Capstone Project

Farouk Jaziri

10 June 2024

OUTLINE

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

EXECUTIVE SUMMARY

- People are driven to learn data science (python, SQL) and web development tools (Javascript, HTML/CSS)
- People lean towards easy to learn and efficient databases
- Most of the respondents are Men, these statistics cannot be relevant if we don’t take the same portion of demographics in our calculation, this will affect our decision.
- The data is missing Respondent_id, this will help us to differentiate between the men and women stats.

INTRODUCTION

- In this presentation, we used 2 datasets:
  - Data Technologies Normalized
  - Data Demographics
- We want to see the change between this year and the next one in terms of preferred technologies to learn
- From this we will analyze and understand the data we will use

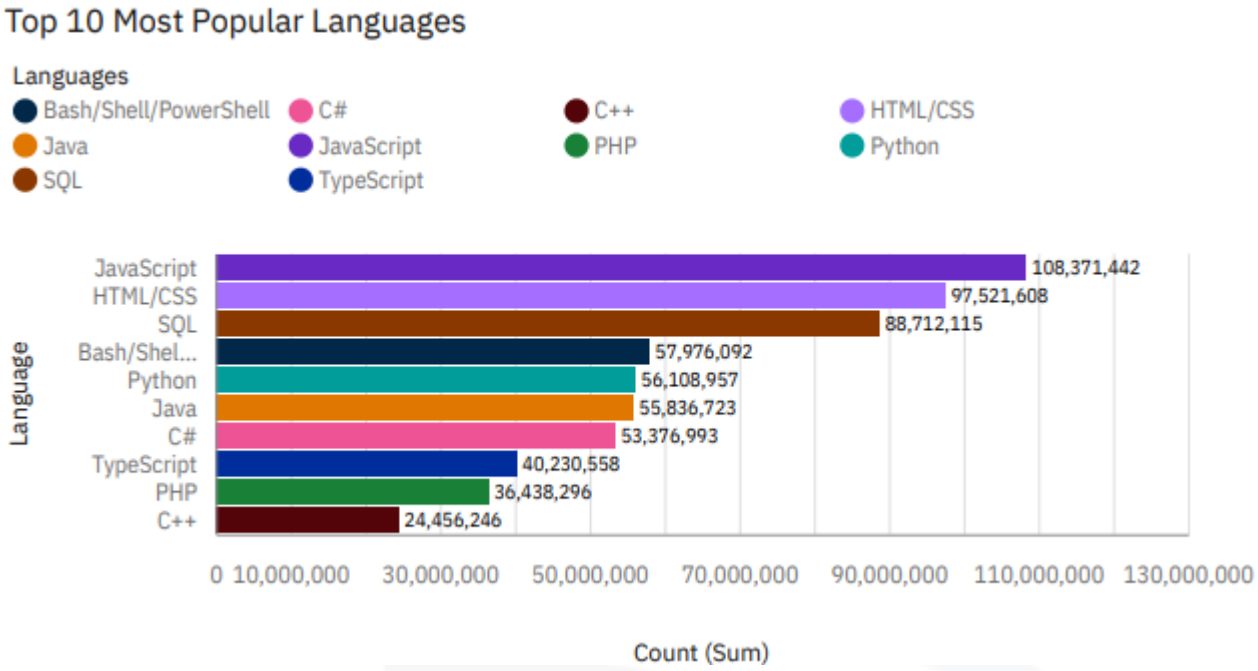
METHODOLOGY

- Analyze the technologies used this year
- Analyze the desired technologies to learn for the next year
- Understand the demographics of the statistics

