IBM Data Analyst Capstone Project

Rodrigue Ndountio Fomena

May 2023

IBM Developer

SKILLS NETWORK

OUTLINE

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

EXECUTIVE SUMMARY

- A global IT consulting firm regularly analyzes data to identify future skill requirements.
- As a Data Analyst, we are tasked with collecting and analyzing data to identify emerging programming skills.
  - We will collect data from job postings, training portals, and surveys.
  - The data will be analyzed to identify trends in programming languages, database skills, and popular IDEs.
  - The data wrangling techniques will be used to make the data ready for analysis.
- We will collect data in various formats like .csv files, excel sheets, and databases.
- Statistical techniques will be used to analyze the data.
- A dashboard will be created by using IBM Cognos Analytics to bring all of our information together.

INTRODUCTION

- In today's fast-paced digital world, technology is constantly evolving, and organizations must keep pace with these changes to remain competitive in the market.
- The project is being carried out by a data analyst who has been hired by a global IT consulting firm.
- The aim of the project is to identify emerging programming skills by analyzing data collected from various sources.
- The project will involve data collection, data analysis, and creating a dashboard to present the findings.
  - Data will be collected from job postings, training portals, and surveys.
  - Statistical techniques will be used to analyze the data and identify trends in programming languages, database skills, and popular IDEs.

METHODOLOGY

- Data collection
  - Collect data from various sources using APIs and Webscraping.
  - Explore and understand the collected data.
- Data Wrangling
  - Identify and handle missing values in the data.
  - Find and remove duplicates.
  - Normalize the data as needed.
- Exploratory Data Analysis (EDA)
  - Examine the distribution of the data.
  - Identify and handle outliers.
  - Determine correlations between variables.
- Data Visualization
  - Use appropriate visualization techniques to show distribution, relationship, composition, and comparison of the data.
- Dashboard Creation
  - Create a dashboard using data visualization tools like IBM Cognos Analytics to display the insights and findings from the data.
- Interpret the findings
  - Interpret the findings and draw conclusions based on data analysis.

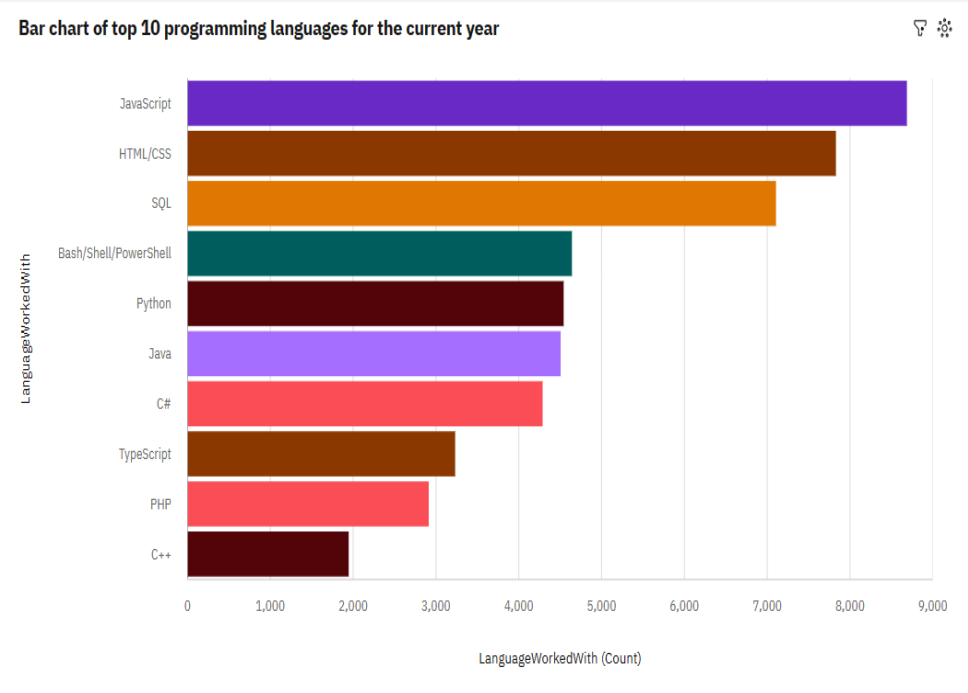
RESULTS

The results of this project are based on the analysis of data collected from various sources in order to identify the top programming skills that are most in demand.

