IBM DATA ANALYST CAPSTONE PROJECT ON

STACK OVERFLOW DEVELOPER SURVEY

BY NIKHIL PANGAONKAR

IBM Developer

SKILLS NETWORK

OUTLINE

- Executive Summary
- Introduction
- Methodology
- Results
  - Visualization – Charts
  - Dashboard
- Discussion
  - Findings & Implications
- Conclusion
- Appendix

EXECUTIVE SUMMARY

- Data contextualization and analysis goal.
- Methodology description.
  - Data gathering.
  - Data analysis.
  - Data visualizations.
- Results presentation supported with graphs and trends.
- Discussion of overall findings and implications regarding the results previously exposed.
- Final conclusions of the carried out research.

INTRODUCTION

- Stack Overflow's annual Developer Survey is the largest and most comprehensive survey of people who code around the world.
- Results don't represent everyone in the developer community evenly.
- Nearly 90,000 developers.
- Trends to predict where the developers are going.
- Characterization of developers around the globe.

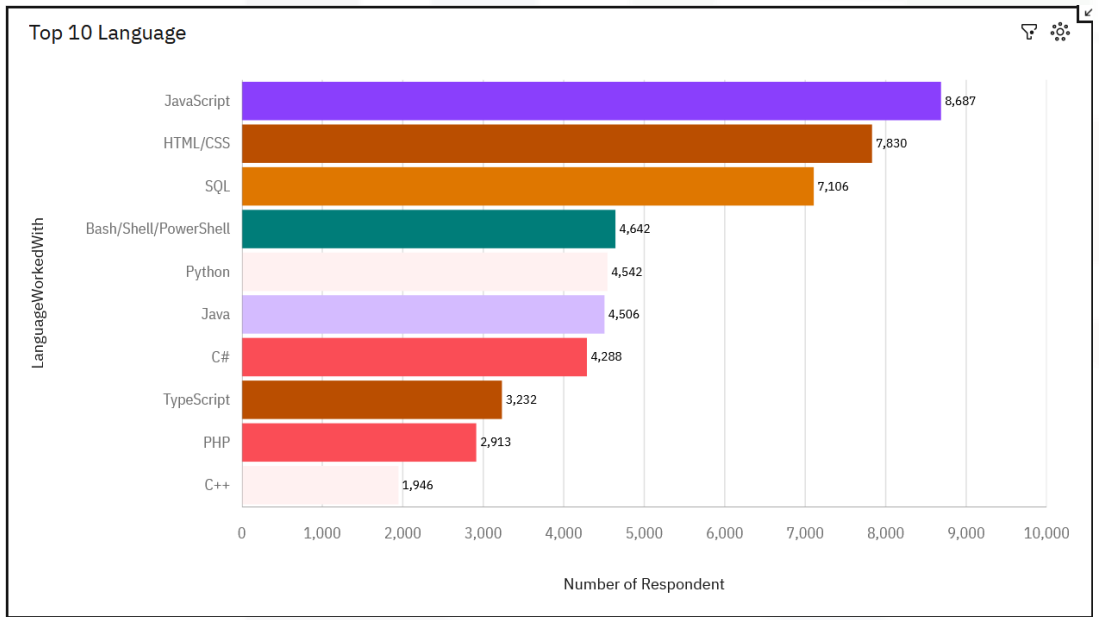
METHODOLOGY

- Collect survey data & explore its content
  - Web Scraping
  - APIs.
  - Request library.
- Data Wrangling
- Exploratory data analysis
  - Analyzing data distribution.
  - Handling outliers.
  - Correlations.
- Data Visualization
  - Highlight distribution of data, relationships, the composition and comparison of data.
- Dashboards

