Stackoverflow Developer Survey Data Analysis

Duong Van Nhat Long

Date: 04/03/2024

IBM Developer

SKILLS NETWORK

OUTLINE

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

EXECUTIVE SUMMARY

Current Technology Usage

- Top 10 Languages Worked with
- Top 10 Database Worked with
- Platform worked with Word-Cloud
- Top 10 Web-frame Worked with

Future Technology Trend

- Top 10 Languages Desired Next Year
- Top 10 Database Desired Next Year
- Platform worked with Desired Next Year
- Top 10 Web-frame Desired Next Year

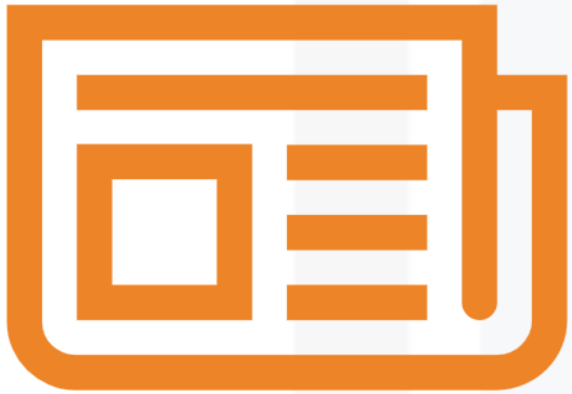
Demographics

- Respondents Classified by Gender
- Respondent Count for Countries
- Respondent Count by Age
- Respondent Count by Gender by Formal Education Level

INTRODUCTION

- This presentation is designed for stakeholders and business decision-makers within a global IT and business consulting services firm.
- Its purpose is to pinpoint forthcoming skill prerequisites within the global IT sector essential for the company to stay abreast of evolving technologies and sustain competitiveness.
- Recommendations will be outlined based on the analysis conducted.

