

# Data Analytics Capstone Project

GIRIJA O K

24-11-2024



IBM Developer  
SKILLS NETWORK



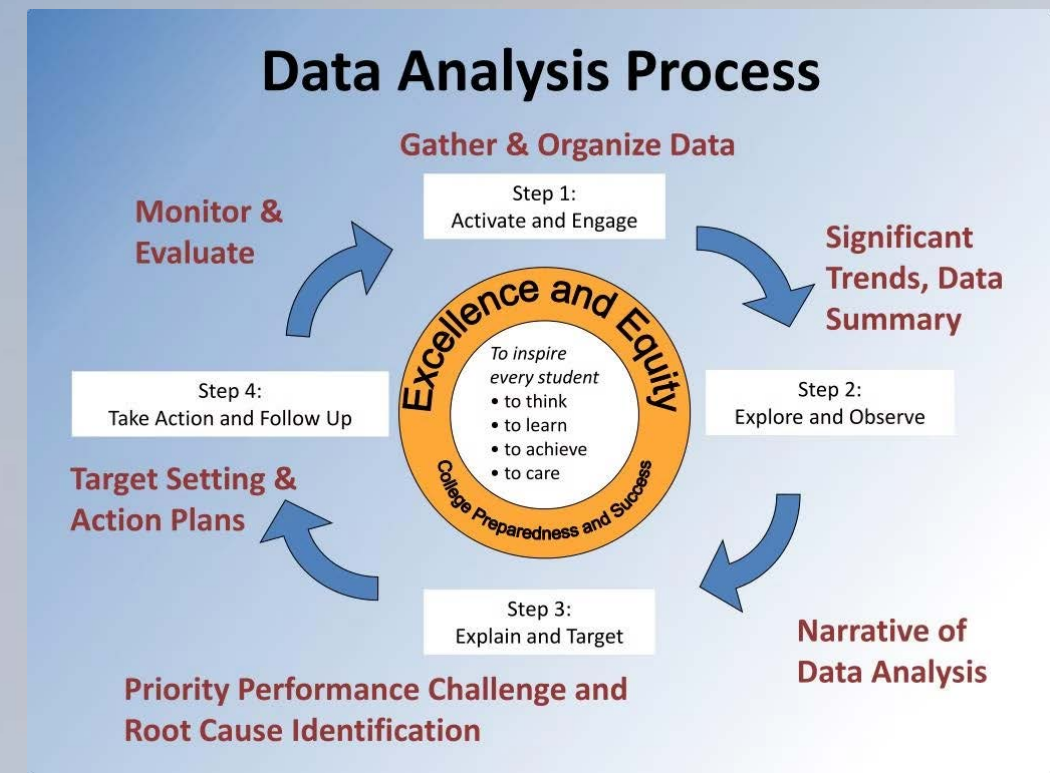
# Outline

- Executive Summary
- Introduction
- Methodology
- Results
  - a. Visualization - Charts
  - b. Dashboard
- Conclusion
- Appendix



# Executive Summary

- Collecting demand of jobs open from various technology and locations.
- Analyzing the data and identify insights and trends.
  - a. The top programming languages
  - b. The top databases skills
  - c. The popular IDEs
- Bringing the information together with dashboard showing current skills, future trend skills, and Demographics





# Introduction

This presentation report uses data analytics to highlight current and projected trends in the need for skills related to programming languages, databases, platforms and web frames.

The following inquiries were investigated using the data:

1. Which programming languages are most in demand today?
2. What are the most in-demand database skills?
3. What popular IDEs or Web frames are there?

The target audience for this research are IT professionals, HR managers, and anybody else with an interest in the IT sector who wants to learn about the top on-demand IT skills in their respective sectors that will also still be relevant in the future.



# Methodology



Collect survey data and explore its contents

- Web Scraping
- APIs
- Request the library

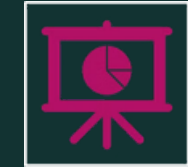


## Data Wrangling

Data wrangling, also known as data munging, refers to the process of transforming and preparing raw data into a format suitable for analysis. This involves cleaning, structuring, and enriching the data to ensure it is consistent, complete, and ready for use in various applications such as data analysis, machine learning, or reporting.



Python was used for exploratory data analysis. To assess the distribution of data, the presence of outliers, and the correlation between various columns in the dataset, an exploratory data analysis was carried out.