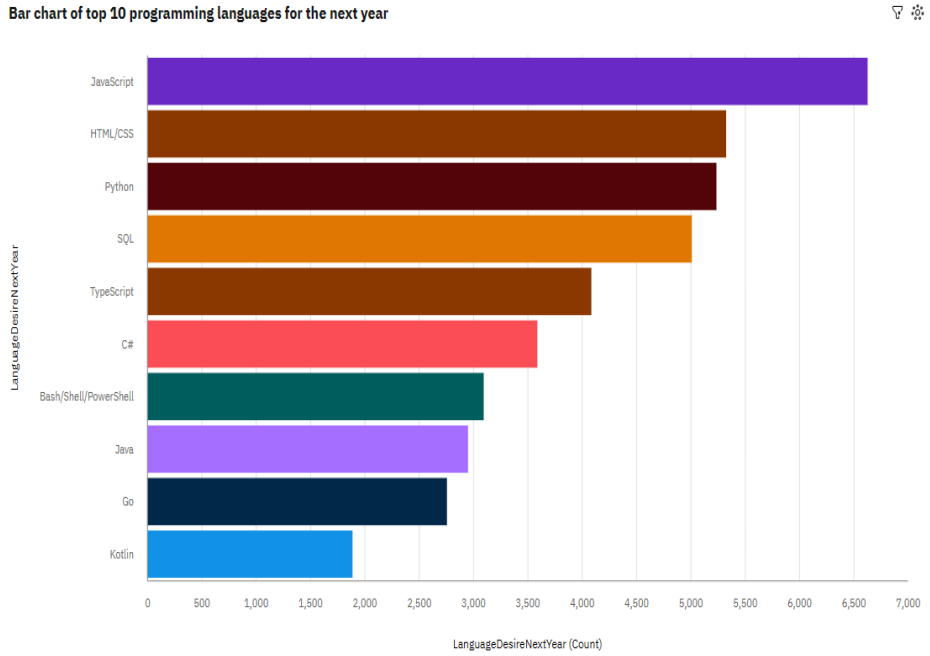
The insights and trends identified from this analysis may include the top programming languages in demand, the top database skills in demand, and the popular IDEs. The final results are presented in a dashboard.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

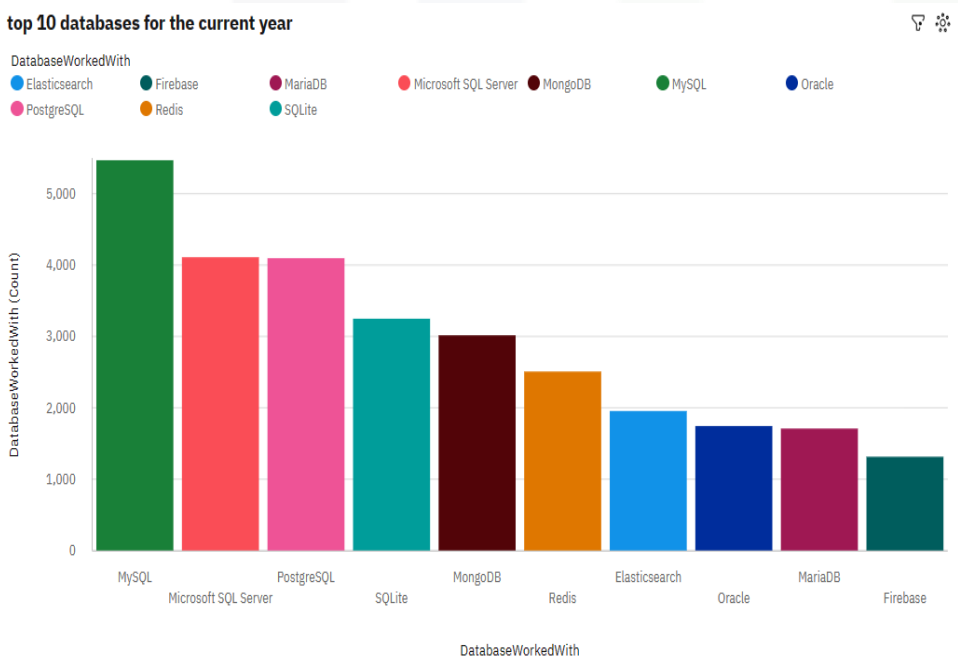
- Current year top 3 languages
  - Javascript
  - HTML/CSS
  - SQL
- Next year top 3 languages
  - Javascript
  - HTML/CSS
  - Python