METHODOLOGY

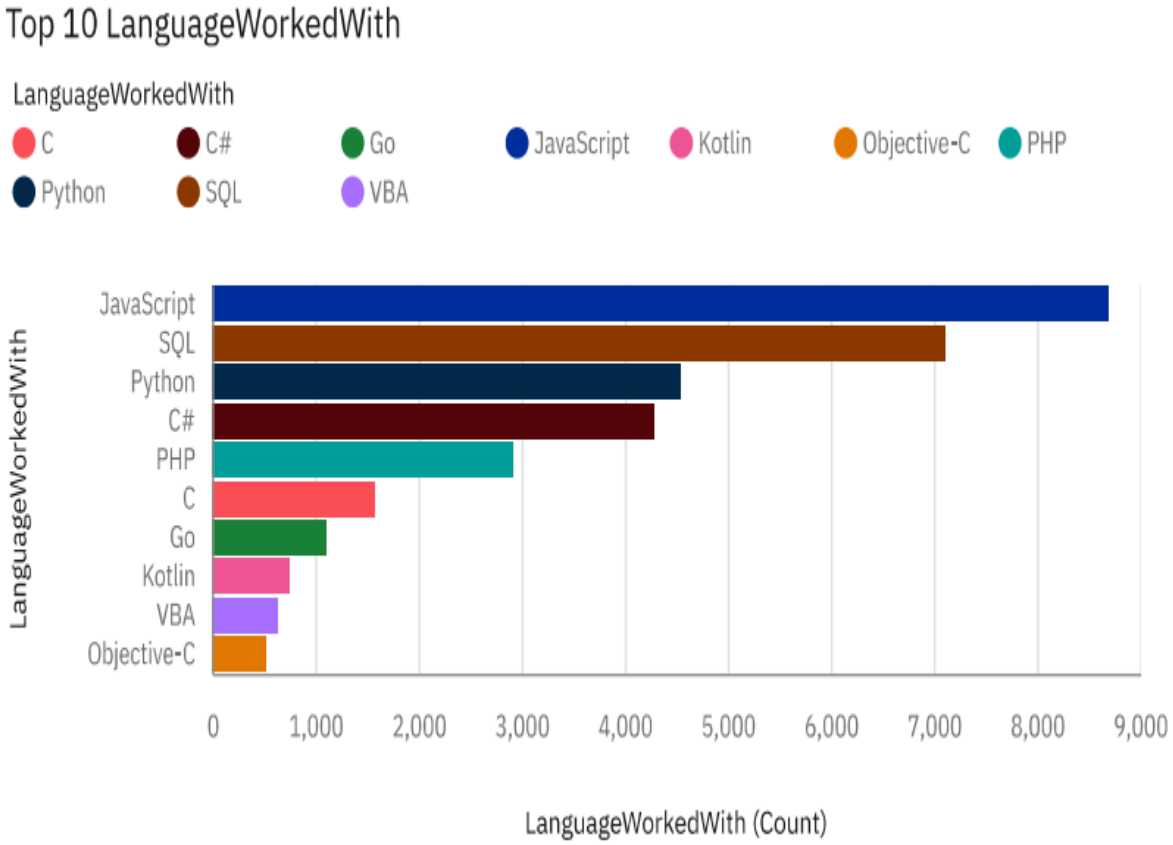


- **Collecting Data:** Utilizing the Stack Overflow Developer Survey API, Employing web scraping techniques to gather data.
- **Exploring Data**: Investigating the collected data to understand its structure and content.
- **Finding Duplicates & Removing Duplicates:** Identifying duplicate entries within the dataset and eliminating them to ensure data integrity.
- **Finding & Imputing Missing Values**: Detecting missing values in the dataset and implementing strategies to fill in or impute these missing values.
- **Normalizing Data**: Standardizing the data to a common scale to facilitate comparisons and analyses.
- **Handling Outliers:** Identifying outliers within the dataset and employing appropriate methods to handle them, such as removing or transforming them.
- **Correlation**: Analyzing the relationships between different variables in the dataset to understand their correlation and dependencies.
- **Visualization Data:** Creating visual representations of the data to gain insights and communicate findings effectively.
- **Building Dashboard:** Developing a dashboard interface to present the analyzed data in an interactive and user-friendly manner for further exploration and decision-making.

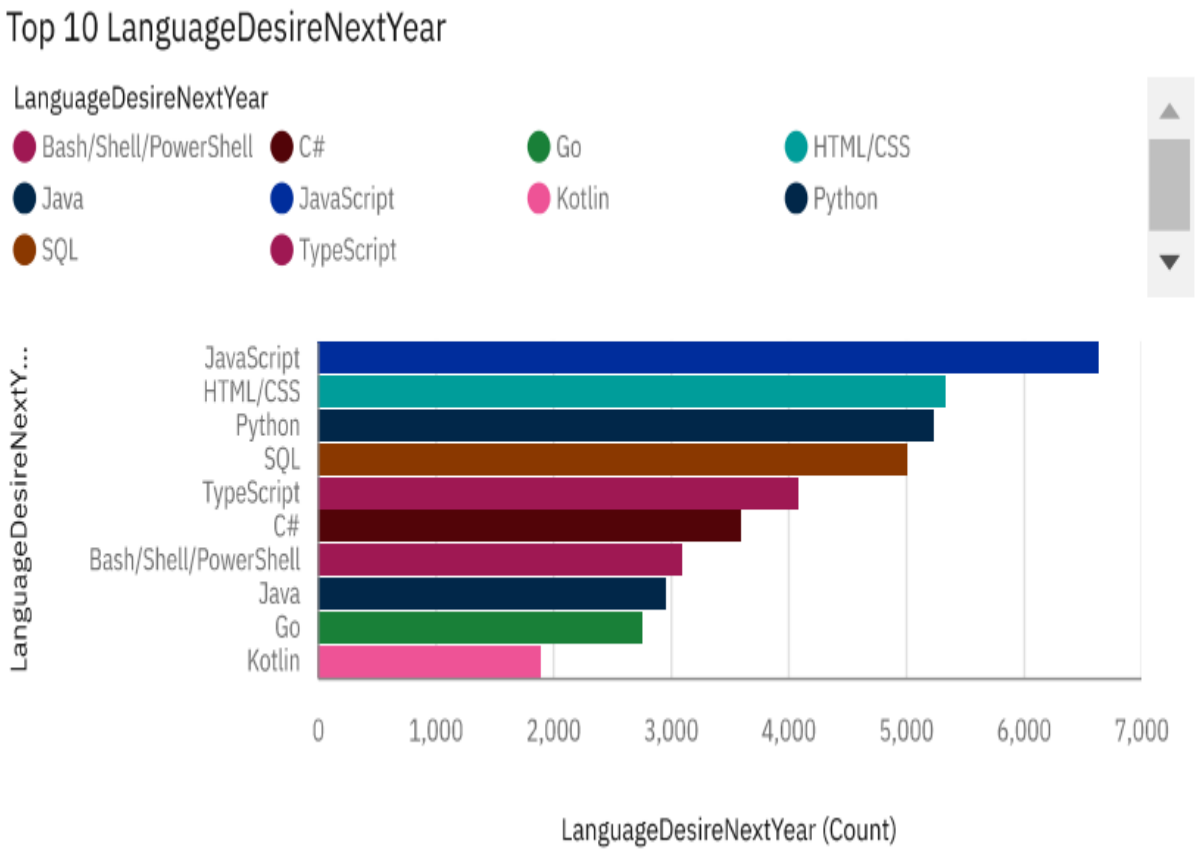
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

