IBM Data Analyst Capstone Project: Analysis on Technology Skills and Trends

Ing. Anna Rývová

April 2024

OUTLINE

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

IBM Developer

SKILLS NETWORK

EXECUTIVE SUMMARY

- We use data analysis to analyze the current state of technology used in IT and the trends for the next year.

- Data was gathered from a Stack overflow survey, IBM site and Github job postings. It was collected, cleaned, subjected to exploratory analysis and visualized on Cognos dashboards.

- Findings:
  - The most used DB is MySql and the Javascript programming language.
  - Next year, most users want to learn PostgreSql database and javascript programming language.
  - Majority of the respondents are males from USA, 28 years of age.

IBM Dev loper

SKILLS NETWORK

INTRODUCTION

- Using data analysis, we defined current and future trends in the necessary knowledge in the field of programming languages, databases, development platforms and web frameworks.

- The following inquiries were investigated using the data:

- Point3
  - What is the most used programming language today?
  - What is the most used database today?
  - What are currently popular development platforms and web frameworks?

- The work is intended for IT professionals, HR managers and all other interested parties in the IT sector who are interested in current and future requirements for IT skills.

IBM Developer

SKILLS NETWORK

METHODOLOGY

- Data in several formats were gathered using the Github jobs **API on Python.**

- To obtain the names of the programming languages and their yearly salaries, the IBM **website was scrapped.** The dataset from a 2019 Stack Overflow developer survey was downloaded and saved.

- **Python** was used to clean and analyze the data.

- **Exploratory data analysis** was performed to assess data distribution, presence of outliers, and correlation between different columns in the data set.

- Charts, graphs and dashboards were created using Python and **Cognos analytics** to visualize the data.
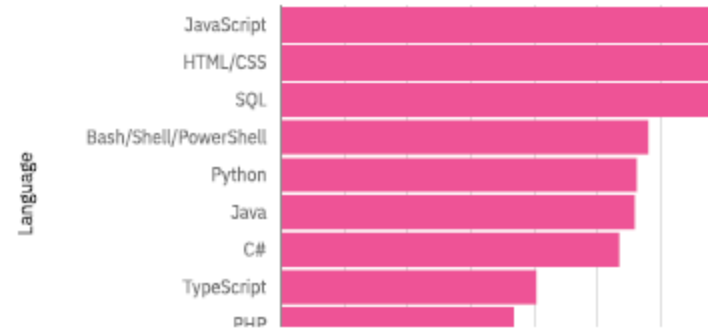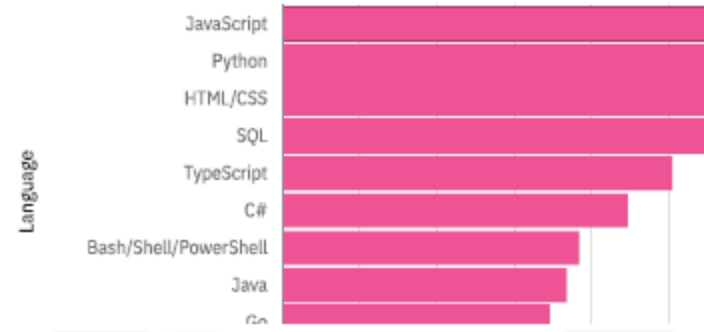
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year

Next Year

Top 10 Languages This Year

Language (y-axis)
- JavaScript
- HTML/CSS
- SQL
- Bash/Shell/PowerShell
- Python
- Java
- C#
- TypeScript
- PHP

Top 10 Languages Desire Next Year

Language (y-axis)
- JavaScript
- Python
- HTML/CSS
- SQL
- TypeScript
- C#
- Bash/Shell/PowerShell
- Java
- Go

