Stack Overflow Developer Survey 2024 IBM Data Analyst Capstone Project

By Ryan Welsh

21/01/2025

OUTLINE

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

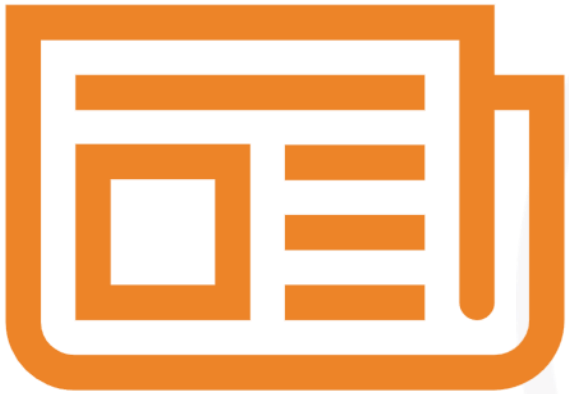
EXECUTIVE SUMMARY

- Introduction
- Methodology
  - **Data Collection**: Downloading and Importing Survey results from the 2024 Stack Overflow Developer Survey.
  - **Data Cleaning and Wrangling**: The collected data was cleaned and processed to ensure accuracy and consistency.
  - **Data Analysis**: The cleaned data was analyzed using statistical and analytical techniques to derive meaningful insights.
- Results
- Discussion
- Conclusion

INTRODUCTION

- **Objective**: Provide a comprehensive analysis of the current trends in programming languages and databases, and their implications for new entrants.
- **Sources**: Stack Overflow Developer Survey, IBM Data Analyst Capstone Project
- **Methodology**: Data collection, Data analysis
- **Content**
  - Present results through visualizations
  - Discuss key findings and their implications
- Goal: By the end, you'll have a clear understanding of the current landscape and future trends in the tech industry.

METHODOLOGY



- Downloading and Importing The Dataset from the Stack Overflow Developer Survey
- As the Data was collated by a third party source I determined I needed to perform Data Cleaning and Wrangling Techniques before performing analysis.
- Data Cleaning – Reviewing the dataset to see the datatypes included, and if the data within was sufficient for my analysis. I then checked for duplicate and null values, there was no duplicates but was null values, these missing values were handled by replacing the null values with the average (mode,median) of the particular column.
- Data Wrangling – Splitting columns to parse Json string so that the results from the survey could be collated correctly prior to analysis, mapping certain parameters in the dataset to categorical values so that I could group respondents into categories.
- Explanatory Data Analysis – Looking for correlations and dependencies between variables. Using cross-tabulations, correlation matrices, and other techniques to explore relationships. Creating visualisations such as histograms, box plots, scatter plots, and bar charts to showcase these correlations.
- Creating and Presenting Findings Via Dashboards
  - IBM Cognos Analytics
  - Google Looker Studio

RESULTS

Top Programming Languages

- **JavaScript**: 2,421 responses
- **HTML/CSS**: 2,037 responses
- **SQL**: 2,018 responses

Top Databases

- **PostgreSQL**: 3,125 responses
- **MySQL**: 2,289 responses
- **SQLite**: 1,956 responses

Top Platforms

- **React**
- **jQuery**

Desired Programming Languages for Next Year

- **JavaScript**: 2,035 responses
- **SQL**: 1,964 responses
- **TypeScript**: 1,863 responses

Desired Databases for Next Year

- **PostgreSQL**: 3,487 responses
- **SQLite**: 1,859 responses
- **Redis**: 1,843 responses

