Stack Overflow Survey 2019 Analysis

Anh Hoang

April 22nd, 2024

OUTLINE

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

EXECUTIVE SUMMARY

- In 2019 Stack Overflow users were given a survey related to their experiences with software development.
- The survey questioning was broken down into 3 categories:
  - Current Technology Usage
  - Future Technology Usage
  - Demographics
- At the time of the survey, web development – like JavaScript, HTML, and CSS languages are highly used.
- Post-2019, web development interest is also strong, but users are also looking towards Python to learn.
- Most responders are men located in the United States and India.

INTRODUCTION

- Stack Overflow is a very popular Q&A platform for students and professionals engaged in software development.
- Software development is a vast field that encompasses web, mobile, backend, and data science. It was in 2019 and still is (as of 2024) a high-paying field experiencing faster-than-average growth in the United States and abroad.
- The survey asked questions related to the current state and future trends of various technologies within IT, usage of the site, and demographics of the respondents.
- The analysis will attempt to provide readers with an understanding of the field and help in forecasting future trends.

METHODOLOGY

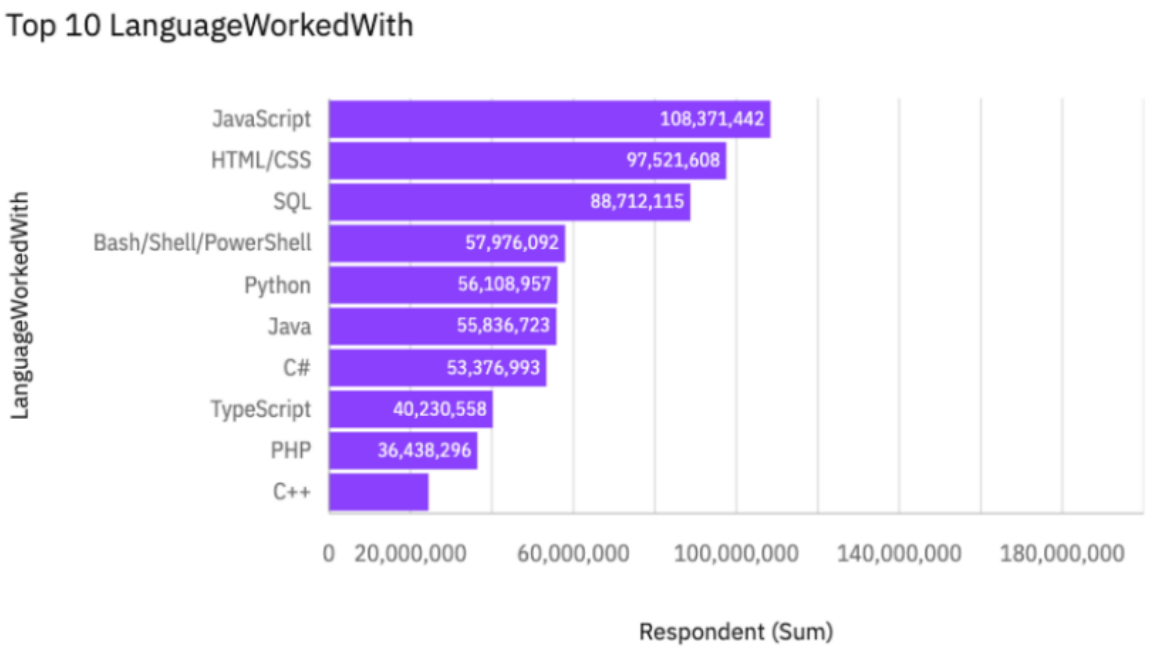
- Data Source:
  - https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DA0321EN-SkillsNetwork/LargeData/m1_survey_data.csv
- JupyterLite was used for the following and uploaded to GitHub under the below links:
- Data Cleansing:
  - https://github.com/AnhHoangDataAnalyst/Stack-Overflow-Survey-2019-Analysis/blob/main/M2DataWrangling-lab.ipynb
- Data Exploration :
  - https://github.com/AnhHoangDataAnalyst/Stack-Overflow-Survey-2019-Analysis/blob/main/M3ExploratoryDataAnalysis-lab.ipynb
- Data Visualization
  - https://github.com/AnhHoangDataAnalyst/Stack-Overflow-Survey-2019-Analysis/blob/main/M3ExploratoryDataAnalysis-lab.ipynb
- IBM Cognos Analytics was used to create dashboards

