Data Analyst Capstone Presentation

By Jens Greve

Jan 11, 2021

OUTLINE

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

IBM **Dev**oper

SKILLS NETWORK

EXECUTIVE SUMMARY

- This Capstone Project has been created to identify trends of
    - Programming languages
    - Databases
- Data from Github job postings and Stack Overflow Developer Survey 2019 have been used for the analysis with input from more than 11.000 respondents.
- Developers is a man's world
    - Number of male developers outweighs female developers 14:1
- Technologies are changing
    - C# still most common language this year, but JavaScript, SQL and HTML/CSS are emerging as most desired languages
    - Increasing desire in Python and with top salary
    - Increasing interest in MongoDB and PostgreSQL
    - jQuery is on the decline and overtaken by React and Angular, but asp.net maintain popularity ranking
- Platforms remain stabile amongst developers
    - Linux, Windows, Docker and AWS are still in the top

INTRODUCTION

- In this report we have collected and analyzed data to identify trends on emerging skills for IT developers based on Stack Overflow's annual Development Survey, the largest developer community in the world.

- The audience include developers and IT consultancy companies

- Key questions the report is trying to answer include
  - What are the top programming languages in demand?
  - What are the top database skills in demand?
  - What are the popular Platforms/IDEs?
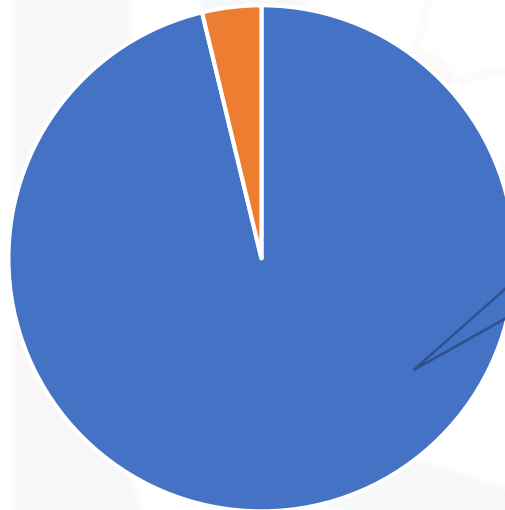 - What are the developer demographics?

METHODOLOGY

1. We have collected the top programming skills that are most in demand from various sources using APIs, web scraping and CSV files, including:
   - Github job postings
   - Training portals
   - Surveys amongst Github developers
2. We have wrangled the data to remove duplicates and correct missing values
3. We have done an initial exploratory data analysis to identify data distribution, outliers and correlation
4. We have then visualized the distribution of data, relationship, composition and comparison
5. Finally we have created a Dashboard
6. Tools used included IBM Watson Studio, IBM Cognos and Excel.

IBM **Dev**eloper

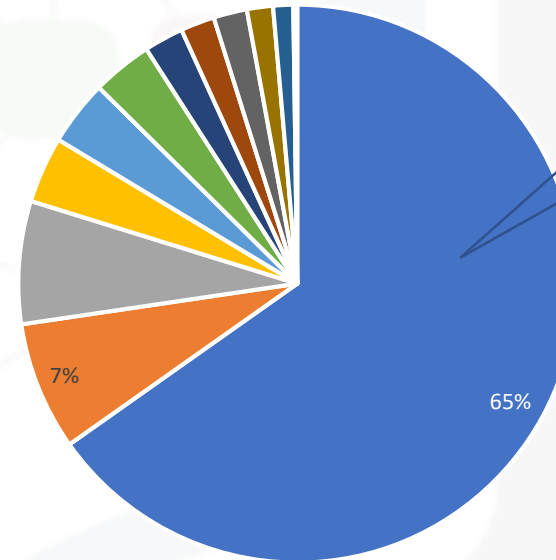SKILLS NETWORK

RESULTS - DEMOGRAPHICS

- The average (mean) age of the developer is 30,8 years old

Employment full-time / part-time

Most developers work full-time !