Implications

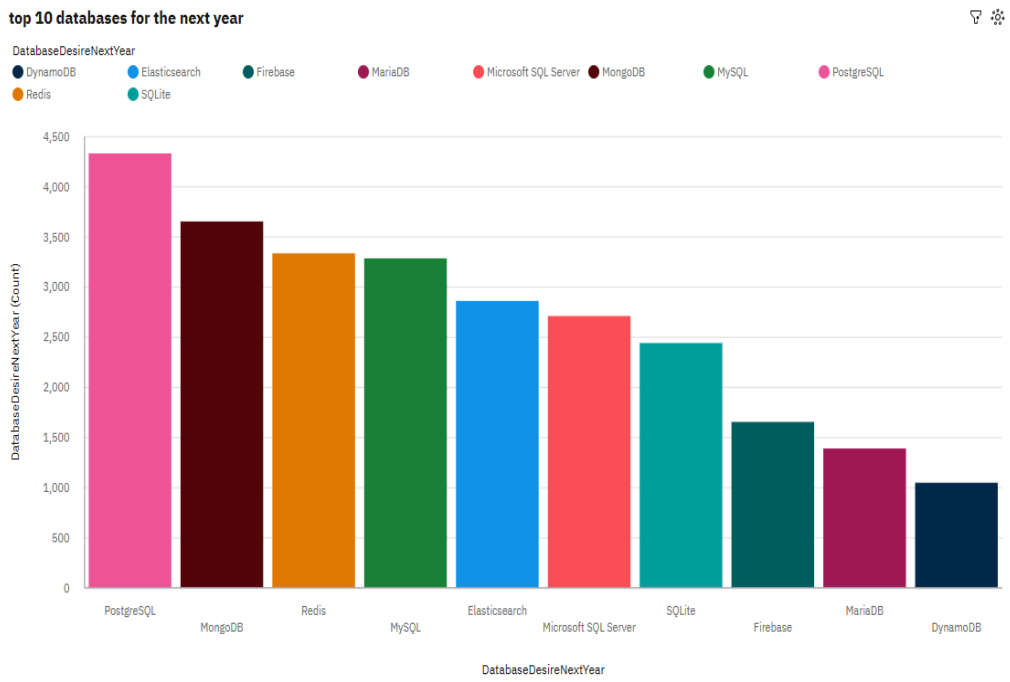
- Javascript and HTML/CSS will remain the 2 best languages as they are currently according to the data, that means that the web development is still in high demand
- While Python is the upcoming trend for programming languages

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Current year top 3 Databases
  - MySQL
  - Microsoft SQL Server
  - PostgresSQL
- Next year top 3 Databases
  - PostgresSQL
  - MongoDB
  - Redis

Implications

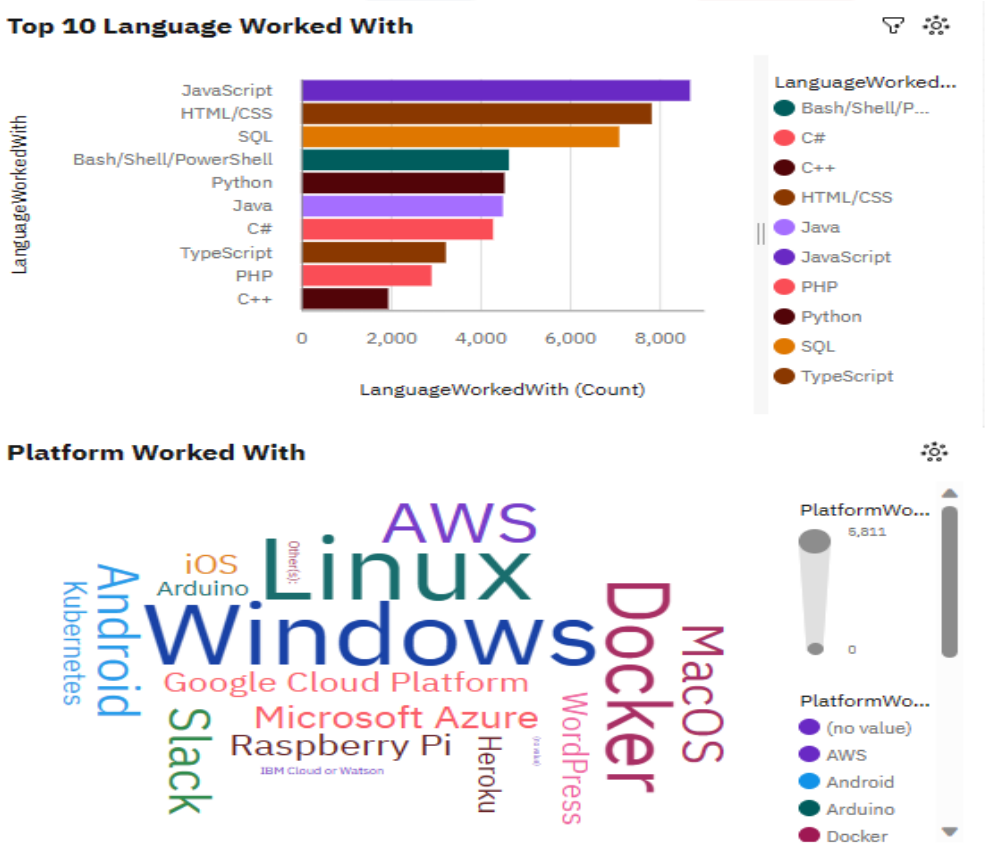
- PostgresSQL will be the most uses database replacing MySQL
- Microsoft SQL Server will no longer be in the top 3, leaving room for non relational open source databases which are: MongoDB and Redis