IBM **Dev**eloper

SKILLS NETWORK

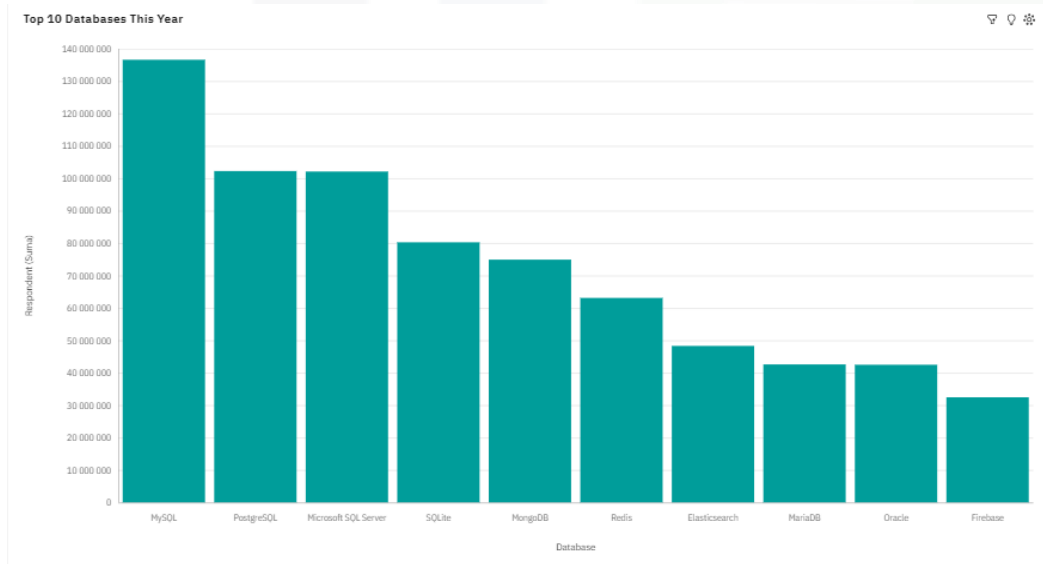# PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Javascript, HTML/CSS and SQL are the most used currently

- The most desired next year languages are Javascript, Python, HTML/CSS and SQL

- In the last 10th place is currently PHP, next year Go, the other 9 languages are the same just in a different order
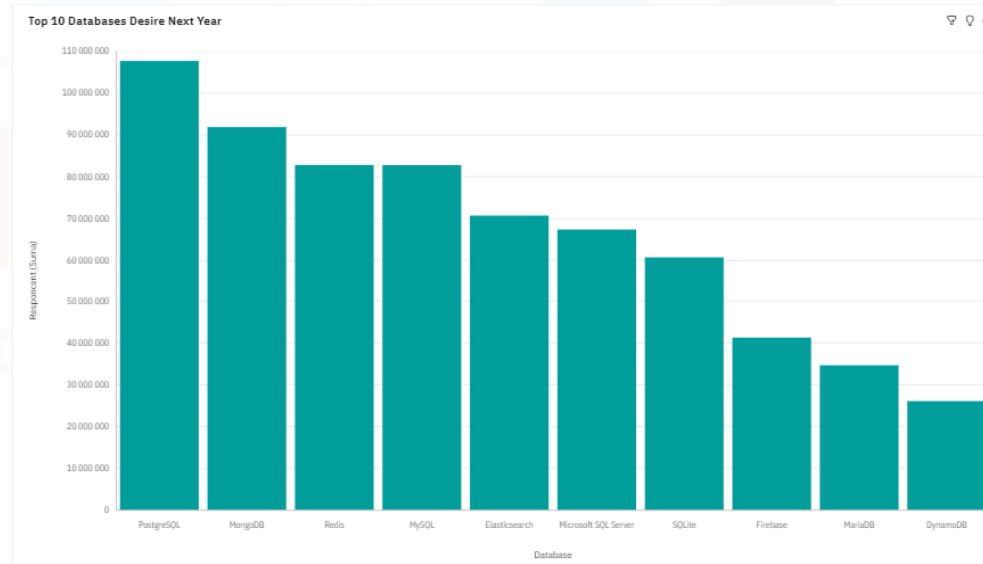
Implications

