Findings and Implications from Stack Overflow Developer Survey 2019

TOM WRIGHT

March 17th 2023

OUTLINE

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

EXECUTIVE SUMMARY

The following slides summarize key findings and set out implications from those findings from data collected from several different sources.

The data analysis carried out gives great insights into several different areas, such as:

- The most popular data languages, databases, and other technologies
- The predicted most popular data technologies
- The demographics (for example: differences in gender amongst developers)

The conclusions and implications from this data is very relevant to current as well as aspiring developers, recruiters, employers using data, teachers and education systems.

INTRODUCTION

The data itself was collected through various ways, such as through APIs and through Webscraping.

Through these processes, I collected the top programming skills that are most in demand from various sources including: Job postings, Training portals, and Surveys.

Once enough data had been collected, analyzing the data began and I started identifying insights and trends that would peak the interest of anyone within the data industry, either just starting in it or employers looking for new recruits.
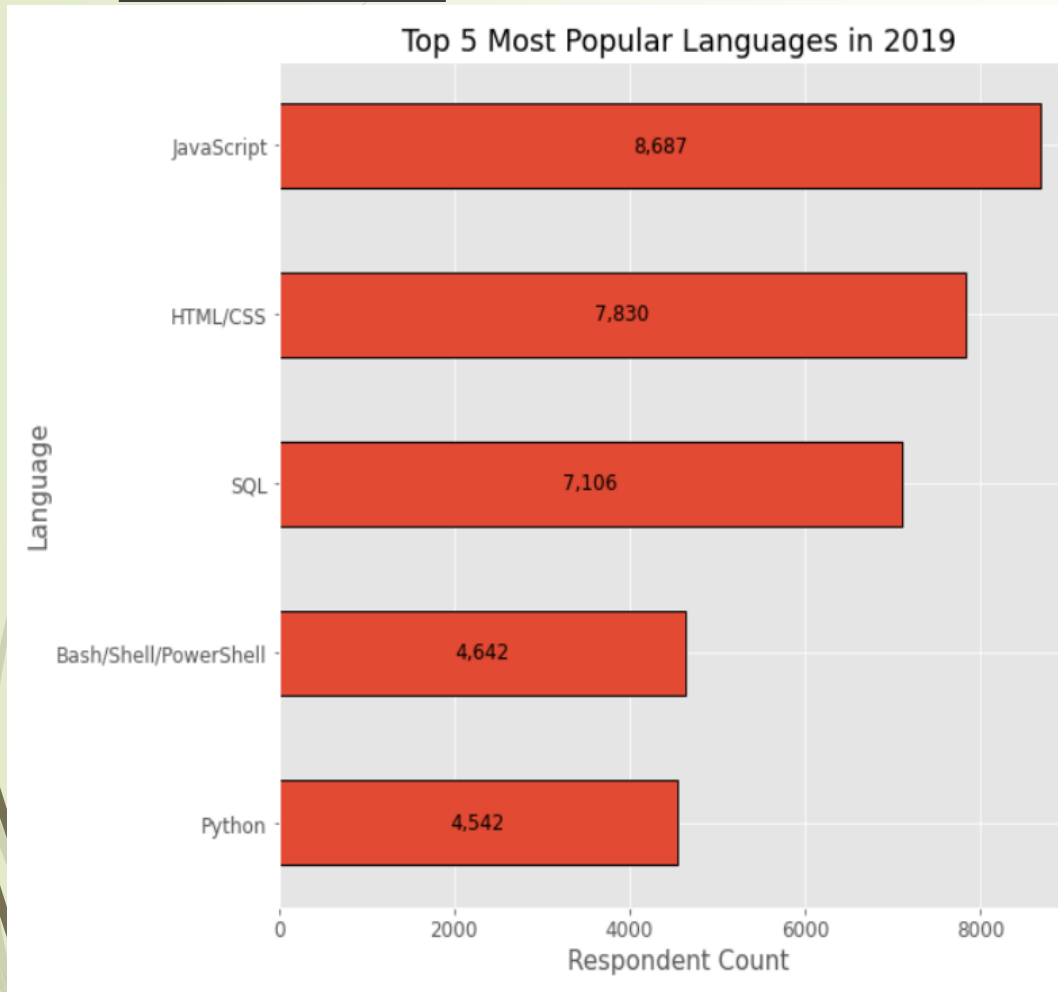
METHODOLOGY

