IBM Data Analyst Capstone Project – Current and Desired Database, Programming Language, Platform and Web Frame Trends for Developers

Jason Bryden

February 19, 2024

IBM Developer

SKILLS NETWORK

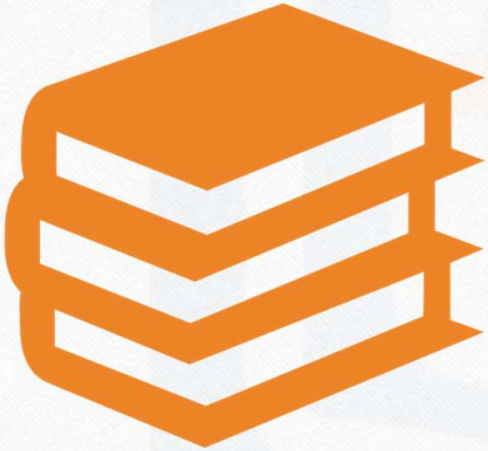
OUTLINE

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

EXECUTIVE SUMMARY

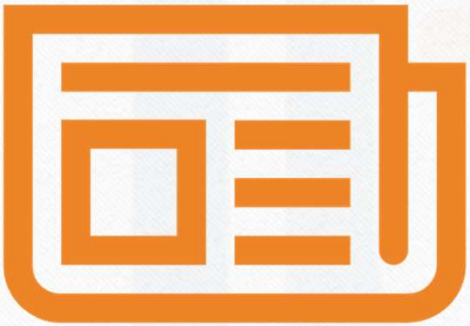
- In today's developer sector, it is very important to stay in tuned with the changing landscape in programming languages, databases, platforms and web frames. The report uses data analytics to feature current and expected trends in the four areas.
- Studies of the demographics of developers in the field were conducted.
- Data was collected from the 2019 Stack Overflow Developer Survey, the IBM site and Jobs API (Application Programming Interface).
- The results showed that Javascript is the top programming language currently and into next year. Unlike programming languages, it is projected that a different database will emerge as the most sough after. MySQL is presently the top database used, but PostgreSQL is expected to become the most appealing database in the future.
- Most of the respondents were male, American and had an average age of 28.

INTRODUCTION



- Data analytics is used in this presentation to highlight current and projected trends to determine what are the most desirable programming languages, databases, platforms and web frames to learn.
- The target audiences were developers,, data professionals, Human Resource (HR) managers and people interested in learning what the top on-demand skills currently and in the future.
- Respondents were asked a series of questions on what they use currently, what they seek to learn next year, where they are from, gender, age and highest education completed.
- The following were examined using the data collected.
  - Which programming languages, databases, platforms and web frames are popular right now?
  - What are the most sought after next in the four categories and how do they differ from this years?