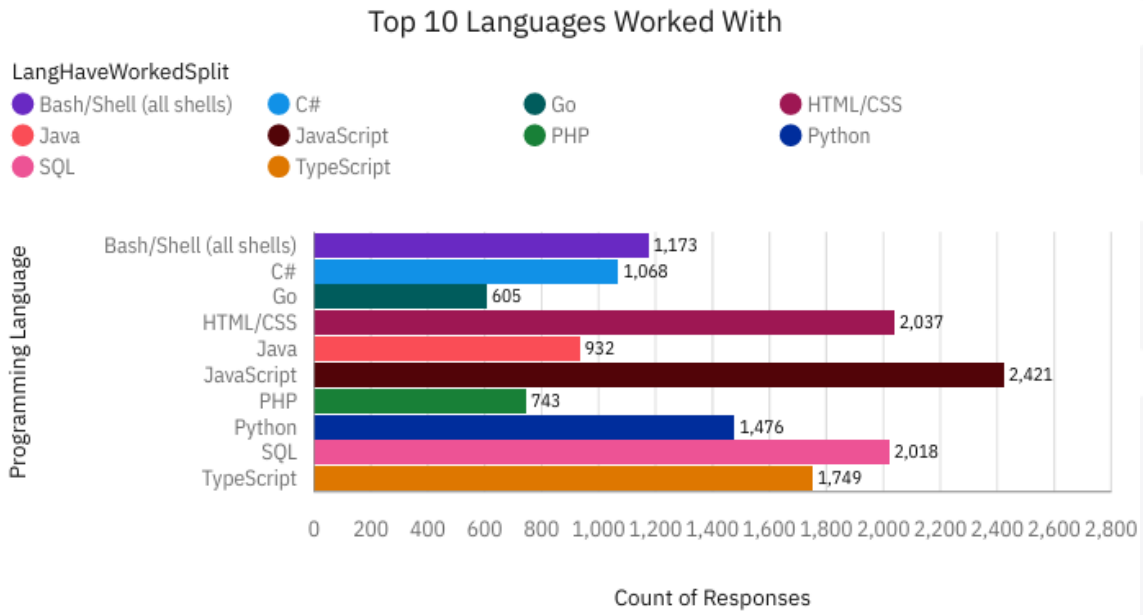
Respondent Demographics

- **Age**: Majority between 25-34 years old (41.3%)

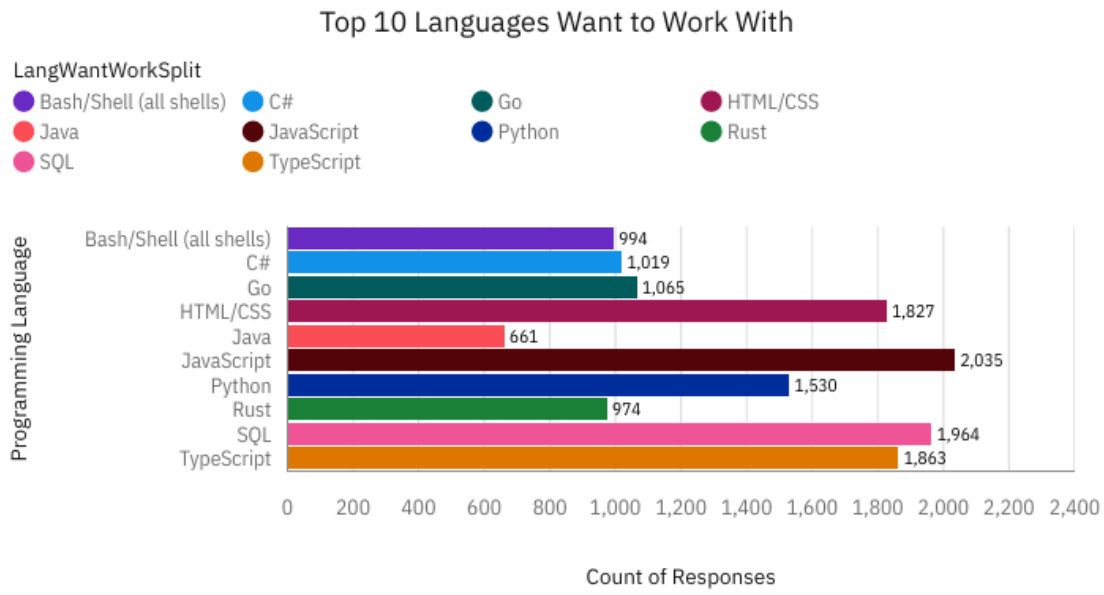
- **Education**: Most have a Bachelor's degree (8,629 responses)