- Data Collection and Wrangling:
  - Collecting Data using both Webscraping and APIs using *SQL*
  - Exploring the Data and cleaning it by: Finding and determining missing values, Removing duplicates, and Normalizing Data, using *Python's pandas library*
- Exploratory Data Analysis and Visualizations
  - Looking at distribution, outliers, and correlation, as well as visualizing the distribution of data, inspecting relationships, compositions, and comparisons of different subjects within the data using *Python's pandas library*
  - Creating dashboards to display the data more clearly using *IBM Cognos Embedded*
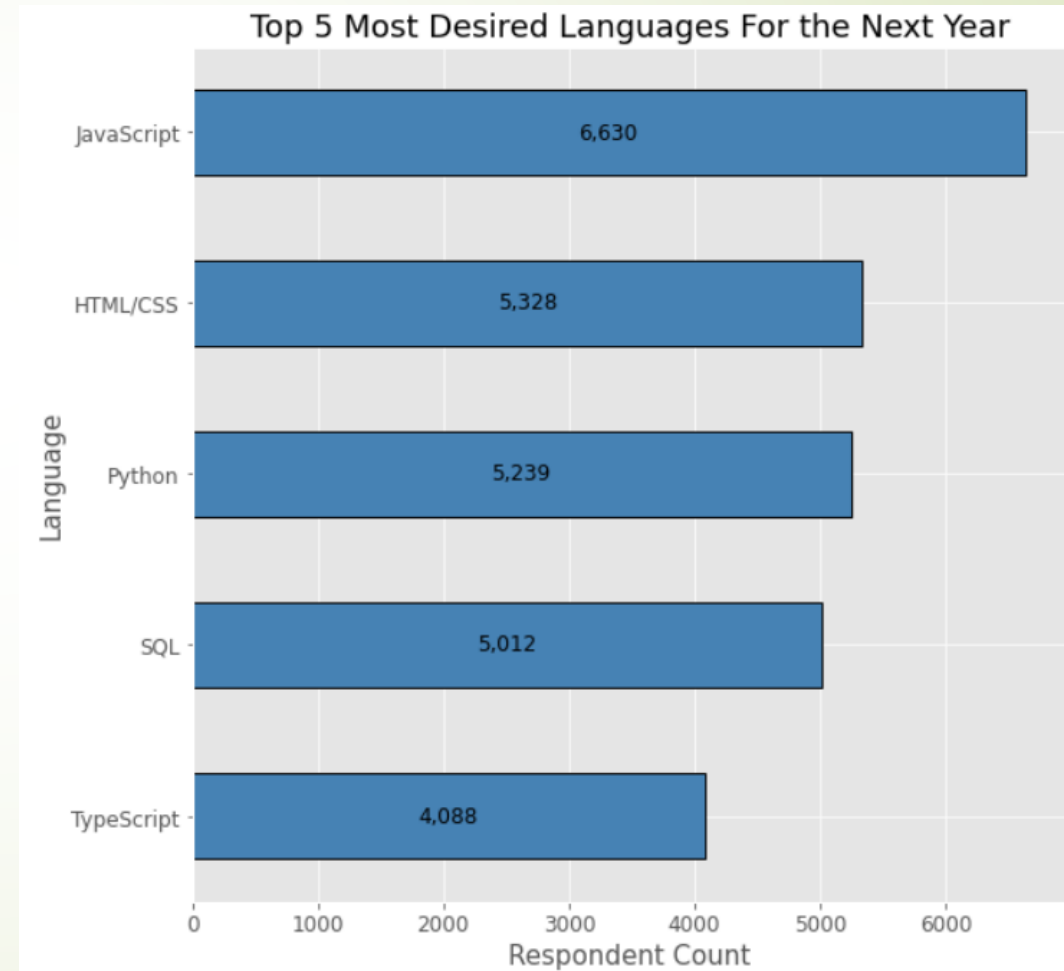
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year

