	US /US	US/ US
English Speaking	Very Confident	Very confident
Age	35 to 44	25 to 34
Gender	Man	Man
Education	Masters	Masters
Mentor	No	yes

Challenged dataset give profiles with similar attributes as Majority profiles

- weighted
- non weighted

Dropped challenged partitioned dataset



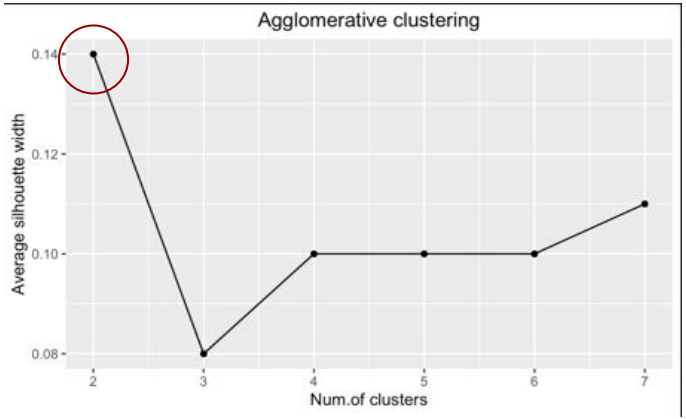
Results: Gender dataset; Majority Profiles

Gender Partition	Cluster 1	Cluster 2
Years at ASF	Over 10 year	Over 10 years
Compensation	A mix but mostly paid	Unpaid only
Volunteering	1 to 2 hours	3 to 5 hours
Challenges	Few	No
Residence/Culture	US/US	US/US
English Speaking	very confident	Very confident
Age	35 to 44	25 to 34
Gender	Women	Women
Education	Undergrad/Masters	Masters
Mentor	No	No

Gender dataset give profiles with similar attributes as Majority profiles

- weighted
- non weighted

Dropped gender partitioned dataset



Outliers: Partitioning Data

Factors: Challenges, Gender, Culture, English Speak

- Challenges <> “No” and “Few” => 51 participants (19 provided emails)
  - Further cluster into 2 or 7 clusters?
- & Gender <> “Man” => 15 participants (6 provided their emails)
- & Culture <> US and Germany => 9 participants (3 provided their emails)
- & English speak <> “Very Confident (fluent)” and “Confident” => 1 participant (did not provide their email)

Examples of the outliers

	Profile
Years at ASF	Over 10 year
Compensation	A mix but mostly paid
Volunteering	10+ hours
Challenges	Many
Residence/Culture	Netherland/ Italy
English Speaking	Average
Age	45 to 54
Gender	Prefer not to say
Education	Masters
Mentor	No
Contact	No

Challenges <> “No” and “Few” & Gender <> “Man” & Culture <> US and Germany & English speak <> “Very Confident (fluent)” and “Confident” => 1 participant

The three examples that provided contact info out of the nine profiles

	Profile 1	Profile 2	Profile 3
Years at ASF	6 to 10 years	over 10 years	over 10 years
Compensation	unpaid only	unpaid only	a mix but mostly paid
Volunteering	I don't have time	3 to 5 hours	6 to 10 hours
Challenges	Several	Many	Several
Residence/Culture	US/United Arab Emirates	US/ India	US/Canada
English Speaking	Very confident	Very confident	Very confident
Age	25 to 34	45 to 54	25 to 34
Gender	Woman	Prefer not to say	Tran woman
Education	Masters	Undergrad	Undergrad
Mentor	No	No	Yes
Contact	Yes	Yes	Yes

Challenges <> "No" and "Few" & Gender <> "Man" & Culture <> US and Germany => 9 participants

The three examples that provided contact info out of the 15 profiles

	Profile 1	Profile 2	Profile 3
Years at ASF	Over 10 years	3 to 5 years	6 to 10 years
Compensation	An equal mix, paid and unpaid	unpaid only	unpaid only
Volunteering	6 to 10 hours	10+ hours	3 to 5 hours
Challenges	Several	Several	Several
Residence/Culture	Ireland/ US	Germany/US	US/US
English Speaking	very confident	confident	very confident
Age	55 to 64	35 to 44	55 to 64
Gender	Woman	woman	Gender varian / Non-conforming / Non-binary
Education	undergrad	undergrad	masters
Mentor	No	yes	No
Contact	Yes	Yes	Yes