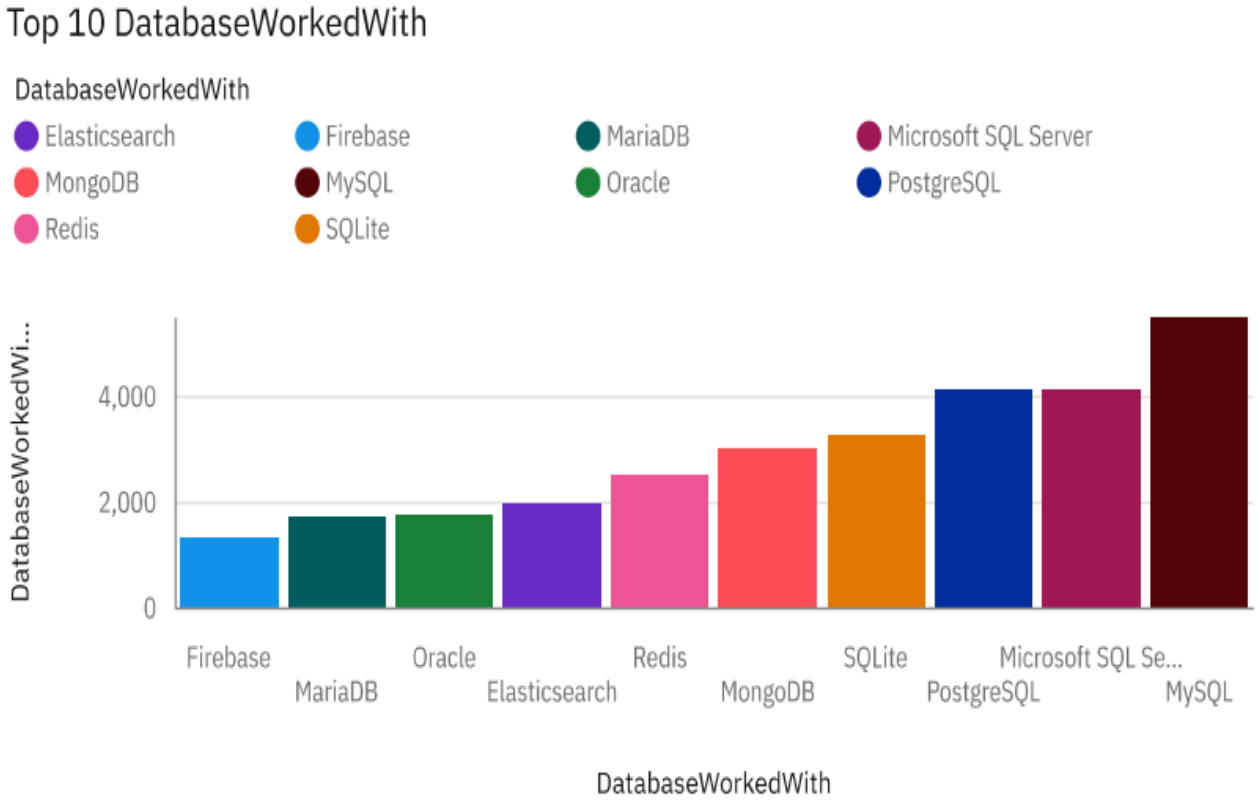
- JavaScript and SQL retain their positions as the top two most popular programming languages for the current and JavaScript and HTML/CSS upcoming year.
- Python and TypeScript are experiencing increased interest for the following year.
- Conversely, interest in SQL and Bash/Shell/PowerShell has declined.