Next Year

Top 5 Most Popular Languages in 2019

Language	Respondent Count
JavaScript	8,687
HTML/CSS	7,830
SQL	7,106
Bash/Shell/PowerShell	4,642
Python	4,542

Top 5 Most Desired Languages For the Next Year

Language	Respondent Count
JavaScript	6,630
HTML/CSS	5,328
Python	5,239
SQL	5,012
TypeScript	4,088

PROGRAMMING LANGUAGE TRENDS – FINDINGS & IMPLICATIONS

Findings

- Javascript and HTML/CSS were the most popular and will likely remain so for next year. However, the number of recipients desiring these 2 languages both decrease by around 2,000.

- SQL was popular also and will stay in the top 5 for the next year.

- Increasing interests in Python and TypeScript.

- Decreasing interest in Bash/Shell/PowerShell.
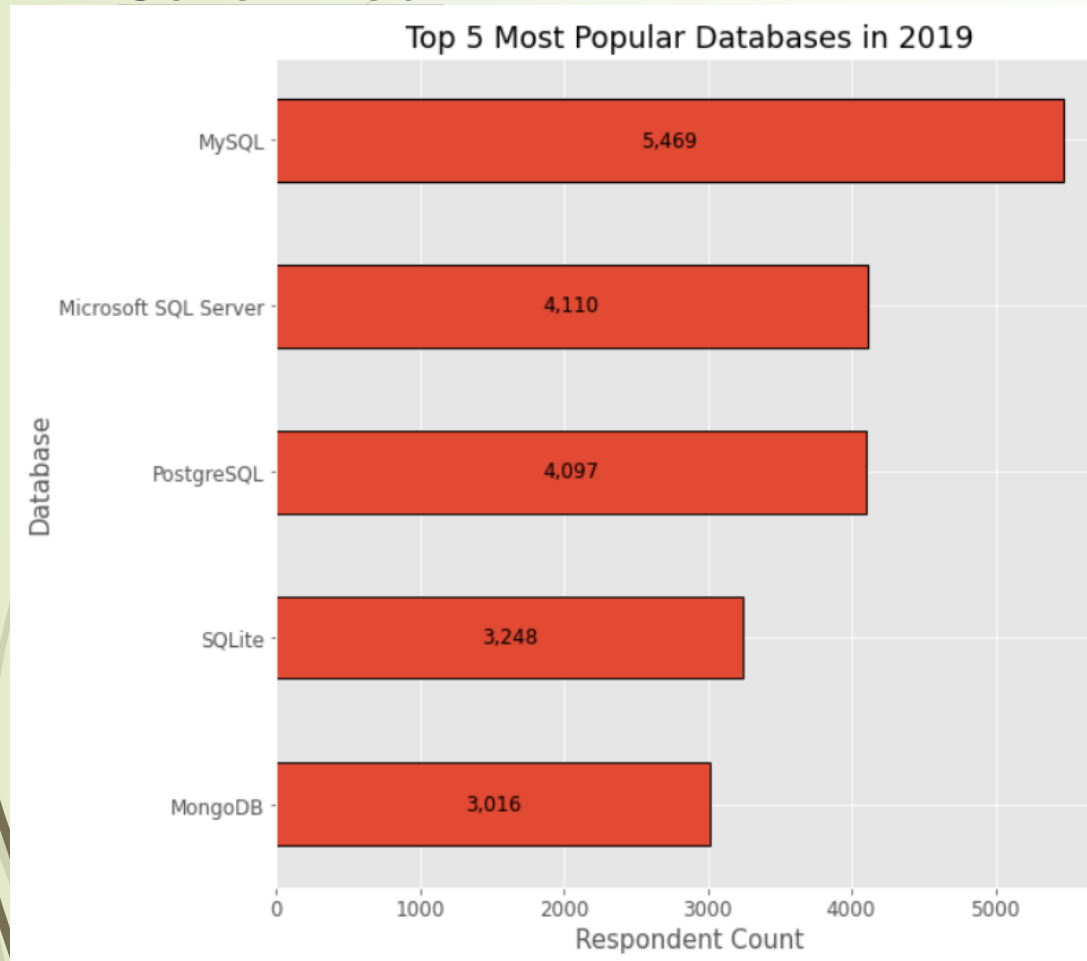
- The number of recipients desiring the top 5 languages decrease from this year to next.