Challenges <> “No” and “Few” &

Gender <> “Man” => 15 participants

Other results and Insights

Contributor Roles and Tenure

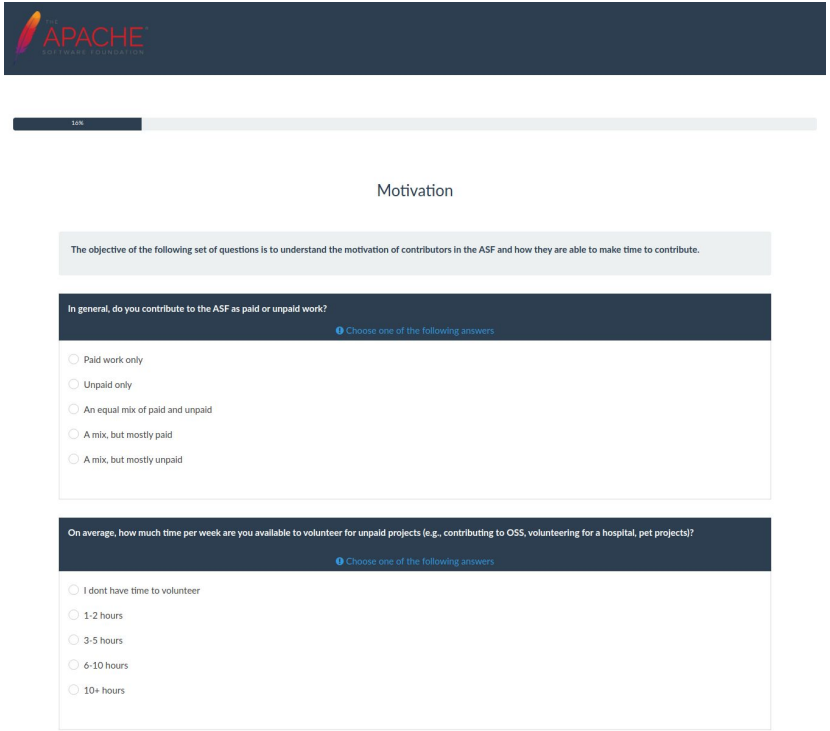
The objective of the following set of questions is to understand the different contributor roles and their experience in the ASF.

- How long have you been involved with the ASF?
- Which of the following tasks did you start with when you got involved in the ASF? Select as many as are relevant.
- How often do you engage in the following activities in the ASF?
- How long (in years) have you been in any of the roles in the ASF?

Motivation

The objective of the following set of questions is to understand the motivation of contributors in the ASF and how they are able to make time to contribute.

- In general, do you contribute to the ASF as paid or unpaid work?
- On average, how much time per week are you available to volunteer for unpaid projects (e.g., contributing to OSS, volunteering for a hospital, pet projects)?



Availability of Protocols / Guidelines

The objective of the following set of questions is to find out if the documentation around policies, processes, and guidelines is being referred to by contributors and any problems in accessing these guidelines.

- How often do you consult the following ASF policies, processes, or guidelines regarding...?
- When you need to locate information about ASF processes, policies, or guidelines, which of the following describes your experience?

APACHE SOFTWARE FOUNDATION

33%

Availability of protocols/ guidelines

The objective of the following set of questions is to find out if the documentation around policies, processes, and guidelines is being referred to by contributors and any problems in accessing these guidelines.

How often do you consult the following ASF policies, processes, or guidelines regarding...?

	Never	Rarely (less than once a month)	Sometimes (more than once a month)	Often (once a week or more)
Mailing list communication				
IRC communication				
Performing code reviews				
Process of getting code accepted				
Code of conduct				
Onboarding newcomers				
Licensing, trademark				
Add new committers/PMC members				
Project releases				
Voting process				

When you need to locate information about ASF processes, policies, or guidelines, which of the following describes your experience?