Implications

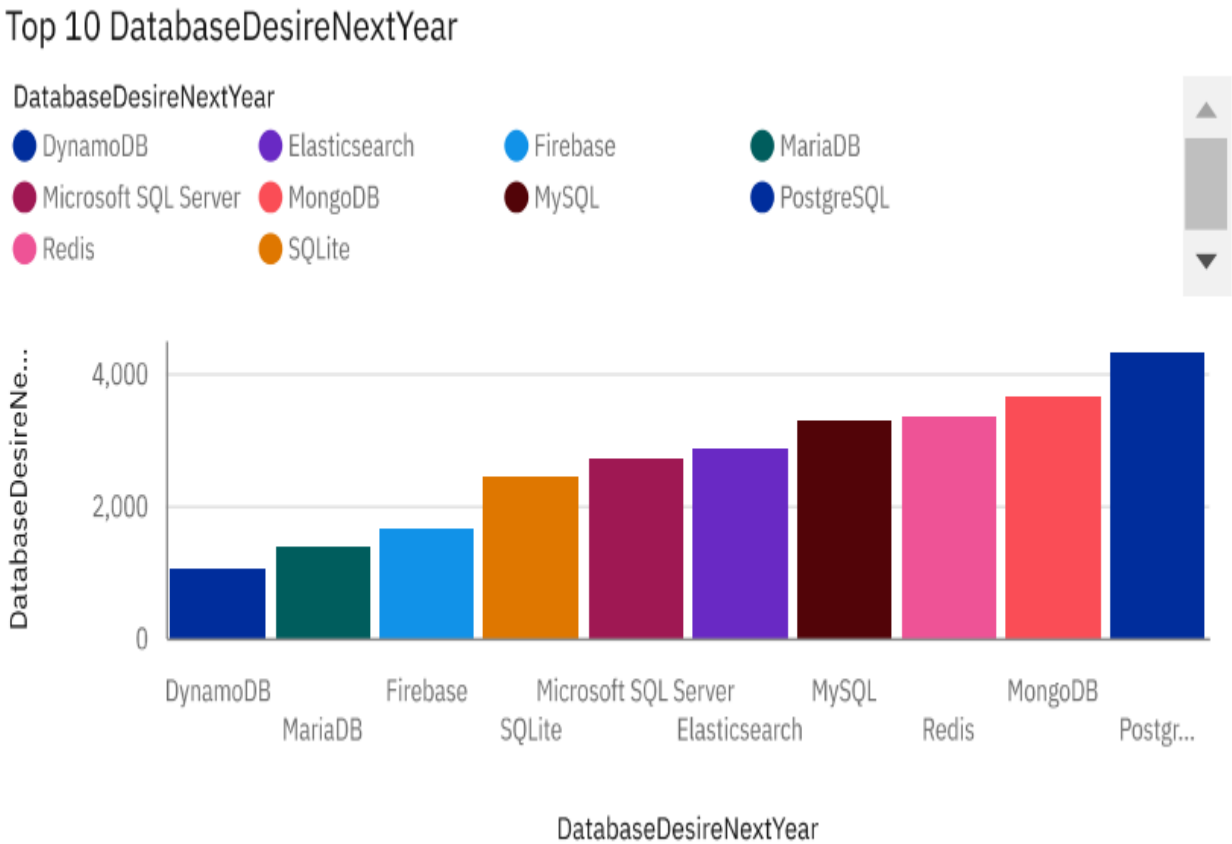
- You can maintain a similar number of employees skilled in JavaScript and HTML/CSS.
- You can increase the number of employees skilled in Python and TypeScript.
- You can decrease the number of employees skilled in SQL and Bash/Shell/PowerShell.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Interest in MySQL, Microsoft SQL Server, and SQLite has decreased for the upcoming year.
- Interest in PostgreSQL and MongoDB has increased compared to the current year.
- There is increased interest in Redis and Elasticsearch for the following year.