- Employed full-time
- Employed part-time

Developer by formal education
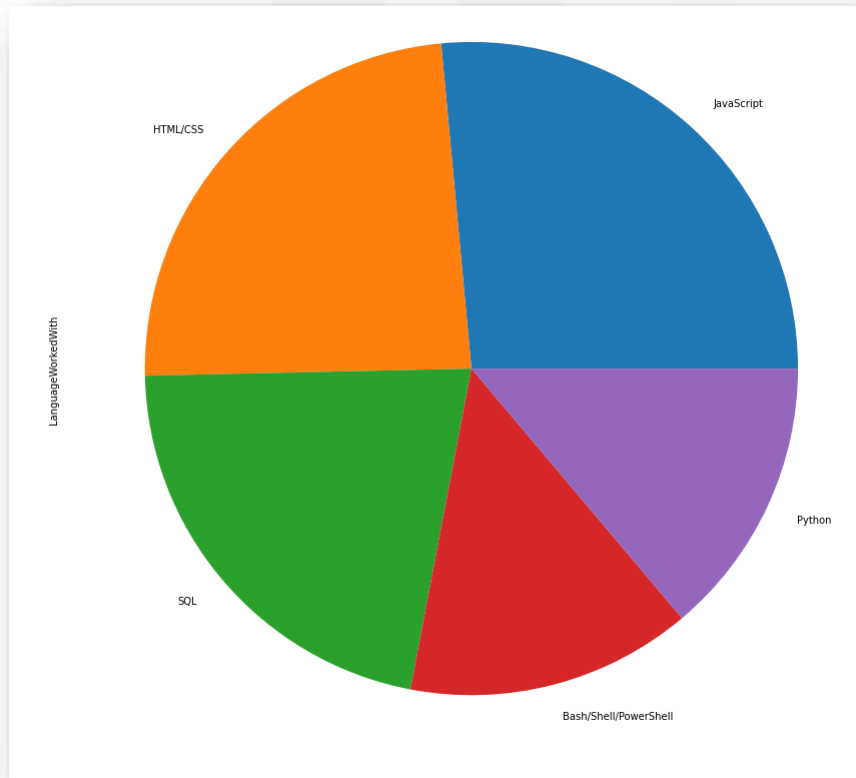
Formal computer science degree is still the most common education !

65%

7%
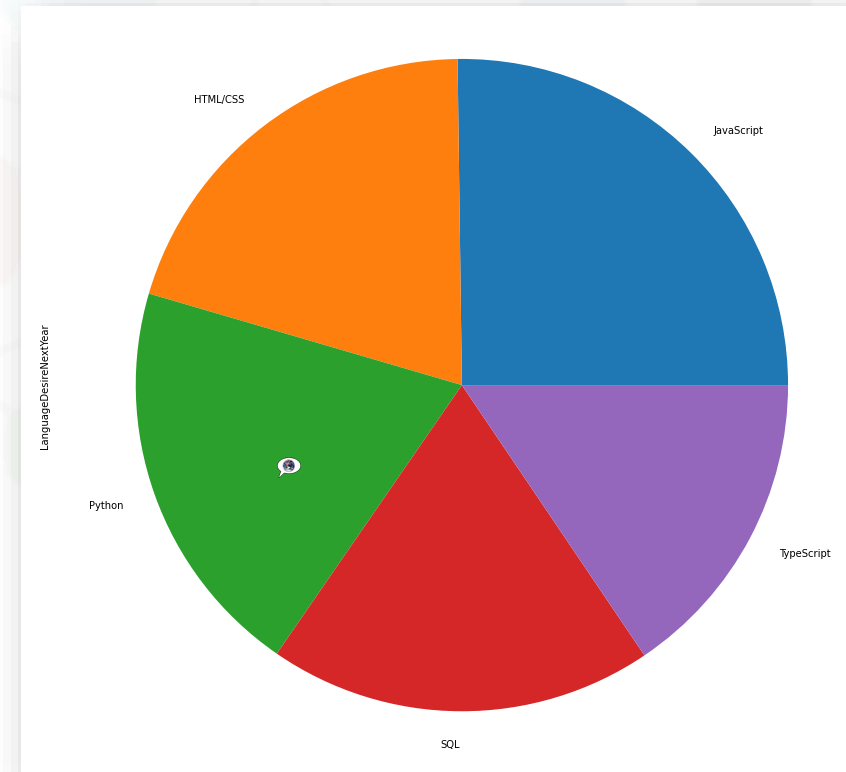
- Computer science, computer engineering, or software engineering
- Information systems, information technology, or system administration