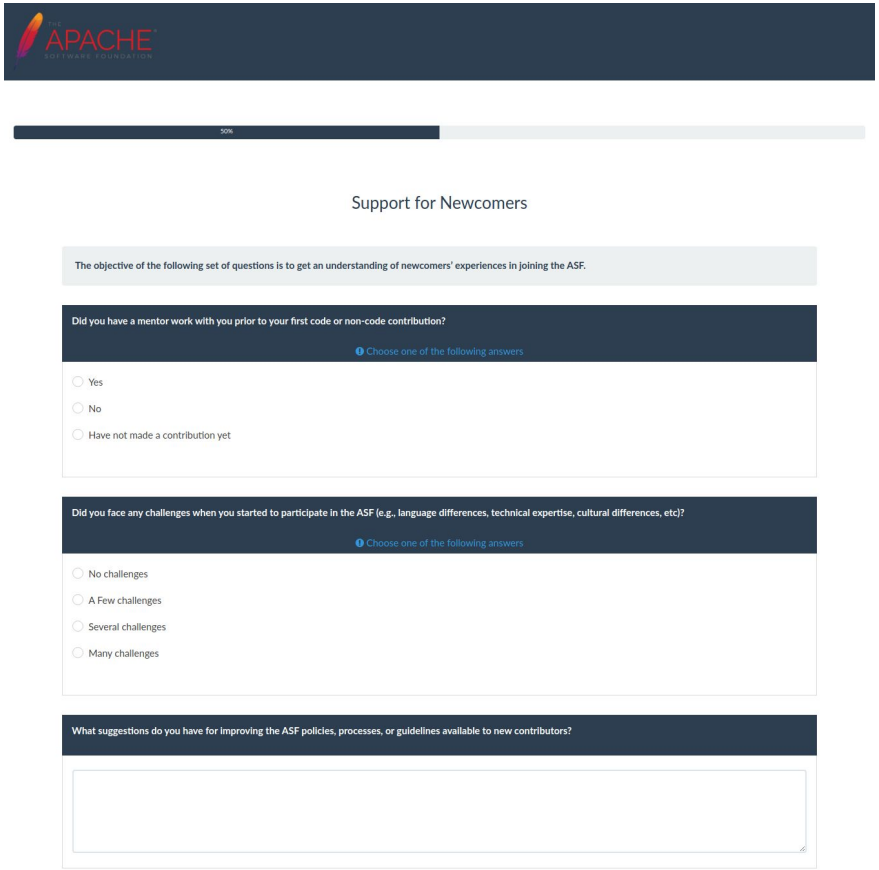
	Always easy to find	Easy to find	A little difficult to find	Quite difficult to find
Mailing list communication				
IRC communication				
Performing code reviews				
Process of getting code accepted				
Code of conduct				
Onboarding newcomers				
Licensing, trademark				
Add new committers/PMC members				
Project releases				
Voting process				

Survey Questions Look & Feel

Support for Newcomers

The objective of the following set of questions is to get an understanding of newcomers' experiences in joining the ASF.