- The greatest interest will continue to be in javascript and HTML/CSS

- Python will become more prominent in the future

- Interest in Bash/Shell/PowerShell scripts will decline in the future

IBM Developer

SKILLS NETWORK

DATABASE TRENDS

Current Year

Next Year

IBM Developer

SKILLS NETWORK

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Currently, the most used DBs are MySQL, PostgreSQL and Microsoft SQL Server.
- In the future, the greatest interest is in DB PostgreSQL, MongoDB, Redis and MySQL.
- In the future, the relatively new DB systems Redis and Elasticsearch will come to the fore, and on the contrary, the traditional Oracle DB will drop out of the TOP 10.
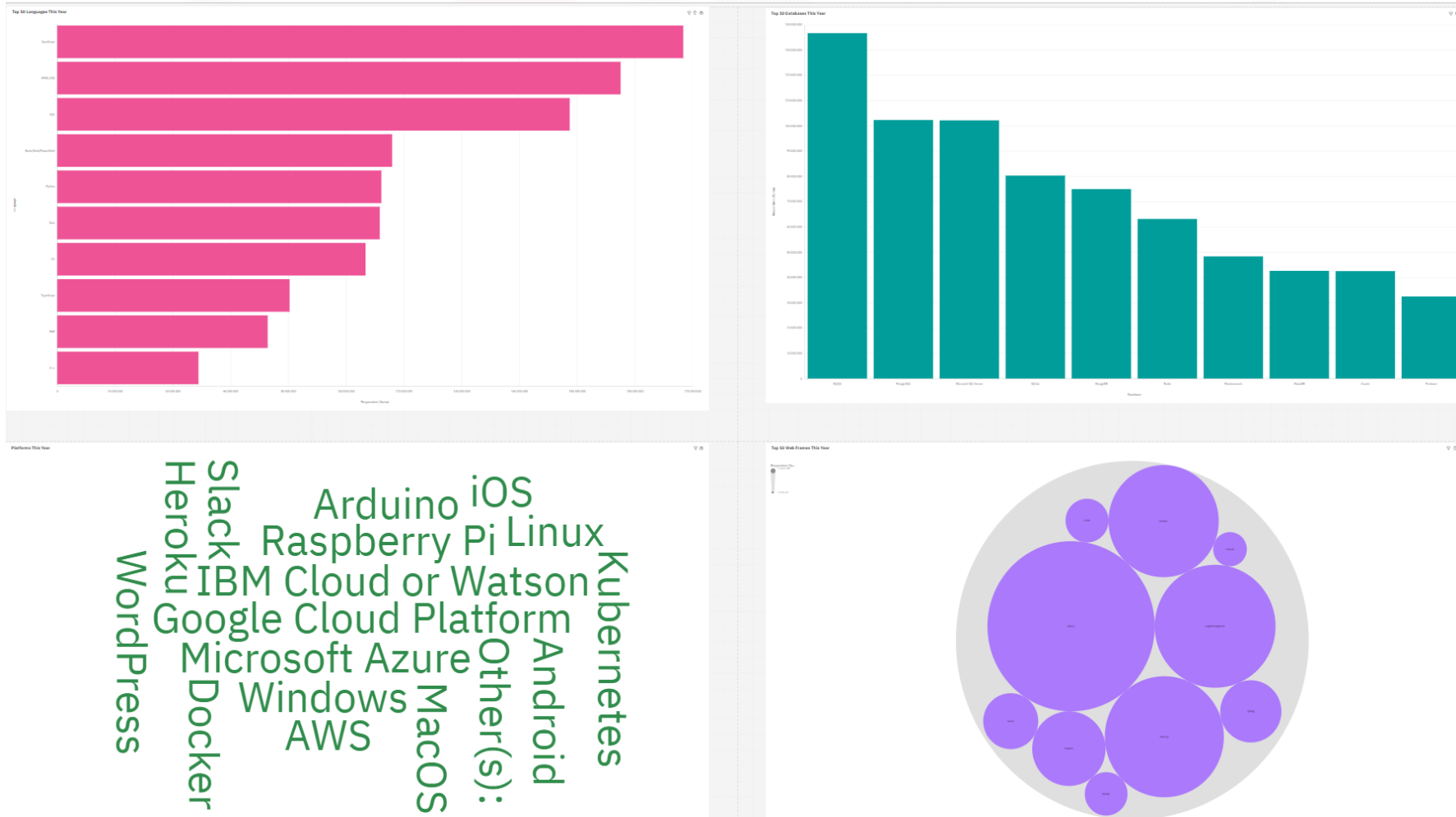
Implications

