IBM DATA ANALYST CAPSTONE PROJECT

Analysis on Emerging Technology Skills and Trends

OUTLINE

EXECUTIVE SUMMARY

- Data contextualization and analysis

- Methodology description

  – Data gathering from IBM site

  – Data analysis by python and MySQL

  – Data visualizations using IBM Cognos

- Results supported with graphs and trends

- Discussion of overall finding and implications regarding the results

- Final conclusions of the study was carried out

INTRODUCTION

- The presentation report uses data analytics to highlight current project trends in need for skills related to programming language, databases and web frames.

- The following were investigated using the data provided

  – Programming languages currently in demand

  – Database skills in demand

  – Popular IDEs or WebFrames currently in demand and in future

- Target audience for this research are IT professionals, HR managers and anyone interested in IT sector.
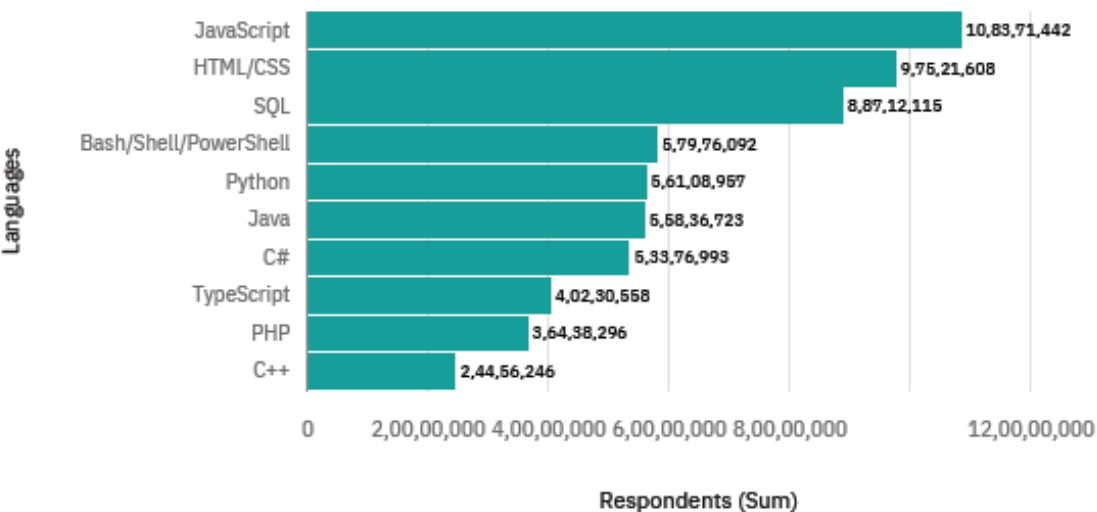
METHODOLOGY
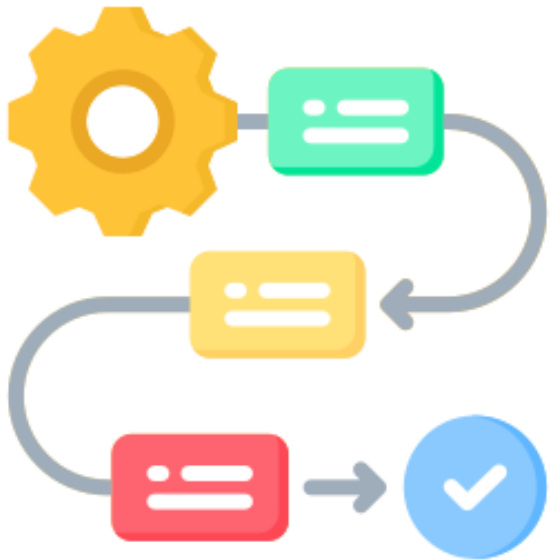
- Collect survey data and understand the data

    – Web scrapping, APIs, request libraries

- Data wrangling

- Exploratory Data Analysis (EDA)