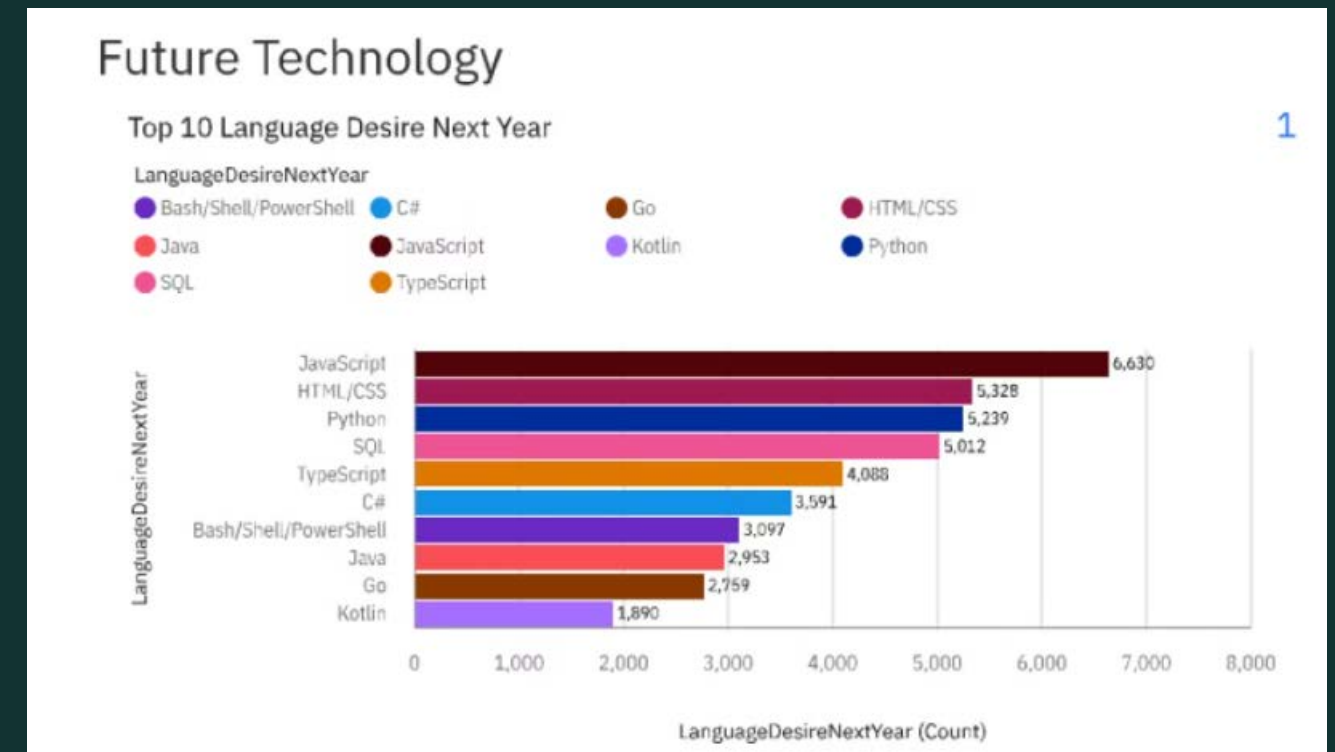
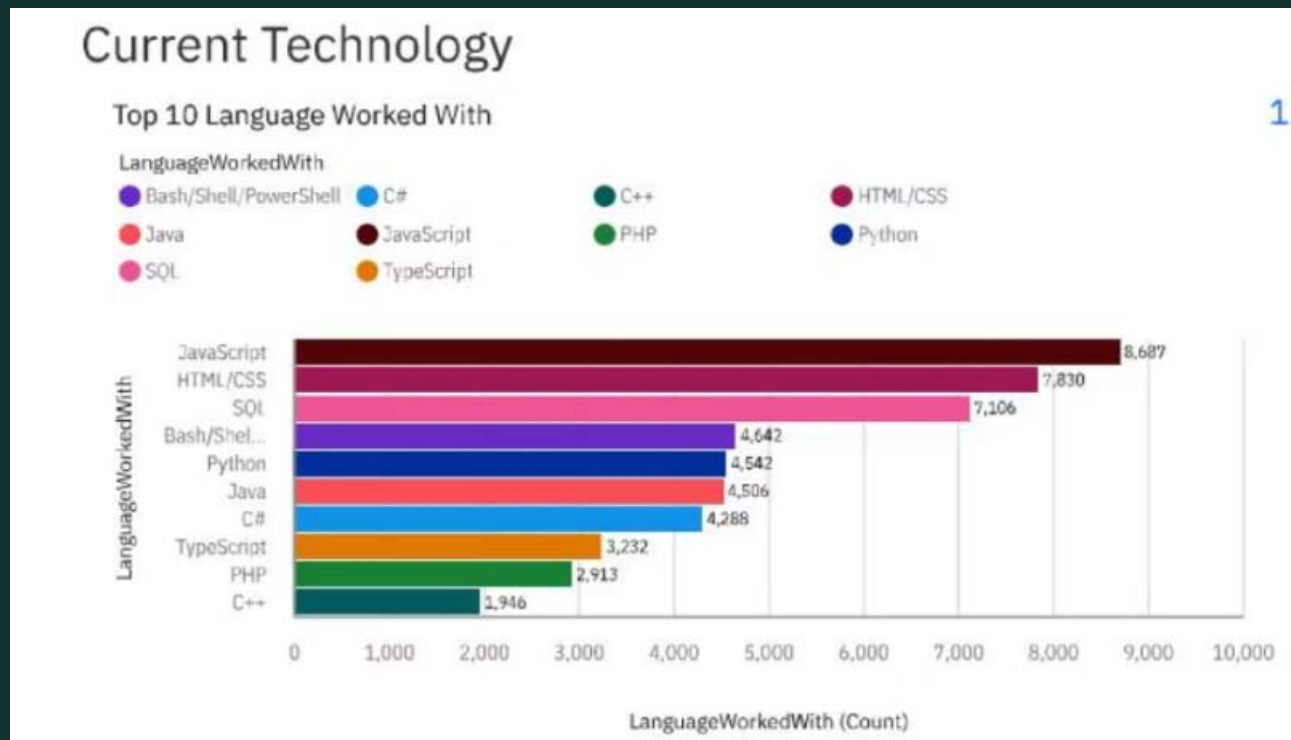
Charts, graphs, and dashboards were created using Python and Cognos analytics to visualize the data. All the python analyses were carried out on Jupyter notebook through visual studio.

