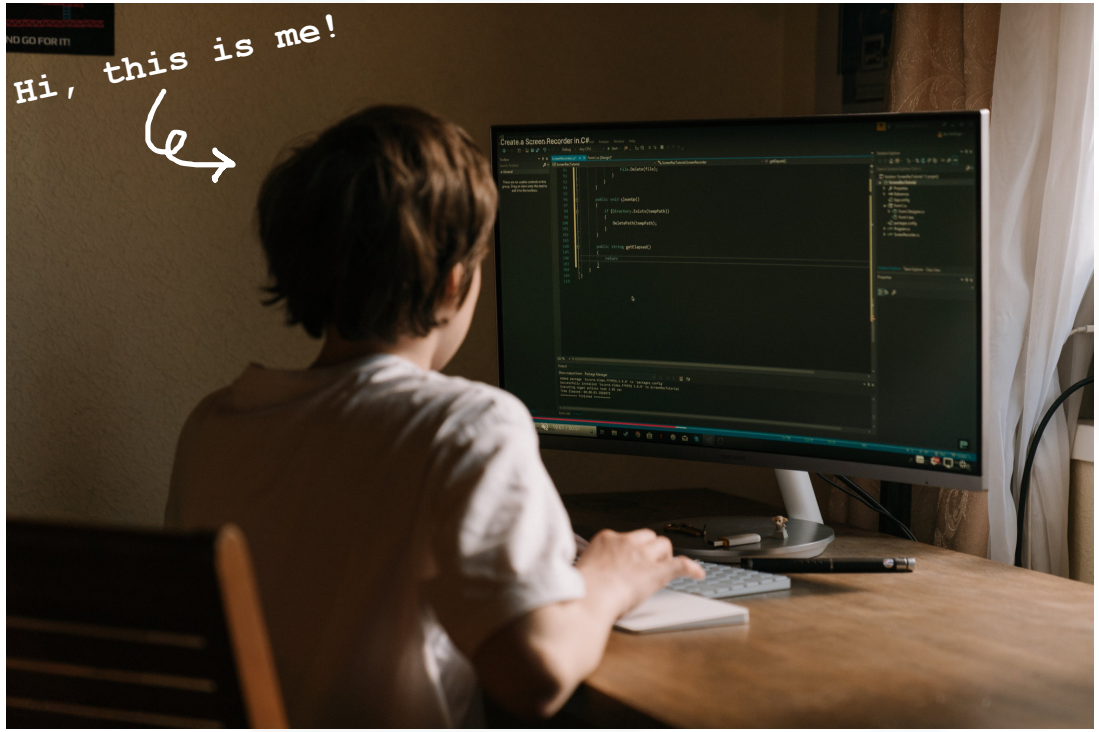
COVER PAGE



Technology trends of developer community on Stack Overflow in 2019

Do Bao Phuc

February, 2023

IBM Developer

SKILLS NETWORK

OUTLINE

EXECUTIVE SUMMARY

- The following slides demonstrate some key findings from an online survey of software developers around the world conducted by Stack Overflow in 2019. The data is available for download under an Open Database License.

- Some significant insights generated include:
  - The technology usage trend (e.g. which languages, platforms, databases,... are used the most) at the time of data collection.
  - The technology usage trend in the next years.
  - Gender gaps in the software developing industry.

INTRODUCTION

- The main objective of this report is to analyze technology usage trends as well as demographic features in the software developing work fields.
- Previous major reports on this topic may include "***An empirical study of Programming Language Trends***" by Y. Chen, R. Dios, A. Mili, L. Wu, K. Wang, which covers programming language trends from 1993 – 2003. While the report is intensively conducted, its methodology is acknowledged by the authors to be quite arbitrary.
- This report shows an update to 2019 with a large data sample collected from one of the most popular sites of developers around the world, Stack Overflow, with a widely approved methodology.
- The findings may bring about considerable referencing benefits for:
  - Aspiring developers, IT students
  - HR, Corporate managers, Policy makers