    – Analyze data distribution

    – Handling outliers

    – Correlations

- Data visualization

    – Highlight distribution of data, relationships and comparison of data

- Creation of dashboards

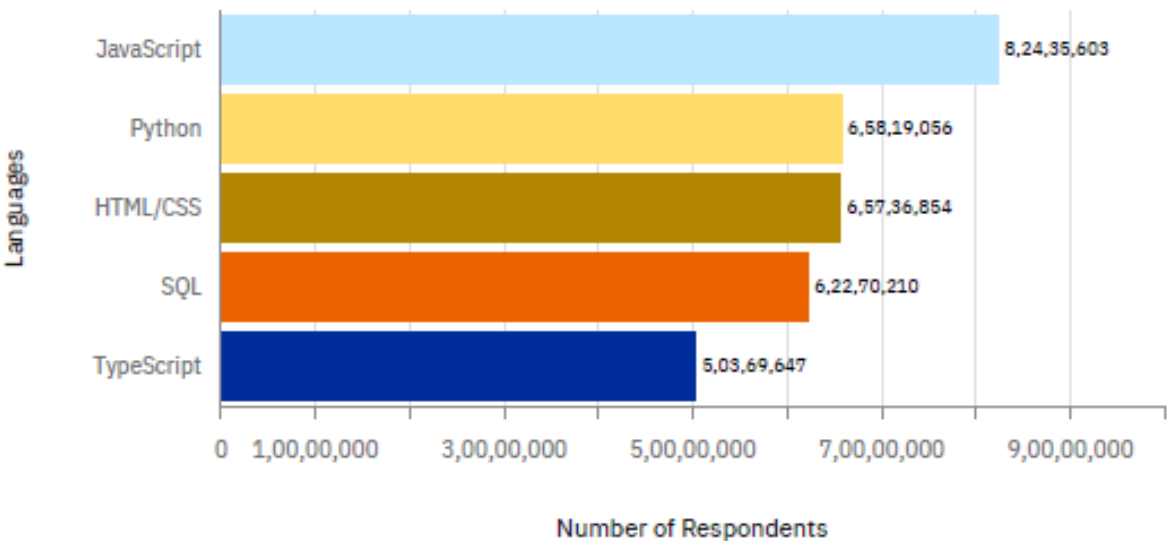# RESULTS

PROGRAMMING LANGUAGE TRENDS

CURRENT YEAR

NEXT YEAR

Top 10 Languages worked with

Languages	Respondents (Sum)
JavaScript	10,83,71,442
HTML/CSS	9,75,21,608
SQL	8,87,12,115
Bash/Shell/PowerShell	5,79,76,092
Python	5,61,08,957
Java	5,58,36,723
C#	5,33,76,993
TypeScript	4,02,30,558
PHP	3,64,38,296
C++	2,44,56,246

Top 5 Desired Languages to learn

Languages	Number of Respondents
JavaScript	8,24,35,603
Python	6,58,19,056
HTML/CSS	6,57,36,854
SQL	6,22,70,210
TypeScript	5,03,69,647

Programming Languages Trends – Findings and Implications

Findings

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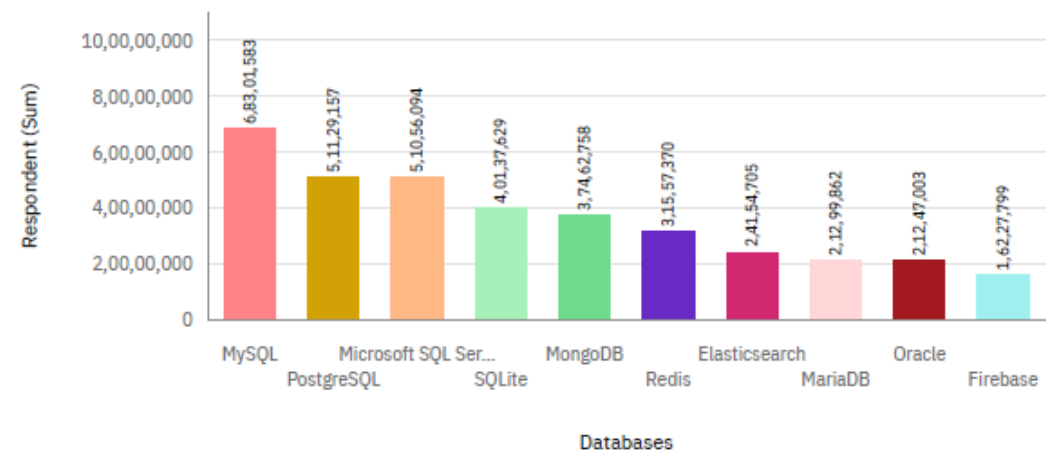
Implications