RESULTS

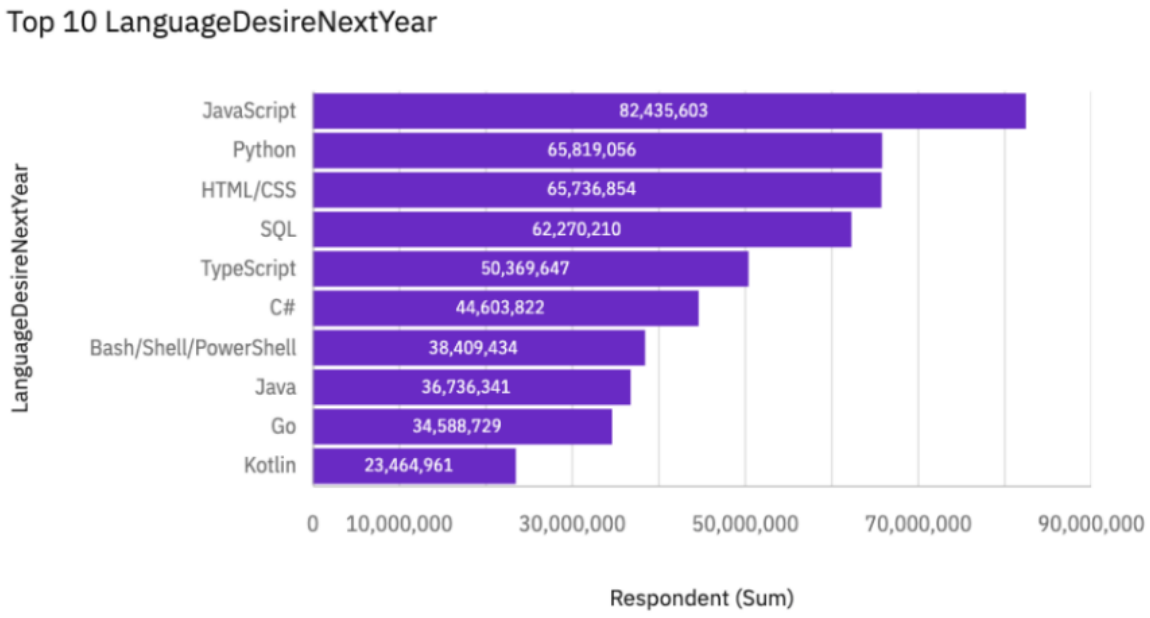
- PROGRAMMING LANGUAGE TRENDS
- DATABASE TRENDS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

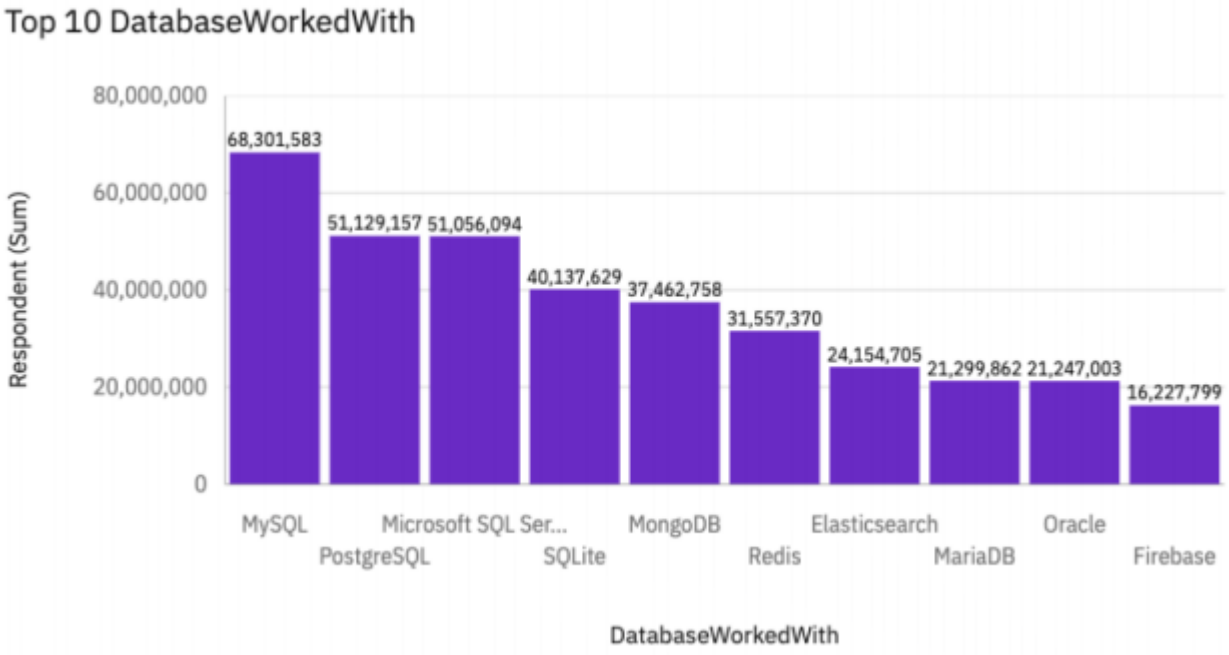
- JavaScript is the most popular to use currently and in the future.
- HTML/CSS is losing its popularity
- Python has a rising trend and is likely to become the second most popular programming language.