METHODOLOGY

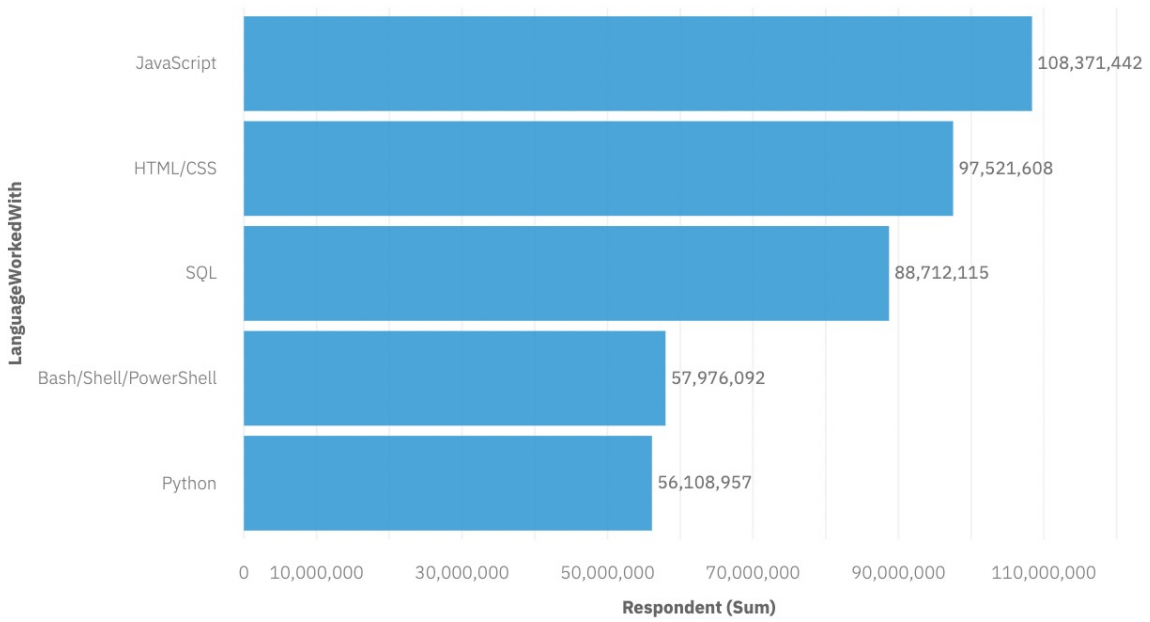
- **Data source: Stack Overflow Developer Survey 2019**
  - Link: https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/
- **Data collection:**
  - The dataset was loaded from IBM Cloud into a data frame by using Pandas. The data visualized in the Appendix was collected by web scraping and APIs.
  - Link on IBM cloud: https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv
- **Data wrangling:**
  - Duplicate removal, missing values imputation and data normalization were conducted on the dataset using Python.
- **Analysis & Visualization**
  - Possible factors were visualized, then outliers were removed before correlation was calculated between the targeted variable and other numerical variables by using Seaborn, Matplotlib and Pandas.
  - A Cognos dashboard was made visualizing the key insights about technology usage trends and demographics.

RESULTS

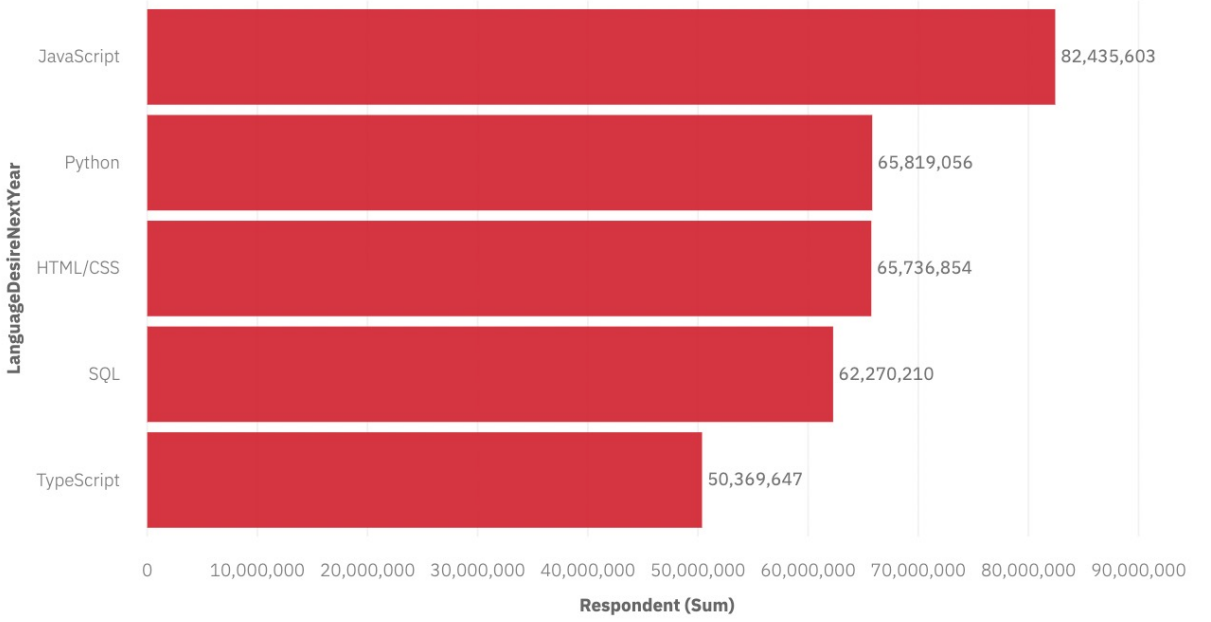
1. Programming language trends
2. Database trends
3. Dashboard