- SQL DB will continue to be the TOP tool for data analysts.
- The greatest interest is and will continue to be in open source systems.
- Microsoft SQL Server and Oracle DB will lose influence, Oracle will even drop out of the TOP 10

IBM Developer

SKILLS NETWORK

DASHBOARD

https://eu2.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FIBM%2BData%2BAnalyst%2BCapstone%2BProject&action=view&mode=dashboard&subView=model0000018e8a2d3775_00000000
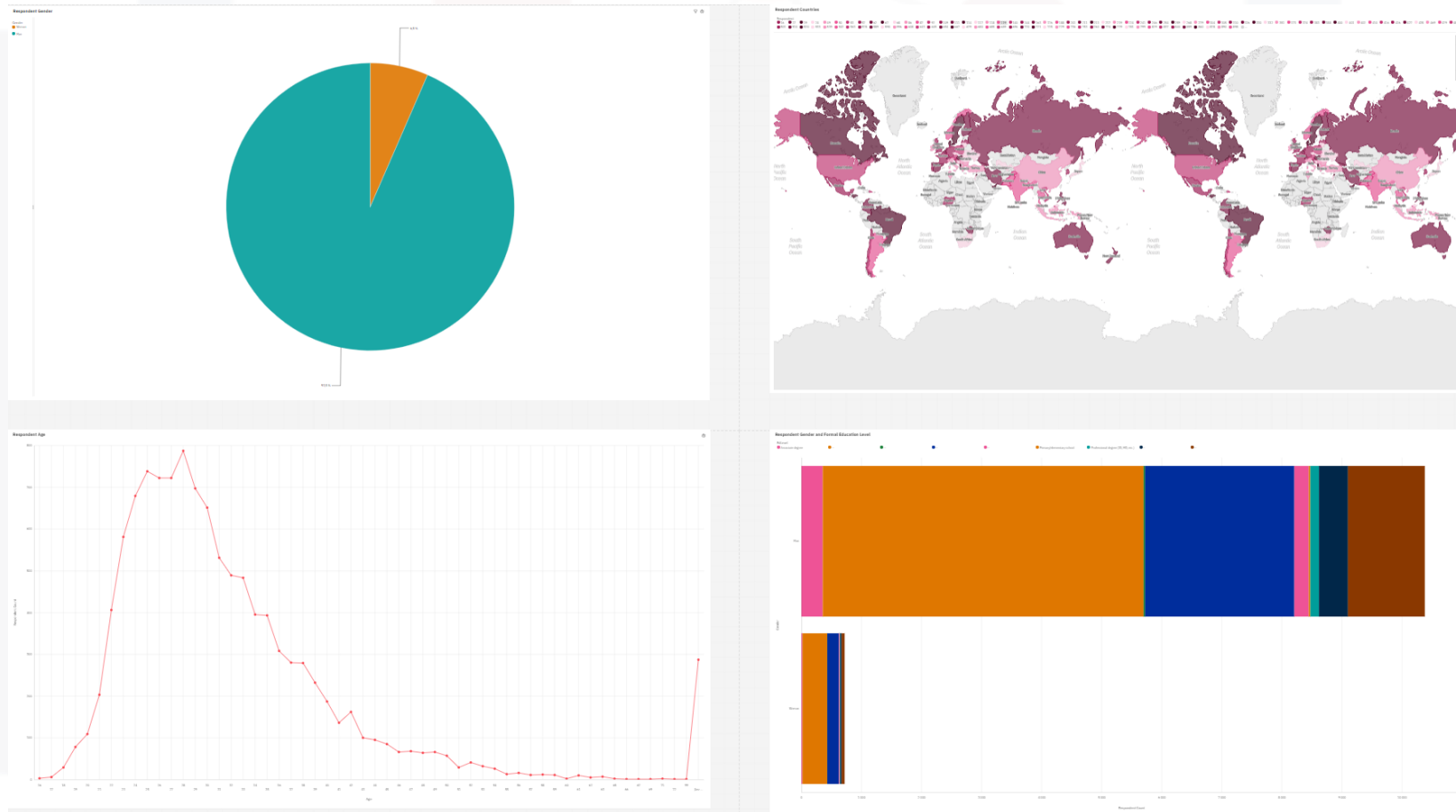
DASHBOARD TAB 1

IBM Developer

SKILLS NETWORK

DASHBOARD TAB 2

DASHBOARD TAB 3

DISCUSSION

- The rising popularity of mobile app development – Kotlin

- What will be the importance of Oracle in the future?