DASHBOARD

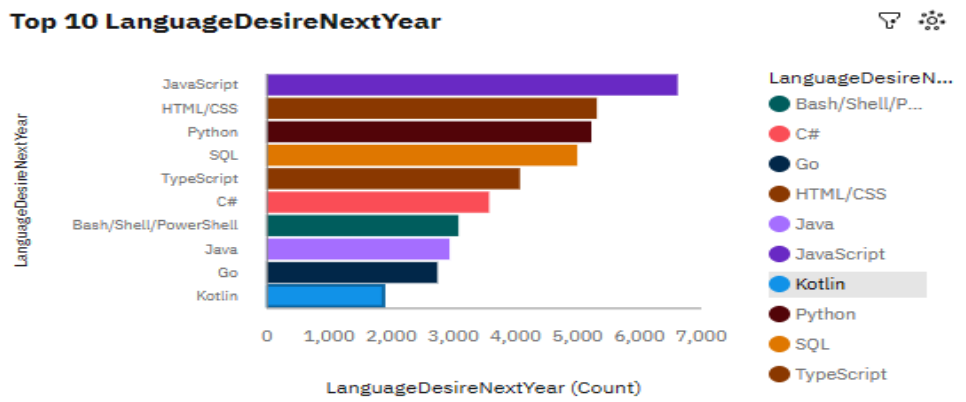
Permanent link of the read-only view of the Cognos dashboard:

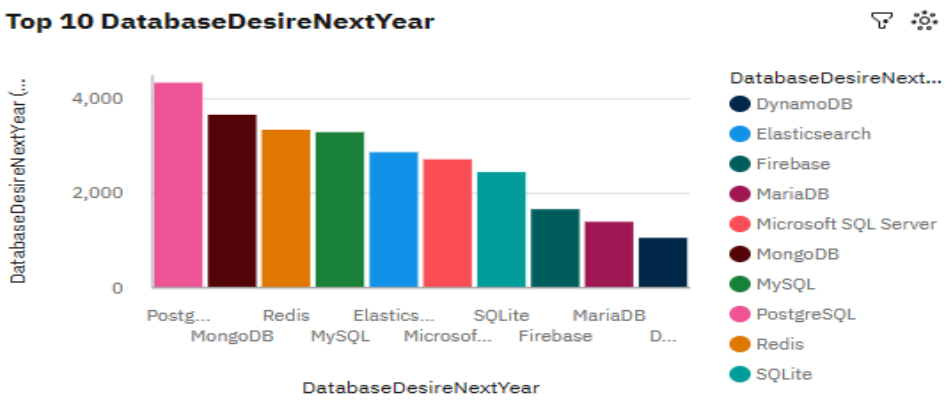
https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FCapstone_Dashboard&action=view&mode=dashboard&subView=model00000187dcdefe72_00000000