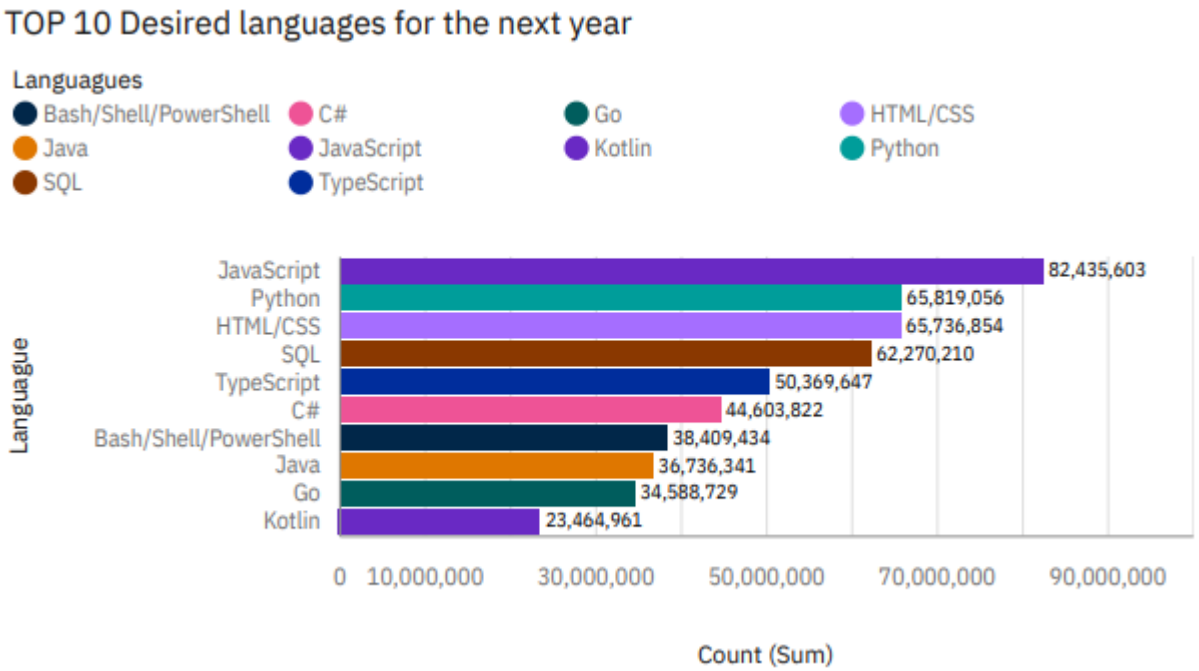
RESULTS

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

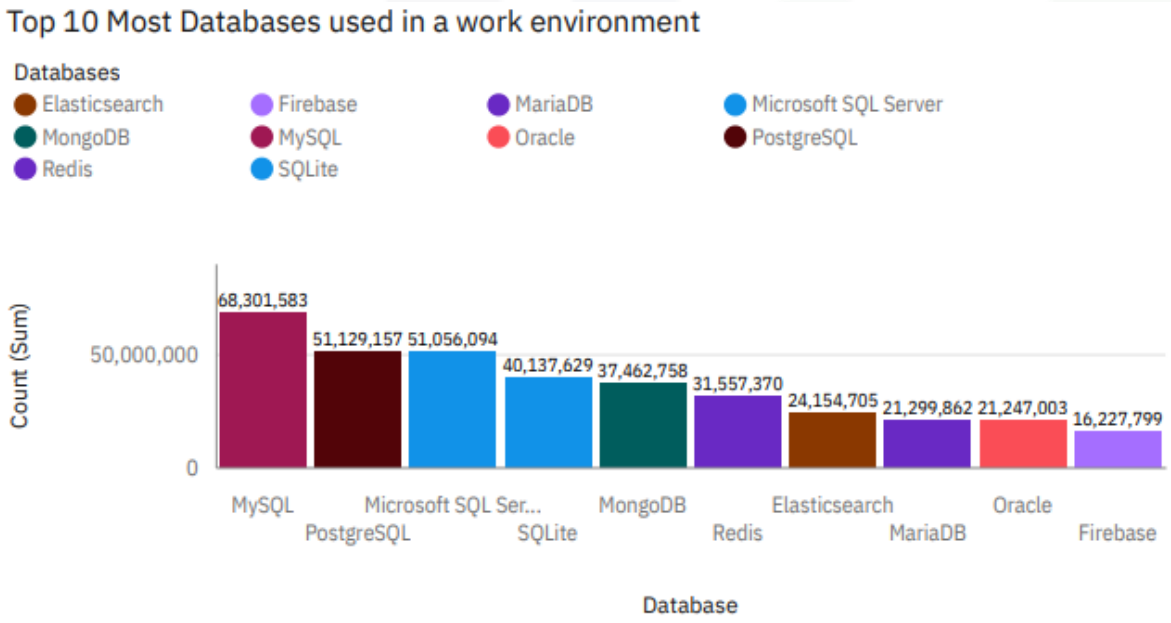
- Javascript remains top 1 even with a loss of 24 percentage points
- SQL gains more than 50 % increase of change