METHODOLOGY

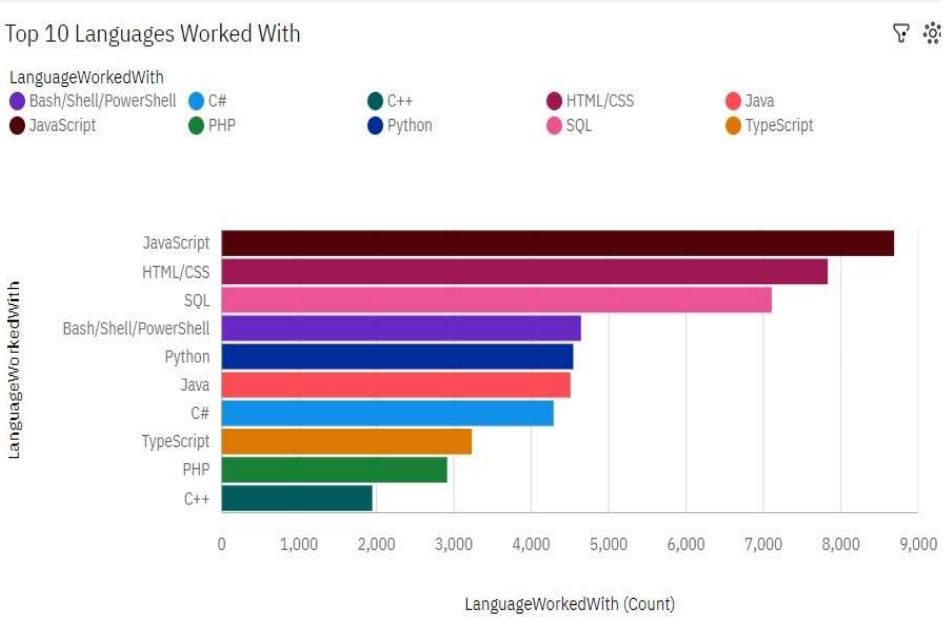


- Job location data was obtained by using Jobs API and then stored in an Excel file.
- The name of the programming language and average annual salary were scraped to obtain the data from the survey.
- Python was used to clean and analyze the data. Exploratory data analysis was done to assess the correlation.
- Charts, graphs and dashboards were created using Python and IBM Cognos.

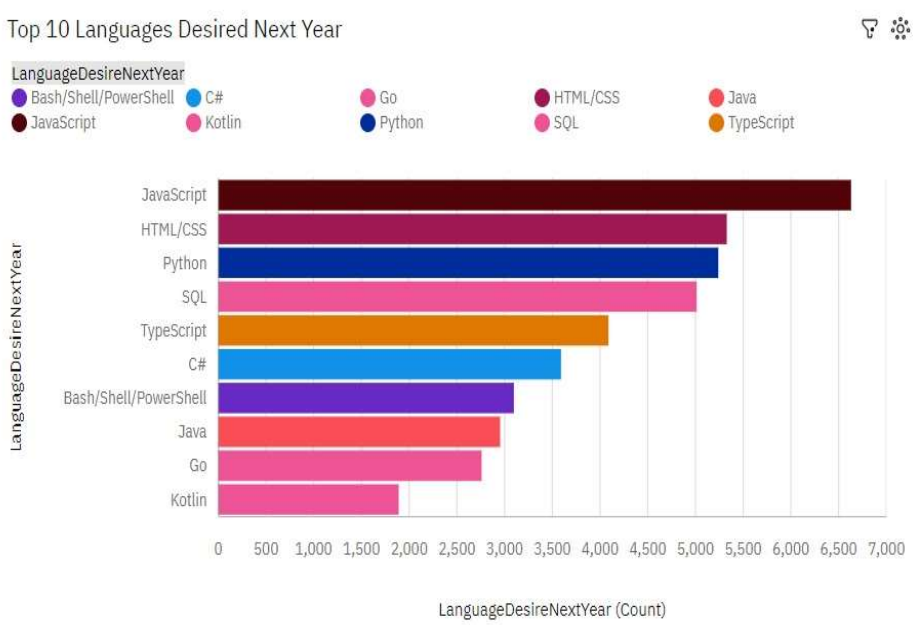
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



