IBM DATA ANALYST
CAPSTONE PROJECT :STACK OVERFLOW DEVELOPER SURVEY

Dipanjan Maity

13th August 2022

IBM Developer

SKILLS NETWORK

OUTLINE

Executive Summary

Introduction

Methodology

Results
◦ Visualization – Charts
◦ Dashboard

Discussion
◦ Findings & Implications

Conclusion

Appendix

IBM **Developer**

SKILLS NETWORK

EXECUTIVE SUMMARY

The following slides summarize key findings from an analysis of Stack Overflow as part of the 2022 Data Analyst Capstone Project.

- The most popular languages, databases, developing platforms, web frame and other technologies

- Attitudes reflecting which technologies will become most popular in the future

- Demographics (e.g., the gender gap among developers)

These findings are relevant particularly to current and aspiring developers, recruiters, educators, and policy makers

INTRODUCTION

- This presentation has been created for stakeholders and business decision makers within global IT and business consulting services firm

- The primary objective of the annual surveys is to gather data regarding technology usage and trends among developers.

The comparison between current & next year technology trend will provide great inside into the market.

- The study consist of 4 main parts:
  - Top language using
  - Top database using
  - Top platform using
  - Top web frame using

METHODOLOGY

- **Data Wrangling:** A portion of the dataset (provided by IBM) was loaded and cleaned using SQL and Python's pandas library.
  - Cleaning procedure: Duplicates removal, data imputation, data normalization

- **Analysis & Visualization:** EDA and data visualization were conducted using various Python libraries and Cognos. Specifically, the following measures were examined:
  - Technologies (i.e., languages, databases, platforms, and web frames) used in 2019)
  - Technologies most desired for the next year
  - Demographics (i.e., gender, country, age, and education).
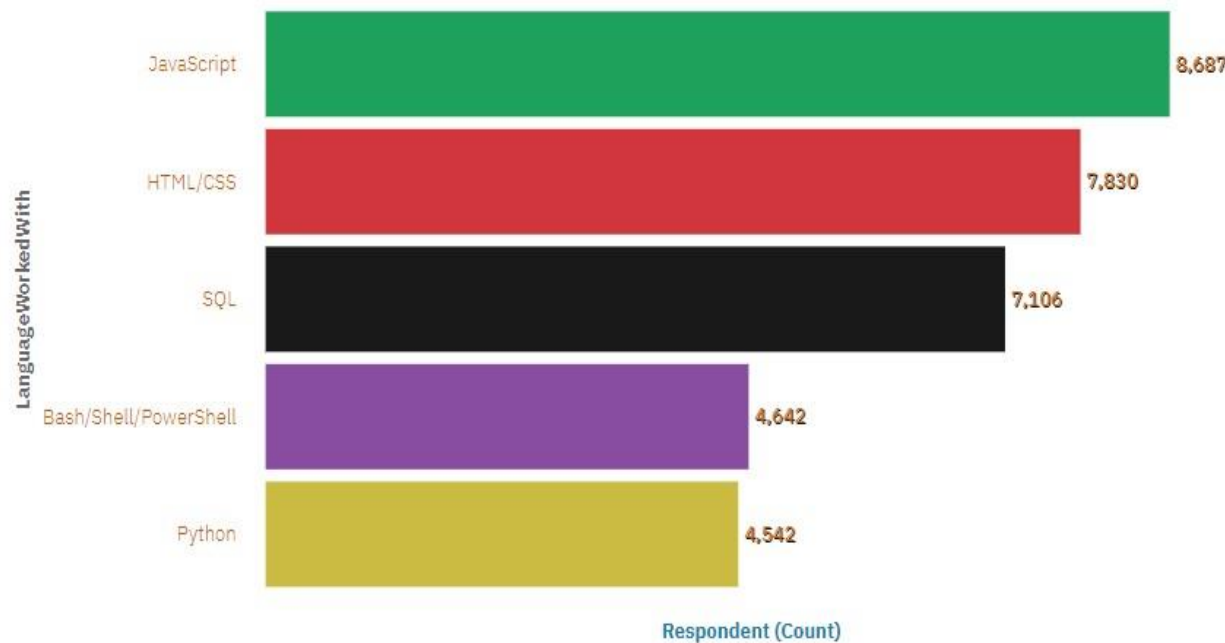
RESULTS