PROGRAMMING LANGUAGE TRENDS

Top 5 most popular languages currently



Top 5 most popular languages in the next year



PROGRAMMING LANGUAGE TRENDS – FINDINGS & IMPLICATIONS

Findings

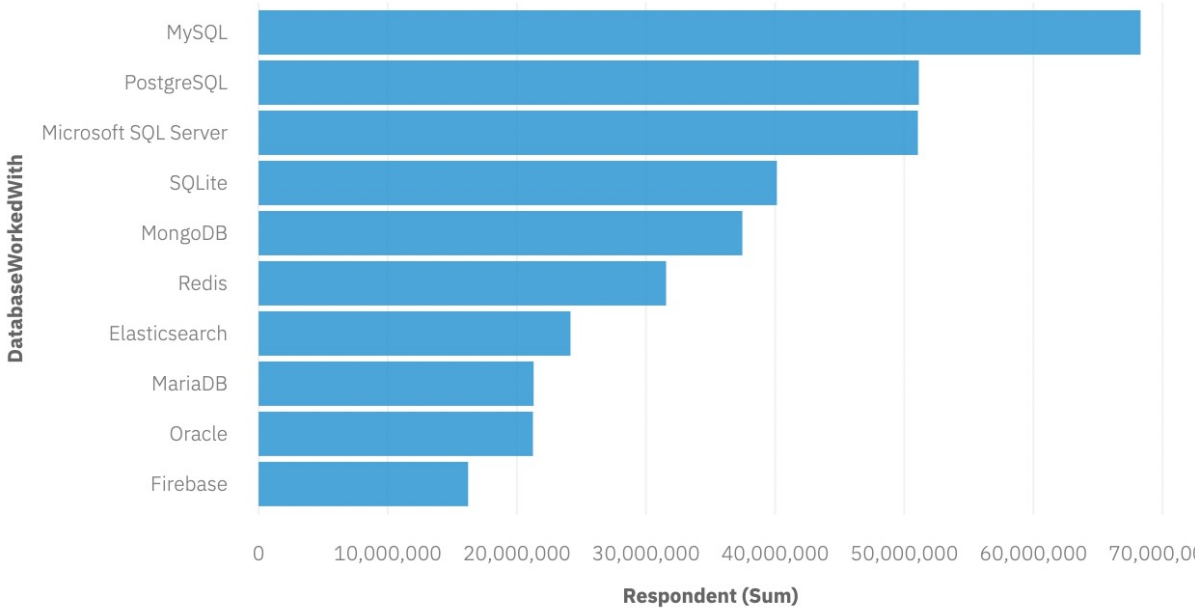
- Python, though stood at the bottom in 2019, was expected to be the second most loved language in the following year.
- JavaScript and HTML/CSS figures, though witnessing a steady decrease, are predicted to remain their popularity in the IT community.
- TypeScript emerged as a new trend in the next few years.

Implications

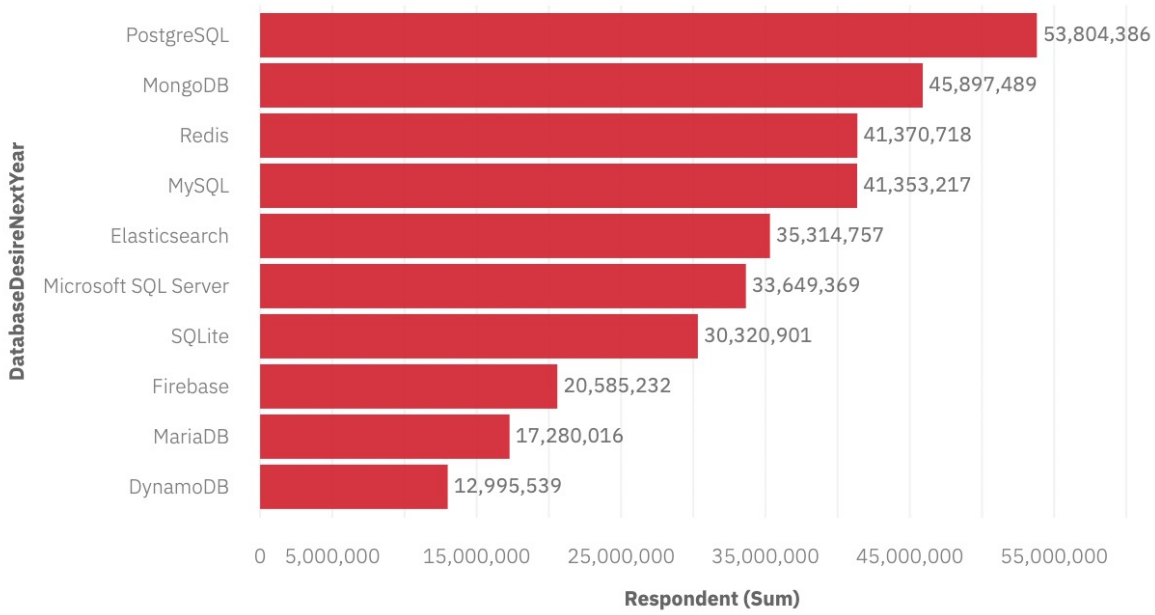
- The high demand for data analytics and machine learning job is still on a surge.
- JavaScript and HTML/CSS are still on demand to help business websites attract their customers' attention.
- Despite being a core of web programming, JavaScript will have a new competitor – TypeScript, which is widely known and is becoming more and more important.