- Javascript and HTML are used for web development which means that web development skill has the highest demand.
- Python is gaining more traction due to increase in demand for AI and ML skills
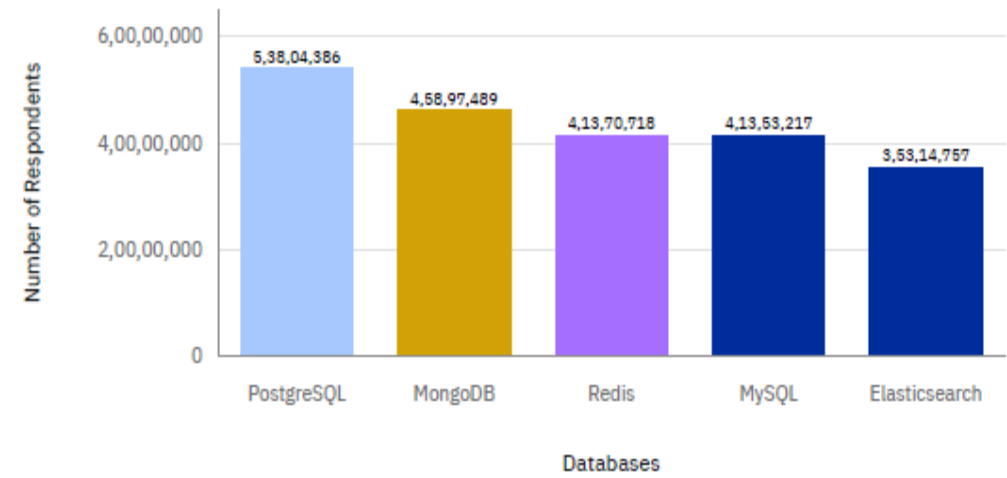- SQL is the most relevant language for data professionals.

DATABASE TRENDS

CURRENT YEAR

Top 10 Databases worked with

NEXT YEAR

Top 5 desired Databases to Learn

Database – Findings and Implications

Findings

➢ MySQL, Microsoft SQL server, PostgreSQL, SQLite and MongoDB are the top 5 most used databases at the moment.

➢ Redis and Elasticsearch are relatively new tools and are set to gain more traction in the new IT space.

➢ PostgreSQL remain at the top of list in both currently used and future trend databases.

Implications

➢ PostgreSQL is the most desired database to learn next year, even though it ranks second in current usage. This might indicate a growing preference for open-source relational databases that offer versatility and scalability.

➢ Companies still prefer open source databases.

➢ Despite its current lower usage, the strong interest in learning Elasticsearch suggests a growing trend toward full-text search engines and analytics tools in modern applications.
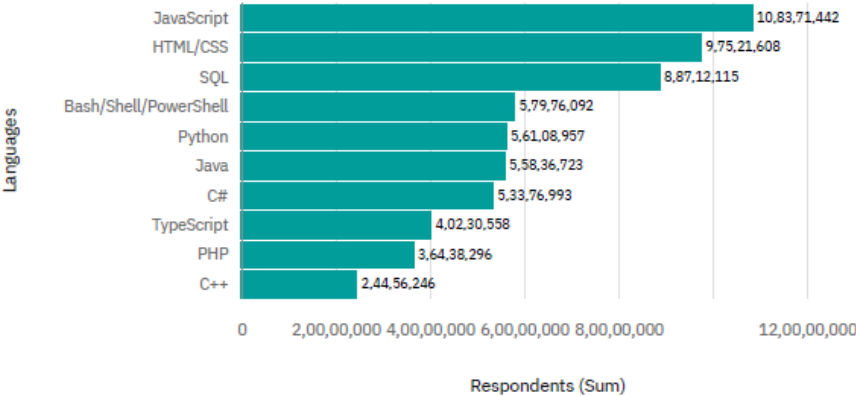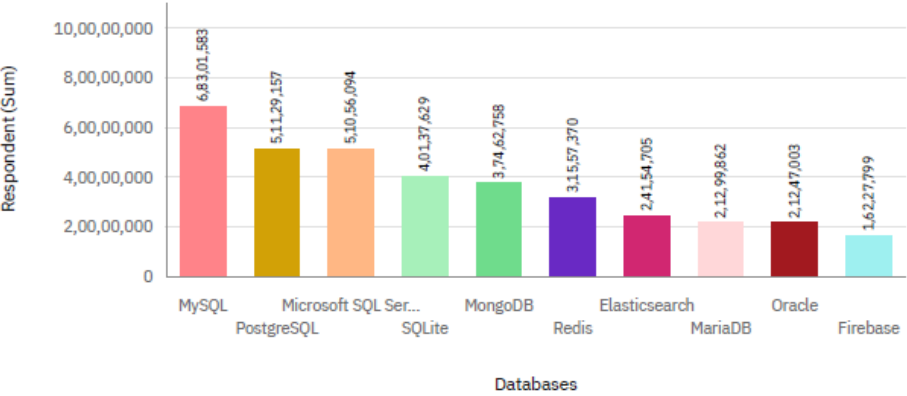
DASHBOARD LINK