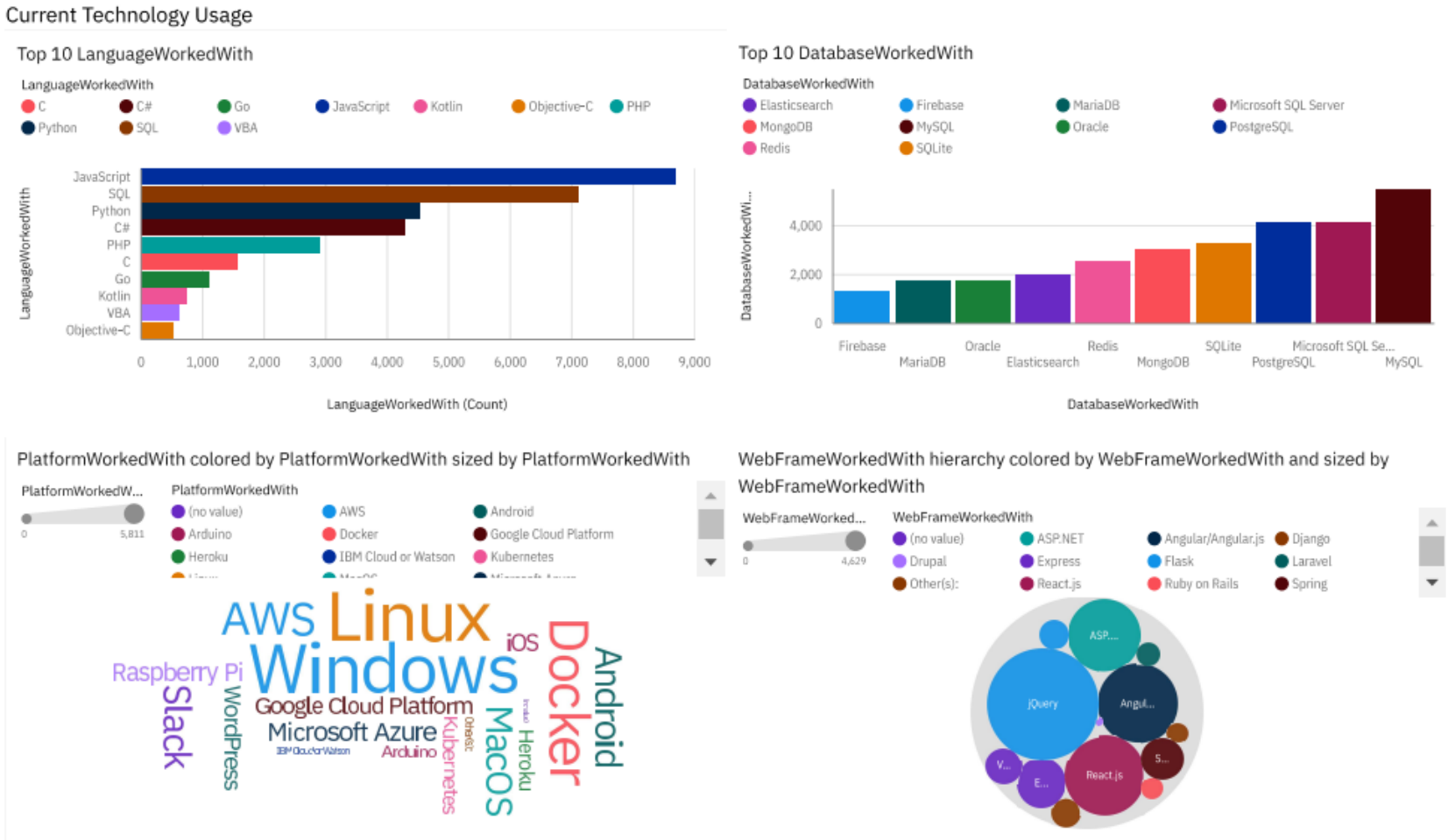
Implications

- You can decrease the number of employees skilled in MySQL, Microsoft SQL Server, and SQLite.
- You can increase the number of employees skilled in PostgreSQL and MongoDB.
- You can also employ more people skilled in Redis and Elasticsearch.

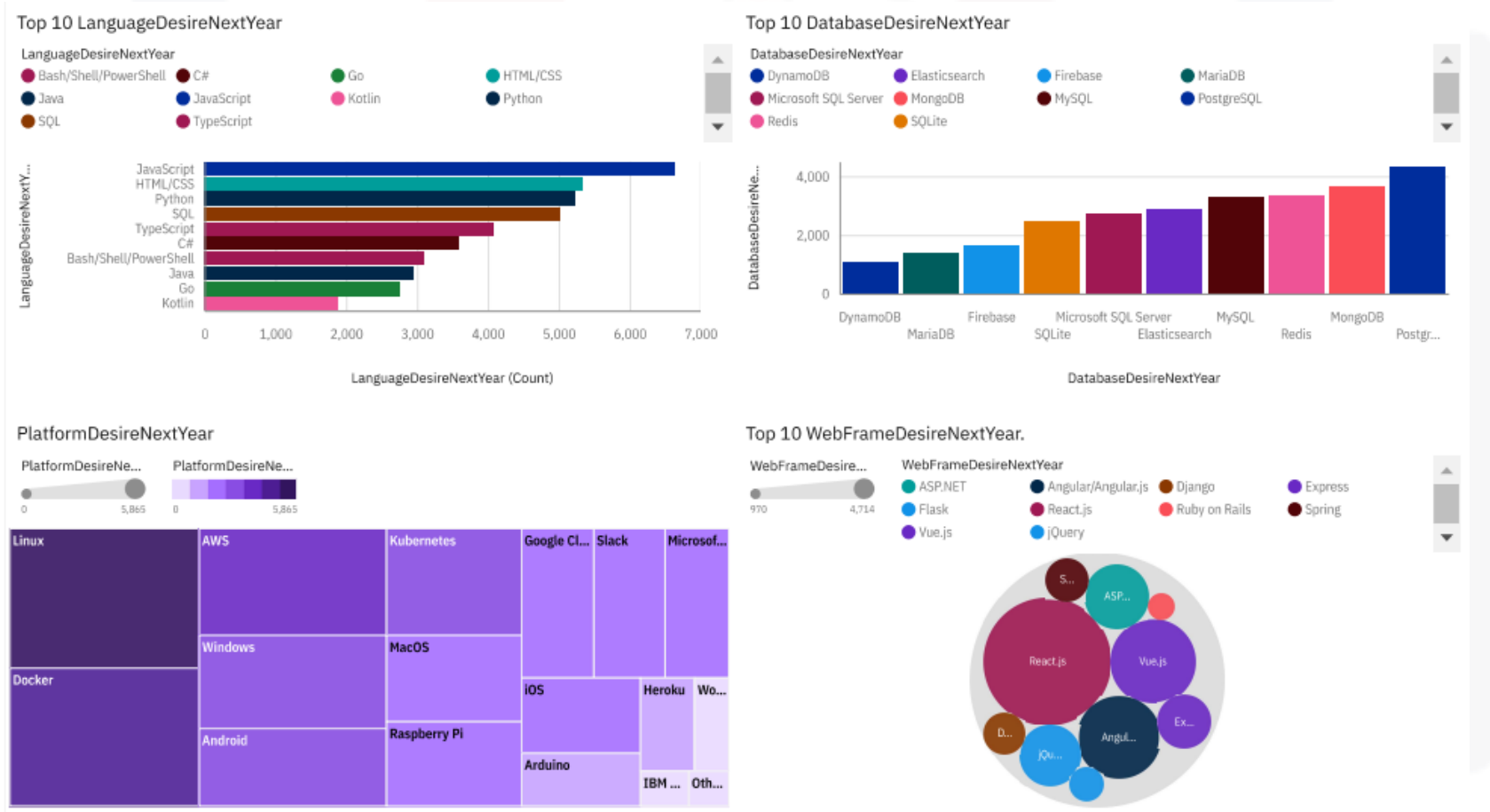
DASHBOARD

https://ap1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FIBM%2BData%2BAnalyst%2BCapstone%2BProject_WEEK5&action=view&mode=dashboard