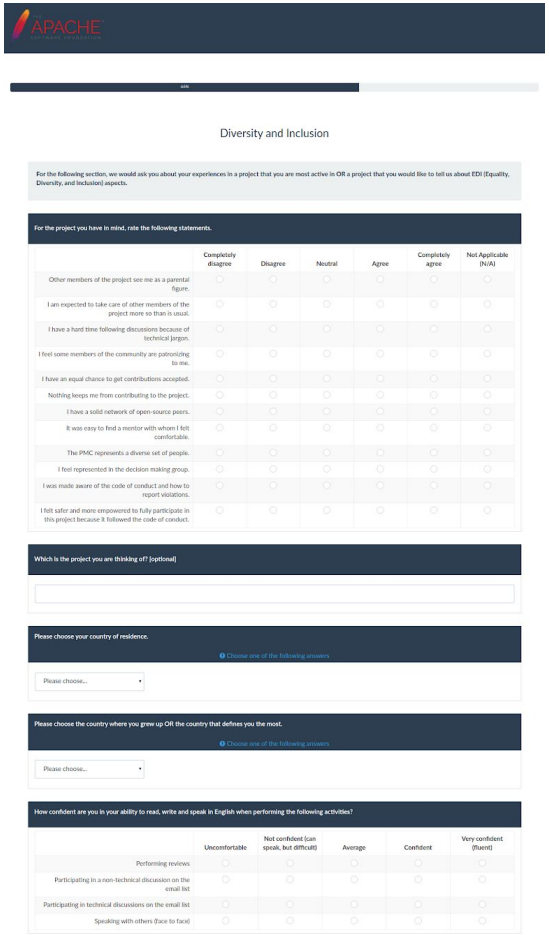
- Did you have a mentor work with you prior to your first code or non-code contribution?
- Did you face any challenges when you started to participate in the ASF (e.g., language differences, technical expertise, cultural differences, etc)?
- What suggestions do you have for improving the ASF policies, processes, or guidelines available to new contributors?



Diversity and Inclusion (1/2)

For the following section, we would ask you about your experiences in a project that you are most active in OR a project that you would like to tell us about EDI (Equality, Diversity, and Inclusion) aspects.

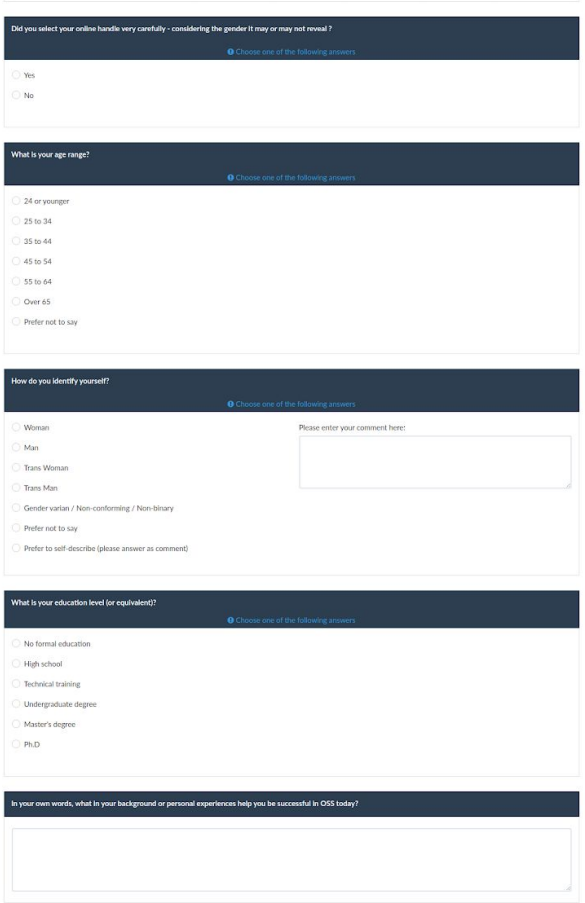
- For the project you have in mind, rate the following statements.
- Which is the project you are thinking of? [optional]
- Please choose your country of residence.
- Please choose the country where you grew up OR the country that defines you the most.
- How confident are you in your ability to read, write and speak in English when performing the following activities?
- [...]



Diversity and Inclusion (2/2)

