Stack OverFlow Developer Analysis

An Exploratory Data Analysis of Global Developer Trends

Rodrigo Arriaza Rojas

August 1st, 2025

OUTLINE

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

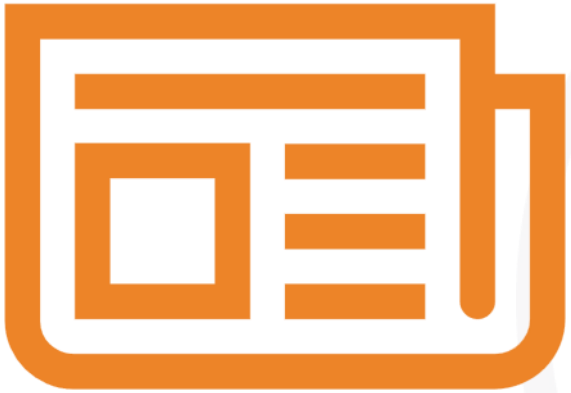
EXECUTIVE SUMMARY

- Conducted data analysis on programming languages and their average annual salaries.
- Performed web scraping and data wrangling to build a clean dataset.
  - Created job-postings.xlsx file using API calls.
  - Created popular-language.csv using Beautiful Soup.
  - Visualizations were derived from the aforementioned files.
- Visualized salary distribution using Python and data visualization libraries.
- Found that Swift and C++ lead the average salary.
- Highlighted insights can support strategic decisions for career and business planning.

INTRODUCTION

- Stack Overflow Developer Survey is one of the largest annual surveys of software developers worldwide.
- This project focuses on analyzing 2023 edition, with over 70,000 responses over multiple countries and roles.
- The goal of this project is to explore current and future trends in software development technologies.
- Key questions:
  - Which technologies are most popular across different age groups?
  - How does experience relate to salary?
  - Are there differences between the tools used by professional's vs learners?