https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&id=iDCB4187228F54F5DB655A1B27CEC4820&objRef=iDCB4187228F54F5DB655A1B27CEC4820

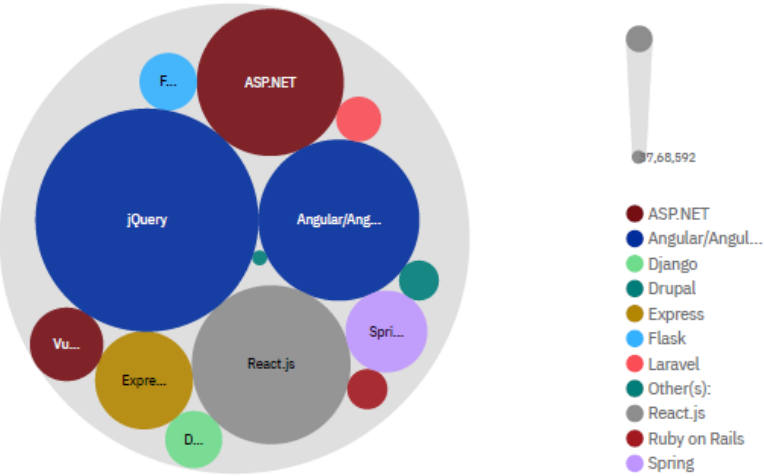# CURRENT TECHNOLOGY USAGE – Dashboard

Top 10 Languages worked with

Language	Respondents (Sum)
JavaScript	10,83,71,442
HTML/CSS	9,75,21,608
SQL	8,87,12,115
Bash/Shell/PowerShell	5,79,76,092
Python	5,61,08,957
Java	5,58,36,723
C#	5,33,76,993
TypeScript	4,02,30,558
PHP	3,64,38,296
C++	2,44,56,246

Top 10 Databases worked with

Database	Respondent (Sum)
MySQL	6,83,01,583
PostgreSQL	5,11,29,157
Microsoft SQL Ser...	5,10,56,094
SQLite	4,01,37,629
MongoDB	3,74,62,758
Redis	3,15,57,370
Elasticsearch	2,41,54,705
MariaDB	2,12,99,862
Oracle	2,12,47,003
Firebase	1,62,27,799

Platforms Interacted

aws Linux Windows Android Heroku Docker WordPress Google Cloud Platform Microsoft Azure MacOS Kubernetes Raspberry Pi Arduino iOS Slack IBM Cloud or Watson OpenshiFt

Top 10 WebFrames used

- ASP.NET
- Angular/Angul...
- Django
- Drupal
- Express
- Flask
- Laravel
- Other(s):
- React.js
- Ruby on Rails
- Spring

7,68,592

FUTURE TECHNOLOGY TREND – Dashboard

Top 5 Desired Languages to learn

Language	Number of Respondents
JavaScript	8,24,35,603
Python	6,58,19,056
HTML/CSS	6,57,36,854
SQL	6,22,70,210
TypeScript	5,03,69,647

