



IBM Data Analyst Capstone

Jeff M.

4/7/2024

OUTLINE



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

EXECUTIVE SUMMARY



- In 2019 StackOverFlow 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.
- The vast majority of users are men located in the United States and India.

INTRODUCTION



- StackOverFlow 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 an understanding of the field and help in forecasting future trends.

METHODOLOGY



- Data Source: Subset of the following file
 - <https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/>
- JupyterLite was used for the following and uploaded to GitHub under the below links:
- Data Exploration:
 - https://github.com/jeffdg1986/coursera-py-ds/blob/Capstone/Lab5_week1_explore%20the%20dataset.ipynb
- Data Cleansing:
 - <https://github.com/jeffdg1986/coursera-py-ds/blob/Capstone/M2DataWrangling-lab.ipynb>
- Data Analysis:
 - <https://github.com/jeffdg1986/coursera-py-ds/blob/Capstone/M3ExploratoryDataAnalysis-lab.ipynb>
- Data Visualization
 - <https://github.com/jeffdg1986/coursera-py-ds/blob/Capstone/M4DataVisualization-lab.ipynb>

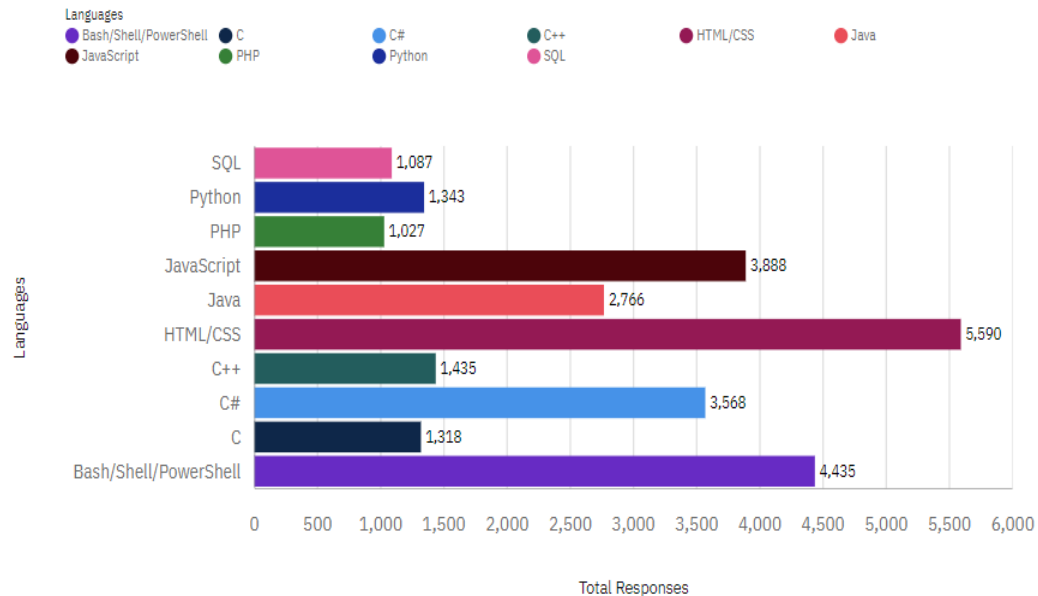
RESULTS



PROGRAMMING LANGUAGE TRENDS

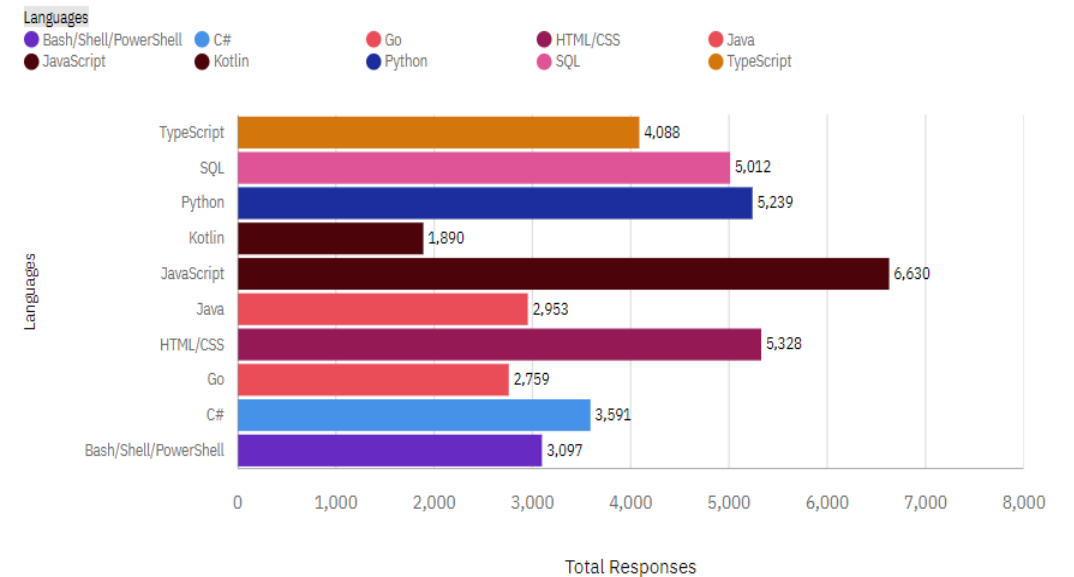
Current Year

Top 10 Languages Worked With



Next Year

Top 10 Languages Desired to Learn Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- High level programming languages are most frequently used.
- Web development languages like JavaScript and HTML are popular.
- Bash / Power Shell scripts are used often in professional settings.

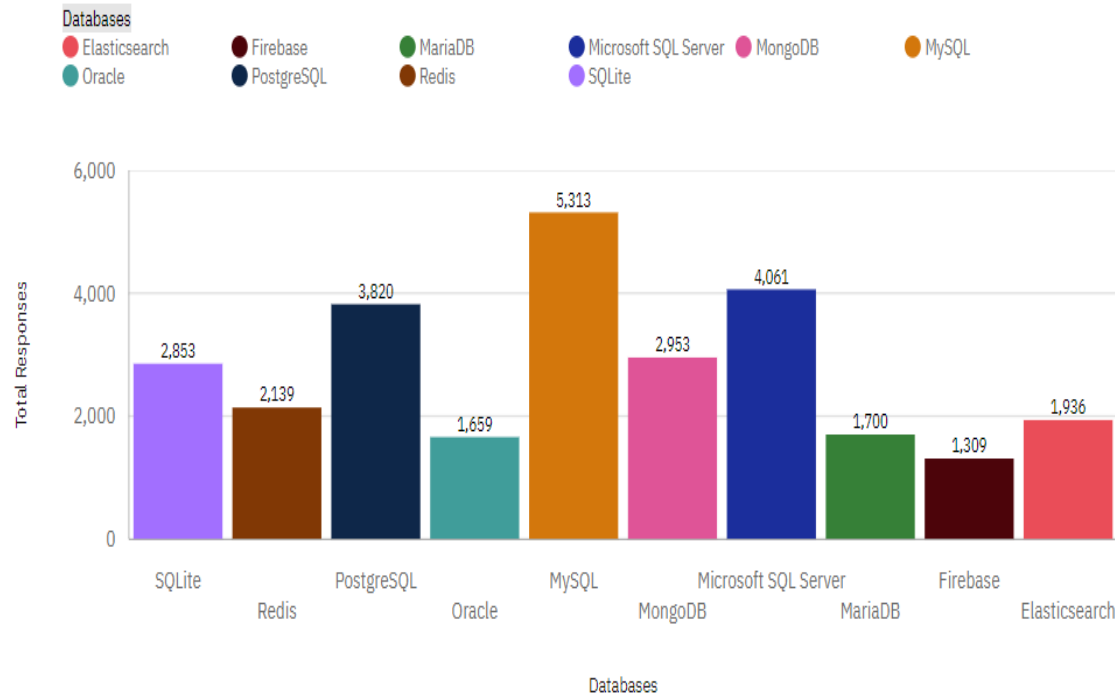
Implications

- The barrier to entry is diminishing and in the future there may be a skill shortage in low level languages like C, C++, and Assembly
- In 2019 there was a high demand for web development.
- Businesses are investing in automation.

