

IBM Data Analyst Capstone Project

Veranika Skachkova
November 2022

OUTLINE



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

EXECUTIVE SUMMARY



- The most popular languages are scripting languages: JavaScript and Python.
- An interest to the markup languages HTML/CSS is stays on the high level. It can be explained by web-development relevance
- Linux family of the platforms are the most popular. It may be because of their availability as so as their usability for the developing
- There are significant gender difference in the IT-sphere
- The majority of the respondents has a Bachelor Degree

INTRODUCTION



- Labor market in the IT-sphere has close connection with this branch development
- An increasing of the data amount and database complicity leads to increase of the demands of the tools for processing of this data
- In this connection new more convenient and powerful tools emerge all the time

METHODOLOGY



- The dataset used for this report is the randomized subset that is $1/10^{\text{th}}$ of the initial data set from:

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

- Data was being wrangled and analyzed using **Pandas**, **MathPlotLib** and **NumPy** Python libraries
- For visualization **IBM Cognos Analytics** was used

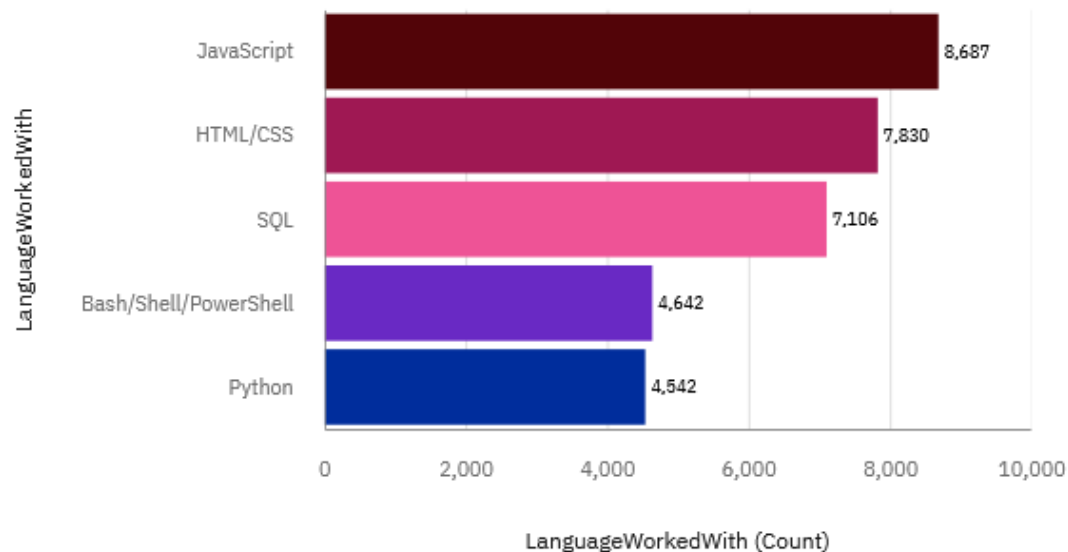
PROGRAMMING LANGUAGE TRENDS

Current Year

Top 5 Language Worked With

LanguageWorkedWith

● Bash/Shell/PowerShell ● HTML/CSS ● JavaScript ● Python ● SQL

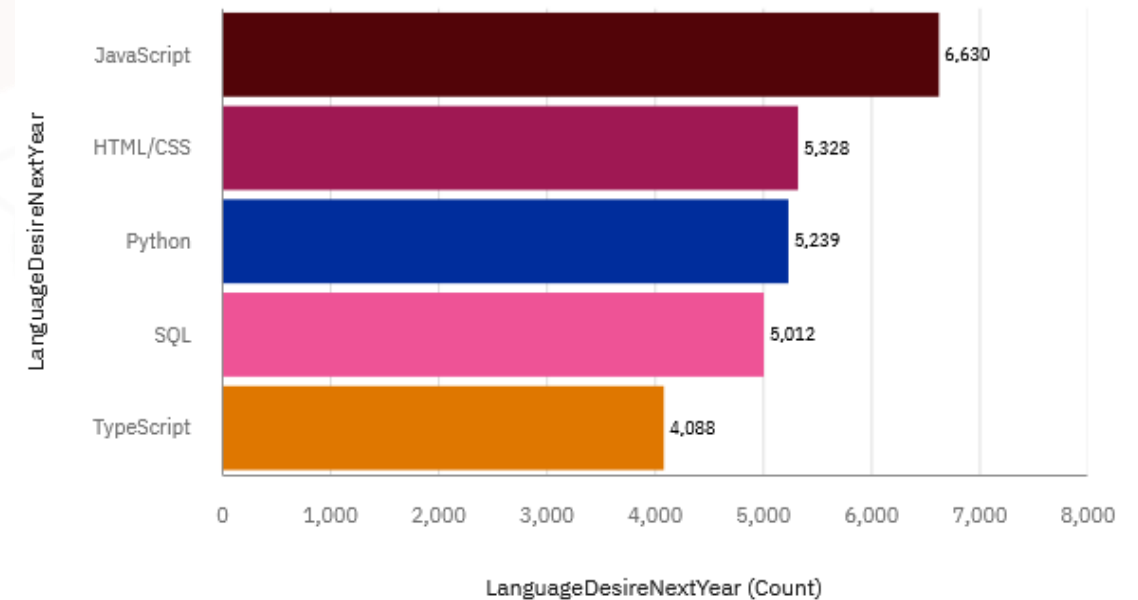


Next Year

Top 5 Language Desire Next Year

LanguageDesireNextYear

● TypeScript ● SQL ● Python ● JavaScript ● HTML/CSS



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Scripting languages (**JavaScript** and **Python**) are the most popular
- Markup language **HTML/CSS** remains relevant
- We can see that an interest to the **Python** is growing
- One of the top 5 programming languages is the query language **SQL**

Implications

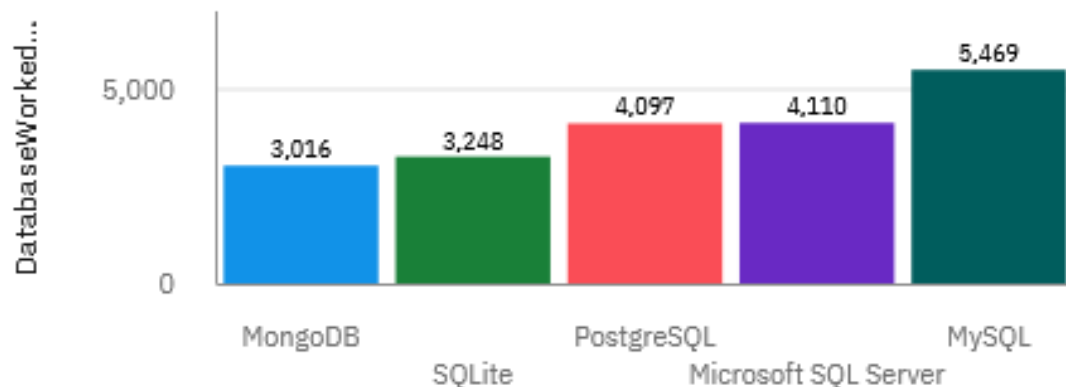
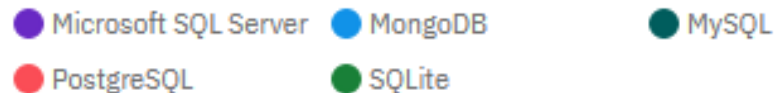
- High demand of the **JavaScript** and **HTML/CSS** reveals the web-development relevance
- The interest growth to the **Python** is possibly because it has convenient libraries that allow use **Python** for different aims from high performance computing, machine learning and big data processing
- **SQL** popularity is related to the constant growth of amount of data