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

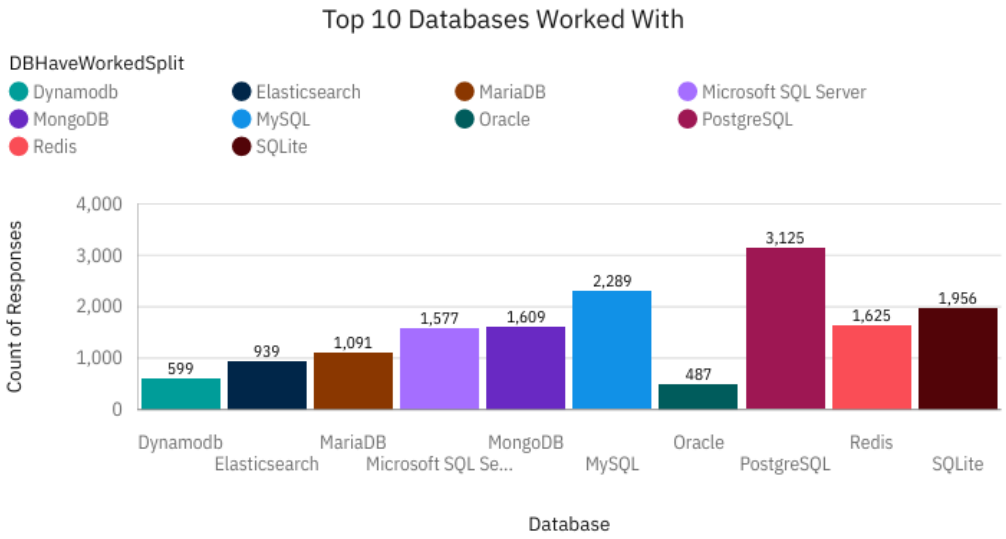
- JavaScript is the most dominant Programming Languages as it features top in both currently used and most desired.
- The languages most frequently used currently are the most desired next year
- Rust is the exception to other languages that it isn't present in the most currently used.

Implications

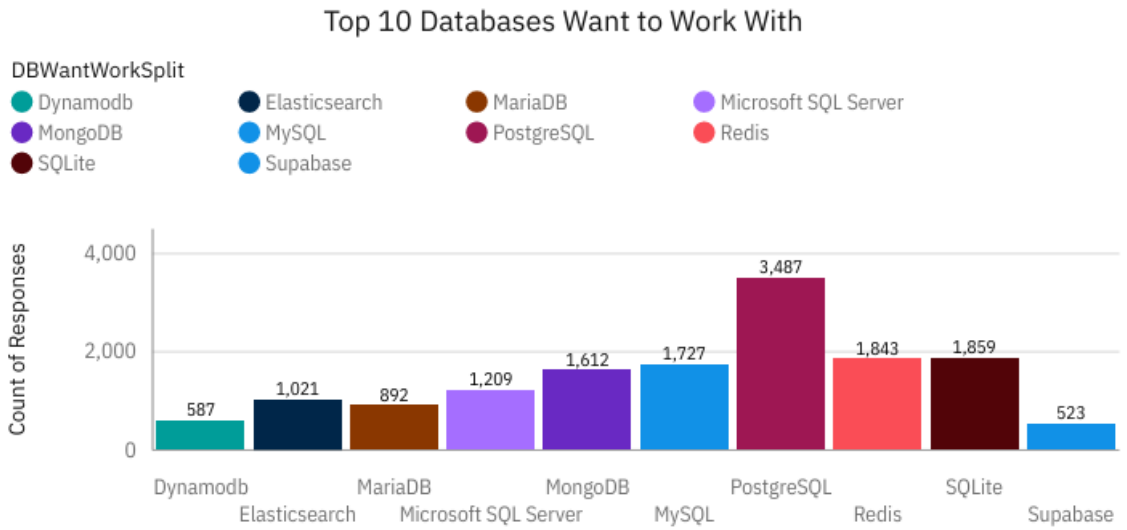
- A new person wanting to enter the field should prioritize learning JavaScript.
- The trend amongst languages overall show that most of the skillset for Developers will remain unchanged
- Rust seems like it could become more prevalent in future projects.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- PostgreSQL is the most dominant Database which top in both currently used and most desired.
- The Databases most frequently used currently are the most desired next year
- People seem to want to deviate from using Oracle as it isn't shown in the next years most desired.