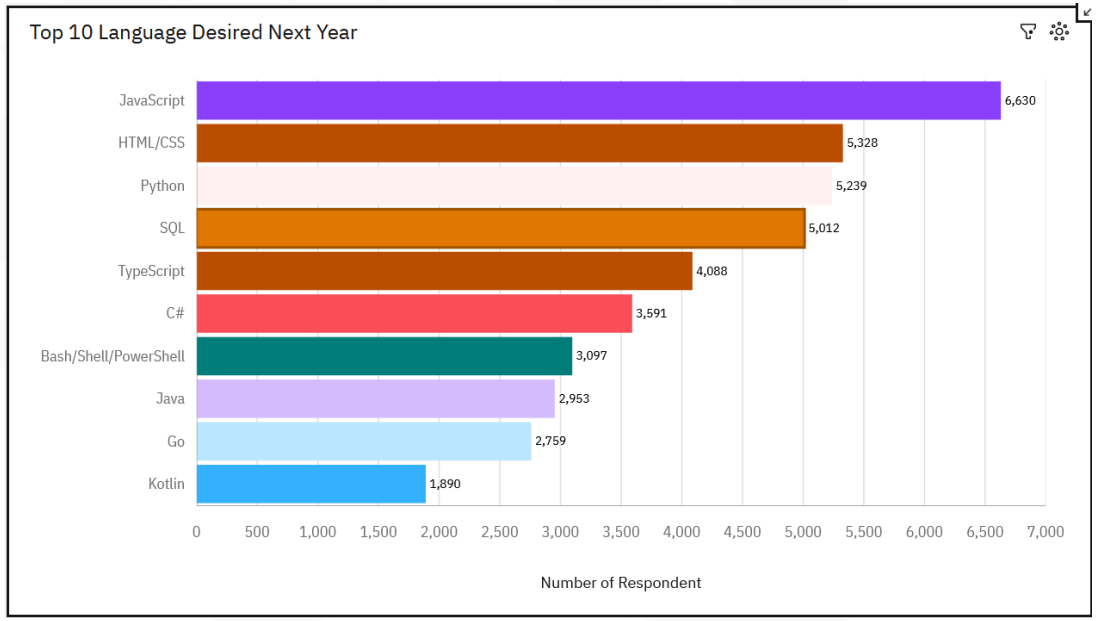
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

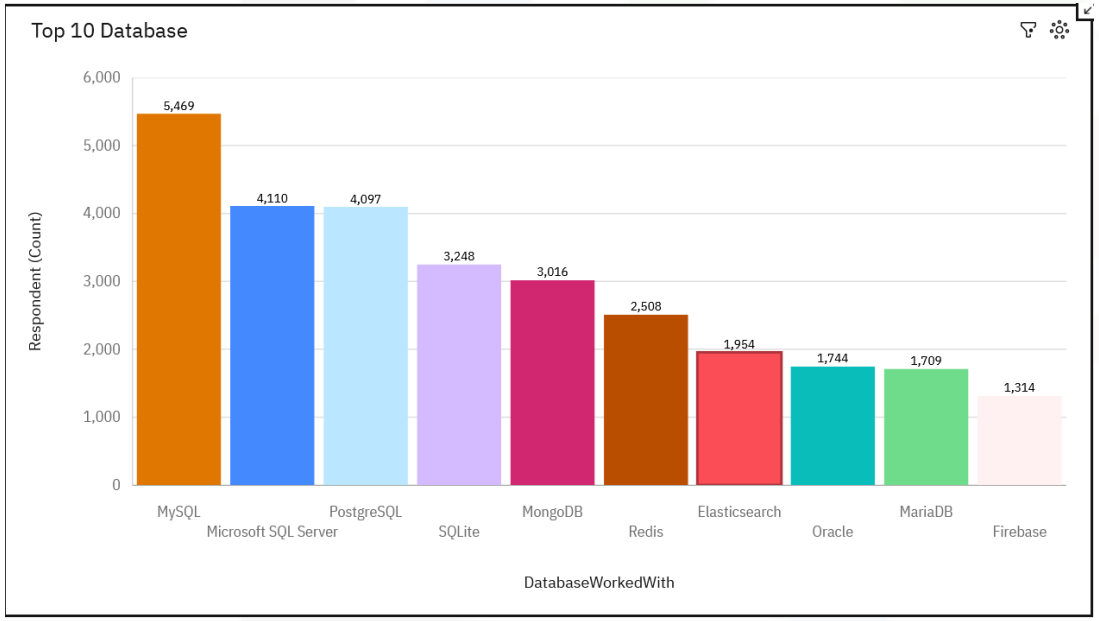
- JavaScript seems to keep as leading language.
- Python fastest-growing.
- Great interest in TypeScript.

