



Data Analyst Capstone Project

Arthur Thouvenin

06/09/2021

OUTLINE



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

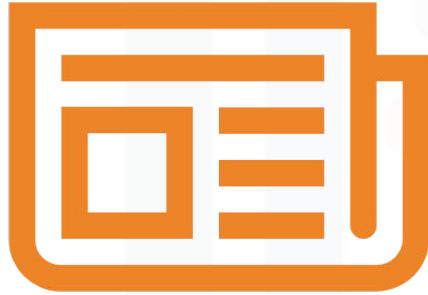


- During this IBM-Coursera course we used a lot of methodologies such as :
 - Data Collection with APIs, web scrapping, SQL connections
 - Data Wrangling with Pandas, Numpy, SQL
 - Data Visualization with Plotly, Seaborn and Dashboards
- Summary of all results
 - Thanks to all of this projects we were able to have a better understanding of the current technologies and future skills requirements.

INTRODUCTION



- The goal of this project was to use every side of data analysis to study the evolution of technologies in the field of data analysis.
- First collect the top programming skills, but also analyzing the data and identify insights and trends.

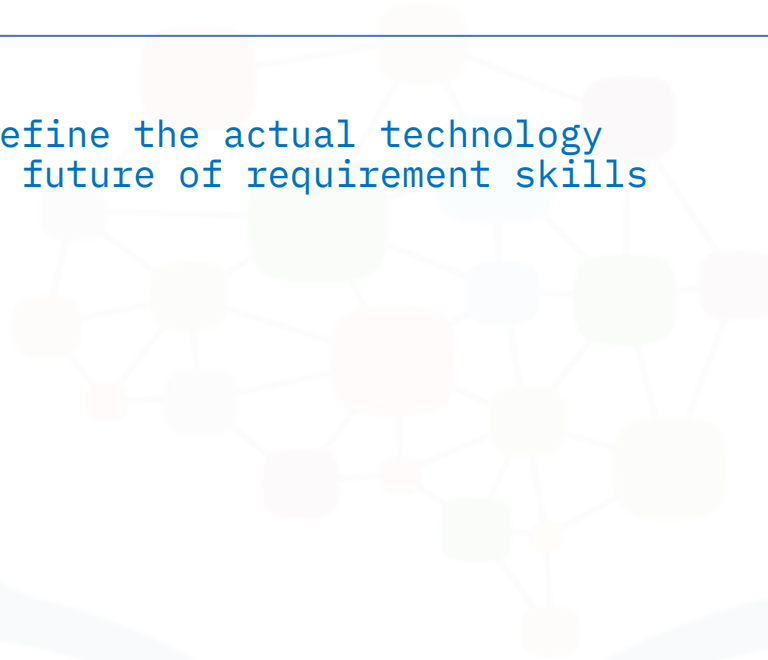


Executive Summary

- Data collection methodology:
 - *We used several methods from web scrapping, to SQL queries and APIs*
- Perform data wrangling
 - *The data were most of the time standaradize and processed with Numpy and Pandas*
- Perform exploratory data analysis (EDA) using visualization and SQL, Excel
- Perform interactive visual analytics using dashboards and charts

RESULTS

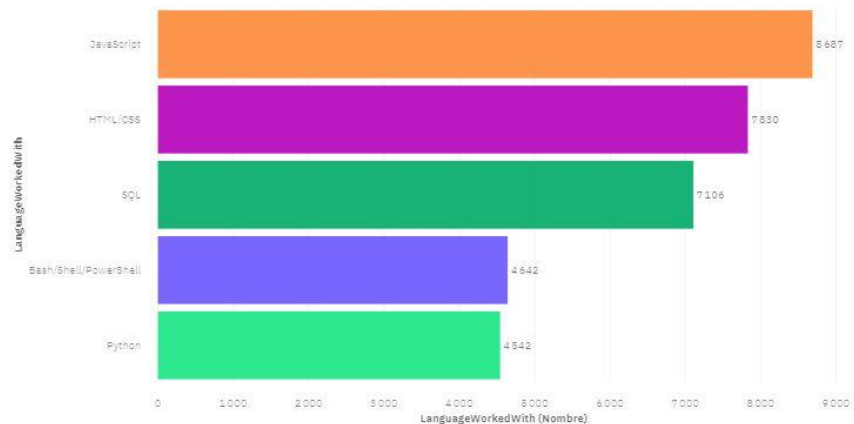
We were able to define the actual technology situation and the future of requirement skills