Implications

- Web development remains highly desired, with both Javascript and HTML/CSS remain the most popular languages. However, TypeScript may begin catching up in the next few years.

- SQL remains to be a popular language, especially for big data storage and querying.

- Python's rising popularity likely employers increasing usage of the language in their systems.

- More languages are being made and implemented, spreading the recipients amongst more languages than the year before.
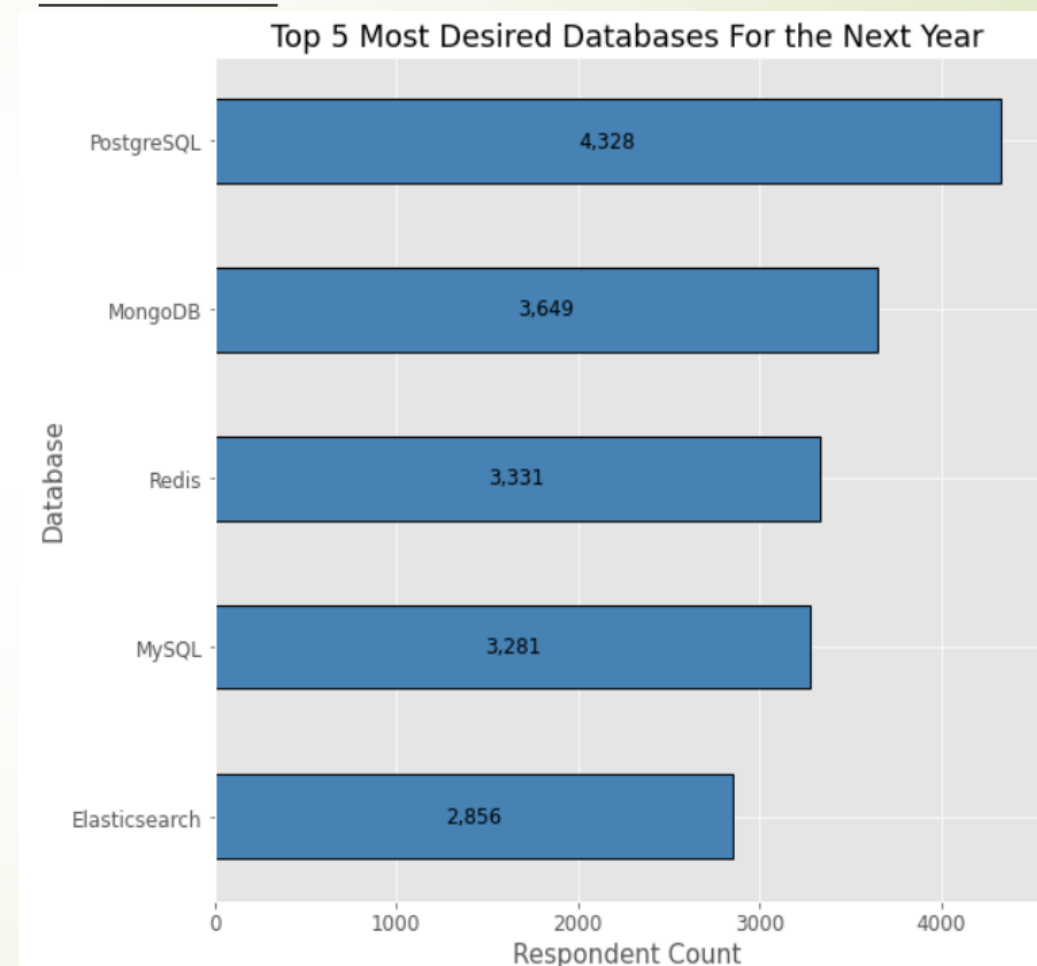
DATABASE TRENDS