- Pay attention to education and development in Africa, Southeast Asia and some areas of Europe.

- Wide gender gap in the Technology sector.

IBM **Developer**

SKILLS NETWORK

OVERALL FINDINGS & IMPLICATIONS

Findings

- Most respondents are aged 24-28

- The most popular languages are for web development

- Most users are under 40 years old

- Most users want to learn javascript, Python, PostgreSQL and Mongo DB in the future.
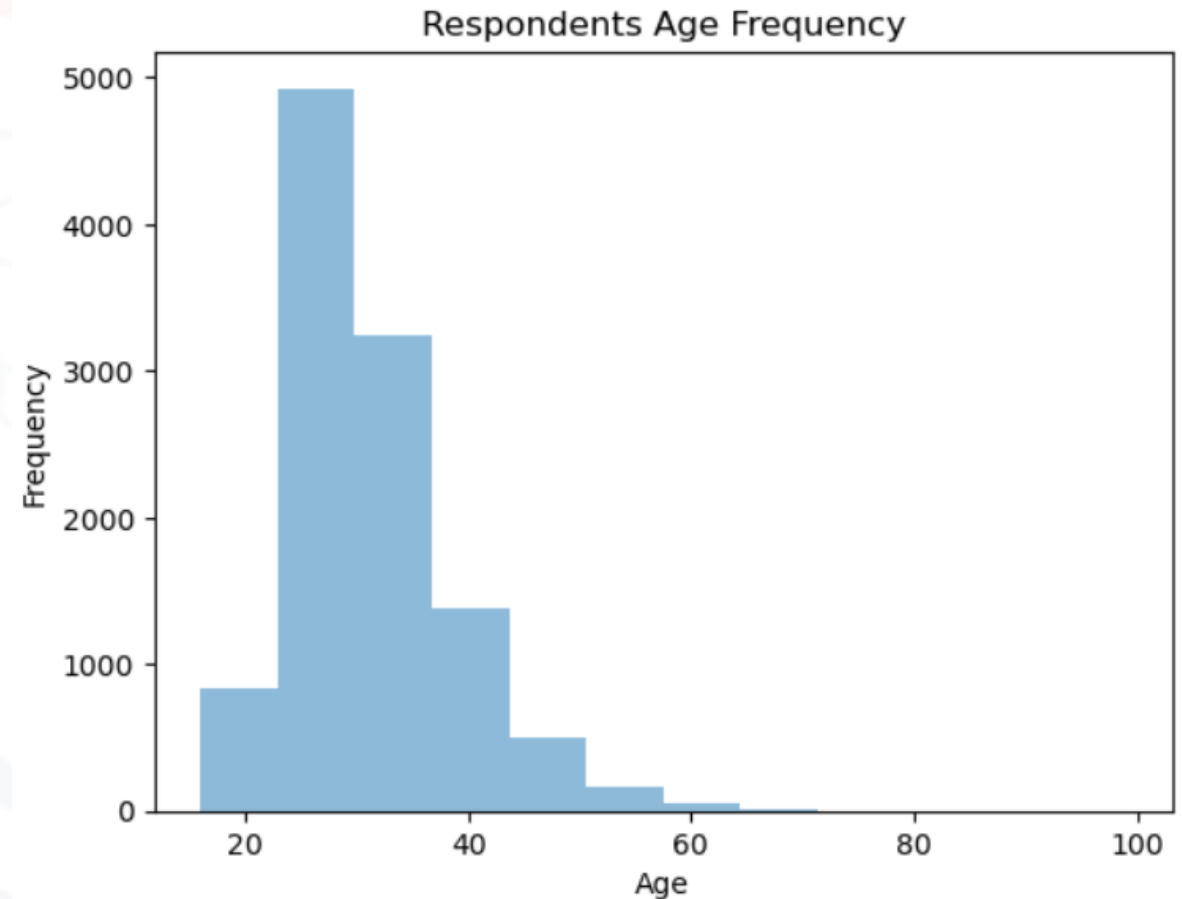
Implications

