

CS 458-558 Project 1
2020 Stack Overflow Developer Demographics
Milan Thakkar, Conner Rhea, Grace Todd

Introduction

Stack Overflow is a powerful tool for professional programmers and hobbyists alike that provides a forum for users to ask questions and seek answers for their questions and struggles. Hundreds of thousands of users make use of its service and its status as a quick fix when debugging has transcended into legendary status. Every year Stack Overflow polls its users for data about its contributors and what kind of developers they are, giving users insight into the kind of people answering their questions. This document primarily analyzes data about education level, the age of Stack Overflow users, and their location to see the effects if any of these parameters imply for programmers.

Number of Respondents v Age

Number of Respondents per Age Group

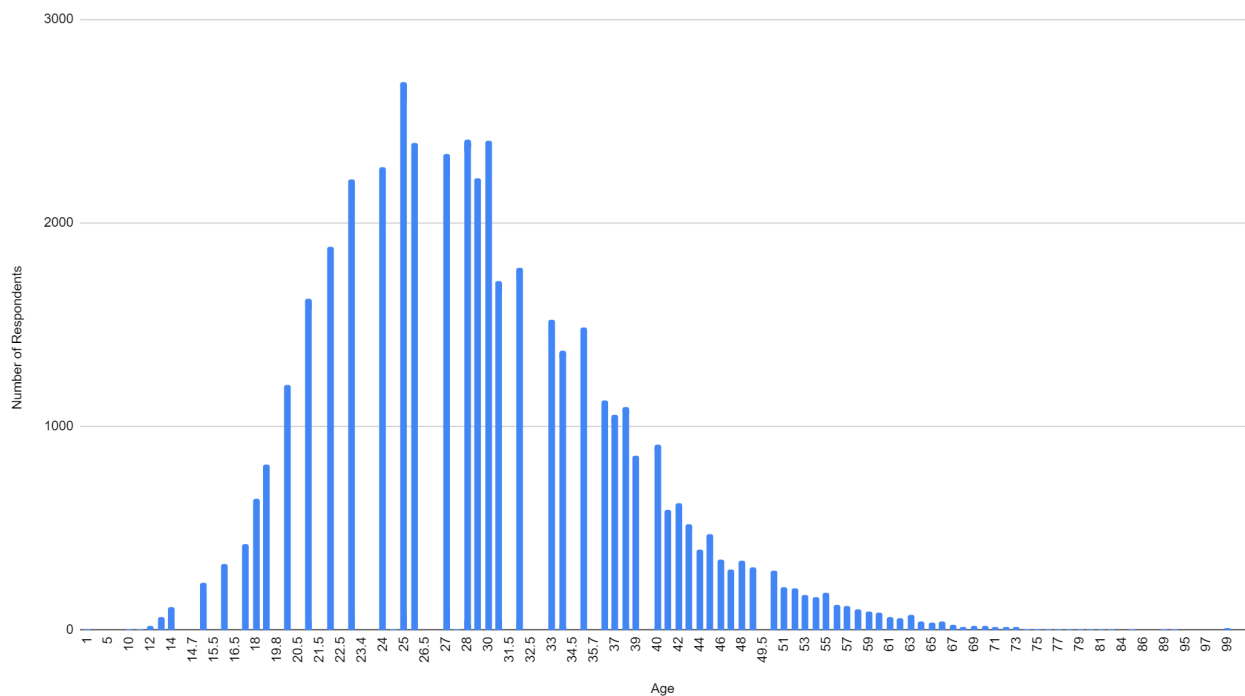


Figure 1: Number of survey respondents per age group

For the respondents of the 2020 Stack Overflow survey, the youngest known participants reported being 12 years old, whereas the oldest known participants reported being 99 years old. The greatest number of participants reported being 24 years old, at a total of 2275 24-year old developers. If we assume that respondents answered the survey honestly, we can determine that there is a fairly broad age range for Stack Overflow developers. The average age of a developer can speak volumes about the level of experience and generational culture around the developers on this site. For example, what kinds of programming languages were popular when they were learning to code, when they attended college, what kinds of degrees they could have or are currently earning, etc.

Education Level v Age First Coded

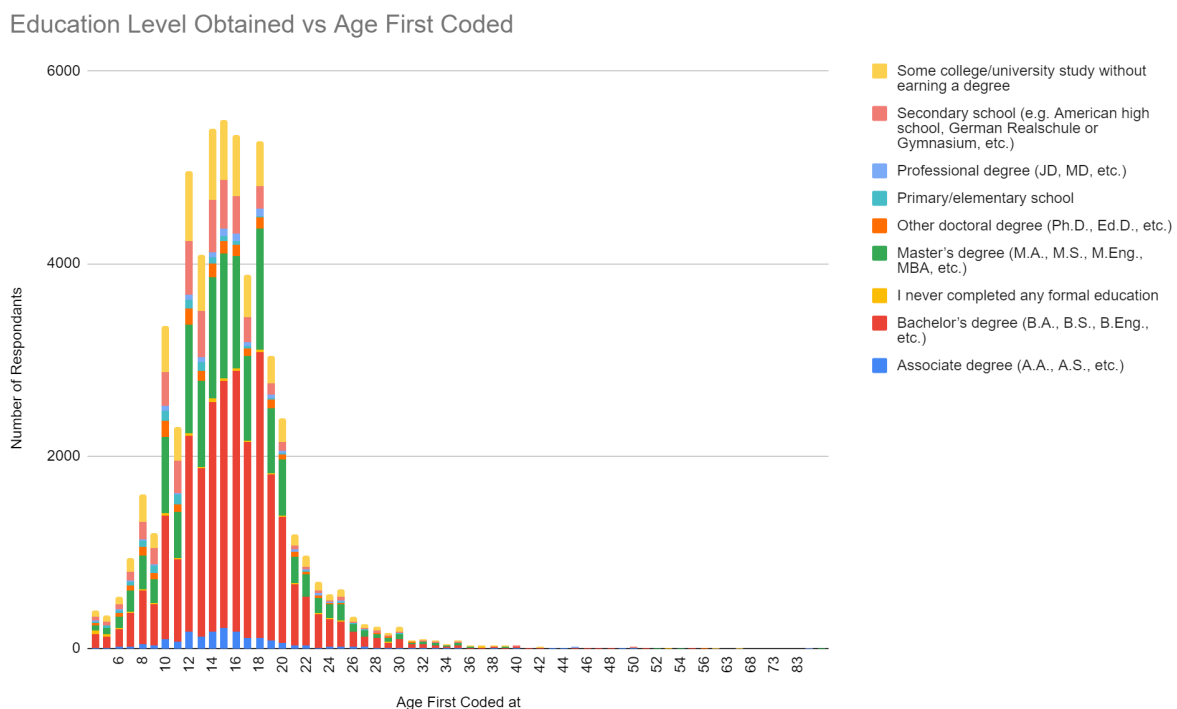


Figure 2: Education level obtained versus age first coded

A common concern amongst college students in programming degrees is that coders who started earlier in life have a major advantage over coders that didn't start until later in life. To this end, an analysis of Stack Overflow's data can provide valuable insight on whether people who started to code earlier in life are more likely to acquire a higher education level than those who did not. However, the data does not back up this common fear of the average college student.

According to the Stack Overflow data, while higher level degrees are more likely to be held by people that start their coding careers early on, the majority of Bachelor and Master degree holders in computer science started in their late teenage years and early 20's. While the data does taper off as the users get older, there were also far fewer responders in older age groups compared to younger ones, likely due to programming becoming more and more accessible as time goes on to the current year.

So the data seems to suggest that even if any given person started coding in their mid to late 20's they could still gain a degree in programming and have a successful career as a professional programmer.

Years Coded v Education Level

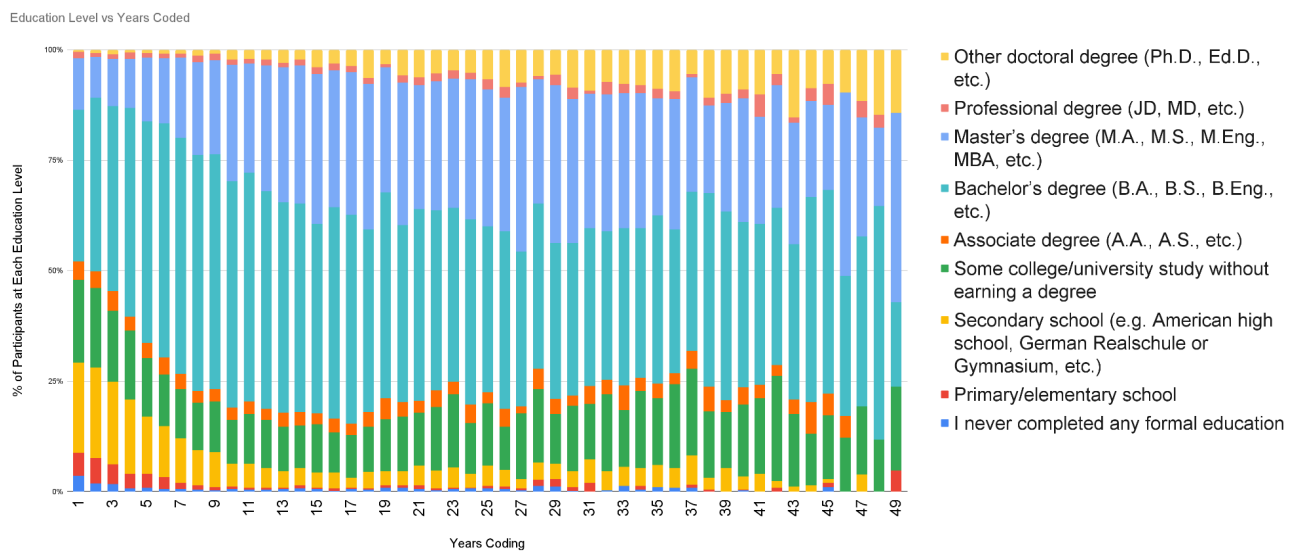


Figure 3: Education level of stack overflow coders versus years spent coding

Another interesting statistic to look at was the education level of each participant per the years they had been coding. For the survey participants that had a total of 8 years experience coding or fewer, a large majority either started college or finished a bachelor's degree, however past 8 years of experience the percentage of people who had not finished their degree showed a downward trend, while the percentage of those who completed a master's degree or higher showed an upward trend in percentage. Meanwhile, the percentage of people who had completed a bachelor's degree remained about the same as the number of years coding increased. The data is

interesting, and may imply that those who don't at least attempt a degree in CS may not stick with the career of coding as long as those who earn a bachelor's or graduate degree. It would be both relevant to the data and interesting to know how many of those with college degrees or higher did their college degree in computer science, or if they switched fields after graduating.

Population of Participants Per Country

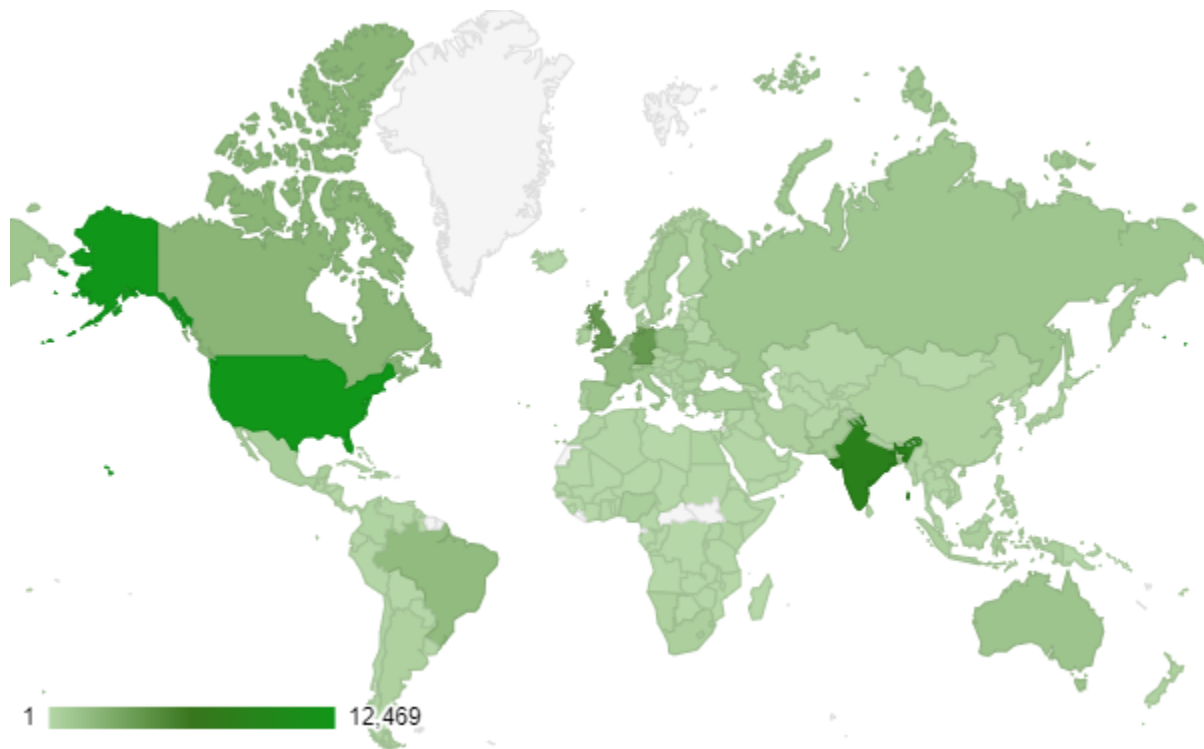


Figure 4: Population of participants per country

The survey demographic reaches a significant number of countries, but the countries that are home to the greatest number of participants are the United States with approximately 12,469 reported developers, and India with approximately 8,403 developers. The number of developers in less popular countries was somewhat surprising, with countries like Poland and Indonesia housing more than 400 developers each. The geographic reach of developers for Stack Overflow is expansive, and potentially allows for more diversified approaches to answers for questions posed on the site. Developers from different countries, who speak different languages and have different educational backgrounds can help to provide a stronger support system for users of Stack Overflow.