- The results consists of 4 main parts:
  - Programming language trend
  - Database trend
  - Platform trend
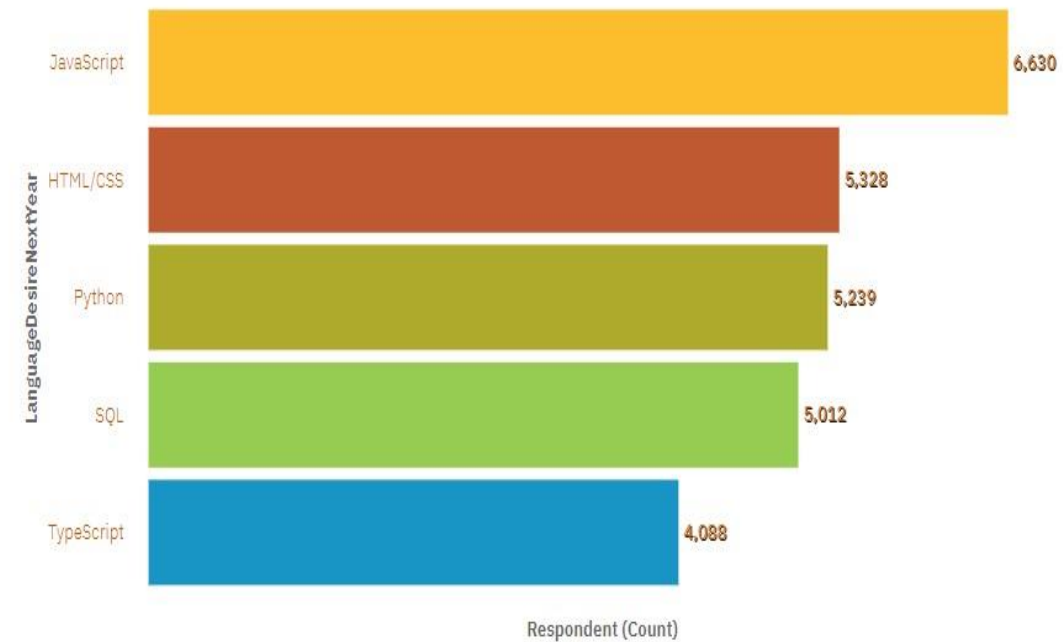 - Web frame trend

PROGRAMMING LANGUAGE TRENDS

Current Year

Next Year

Top 5 Most Popular languages in 2019

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

Top 5 Languages Desired for The Next Year

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

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

1. JavaScript & HTML/CSS are top programming languages for both current and future trend.

2. SQL was popular in 2019 and will likely remain.

3. Popularity of python is increasing day by day

4. Decreasing interest in PowerShell/Bash
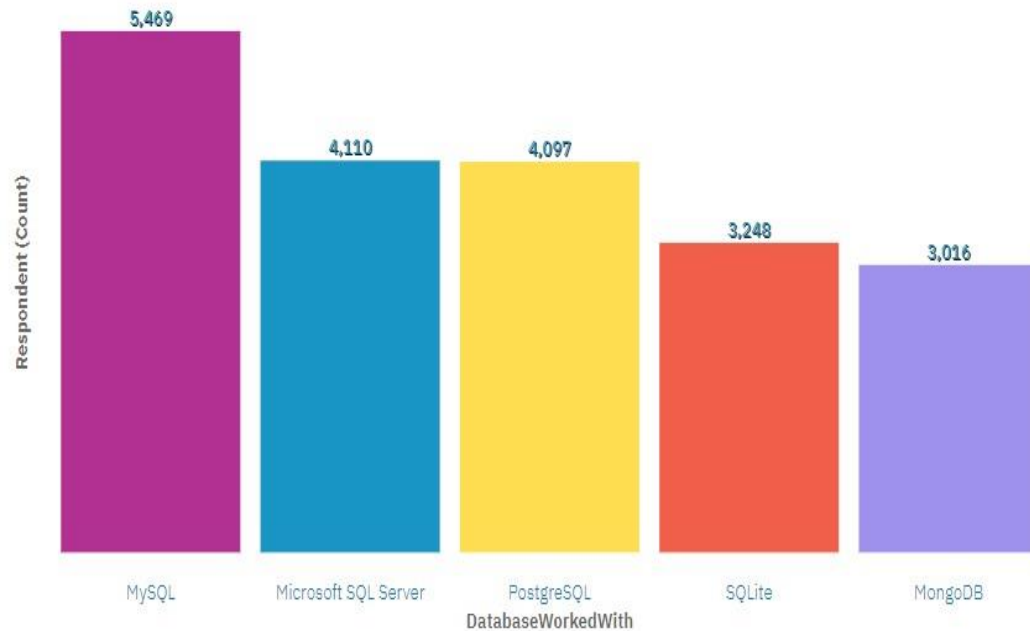
5. C++ and PHP are disappearing from top languages

Implications

- Web programming languages still stay dominance in near future and JavaScript and HTML/CSS remain the dominant languages—however, TypeScript may catch up in the future.