IBM **Dev**eloper

SKILLS NETWORK

PROGRAMMING LANGUAGE TRENDS

Current Year

Next Year

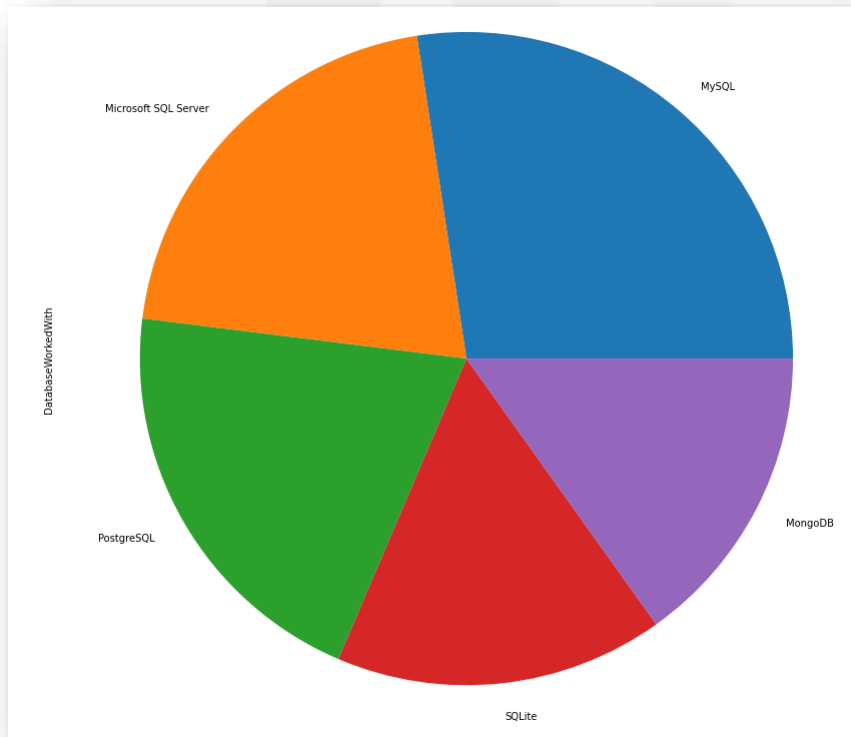# PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- HTML/CSS is remains strong
- Javascript remains strong
- Python growing in popularity
- SQL new in top 5
- Bash/shell/powershell no longer in top 5
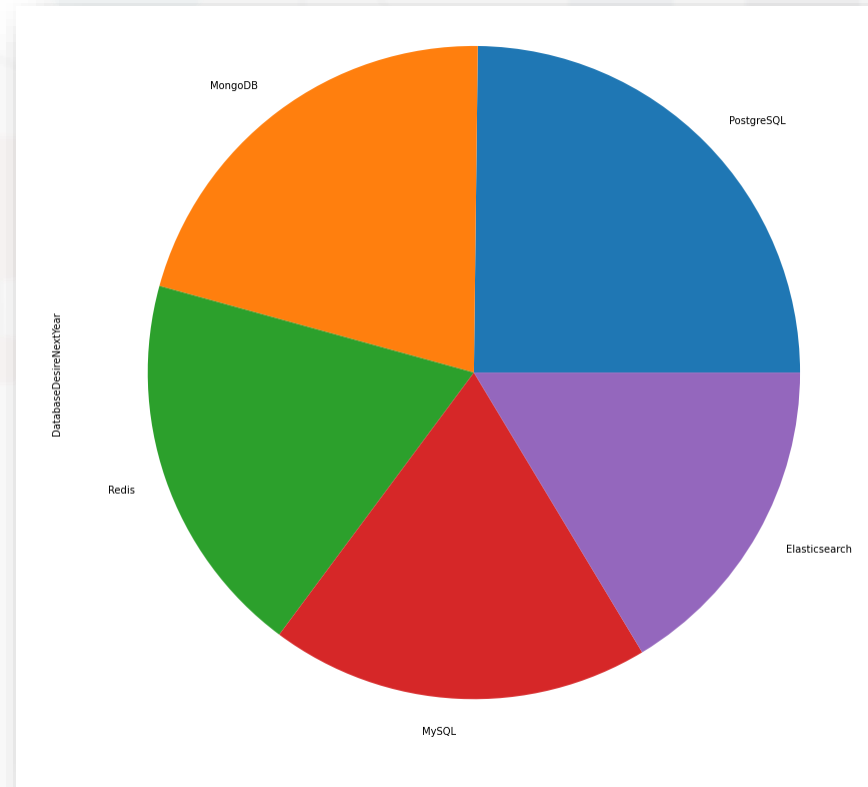
Implications