# RESULTS

# PROGRAMMING LANGUAGE TRENDS

Current year

Next year



# PROGRAMMING LANGUAGE TRENDS – FINDINGS & IMPLICATIONS

## Findings:

- Javascript, HTML/CSS, SQL, Shell languages and Python are the most used languages currently.
- Javascript, HTML/CSS, Python, SQL, and Typescript will be the most used languages next year and future years.
- Python will have more demand than SQL next year.

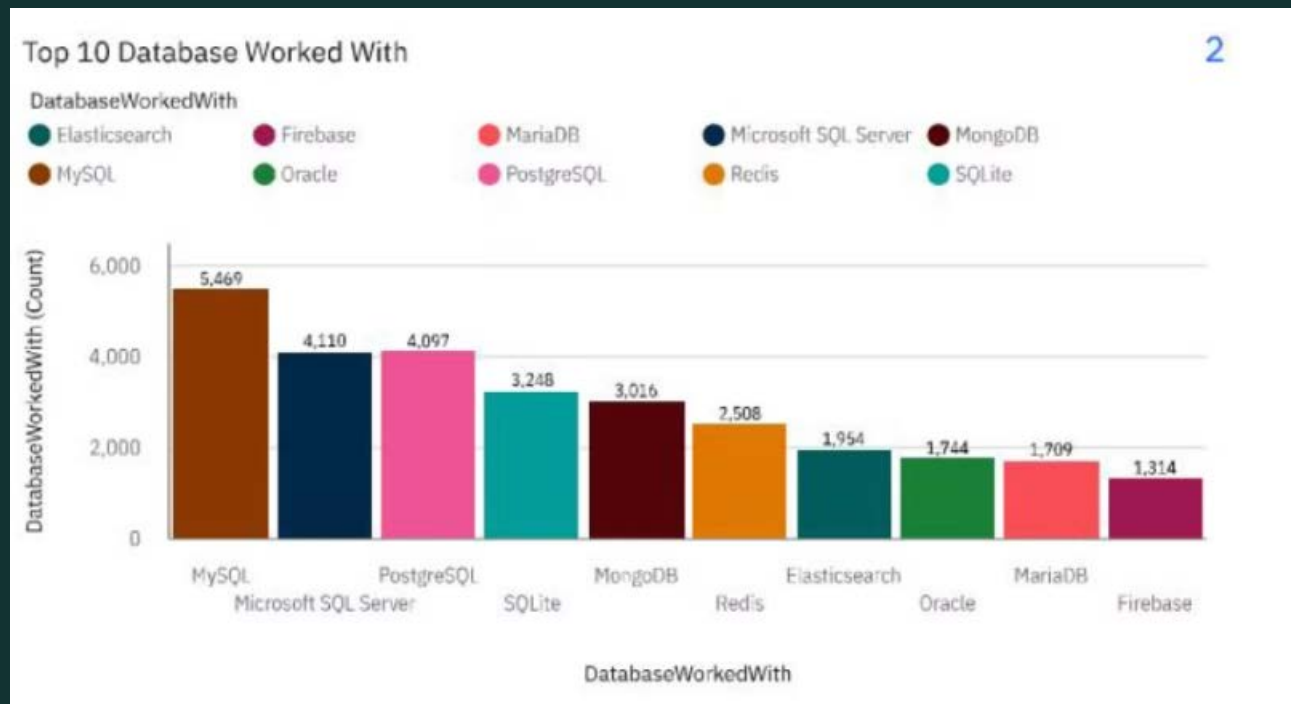
## Implications:

- Javascript and HTML are used for web development which means that web development as a tech skill has the highest demand, especially as Typescript is getting viral.
- Python is gaining more and more traction due to the increase in demand for AI and ML skills.
- SQL is still the most relevant language for data professionals. It is important for aspiring data analysts, scientists, business analysts etc to have SQL skills.