Top 5 desired Databases to Learn

Database	Number of Respondents
PostgreSQL	5,38,04,386
MongoDB	4,58,97,489
Redis	4,13,70,718
MySQL	4,13,53,217
Elasticsearch	3,53,14,757
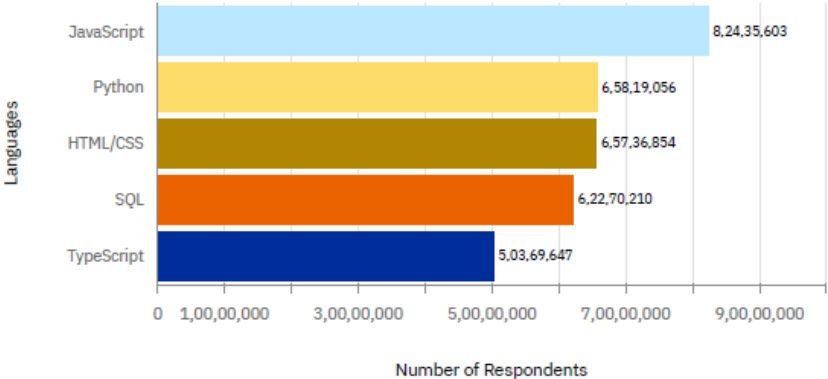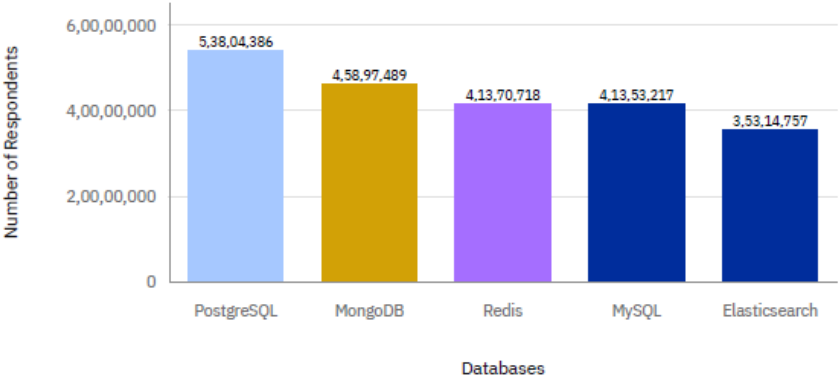
Desired platform - Next year

Linux, Docker, AWS, Windows, Android, Kubernetes, MacOS, Raspberry Pi, Google Cloud Platform, Slack, iOS, Heroku, Microsoft ..., Arduino, WordPress, IBM Clou...
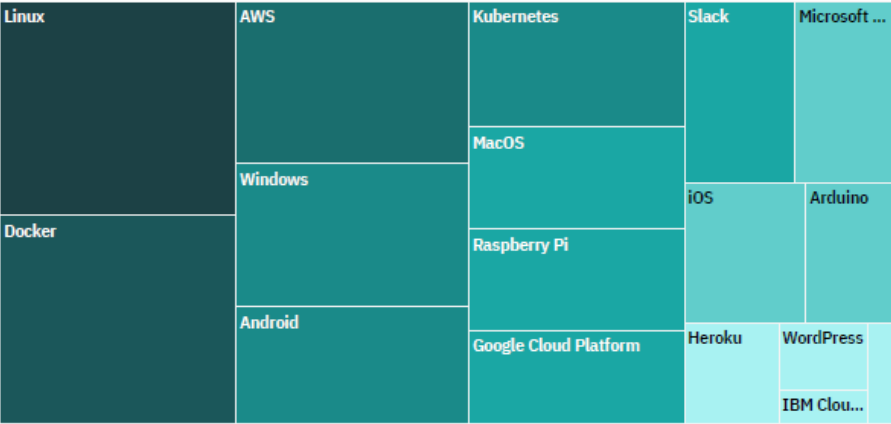
Top 10 desired Webpage