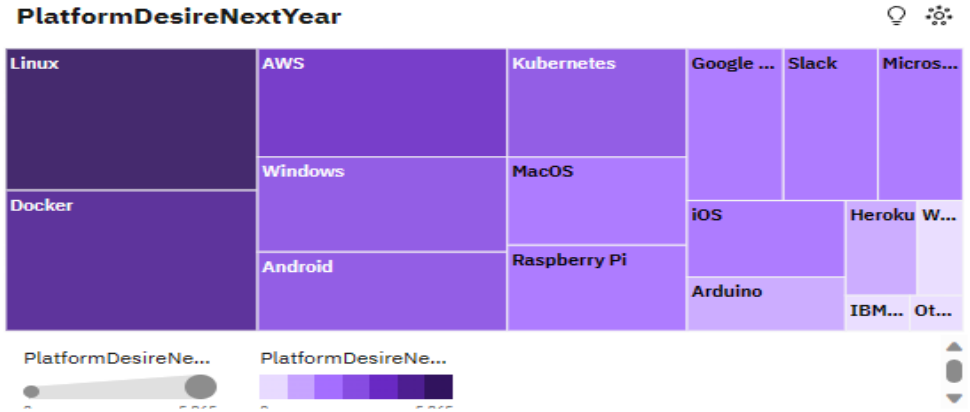
CURRENT TECHNOLOGY USAGE

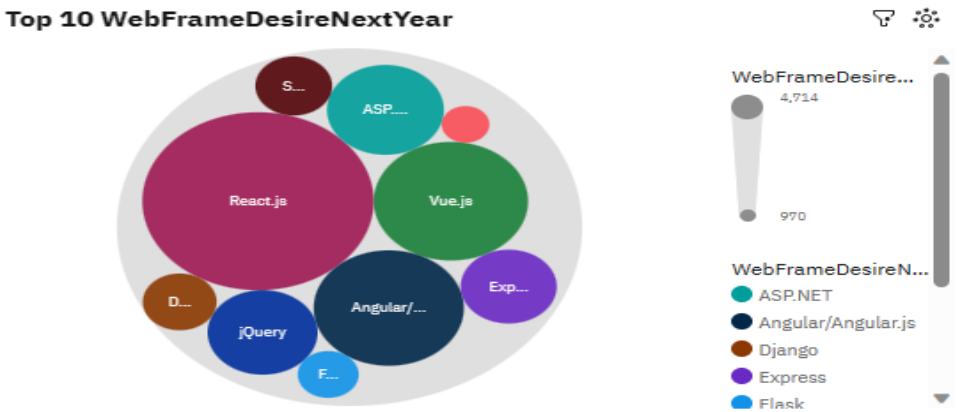


FUTURE TECHNOLOGY TREND







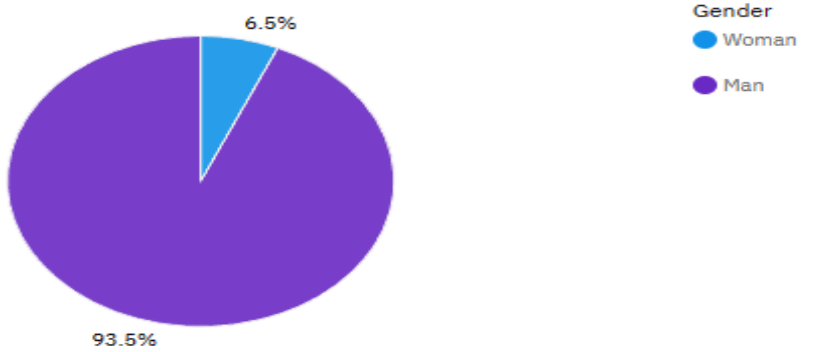


IBM Developer

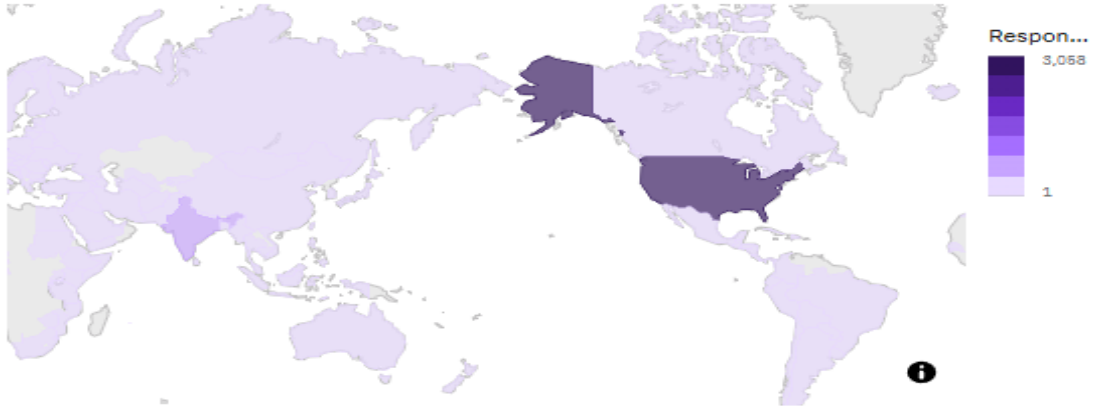
SKILLS NETWORK

DEMOGRAPHICS

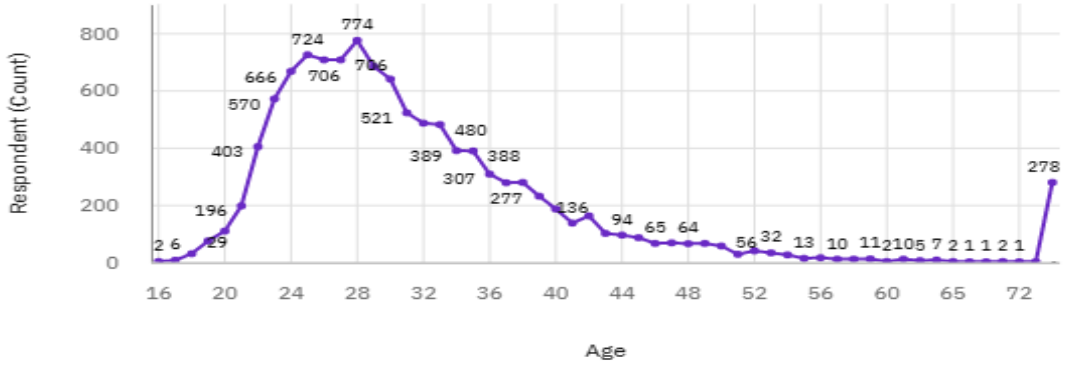
Respondent by Gender



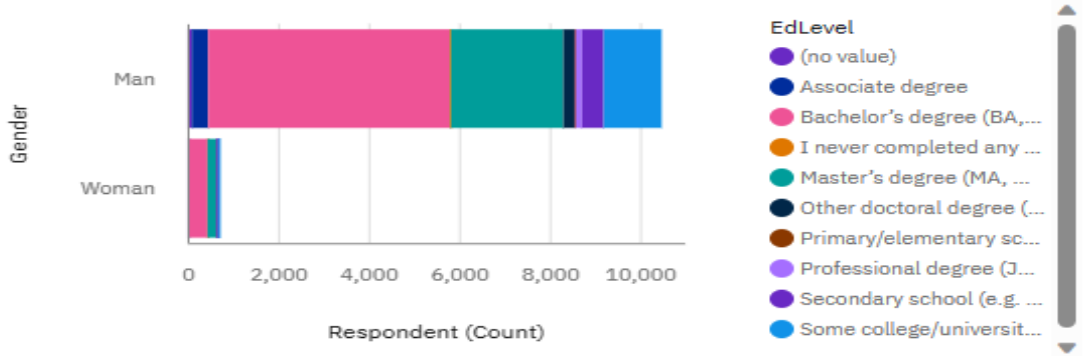
Respondent for Country regions



Respondent by Age



Respondent Count by Gender, classified by Formal Education Level



DISCUSSION

- The survey recorded 11,398 respondents, the majority of whom were developers by profession.
- The mean age of respondents was 30 years old, with most falling between the age range of 20-40 years old.
- The technology field is dominated by males with a ratio of 93.5% to 6.5% females.
- Most professionals have at least a degree for educational level.
- Compensation has a positive correlation with age for the age range of 25-30 years old, and the normalized annual salary of respondents is USD 103,000.
- The discussion then shifts towards technology trends and the future, training and reskilling workers, increasing female participation in technology, bridging the technology gap in developing countries, and eliminating age and education discrimination in employment.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Web page interactive languages are still in demand, with Python being the current trend.
- SQL databases are still in demand and widely used.
- Linux is the preferred platform for development and deployment.
- React.js is a future trend for web frameworks, while Docker and AWS are growing in popularity. Technology jobs are concentrated in countries like the USA and India, but a gender gap persists.