DATABASE TRENDS

Current Year

Top 5 DatabaseWorkedWith

DatabaseWorkedWith

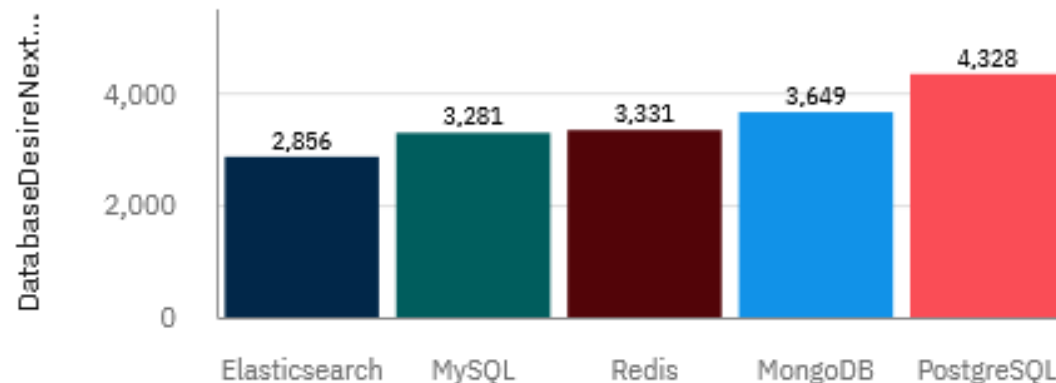


DatabaseWorkedWith

Next Year

Top 5 DatabaseDesireNextYear

DatabaseDesireNextYear



DatabaseDesireNextYear

DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Most popular databases are MySQL, PostgreSQL and MongoDB
- Microsoft SQL Server and MySQL are losing popularity
- There is a growing interest to scalable and high performance databases

Implications

- Popularity growth of the PostgreSQL associates with its cross-platform and high operating speed in contrast with Microsoft SQL Server and MySQL
- The growth of the MongoDB relevance can be explained by its great compatibility with the most common programming languages and its high performance
- As can be seen from the charts before, there is a growing demand on noSQL databases which allow to avoid cumbersome complex queries while working with big databases