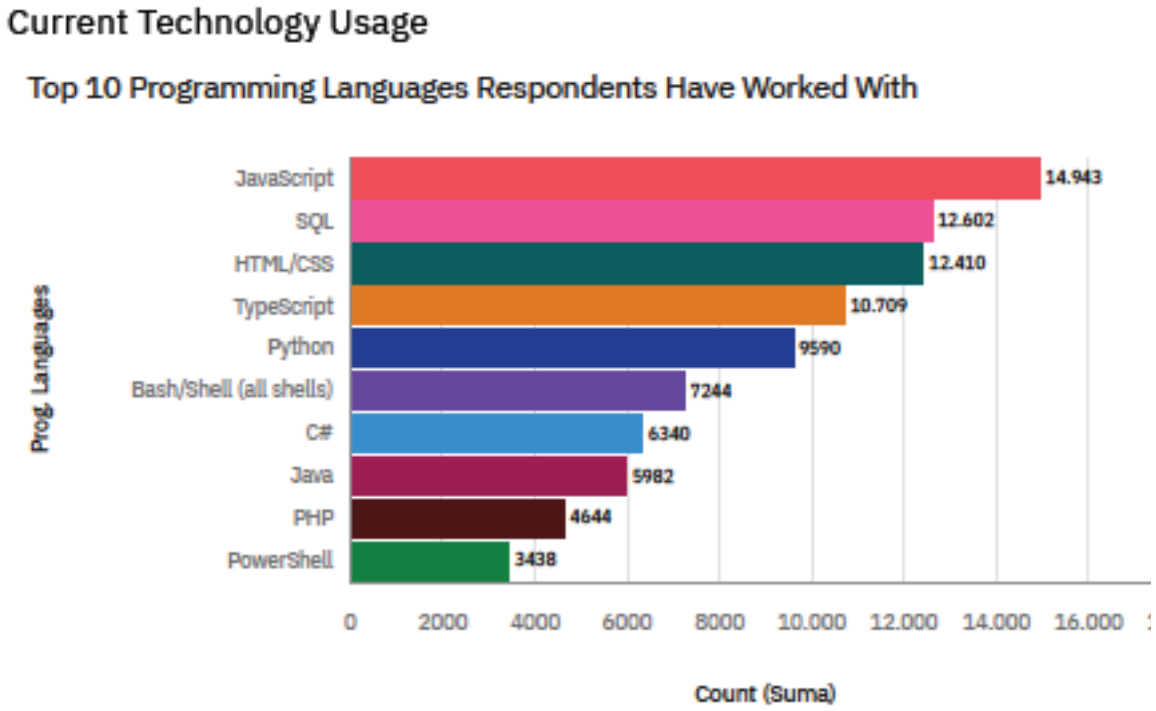
METHODOLOGY



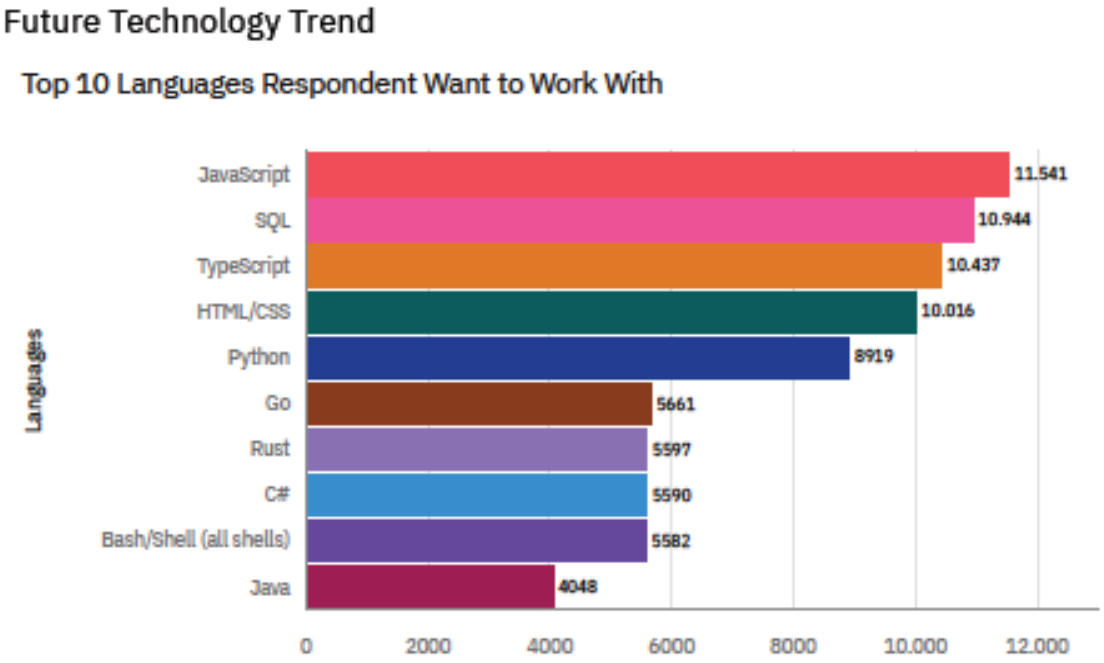
- Data Sources: The primary dataset used was Stack Overflow Developer Survey 2023 (~70,000 responses).
- Web Scraping: practiced in early labs to collect supplementary data.
- Data Processing: Cleaning, renaming columns, handling missing values, and filtering data were performed using Python.
- Exploratory Data Analysis: Visualizations were created using Seaborn and Matplotlib to explore trends in technology usage, experience, compensation, and demographics
- Visualization
  - Matplotlib and Seaborn to analyze data
  - Cognos dashboard to showcase current and future trends

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

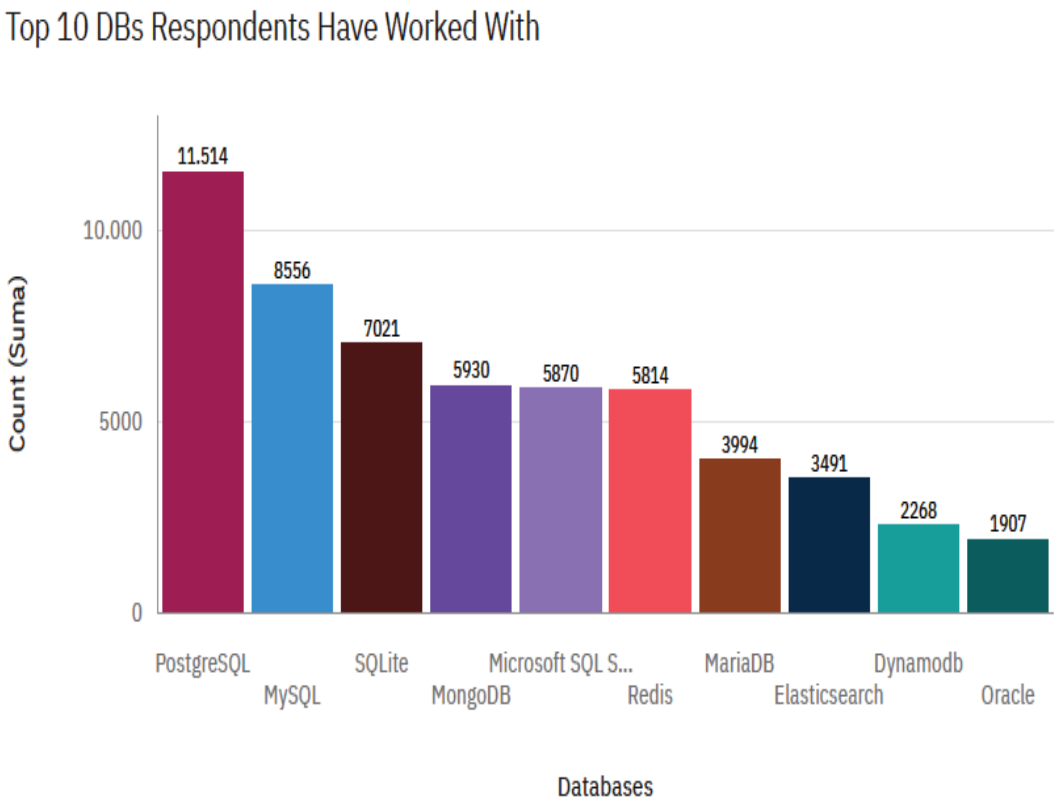
- JavaScript and SQL remain consistently relevant across current usage and future interest.
- Typescript is gaining popularity, especially in future trends
- Python is widely used currently but shows a slightl decline in future trends.