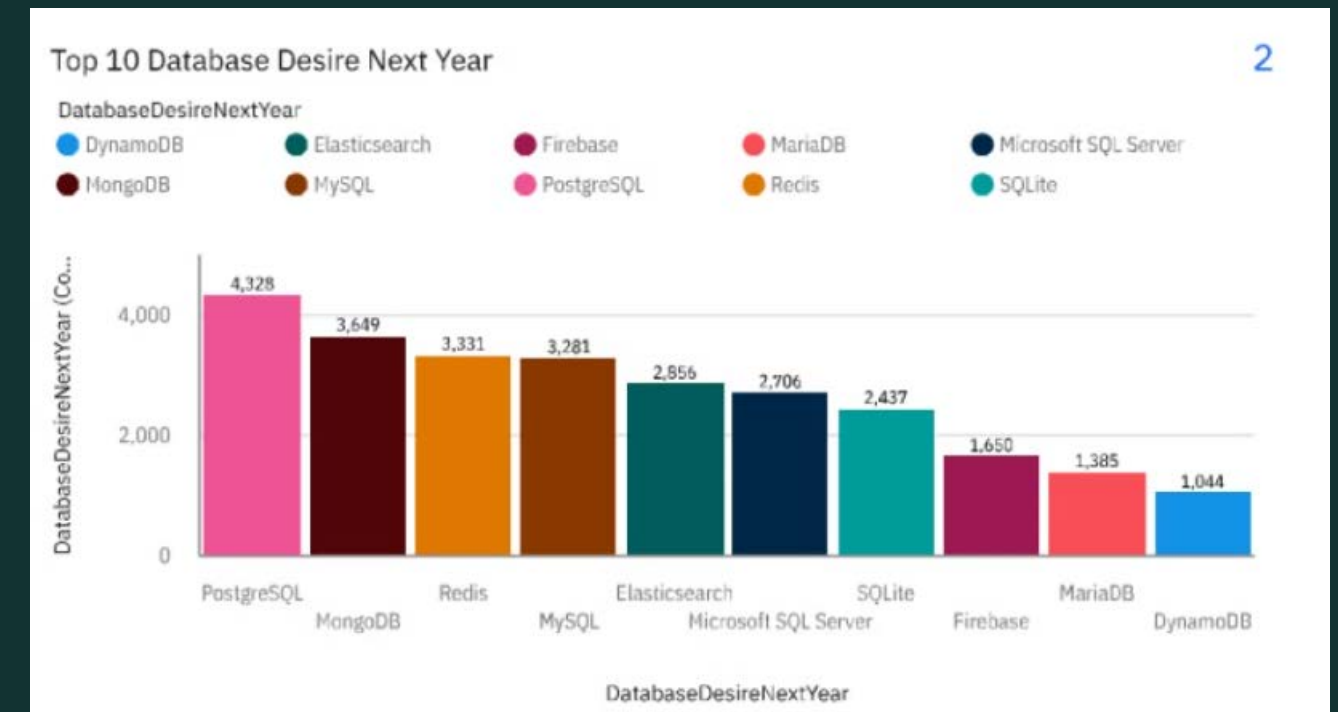


# DATABASE TRENDS

Current Year:



Next Year:



# DATABASE TRENDS – FINDINGS & IMPLICATIONS

## Findings:

- MySQL, Microsoft SQL server, Postgre SQL, SQLite and MongoDB are the top 5 most used databases at the moment.
- However, Postgre SQL, MongoDB, Redis, MySQL and Elasticsearch are projected to become more popular in the future.
- Redis and Elasticsearch are relatively new tools and are set to gain more traction in the IT space.

## Implications:

- SQL is still a top tool to watch out for in data specialists.
- Companies still prefer Open source databases.
- Oracle SQL was not among the top 5. It is losing relevance as time passes.