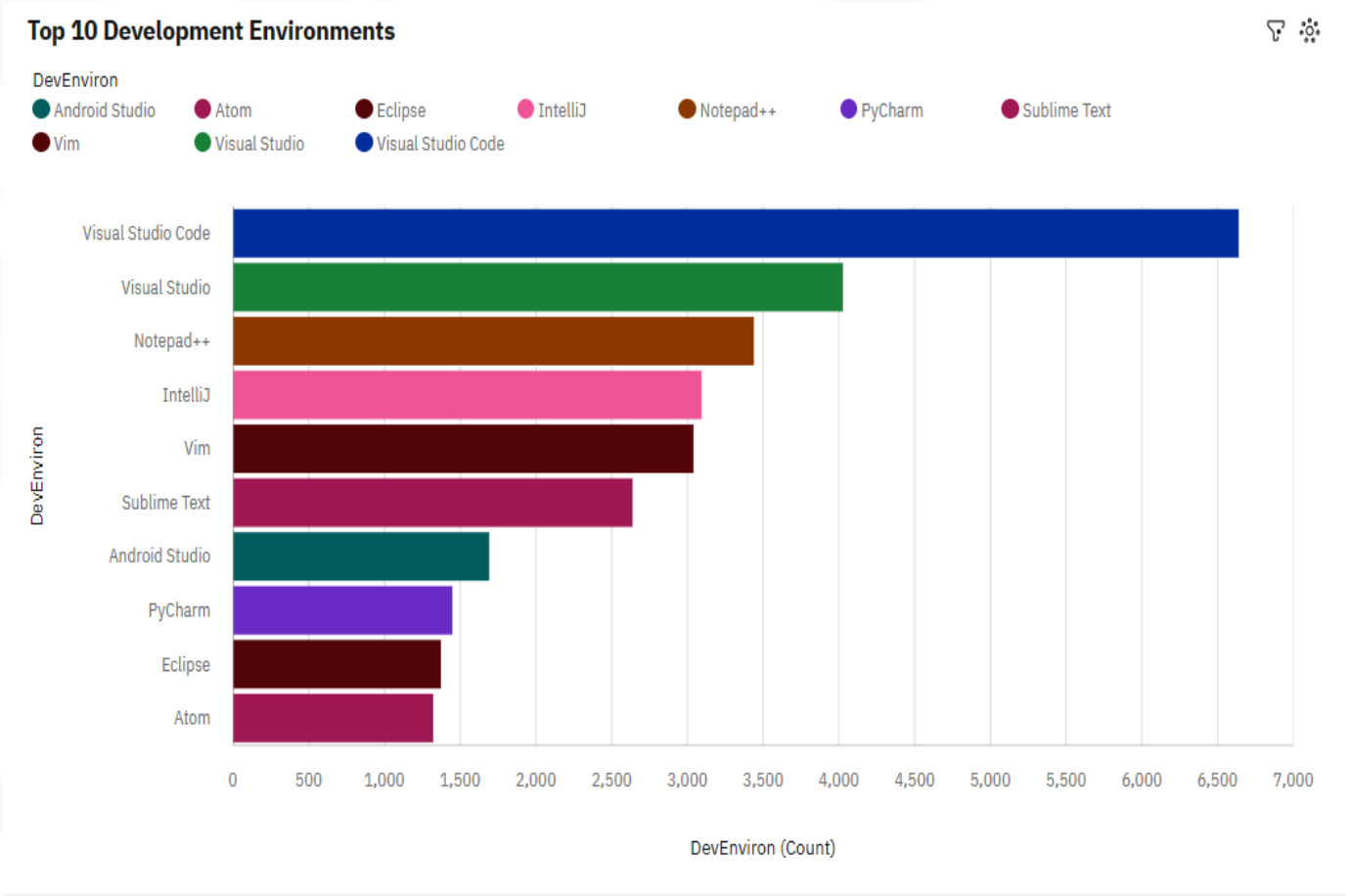
Implications

- Digitalization continues to drive demand for web page interactive languages.
- The trend towards using Python is growing, and companies should consider adopting it.
- Companies should be open to adopting new technologies and frameworks as professionals migrate away from legacy platforms and frameworks.
- The shift to faster app deployments and cloud services in the future requires companies to be flexible and adaptable, while also ensuring equal access to technology and job opportunities for everyone.