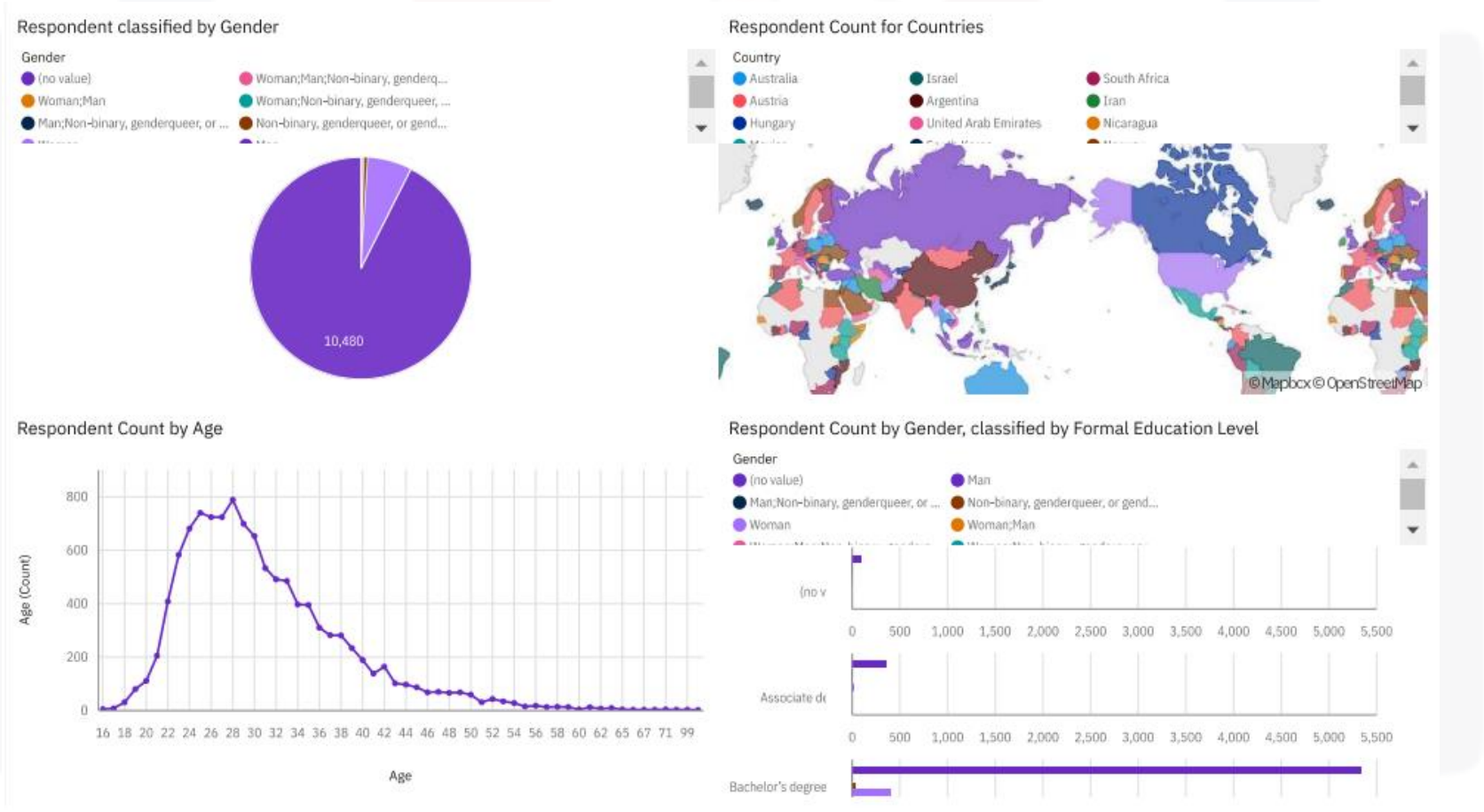
DASHBOARD TAB 1



DASHBOARD TAB 2



DASHBOARD TAB 3



DISCUSSION

- The survey included respondents of all ages.
- All genders were considered in the survey.
- Participants from various countries were asked to participate in the survey.
- Individuals with different education levels were also included in the survey.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Over 90 percent of the respondents in the survey are male.
- The average age of the participants is 30.
- There is a notable increase in interest in TypeScript, alongside continued growth in Python.
- Redis, Elasticsearch, PostgreSQL, and MongoDB are experiencing heightened interest.
- Interest in Slack and Windows is notably declining.
- Vue.js is seeing a significant rise in interest, while React.js continues to grow steadily.