Implications

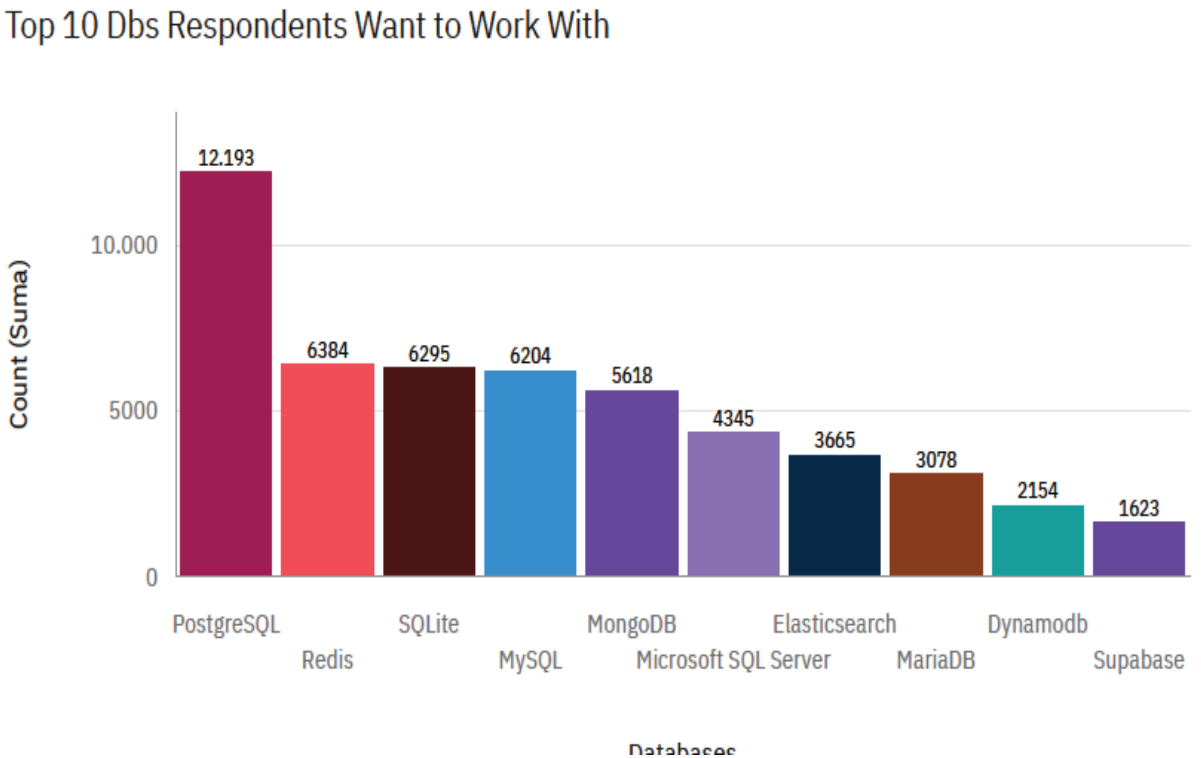
- Despite the upsurge of new technologies, JavaScript and SQL continues to be foundational skills for developers.
- TypeScript rising popularity suggest a shift towards strongly type JavaScript alternatives.
- Developers may need to diversify skills beyond Python to stay aligned to future demans.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- PostgreSQL ranks first in both current and future use.
- Redis is gaining popularity in developers for future use.
- Oracle ranks last in current use, and does not appear in the top 10 for future databases.

Implications

- PostgreSQL remains the preferred DB among developers and is expected to stay relevant in the future.
- For the growing demand for NoSQL and AI applications, Redis is becoming increasingly popular.
- Due to strong competition from modern cloud-based solutions, Oracle is losing relevance.