- Web development remains very lucrative

- It is important to master SQL DB (PostgreSQL) and NoSQL DB (Mongo DB)Implication 2

- The need to support education in less developed countries
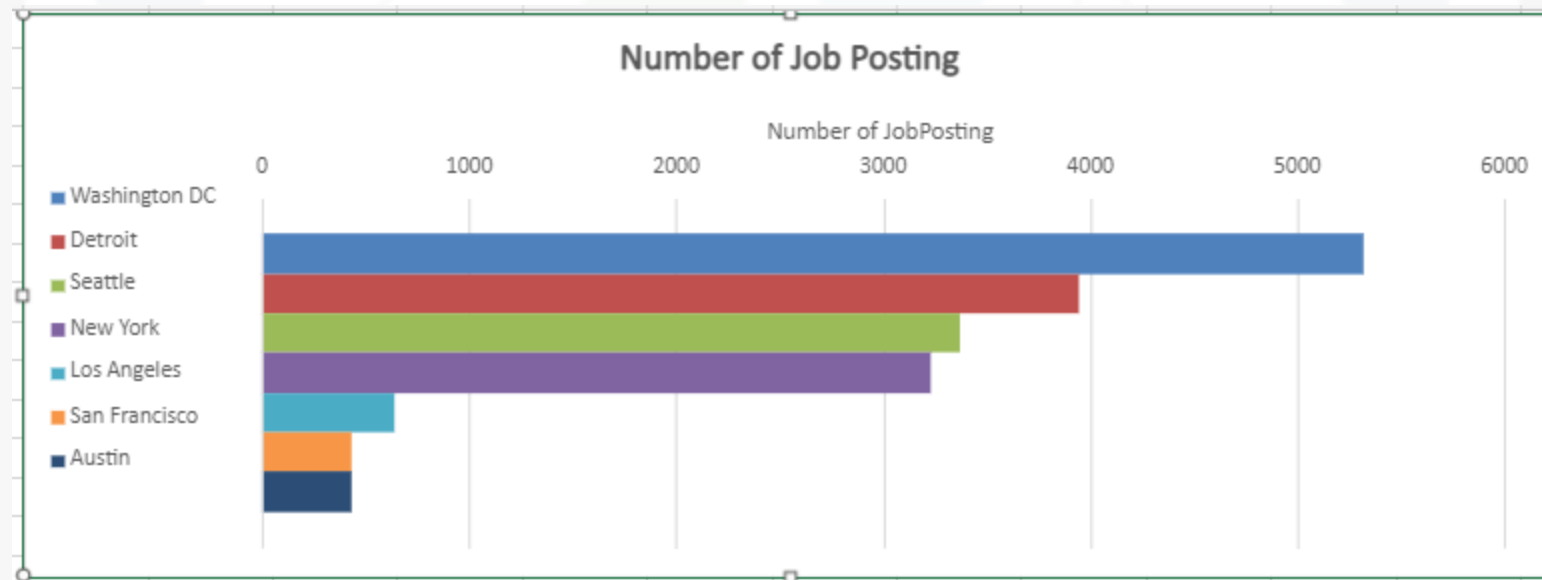
IBM **Dev**eloper

SKILLS NETWORK

CONCLUSION

- Due to the constant development of trends in the Tech sector, it is very important to always follow the development trends.

- The need to support education in less developed countries

- Wide gender gap in the Technology sector.

- Traditional DB systems (Microsoft SQL Server, Oracle DB) are losing their position (Oracle even drops out of the TOP 10)

IBM Developer

SKILLS NETWORK

APPENDIX

Respondents Age Frequency

JOB POSTINGS

Number of Job Posting

Number of JobPosting

	0	1000	2000	3000	4000	5000	6000

- Washington DC
- Detroit
- Seattle
- New York
- Los Angeles
- San Francisco
- Austin

POPULAR LANGUAGES

Annual Average Salary

Annual Average Salary

| | $0,000 | $20,000 | $40,000 | $60,000 | $80,000 | $100,000 | $120,000 | $140,000 |

Swift
Python
C++
Javascript
Java
Go
R
C#
SQL
PHP

IBM **Dev** oper

SKILLS NETWORK