DATABASE TRENDS

Top 10 most popular databases currently



Top 10 most popular databases next year



DATABASE TRENDS – FINDINGS & IMPLICATIONS

Findings

- MySQL and SQLite usage were expected to fall considerably, while PostgreSQL ranked 2nd in 2019 and was expected to lead the rank in the following years.
- MongoDB and Redis was predicted to have a considerable jump to the 2nd and 3rd position respectively in the coming years.
- A new player – Elasticsearch would join the market in the following years.

Implications

- Developers was turning into an open-source DBMS that store large, sophisticated datasets and was completely ACID.
- There were a high demand of working with NoSQL DBMS to deal with non-relational and unstructured data.

DASHBOARD

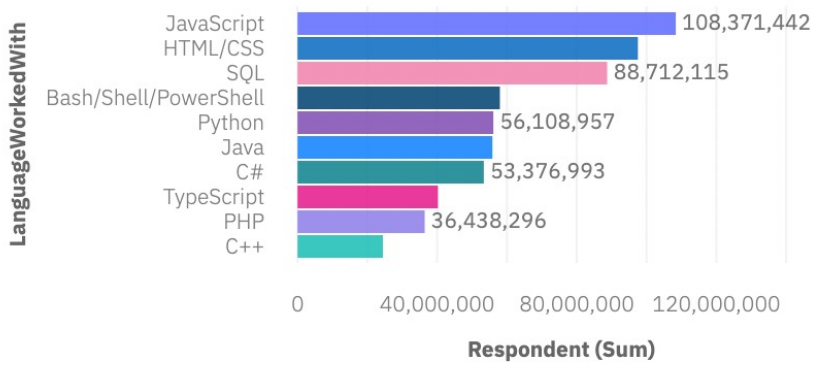
The Cognos dashboard below visualized key insights about technology usage trends in the current year (tab 1), in the next year (tab 2) and demographics in the software developing work field (tab 3).

Click here to get access to the Cognos Dashboard.

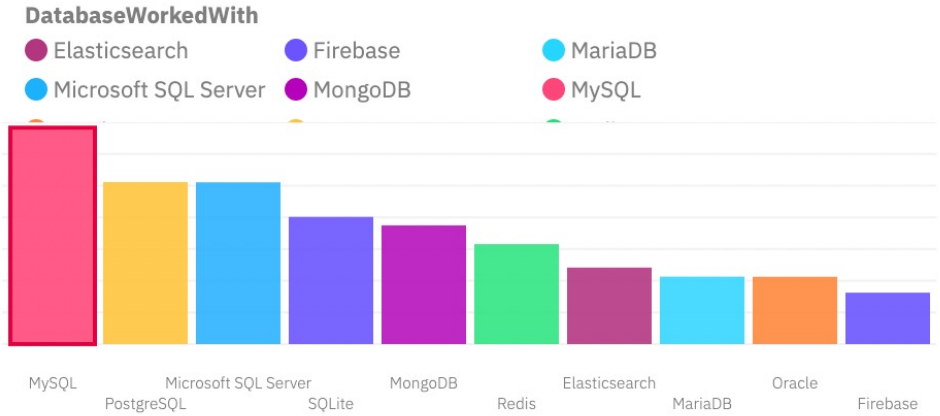
Note: You have to register/ sign in IBM Watson Studio to view the dashboard.