DASHBOARD

Dashboard Overview

This dashboard was created using IBM Cognos Analytics; it presents demographic insights from developer survey, including age distribution, education level, and geographic location of respondents. The following slides contain detailed visualizations extracted from the interactive dashboard.

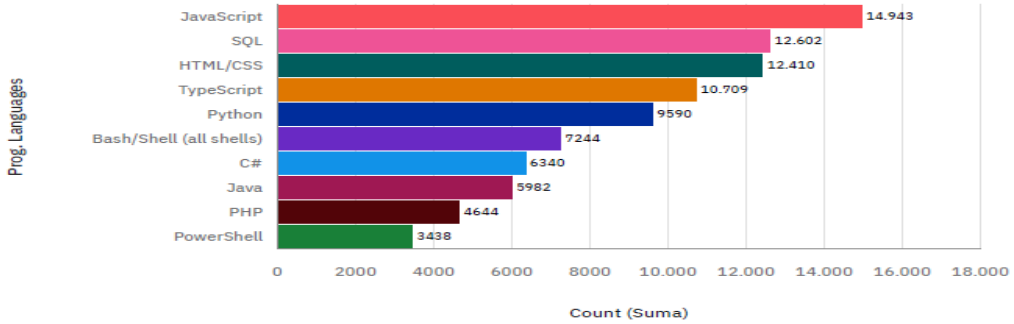
DASHBOARD TAB 1

2/8/25, 1:58

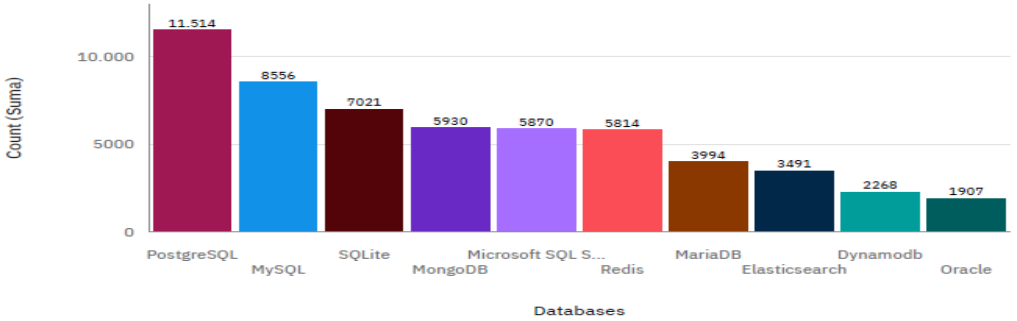
Cognos Dashboarad - Capstone Project

Current Technology Usage

Top 10 Programming Languages Respondents Have Worked With



Top 10 DBs Respondents Have Worked With



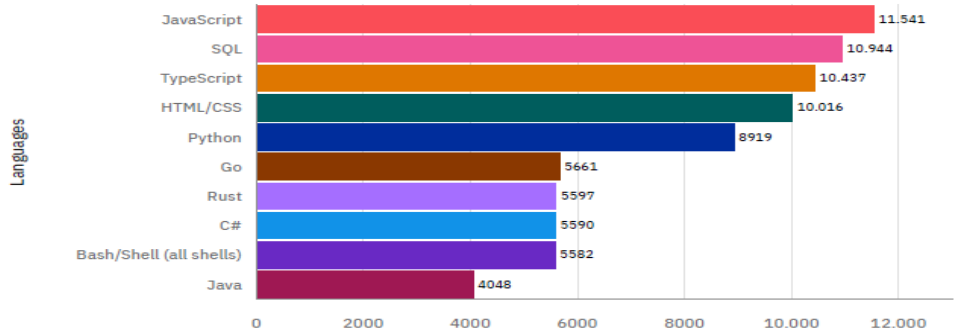
Top 10 Platforms Respondents Have Worked With