Implications

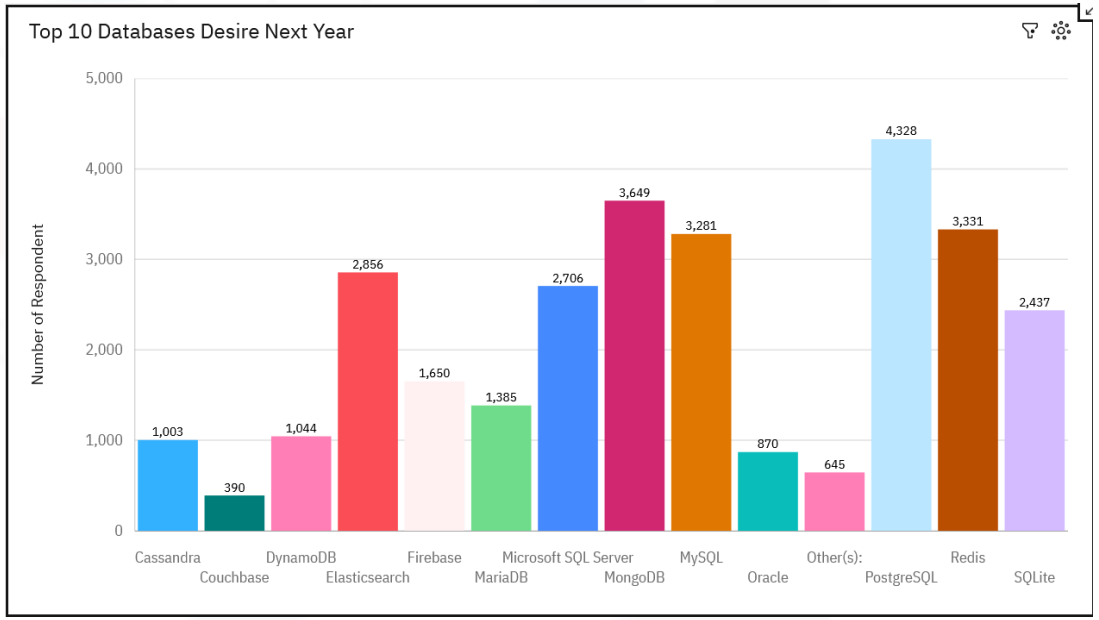
- Possible developers migration from JavaScript to TypeScriptD.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- MySQL as most used database.
- Lack of interest in Microsoft SQL Server and SQLite.
- Increasing interest in PostgreSQL and MongoDB.

Implications

- Microsoft SQL Server and SQLite losing ground in the market.
- PostgreSQL and MongoDB establishment in the market.

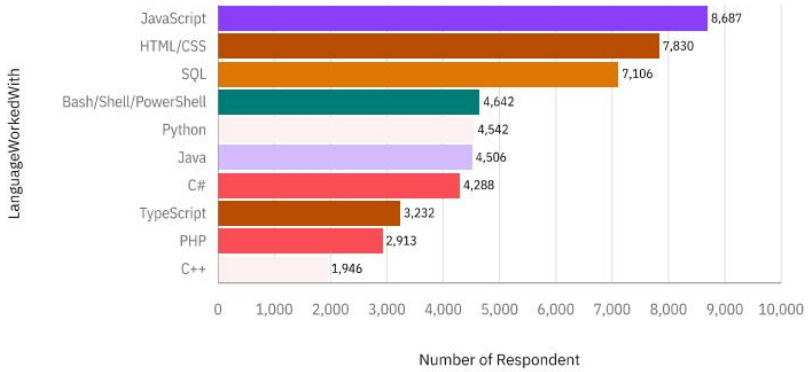
DASHBOARD

https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCapstone%2Bproject%2BPeer%2BAssignment%2BPart%2B1&action=view&mode=dashboard&subView=model0000018bb4817c70_00000000