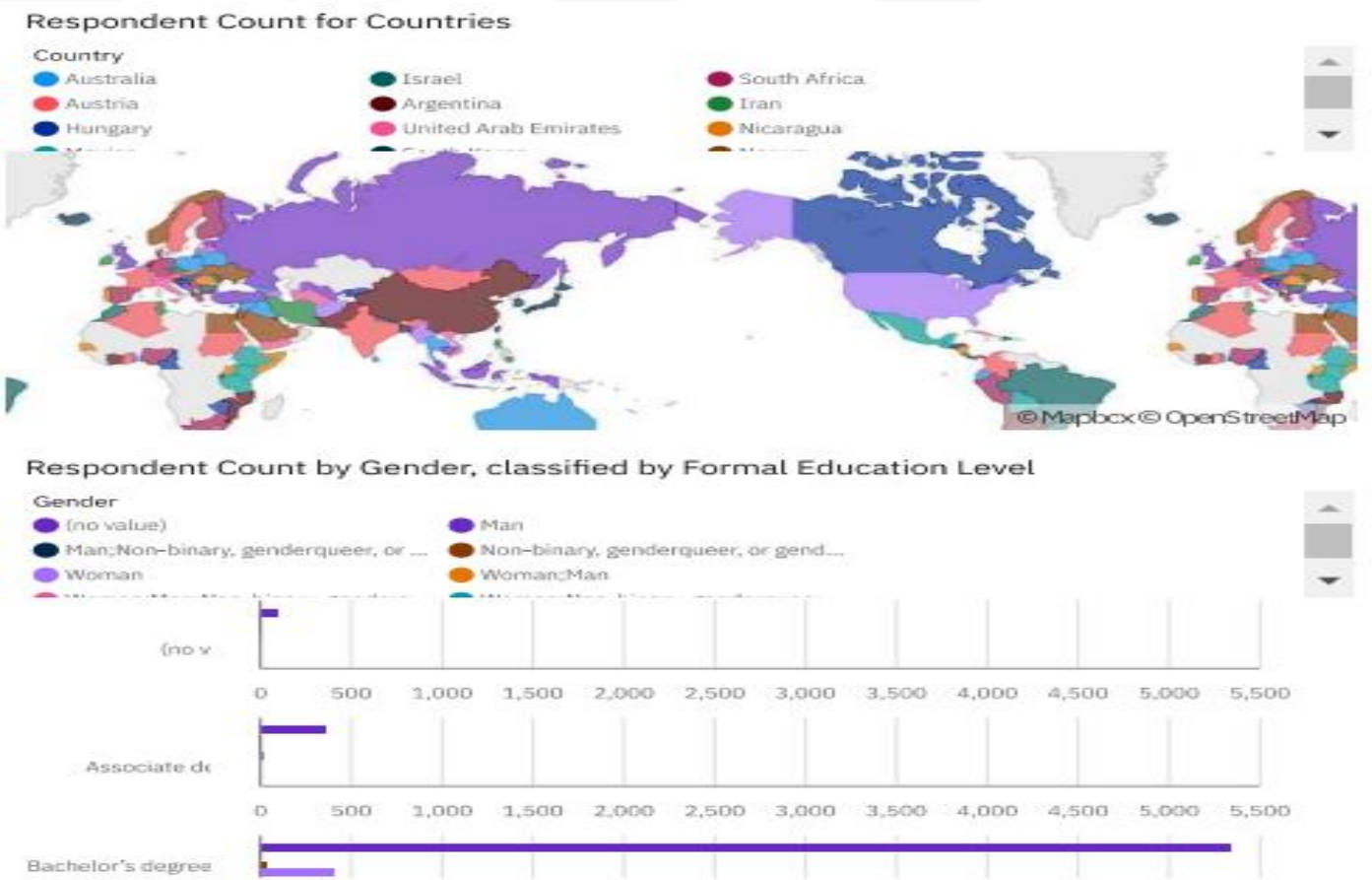
Implications

- The software sector continues to exhibit a male-dominated landscape.
- Locating highly experienced professionals in development remains a persistent challenge.
- It's advisable to maintain a stable workforce proficient in JavaScript and HTML/CSS while increasing recruitment for Python and TypeScript expertise.
- Furthermore, consider bolstering your team with more talent specialized in PostgreSQL, MongoDB, Elasticsearch, and Redis.
- In the realm of web development, ensure sufficient staffing for ASP.NET while augmenting the team with skilled Vue.js and React.js practitioners.
- In terms of infrastructure, sustain your Linux workforce while expanding the team with Docker, AWS, and Android experts, while scaling back reliance on Slack and Windows.