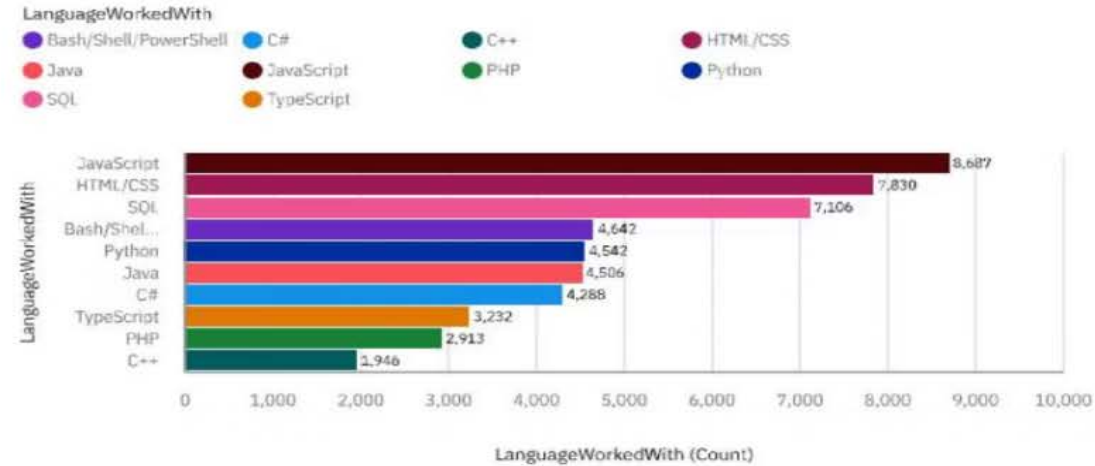
# DASH BOARDS



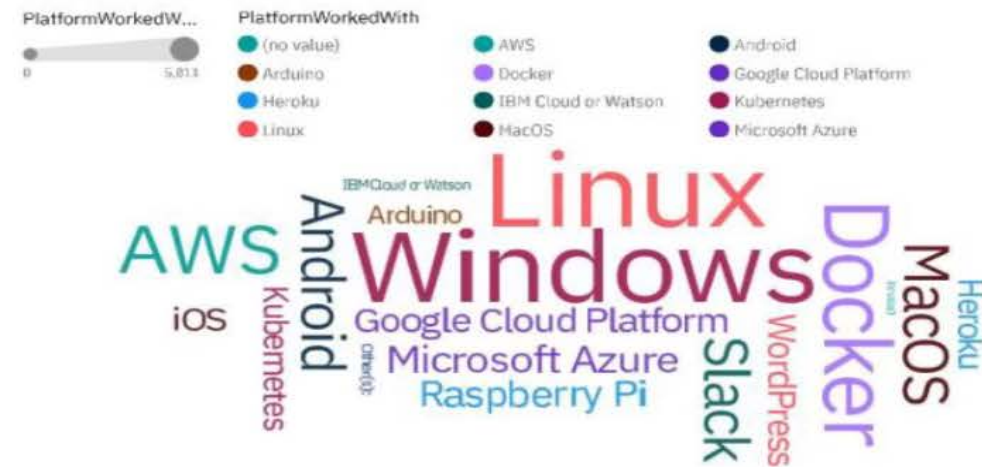
# DashBoard 1- Current Technology

## Current Technology

### Top 10 Language Worked With

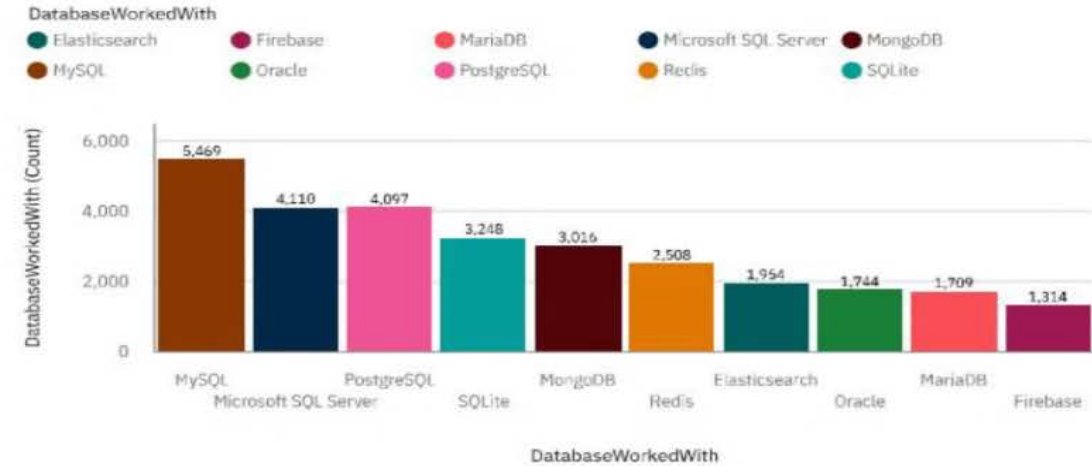


### Platform Worked With



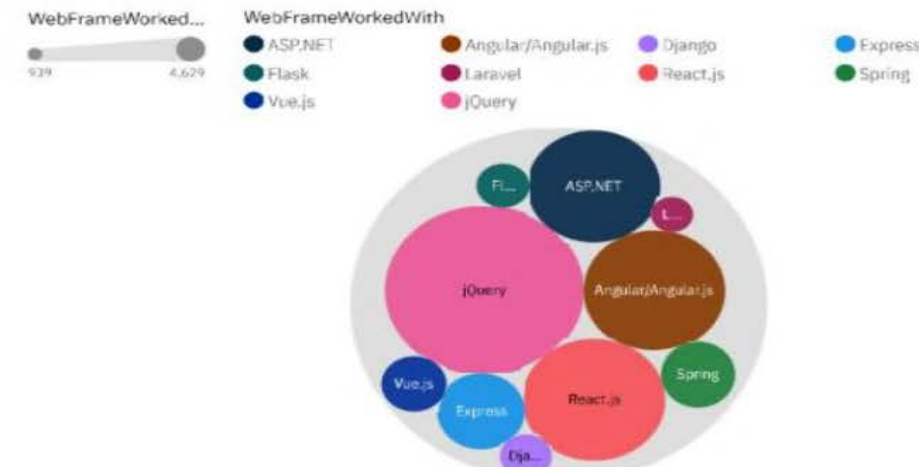
1

### Top 10 Database Worked With



2

### Top 10 Web Frame Worked With



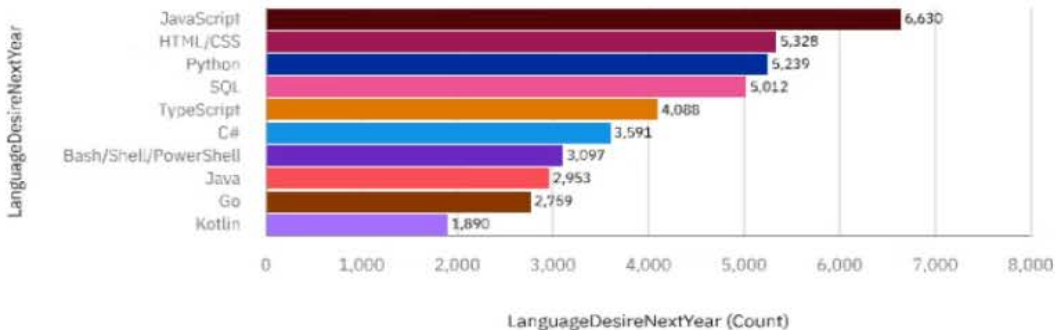
3

# DashBoard 2- Future Technology

## Future Technology

### Top 10 Language Desire Next Year

LanguageDesireNextYear



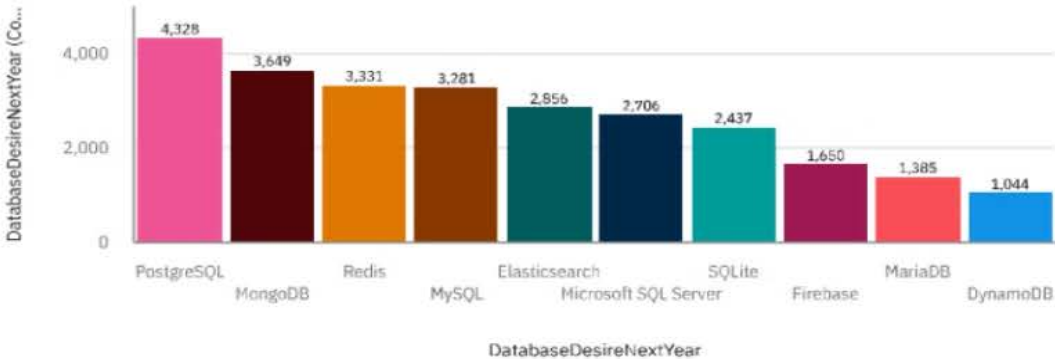
### Platform Desire Next Year



1

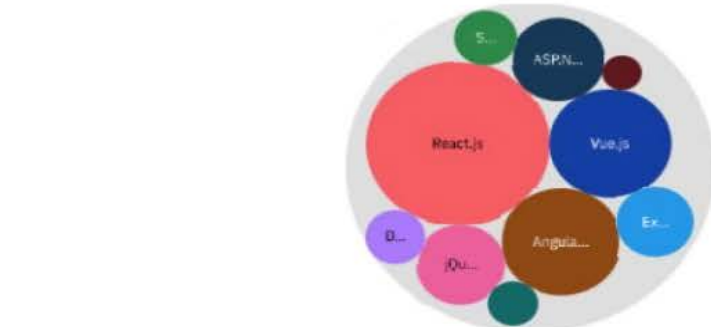
### Top 10 Database Desire Next Year

DatabaseDesireNextYear



2

### WebFrameDesireNextYear hierarchy colored by WebFrameDesireNextYear and sized by WebFrameDesireNextYear



3