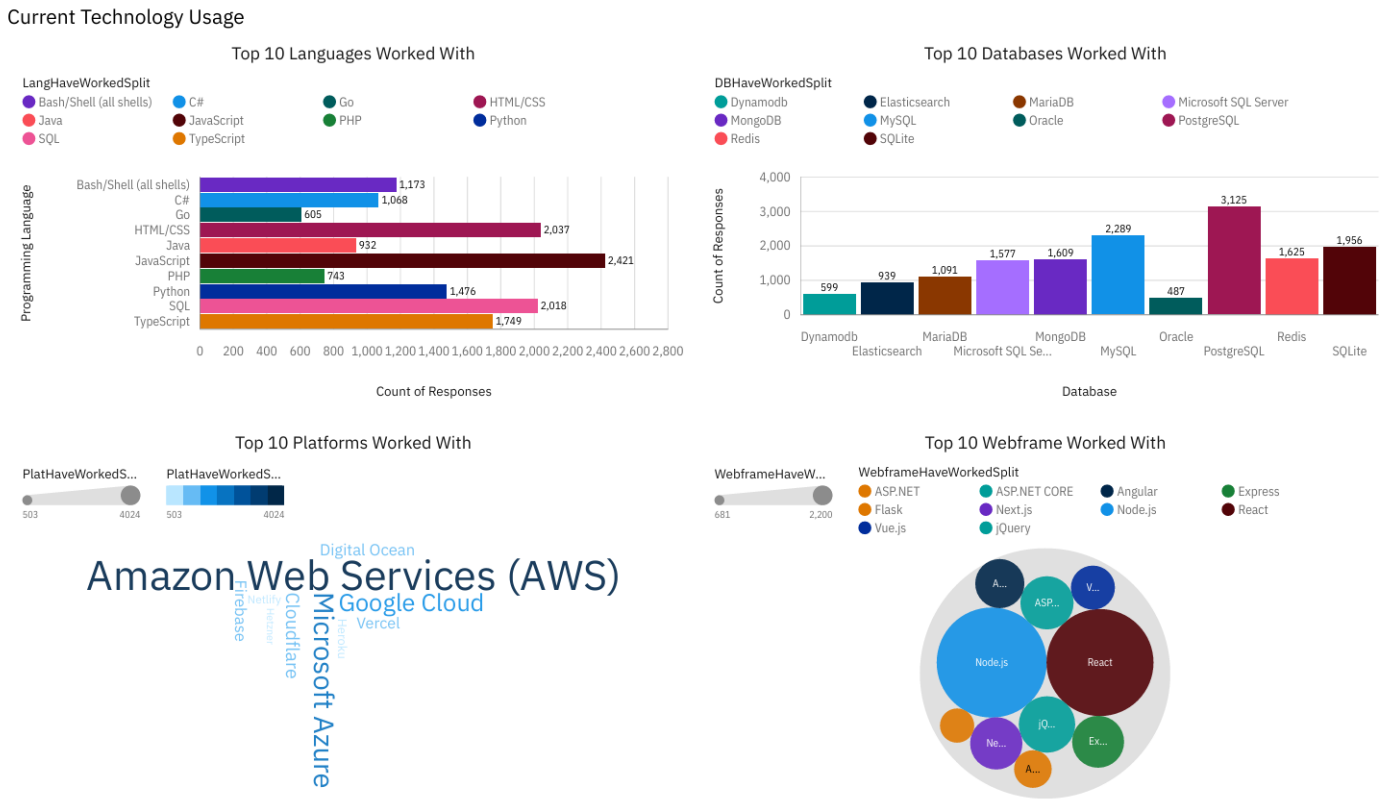
Implications

- New people wanting to enter the Career should try to learn how to use PostgreSQL
- The trend amongst Databases overall show that most of the skillset for Developers will remain unchanged
- It seems it may be time for people to develop skills outside of the Oracle Database