DATABASE TRENDS

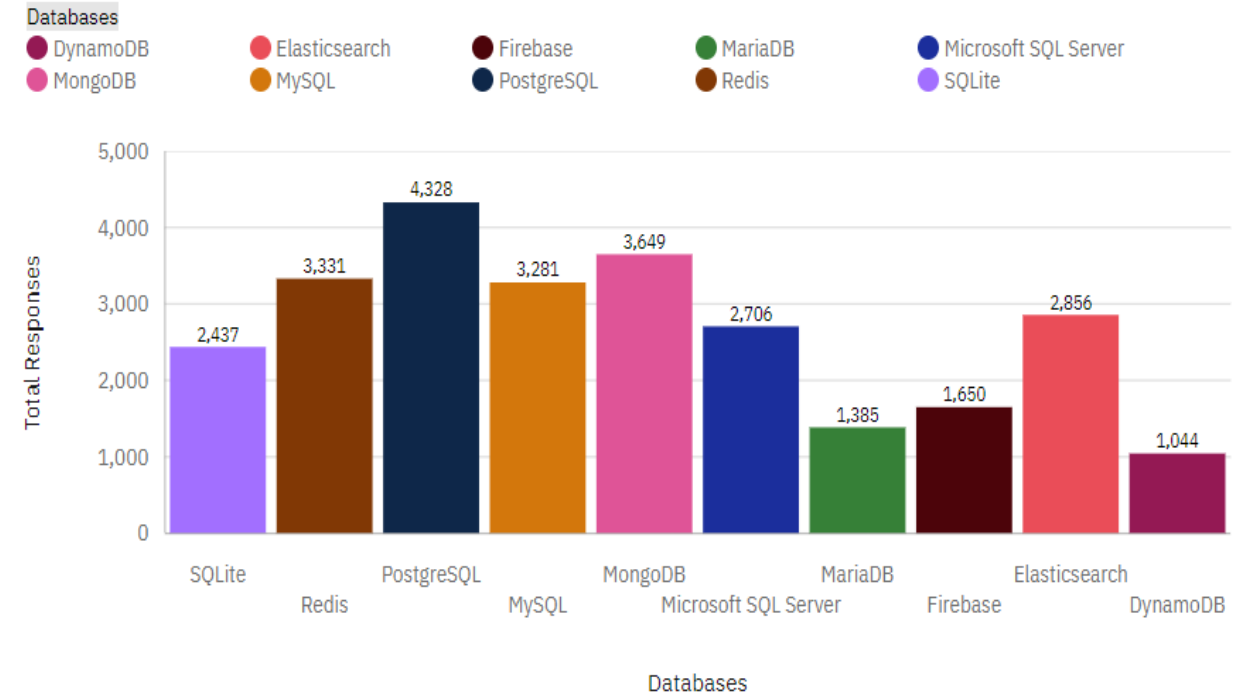
Current Year

Top 10 Databases Worked With



Next Year

Top 10 Databases Desired to Learn 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 NoSQL databases
- Amongst SQL databases, PostgreSQL interest is expected to grow in the future.

Implications

- In 2019 most business were still user 'older' technology that is more transactional based
- There's a growing demand for scalability and IOT applications
- Scalability and a strong open-source community are important aspects of databases of the future.

DASHBOARD



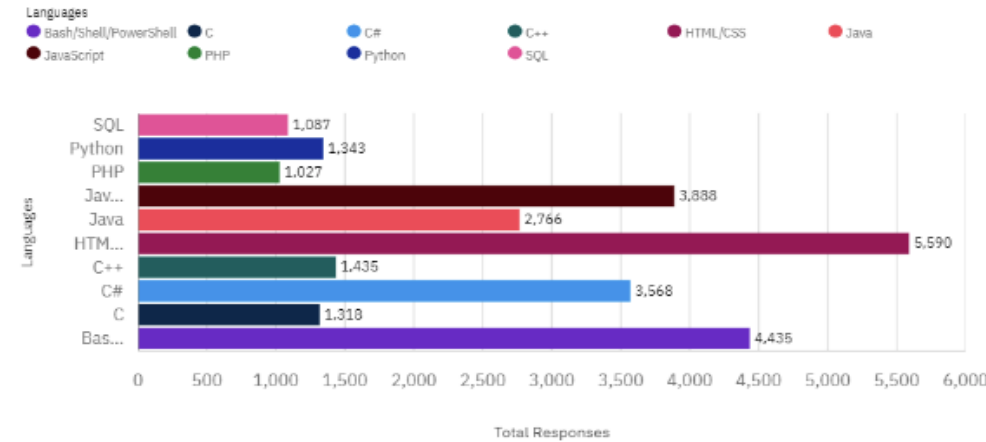
IBM Cognos Data Visualizations:

<https://github.com/jeffdg1986/coursera-pyds/blob/Capstone/Capstone%20Week%205%20Project%20-%20Cognos.pdf>

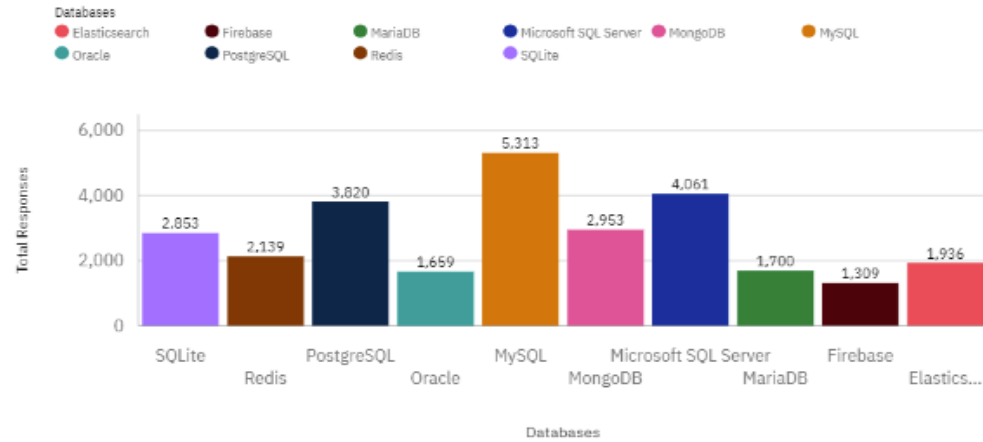
DASHBOARD TAB 1 - Current Technology Usage