Implications

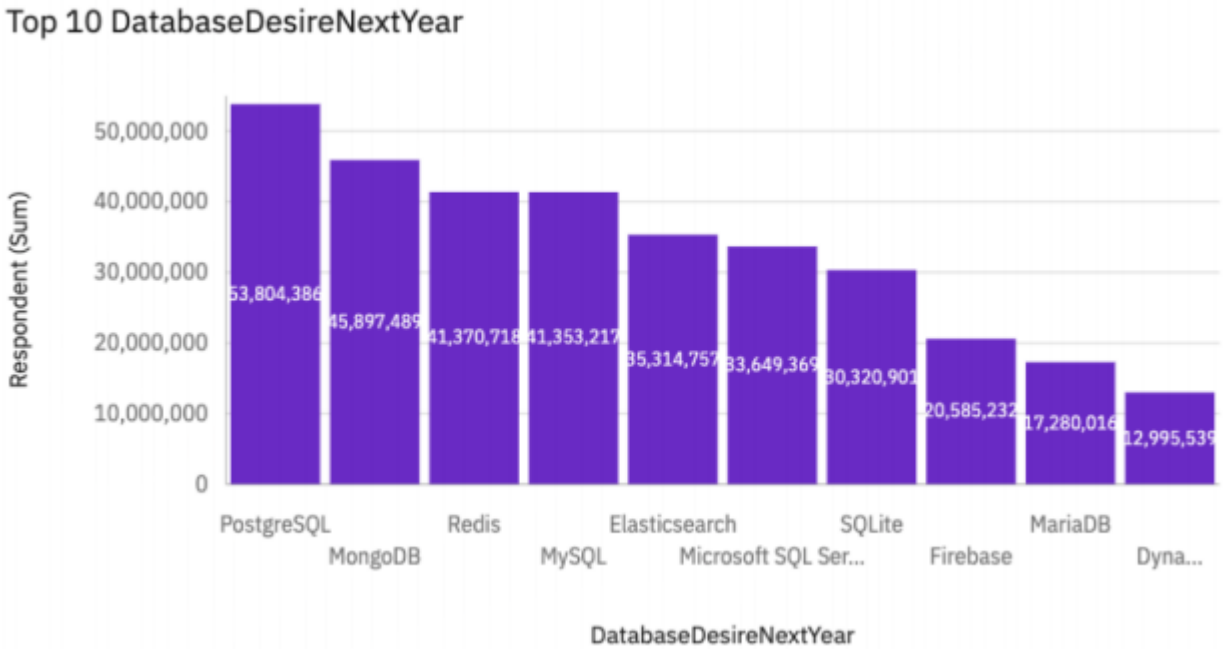
- JavaScript and HTML / CSS have current demand now and in the future
- Python is catching up in demand and will have even more opportunities
- C++ and Kotlin are much less used which means they aren't as utilized for most jobs available compared to the most popular languages

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- In 2019 most users were operating in SQL-type databases
- Most users plan on growing their skillset in PostgreSQL databases
- SQLite will have fewer users in future.