DASHBOARD

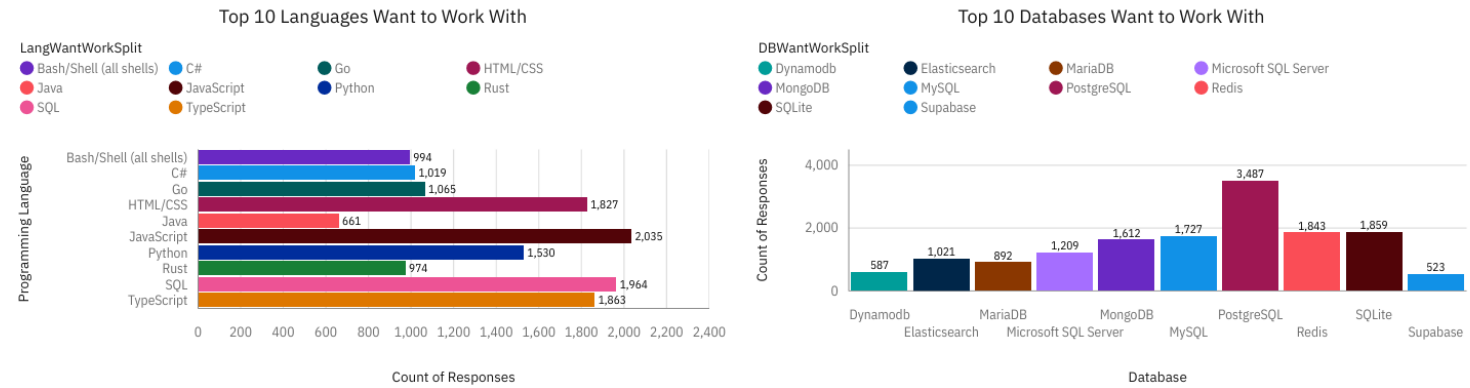
[IBM-Projects/Stack Overflow Survey Dashboard.pdf at main · Rywel94/IBM-Projects]