DASHBOARD TAB 1

Top 10 Languages Worked With

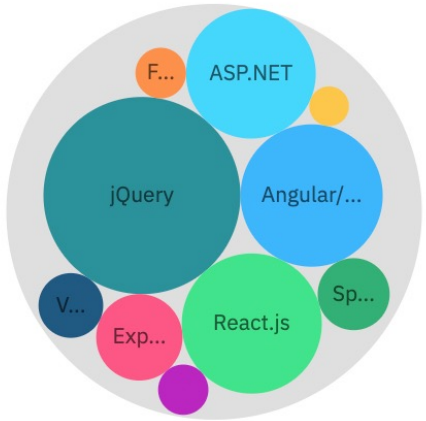


Top 10 Databases Worked With

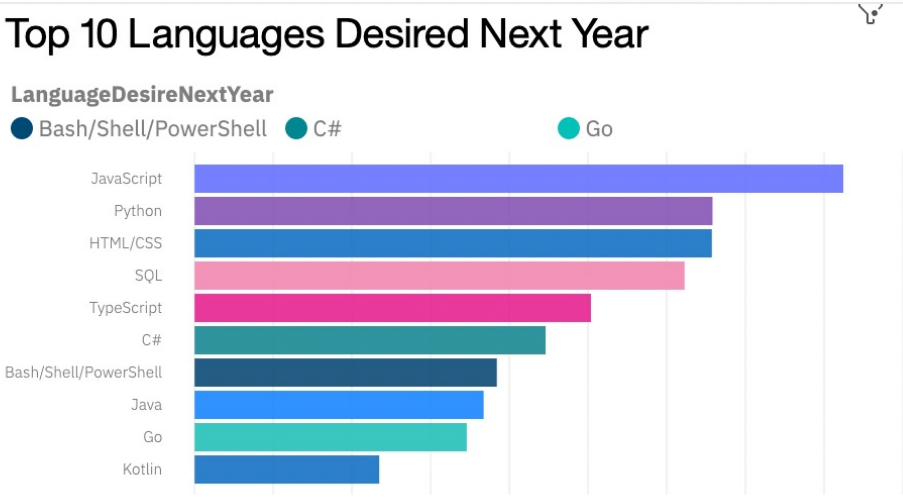


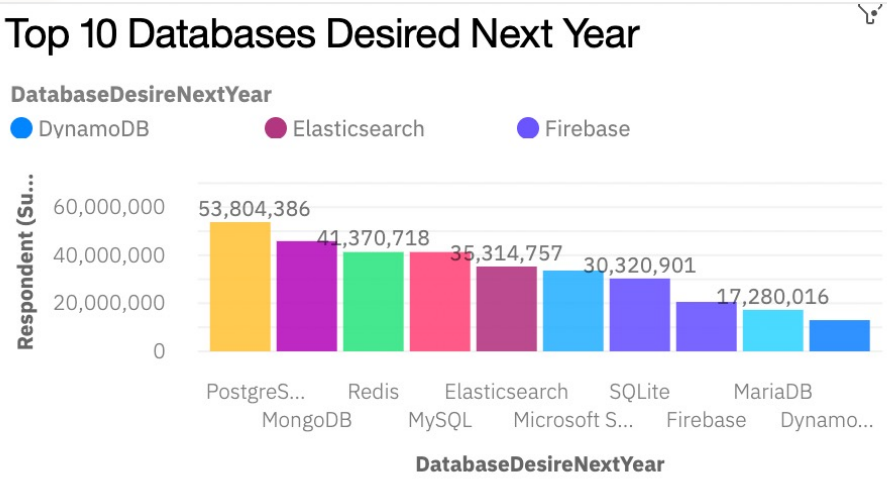
Platforms Worked With

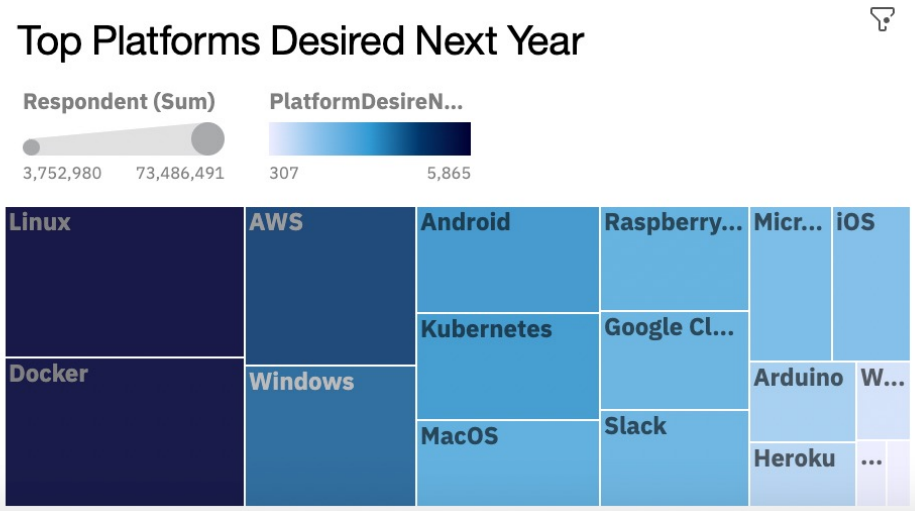
Top 10 Web Frames Worked With

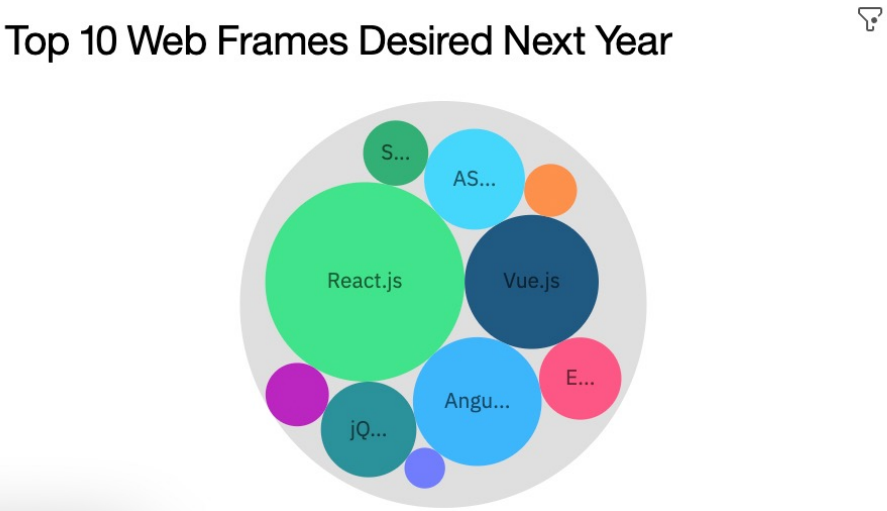


DASHBOARD TAB 2



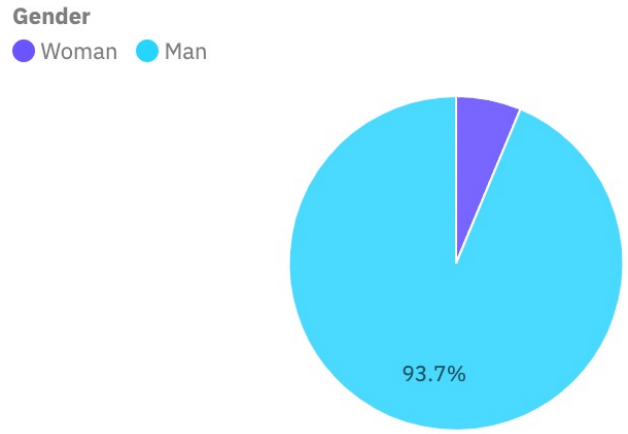




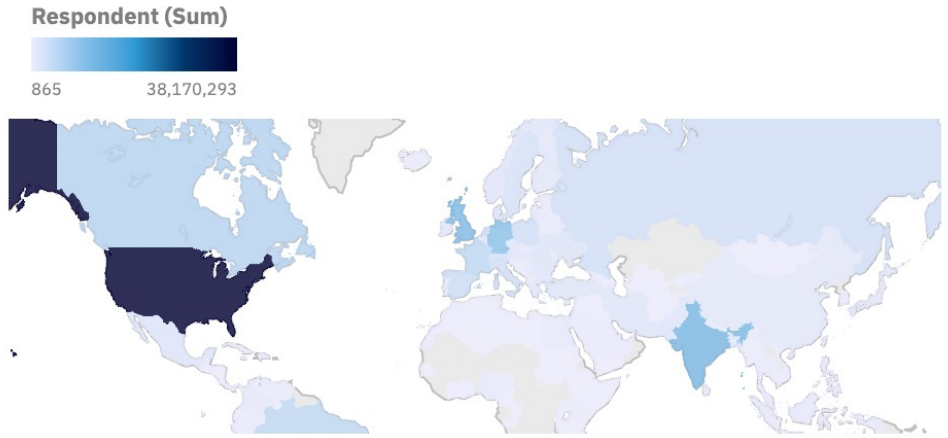


DASHBOARD TAB 3

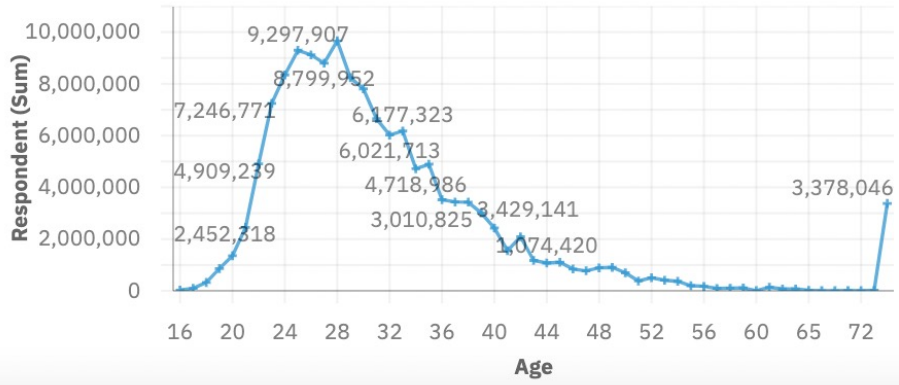