Implications

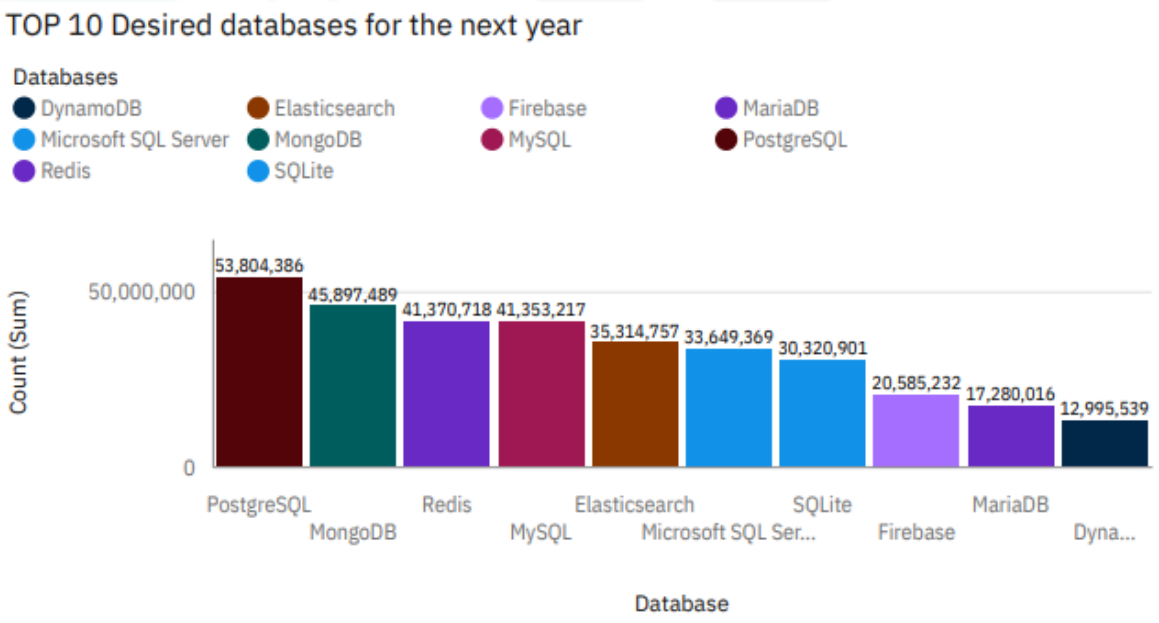
- People are driven to learn data science (python, SQL) and web development tools (Javascript, HTML/CSS)

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- PostgreSQL is the TOP 1 Database to learn even though it is already known for most people this year