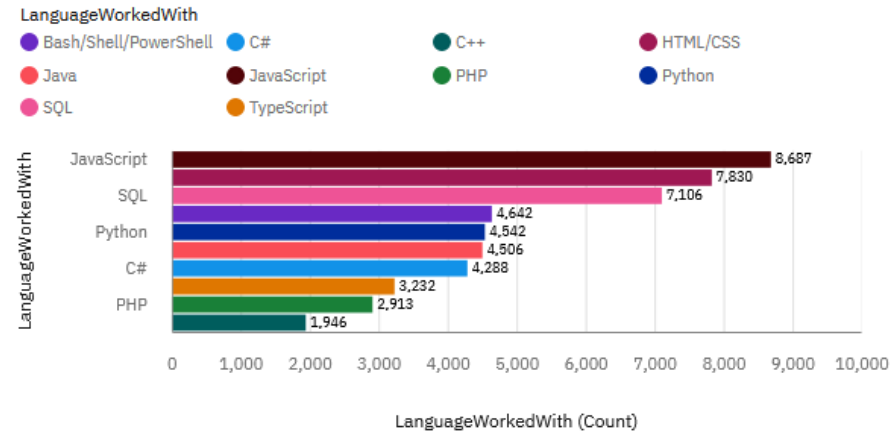
DASHBOARD



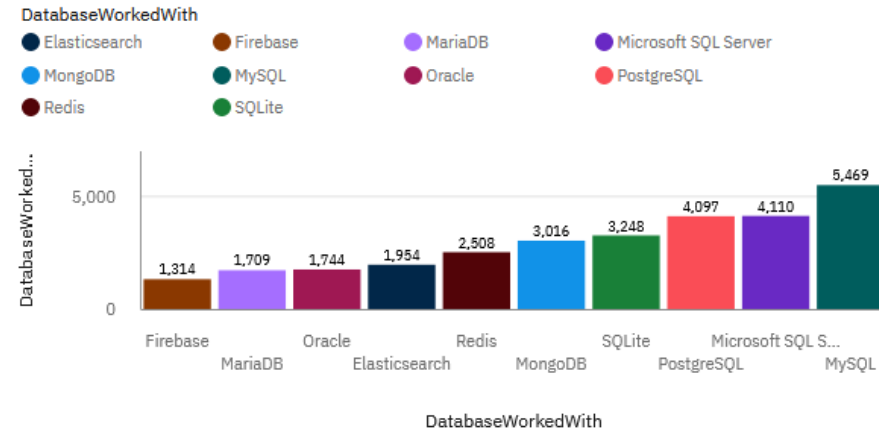
[Link to Dashboard](#)