Implications

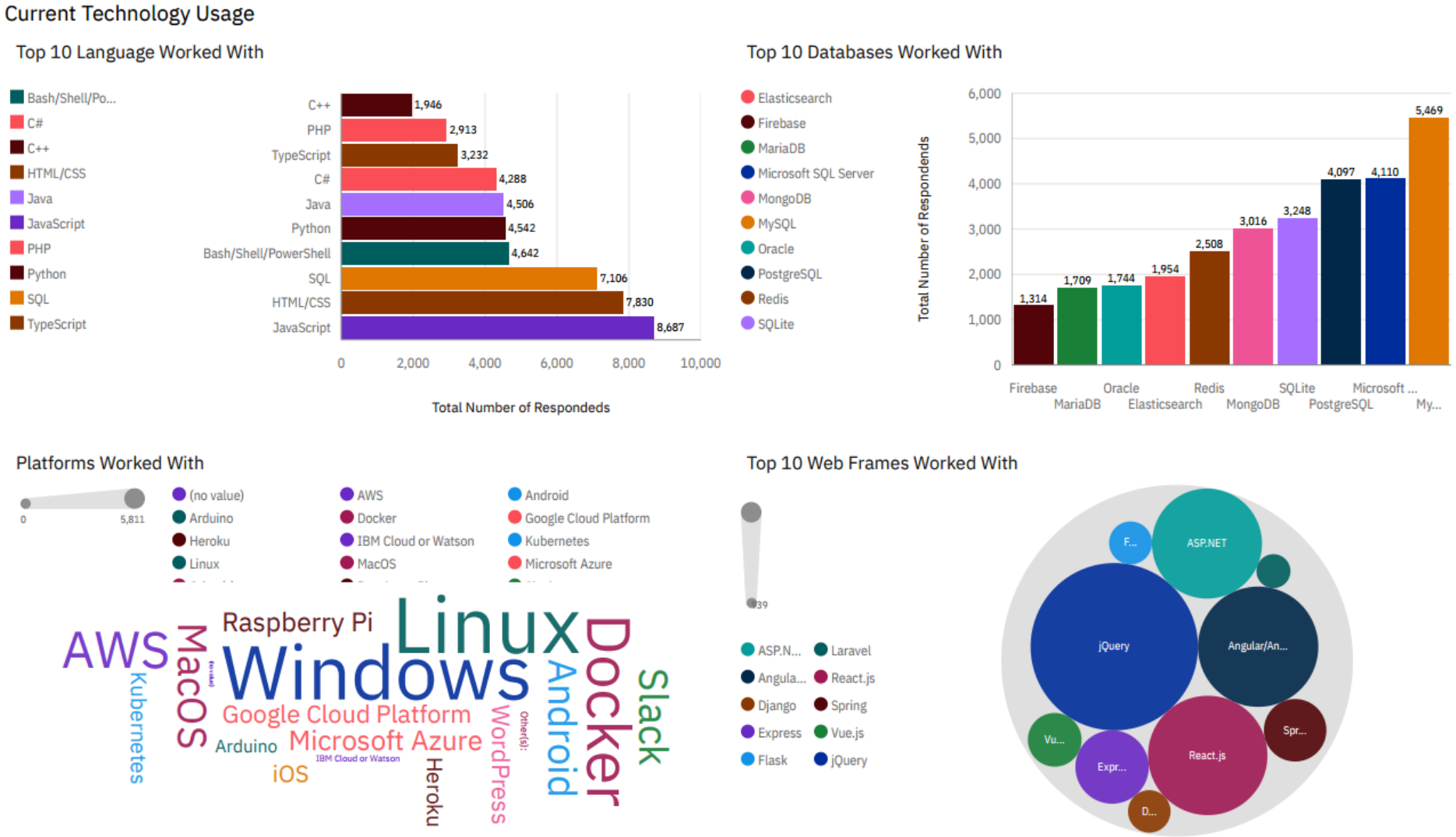
- People are very interested to learn non relational databases
- Big data will be even more dominant in the future
- Scalability and a strong open-source community are important aspects of databases of the future.