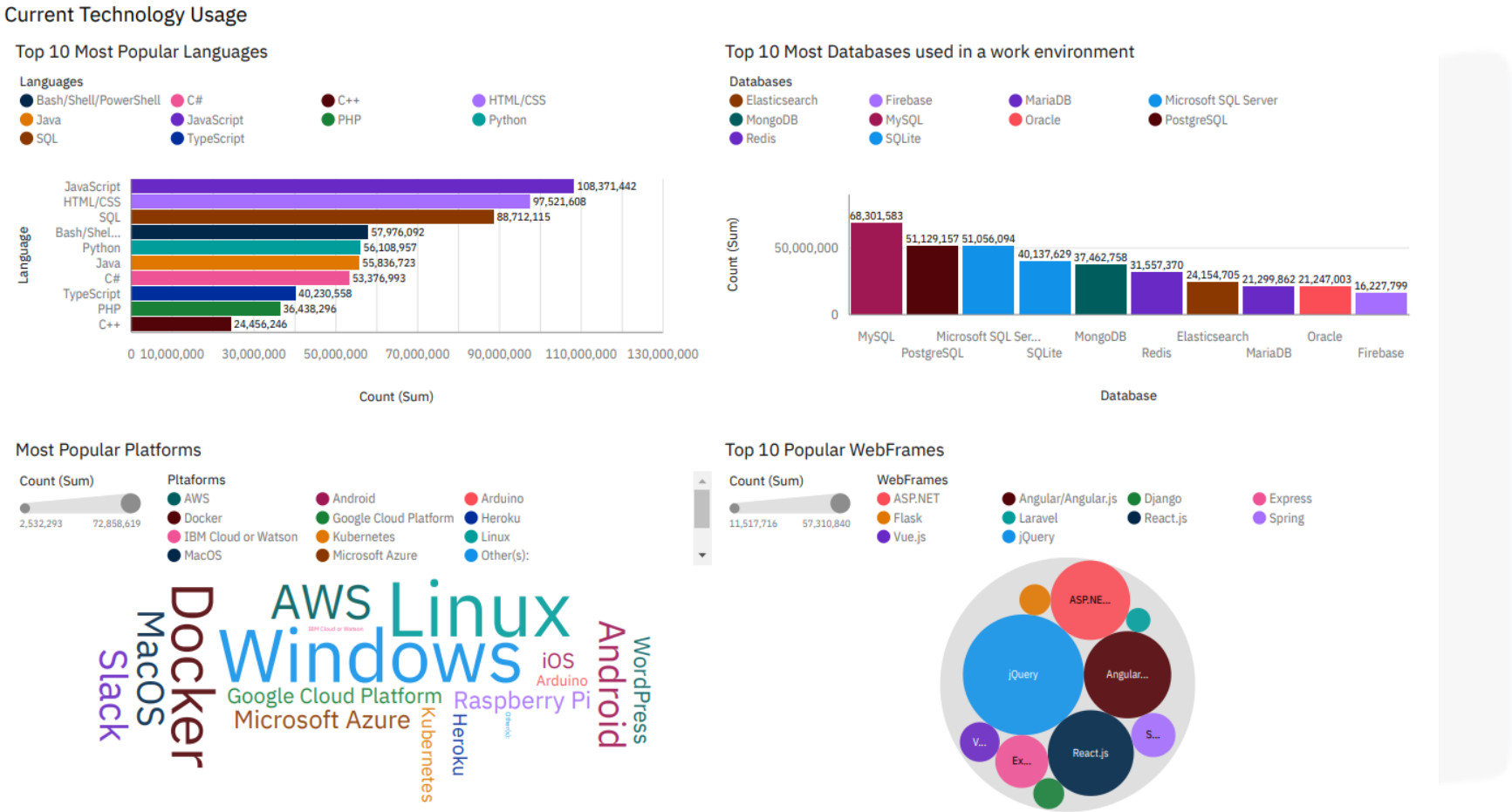
Implications

- People lean towards easy to learn and efficient databases