DASHBOARD TAB 1

Top 10 Language Worked With



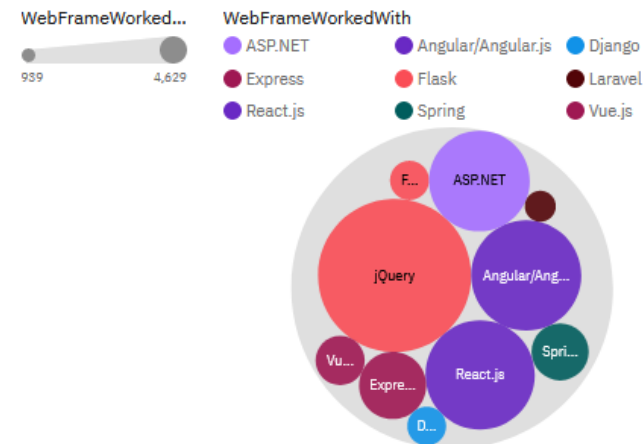
Top 10 DatabaseWorkedWith



PlatformWorkedWith

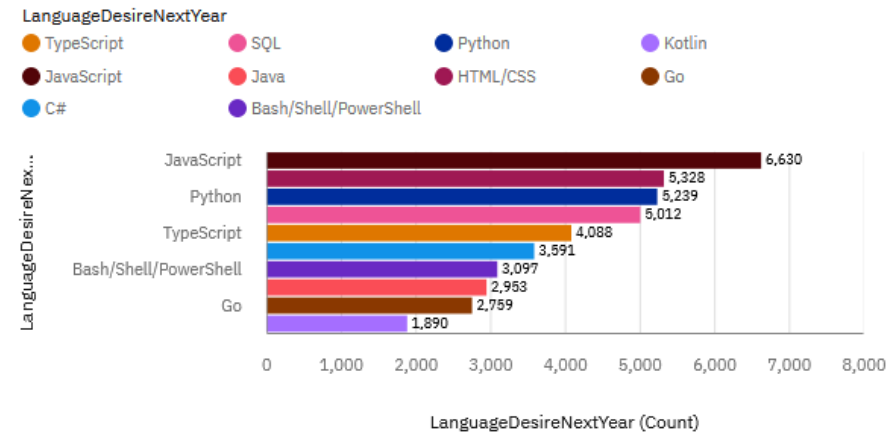


Top 10 WebFrameWorkedWith

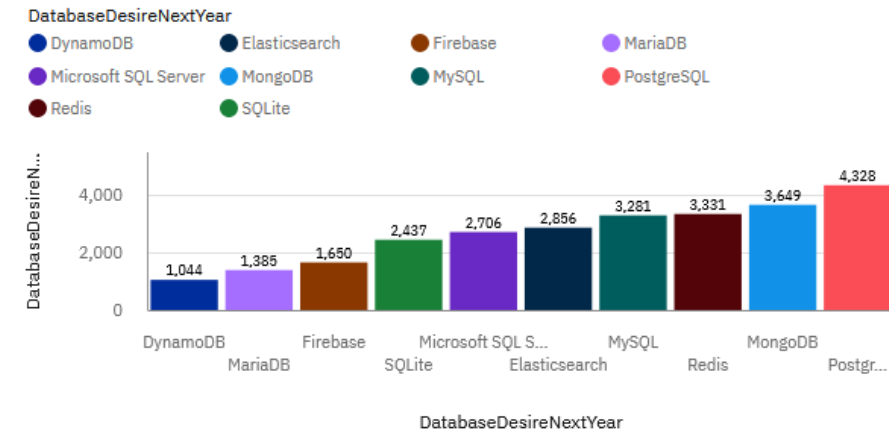


DASHBOARD TAB 2

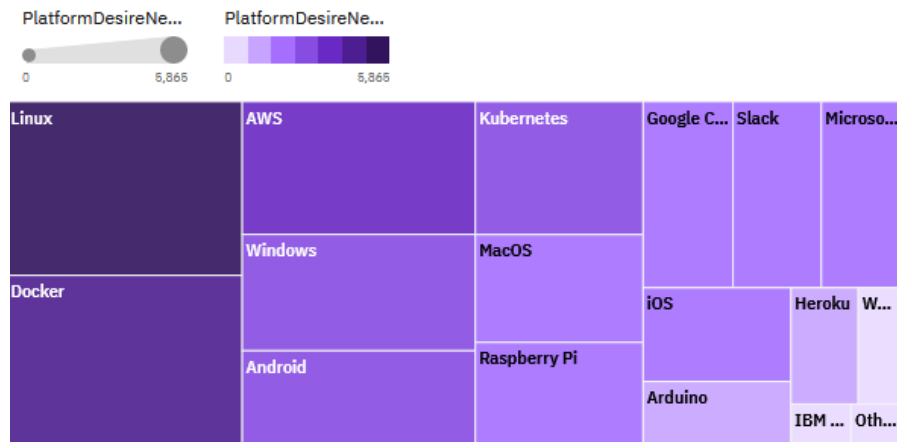
Top 10 LanguageDesireNextYear



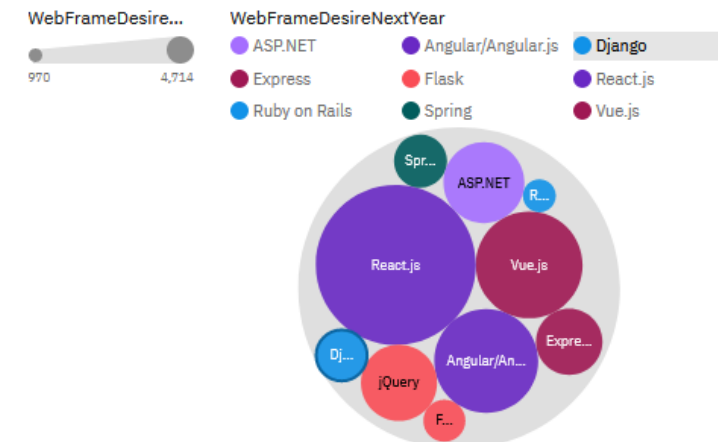
Top 10 DatabaseDesireNextYear



PlatformDesireNextYear



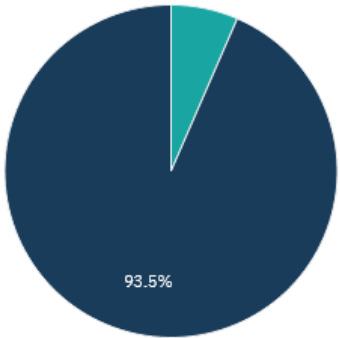
Top 10 WebFrameDesireNextYear



DASHBOARD TAB 3

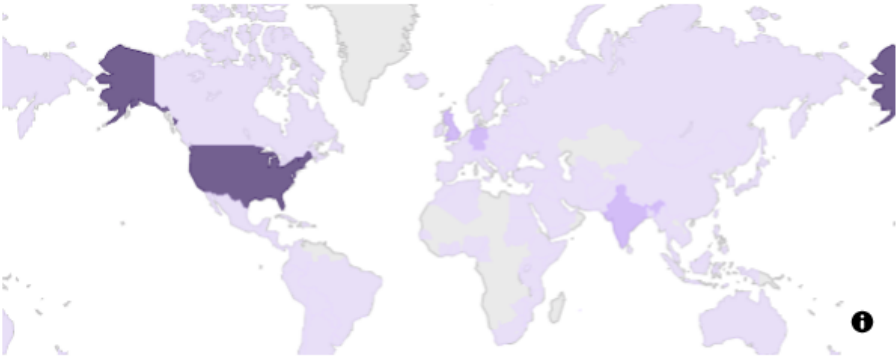