Top 10 Web Framworks Respondents Have Worked With

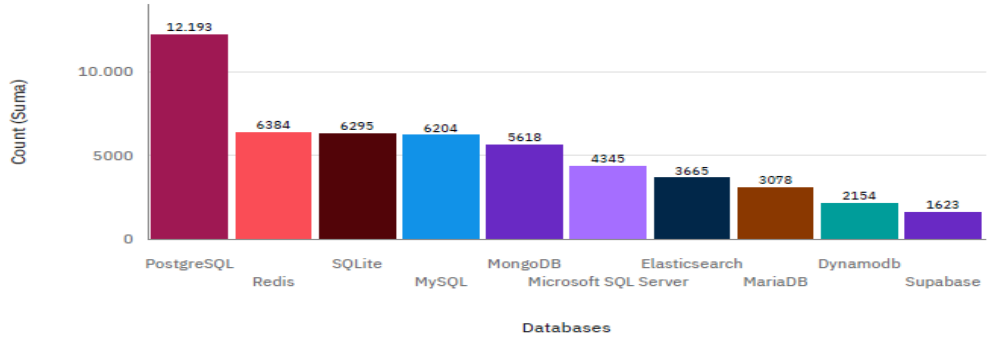
DASHBOARD TAB 2

Future Technology Trend

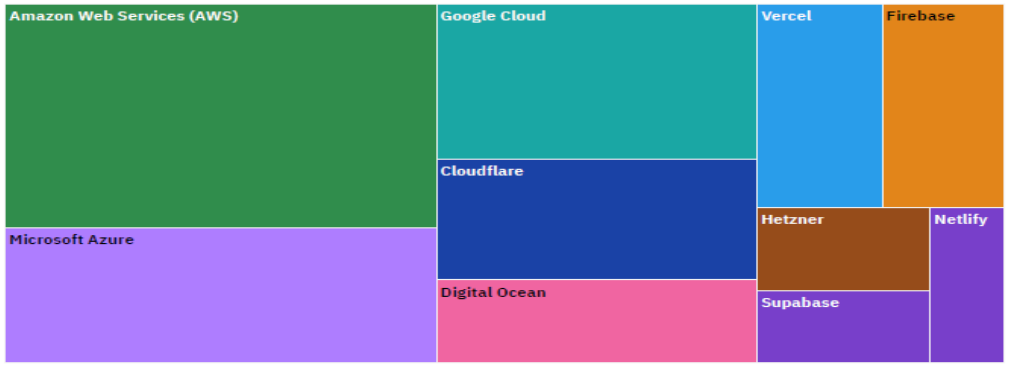
Top 10 Languages Respondent Want to Work With



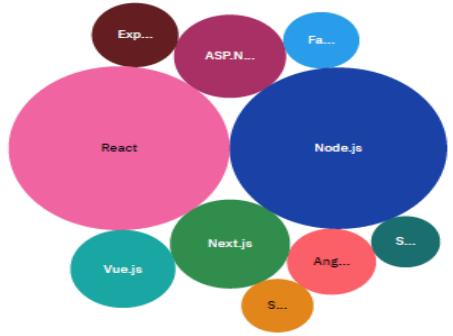
Top 10 Dbs Respondents Want to Work With



Tree Map of Top 10 Platforms Respondent Want to Work With



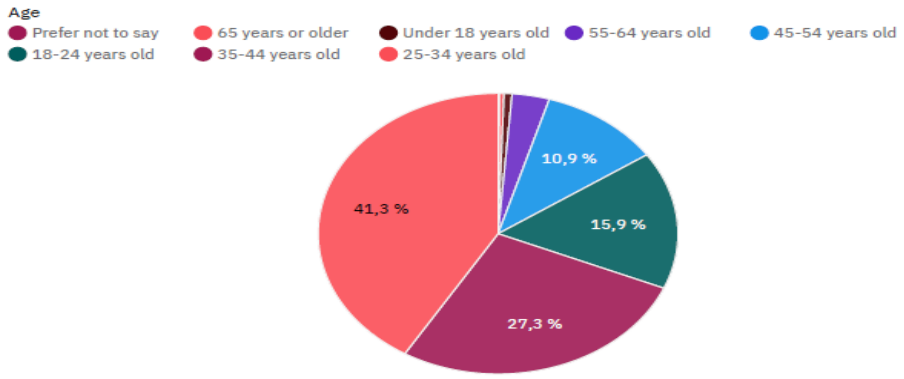
Top 10 Web Frameworks Respondent Want to Work With