- Web development with HTML/CSS and Javascript remains strong
- Developers need to learn Python
- Is Python replacing Bash/Shell/Powershell?
- Developers should add SQL to their skillset

IBM Developer

SKILLS NETWORK

DATABASE TRENDS

Current Year

Next Year

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- MS SQL Server and SQLlite no longer in top 5
- MongoDB is now most desired database technology
- Redis and Elastic Search are newcomers in top 5
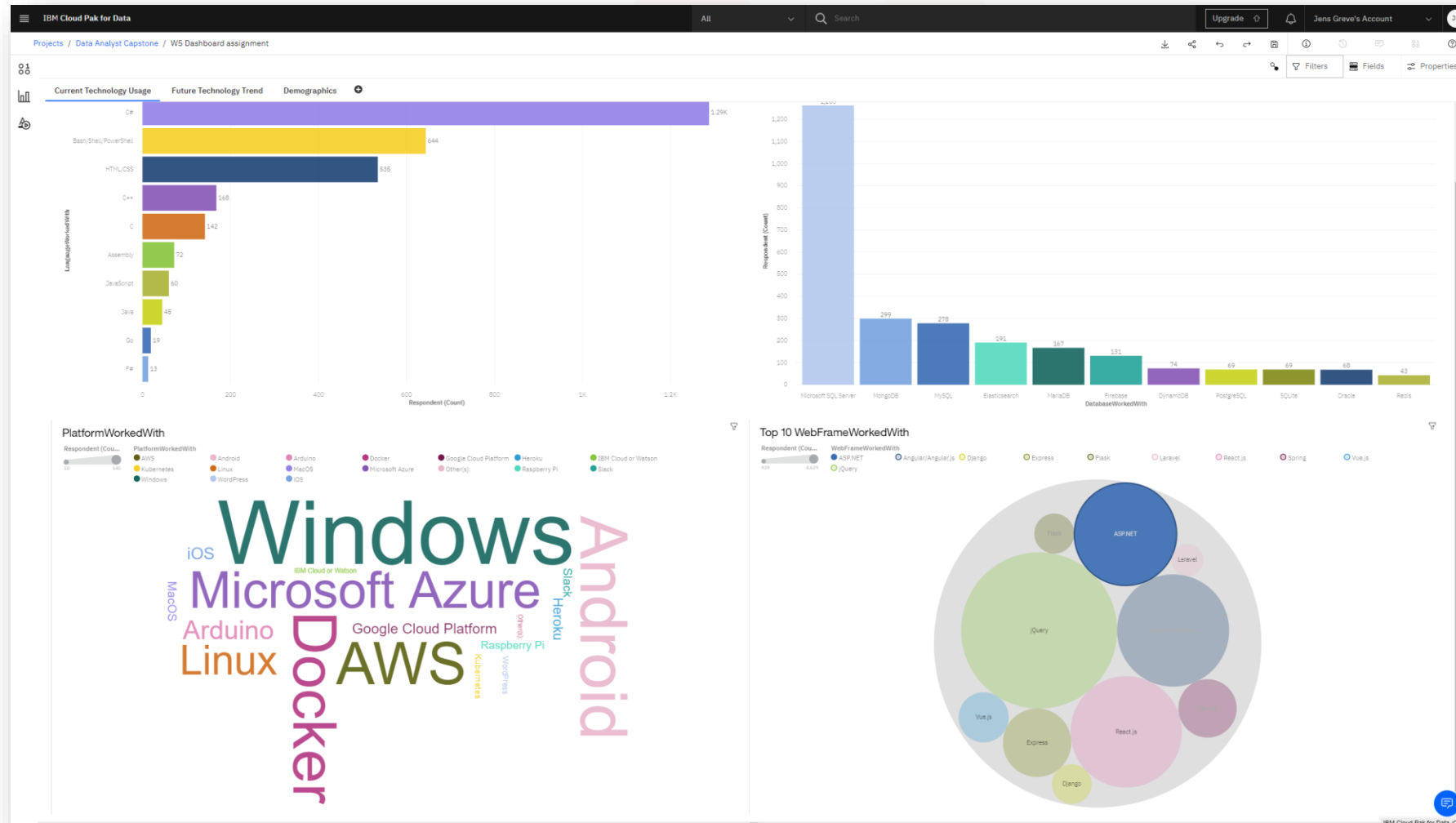- Oracle not on the list !