Respondent classified by Gender

Gender
● Woman ● Man

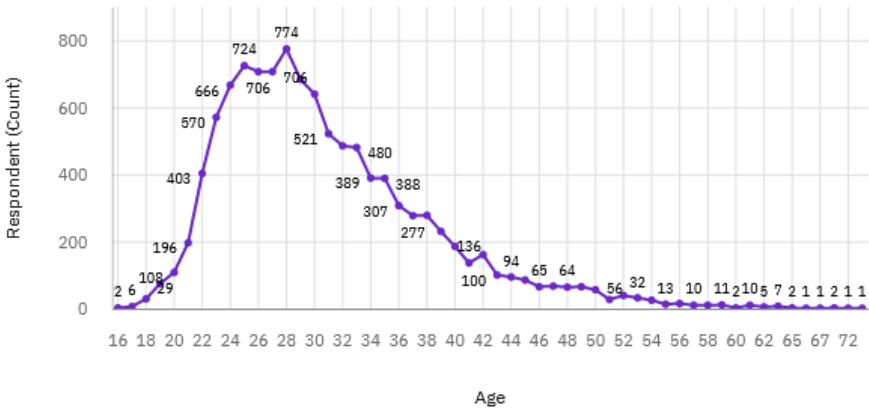


Respondent Count for Countries

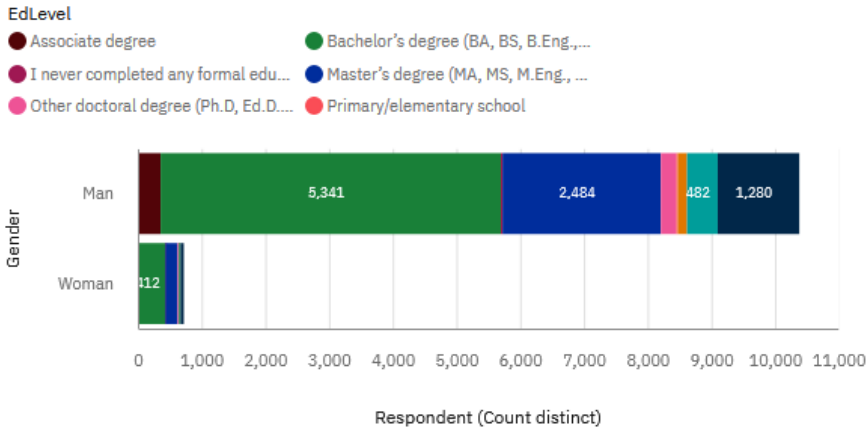
Country (Count)
1 3,068



Respondent Count by Age



Respondent Count by Gender, classified by Formal Education Level



DISCUSSION



- As we can see different technologies in IT-sphere replace each other adjusting under growing business demand
- There is a big gender difference in this industry, but this difference in employee amount is almost not treat to the salary