- Python is becoming more and more important.

- and JavaScript and HTML/CSS remain the dominant languages—however, TypeScript may catch up in the future.

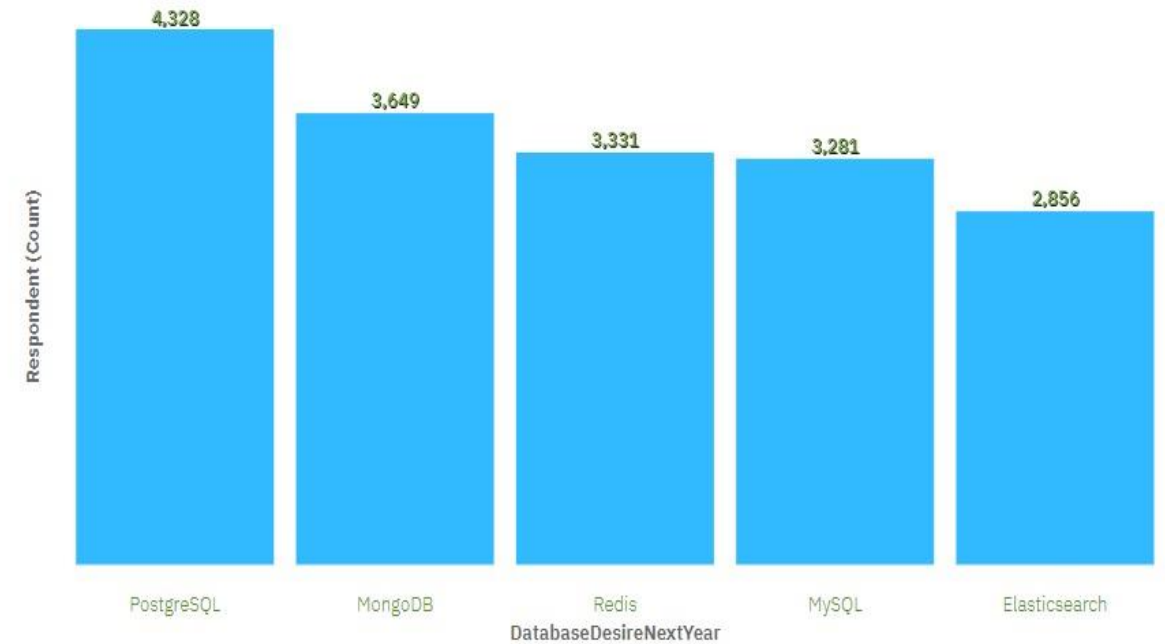- Legacy language such as C++, PHP etc. are gradually falling out of interest.

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

DatabaseWorkedWith

Top 5 Databases Desired for The NextYear

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

DatabaseDesireNextYear

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

1. SQL database programs were the most popular in 2019, with MySQL in the lead

2. Popularity of PostgreSQL is increasing and most desired database

3. Increasing interest in Elasticsearch

4. MongoDB was popular in 2019 and gaining interest.

Implications

1. Open source databases are getting popular.

2. NoSQL database programs are gaining popularity, which likely reflects a growing need to handle non- relational and unstructured data.

3. Current and aspiring data analysts should develop competence in NoSQL in addition to SQL database programs

PLATFORM TRENDS

PLATFORM TRENDS- FINDINGS & IMPLICATIONS

Findings

- Windows and Linus is most dominating platform in the current scenario

- Cloud platform such as Kubernetes, Docker, AWS is getting more and more market share.

