JUNIOR DATA ANALYST FINAL PROJECT.

BY:

Victor Nwaezeapu

12/08/2024

IBM Developer

SKILLS NETWORK

OUTLINE

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

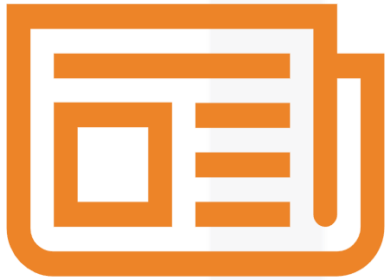
EXECUTIVE SUMMARY

- **Overview:** The project involved the creation of a dashboard with three distinct tabs using IBM Cognos Analytics. Based on the survey results, the dashboard is designed to provide insights into current technology usage, future technology trends, and demographic data.
- **Current Technology Usage:**
  - **Top 10 Languages Worked With:** A bar chart was utilized to display the most commonly used programming languages among respondents.
  - **Top 10 Databases Worked With:** A column chart showcased the leading databases that respondents have experience with.
  - **Platforms Worked With:** A word cloud visualized the various platforms respondents have used.
  - **Top 10 Web Frameworks Worked With:** A hierarchical bubble chart highlighted the most popular web frameworks.
- **Future Technology Trend:**
  - **Top 10 Desired Languages:** A bar chart outlined the programming languages that respondents aspire to work with in the next year.
  - **Top 10 Desired Databases:** A column chart presented the databases respondents wish to explore further.
  - **Desired Platforms:** A tree map chart provided an overview of the platforms respondents are interested in adopting.
  - **Top 10 Desired Web Frameworks:** A hierarchical bubble chart depicted the web frameworks that are gaining interest.
- **Demographics:**
  - The demographics dashboard included insights into the distribution of respondents by gender, age, and education level, visualized through various charts like pie, map, line, and stacked bar charts.
- **Conclusion:** The dashboards collectively provide a comprehensive view of the current state and future aspirations in the technology landscape, backed by detailed demographic insights.

INTRODUCTION

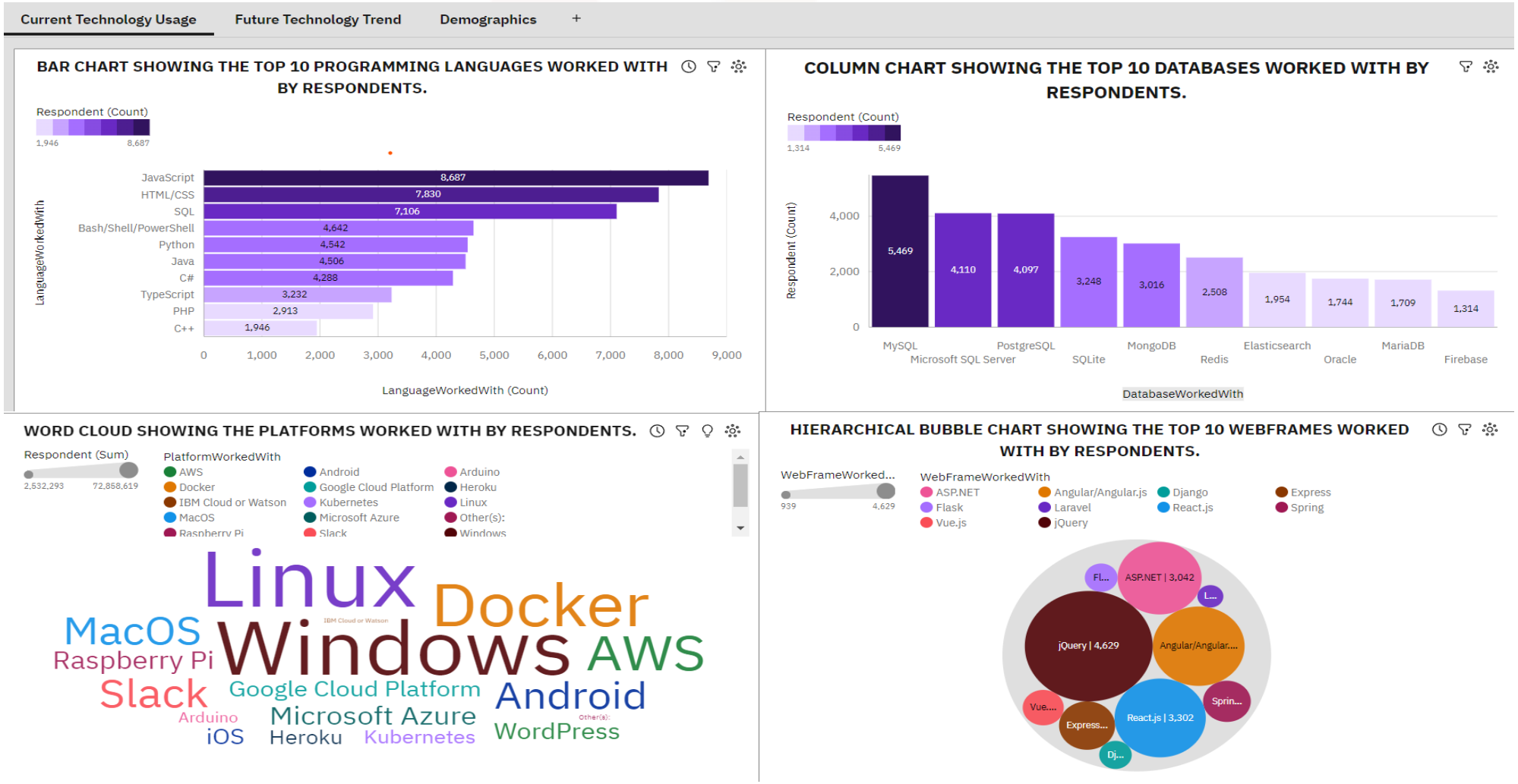
- **Project Objective:** The primary objective of this project was to leverage IBM Cognos Analytics to create a series of dashboards that provide actionable insights into various technological trends and demographic distributions.
- **Data Sources:** The project utilized two main datasets: one focusing on technology usage and preferences, and another on demographic information. These datasets were crucial in constructing visual representations that align with the project's goals.
- **Dashboard Design:** The dashboards were designed to present data clearly and concisely, ensuring that users can quickly understand key trends. The dashboards are divided into three main sections: Current Technology Usage, Future Technology Trends, and Demographics.
- **Analysis Focus:**
  - **Current Technology Usage:** This section of the dashboard offers insights into the technologies currently being used by professionals.
  - **Future Technology Trends:** This section focuses on the technologies that respondents intend to use in the future.
  - **Demographics:** The final section provides an overview of the demographic composition of the respondents, segmented by various factors such as age, gender, and education level.