OVERALL FINDINGS & IMPLICATIONS

Findings

- The most popular programming languages are: JavaScript, HTML/CSS and SQL
- Open Source Platforms (Linux, Docker) gain popularity, because open code is useful and convenient for developing

Implications

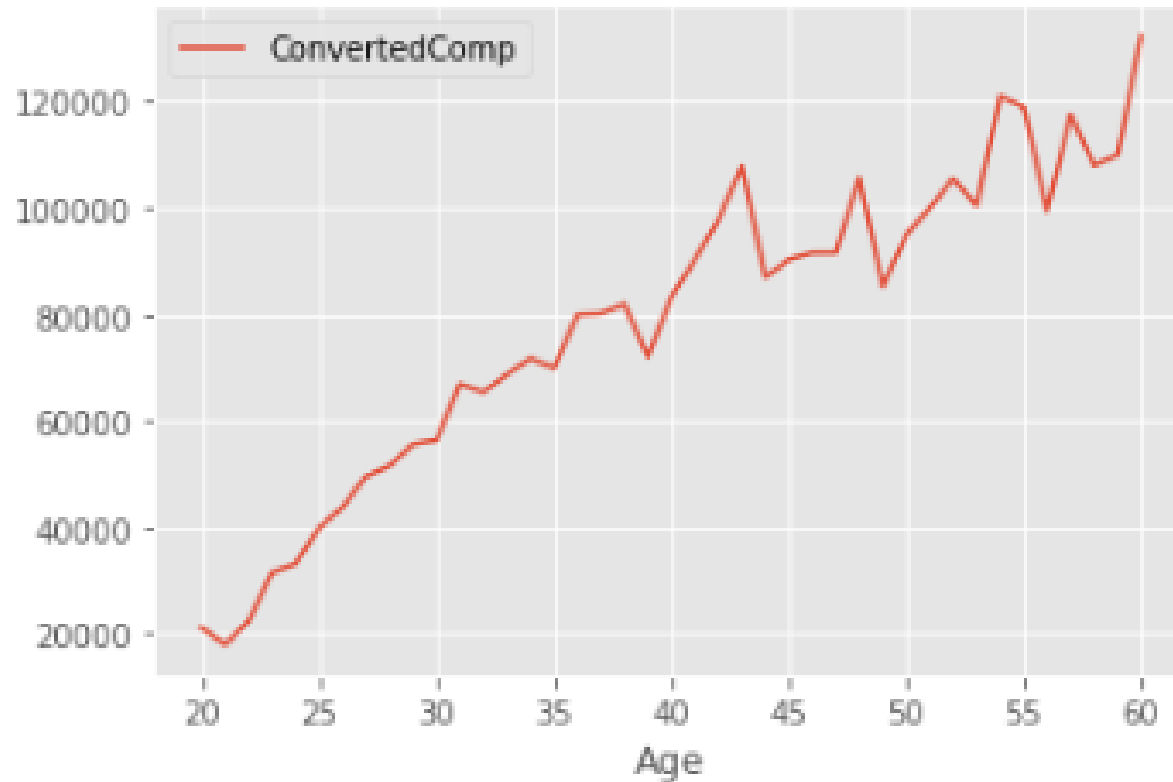
- With the increase in the amount of information, technologies that can process large amounts of data quickly and efficiently on different platforms are gaining popularity
- There is a big gender gap in the IT-sphere