PROGRAMMING LANGUAGE TRENDS

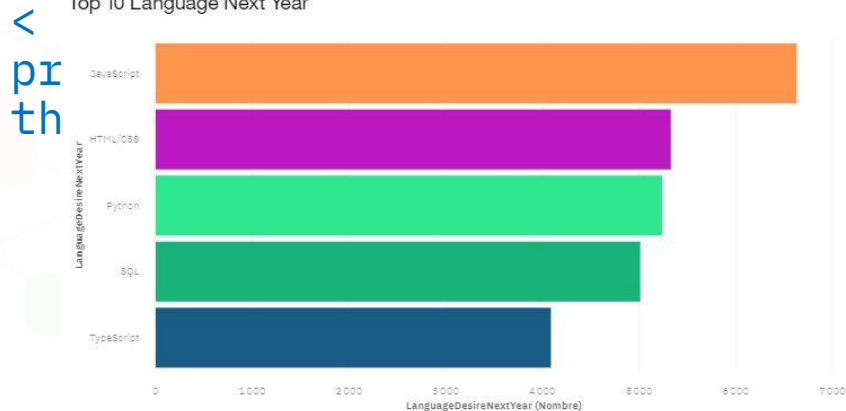
Current Year

Top 10 Language Worked With



Next Year

Top 10 Language Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

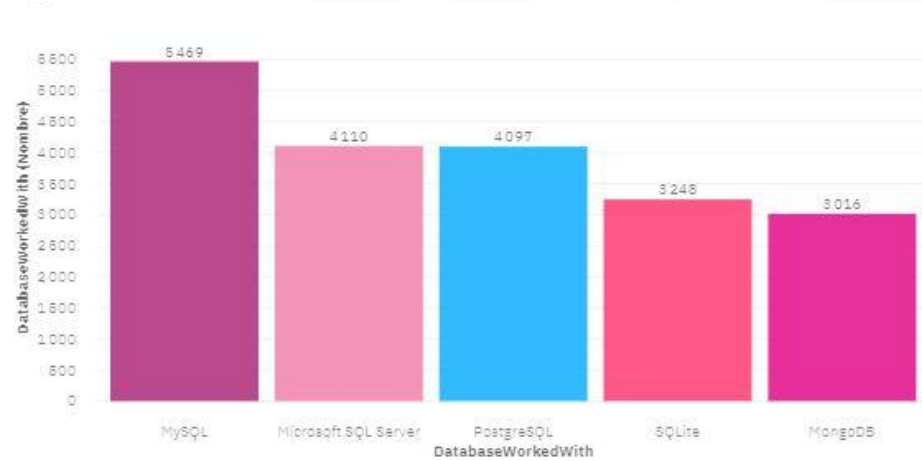
- First language Javascript this year and Javascript next year
- Second language HTML/CSS this year and HTML/CSS next year
- Third language SQL this year and Python next year

Implications

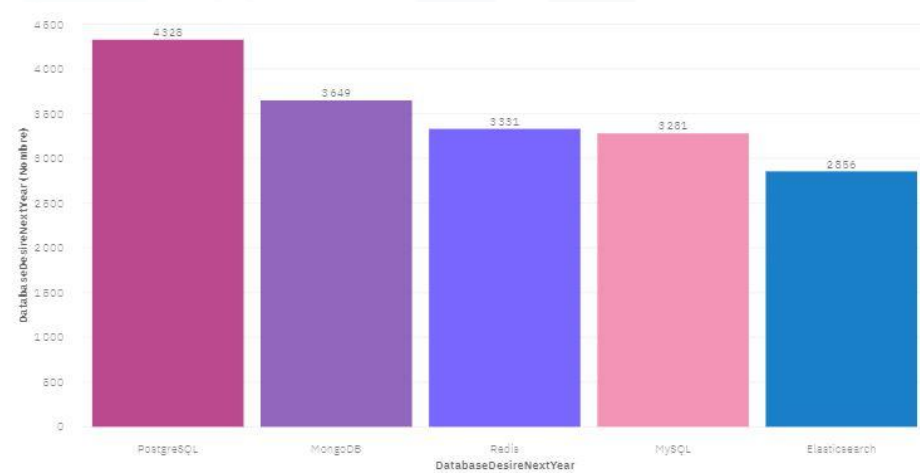
- Javascript and HTML/CSS are key languages to have in programming skills
- For the following year it would be important to select candidates who master Python

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- We can note that the Microsoft SQL server completely disappear from the top 5 database
- MongoDB, Redis, and ElasticSearch are promising databases