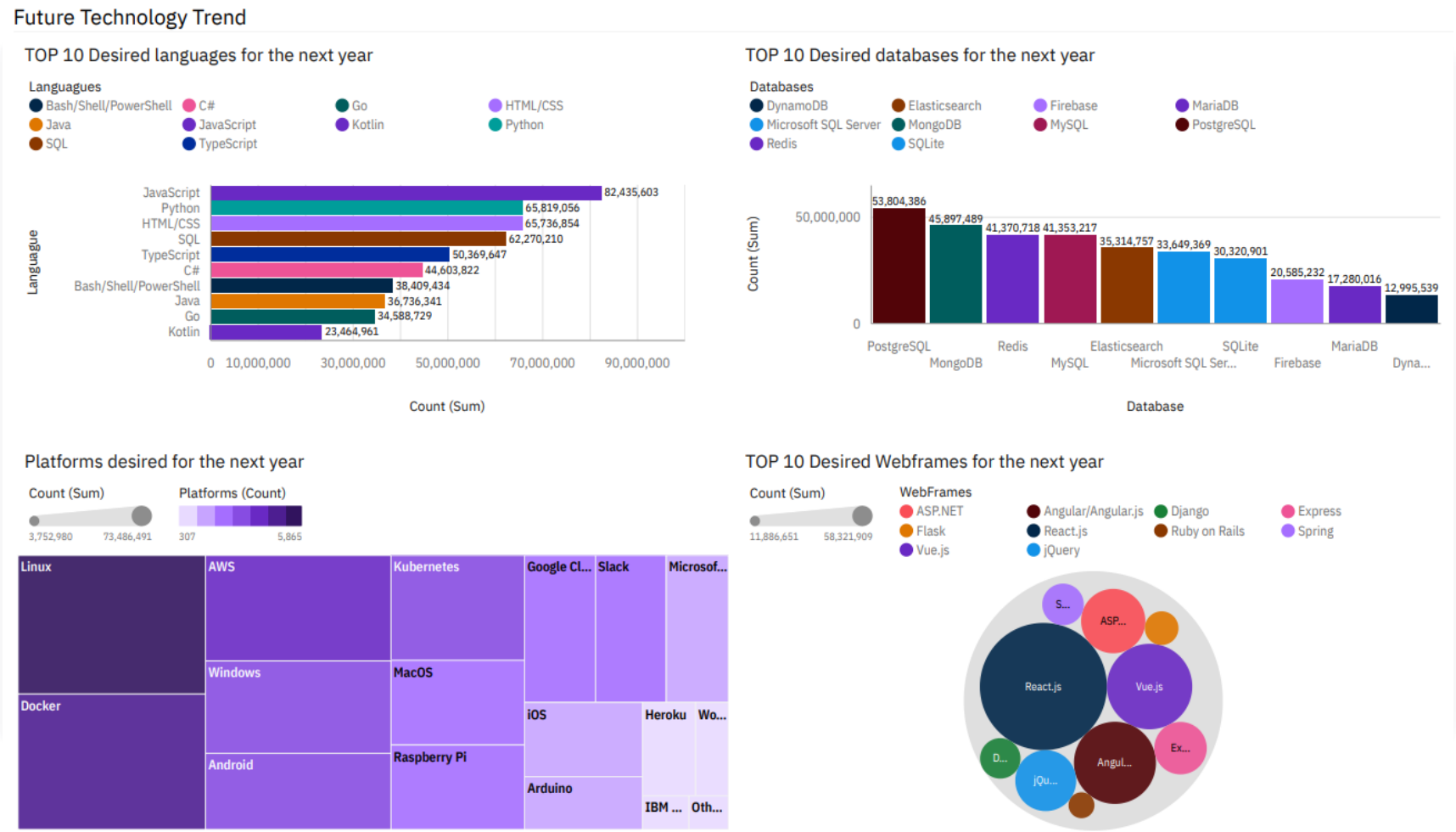
DASHBOARD

https://github.com/djazirifarouk/Data-Analytics-and-Visualization-Capstone-Project.git

