

This document presents the methods used in the Python study (IRB approved) that resulted in this public dataset: participants demographics, recruitment methods and participant management. The recruitment phase started in August 2017 and lasted until August 2018.

**Recruitment Methods.** Given the specialized cohort of participants needed for this study (developers, mostly professional developers working in the industry) recruitment mainly leveraged the networking connections and mailing lists available to PI Oliveira, PI Cappos, PI Brun, and to computer science students involved in the project. We leveraged mailing lists, word-of-mouth, github lists, and Facebook groups of Python developers.

Individuals were recruited across the world from North and South America ( $n = 110$ , 56.9%), Asia ( $n = 62$ , 32.1%), and Europe ( $n = 17$ , 8.8%), from the following countries: United States, Brazil, India, Bangladesh, Pakistan, Malaysia, Greece, Poland, France, England, Germany and the Middle East. Professional developers who work or have worked in the industry were compensated with a \$50 gift card. Senior undergraduate students majoring in computer science, computer engineering and related fields were compensated with a \$20 gift card.

In total we recruited ( $N = 193$ ) developers age 18 and 71 years ( $M = 28.24$ ,  $SD = 7.30$ ) who successfully proceeded to all phases of the study . Most of the developers were male ( $n = 173$ , 89.6%) and work in the industry: 171 (68.9%) professional developers vs 23 (31.1%) senior undergraduate or graduate students . Most developers ( $n = 171$ , 88.6%) had been programming in Python for more than two years.