Current Technology Usage

Top 10 Languages Worked With



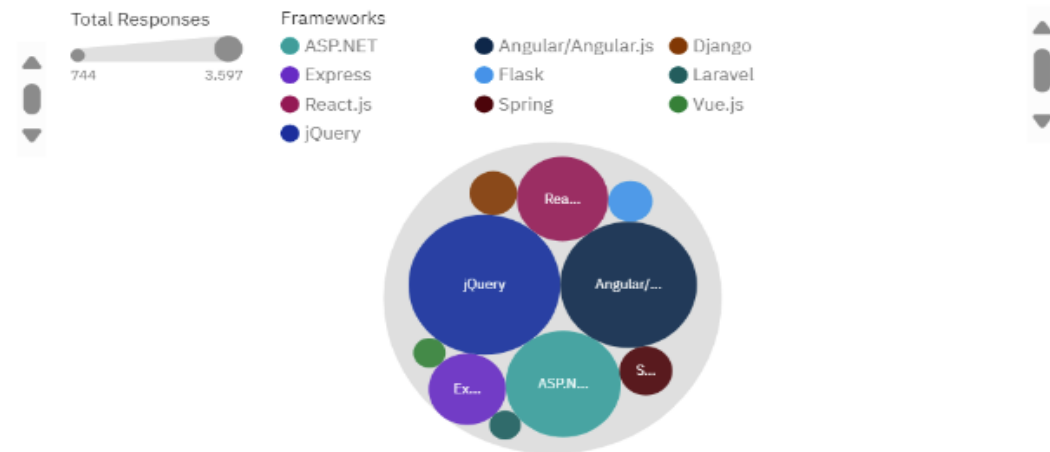
Top 10 Databases Worked With



Word Cloud: Popular Platforms Used



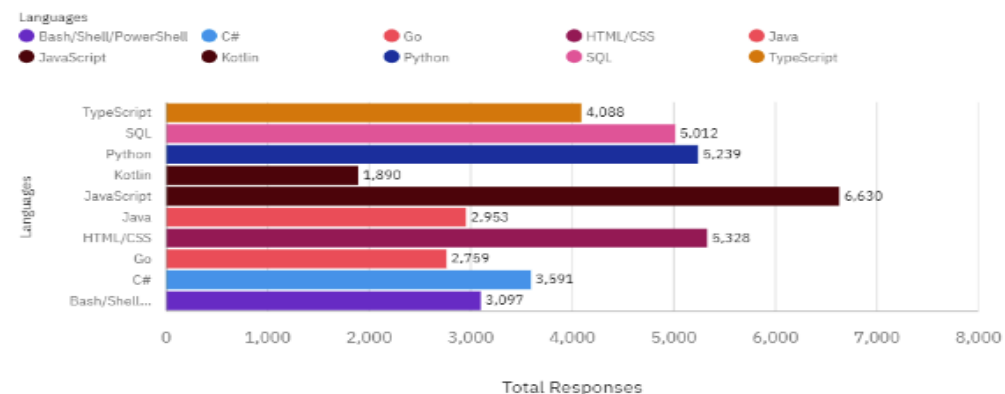
Hierarchy Map: Popular Web Frameworks Used



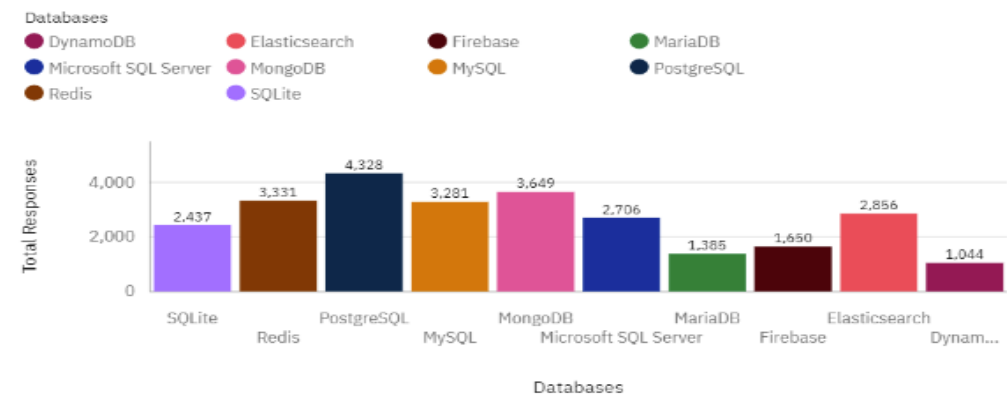
DASHBOARD TAB 2 – Future Technology Trend