Implications

- Windows and Linux will be the most developing platform in a foreseeable future

- Cloud computing platform is getting popular day by day.
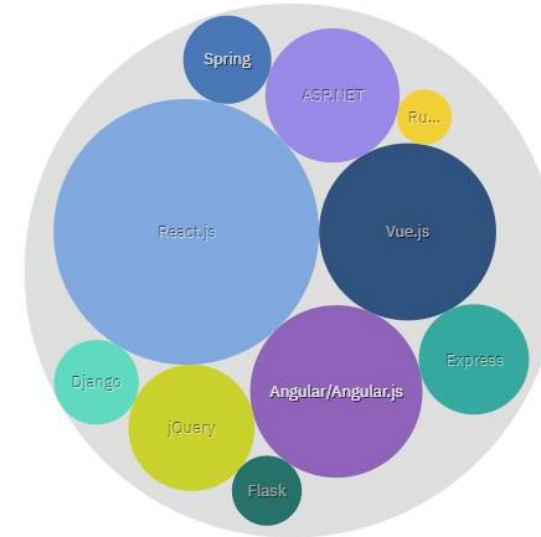
WEB FRAME TRENDS

Current Year

Next Year

Top 10 Web Frames used by Developers in 2019

Flask
ASP.NET
Lara...
jQuery
Angular/Angular.js
Vue.js
Spring
Express
React.js
Djang...

Top 10 WebFrameDesireNextYear

Spring
ASP.NET
Ru...
React.js
Vue.js
Django
Express
jQuery
Angular/Angular.js
Flask

IBM **Developer**

SKILLS NETWORK

WEB FRAME TRENDS- FINDINGS & IMPLICATIONS

Findings

- Popularity of jQuery is getting decreased .

- React.js becomes most desired webframes for the next year

Implications

- jQuery quickly loses market share show lack of interest in its future development.

- React.js and Vue.js becomes the new standard for web frame development.

DASHBOARD

The following link contains the full, interactive Cognos dashboard summarizing
(a) current technology use,
(b) future technology trend,
(c) demographics of the survey respondents

Click Here to open dashboard(Ctrl+Click)

Screenshots of the dashboard are shown in the next three slides.

