IBM Data Analyst Capstone Project

Saurabh Narendra Kulkarni

July 2024

IBM Developer

SKILLS NETWORK

OUTLINE

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

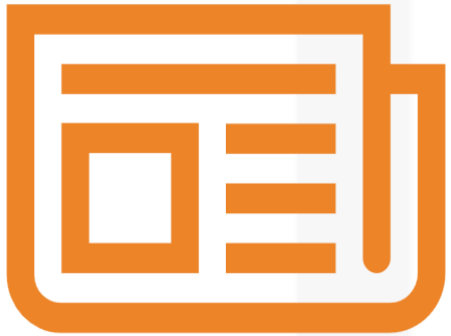
EXECUTIVE SUMMARY

- **Analysis of Current Technology Usage**
  - JavaScript, HTML/CSS, SQL, Bash/Shell/PowerShell, Python, Java, C#, TypeScript, PHP, and C++ are the top 10 programming languages used.
  - MySQL, Microsoft SQL Server, PostgreSQL, MongoDB, SQLite, Redis, Elasticsearch, MariaDB, Oracle, and Firebase are the top 10 databases used.
  - Platforms worked with include Docker, Linux, AWS, Android, Windows, Google Cloud Platform, MacOS, Microsoft Azure, Heroku, and iOS. The top 10 web frameworks worked with are jQuery, Angular/Angular.js, ASP.NET, React.js, Express, Django, Spring, Flask, Laravel, and Vue.js.
- **Future Technology Trends**
  - HTML/CSS, Bash/Shell/PowerShell, C#, JavaScript, Go, Java, C++, Python, C, and Elixir are the most desirable programming languages for the coming year..
  - Elasticsearch, MongoDB, PostgreSQL, Microsoft SQL Server, MySQL, Firebase, MariaDB, Cassandra, DynamoDB, and Redis are the most sought-after databases for the next year.
  - Platforms respondents are interested in include AWS, Docker, Android, Linux, Arduino, Google Cloud Platform, Kubernetes, Windows, iOS, and Heroku. The most desired web frameworks for the upcoming year are Angular/Angular.js, React.js, ASP.NET, Express, Django, jQuery, Vue.js, Flask, Spring, and Laravel.l..
- **Demographics Analysis**
  - The gender distribution among respondents shows a significant gap, with males representing 93.7% and females 6.3%.
  - Visualizations depict the distribution of respondents across various countries.
  - Respondent age distribution spans from 16 to 99 years. Additionally, the analysis includes respondent counts categorized by gender and formal education level.
- **Emerging Technology Trends:**The analysis reveals a clear shift in technology preferences. The top programming languages, databases, platforms, and web frameworks for the next year indicate a growing interest in modern, scalable technologies and tools. This trend underscores the importance for organizations and professionals to adapt to these evolving technologies to stay competitive.
- **Gender Representation and Global Reach:**The significant gender gap in the IT industry highlights the need for more inclusive practices and initiatives to attract a diverse talent pool. The global distribution of respondents and their varied age and educational backgrounds suggest that technology adoption is widespread and crosses various demographics, indicating a broad and evolving technological landscape.

INTRODUCTION

- Stack Overflow, a renowned platform for developers, conducted a 2019 survey to gather global insights on technology trends and developer preferences.
- The assignment utilizes the free trial version of IBM Cognos Analytics for creating visualizations and dashboards.
- Access to IBM Cognos Analytics is required. A Cognos lab will guide you on how to obtain access and use the tool effectively.
- Target Audience
  - **IT Professionals:** Individuals working in or interested in technology and development.
  - **General Audience:** Anyone interested in understanding technology trends and developer preferences.