Future Technology Trend

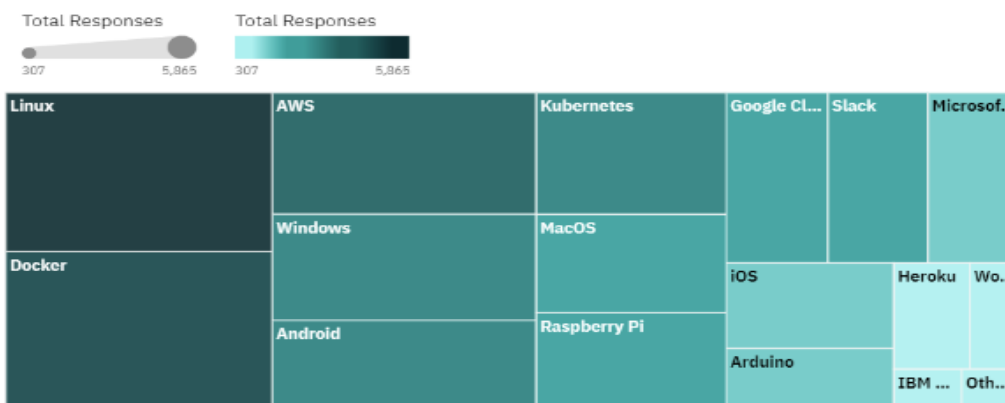
Top 10 Languages Desired to Learn Next Year



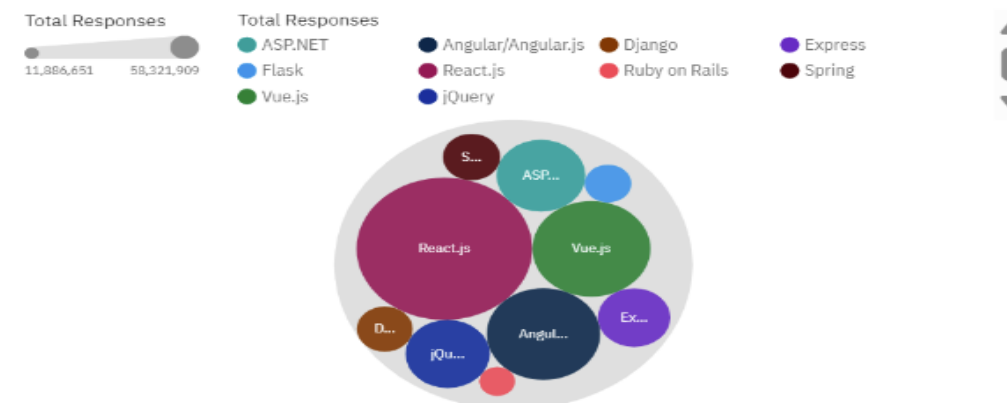
Top 10 Databases Desired to Learn Next Year



Heat Map: Platform Desired Next Year



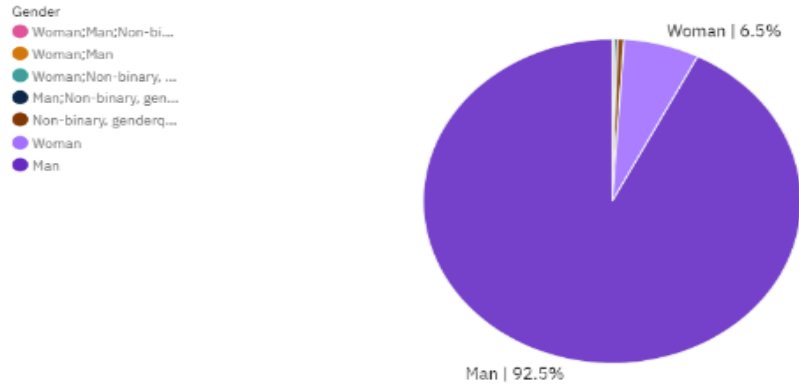
Hierarchy Bubble Chart: Top 10 Web Frameworks Desired Next Year



