Stack Overflow survey
Judah Mwatee
1st June 2025

Skills Network

IBM

OUTLINE

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

EXECUTIVE SUMMARY

In this project, we analyze data to answer key questions about the use of technology. we analyze trends in programming languages, databases, platforms, and frameworks, focusing on those technologies professionals currently use and those they aim to learn in the future.
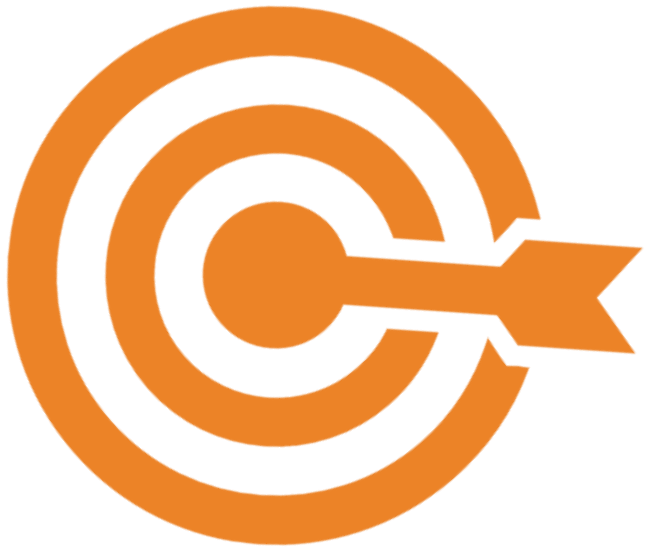
Summary of Methodologies

- Data Collection

- Data Wrangling

- analyze data from a relational database using SQL and Python

- Pandas and NumPy to explore the data

- Visualize data using Jupyter Notebook

- create a dashboard using IBM Cognos Analytics

INTRODUCTION

This capstone project is part of the IBM Data Science Professional Certificate. The goal of the project is to demonstrate proficiency in data science and machine learning techniques using a real-world data and to summarize the results in a report.

The primary task is to analyze the data and obtain valuable insights into current and future technological trends using the latest Stack Overflow Developer Survey dataset. These insights will play a key role in guiding strategic business decisions. Through careful analysis, We extract insights that highlight emerging patterns and shifts in the tech landscape. These findings will then be presented to various stakeholders in the organization for informed decision making.

This project aims to extract, transform, analyze data from a relational database using SQL and Python. Utilize Python's data analysis libraries such as Pandas and NumPy to explore the data. The extracted data will be processed and visualized using Jupyter Notebook. Create a dashboard using IBM Cognos Analytics or Google Looker Studio.

IBM

METHODOLOGY

Collect survey data

Web scraping
APIs
Request library

Data wrangling

Exploratory Data Analysis (EDA)