Implications

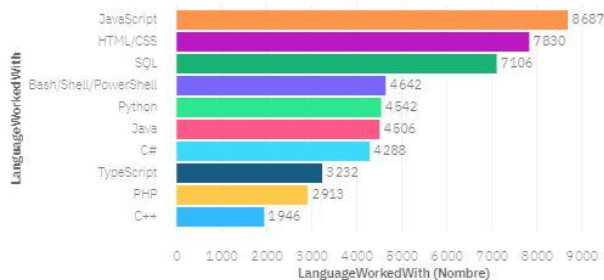
- Hiring developer that master MongoDB, Redis and ElasticSearch would be a good idea for the next year
- Training current developer on those databases would help companies stay on track with the best technologies



- Current Year : [LINK](#)
- Next Year : [LINK](#)
- Demographics : [LINK](#)

DASHBOARD TAB 1

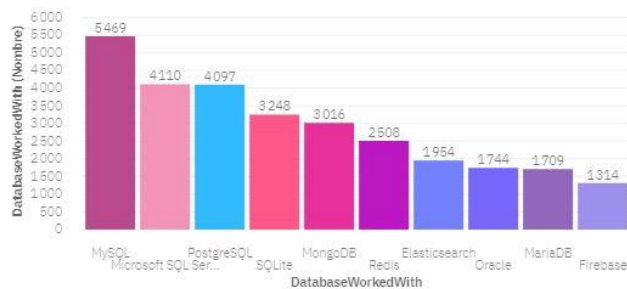
Top 10 Language Worked With



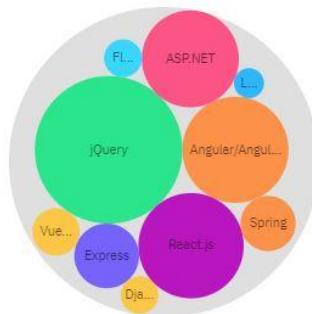
Platform Worked With



Top 10 Database worked with

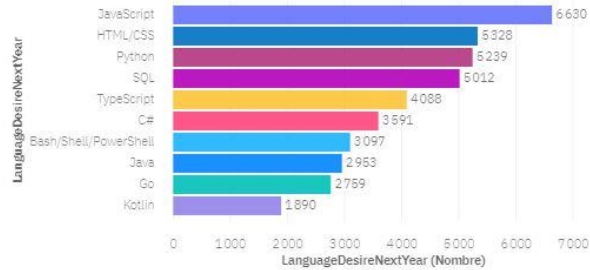


Top 10 Web Frameworks Worked With



DASHBOARD TAB 2

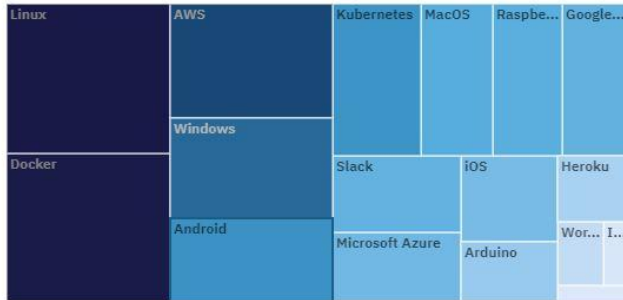
Top 10 LanguageDesireNextYear



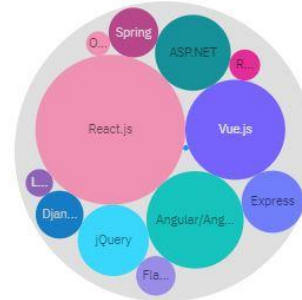
Top 10 DatabaseDesireNextYear



PlatformDesireNextYear



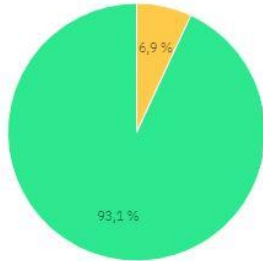
Top 10 WebFrameDesireNextYear



DASHBOARD TAB 3

Respondent Gender

Gender
● Woman ● Man



Respondent Age



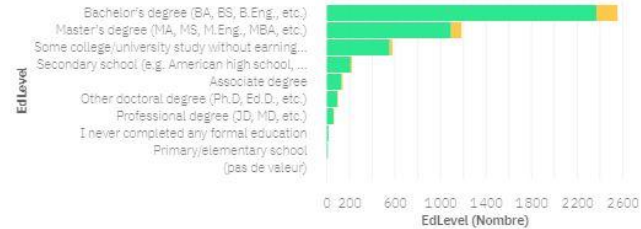
Respondent Location

Country (Nombre)