Implications

- Is the trend driving towards MongoDB and Elastic Search due to cloud or their capabilities
- Is trend shift driven by preference for open source?
- Is MS SQL server loosing marketshare?
- Is SQLlite loosing marketshare or equally needed for mobile?

DASHBOARD

[https://eu-gb.dataplatform.cloud.ibm.com/dashboards/913ab48d-4cec-454c-843c-b0d094e9dc44/view/5267c42c63b73dcd49deb5e407cc2d0f7437740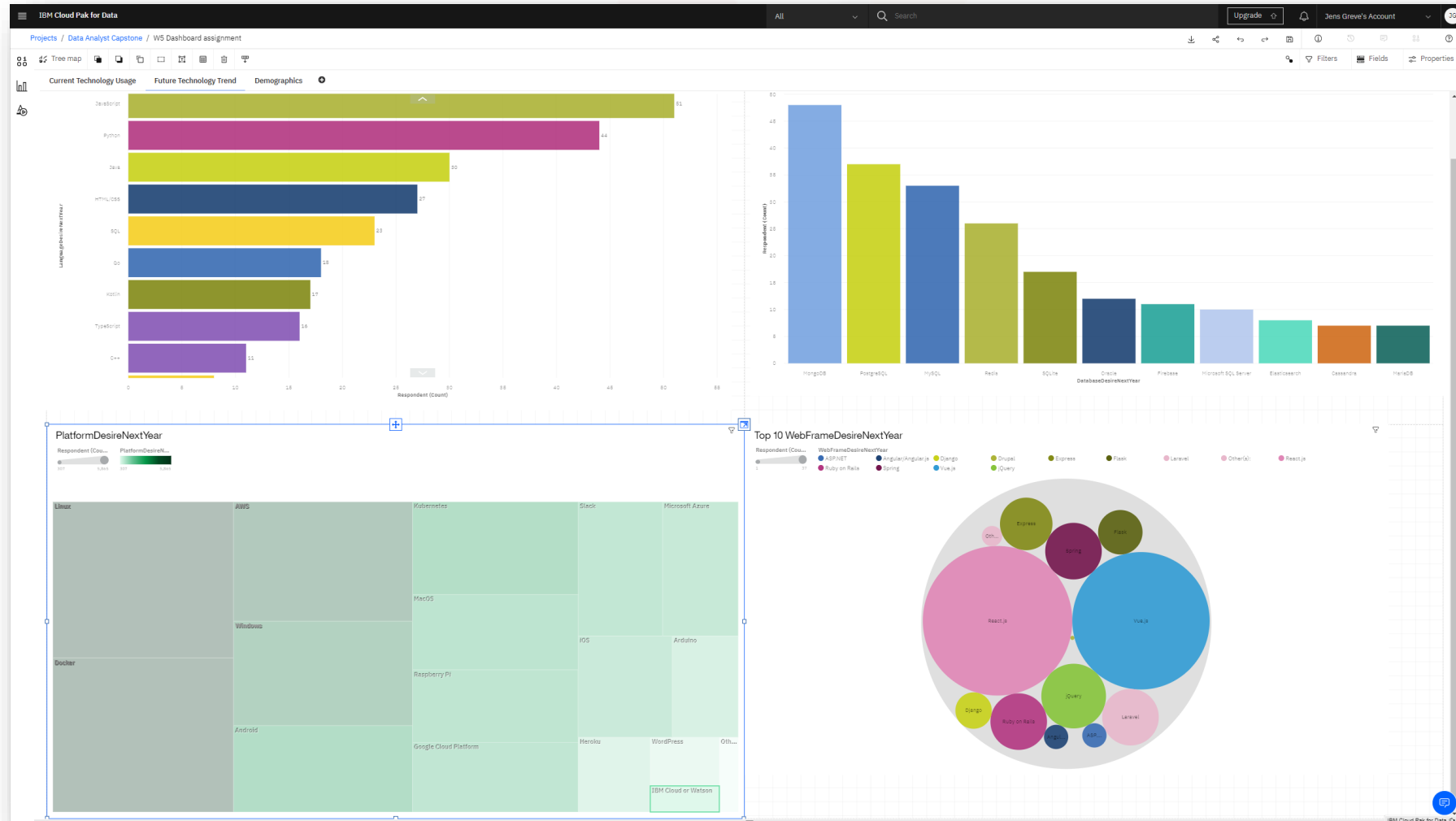cb5bb8052d2d17b4907612397a8601193c87b430edd160c67a2b8475998](https://eu-gb.dataplatform.cloud.ibm.com/dashboards/913ab48d-4cec-454c-843c-b0d094e9dc44/view/5267c42c63b73dcd49deb5e407cc2d0f7437740cb5bb8052d2d17b4907612397a8601193c87b430edd160c67a2b8475998)
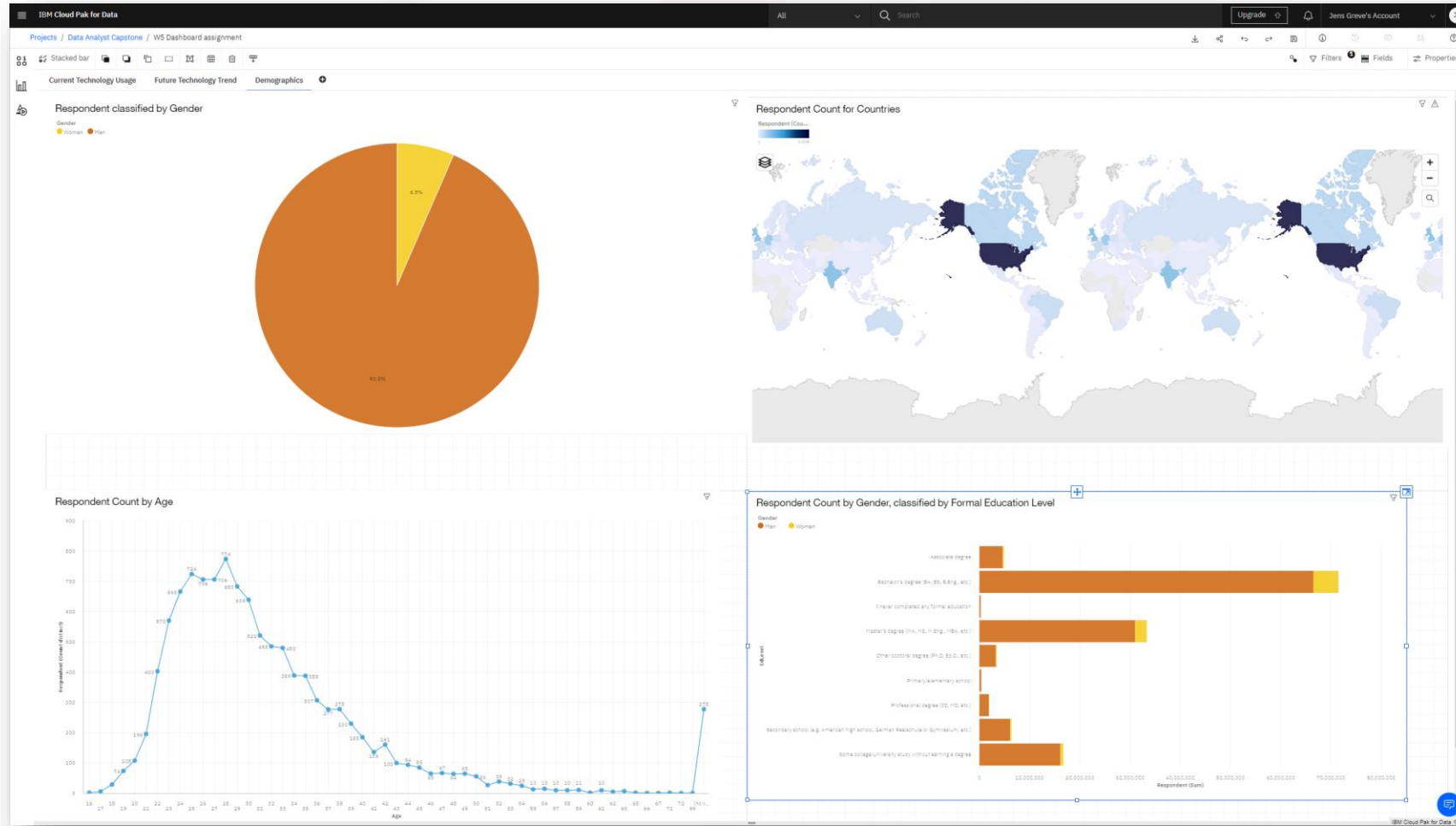
IBM **Developer**

