Stack Overflow Online Survey Analysis

Ibrahim Abbas

7/Jan/2024

IBM Developer

SKILLS NETWORK

TABLE OF CONTENTS

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

EXECUTIVE SUMMARY

- Collecting and exploring the data
- Cleaning the data
  - Removing Duplicates
  - Dealing with missing values
  - Normalizing the data
- Analyzing the data
- Visualizing the data
- Building the dashboard

INTRODUCTION

- An online survey was conducted and it asked software professionals across the world technical-related questions
- In this report we'll discover this data and ask questions that will be answered through the analysis of the data
- The questions to be asked:
  - What is the Gender, Age, Country, and education level of the respondents?
  - What database and programming language do the respondents mostly use?
  - What is the most desired database and programming language among the respondents
- Finally we'll discuss the:
  - Findings
  - Implications

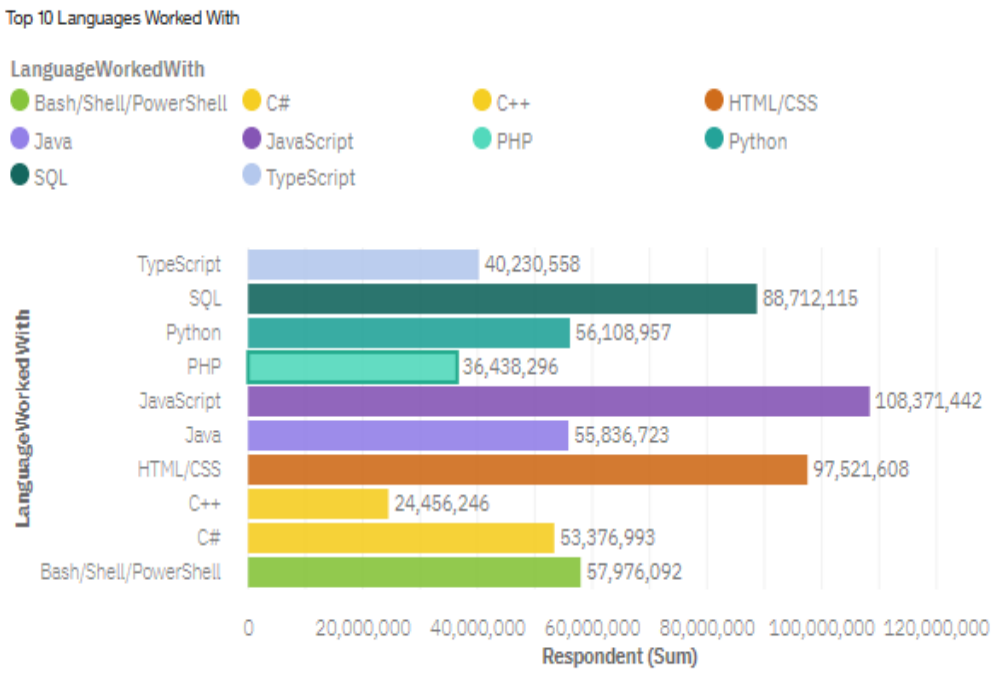
METHODOLOGY

- Data Collecting
  - The data was extracted from the Stack Overflow website and was read in Python by Jupyter Notebook
- Data Cleaning
  - Duplicate rows were identified and removed using the pandas method drop_duplicates
  - Missing values were identified and replaced by the mode
  - Certain columns were normalized
- Data analysis
  - The distribution of the data was visualized using a histogram
  - Outliers were identified and dealt with using the interquartile range
  - Correlation among the columns was identified
- Data Visualization and Dashboard
  - The data was extracted from an SQL server and was visualized
  - A dashboard was built