Respondent classified by Gender



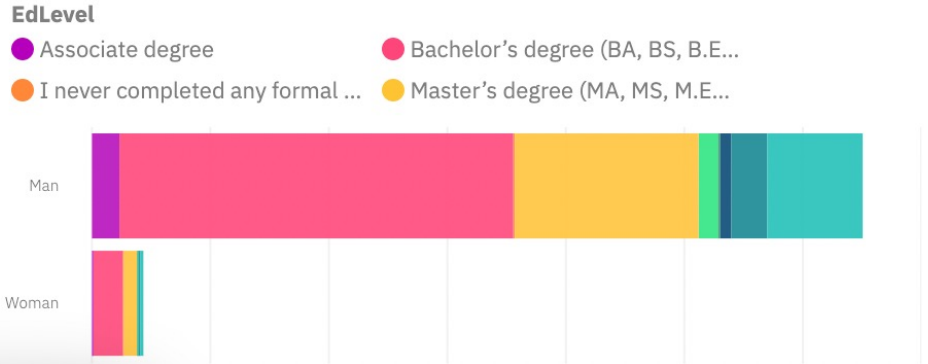
Respondent Count for Countries



Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level



DISCUSSION

It can be seen that the dashboards present insightful answers to some critical questions for all kinds of mentioned viewers:

1. For aspiring developers and IT students

+ What will be the trends in technology usage in the software development industry and what could be implicated from the trends?