DASHBOARD A: CURRENT TECHNOLOGY USE

Top 10 Most Popular languages in 2019
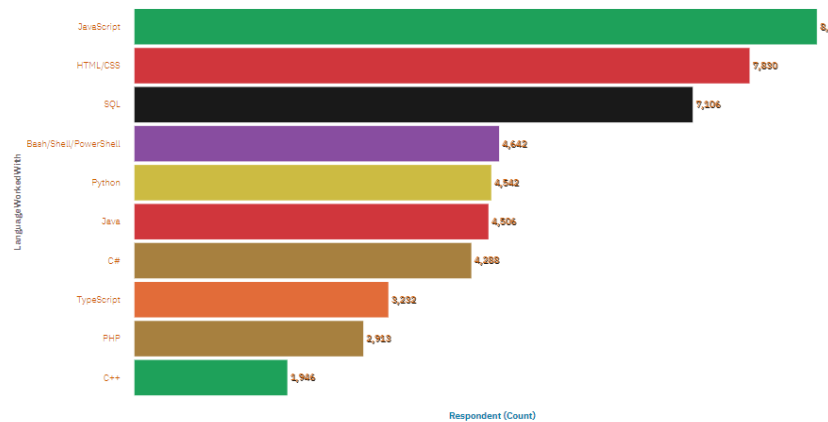
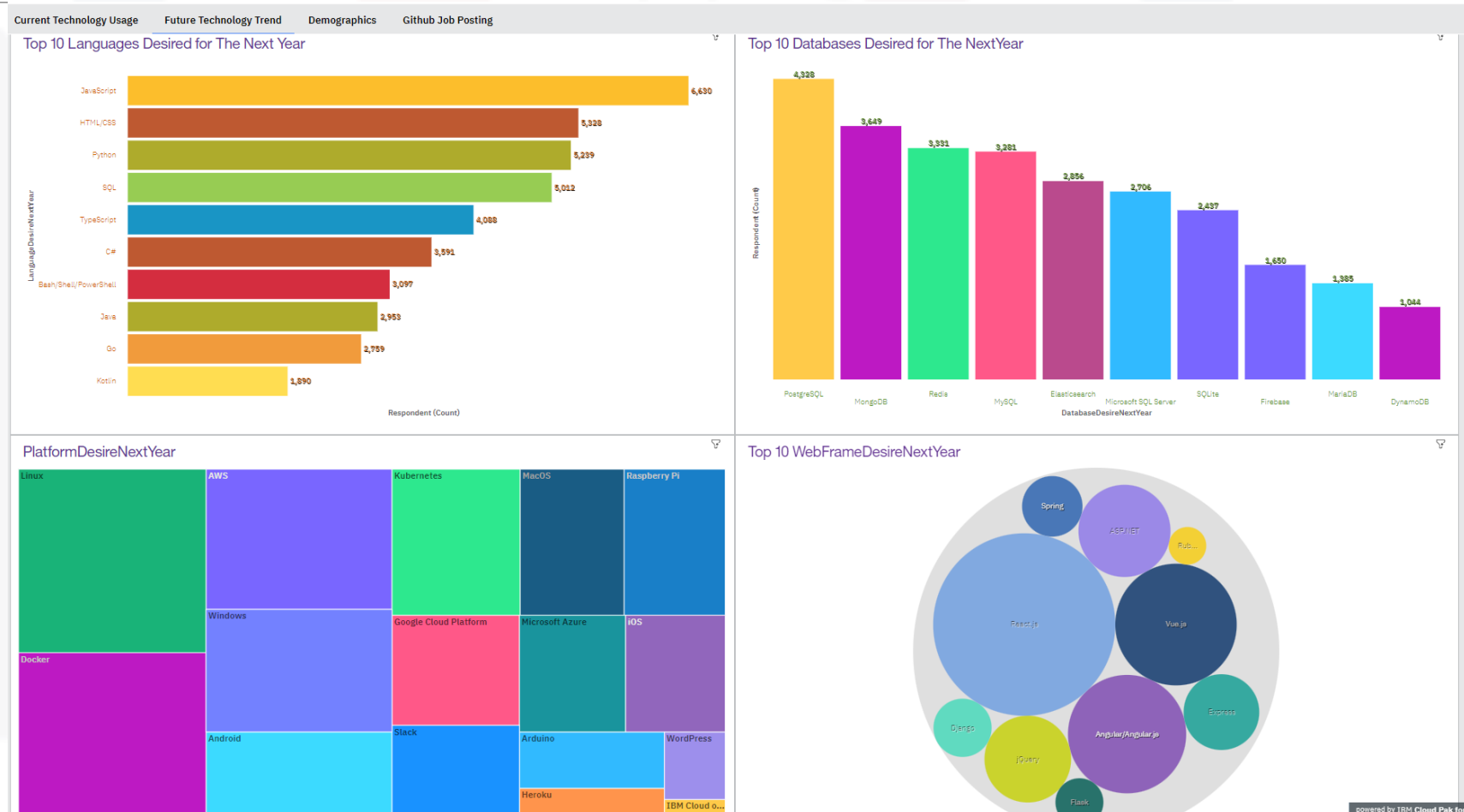Top 10 Most Popular Databases in 2019
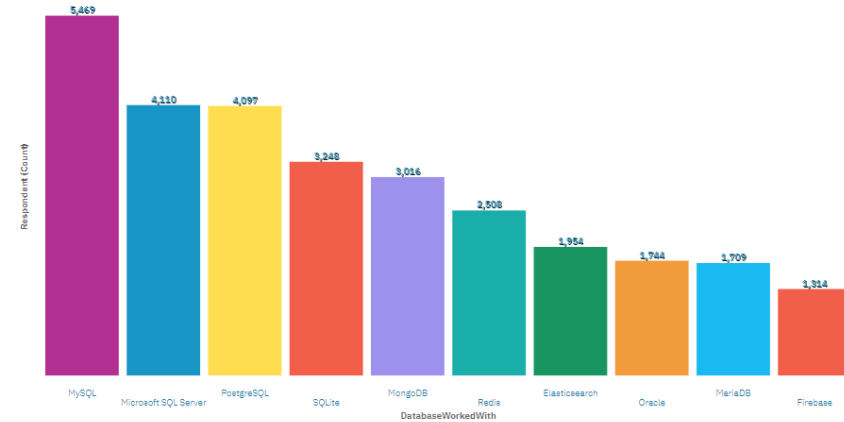
Platform Used by Developers in 2019