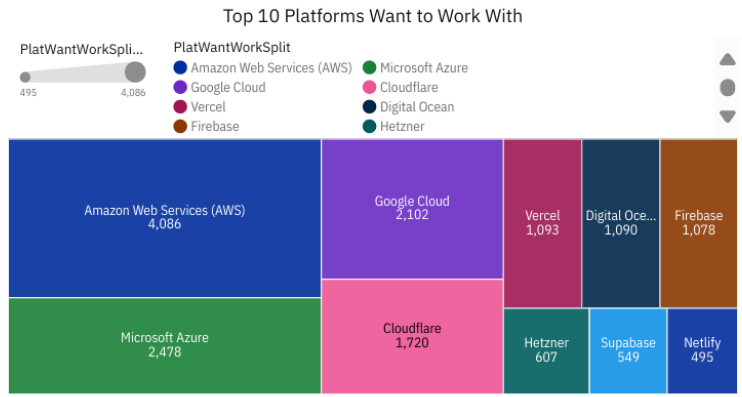
DASHBOARD TAB 1

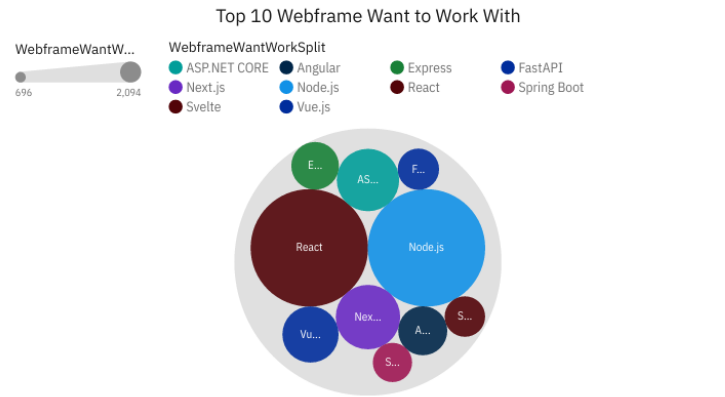


DASHBOARD TAB 2

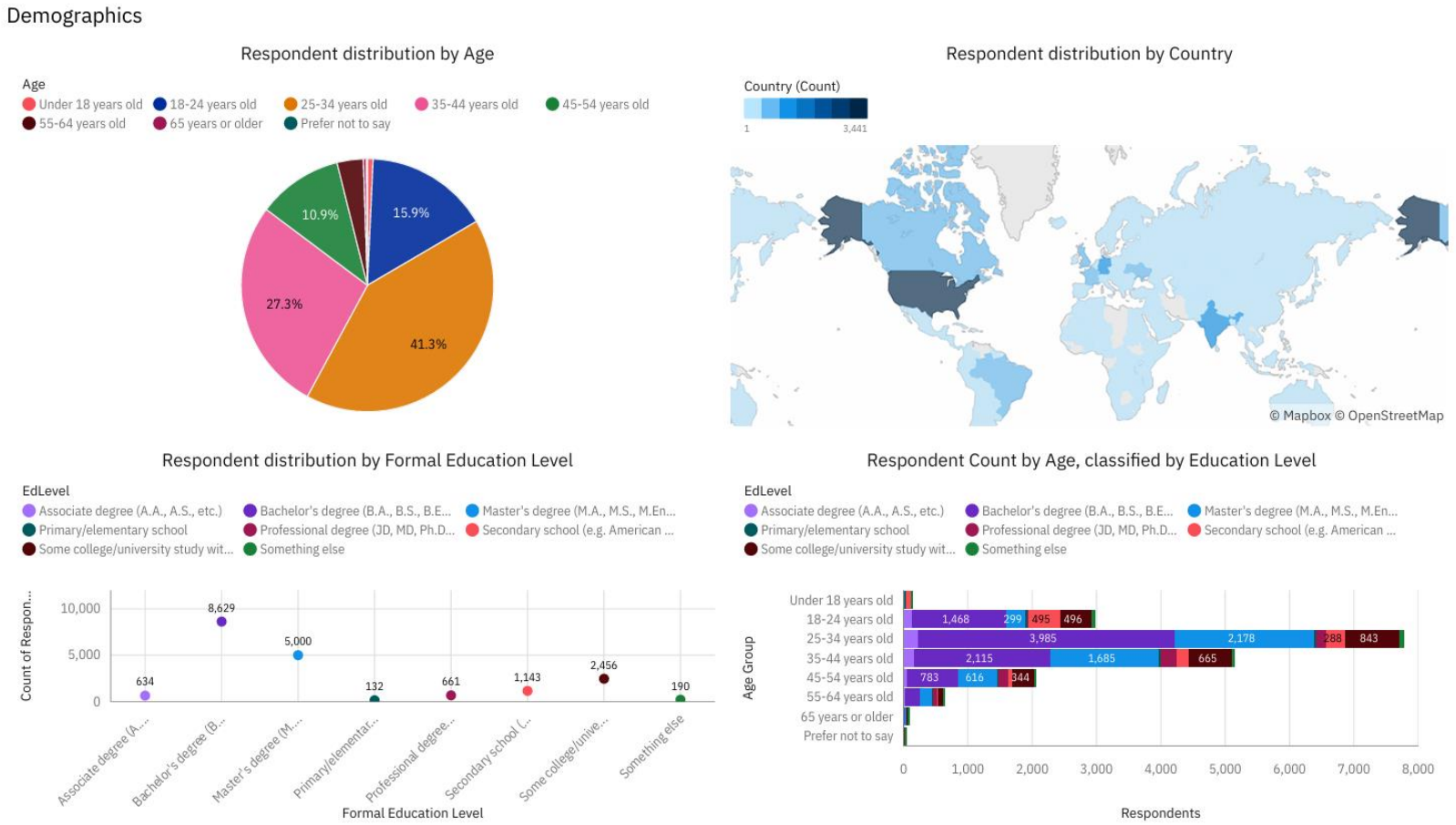
Future Technology Trend







DASHBOARD TAB 3



DISCUSSION

- What can we gather from these findings?
- Are the results as you expected after the initial introduction of this presentation?
- What metrics do you feel are the most important from the data gathered?