IBM Developer

SKILLS NETWORK

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

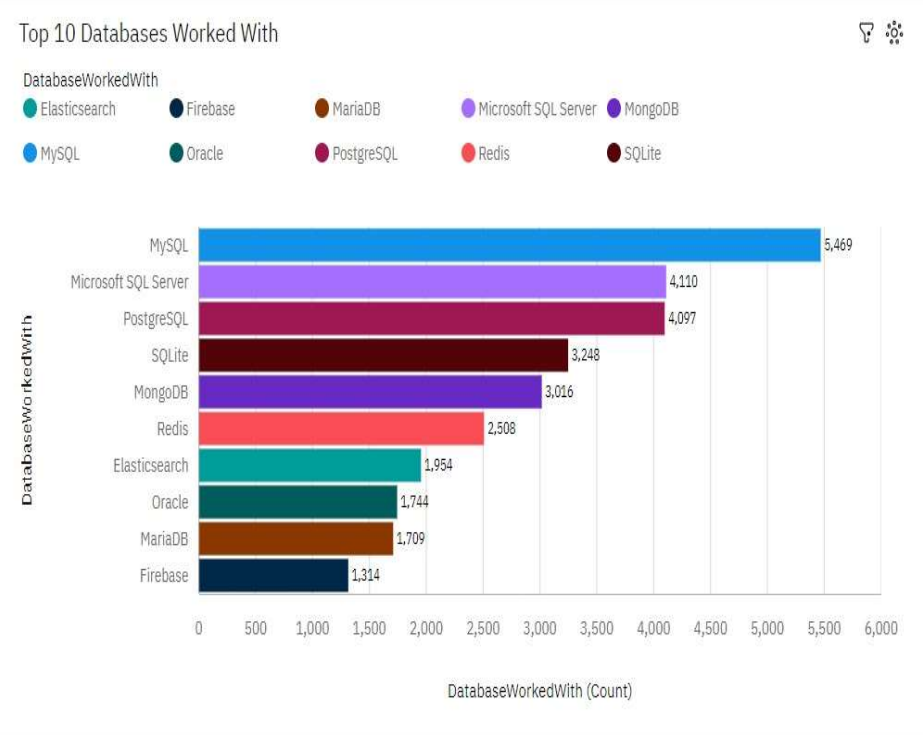
Findings

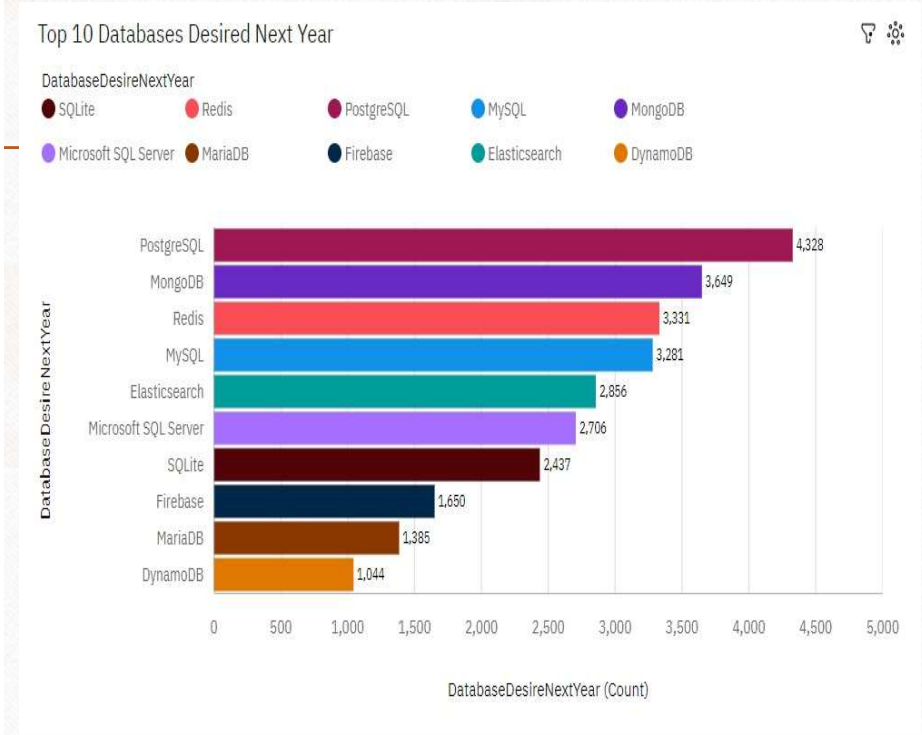
- Javascript will continue to be top programming language going forward followed by HTML/CSS. SQL and Python remain in the top five, with Python jumping SQL in the future.
- Bash/Shell/Power Shell, Java and C++ are losing its popularity, while Typescript, Go and Kotlin are gaining in prevalence.
- Typescript is becoming a more popular language as it moves into the top five desired programming languages for next year.
- C++ and PHP do make the top 10 desired programming languages for next year, replaced by Go and Kotlin.

Implications

- Javascript and HTML/CSS being first and second respectively show that web development is still a lucrative tech skill to have. Typescript moving into the top five further reinforces this.
- With Python being the program for Artificial Intelligence (AI) and Machine Learning (ML), it has gained in demand going from fifth to third.
- SQL is still a coveted language to learn for current and those seeking to become data professionals as it remains in the top five at fourth.
- Learning mobile technology language is becoming more important.

DATABASE TRENDS





IBM Developer

SKILLS NETWORK

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- SQL based databases (My SQL, Microsoft SQL Server, PostgreSQL and SQLite) are the most used type of databases as they occupy the top four spots.
- While My SQL is currently the most popular database, PostgreSQL is the most sought out database in the future.
- NoSQL databases such as MongoDB, Redis and Elasticsearch are trending up and are the second, third and fifth most coveted databases to learn next year, respectively.
- Nine of the top 10 databases make the desired databases next year chart. Oracle is out and DynamoDB is in.