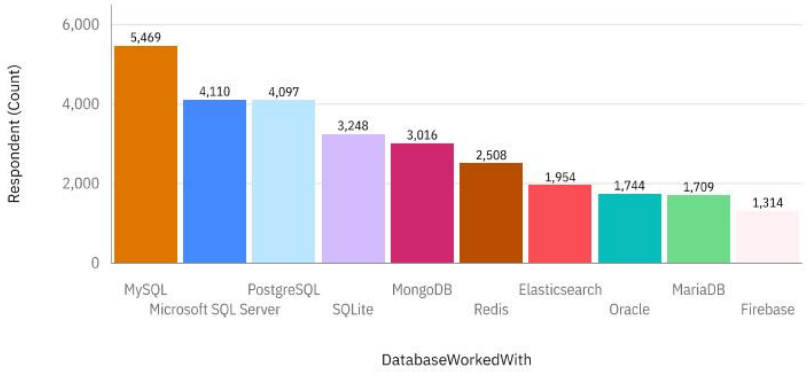
DASHBOARD TAB 1

Current Technology Usage

Top 10 Language

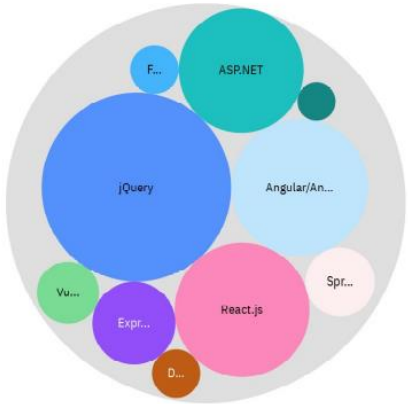


Top 10 Database



Platform Interacted

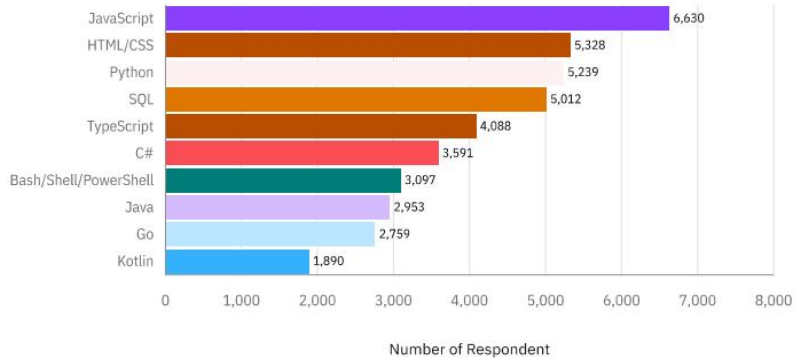
Top 10 Web Frame Works



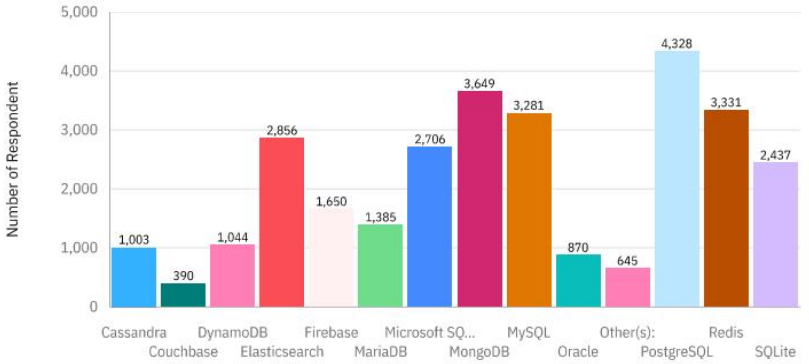
DASHBOARD TAB 2

Future Technology Trend

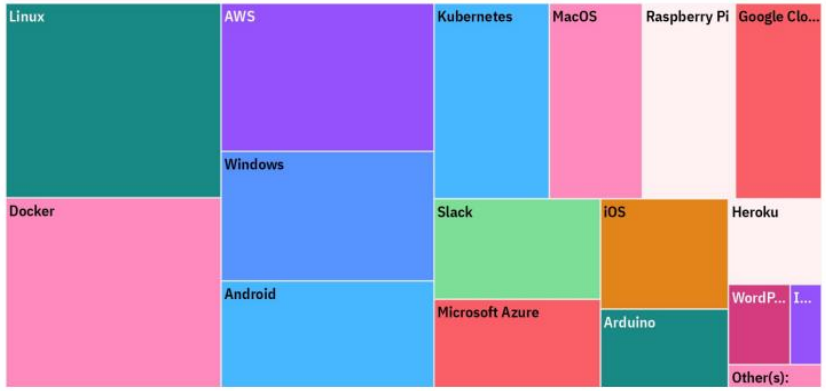
Top 10 Language Desired Next Year



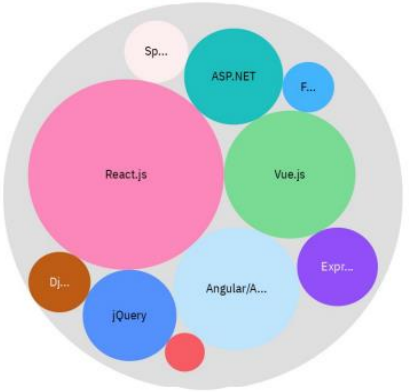
Top 10 Databases Desire Next Year



Platform Desired for Next Year



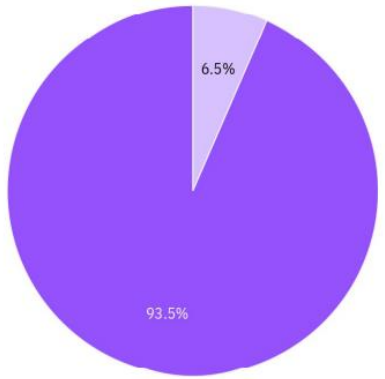