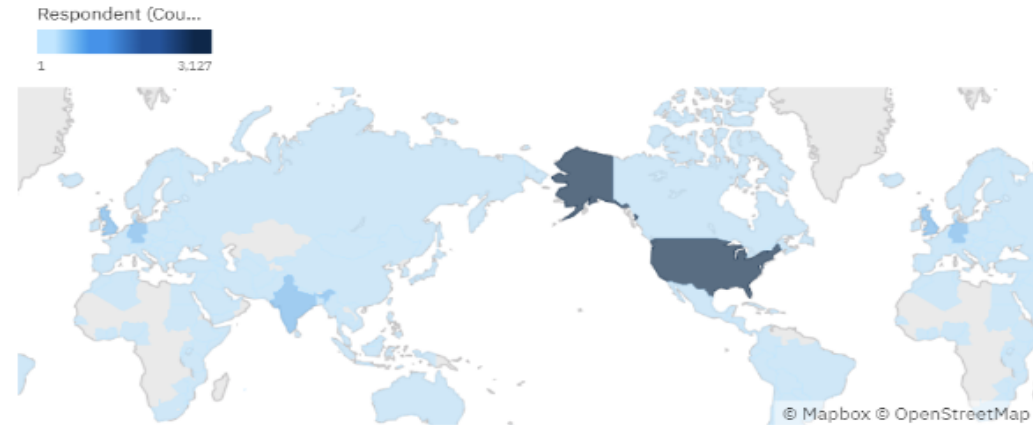
DASHBOARD TAB 3 - Demographics

Demographics

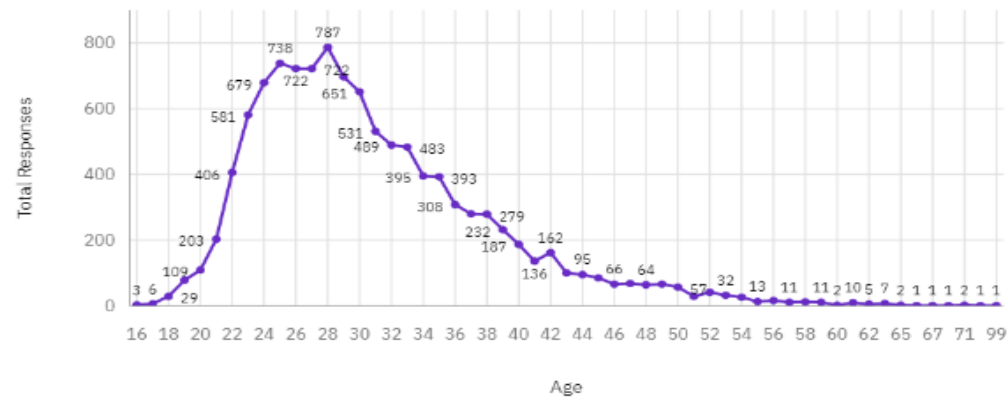
Respondents by Gender



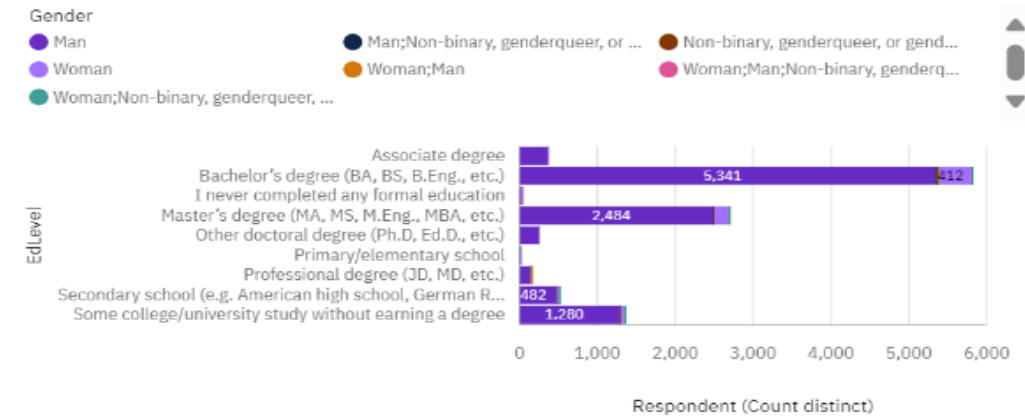
Survey Respondents by Country



Respondent Count by Age



Survey Respondent Formal Education by Gender



DISCUSSION



It's important to continuously improve your skillset in this field. This is true both in a way to advance your career and to ascertain what types of technologies and applications are finding their way into business plans.

Most of the world's tech talent is dominated by males in the United States and India, but in the last 10 years that has started to change.