Top 10 Web Frames used by Developers in 2019
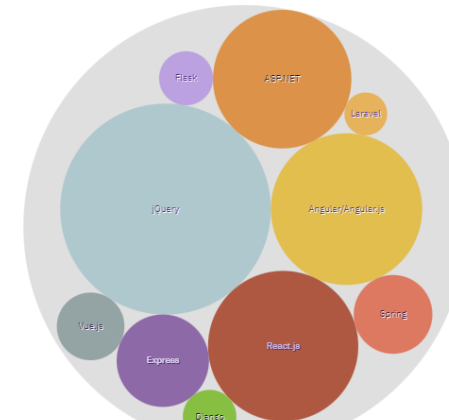
DASHBOARD B: FUTURE TECHNOLOGY TREND
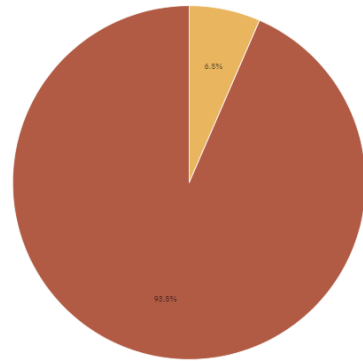
DASHBOARD C: DEMOGRAPHICS
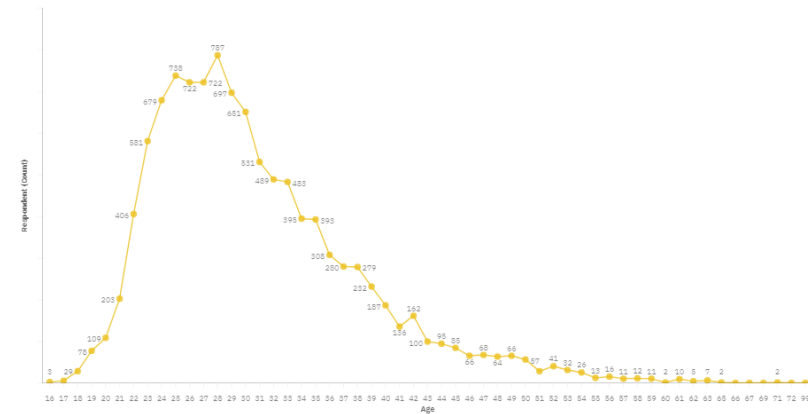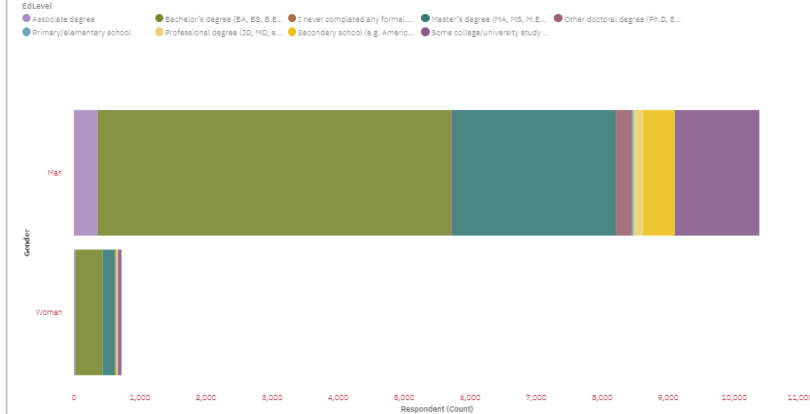
DISCUSSION

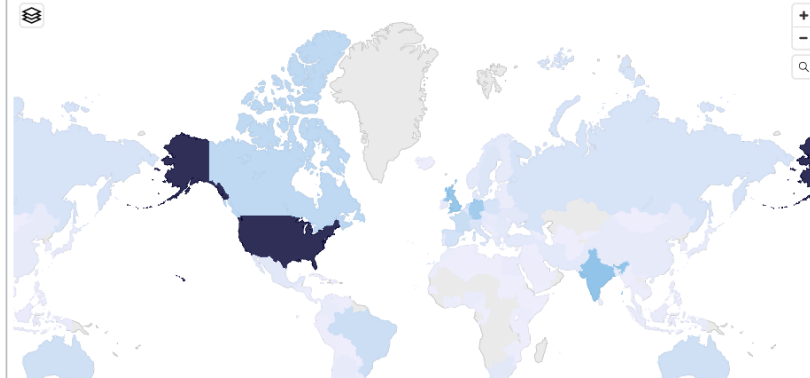Taken together, the findings yield insights into the following questions:
1. What kinds of developer technologies are in top demand?
2. Which technologies should prospective developers and data professionals be learning?
3. Which technologies should educators place more emphasis on teaching in upcoming years?
4. What does the distribution of annual compensation for developers look like?
5. What is the developer demographic like? Is there a gender representation gap?