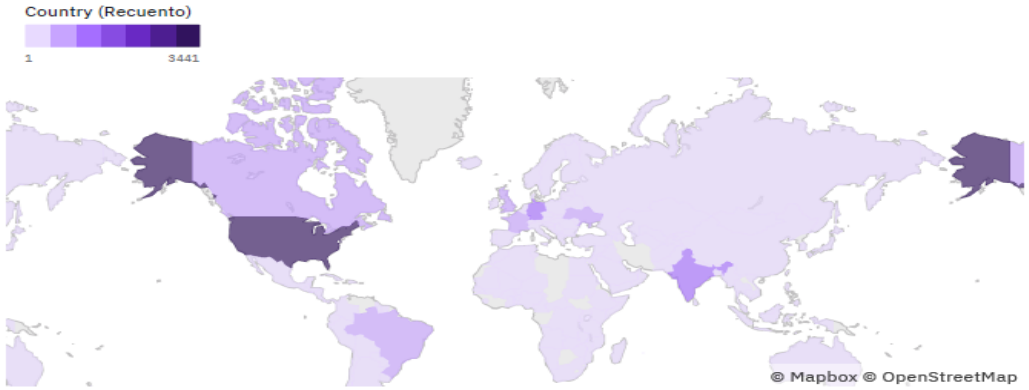
DASHBOARD TAB 3

Demographics

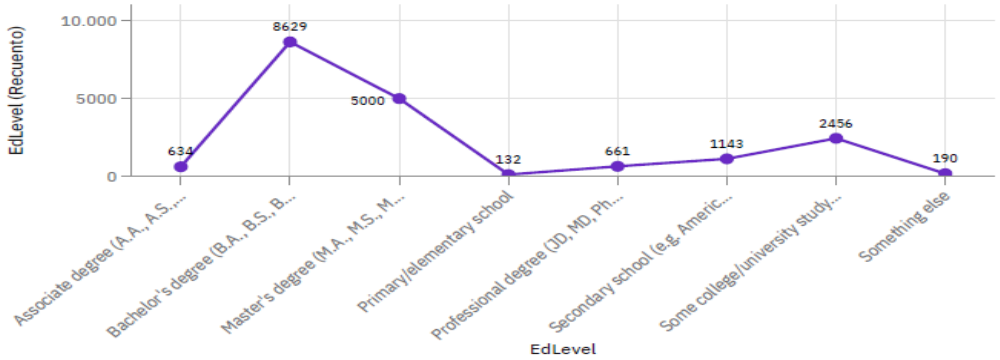
Respondents Age Distribution



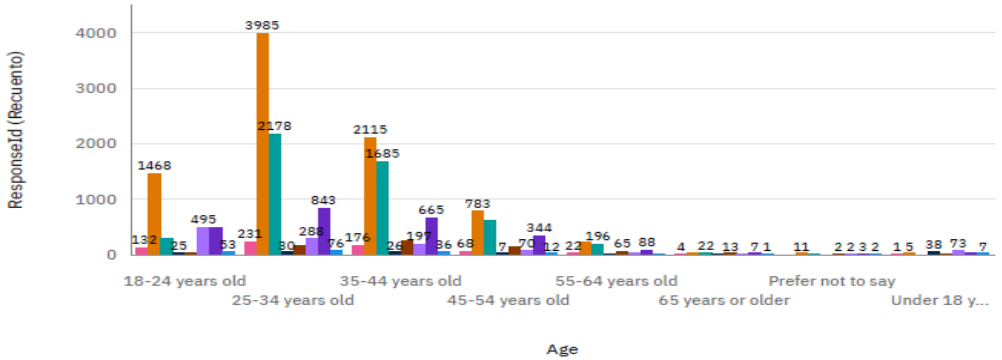
Number of Respondents by Country



Respondents Distribution by Education Level



Number of Respondents by Age and Ed Level



DISCUSSION

- The majority of the respondents are in the 25-34 years old age group (41.3%)
- United States is by far the country with the most respondents.
- AWS dominates as the most-used and most desired platform, followed by Azure and Google Cloud.

OVERALL FINDINGS & IMPLICATIONS

Findings

- The majority of the respondents are between 25-34 years old.
- United States has the highest number of respondents
- Most respondents hold Bachelor's degree or Master's degree, suggesting the sample is highly educated.

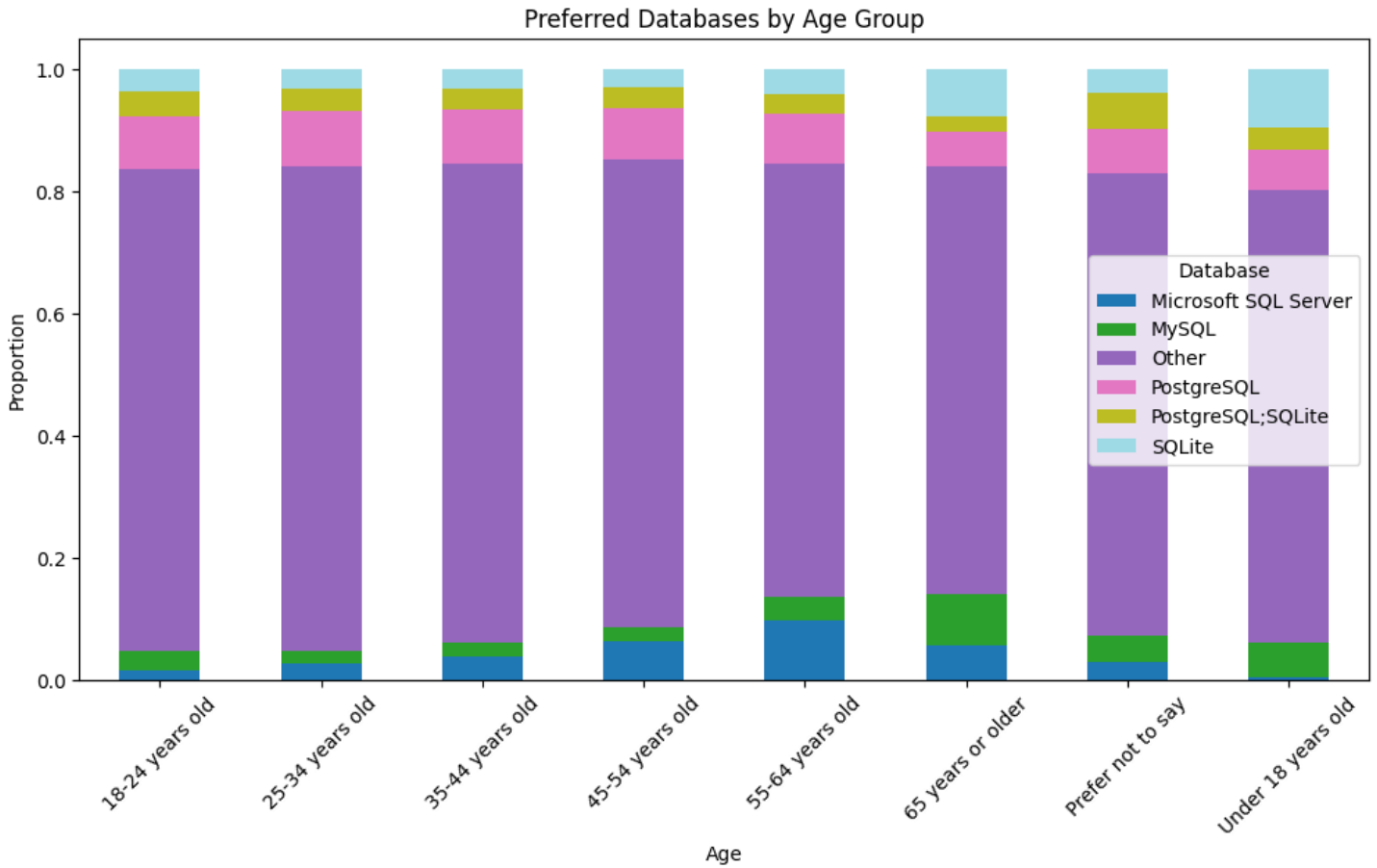
Implications

- The rising need for data and AI is creating more careers centered in data and software development.
- In the USA, data and development careers are well established and growing rapidly.
- As data, AI and development expand quickly, there is a rising need for highly educated professionals.