Top 10 Web Frame Works Desired Next Year



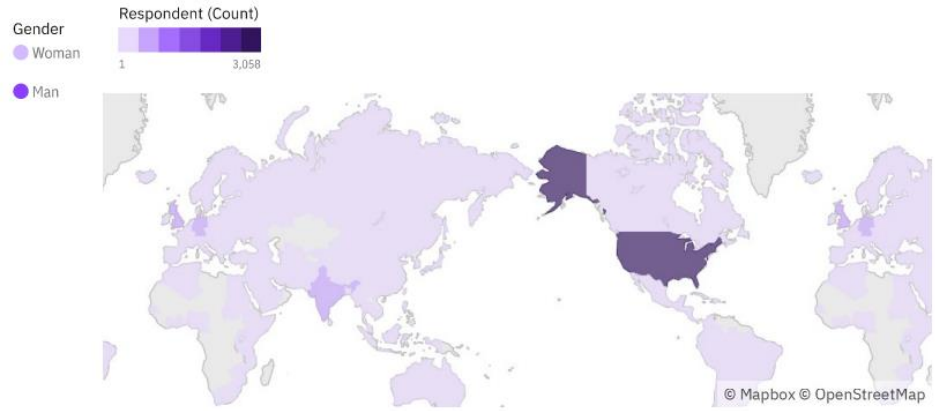
DASHBOARD TAB 3

Demographics

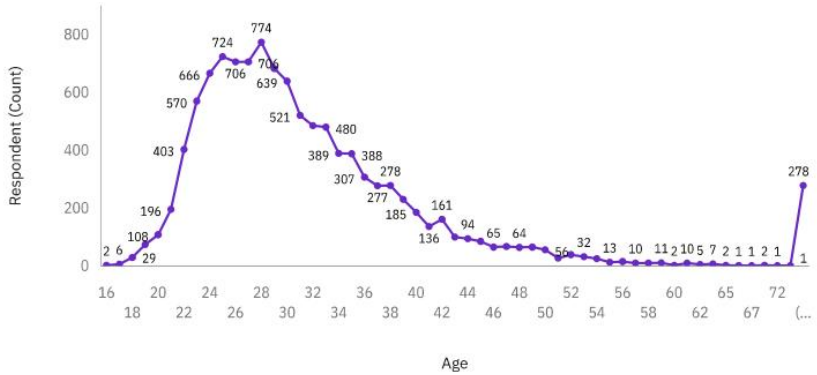
Respondent in % by Gender



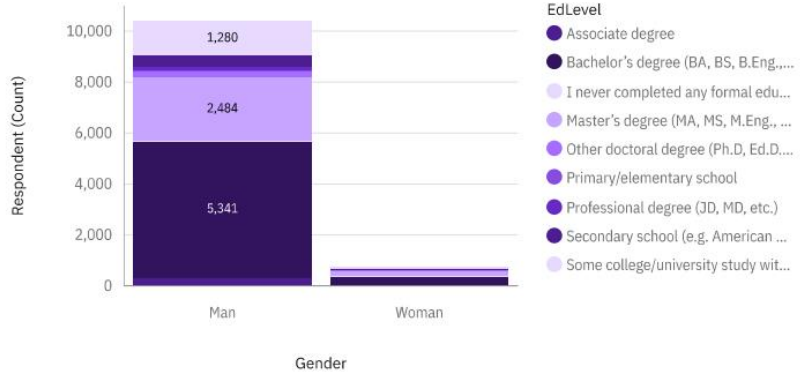
Respondent for Country regions



Respondent by Age



Respondent by Gender Classified by Education Level



DISCUSSION

OVERALL FINDINGS & IMPLICATIONS

Findings

- JavaScript widely used and TypeScript getting popular.
- Over 90% young male developers.
- Developers mostly located in developed countries.