Distribution
Outlier Handling
Correlation

Visualization

Distribution of data
Relationship
Comparison
Composition

Dashboards

RESULTS

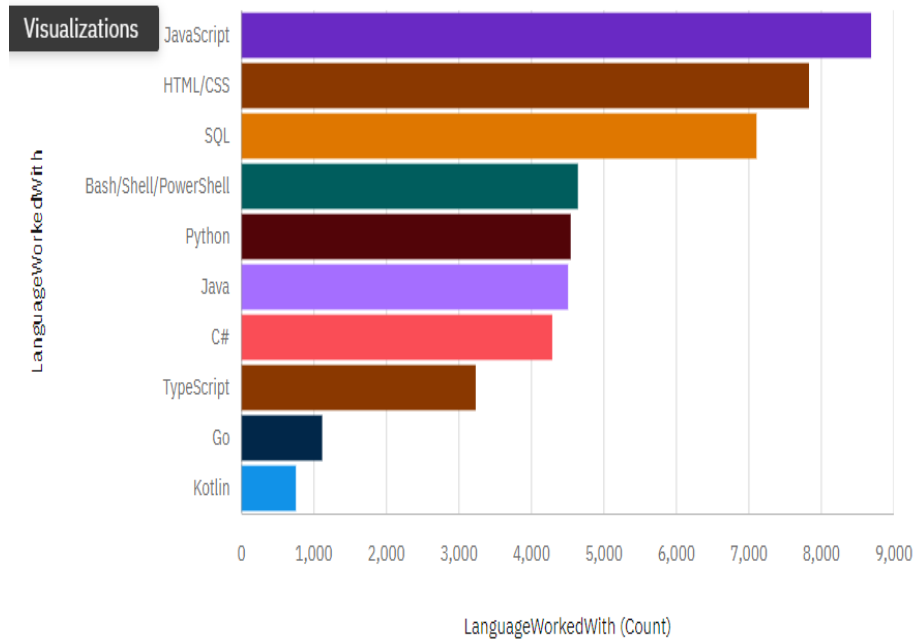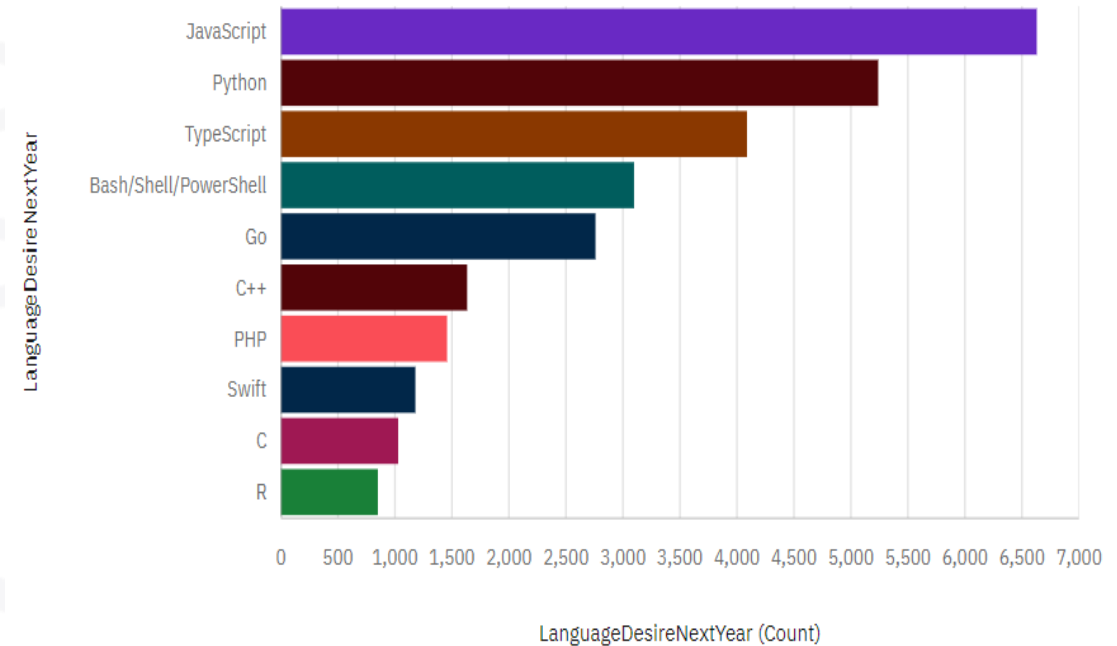**Normalize Data**

PROGRAMMING LANGUAGE TRENDS

PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

- Javascript leading
- HTML,SQL might not use very much
- Python might lead but Kotlin will jump it's place significanlty

Implications

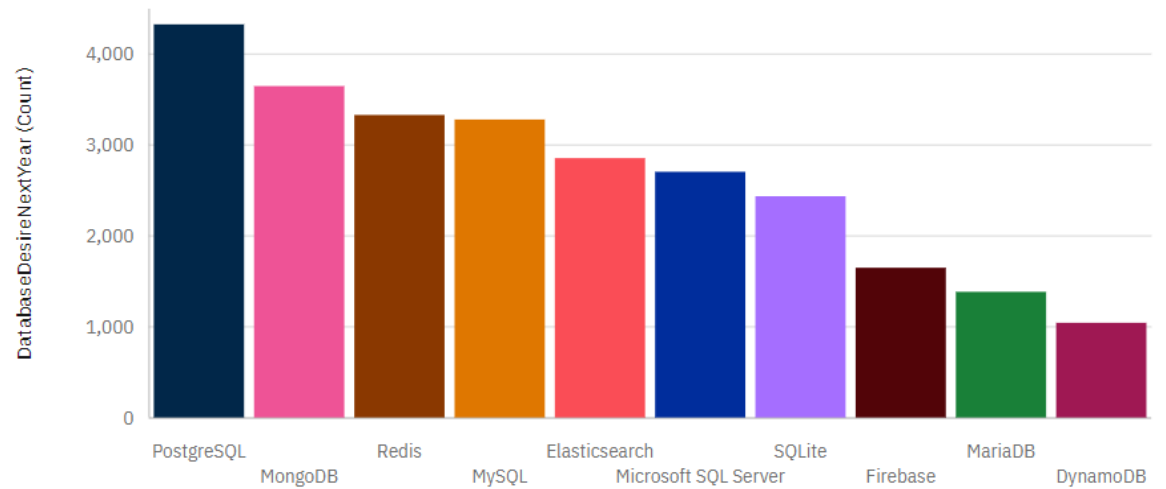- Developers will need to learn new languages in the near future