DASHBOARD

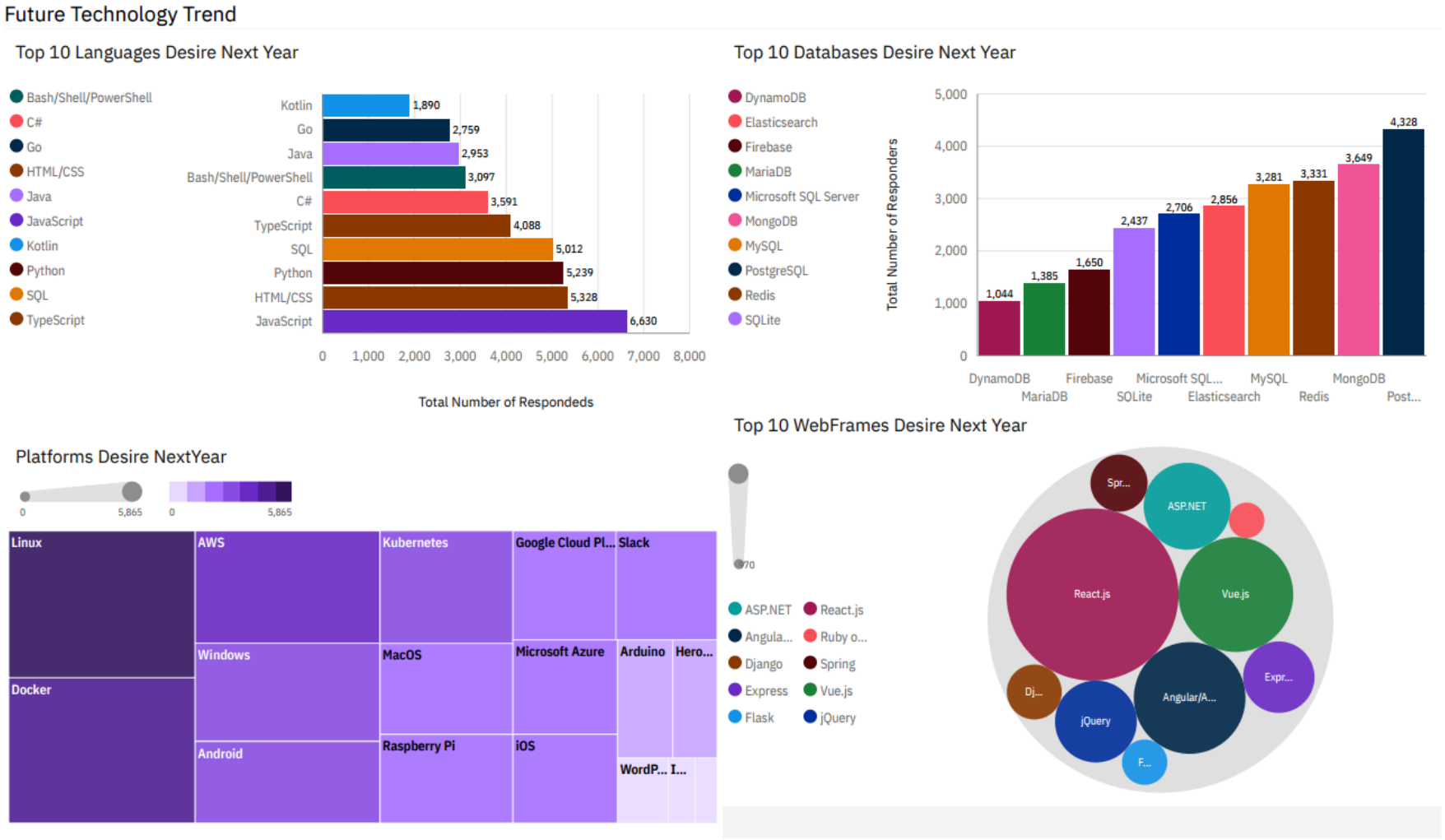
IBM Cognos Data Visualizations:

https://github.com/AnhHoangDataAnalyst/Stack-Overflow-Survey-2019-Analysis/blob/main/M5%20IBM%20COGNOS%20W5%20M9%20AH.pdf

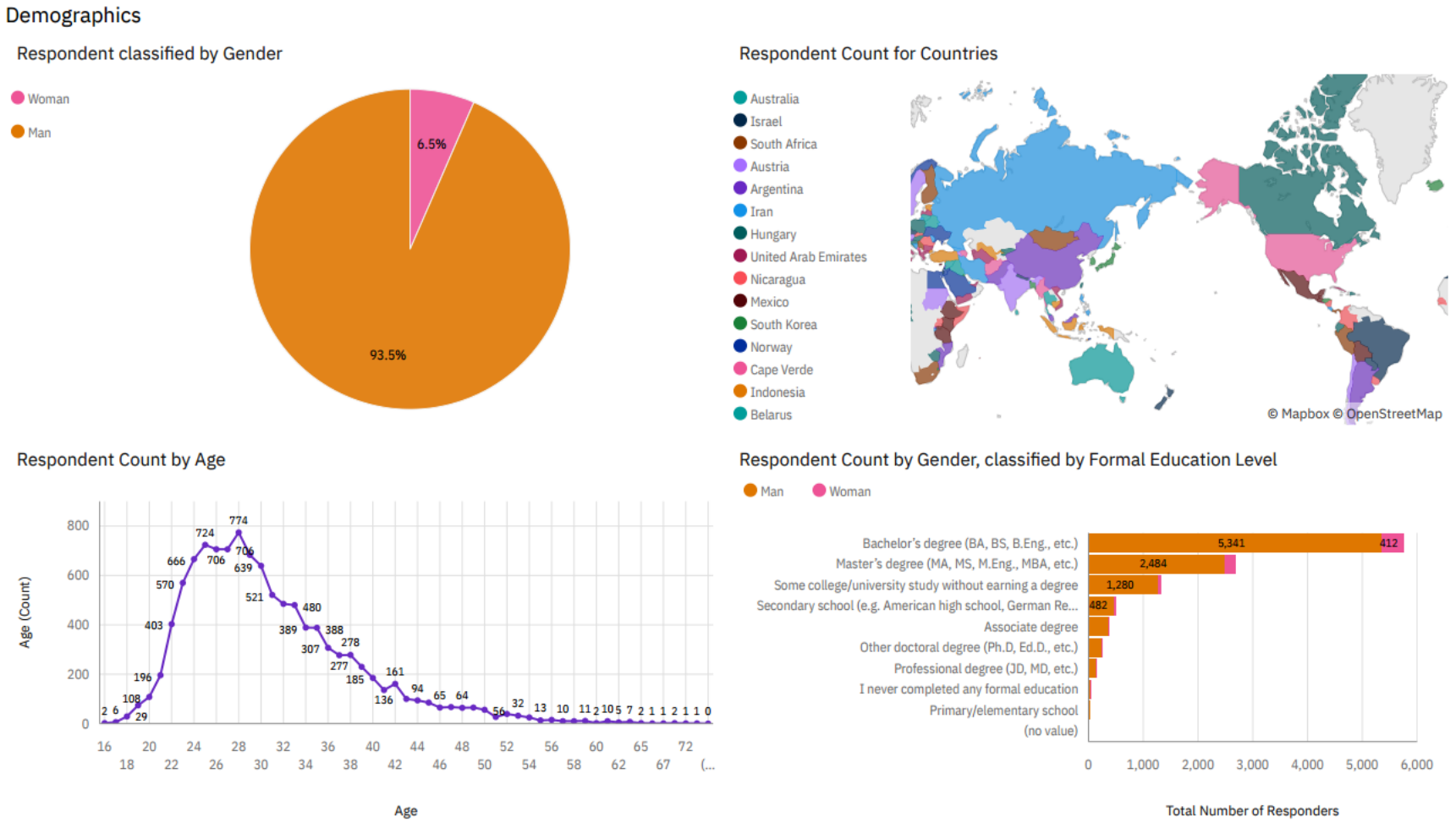
DASHBOARD TAB 1 - Current Technology Usage



DASHBOARD TAB 2 – Future Technology Trend



DASHBOARD TAB 3 - Demographics



DISCUSSION

- Where to invest your time and money for improving skills in competitive IT industry?
- Is it worth to constantly study?
- Do women have more chance in getting a job in IT?
- What are the main IT tools to focus on?