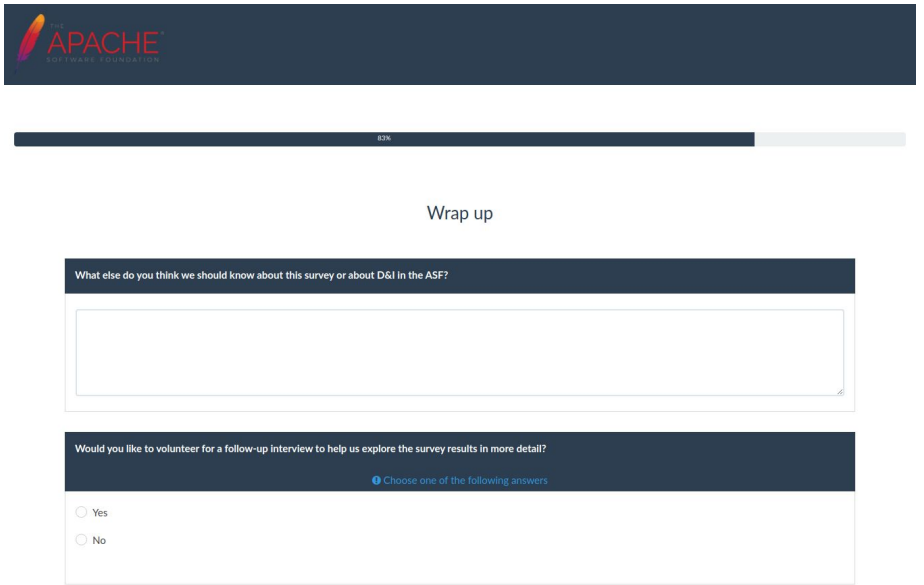
For the following section, we would ask you about your experiences in a project that you are most active in OR a project that you would like to tell us about EDI (Equality, Diversity, and Inclusion) aspects.

- [...]
- Did you select your online handle very carefully - considering the gender it may or may not reveal ?
- What is your age range?
- How do you identify yourself?
- What is your education level (or equivalent)?
- In your own words, what in your background or personal experiences help you be successful in OSS today?



Wrap Up

- What else do you think we should know about this survey or about D&I in the ASF?
- Would you like to volunteer for a follow-up interview to help us explore the survey results in more detail?



Age per years at ASF

Row Labels	Count of Age	% of Age2
3 to 5 years	**164**	**26.71%**
35 to 44	76	46.34%
25 to 34	50	30.49%
45 to 54	25	15.24%
55 to 64	8	4.88%
Over 65	3	1.83%
Prefer not to say	1	0.61%
24 or younger	1	0.61%
Over 10 years	**158**	**25.73%**
45 to 54	61	38.61%
35 to 44	59	37.34%
55 to 64	18	11.39%
Over 65	9	5.70%
Prefer not to say	7	4.43%
25 to 34	3	1.90%
24 or younger	1	0.63%
6 to 10 years	**127**	**20.68%**
35 to 44	53	41.73%
45 to 54	35	27.56%
55 to 64	19	14.96%
25 to 34	11	8.66%
Over 65	6	4.72%
Prefer not to say	3	2.36%
1 to 2 years	**99**	**16.12%**
25 to 34	50	50.51%
35 to 44	27	27.27%
45 to 54	9	9.09%
24 or younger	5	5.05%
55 to 64	5	5.05%
Prefer not to say	2	2.02%
Over 65	1	1.01%
Less than 1 year	**66**	**10.75%**
25 to 34	33	50.00%
35 to 44	16	24.24%
45 to 54	7	10.61%
24 or younger	5	7.58%
55 to 64	3	4.55%
Prefer not to say	1	1.52%
Over 65	1	1.52%
Grand Total	**614**	**100.00%**

Immigrants data: 134 data points

Row Labels	Count of residence	% of residence
Albania	1	0.75%
Australia	2	1.49%
Austria	1	0.75%
Belarus	1	0.75%
Bulgaria	1	0.75%
Canada	6	4.48%
Czech Republic	1	0.75%
France	5	3.73%
Germany	12	8.96%
Hungary	1	0.75%
Ireland {Republic}	2	1.49%
Japan	2	1.49%
Luxembourg	1	0.75%
Malaysia	1	0.75%
Morocco	1	0.75%
Netherlands	5	3.73%
New Zealand	3	2.24%
Norway	1	0.75%
Russian Federation	1	0.75%
Singapore	3	2.24%
Slovakia	1	0.75%
Spain	2	1.49%
Sweden	3	2.24%
Switzerland	4	2.99%
United Kingdom	8	5.97%
United States	63	47.01%
Vietnam	2	1.49%
Grand Total	**134**	**100.00%**