METHODOLOGY

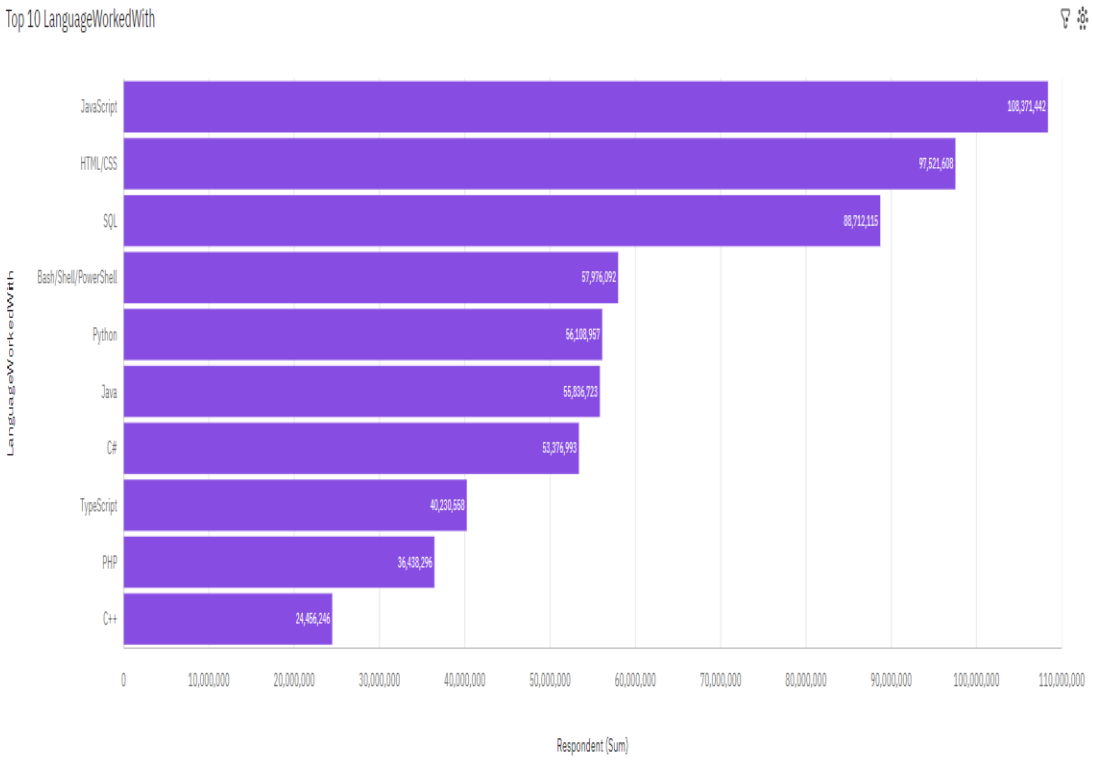


- **Data Sources:** Stack Overflow Developer Survey 2019, GitHub Job Postings, Programming Language Annual Salary.
- Data Exploration and Cleaning: Python
- Data Visualization: IBM Cognos Analytics
- Presentation
  - **Data Exploration and Cleaning**: Python used for cleaning and exploration of data.
  - **Data Visualization**: IBM Cognos Analytics employed for creating dashboards and visualizations.
  - **Presentation**: Compile the final visualizations into a professional presentation.

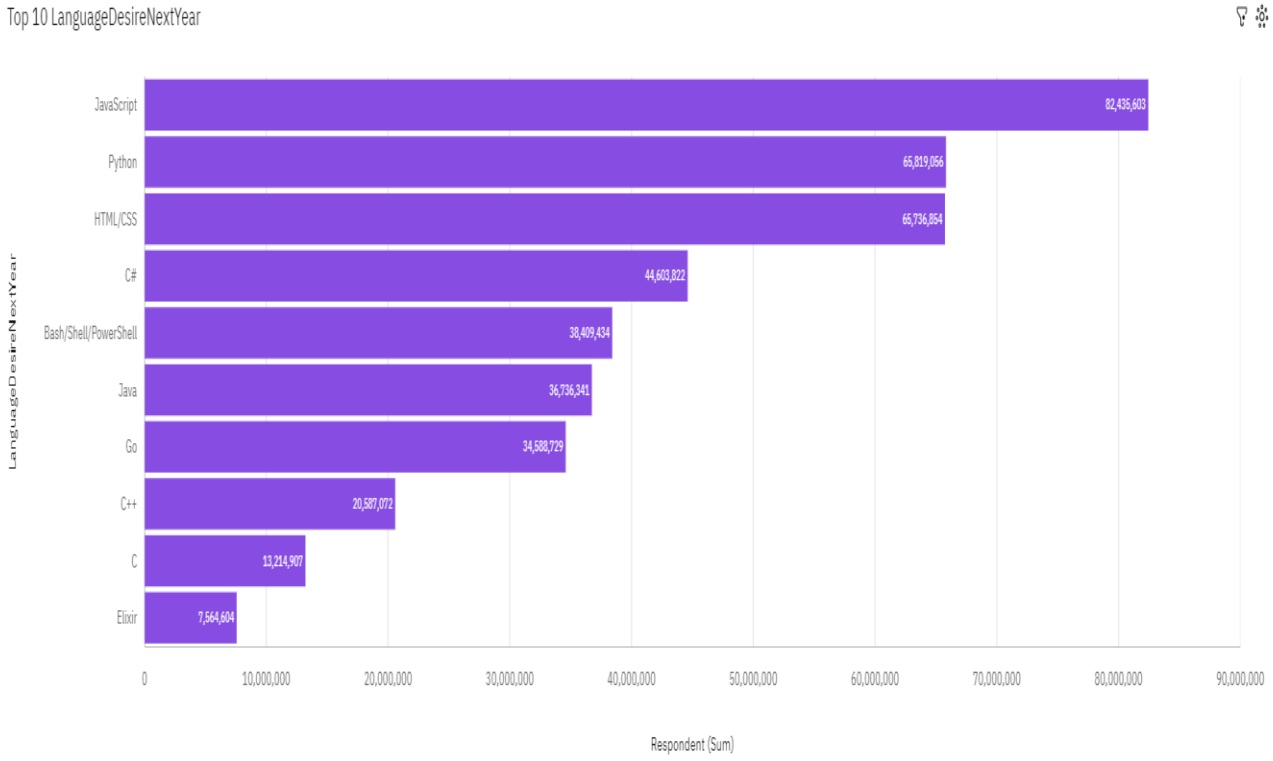
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Top Programming Languages Used: JavaScript, HTML/CSS, SQL, Bash/Shell/PowerShell, Python, Java, C#, TypeScript, PHP, and C++ are the most widely used programming languages currently.
- Desirable Programming Languages for Next Year: HTML/CSS, Bash/Shell/PowerShell, C#, JavaScript, Go, Java, C++, Python, C, and Elixir are expected to be the most desirable programming languages for the upcoming year.
- Technology Adoption Trends: There is a notable shift towards newer and versatile programming languages, indicating an evolving landscape where flexibility and multi-functional tools are increasingly valued.

Implications

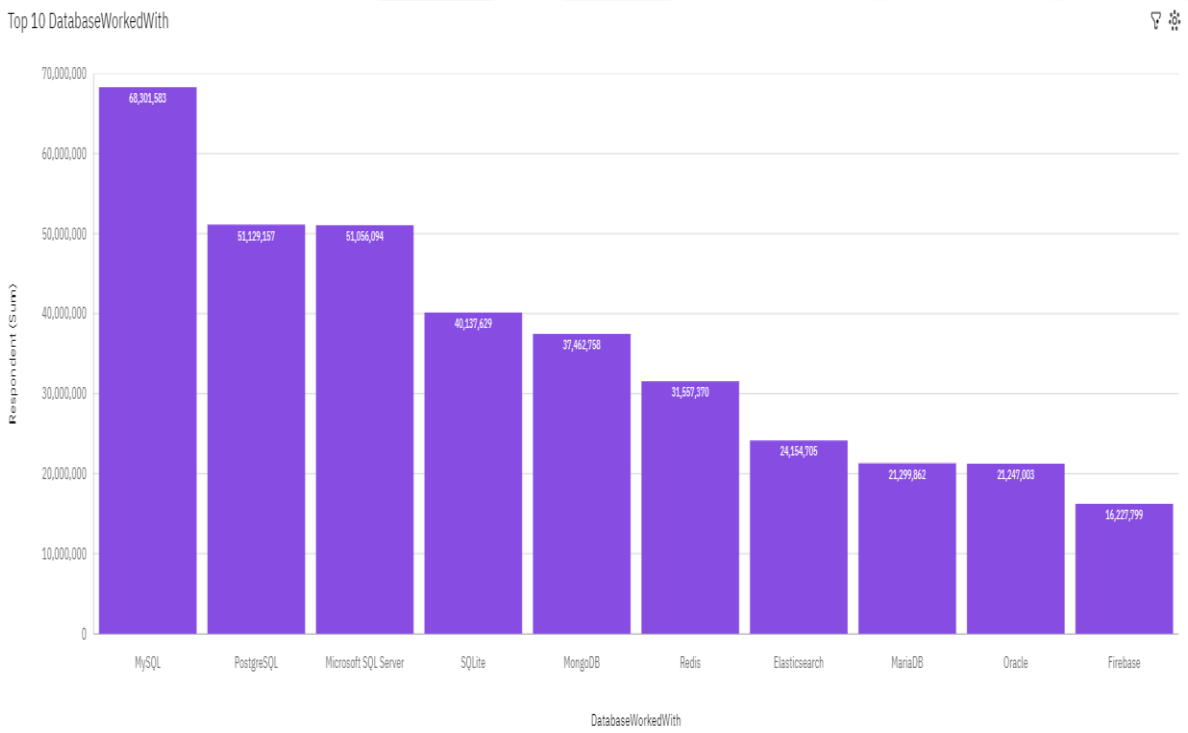
• Skill Development: Developers should focus on learning and mastering the most desirable programming languages for the next year to stay competitive and relevant in the job market.

• Hiring and Recruitment: Organizations should prioritize candidates with skills in the top programming languages and emerging technologies to align with current and future technology trends.