CONCLUSION



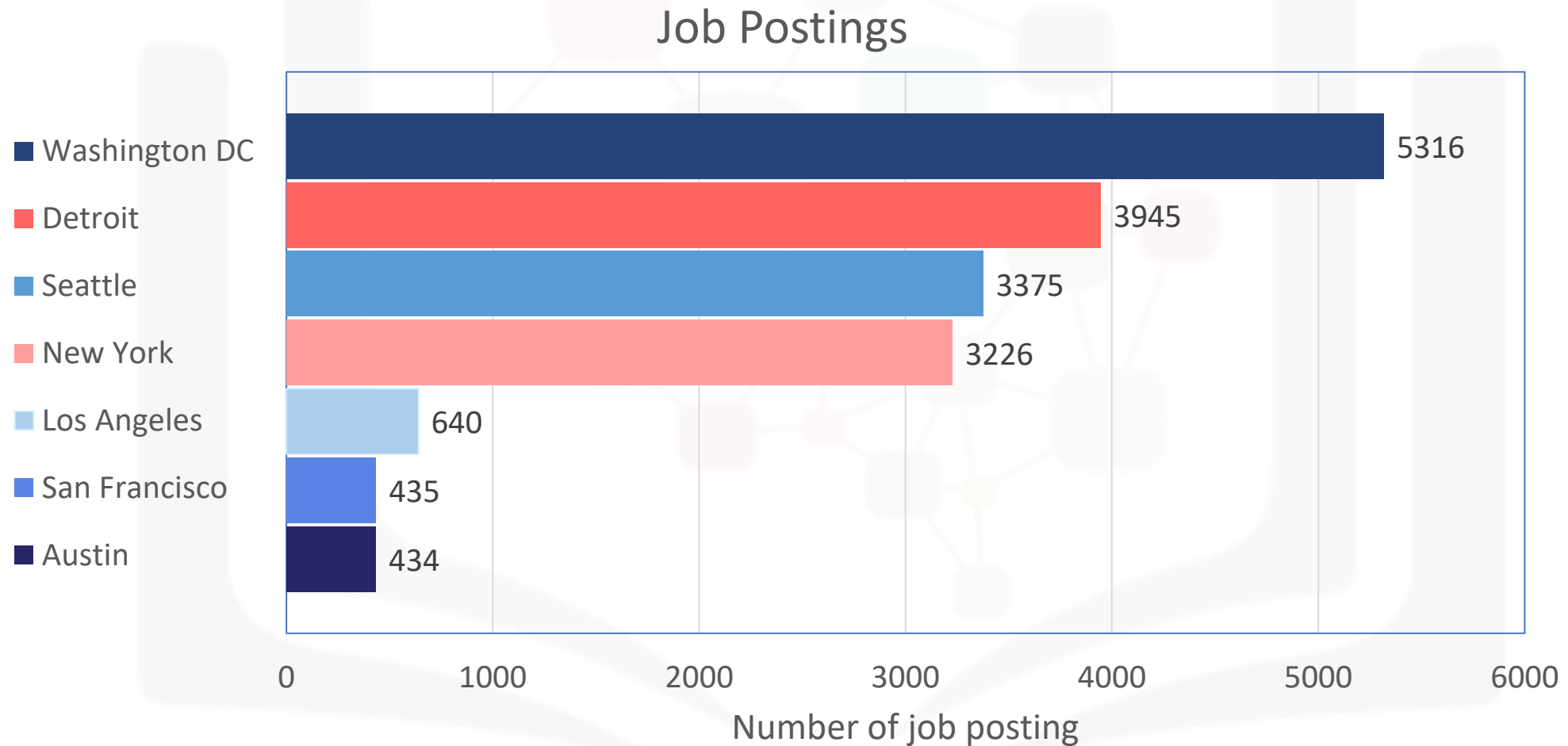
- Increasing of the data amount and demands to the software productivity leads to the popularity growth of the instruments that can satisfy all the above
- Wide spreading of the Internet technologies leads to growing of the interest to the programming languages that are used for web-developing (HTML/CSS, JavaScript, Python, SQL)

Demographic Analysis



- With growth of the age of the employee, their salary also is growing. It can be explained by experience and knowledge of the employee
- The annual salary median for man and woman don't have a big difference:
57 744\$ vs. 57 708\$

JOB POSTINGS



POPULAR LANGUAGES