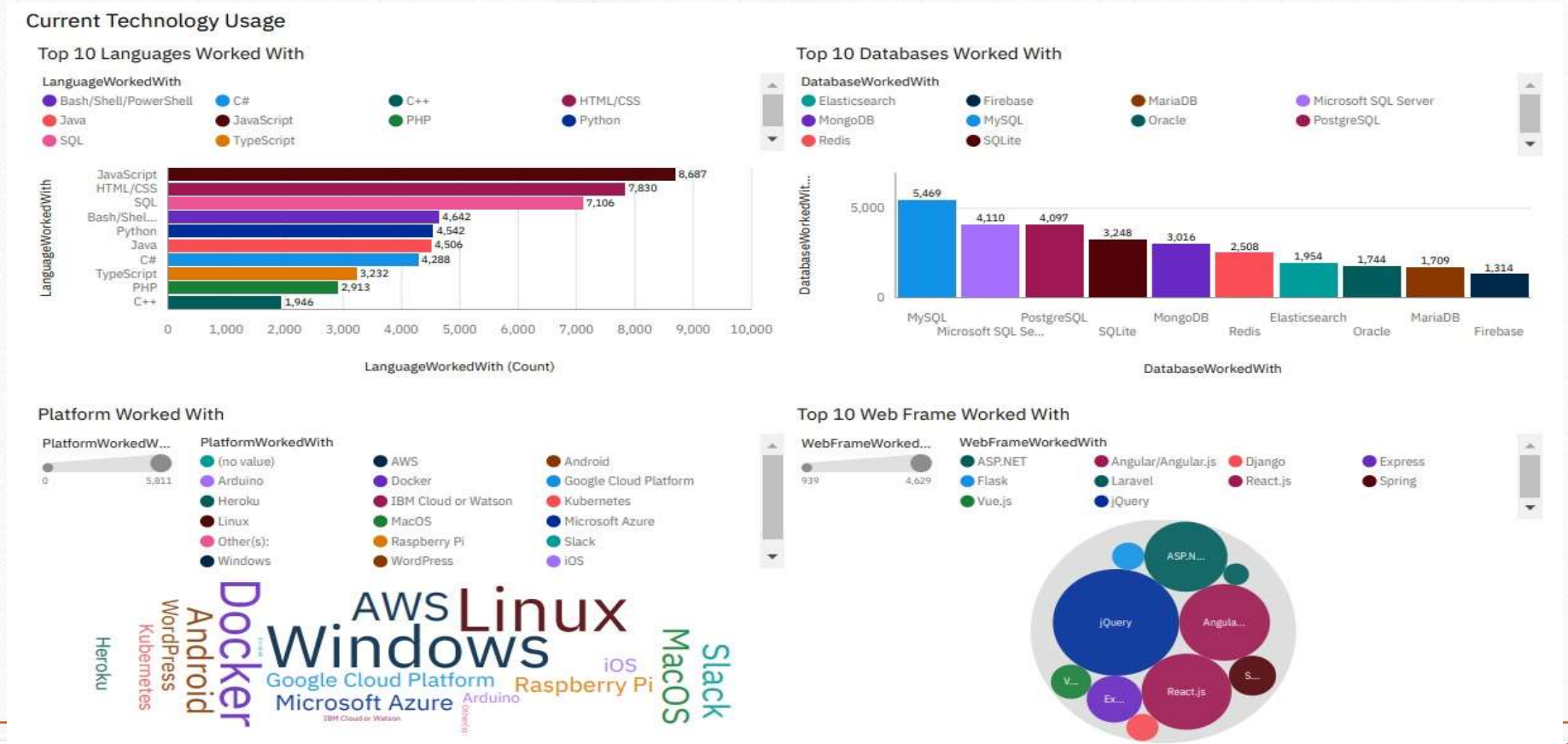
Implications

- Open-source databases are still the most popular type to use and learn as employers seek people who are proficient in them.
- NoSQL databases are becoming as sought after to learn as SQL databases.
- Oracle is losing its footprint in the database market, going from seventh currently to not even making the top 10 for desired databases next year.
- On the other hand, DynamoDB makes its entrance in the top 10 for desired databases next year in the tenth spot.