Formal Education by Gender

Gender
● Man ● Woman



DISCUSSION

Here as we can see Woman are still badly represented in computing fields

However some points can be discussed such as the fact that only 5000 rows of the demographic dataset has been used for this dashboard

OVERALL FINDINGS & IMPLICATIONS

Findings

- Woman are badly represented in computing fields
- Repondent of those datasets are mainly based in one country which may not represent an interesting overview of developer throughout the world

Implications

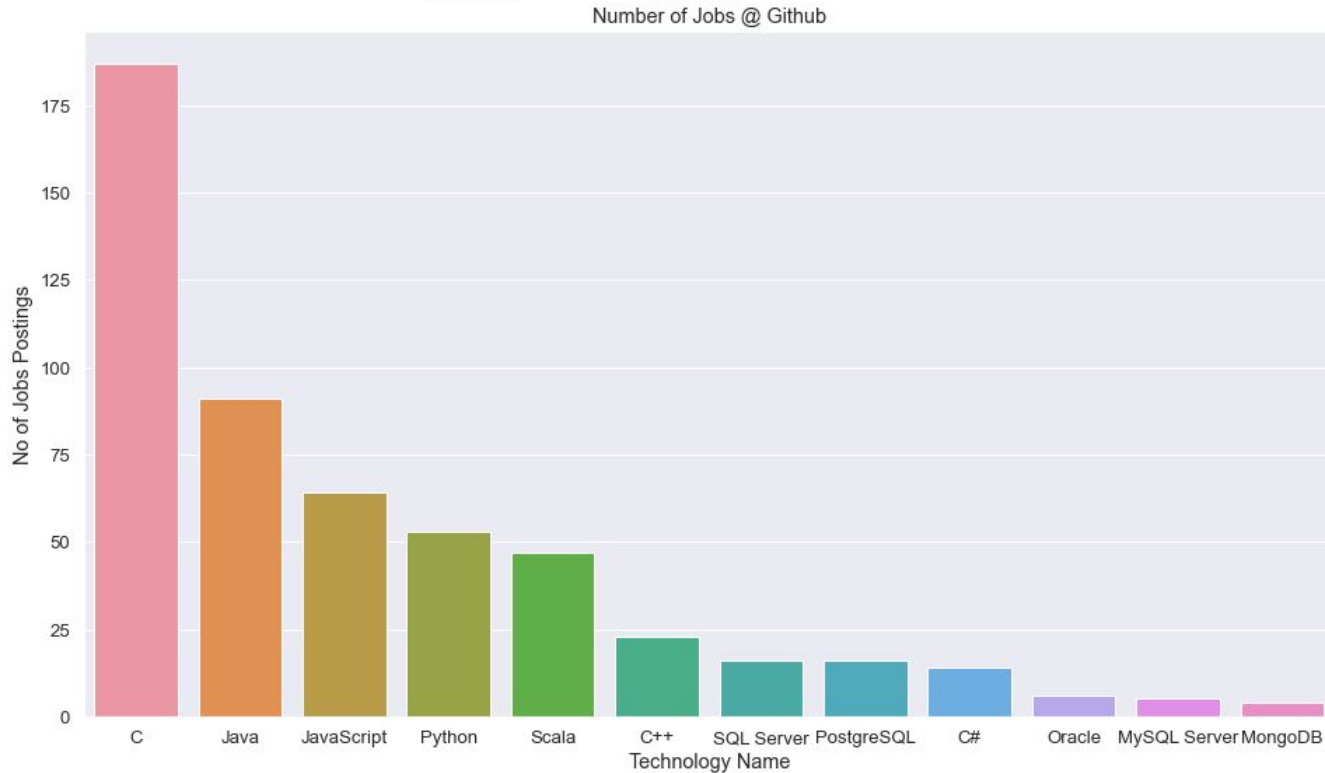
- Learn computing skills to women can be a huge step to bring more women to be developer themselves
- Asking more developer in other countries could also be interesting to compare the trend between countries

CONCLUSION



- Developer is a work where technologies moves fast and you need to keep track of those evolution
- Women are still badly represented in this field
- It would be interesting to ask more developer of other countries to compare them with US
- Finally those dashboards and this analysis helps to track the current technologies used by developer and technologies of the future (next year)

GITHUB JOB POSTINGS



POPULAR LANGUAGES