Current Year

Next Year

Top 5 Most Popular Databases in 2019

Database	Respondent Count
MySQL	5,469
Microsoft SQL Server	4,110
PostgreSQL	4,097
SQLite	3,248
MongoDB	3,016

Top 5 Most Desired Databases For the Next Year

Database	Respondent Count
PostgreSQL	4,328
MongoDB	3,649
Redis	3,331
MySQL	3,281
Elasticsearch	2,856

DATABASE TRENDS – FINDINGS & IMPLICATIONS

Findings

- SQL database programs were the most popular, taking up the top 4 places, with MySQL being the most popular out of them.

- PostgreSQL is gaining popularity the most above any other SQL database program and replaced MySQL as the predicted most desired database program for next year.

- Increasing popularity in Elasticsearch, Reddis, and MongoDB.

Implications

- NoSQL database programs are gaining in popularity, with next year's list seeing 2 of the 4 SQL database programs in this years top 5 replaced by NoSQL ones. This reflects a growing need to handle nonrelational and unstructured data more often.

- A tip to aspiring and current data analysts, develop competence in NoSQL database programs as well as SQL database programs.

DASHBOARD

The link below takes you to a fully interactive dashboard, made on IBM Cognos, consisting of 3 tabs, each summarizing different topics:
- current technology use
- future technology trend
- demographics of the survey respondents

Screenshots of each slide of the dashboard are available in the next three slides.
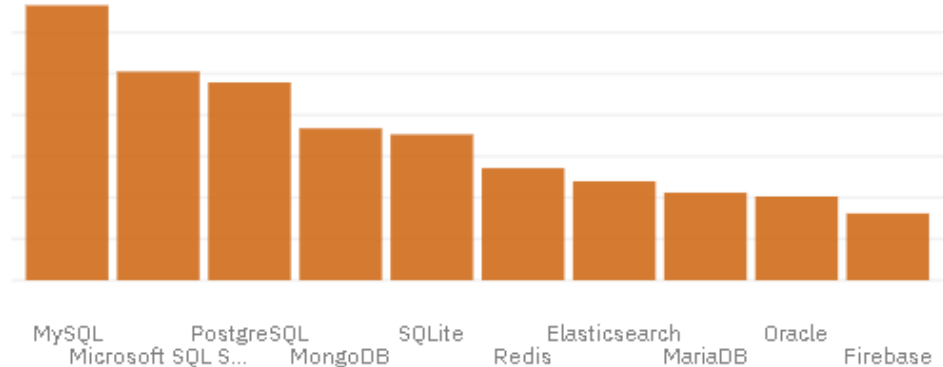
My Cognos Dashboard

TAB 1 – Current Technology Used
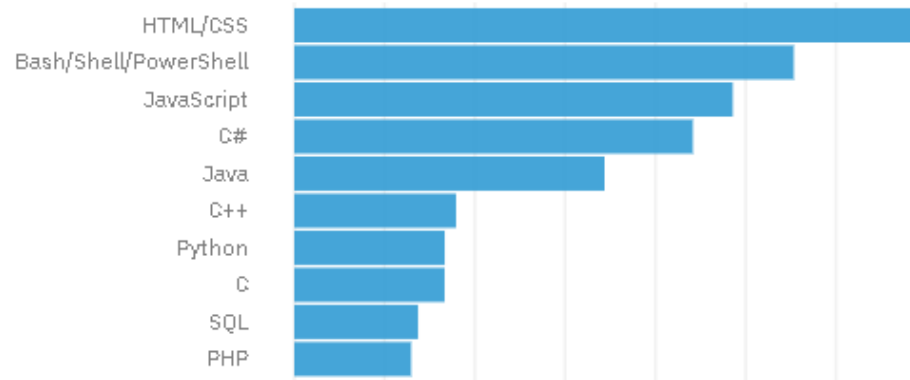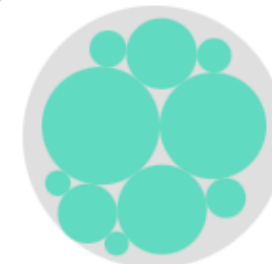
TAB 2 – Future Technology Trends