DATABASE TRENDS

Current Year

Next Year

Skills Network
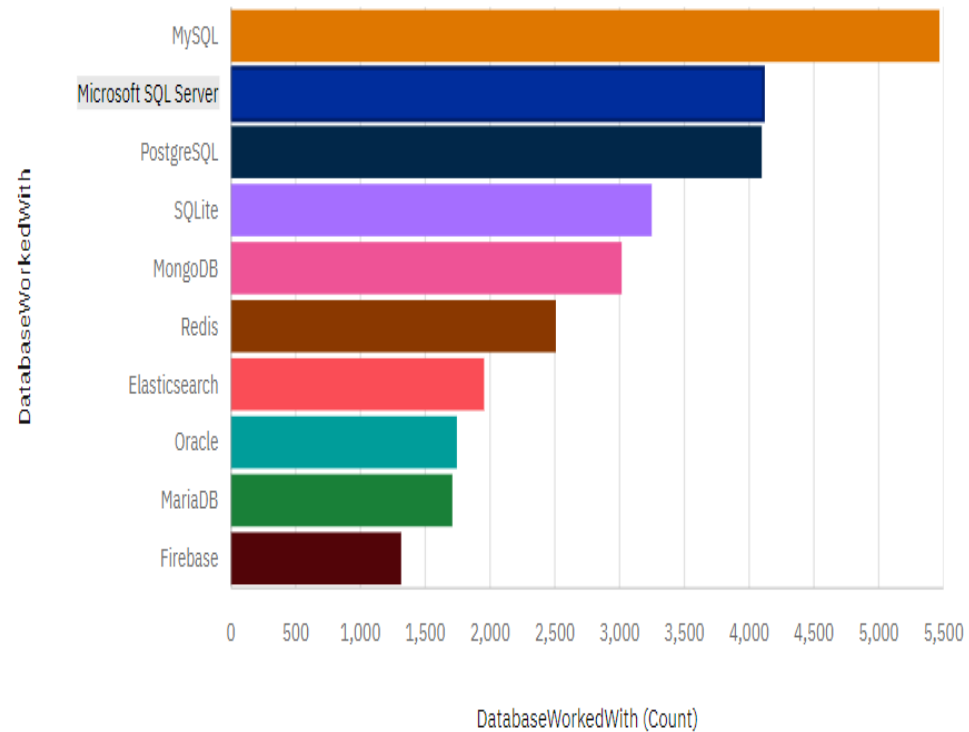
DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- MySQL as most used database.
- Lack of interest in Microsoft SQL Server and SQLite.
- Increasing interest in PostgreSQL and MongoDB.