2. For HR, industry manager

+ Which age group contains the most crowded number of potential staffs in the industry?

+ Where are the potential staffs located?

3. For Policy makers

+ Are there any barriers for women and other genders to join the IT industry and should we create more environment for them to develop their ability?

OVERALL FINDINGS & IMPLICATIONS

Findings

- Python, JavaScript and SQL are gaining a huge popularity among software developers.
- PostgreSQL, MongoDB and Redis would be the 3 most favorable DBMS in the coming years.
- Most of the respondents were located in the United States.
- Male workers dominated the other genders in the work field.

Implications

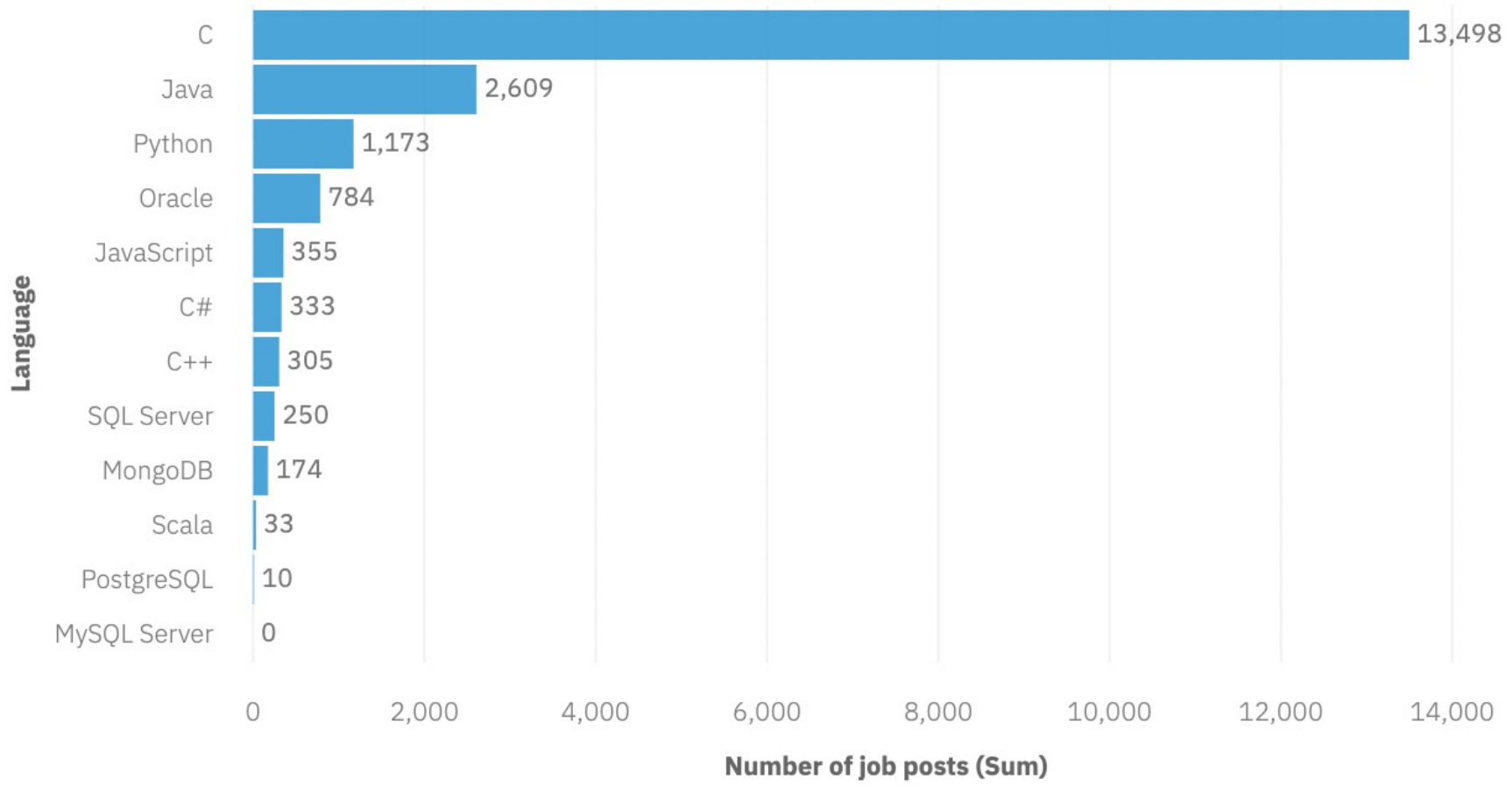
- Jobs relating to Data and ML are on a surging demand, and the same goes with web development.
- Developers was turning into an open-source DBMS that store large and sophisticated dataset (non-relational and unstructured data)
- Due to the predominance of the Americans in the survey, its result possibly does not apply to the social context of countries around the world.
- Policy makers, educators, and organizations should work to minimize the gender representation gap in addition to the technology divide between countries.