Respondent (Sum)
1,18,86,651

WebFrameDesireNextY...
- ASP.NET
- React.js
- Angula...
- Ruby o...
- Django
- Spring
- Express
- Vue.js
- Flask
- jQuery
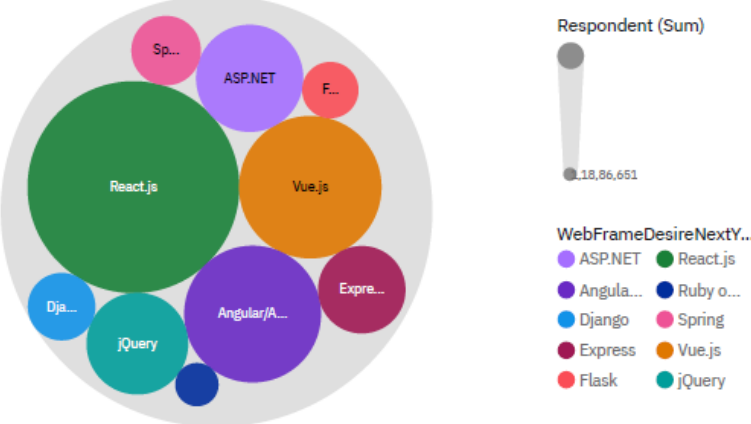
DEMOGRAPHICS- Dashboard
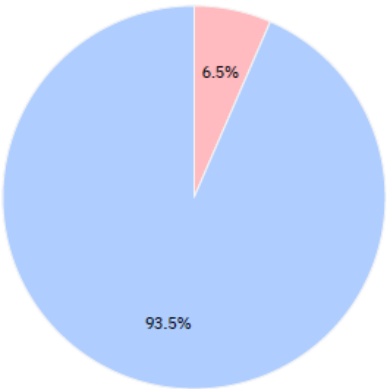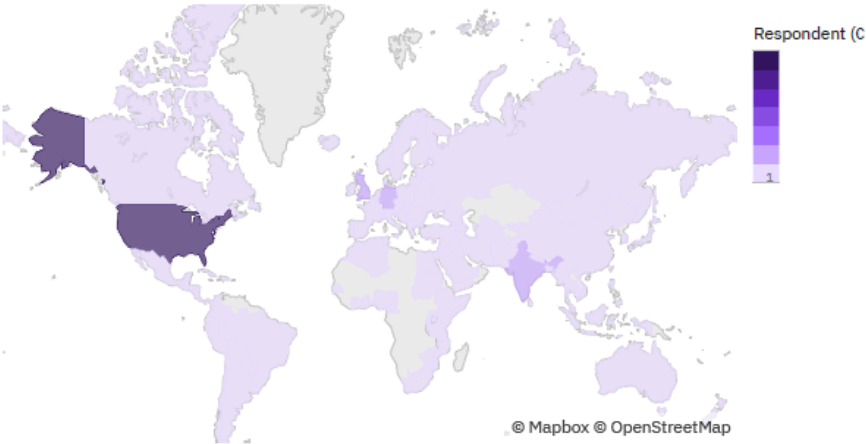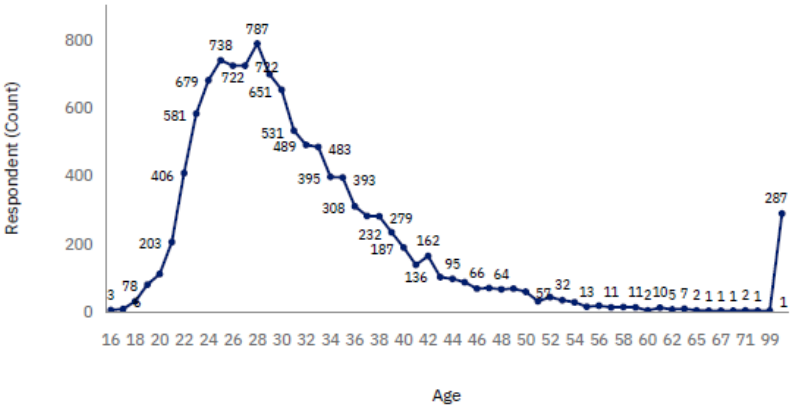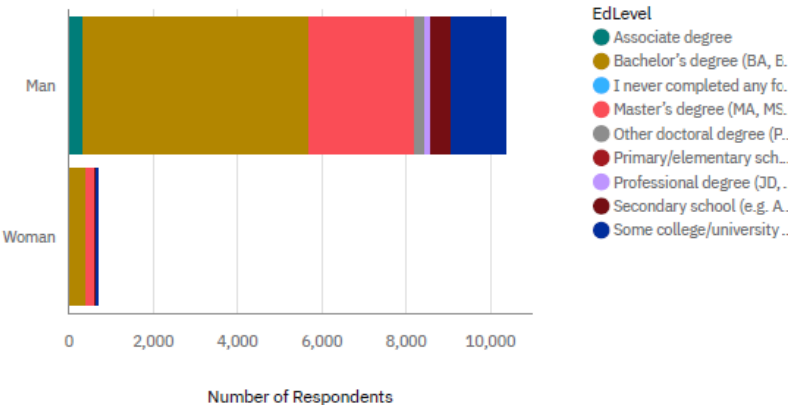
DISCUSSION