Implications

- Microsoft SQL Server and SQLite losing ground in the market.
- PostgreSQL and MongoDB establishment in the market.
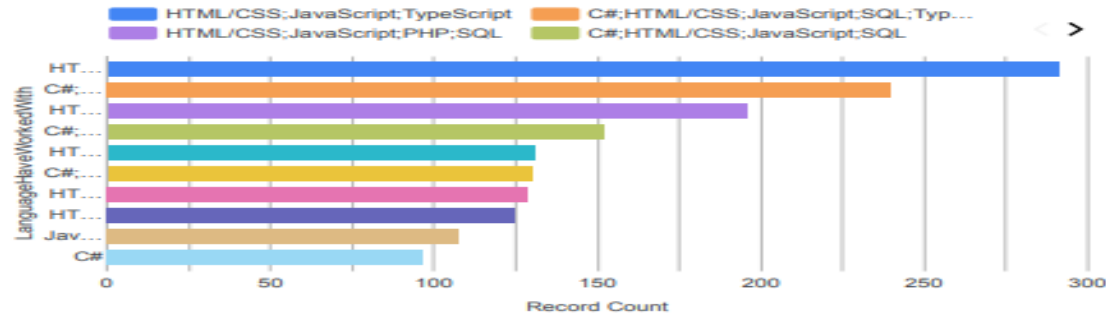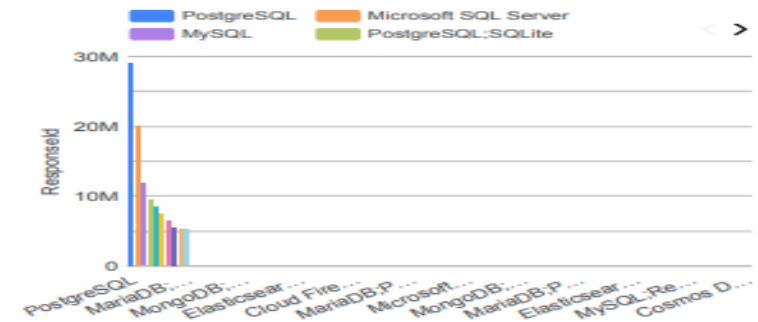
DASHBOARD

https://github.com/JudahMwatee/TestRepo/blob/main/IBM_Project.pdf

DASHBOARD TAB 1
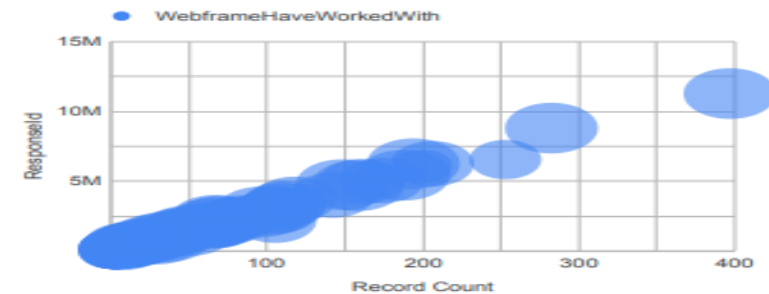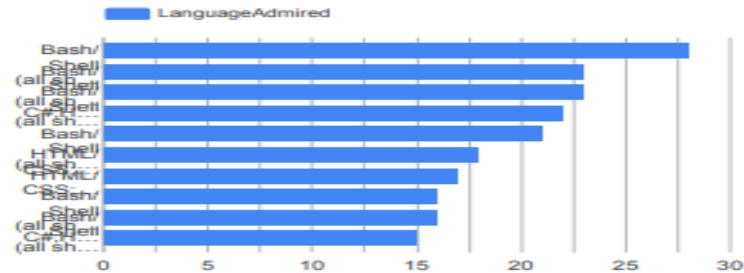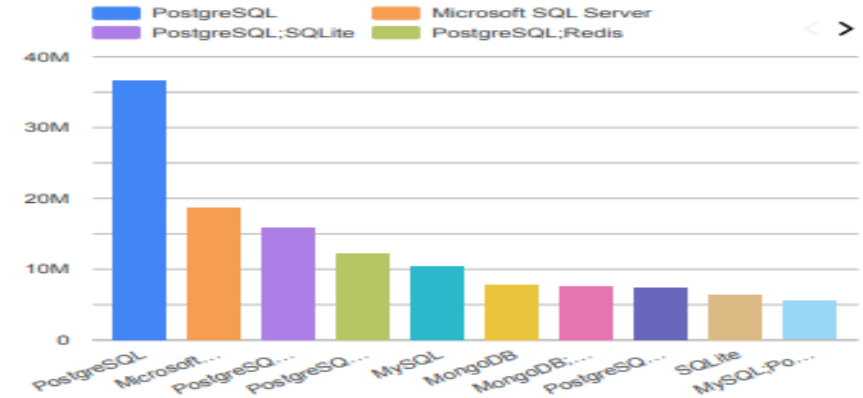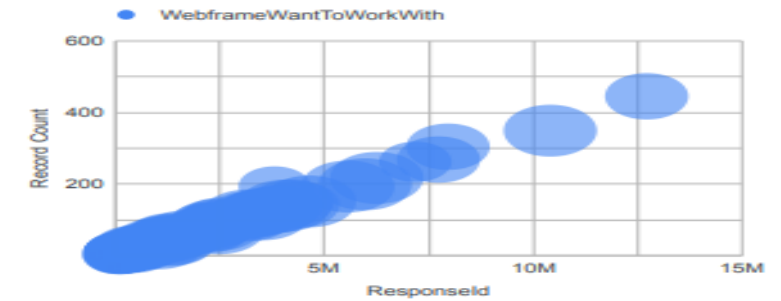
DASHBOARD TAB 2