Row Labels	Count of culture	% of culture
Algeria	1	1
Andorra	1	1
Argentina	2	2
Australia	1	1
Austria	3	3
Belarus	1	1
Belgium	1	1
Brazil	3	3
Cameroon	1	1
Canada	6	6
China	12	12
Cuba	1	1
Denmark	1	1
France	10	10
Germany	5	5
Greece	1	1
India	18	18
Iran	1	1
Italy	1	1
Japan	2	2
Kazakhstan	2	2
Korea South	3	3
Kyrgyzstan	1	1
Lithuania	1	1
Mexico	3	3
Moldova	1	1
New Zealand	1	1
Norway	3	3
Portugal	2	2
Romania	3	3
Russian Federation	10	10
Serbia	1	1
Slovakia	1	1
South Africa	1	1
Spain	5	5
Sweden	2	2
Switzerland	1	1
Taiwan	1	1
Turkey	2	2
Ukraine	2	2
United Arab Emirates	1	1
United Kingdom	6	6
United States	9	9
Grand Total	**134**	**134**

Side by Side: Majority Profiles; Weighted and Non-Weighted

Non-Weighted	Cluster 1	Cluster 2
Years at ASF	3 to 5	3 to 5
Compensation	Unpaid only	Unpaid only
Volunteering	1 to 2 hours	1 to 2 hours
Challenges	No	Few
Residence/Culture	US /US	Germany/ Germany
English Speaking	Very Confident	Confident
Age	35 to 44	35 to 44
Gender	Man	Man
Education	Undergrad	Masters
Mentor	No	No
Careful handle	No	No

Weighted	Cluster 1	Cluster 2
Years at ASF	3 to 5	3 to 5
Compensation	Unpaid only	Unpaid only
Volunteering	1 to 2 hours	1 to 2 hours
Challenges	No	Few
Residence/Culture	US /US	Germany/ Germany
English Speaking	Very Confident	Confident
Age	35 to 44	35 to 44
Gender	Man	Man
Education	Undergrad	Masters
Mentor	No	No

Side-by-Side: Challenged dataset; Non-weighted & Weighted

Non-weighted	Cluster 1	Cluster 2
Years at ASF	3 to 5	Over 10 years
Compensation	Unpaid only	Mix but mostly paid
Volunteering	1 to 2 hours	3 to 5 hours
Challenges	Few	Few
Residence/Culture	US /US	US/ US
English Speaking	Very Confident	Very confident
Age	35 to 44	25 to 34
Gender	Man	Man
Education	Masters	Masters
Mentor	No	yes

Weighted	Cluster 1	Cluster 2
Years at ASF	6 to 10	3 to 5
Compensation	Unpaid only	Unpaid only
Volunteering	1 to 2 hours	1 to 2 hours
Challenges	Few	Few
Residence/Culture	Germany/ Germany	US/ US
English Speaking	Confident	Very confident
Age	35 to 44	*35 to 44
Gender	Man	Man
Education	Masters	Masters
Mentor	No	*No

Side-by-Side: Gender dataset; Non-weighted & Weighted

Non_Weighted	Cluster 1	Cluster 2
Years at ASF	Over 10 year	Over 10 years
Compensation	A mix but mostly paid	Unpaid only
Volunteering	1 to 2 hours	3 to 5 hours
Challenges	Few	No
Residence/Culture	US/US	US/US
English Speaking	very confident	Very confident
Age	35 to 44	25 to 34
Gender	Women	Women
Education	Undergrad/Masters	Masters
Mentor	No	No

Weighted	Cluster 1	Cluster 2
Years at ASF	Over 10 year	1 to 2 years
Compensation	Unpaid only	A mix but mostly paid
Volunteering	3 to 5 hours	1 to 2 Year
Challenges	Few/ No	No
Residence/Culture	US/US	US/India
English Speaking	very confident	confident
Age	35 to 44	25 to 34
Gender	Prefer not to say/ women	Women
Education	Undergrad	Masters
Mentor	No	No