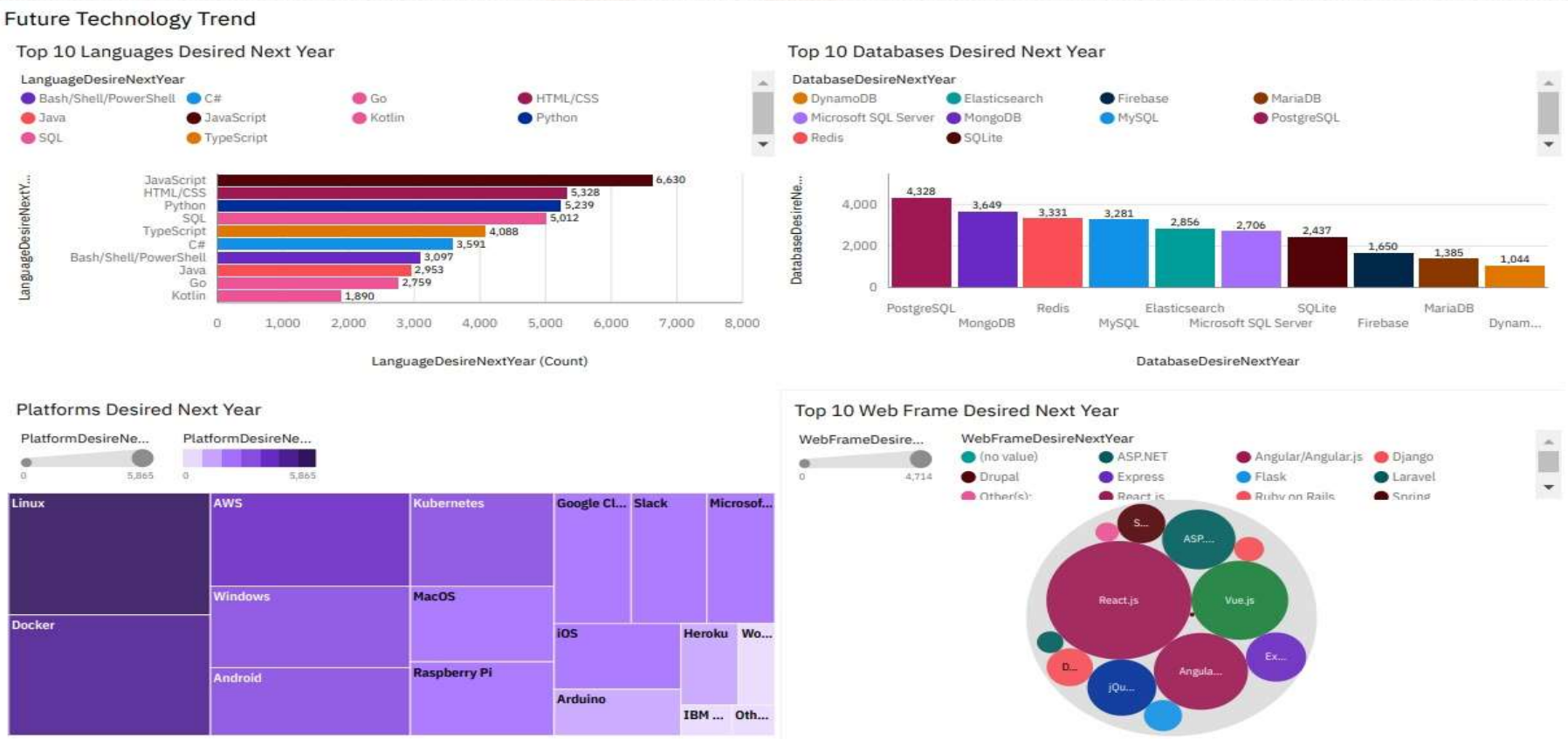
DASHBOARD

GitHub Link Cognos Dashboard:
https://github.com/jhbryden/data-analyst-capstone-project/blob/main/Capstone%20Dashboard%20(3).pdf