DASHBOARD TAB 3
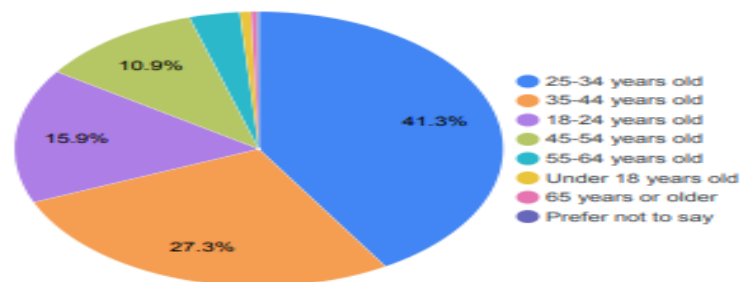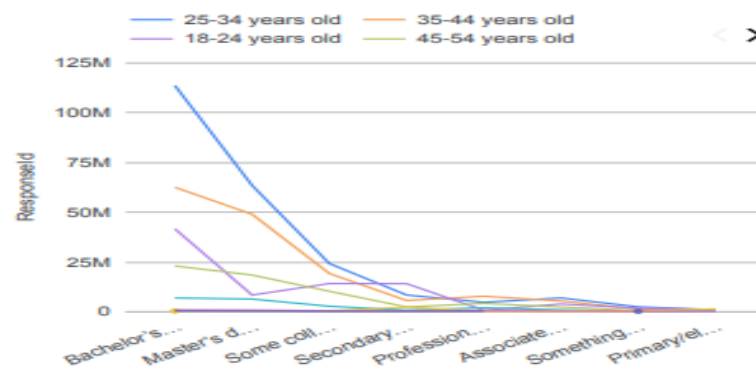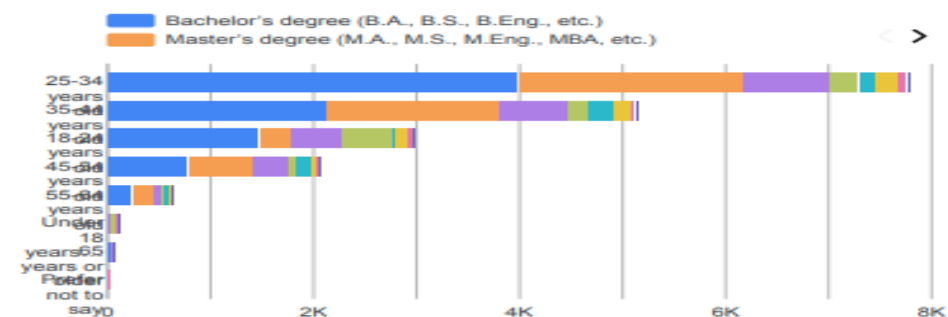
DISCUSSION

OVERALL FINDINGS & IMPLICATIONS

Findings

- JavaScript widely used and TypeScript getting popular.

- Over 90% young male developers.

- Developers mostly located in developed countries.

Implications

- JavaScript and TypeScript web frames gaining followers.

- Global polarization of developers location and gender.