Overall findings and implications

Findings

➢ Most people in the IT profession have a bachelor's degree.

➢ Web development languages are the most popular and on-demand tools in the IT field currently.

➢ The tech sector is majorly filled with young people under 40 years of age.

Implications

➢ It is important for data professionals to develop proficiencies in NoSQL in addition to SQL databases.

➢ Web development is still a very lucrative skill.

➢ Less developed countries need more access to tech trainings and education.
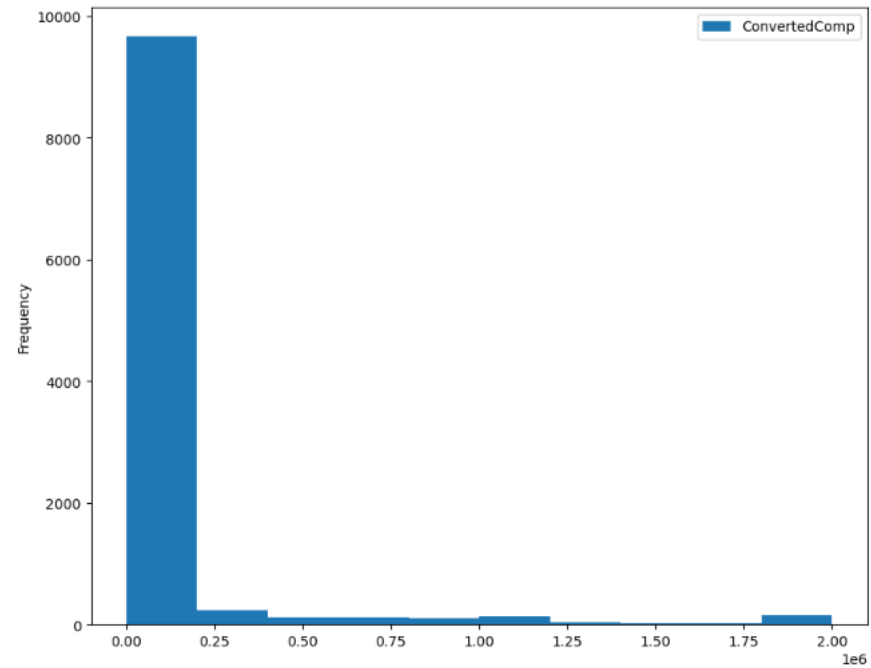
CONCLUSION