DASHBOARD TAB 1



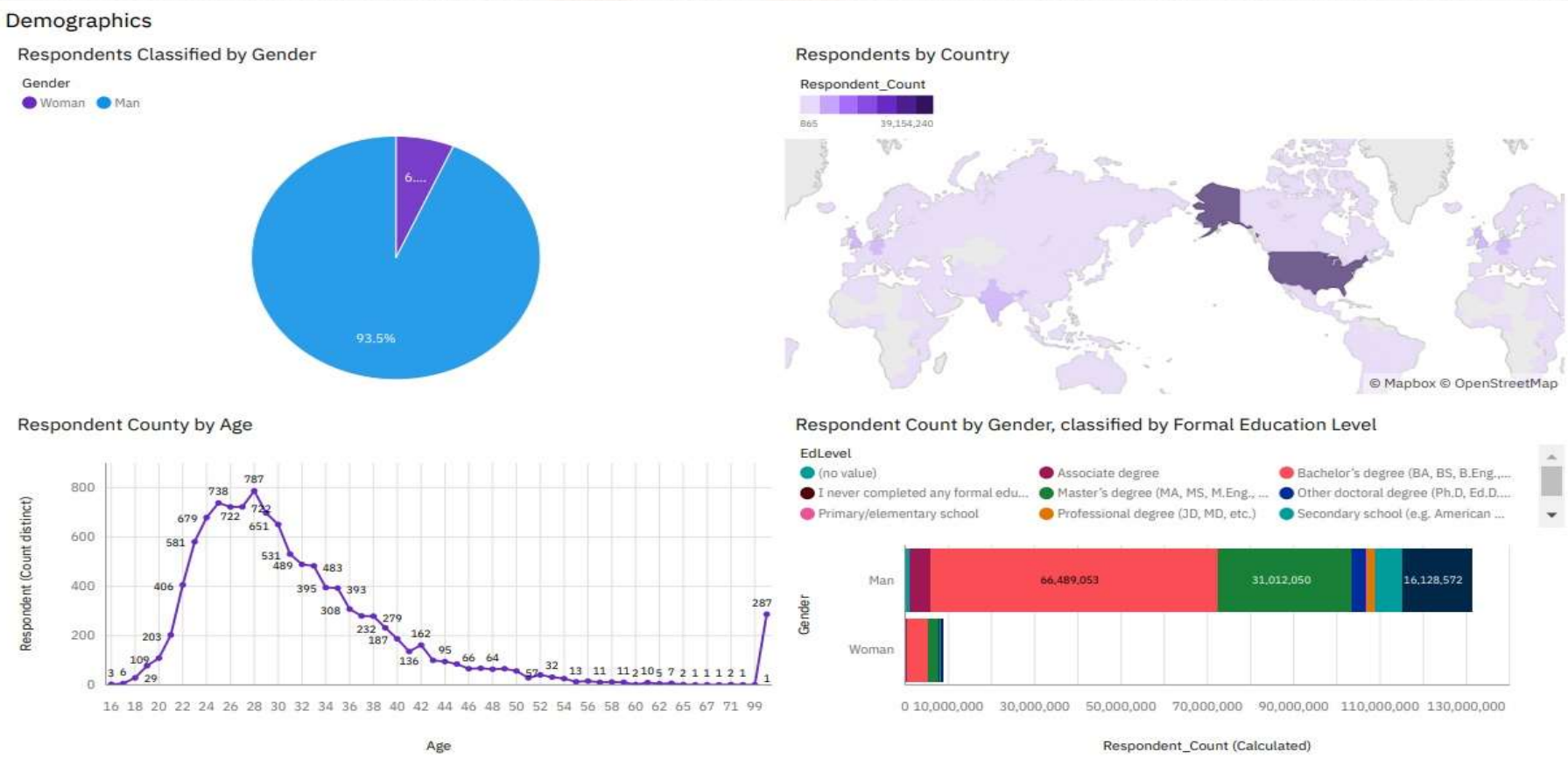
DASHBOARD TAB 2



IBM Developer

SKILLS NETWORK

DASHBOARD TAB 3



IBM Developer

SKILLS NETWORK

DISCUSSION

- How do we close the gender gap in the developer field as 93.5% of the respondents were male?
- Is having a bachelor's degree good enough as most of the respondents have a Bachelors as their highest degree?
- Spreading the tech knowledge and skills to the developing world to close the divide between the United States and the rest of the world.
- With mobile technology becoming a bigger share of the marketplace, programs like Kotlin will become a desired language to learn.
- How can we get older people to be interested in all the trends with the average age being 28 years old?
- Words like Linux, Windows, AWS and Docker will continue to be a major part of the developer and technology lexicon.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Most people in the developer field have a Bachelors' degree and are under the age of 40 with the apex age being 28.
- Web development programs such as Javascript and HTML/CSS are the most popular programs to know or learn in the field.
- Highest number of respondents were from the United States and the majority of the respondents were male.
- Javascript (language), PostgreSQL (databases), Linux (platform) and React.js (web frame) are the most desired in each category to learn for next year.