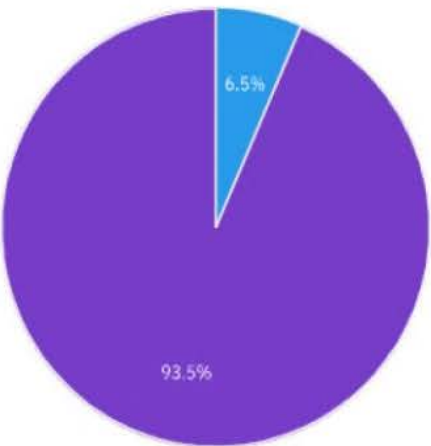


# DashBoard 3- Demographics

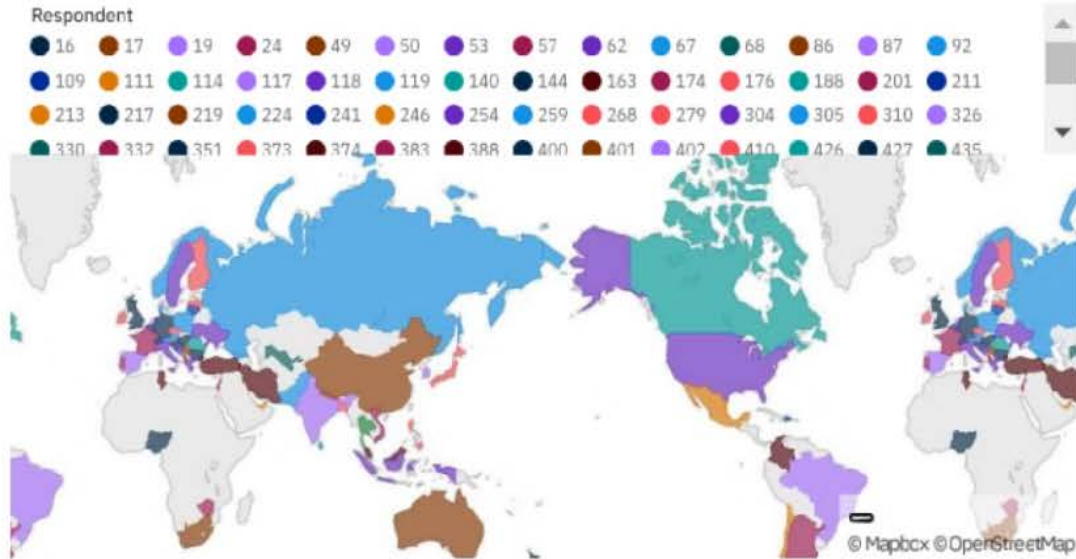
## Demographics

Respondent classified by Gender

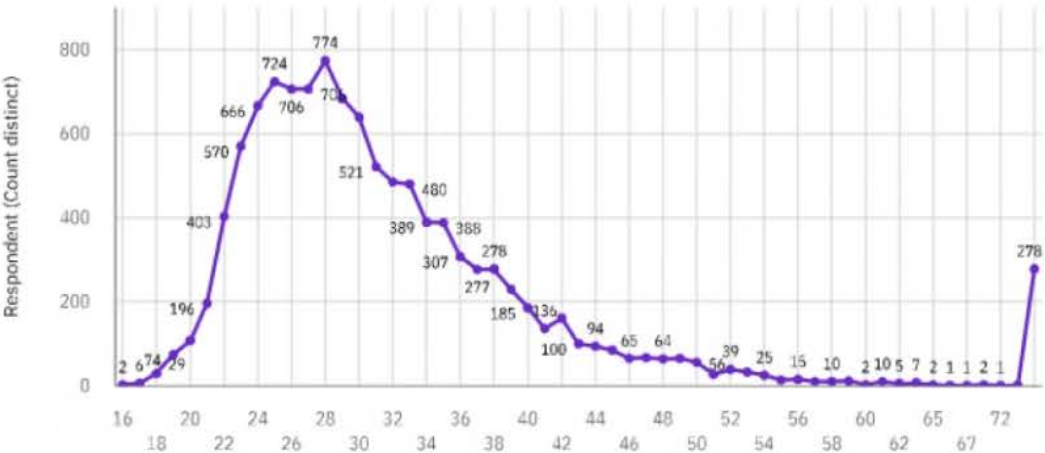
Gender  
● Woman ● Man



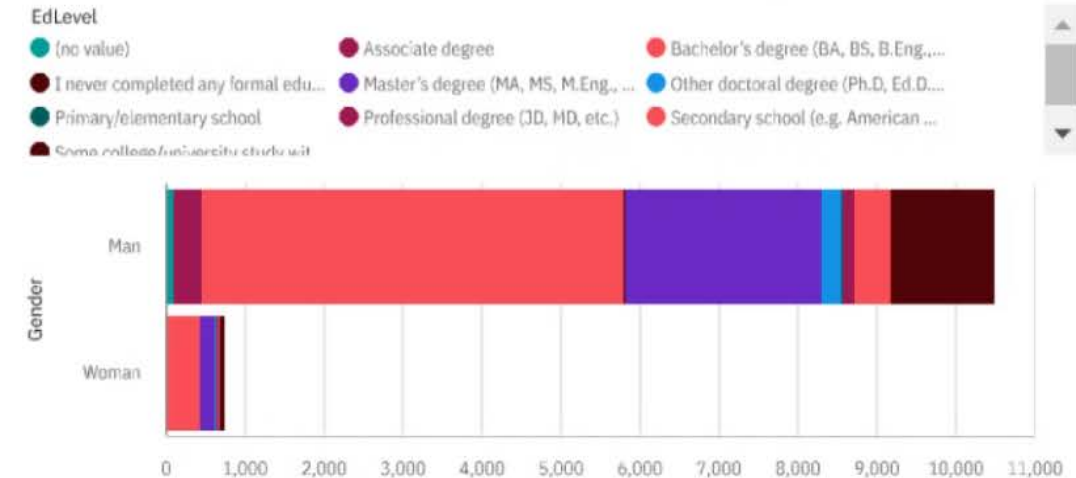
1 Respondent Count for Country



Respondent Count by Age



3 Respondent Count by Gender, classified by Formal Education Level



# Discussion

- Upskilling in the Technology sector.
- How do we close the wide gender gap in the Technology sector?
- Is completing a masters or doctorate degree really a requirement?
- The increasing demand for mobile development as Kotlin is getting popular.
- More tech education, access and development in less developed regions in South east Asia, South America, Africa and some parts of Europe.
- How relevant will Oracle SQL still be in the future?



# OVERALL FINDINGS & IMPLICATIONS

## Findings:

- Most people in the IT field have a Bachelors' degree.
- Web development languages are the most popular and on-demand tools in the IT field currently.
- The Tech sector is filled with majorly young people under 40 years of age.
- Most respondents want to learn Postgre SQL and React JS next year.

## Implications:

- It is important for data professionals to develop proficiency in NoSQL in addition to SQL databases.
- Web development is still a very lucrative skill.
- Less developed countries need more access to tech training and education.

# Conclusion

- Developers are good people with very marked characteristics.
- It is expedient to stay updated in the Tech sector as the trends keep changing over time.

