

# Mathematics Developers Survey 2016

Nejc Ilenič

## 1 Introduction

Anonymised responses from Stack Overflow Annual Developer Survey are published each year along the results to encourage their further analysis. Being curious about where in the world an aspiring data scientist should start his/her career, I have decided to use the shared data in an attempt to answer this question and to learn more about people identifying themselves as mathematics developers.

An important thing to note when interpreting the results is that this data may not be a representative sample from the population of mathematics developers. One should keep in mind that these are developers who were aware of the survey and were willing to answer the questions.

## 2 Data Preparation

The dataset was constructed from survey that took place from January 7 to January 25, 2016, with responses originating from Stack Overflow, Stack Exchange technical sites, Facebook and Twitter. Raw data consists of 56030 samples and 66 features, all of which are optional.

In order to obtain an adequately sizable sample, I have decided to include all respondents that belong to the occupation group of mathematics developers, which includes data scientists, machine learning developers and developers with statistics and mathematics backgrounds. After filtering out other occupations and responses with unknown countries we are left with 2132 samples. In the next section a basic exploratory analysis is carried out.

## 3 Exploratory Data Analysis

Number of mathematics developers per country can be seen in Figure 1. Minimum number of 40 respondents is required to take the country into account and all others are placed into a single group called *Other*. Note that selected countries and number of people may be different when doing inference of specific features due to missing values.

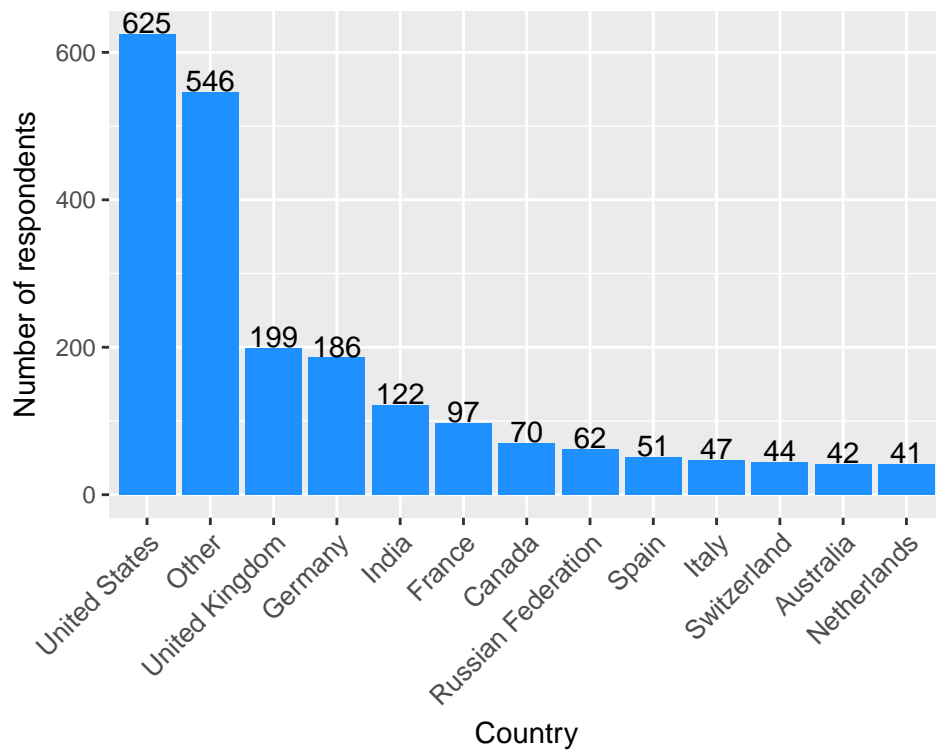


Figure 1: Number of mathematics developers per country.

todo