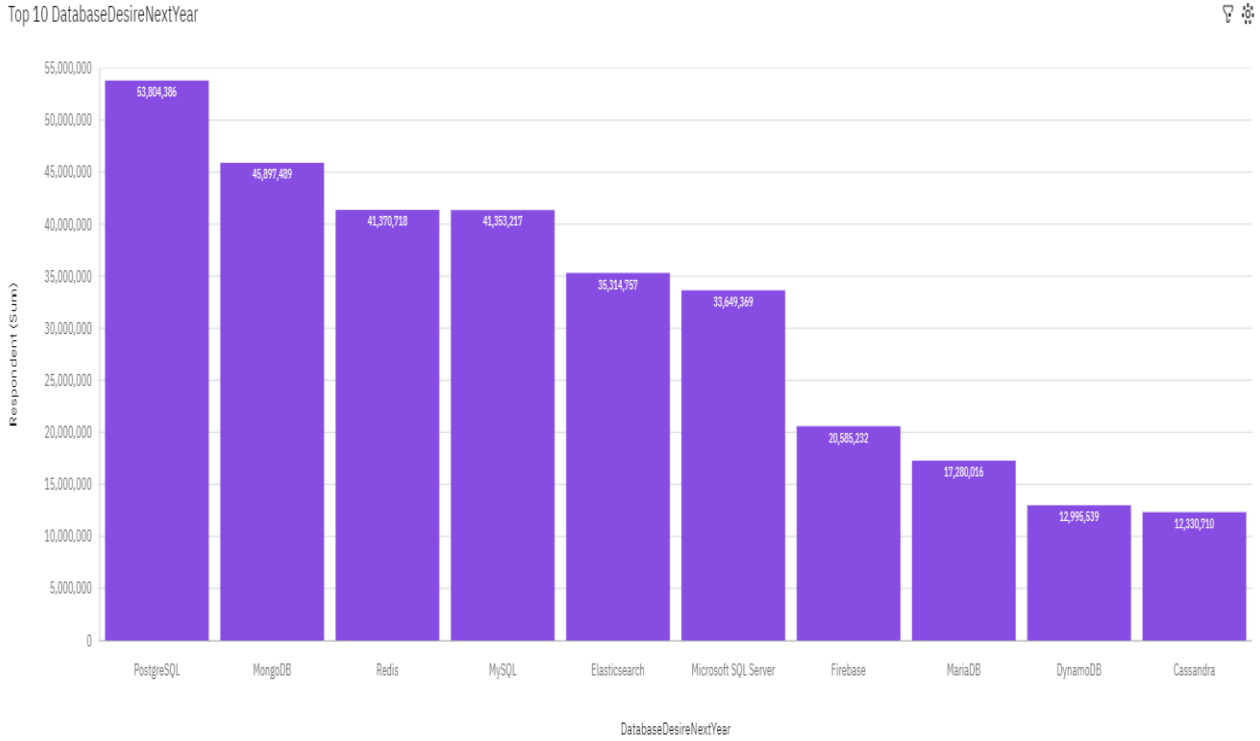
• Strategic Planning: Companies should consider integrating popular and upcoming programming languages into their technology stack and training programs to enhance their product offerings and keep up with industry advancements.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- **Top Databases Used**: MySQL, Microsoft SQL Server, PostgreSQL, MongoDB, SQLite, Redis, Elasticsearch, MariaDB, Oracle, and Firebase are the most frequently used databases in the current landscape.
- **Most Desired Databases for Next Year**: Elasticsearch, MongoDB, PostgreSQL, Microsoft SQL Server, MySQL, Firebase, MariaDB, Cassandra, DynamoDB, and Redis are anticipated to be the most sought-after databases for the upcoming year.
- **Emerging Database Trends**: There is a growing preference for databases that support real-time data processing, scalability, and flexibility, reflecting the industry's shift towards more dynamic and scalable data management solutions.

Implications

- Database Skill Development: Professionals should focus on gaining expertise in the most desired databases to enhance their career prospects and align with future industry requirements.

- Strategic Database Selection: Organizations should consider integrating popular and emerging databases into their technology stack to improve data management capabilities and stay competitive.

- Investment in Data Infrastructure: Companies should invest in databases that offer scalability and real-time processing features to meet evolving data needs and support their growth strategies.