METHODOLOGY



- **Data Collection**: The data for this project were from two primary sources, consisting of survey results focusing on technology usage and demographic information. The datasets used were “m5_survey_data_technologies_normalised.csv” and “m5_survey_data_demographics.csv”, which provided the foundation for all subsequent analyses.
- **Data Preparation**: The collected datasets were uploaded as data assets into IBM Cognos Analytics. Any irrelevant or incomplete entries were excluded from the analysis.
- **Dashboard Development**: The analysis was divided into three key tabs on the dashboard, each focusing on a different aspect of the survey data:
  - Current Technology Usage Tab: This dashboard section utilized the technologies dataset to visualize the most commonly used programming languages, databases, platforms, and web frameworks among respondents.
  - Future Technology Trends Tab: This dashboard section highlights the technologies which the survey respondents desire to use in the coming year, providing insights into emerging trends.
  - Demographics Tab: This dashboard section focused on the demographic characteristics of the respondents, using filters to isolate specific groups and generate relevant visualizations.
- **Visualization Techniques**:
  - Chart Selection: Appropriate chart types were selected based on the nature of the data being visualized. For instance, bar charts were used for ranking items, while pie charts were used for displaying proportions.
  - Customization: Each chart was customized to include value labels, titles, and color schemes that enhance readability and comprehension. This ensured that the visualizations effectively communicated the intended insights.