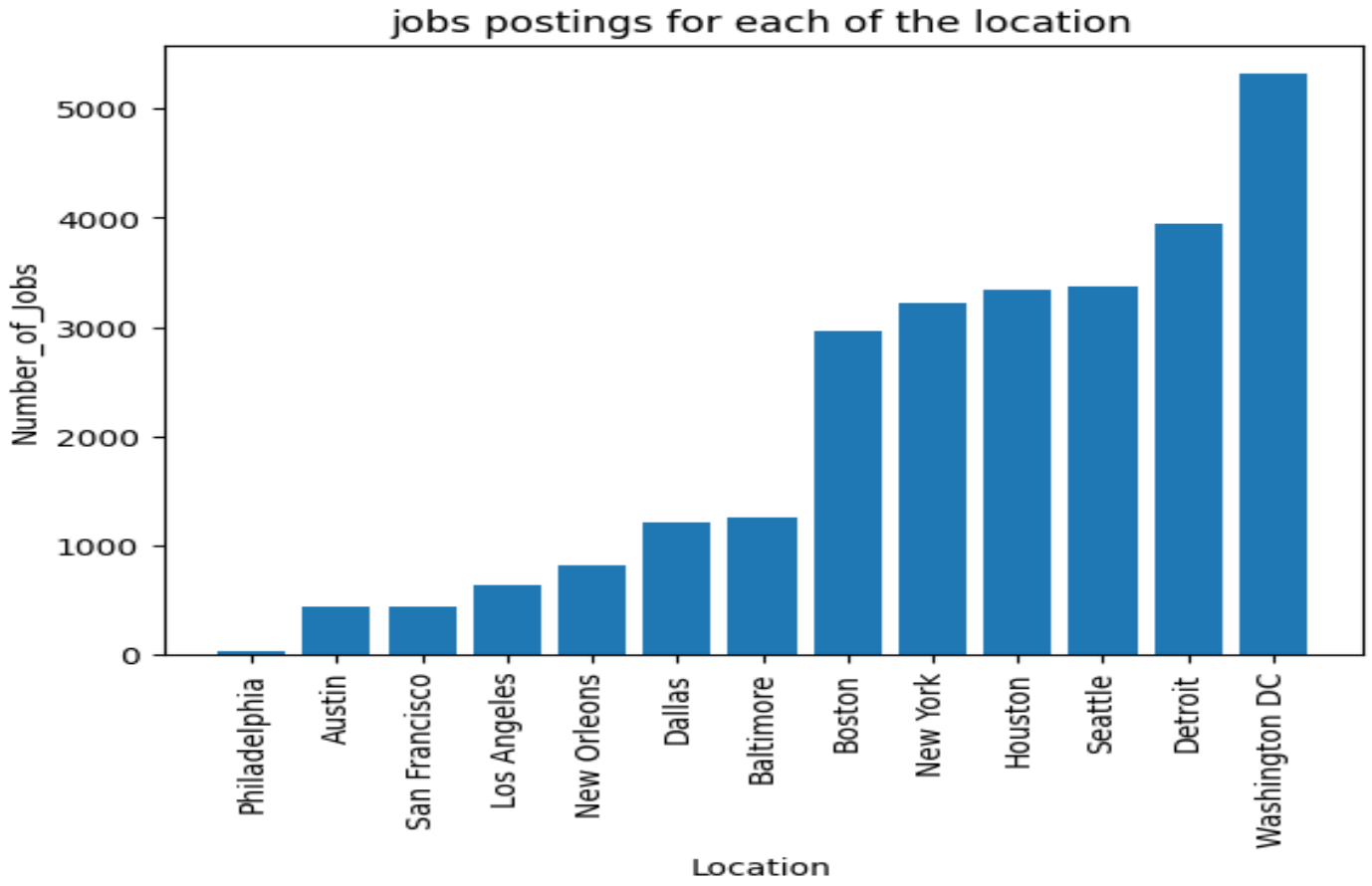
CONCLUSION

- Demand for Javascript and HTML/CSS remains high, with good pay for professionals.
- Python offers high salaries and job opportunities.
- Professionals in the field need to adapt to changing platforms and frameworks.
- Gender diversity is needed in the technology field. Machine learning can help predict trends and salaries. An overview of technology trends, programming languages, databases, and demographics can inform actions to be taken.

APPENDIX


Top 10 Development Environments
DevEnviron
Android Studio
Atom
Eclipse
IntelliJ
Notepad++
PyCharm
Sublime Text
Vim
Visual Studio
Visual Studio Code
DevEnviron
Visual Studio Code
Visual Studio
Notepad++
IntelliJ
Vim
Sublime Text
Android Studio
PyCharm
Eclipse
Atom
0
500
1,000
1,500
2,000
2,500
3,000
3,500
4,000
4,500
5,000
5,500
6,000
6,500
7,000
DevEnviron (Count)

JOB POSTINGS



POPULAR LANGUAGES