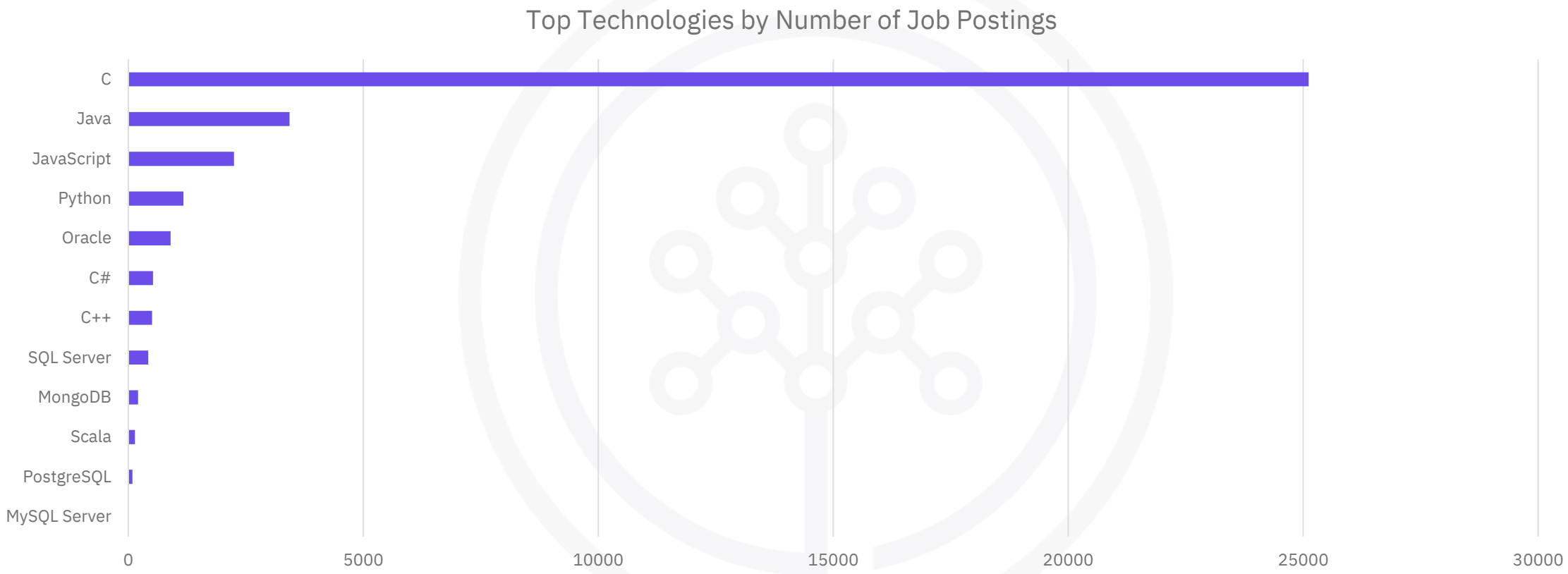
CONCLUSION

- JavaScript and SQL remain among the most relevant languages and for both current use future learning.
- Node.js and React are the leading web frameworks respondents currently use and that want to explore in the future.
- Respondents’ age distribution suggest a growing influx of new learners and professionals entering the tech industry.
- Cloud-based platforms are gaining popularity and reflect a shift towards scalable, modern infrastructure.

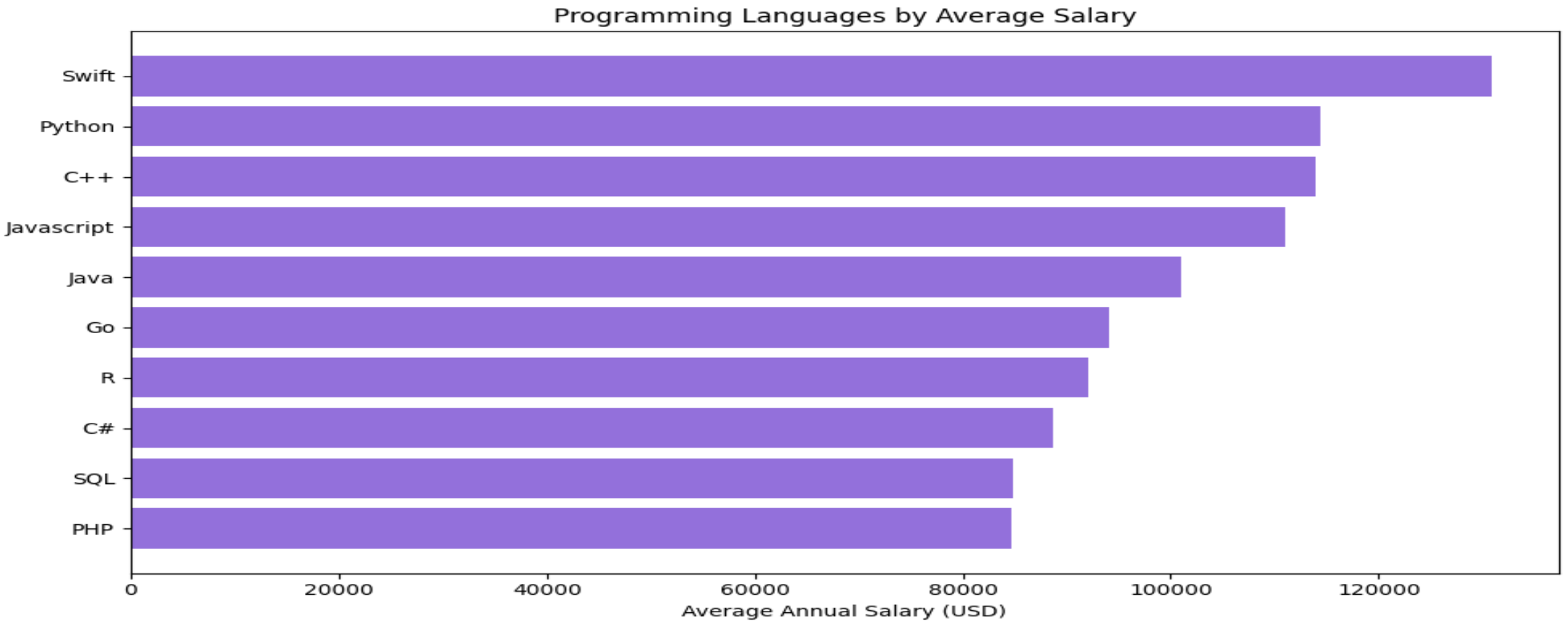
APPENDIX


Preferred Databases by Age Group
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
18-24 years old
25-34 years old
35-44 years old
45-54 years old
55-64 years old
65 years or older
Prefer not to say
Under 18 years old
Age
Database
Microsoft SQL Server
MySQL
Other
PostgreSQL
PostgreSQL;SQLite
SQLite

JOB POSTINGS



POPULAR LANGUAGES



Acknowledgments

I want to thank my mother and wife for their unwavering support during my transition from education to the field of data analytics and data science.
I also thank Kari, my loyal German Shepherd, who stays by my side at night while I learn and code.
Finally, my gratitude to IBM for making these courses accessible to learners around the world.