OVERALL FINDINGS & IMPLICATIONS

Findings

1. Programming Languages typescript is gaining interest and Python continues to grow as well

2. High usage and interest in JavaScript and HTML/CSS remain high usage. There's also increasing interest in Typescript.

3. High usage and interest in SQL. MySQL had the highest usage in 2019, but PostgreSQL is gaining interest and was the overall most desired database program for the next year

4. NoSQL database programs gaining interest, of which MongoDB was the most used in 2019 and desired for the next year

5. In Platforms Windows and Linux is going to dominate for a long time

Implications

1. Web development is still in high demand. Current and prospective developers may consider picking up Typescript in addition to JavaScript and HTML/CSS.

2. With the growing need to handle big data and perform AI and ML work, data professionals should continue to enhance SQL competence but also enhance competence with NoSQL database programs and Python.

3. Policy makers, educators, and organizations should work to minimize the gender representation gap in addition to the technology divide between country
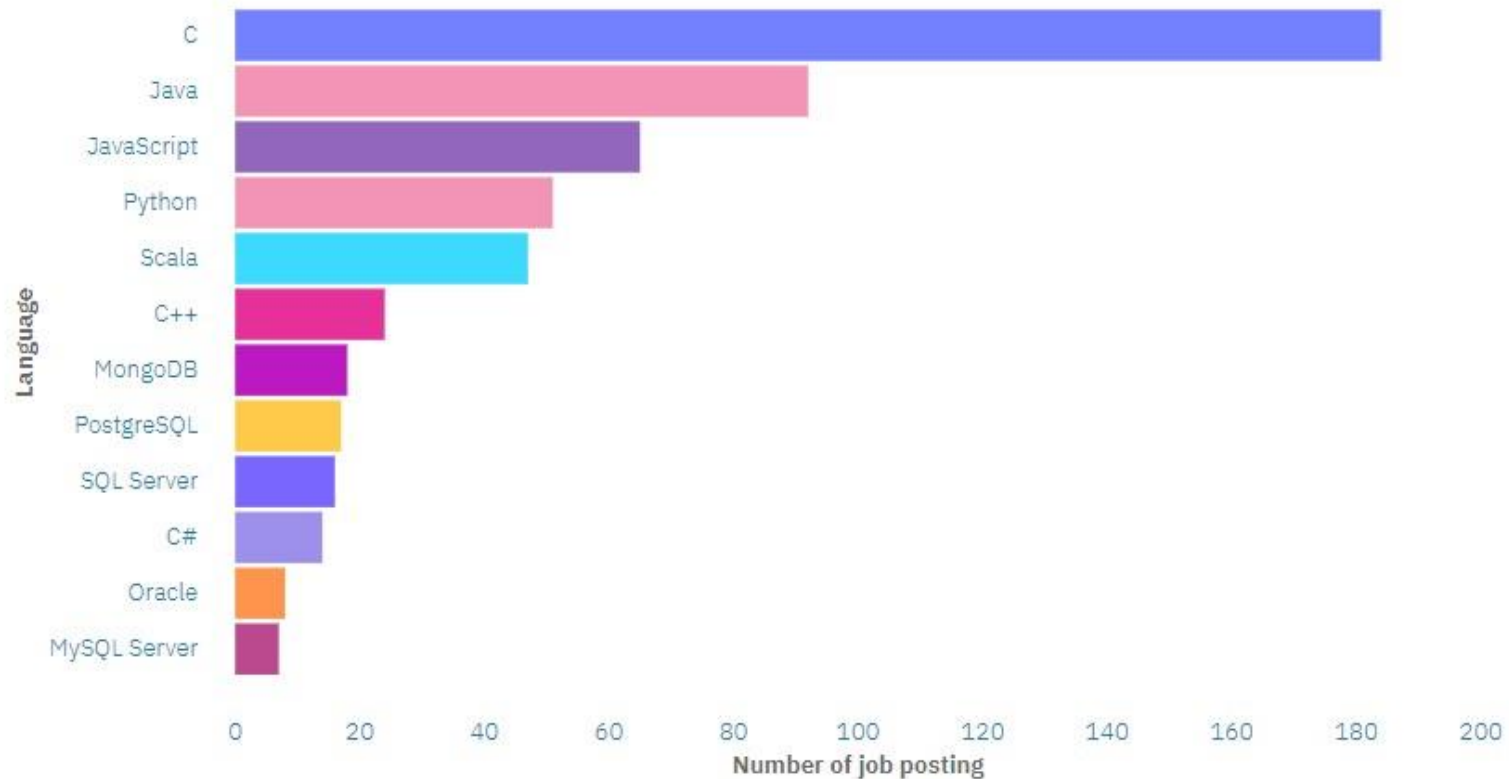
CONCLUSION