TAB 3 - Demographics
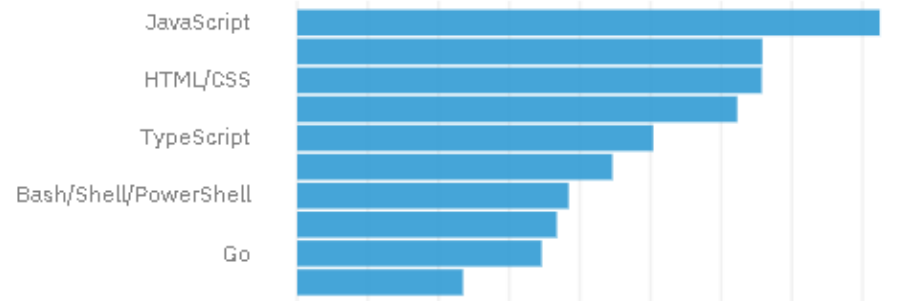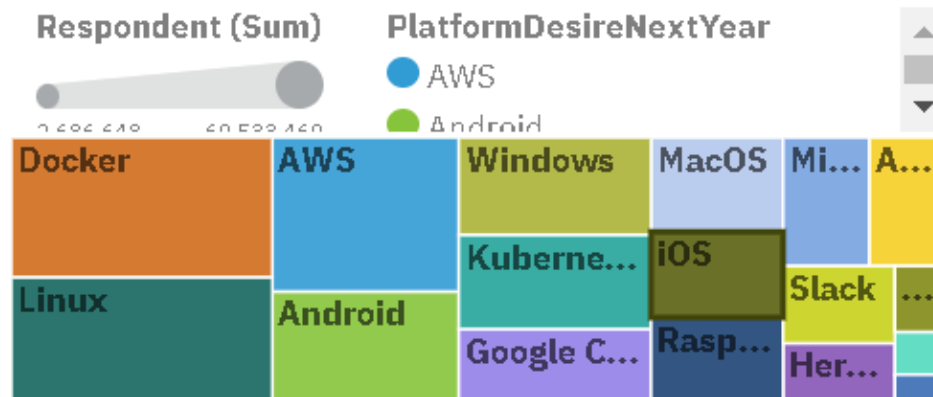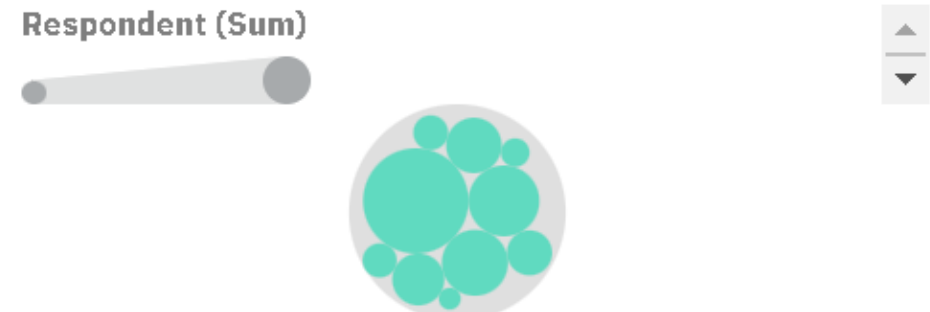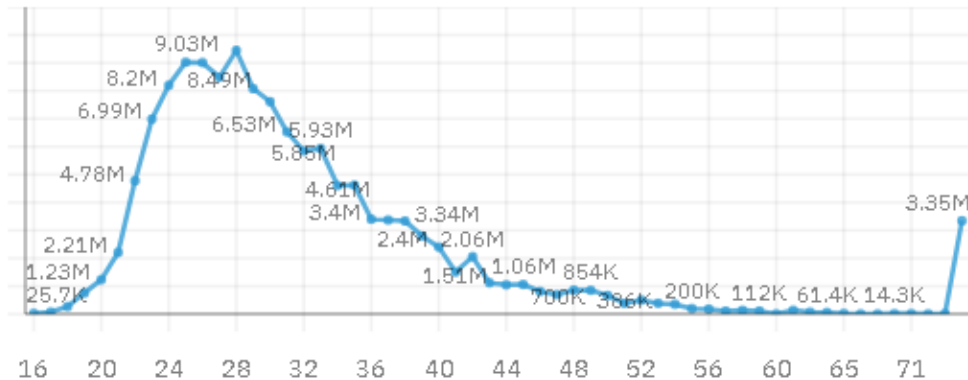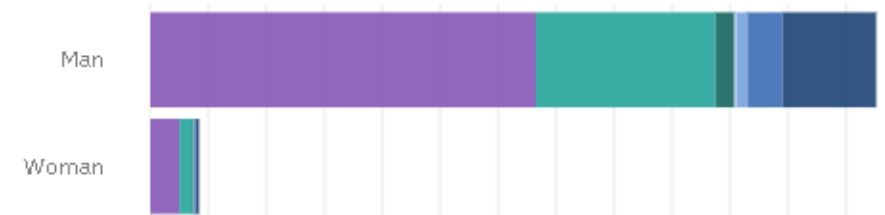
DISCUSSION