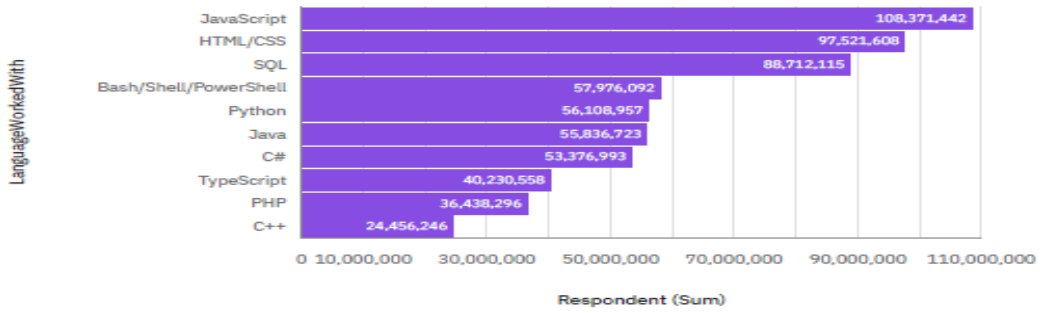
DASHBOARD

https://github.com/SaurabhKulkarni98/Capstone-Project/blob/main/Capstone%20Project.pdf

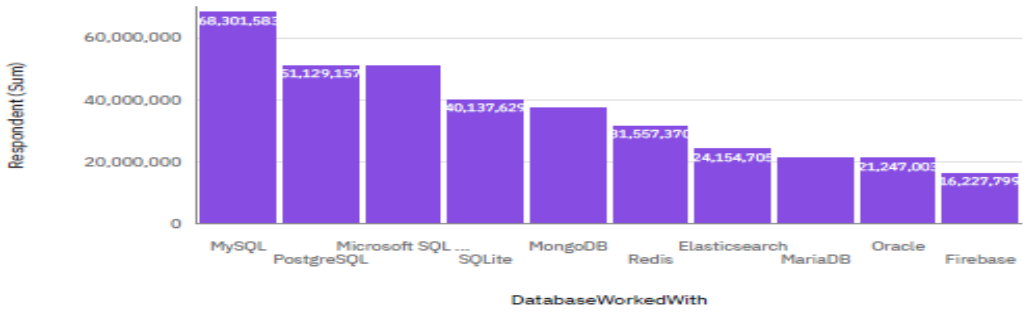
DASHBOARD TAB 1

Current Technology Usage

Top 10 LanguageWorkedWith

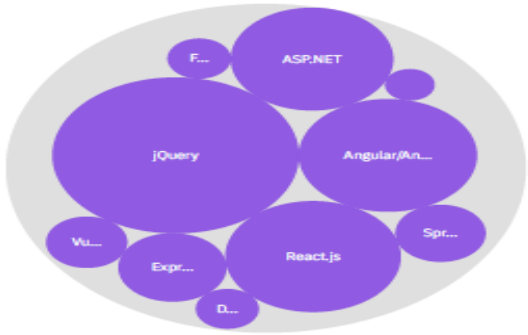


Top 10 DatabaseWorkedWith



PlatformWorkedWith

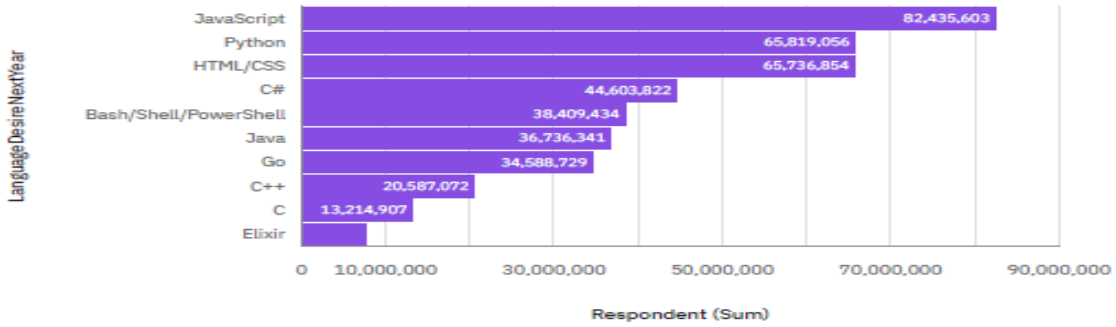
Top 10 WebFrameWorkedWith



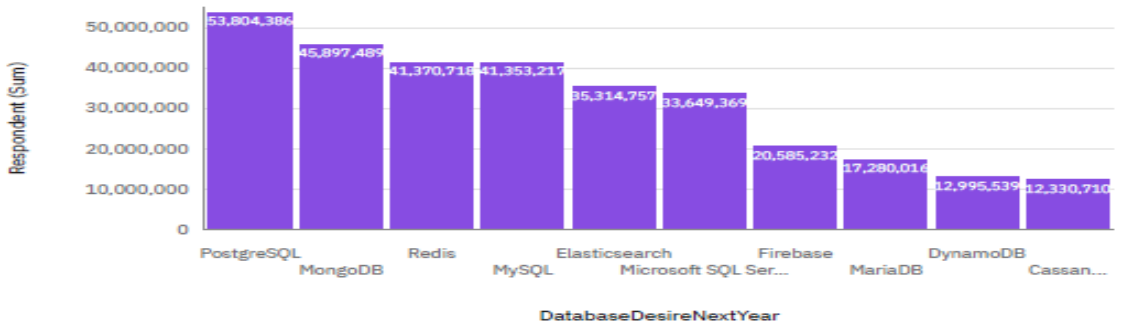
DASHBOARD TAB 2

Future Technology Trend

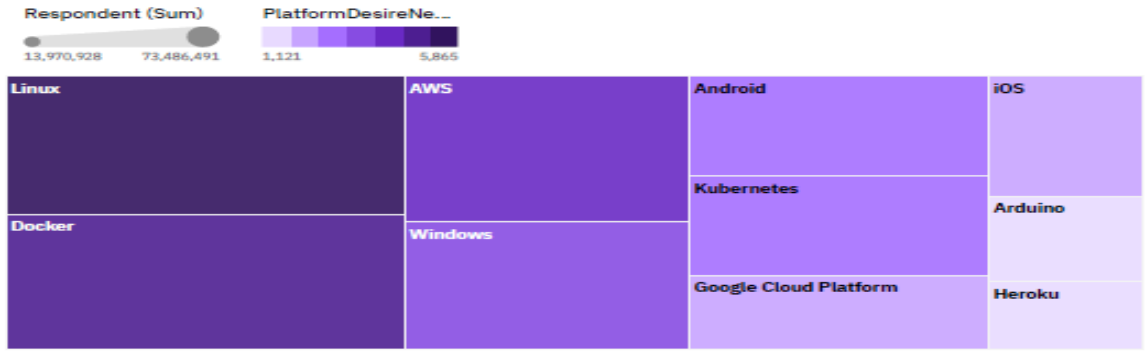
Top 10 LanguageDesireNextYear



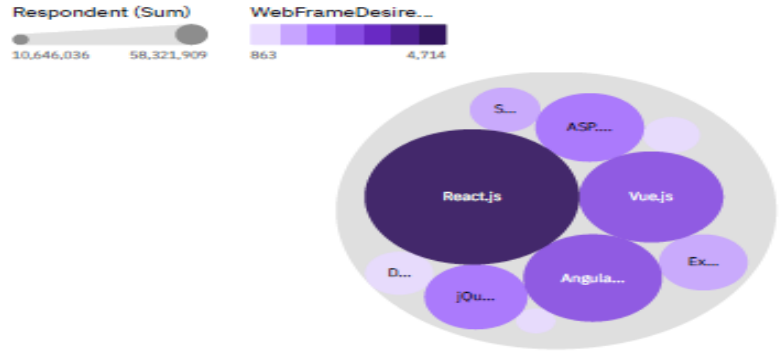
Top 10 DatabaseDesireNextYear



PlatformDesireNextYear



Top 10 WebFrameDesireNextYear



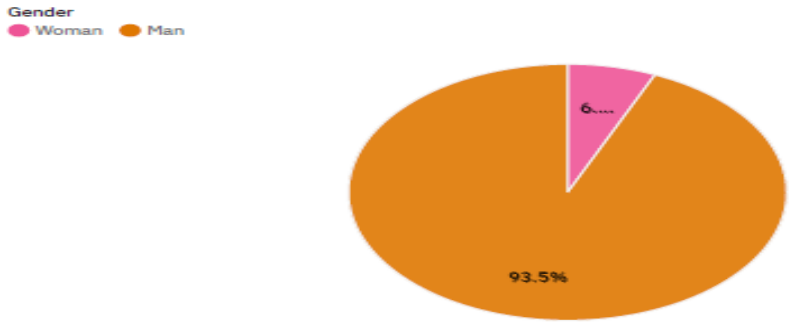
IBM Developer

SKILLS NETWORK

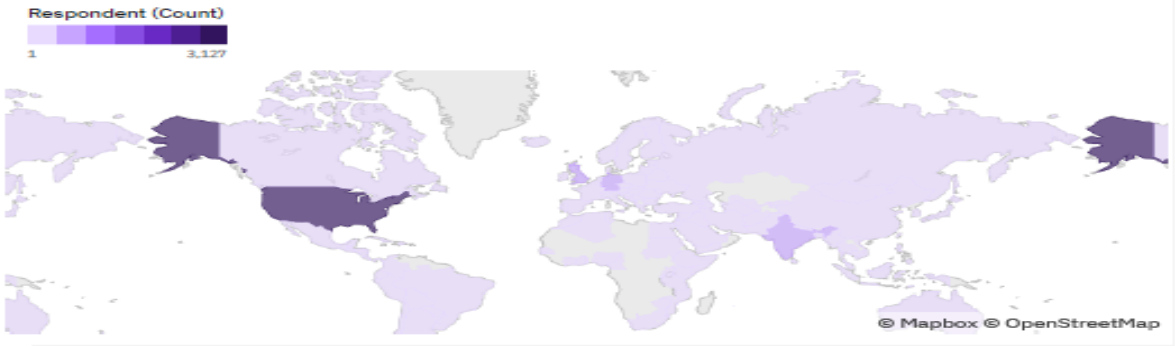
DASHBOARD TAB 3

Demographics

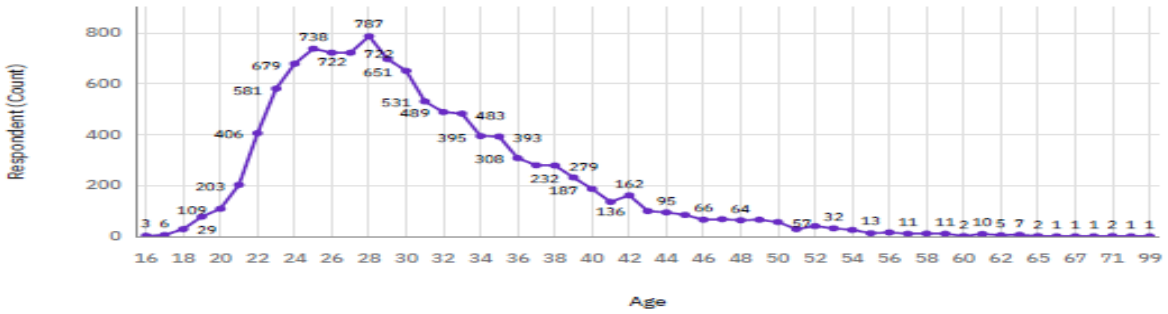
Respondent classified by Gender