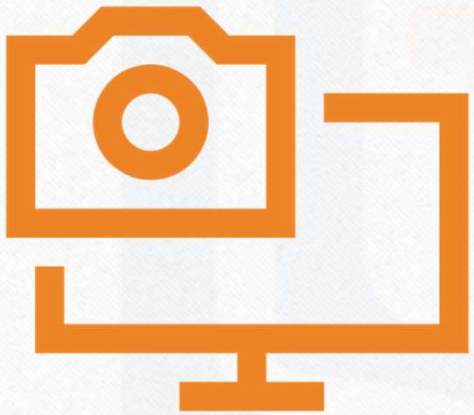
Implications

- Less developed countries need access to learn and develop the skills needed to thrive in the field.
- Web development is still a highly coveted skill.
- Data professionals need to know NoSQL just like SQL.
- Python continues to be important due to it being connected to AI and ML.

CONCLUSION

- It is imperative to stay updated about the trends in the industry as they change over time and in some cases, a very short period, in the developer field.
- The findings show that work needs to be done globally to close the developer gap between the United States and the rest of the world, in addition to closing the gender gap.
- Hiring managers can use this data to seek candidates who have experience using the popular programming languages, databases, platforms and web frames.
- Developers and potential developers can see the results and use this to know what programs to learn or expand their knowledge.

APPENDIX



Correlation

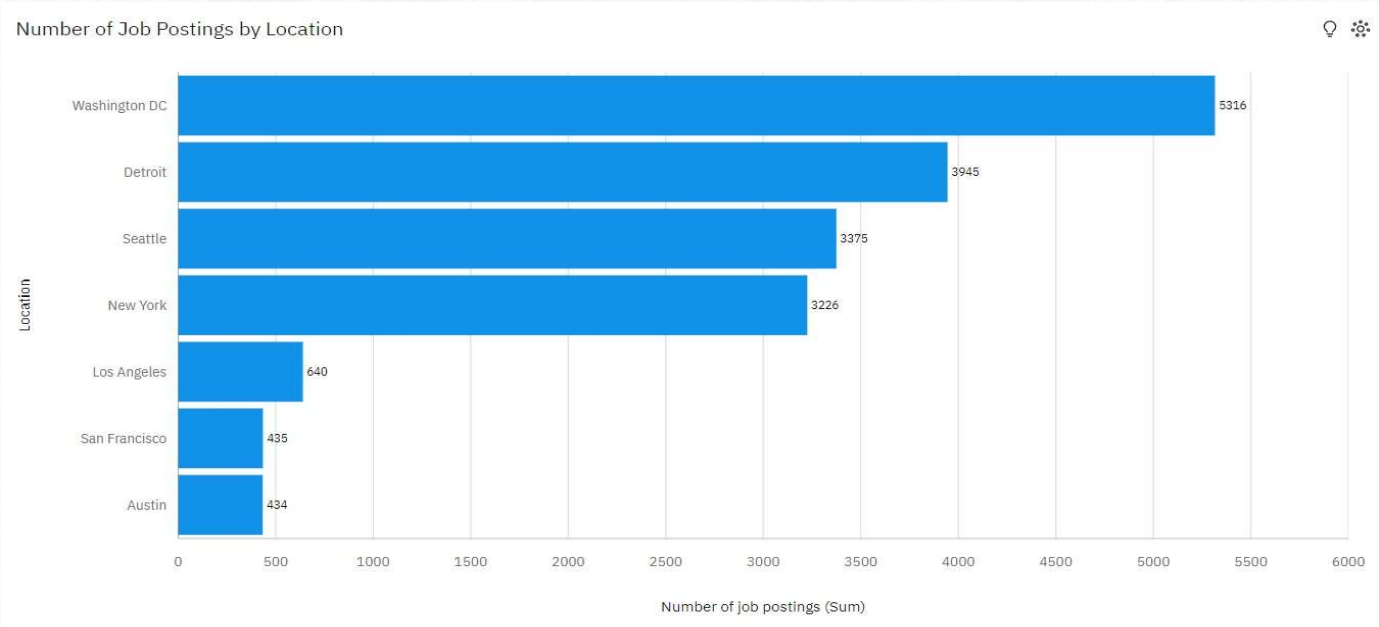
Finding correlation

Find the correlation between `Age` and all other numerical columns.

```
# your code goes here
df.corr()
```