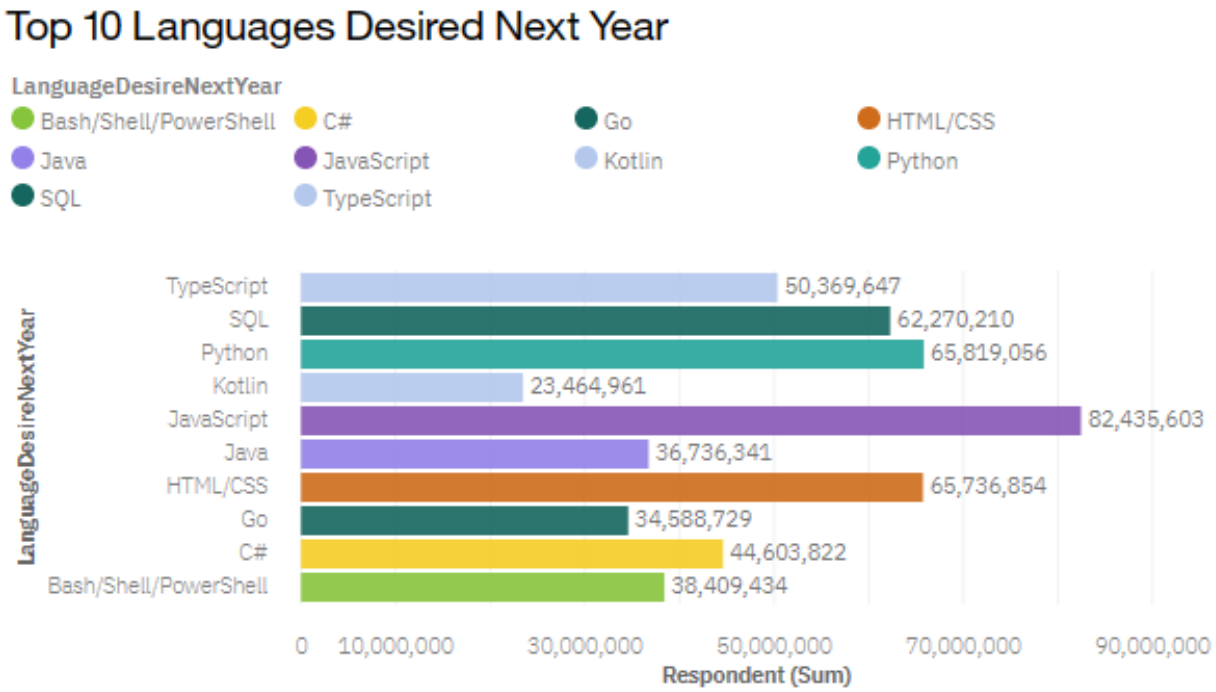
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

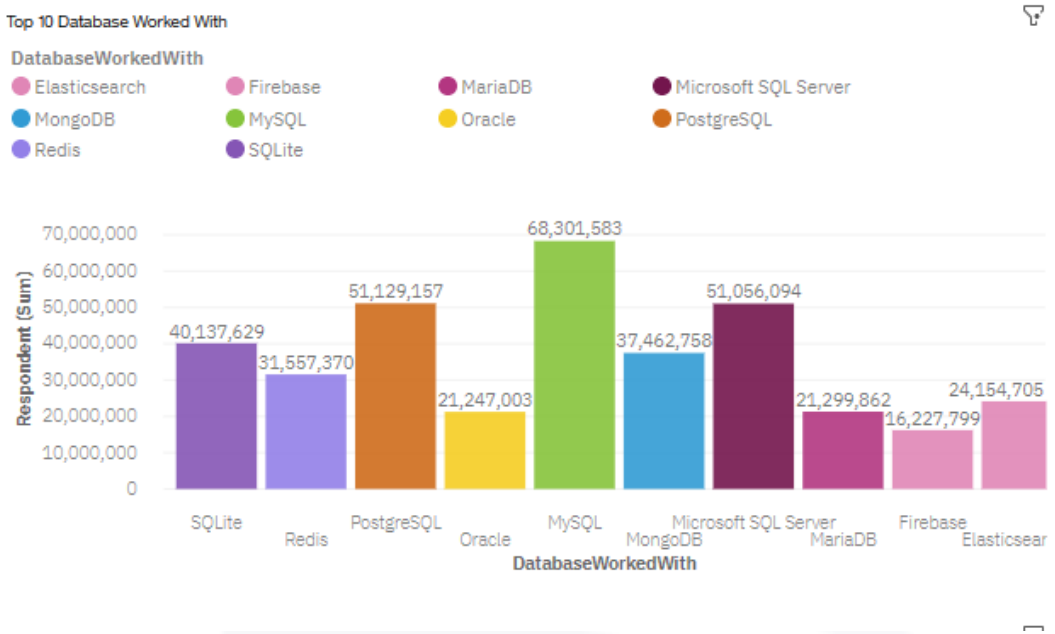
- JavaScript is the most common language worked with and desired, thus having it is a plus
- Although few respondents have little experience with Python, most of them chose it as their desired language to learn
- Many people know the Bash language but very few chose it as a desired language to learn

Implications

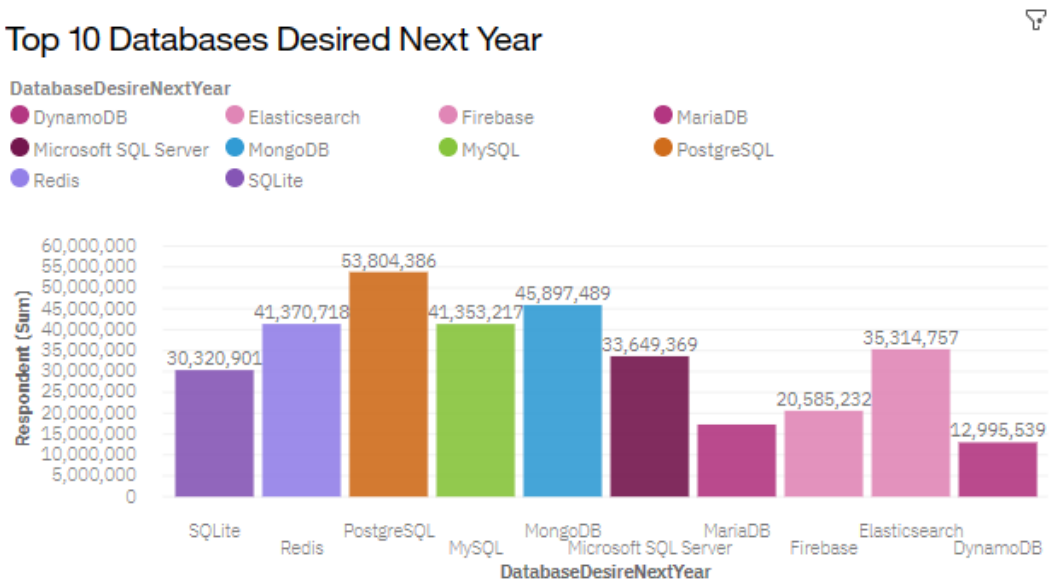
- Not so many are interested in the Bash language thus a decline in Linux users
- A lot of developers are migrating from JavaScript to Python

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Most companies use MySQL but most respondents prefer not to work with it
- PostgreSQL is the most desired database to learn
- A lot of people would like to learn MongoDB but few work with it

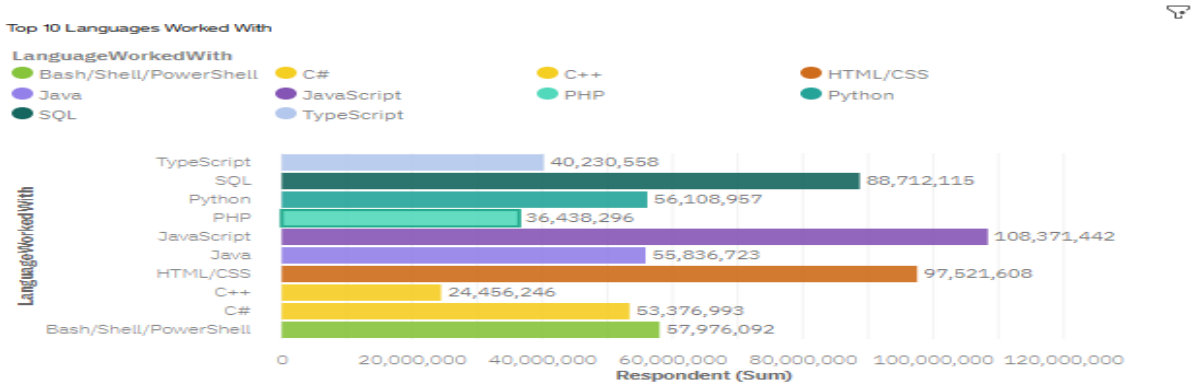
Implications

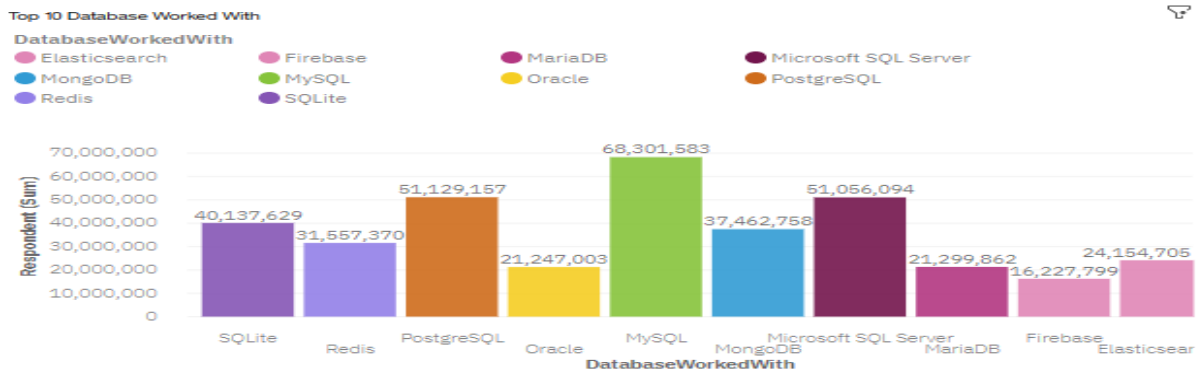
- Institutes would face a lack of MySQL developers since not many are willing to learn it
- Microsoft SQL Server is going to lose popularity because not so many are willing to learn it

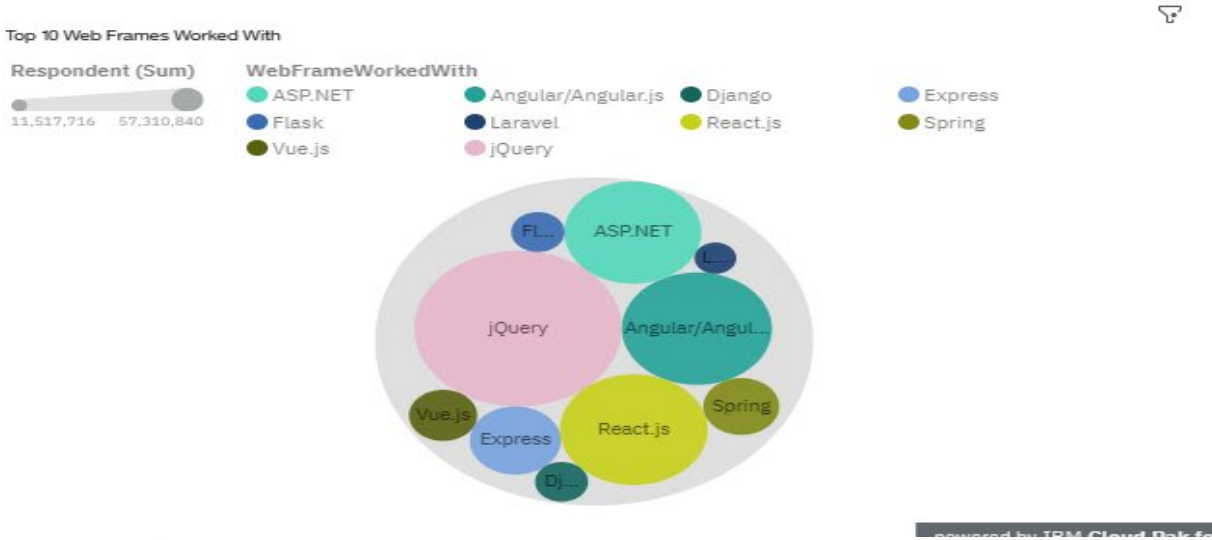
DASHBOARD

[Stack Overflow Survey Dashboard]

CURRENT TECHNOLOGY USAGE



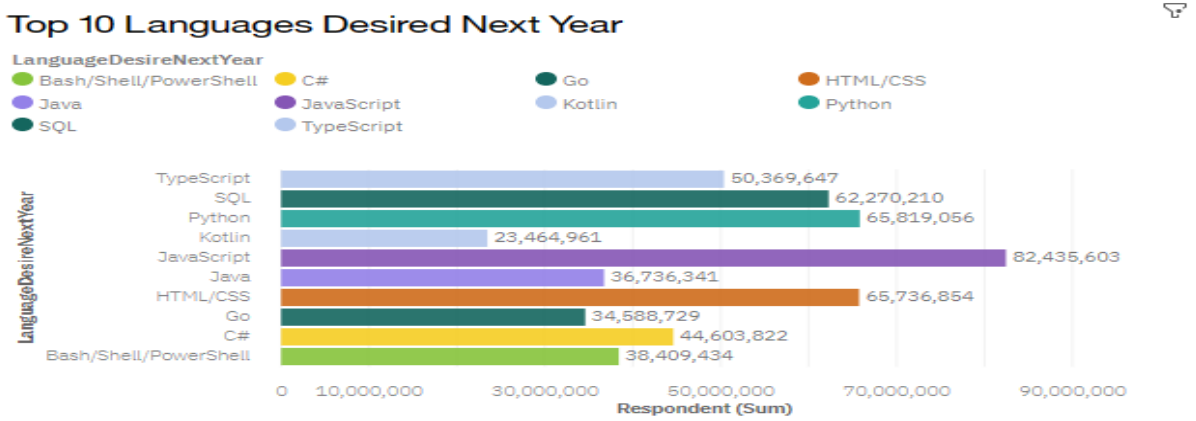


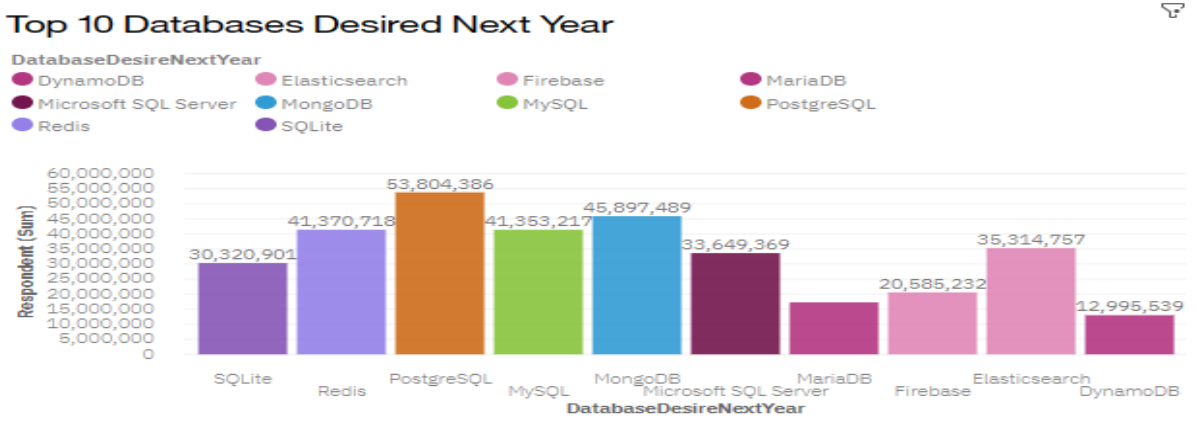


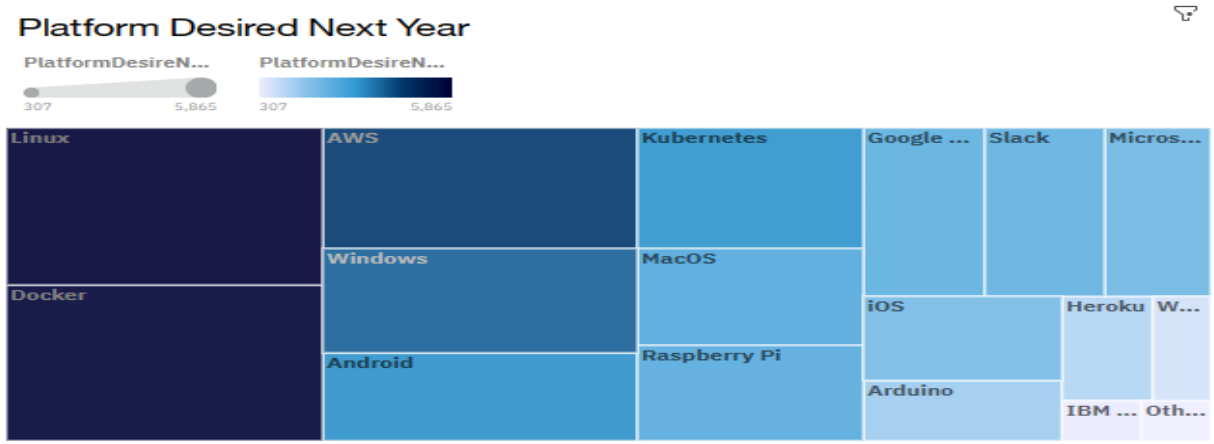
IBM Developer

SKILLS NETWORK

FUTURE TECHNOLOGY TREND



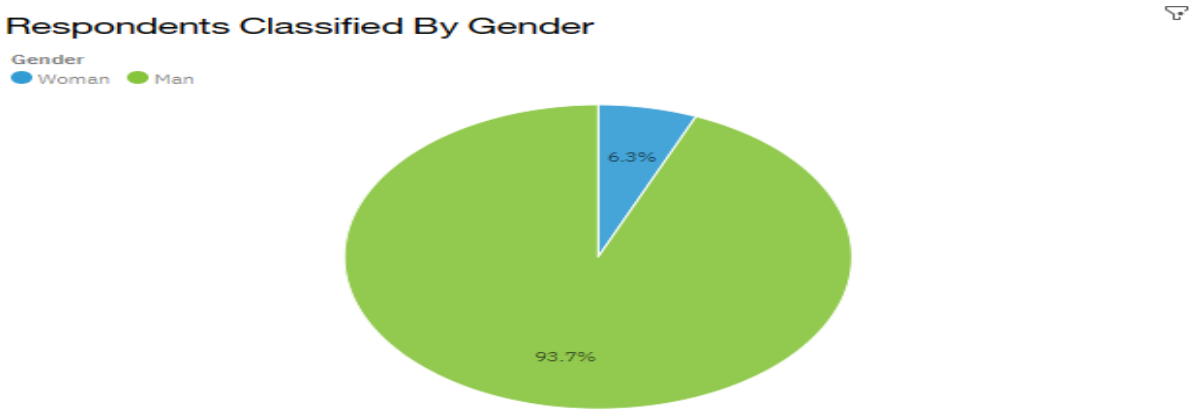


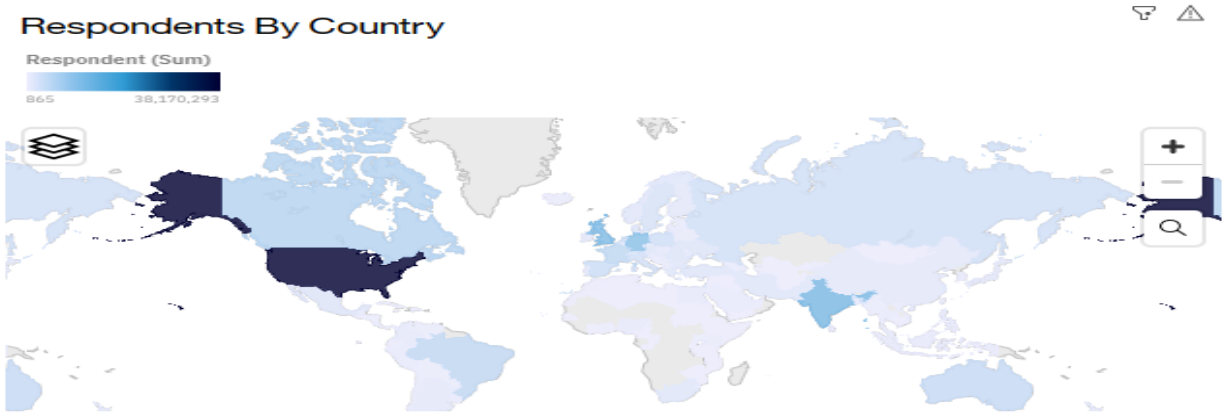


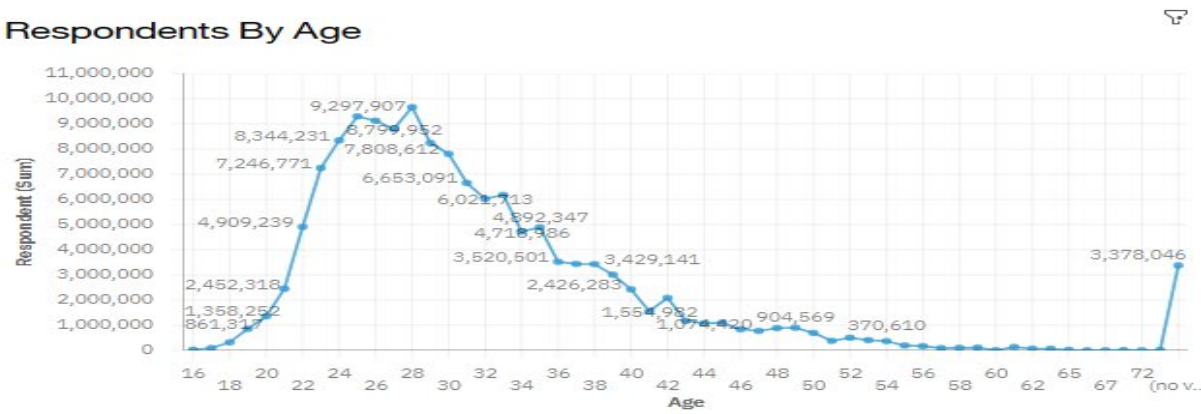
IBM Developer

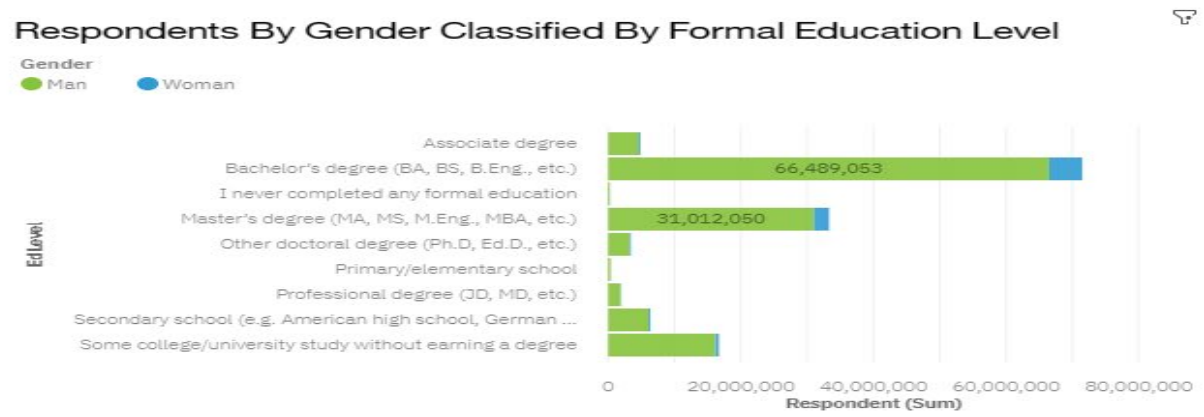
SKILLS NETWORK

DEMOGRAPHICS









DISCUSSION

OVERALL FINDINGS & IMPLICATIONS

Findings

- Python and JavaScript are going to dominate the market
- PostgreSQL is keeping up with the pace as it’s the most desired database language to learn
- Most respondents are male around 20-40 of age

Implications

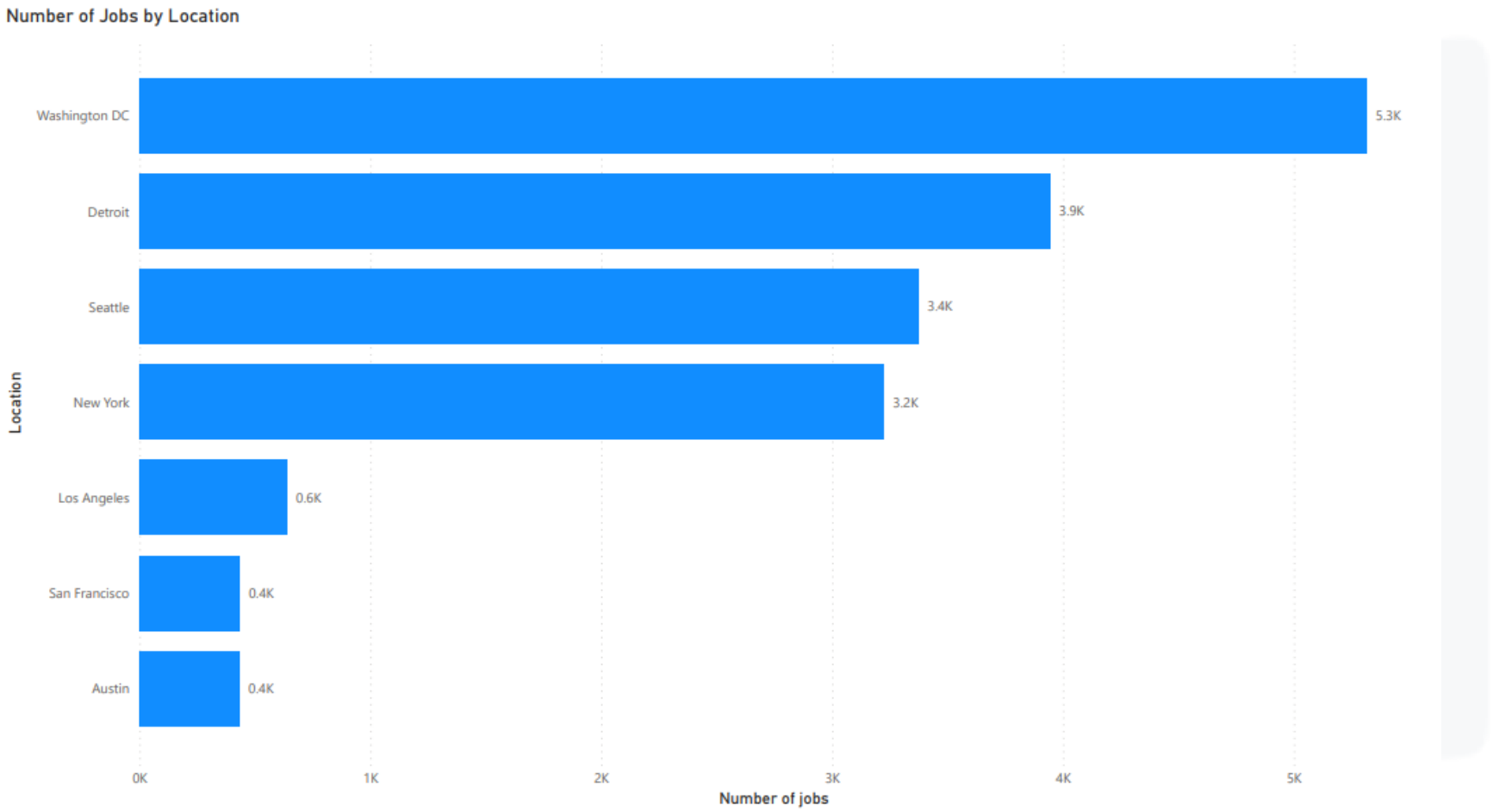
- Microsoft SQL Server might leave the market very soon
- The Bash language isn’t desirable
- Linux and Windows are dominating the OS market

CONCLUSION

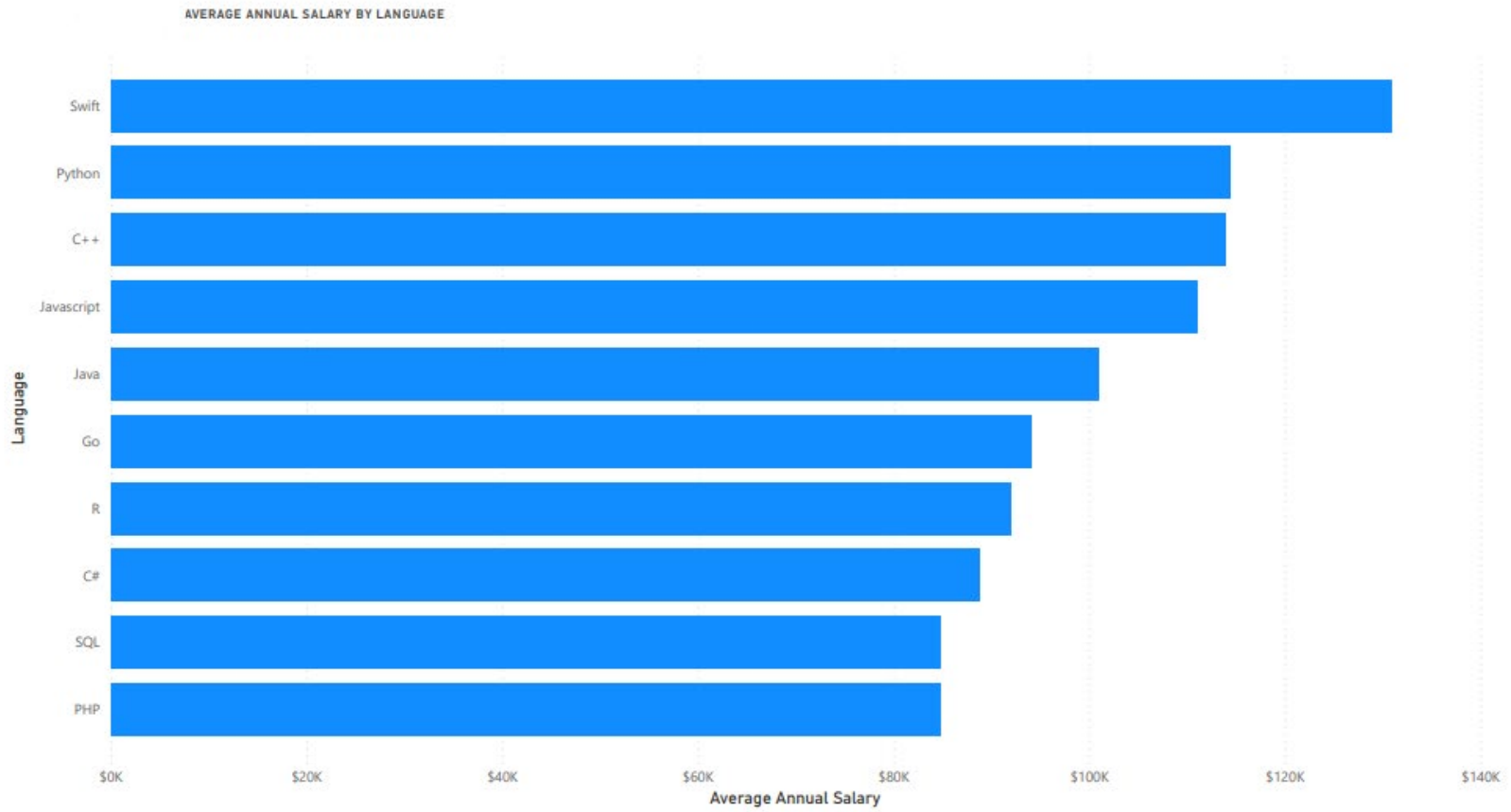
- Learning Python and JavaScript is a winning investment
- PostgreSQL has proven to be efficient and desirable
- Linux and Windows are the most favorable OS for developers
- MongoDB will have a bright future

APPENDIX AND REFERENCES

JOB POSTINGS



POPULAR LANGUAGES



REFERENCES

- Data Sources:
  - Stack Overflow Survey
  - Popular Languages