Here, through the findings and insights found from the dashboard, we can answer questions such as:

- What developer technologies are in demand?

- Which technologies should prospective developers and data professionals be learning?

- Which technologies should data education systems focus on teaching in upcoming years?

- What is the developer demographic like? Is there a gender representation gap within the data industry?

OVERALL FINDINGS & IMPLICATIONS

Findings

- High usage and interest in Javascript and HTML/CSS, as well as increasing interest in Typescript and Python.

- High usage and interest in SQL. MySQL had the highest usage out of all the SQL database programs, however PostgreSQL is the program that is gaining interest most and was the most desired database program for the coming year.

- NoSQL database programs are gaining interest too, with MongoDB being the most used and desired for the coming year.

- Looking at the demographic, a severe gender representation gap, with the majority of developers being men.

- Technology differs from country to country.

Implications

- Web development is still highly sort after. Current and aspiring developers may consider learning Typescript in addition to the more popular Javascript and HTML/CSS.

- With the growing need to handle large datasets, data professionals should continue to be familiar with and develop skills in SQL, however, not restrict themselves to just SQL, but also become competent with NoSQL database programs.

- Being competent at Python is a must!

- The gender representation gap must be addressed to get more women into data and inspire a new generation in a fast-growing industry.

- Businesses need to adapt to changing technology preferences, not just with developers but also what is used around the world to sync with other countries and understand where the industry is moving.

CONCLUSION

- From the various datasets examined, clear findings and insights are visible

- Numerous interesting insights were found into the technologies most popularly used and most desired by developers in addition to findings into the developer demographic and the major gap in representation.

- These insights should be particularly relevant for current as well as aspiring developers that want to move into or up in the data industry and compete with other developers around the world.

- These same findings should put emphasis on businesses themselves to upskill their developers, as well as improve and focus education in data on what programs are most popular right now as well as in the future.