CONCLUSION

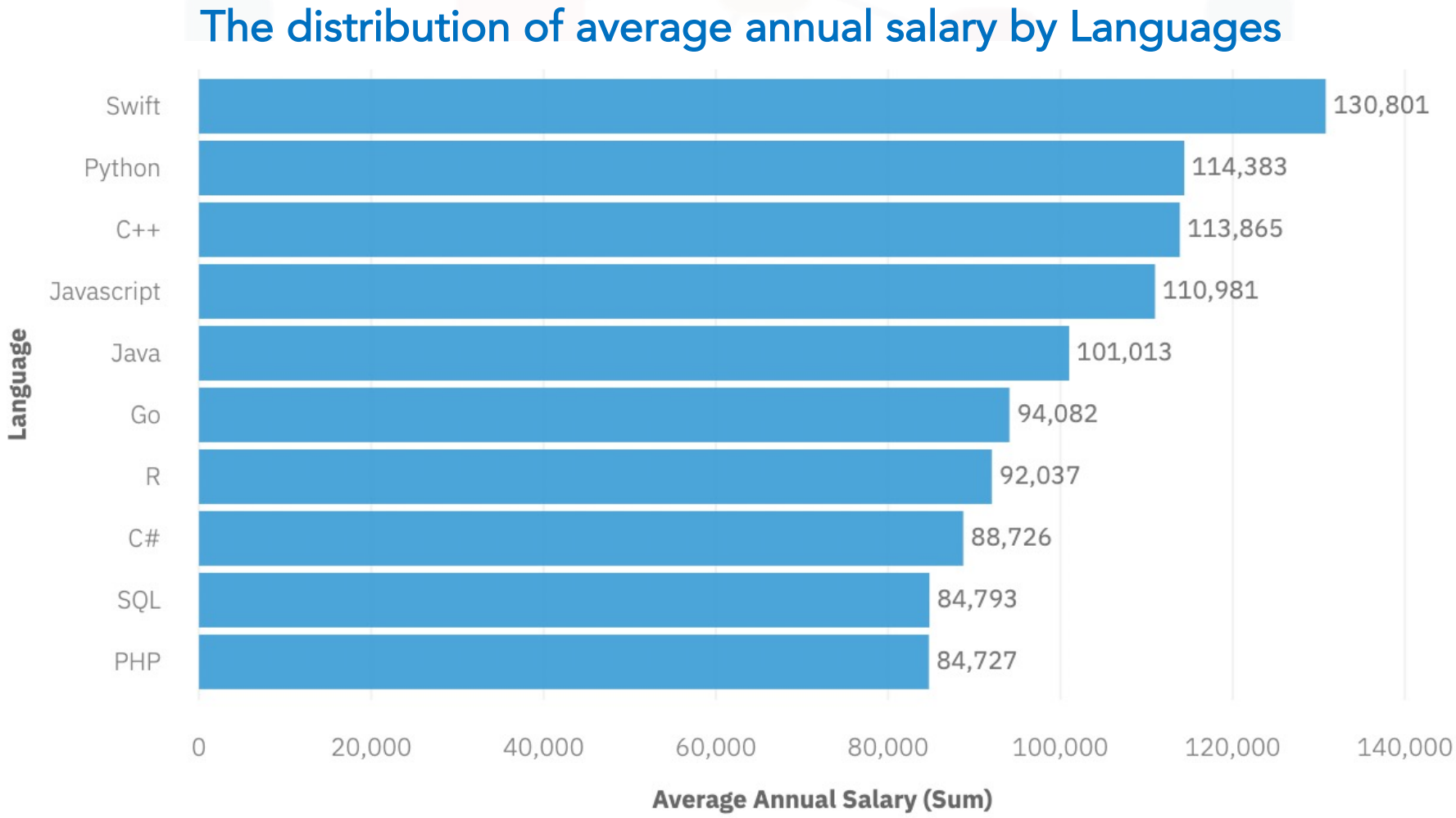
- A subset of data collected from the **Stack Overflow Developer Survey 2019** was examined.
- The findings harvested numerous insights into the trending and desired technology in one of the world's biggest community in addition to their demographics.
- These insights should be particularly relevant for current and aspiring developers aiming to remain competitive, businesses aiming to upskill their talent or get new hires, educators and policy makers aiming to address gender as well as economic issues.

APPENDIX 1 - JOB POSTINGS

The distribution of job opportunities by Languages



APPENDIX 2 - POPULAR LANGUAGES



APPENDIX 3 – JOB LOCATIONS