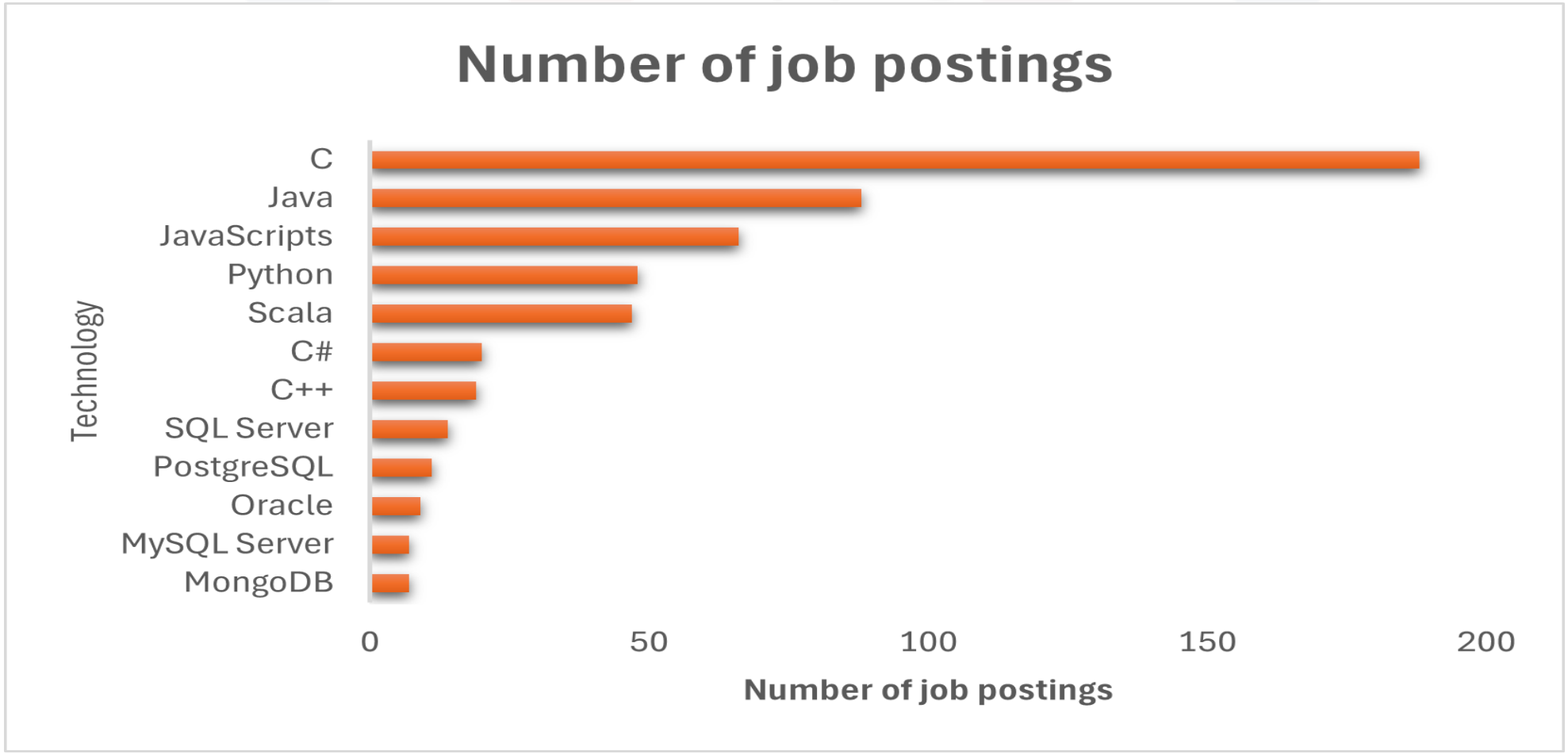
CONCLUSION

- Up to now, the software industry is still largely influenced by male dominance. Finding highly experienced development professionals remains a challenge for many organizations.
- To meet the growing market demand, maintaining a solid team in JavaScript and HTML/CSS is necessary, but there is also a need to increase recruitment of people with skills in Python and TypeScript. This reflects the upward trend of TypeScript and the stability of Python in the industry.
- At the same time, to meet the needs of modern software projects, expanding the team with people skilled in PostgreSQL, MongoDB, Elasticsearch and Redis is necessary. This represents a shift from traditional DBMSs to more popular and high-performance solutions such as MongoDB and Elasticsearch.

APPENDIX


Respondent Count for Countries
Country
Australia
Austria
Hungary
Israel
Argentina
United Arab Emirates
South Africa
Iran
Nicaragua
©Mapbox © OpenStreetMap
Respondent Count by Gender, classified by Formal Education Level
Gender
(no value)
Man;Non-binary, genderqueer, or ...
Woman
Man
Non-binary, genderqueer, or gend...
Woman;Man
(no v
Associate de
Bachelor's degree
0
500
1,000
1,500
2,000
2,500
3,000
3,500
4,000
4,500
5,000
5,500

JOB POSTINGS



POPULAR LANGUAGES