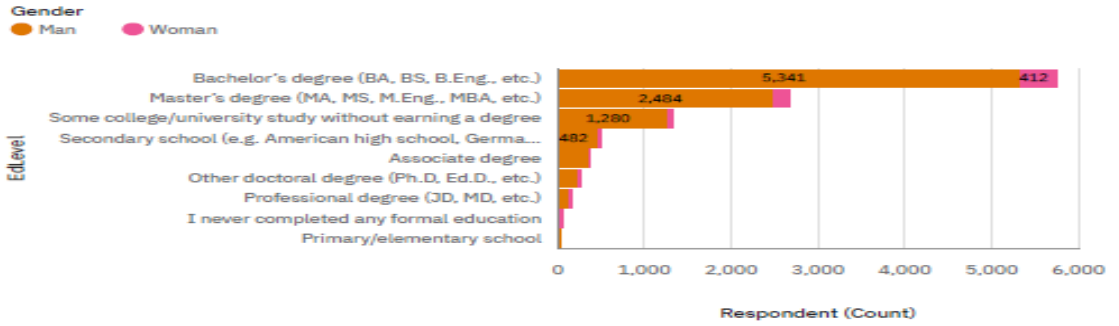
Respondent Count for Countries



Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level



DISCUSSION

- Emerging Technologies and Shifting Demand
  - New technologies continually emerge, driving changes in demand for specific programming languages, databases, platforms, and web frameworks.
  - Organizations and professionals must stay updated with these trends to remain competitive and meet evolving industry requirements.
- Addressing the Gender Gap in IT
  - The significant gender gap, with males representing 93.7% and females 6.3%, highlights the need for more inclusive practices and initiatives in the IT industry.
  - Encouraging diversity and providing opportunities for underrepresented groups can help bridge this gap and create a more balanced workforce.
- Global and Demographic Trends
  - The global distribution of respondents and their varied age and educational backgrounds suggest a broad and evolving technological landscape.
  - Understanding these trends can help in developing targeted strategies for technology adoption and skill development across different demographics.

OVERALL FINDINGS & IMPLICATIONS

Findings

- The most used programming languages are JavaScript, HTML/CSS, SQL, and Python.
- Popular databases among respondents include MySQL, PostgreSQL, and MongoDB.
- Key platforms and web frameworks used are Docker, Linux, AWS, React.js, and Angular/Angular.js.

Implications

- Organizations must adapt to evolving technology trends to stay competitive, emphasizing continuous learning and adaptation.
- Addressing the significant gender gap in the IT industry is crucial for fostering a diverse and inclusive workforce.
- Understanding global and demographic trends helps in developing targeted strategies for technology adoption and workforce development.