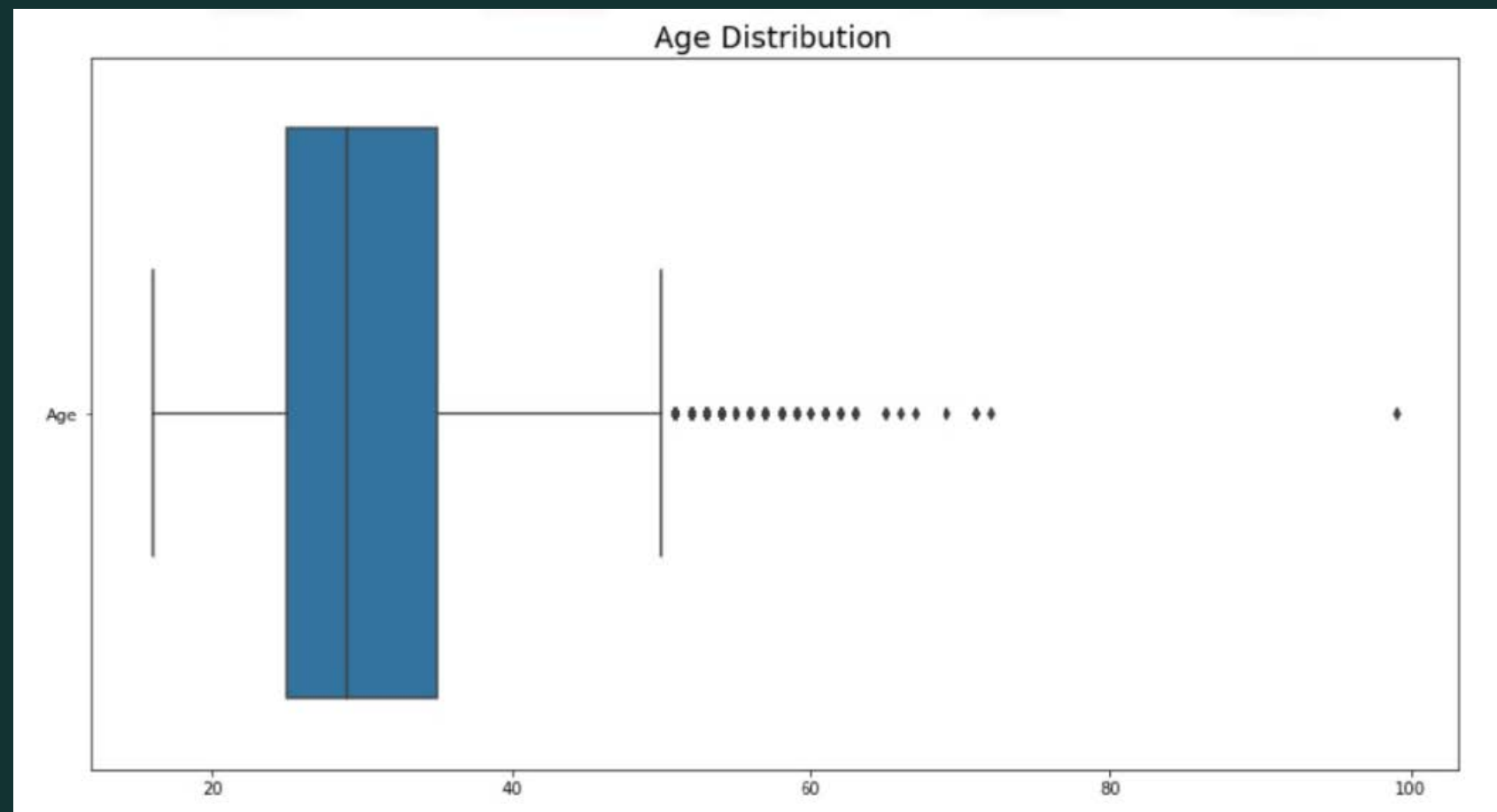


# Appendix

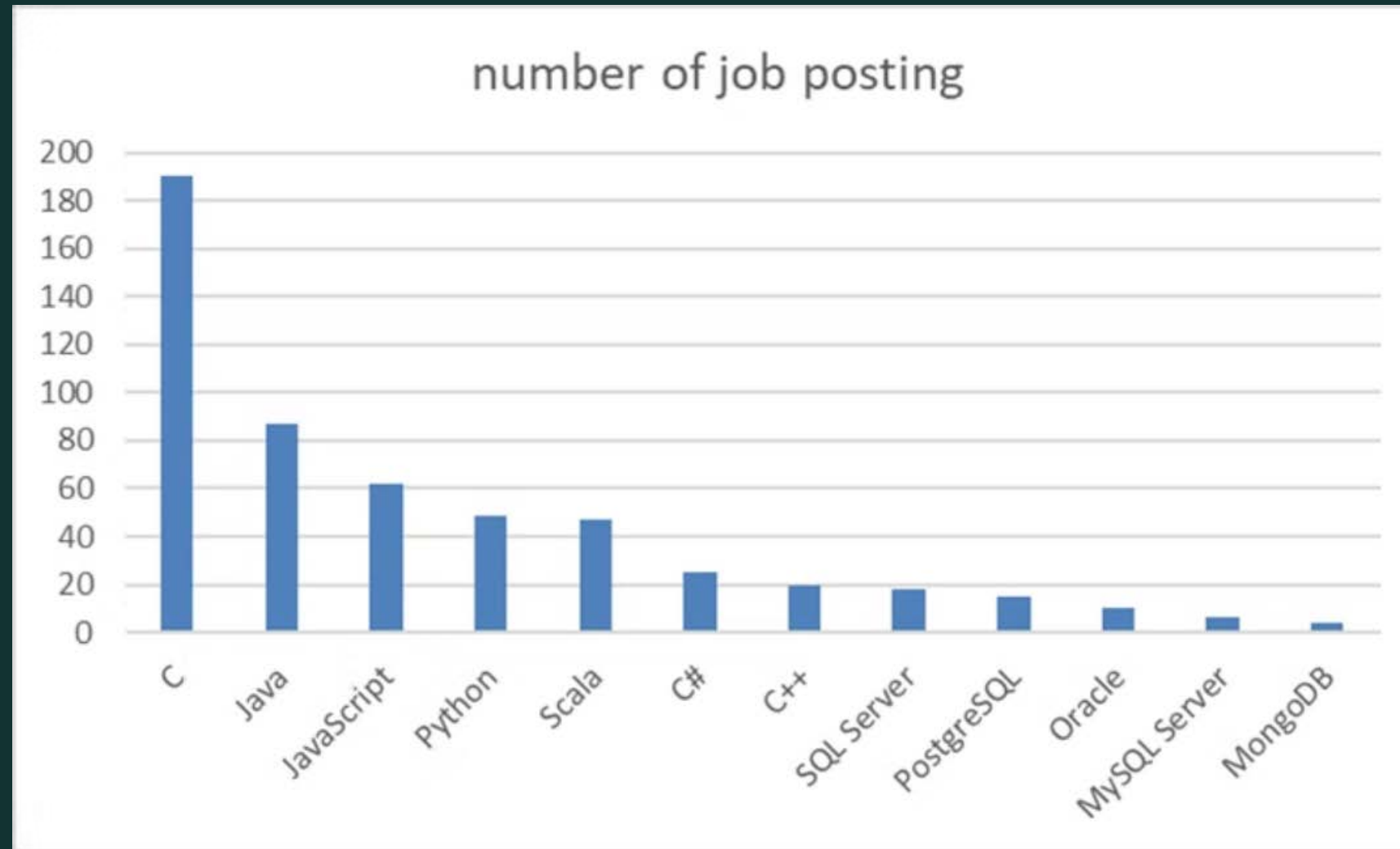




# Age Distribution Box Plot:



# Job Postings:



# Popular Languages:

