

Project: Programming Language - C++ (WG21)
Author: Andrew Tomazos <andrewtomazos@gmail.com>
Date: 2021-04-23

How much C++ code is there? How many C++ programmers?

It would be useful, for various reasons, to know how much C++ code there is in the world and how many C++ programmers there are.

First of all, let us clarify a few things. A lot of code gets written and is used little, or not at all. For example, students write code for practice. Hobbyist programmers write code for fun or to experiment. A lot of code gets written for projects that fail (the product or service fails in the market, or becomes obsoleted). A lot of code is “throw away” code written to be run once (although typically such code is not written in C++ but a different language). A lot of code gets written that is then refactored or factored out.

For the sake of this study, we’re interested only in how much *production* C++ code exists today. We defined production code as code that is both actively maintained and used to build software that is currently used (directly or indirectly) by a significant number of people.

Based on the authors extensive experience working with different codebases both open and proprietary (including but not limited to the clearly largest codebase on the planet) it has been observed that the current size of an organizations production codebase seems to be in (perhaps surprisingly) approximately constant proportion to the number of full-time programmers working on extending and maintaining that codebase. This constant is approximately 10,000 lines of code per current full-time programmer. (It would be helpful if the reader could divide the number of lines of production C++ code they have at their company, or know of, by the number of full-time C++ programmers they have and let the author know the result by email - particularly if it is wildly different from this constant.)

If we can accept this constant ratio, then we only have to answer the second question (How many C++ programmers are there?) and can multiply by 10,000 to get the answer to the first.

To answer how many C++ programmers there are, we are going to take a top-down approach.

We start with the fact that there are 8 billion people in the world.

Let us say that about half of them are too young or too old to be C++ programmers, leaving 4 billion of working age.

Let us say of those about half of them are of an inadequate economic status (say < \$5k GDP/head) to have access to a programming career through lack of resources and other economic pressures. This leaves 2 billion people.

For whatever reason (Authors note: and let me make it very explicitly clear, I do not make any comments on the reasons. I have absolutely no idea WHY this is true.) we observe that currently the vast majority of programmers are male (> 90%). Let us round this up to 100%, leaving half remaining, 1 billion people.

To go further, of the 1 billion people that potentially are C++ programmers, how many actually are?

Let's take that in two steps. How many of those 1 billion people become programmers? And how many programmers are C++ programmers?

So let's conduct a thought experiment, if we think back to the last year of school where people did not have a choice in the subjects they took (usually about 14 years old in most countries) and then think of those that ended up becoming programmers, how many do you know (including yourself if so) ended up becoming programmers and how many did not? From various (admittedly anecdotal) evidence the ratio is about 1 in 50. There are lots of other careers apart from programming. So we are down to 20 million programmers. Most people do not spend their entire career as programmers (moving to and from various support roles, management, or live on the border of programming combined with some other STEM activity), so let's half that again, to 10 million real current working programmers.

Of those 10 million working programmers, how many are primarily reading/writing C++ as opposed to some other language?

So we can appeal to the TIOBE index which approximates the amount of programming related internet activity that is C++ related (from search engines). We can also look at programming job ads and the ratio that ask for a language to how many ask for C++. We're comfortable that a 10% ratio can be supported by evidence - that is 1 in 10 working programmers are primarily reading/writing C++ code.

This leaves us with the answers:

There are approximately 1 million working C++ programmers in the world.

Multiplying that by 10,000 leaves us with:

There are approximately 10 billion lines of production C++ code in the world.