OVERALL FINDINGS & IMPLICATIONS

Findings

- 1. JavaScript Dominance: JavaScript remains the most popular programming language, indicating its continued relevance in web development.
- 2. PostgreSQL Popularity: PostgreSQL is the most used and desired database, reflecting its robustness and versatility.
- 3. Node.js and React: These platforms are highly favoured, showcasing the trend towards JavaScript-based frameworks for both front-end and back-end development.
- 4. Age and Education: The majority of respondents are young professionals (25-34 years old) with a Bachelor's degree, suggesting a well-educated and relatively young developer community.

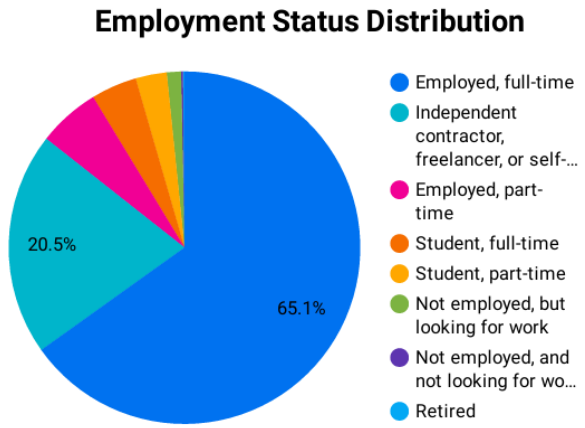
Implications

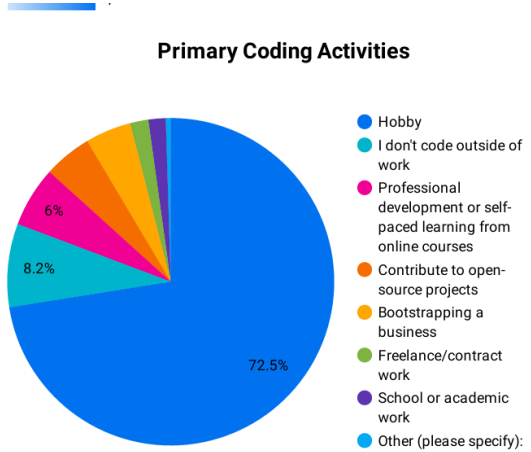
- 1. Skill Development: Developers should focus on mastering JavaScript and PostgreSQL to stay competitive in the job market.
- 2. Tool Adoption: Companies might consider adopting Node.js and React for their projects to align with industry trends and attract top talent.
- 3. Educational Programs: Educational institutions could tailor their curriculum to emphasize these popular technologies, ensuring graduates are well-prepared for the industry.
- 4. Community Engagement: The developer community is young and educated, which could influence the types of events, resources, and support systems that are most effective for engagement and professional development.