SKILLS NETWORK

DASHBOARD TAB 1

DASHBOARD TAB 2

DASHBOARD TAB 3

DISCUSSION

- Even though the survey included more than 11.000 developers, the majority are in North America, so the results may be skewed.

- Github has most jobs for C, but C is not on the top 5 list of languages. This discrepancy should be further investigated.

- Some discrepancies found between popular languages between the sources, e.g. for C#

- Move to cloud databases, but no significant change in platforms/IDEs on cloud.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Average developer is a 31 year old man with a computer science degree and 100kUSD salary

- Web technologies remain desired, but jQuery is on the decline and overtaken by React and Angular, but asp.net maintain popularity ranking

- Increasing desire in Python and with top salary

- Increasing interest in MongoDB and PostgreSQL

- Platforms remain stabile amongst developers (Linux, Windows, Docker and AWS)

Implications

- 14x more men than women in sw development

- Technologies are changing and more towards web and cloud databases

- Do Javascript developers stay with MySQL and PostgreSQL next year?

- Who uses SQLite. Mobile developer and will they keep their preference next year.
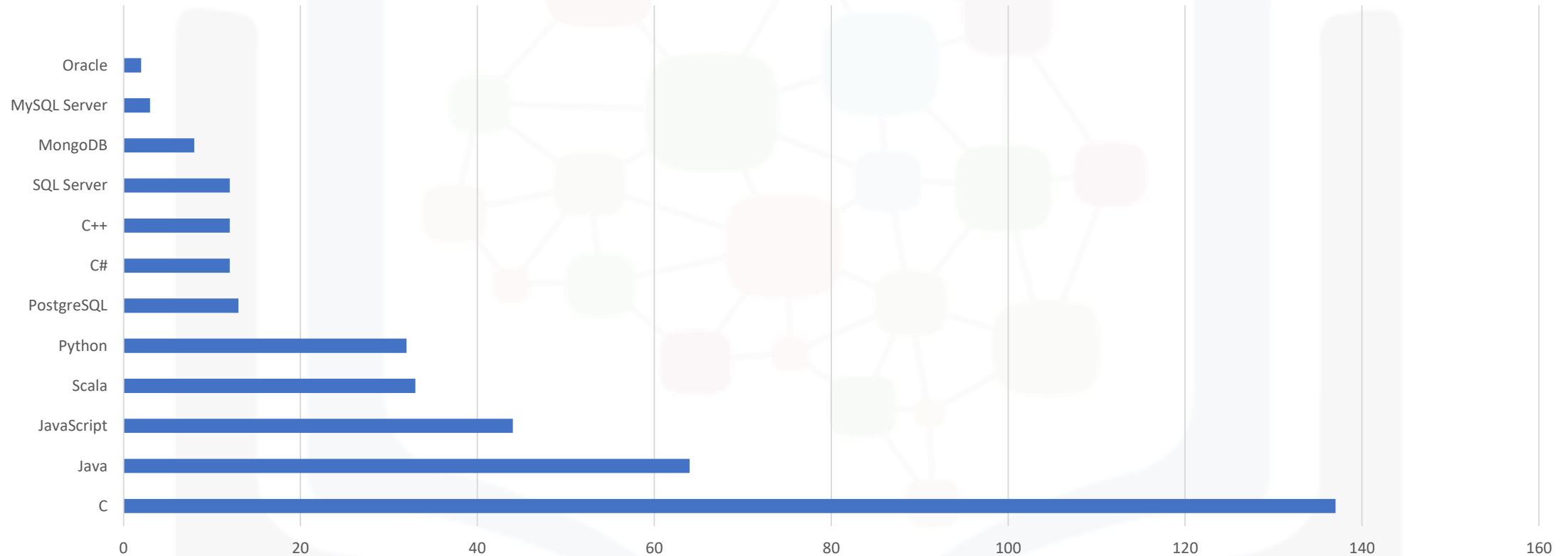
IBM Developer