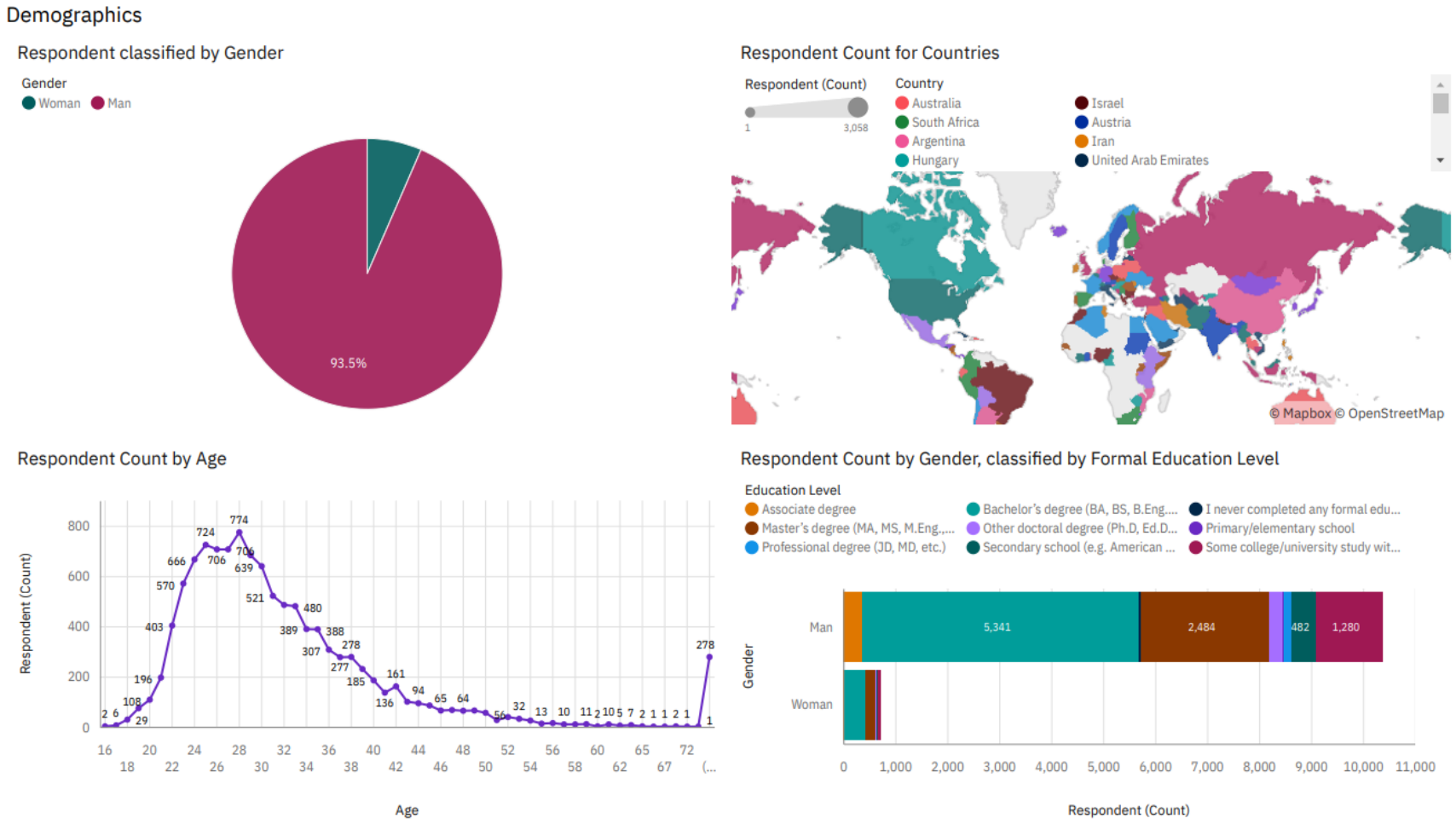
Current Technology Usage



Future Technology Trend



Demographics



DISCUSSION

- Most of the respondents are Men, these statistics cannot be relevant if we don’t take the same portion of demographics in our calculation, this will affect our decision.
- The data is missing Respondent_id, this will help us to differentiate between the men and women stats.

OVERALL FINDINGS & IMPLICATIONS

Findings

- Between 25-44 years old has the densest respondent count
- Cloud technologies are becoming more interested to learn for the next year