It is expedient to stay updated in the tech sector as the trends keep changing over time..
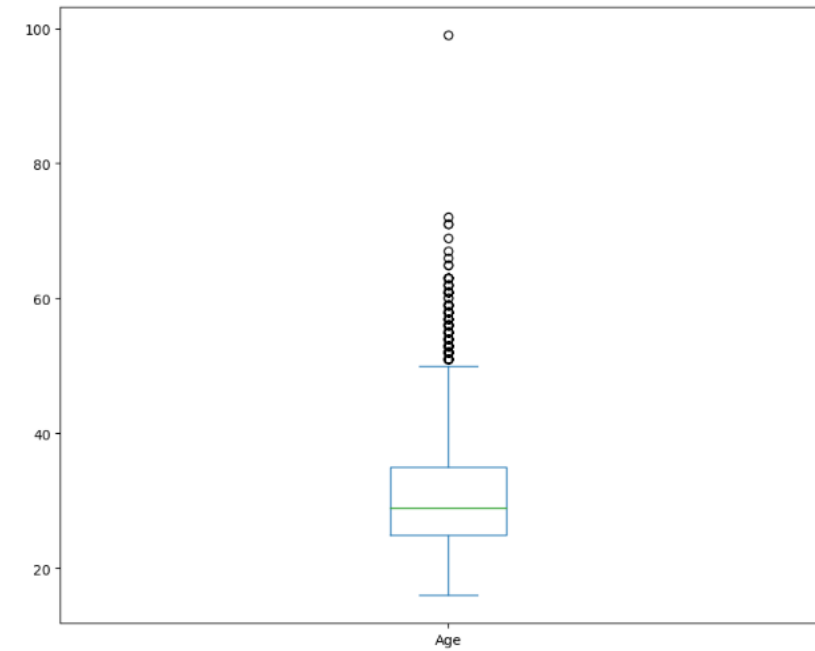
APPENDIX


```
#Plot a histogram of ConvertedComp

QUERY1 = """
SELECT ConvertedComp
FROM master
"""
df = pd.read_sql_query(QUERY1,conn)
df.head()
df.plot(kind='hist', figsize=(10, 8) )

plt.show()
```
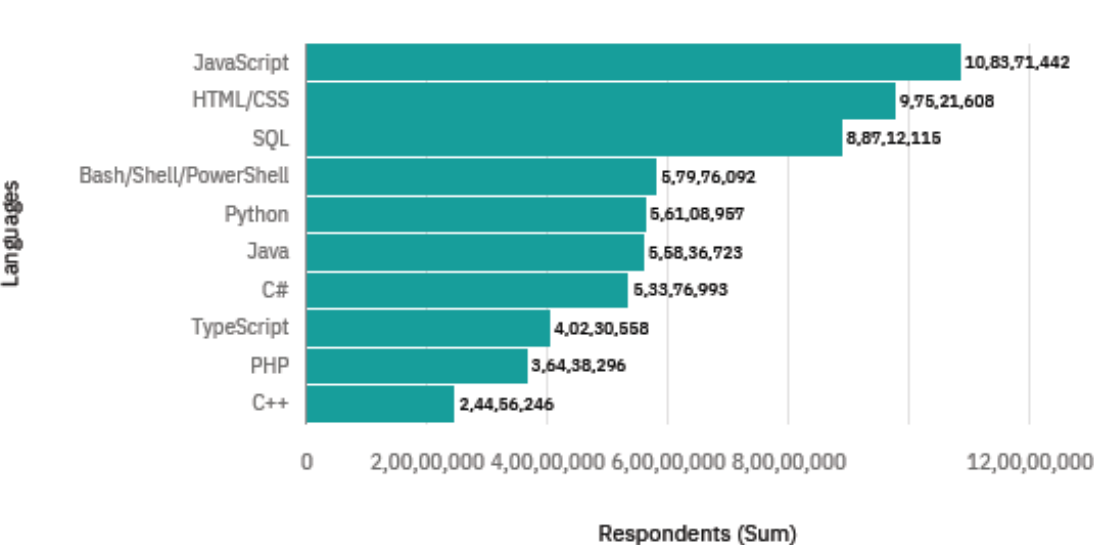


```
#Plot a box plot of Age
QUERY2 = """
SELECT Age
FROM master
"""
df = pd.read_sql_query(QUERY2,conn)
df.head()
df.plot(kind='box', figsize=(10, 8) )

plt.show()
```

Popular languages

Job postings

Top 10 Languages worked with

Language	Respondents (Sum)
JavaScript	10,83,71,442
HTML/CSS	9,75,21,608
SQL	8,87,12,115
Bash/Shell/PowerShell	6,79,76,092
Python	6,61,08,957
Java	6,58,36,723
C#	6,33,76,993
TypeScript	4,02,30,558
PHP	3,64,38,296
C++	2,44,56,246

Location	Number of Jobs
Washington DC	5316
Detroit	3945
Seattle	3375
New York	3226
Los Angeles	640
San Francisco	435
Austin	434