	Respondent	CompTotal	ConvertedComp	WorkWeekHrs	CodeRevHrs	Age
Respondent	1.000000	-0.013490	0.002181	-0.015314	0.004621	0.004041
CompTotal	-0.013490	1.000000	0.001037	0.003510	0.007063	0.006970
ConvertedComp	0.002181	0.001037	1.000000	0.021143	-0.033865	0.105386
WorkWeekHrs	-0.015314	0.003510	0.021143	1.000000	0.026517	0.036518
CodeRevHrs	0.004621	0.007063	-0.033865	0.026517	1.000000	-0.020469
Age	0.004041	0.006970	0.105386	0.036518	-0.020469	1.000000

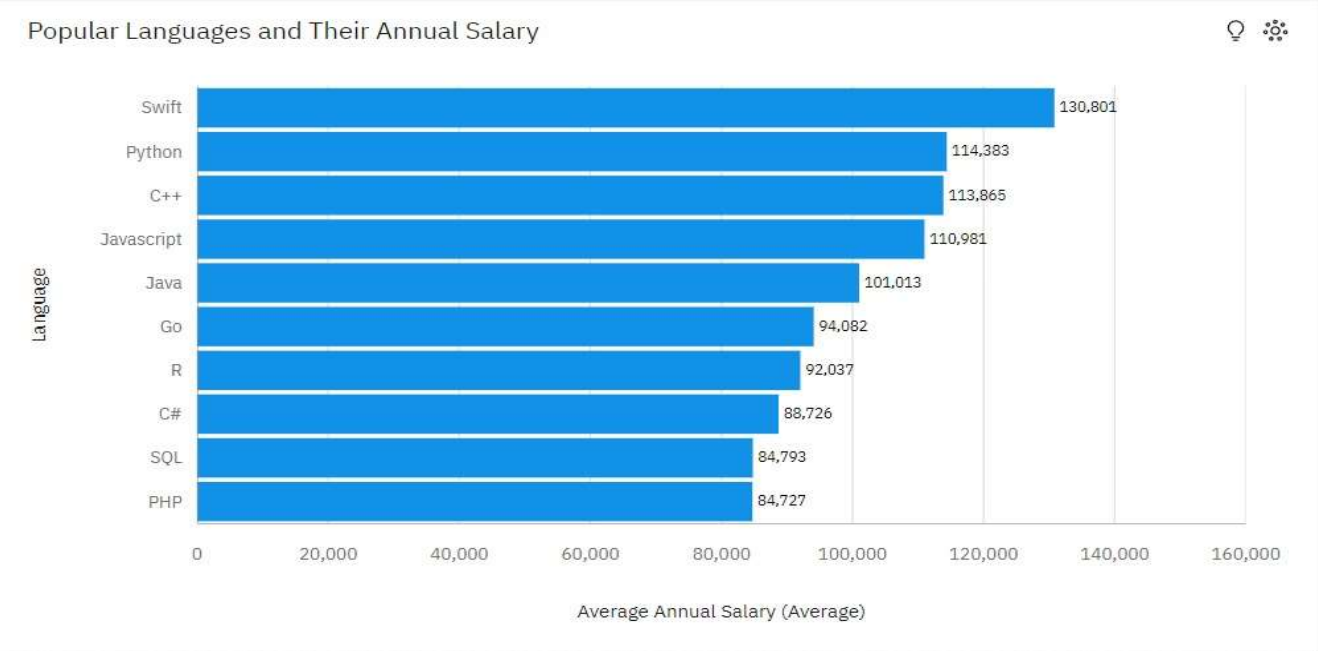
JOB POSTINGS



Bar Chart showing the amount of Job Postings by Location using Job API in descending order. Information courtesy of an Excel file named job-postings.xlsx.

Washington, DC had the most job postings, while Austin, Texas had the fewest.

POPULAR LANGUAGES



The Popular Languages and their annual salary was put together by web scraping using Python's Beautiful Soup. Swift had the highest annual salary of the popular languages, while PHP had the lowest. Salaries ranged from $84,727 to $130,801.

The bar chanted is presented in descending order of salary.

Information was web scraped from the file "popular-languages.csv."