SKILLS NETWORK

CONCLUSION

- JavaScript and HTML would be the most popular programming language next year

- Increasing desire for Python

- Increasing desire for cloud database like MongoDB and ElasticSearch

- Linux, Windows, Docker and AWS remain strong

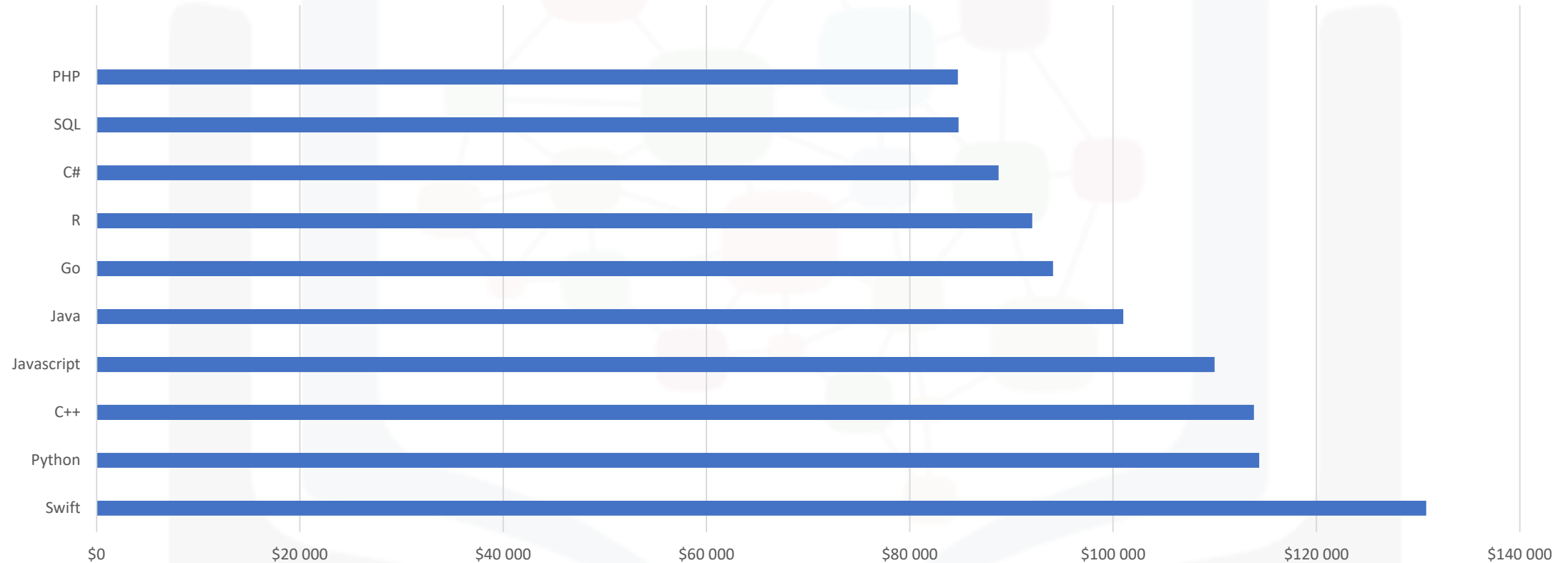- Developers working with JavaScript and HTML prefer working with MySQL

IBM **Dev**eloper

SKILLS NETWORK

APPENDIX

- Include any relevant additional charts, or tables that you may have created during the analysis phase.

GITHUB JOB POSTINGS

Source: https://jobs.github.com/api

POPULAR LANGUAGES BY AVERAGE SALARY

IBM Developer

SKILLS NETWORK

RESULTS

	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