Implications

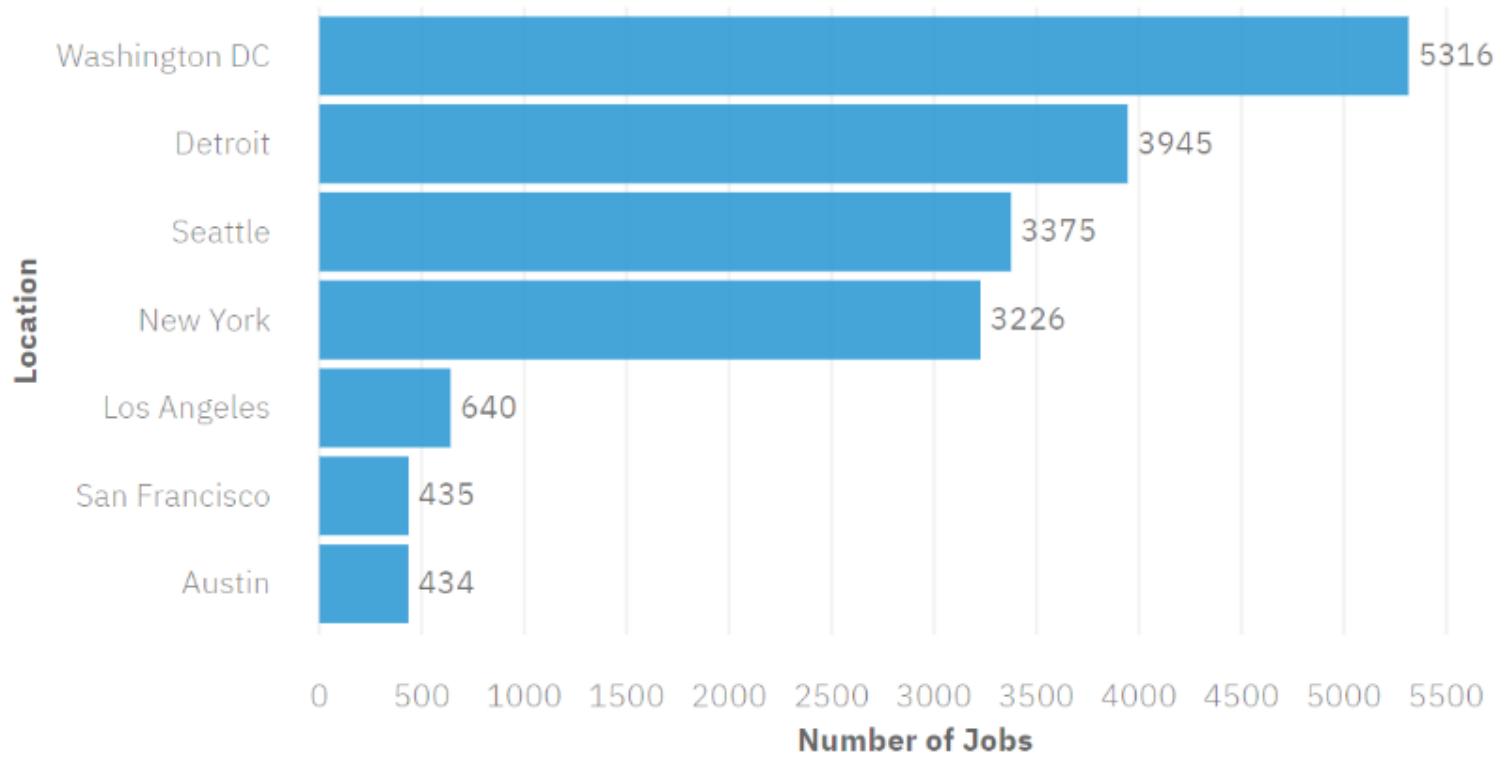
- JavaScript and TypeScript web frames gaining followers.
- Global polarization of developers location and gender.
- Young developers without postgrad studies on its majority.

CONCLUSION

- Developers are people with very marked characteristics.
- A good idea of popularity trends of different tools, platforms and languages can be obtained.
- There is a job to be done to spread accessibility of this labor market to countries in development.

APPENDIX

JOB POSTINGS



POPULAR LANGUAGES