1. To stay relevant in the technology field, it's important to learn web-based language & non-relational database.

2. Among web framework, React.js is the best to learn for the future.

3. Cloud-based service is the right direction for IT investment.

4. Carve out budget in order to hire additional staff with skills needed to fill any gaps.

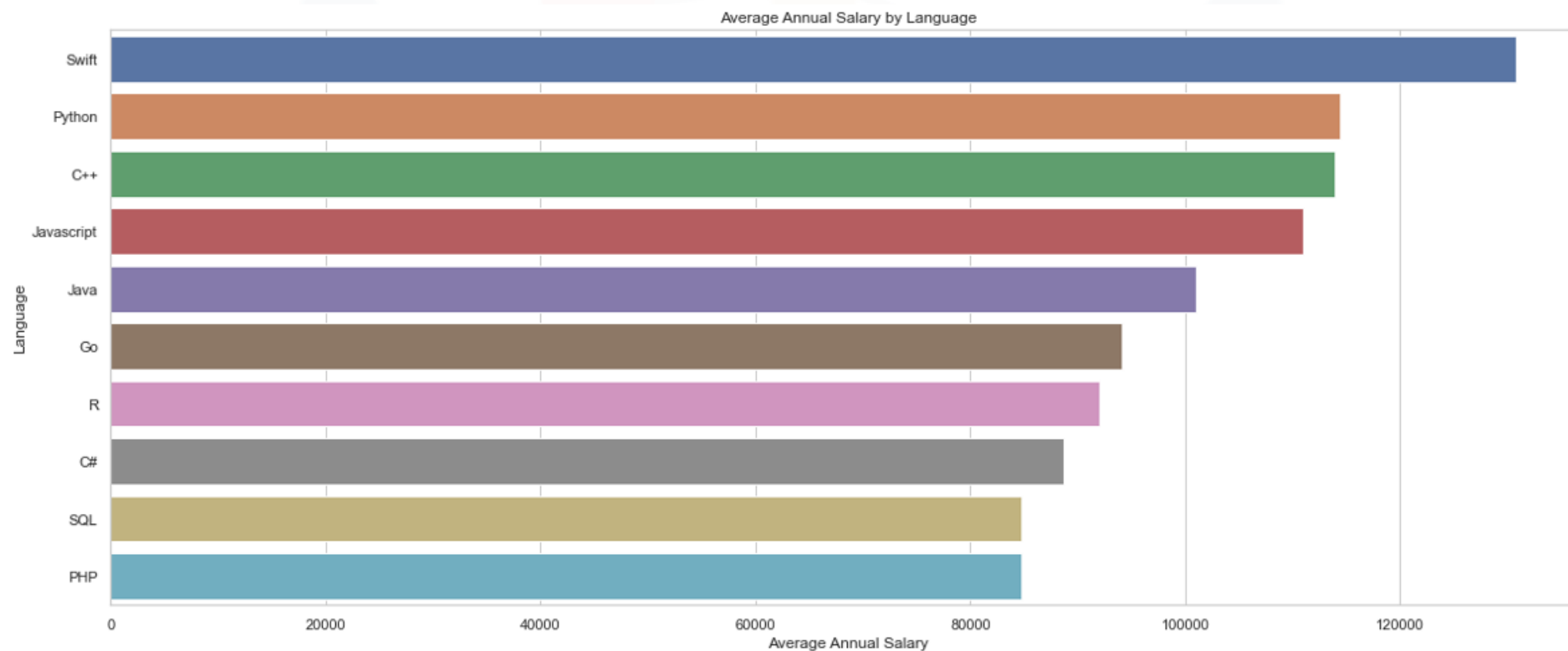5. Make adjustments in staff for those skills no longer in demand

These insights should be particularly relevant for current and prospective developers aiming to remain competitive, businesses aiming to upskill their talent, educators in the field, and policy makers aiming to address gender and economic issues.

APPENDIX A: GITHUB JOB POSTINGS FOR SELECTED TECHNOLOGIES

APPENDIX B: POPULAR LANGUAGES BY SALARY

Average Annual Salary by Language

Resources

1. https://stackoverflow.blog/2019/04/09/the-2019-stack-overflow-developer-survey-results-are-in/

2. m5_survey_data_demographics.csv and m5_survey_data_technologies_normalised.csv

3. Github-job-postings.xlsx