It may be interesting to understand if the change is due to the recent surge in compensation or due to a newly developed interest. Perhaps looking at age at first code by gender would elucidate an understanding.

OVERALL FINDINGS & IMPLICATIONS

Findings

- JavaScript, Python, and Bash are popular languages used
- Interest in SQL type databases is subsiding in favor of NoSQL.
- Males dominate the industry at 93.5% to 6.5%

Implications

- Becoming adept at each of these may improve success in the job hunt.
- Future database job growth may focus on scalability and perhaps within IOT.
- Woman may not have as much of an interest in the field or may be subject to discrimination.

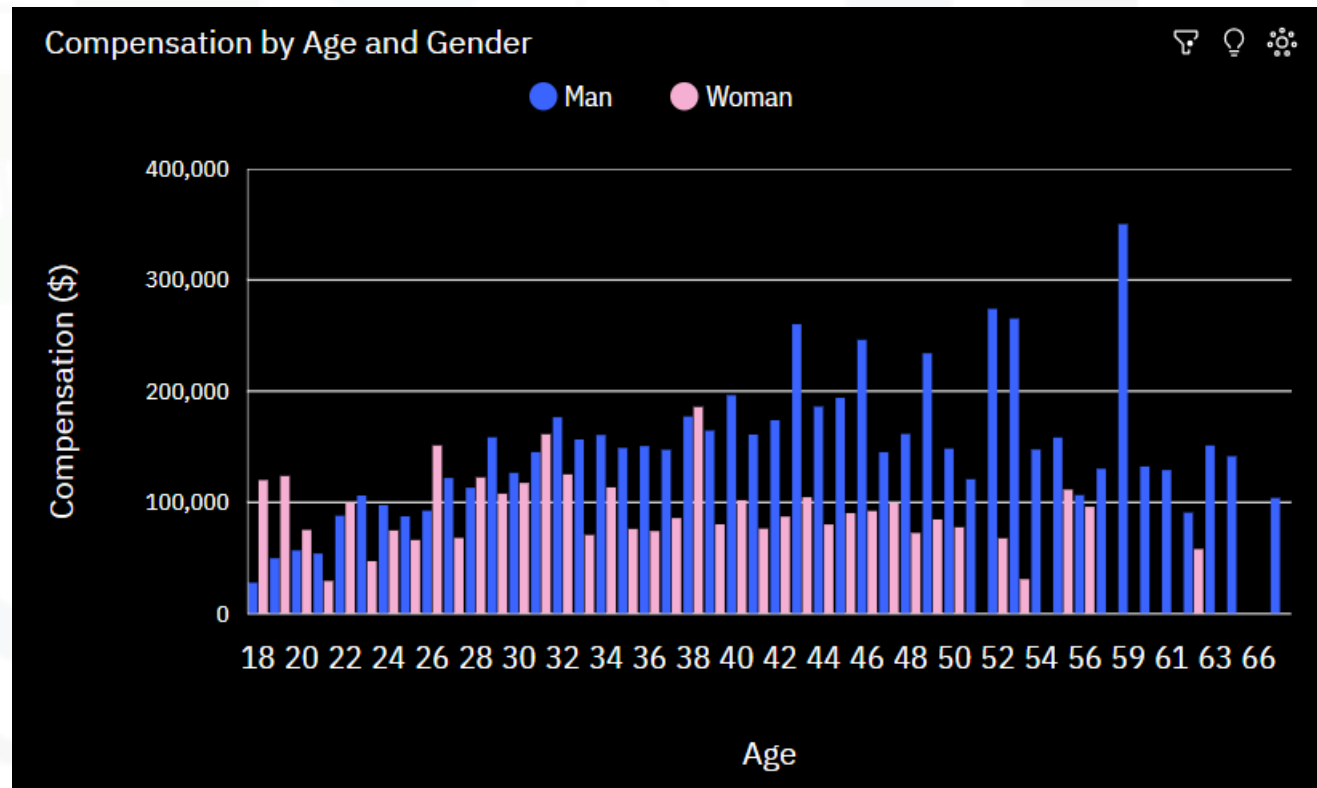
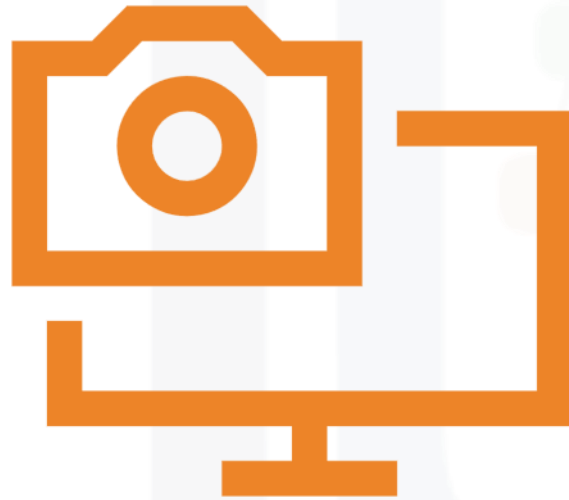
CONCLUSION