CONCLUSION

- **Adaptation to Evolving Technologies:** Staying updated with emerging trends in programming languages, databases, platforms, and web frameworks is essential for maintaining competitiveness and innovation.

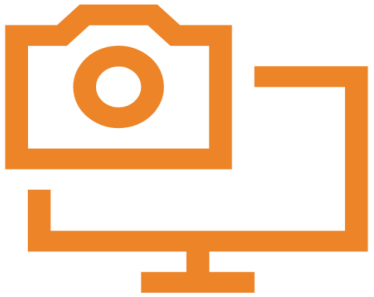
- **Addressing the Gender Gap:** Encouraging diversity and providing equal opportunities for all genders can help bridge the significant gender disparity in the IT industry.

- **Importance of Data-Driven Insights:** Utilizing data from sources like the Stack Overflow Developer Survey enables informed decision-making and alignment with industry trends.

- **Global and Demographic Trends:** Understanding the global distribution and demographic characteristics of respondents aids in developing targeted strategies for technology adoption and workforce development.

APPENDIX

Relevant additional charts, or tables that I have created during the analysis phase.



LanguageWorkedWith

LanguageWorkedWith
Bash/Shell/PowerShell
C#
C++
HTML/CSS
Java
JavaScript
PHP
Python

LanguageDesireNextYear

LanguageDesireNextYear
Bash/Shell/PowerShell
C#
Go
HTML/CSS
Java
JavaScript
Kotlin
Python

DatabaseWorkedWith

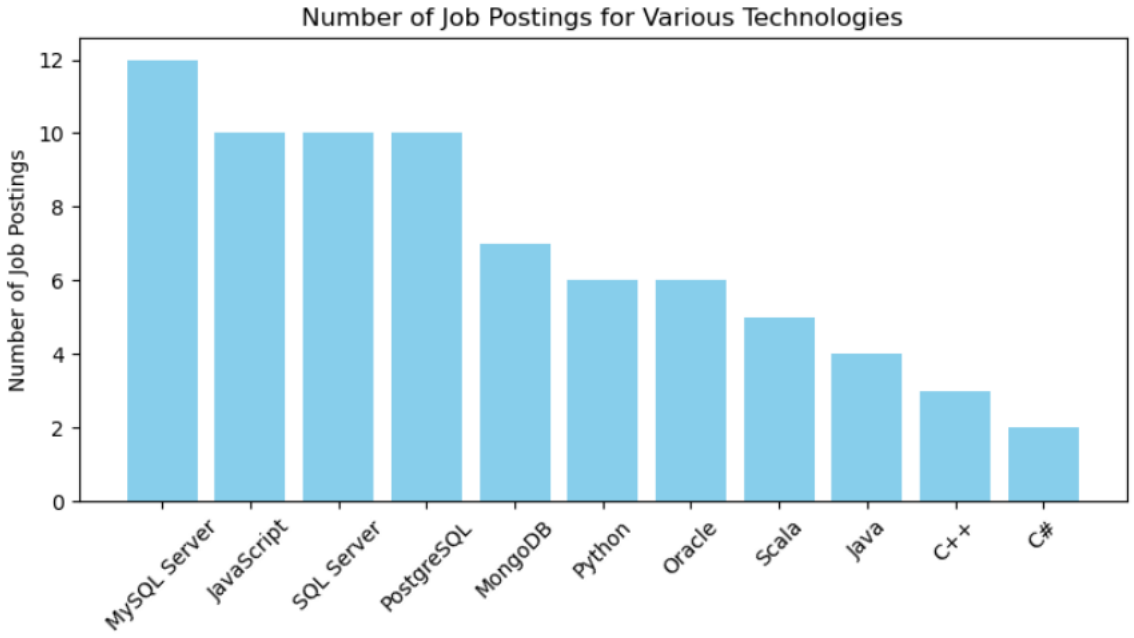
DatabaseWorkedWith
Elasticsearch
Firebase
MariaDB
Microsoft SQL Server
MongoDB
MySQL
Oracle
PostgreSQL

DatabaseDesireNextYear

DatabaseDesireNextYear
DynamoDB
Elasticsearch
Firebase
MariaDB
Microsoft SQL Server
MongoDB
MySQL
PostgreSQL

JOB POSTINGS

In this analysis, we present the job posting data collected from the GitHub Jobs API, saved in the file "job-postings.xlsx." The bar chart below visualizes the number of job postings for various technologies. The data is ordered in descending order, highlighting the technologies with the highest number of job postings.



POPULAR LANGUAGES

In this analysis, we present the average annual salary data for various programming languages, which was collected through web scraping and saved in the file "popular-languages.csv." The bar chart below visualizes the average annual salaries for different programming languages. The data is ordered in descending order of salary to highlight the highest-paying languages.