- Furthermore, more must be done to improve the representation for women in data, from employers to education systems.
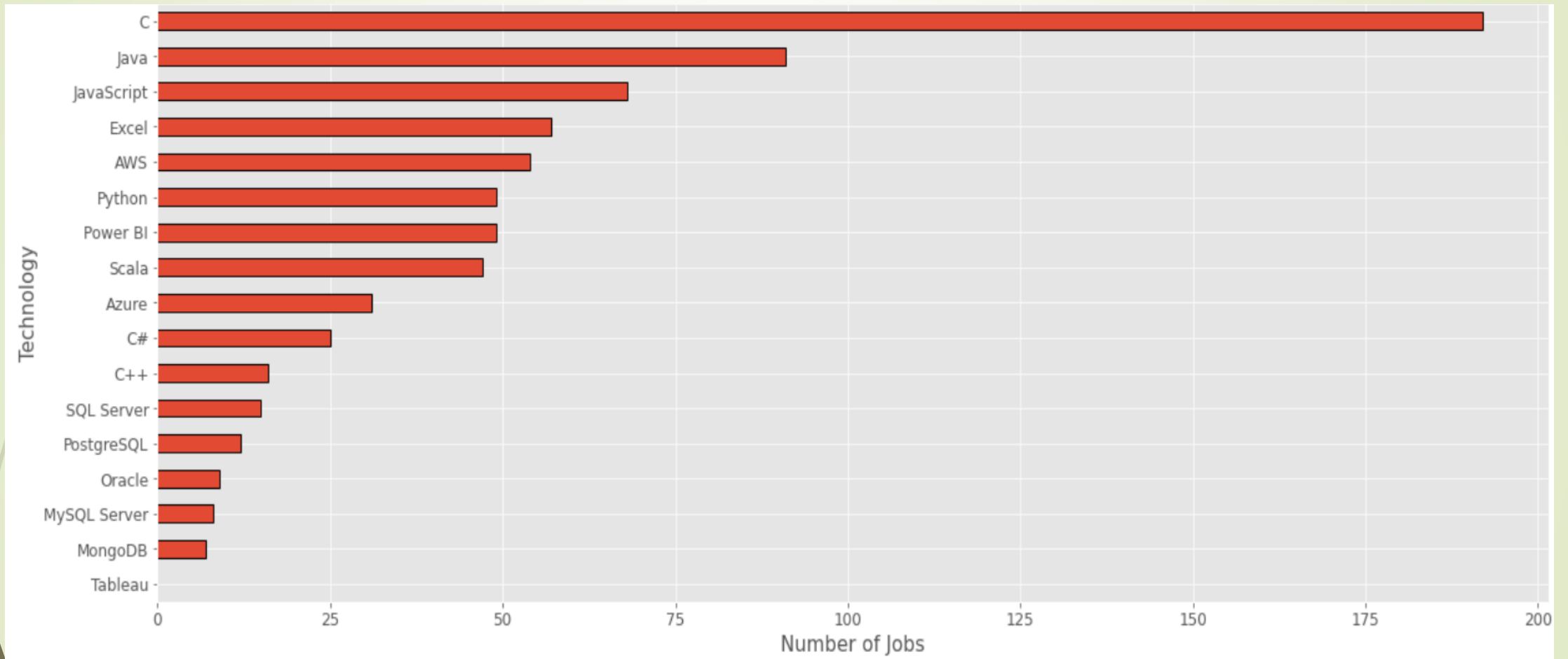
APPENDIX

In this appendix, are 2 graphs, from subsets of data collected as part of the 2019 Stack Overflow Developer Survey that was analyzed.

These graphs will interest current as well as aspiring developers, as well as recruiters and educations sectors in the data field.

The first looks at job postings for different technologies, showing what are the most desired technologies employers look for. The second displays popular languages against average annual salary, displaying what languages are valued highest in a developer.
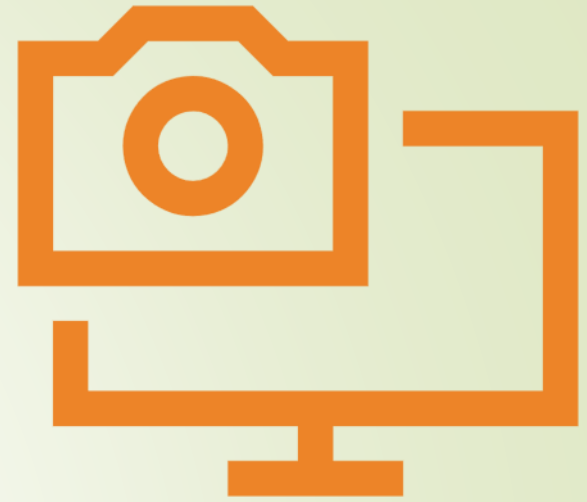
1. JOB POSTINGS FOR DIFFERENT TECHNOLOGIES

2. POPULAR LANGUAGES BY SALARY