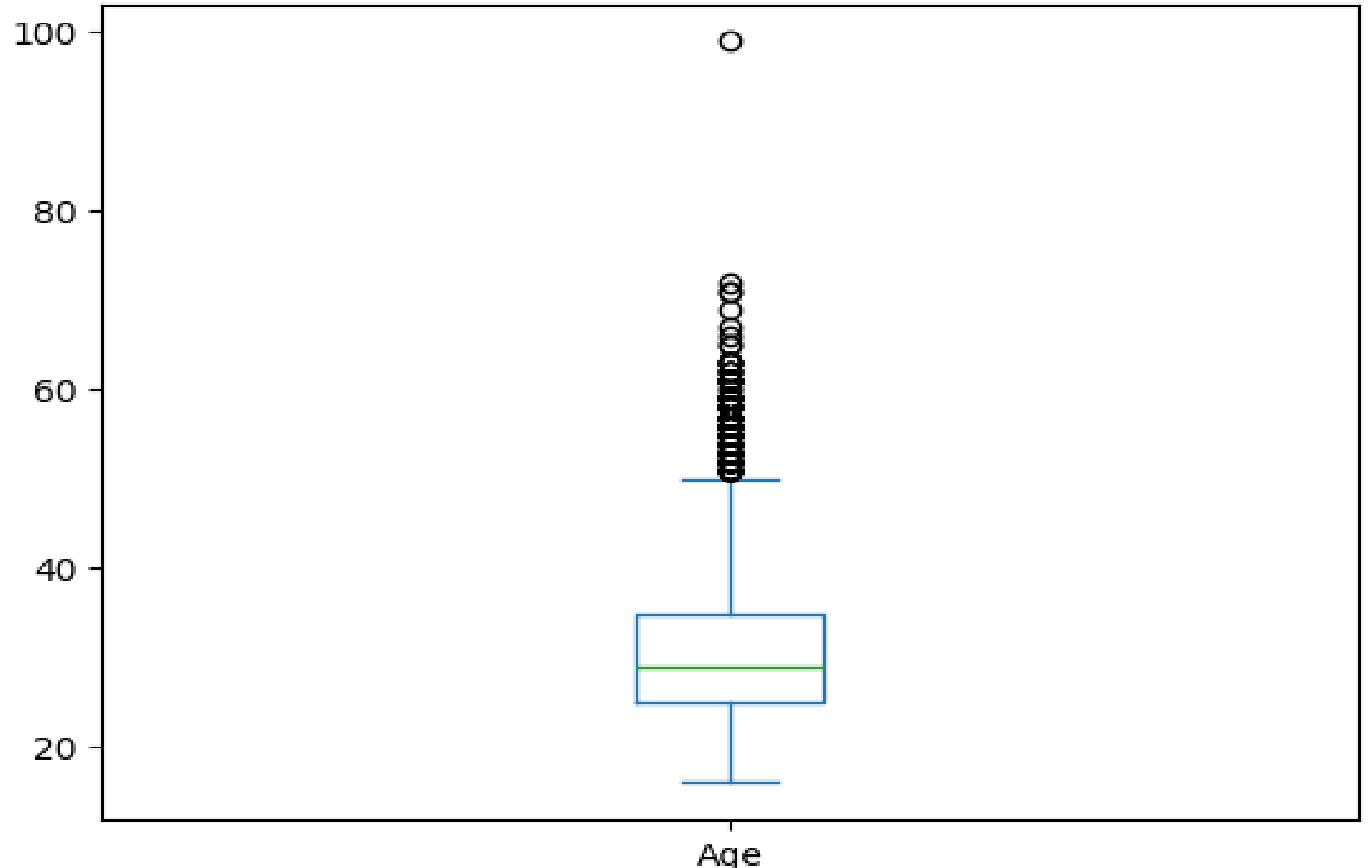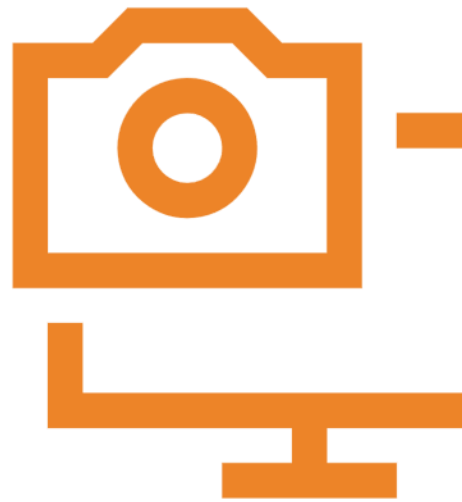- Young developers without postgrad studies on its majority.

CONCLUSION

- Developers are people with very marked characteristics.

- A good idea of popularity trends of different tools, platforms and languages can be obtained.

- There is a job to be done to spread accessibility of this labor market to countries in development.

IBM

APPENDIX

JOB POSTINGS

POPULAR LANGUAGES

Bar chart: Average Annual Salary by Language

Language	Average Annual Salary
Swift	130.801
Python	114.383
C++	113.865
Javascript	110.981
Java	101.013
Go	94.082
R	92.037
C#	88.726
SQL	84.793
PHP	84.727