OVERALL FINDINGS & IMPLICATIONS

Findings

- Non-relational databases are the future and needed to work with unstructured data
- There are almost 10 times less women in the IT industry, so it should be diversified
- IT workers are mostly young age 20-35

Implications

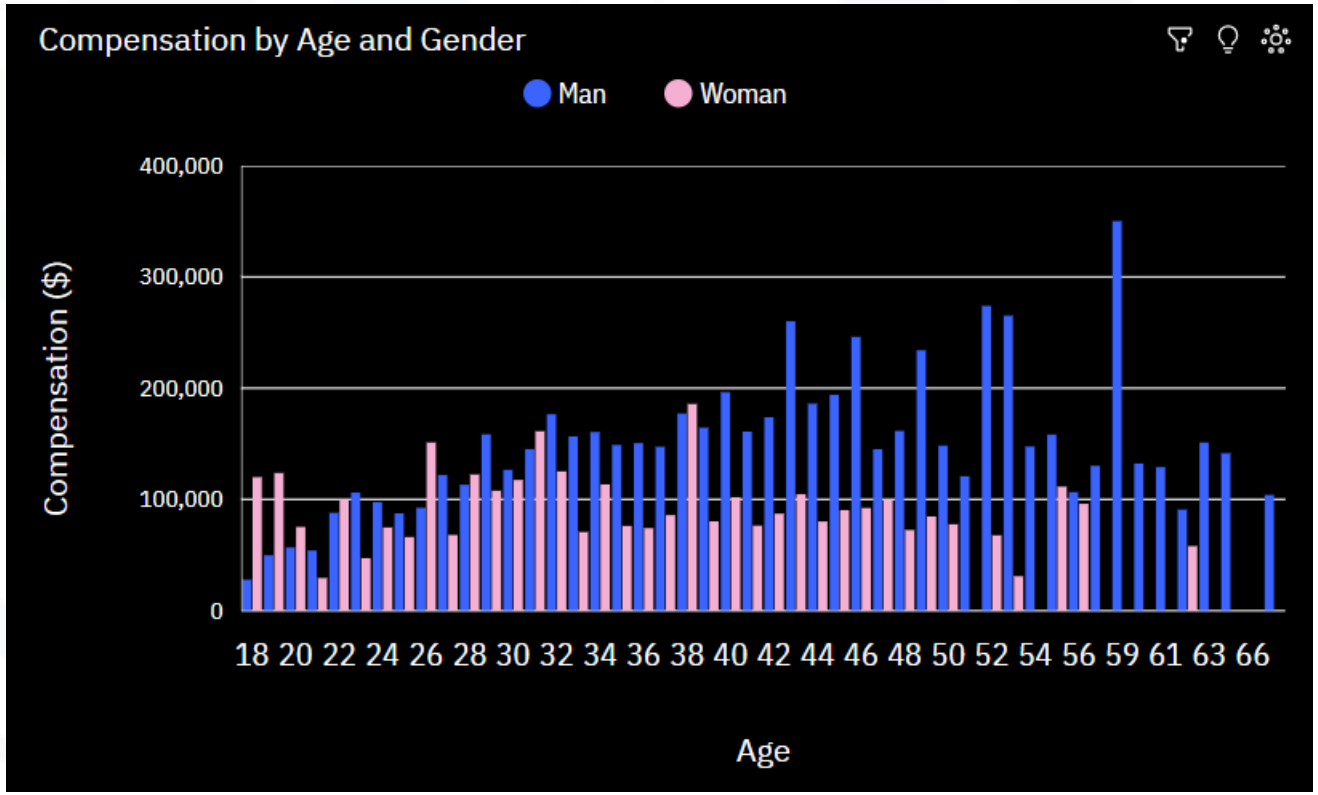
- It is worth investing in learning Python, JavaScript, HTML/CSS to be competitive in the web development industry
- Most professionals are well-educated, so it is worth getting Bachelor's or Master's degree to stay competitive.

CONCLUSION

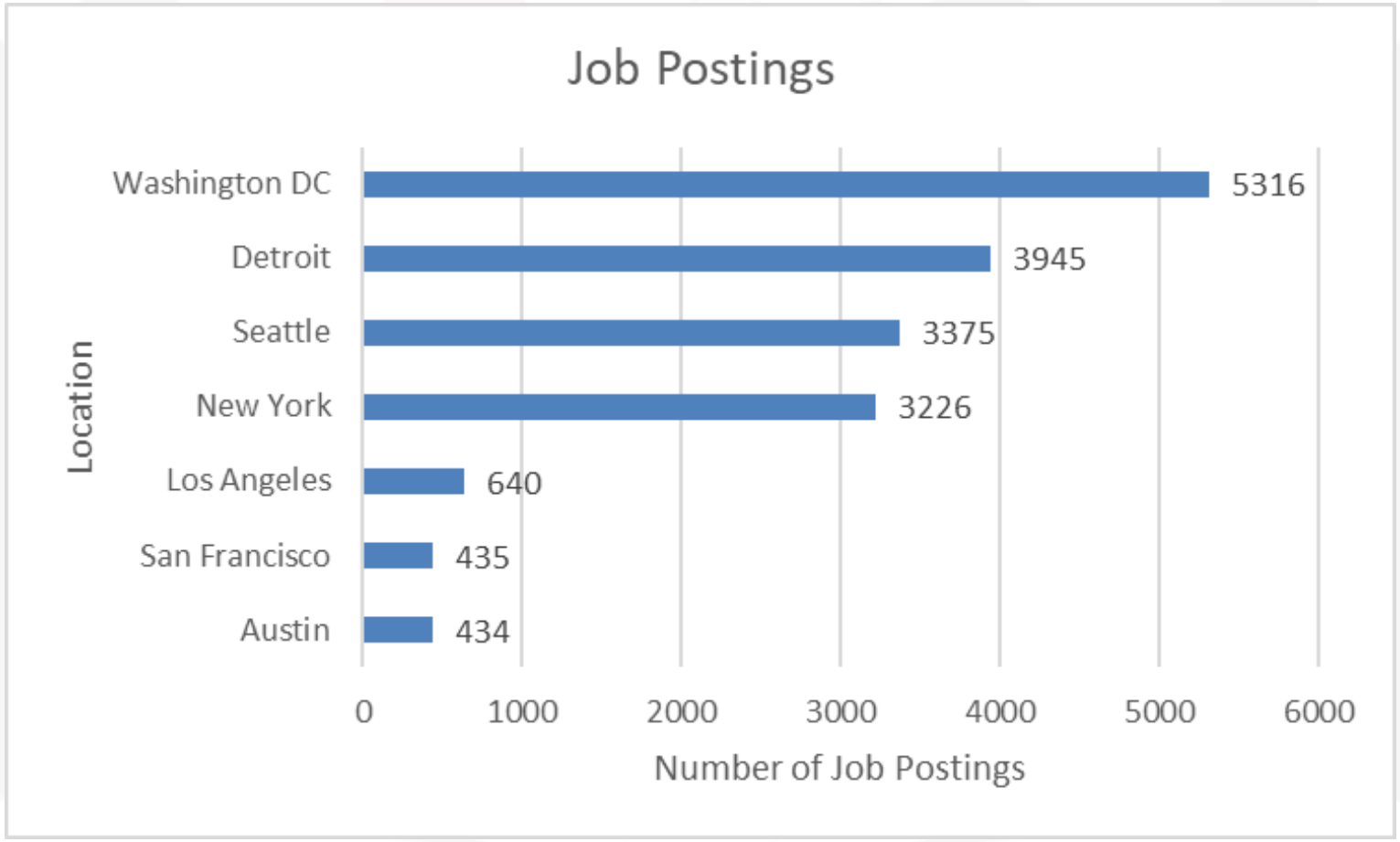
- In order to stay competitive on IT job market JavaScript, Python, non-relational
- databases, HTML/CSS, Cloud Services are must have
- Any technology related to Big Data or unstructured data is in high-demand
- There are mostly men between 22-35 from the USA working in the IT industry
- Higher Education is a base to have for the IT

APPENDIX:

Females are tend to make significantly less than males, but that is improving. What factors are causing this?



JOB POSTINGS



POPULAR LANGUAGES