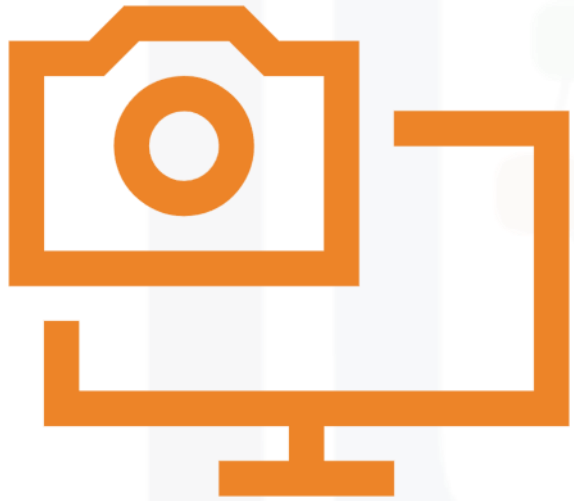
Implications

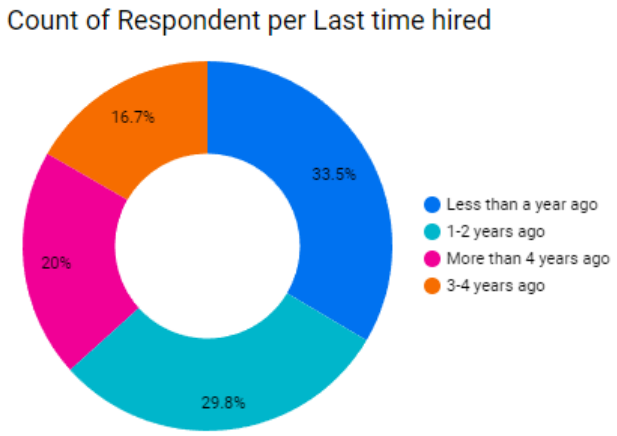
- People are towards 2 directions:
  - Data science tools
  - Web development tools
- People aged below 18 are not interested in this domain

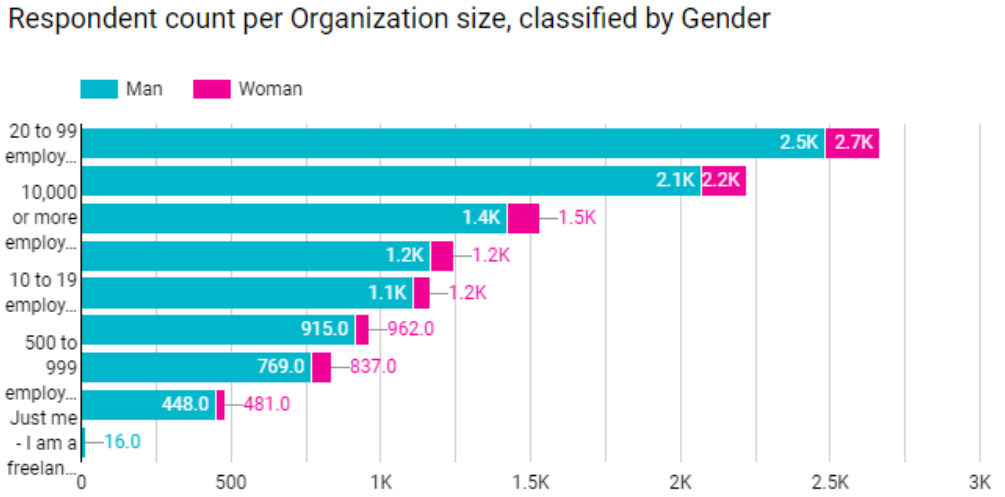
CONCLUSION

- We should differentiate between gender stats to make sure we have clear decision
- Provide knowledge for the desired technologies of the next year
- Encourage people below 18 to start learning about more technologies

APPENDIX




Count of Respondent per Last time hired
16.7%
33.5%
20%
29.8%
Less than a year ago
1-2 years ago
More than 4 years ago
3-4 years ago


Respondent count per Organization size, classified by Gender
Man
Woman
20 to 99 employ...
10,000 or more employ...
10 to 19 employ...
500 to 999 employ...
Just me - I am a freelan...
2.5K
2.7K
2.1K
2.2K
1.4K
1.5K
1.2K
1.2K
1.1K
1.2K
915.0
962.0
769.0
837.0
448.0
481.0
16.0
0
500
1K
1.5K
2K
2.5K
3K

JOB POSTINGS

In Module 1 you have collected the job posting data using Job API in a file named “job-postings.xlsx”. Present that data using a bar chart here. Order the bar chart in the descending order of the number of job postings.

POPULAR LANGUAGES

In Module 1 you have collected the job postings data using web scraping in a file named “popular-languages.csv”. Present that data using a bar chart here. Order the bar chart in the descending order of salary.