CONCLUSION

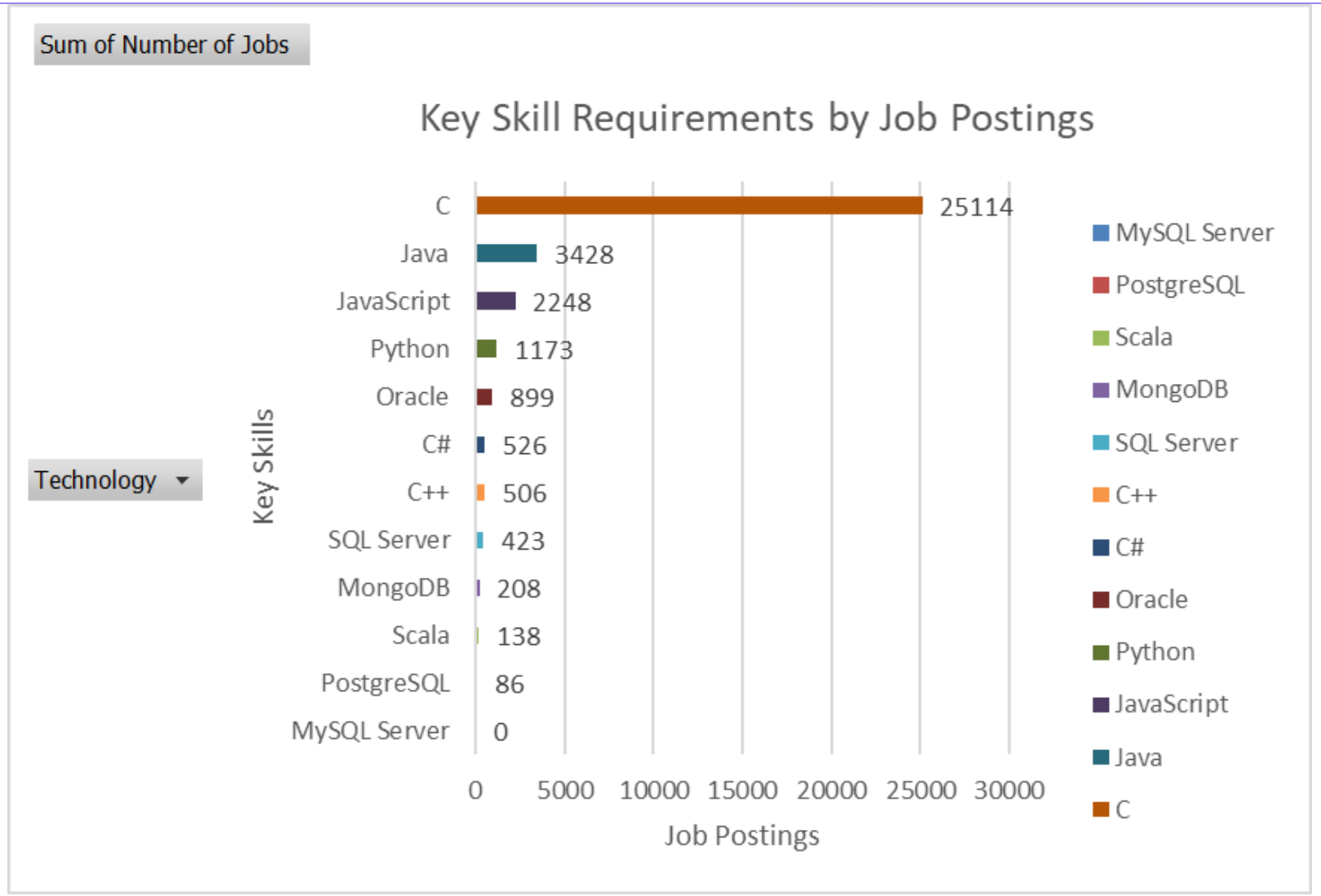
- New Entrants into the field should priorities learning a JavaScript Stack regardless of Front or Back end
- The Job Market is highly competitive with the majority being mid to late 20's with Degree's
- Companies should look at using platforms which JavaScript based in order to attract and retain talent.
- Universities or Online MOOC's should look at providing lessons or courses that include JavaScript as this is the most prevalent skill in use in the sector

APPENDIX


Employment Status Distribution
20.5%
65.1%
Employed, full-time
Independent contractor, freelancer, or self-...
Employed, part-time
Student, full-time
Student, part-time
Not employed, but looking for work
Not employed, and not looking for wo...
Retired


Primary Coding Activities
6%
8.2%
72.5%
Hobby
I don't code outside of work
Professional development or self-paced learning from online courses
Contribute to open-source projects
Bootstrapping a business
Freelance/contract work
School or academic work
Other (please specify):

JOB POSTINGS



POPULAR LANGUAGES