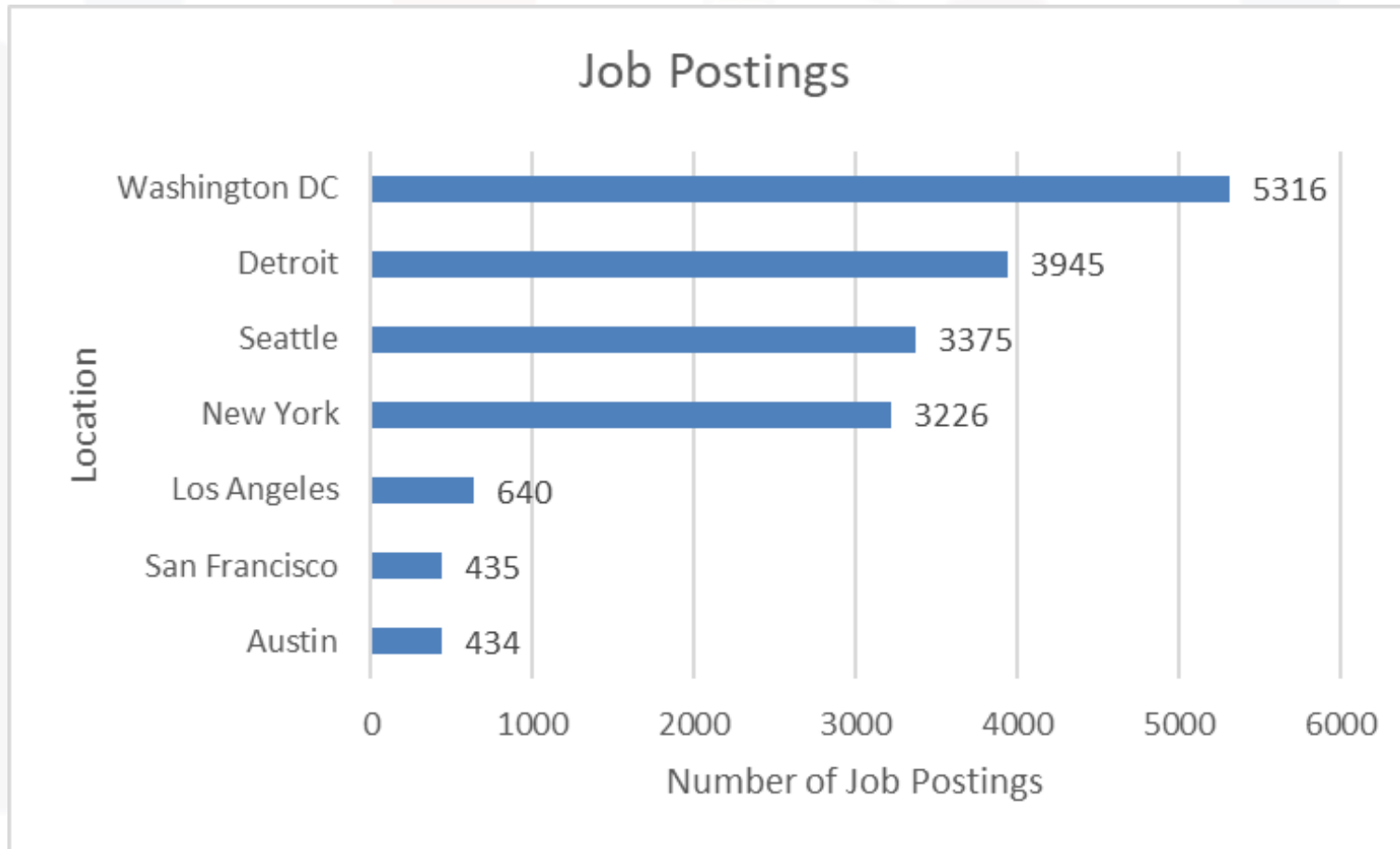
- Web application development was a strongly growing subfield in 2019 with solid projections for that to continue in 2020.
- Software developers on average make 2x the median income
- Python is a necessity to enter the data analysis / data science field
- Females are underrepresented in software development and are also paid significantly less than males, but that has started to change as of 2010.

APPENDIX:

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



JOB POSTINGS



POPULAR LANGUAGES