RESULTS



IBM Developer

SKILLS NETWORK

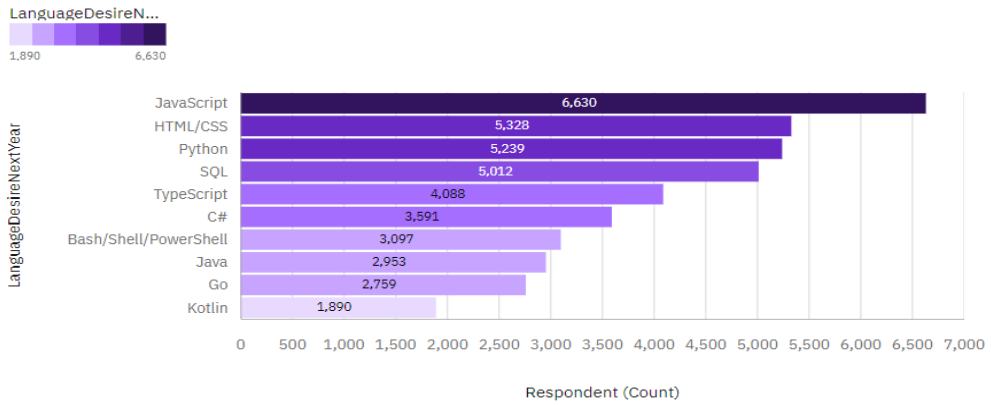
RESULTS

Current Technology Usage

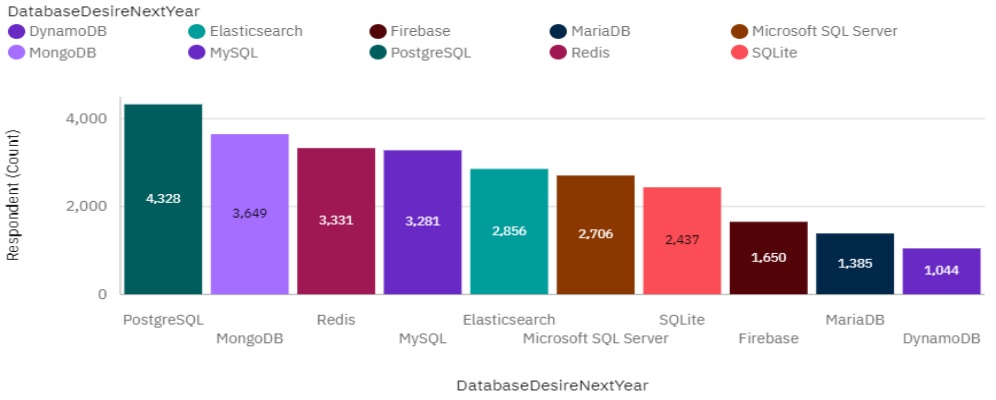
- **Top 10 Languages Worked With:** The analysis revealed that the most commonly used programming languages among the survey respondents included JavaScript, HTML/CSS, and SQL. These languages dominated the technology landscape, indicating their widespread adoption and continued relevance in the industry.
- **Top 10 Databases Worked With:** The database technologies most frequently used by respondents included MySQL, Microsoft SQL Server, and PostgreSQL. These databases are critical components of many organizations' technology stacks.
- **Platforms Worked With:** The word cloud visualization indicated a variety of platforms being used by respondents, with popular choices including Windows, Linux, and AWS. This reflects the diversity in platform preferences among developers.
- **Top 10 Web Frameworks Worked With:** The hierarchy bubble chart showed that frameworks like React, Angular, and Django were among the top web frameworks utilized by the respondents, highlighting their importance in modern web development.

RESULTS

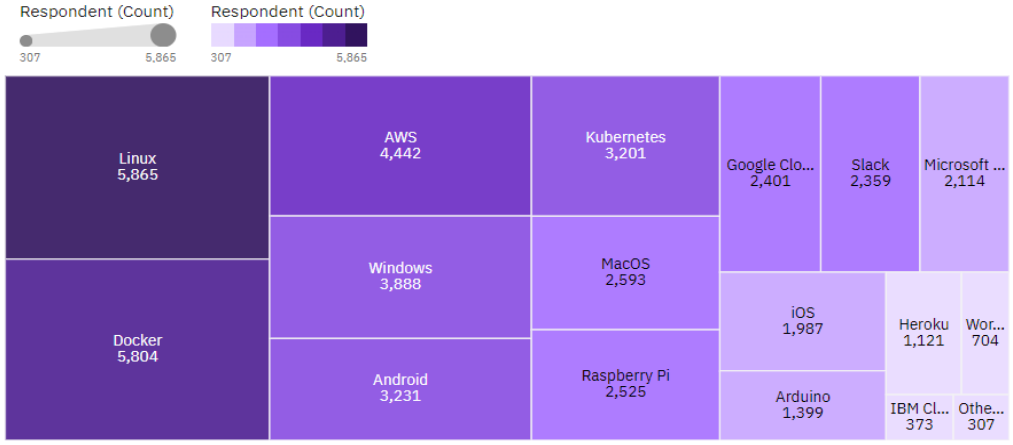
BAR CHART SHOWING THE TOP 10 PROGRAMMING LANGUAGES RESPONDENTS DESIRE TO WORK WITH NEXT YEAR.



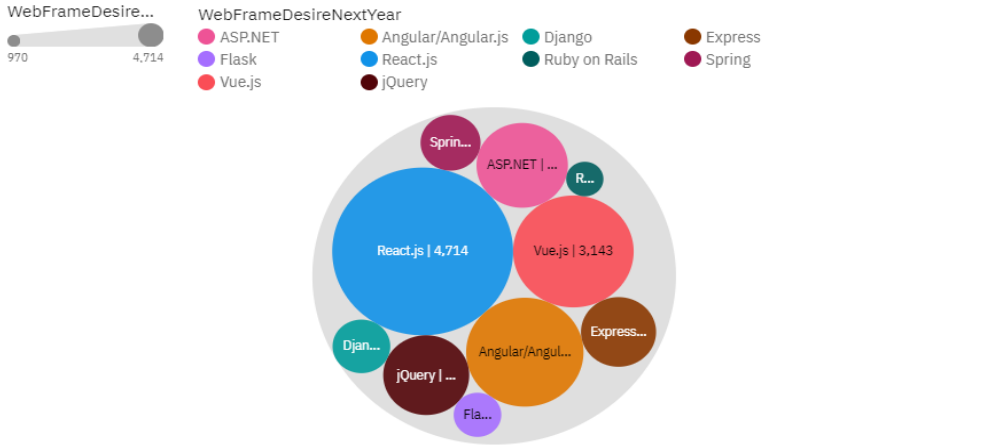
COLUMN CHART SHOWING THE TOP 10 DATABASES RESPONDENTS DESIRE TO WORK WITH NEXT YEAR.



TREE MAP CHART SHOWING THE PLATFORMS RESPONDENTS DESIRE TO WORK WITH NEXT YEAR.



HIERARCHICAL BUBBLE CHART SHOWING THE TOP 10 WEBFRAMES RESPONDENTS DESIRE TO WORK WITH NEXT YEAR.



IBM Developer

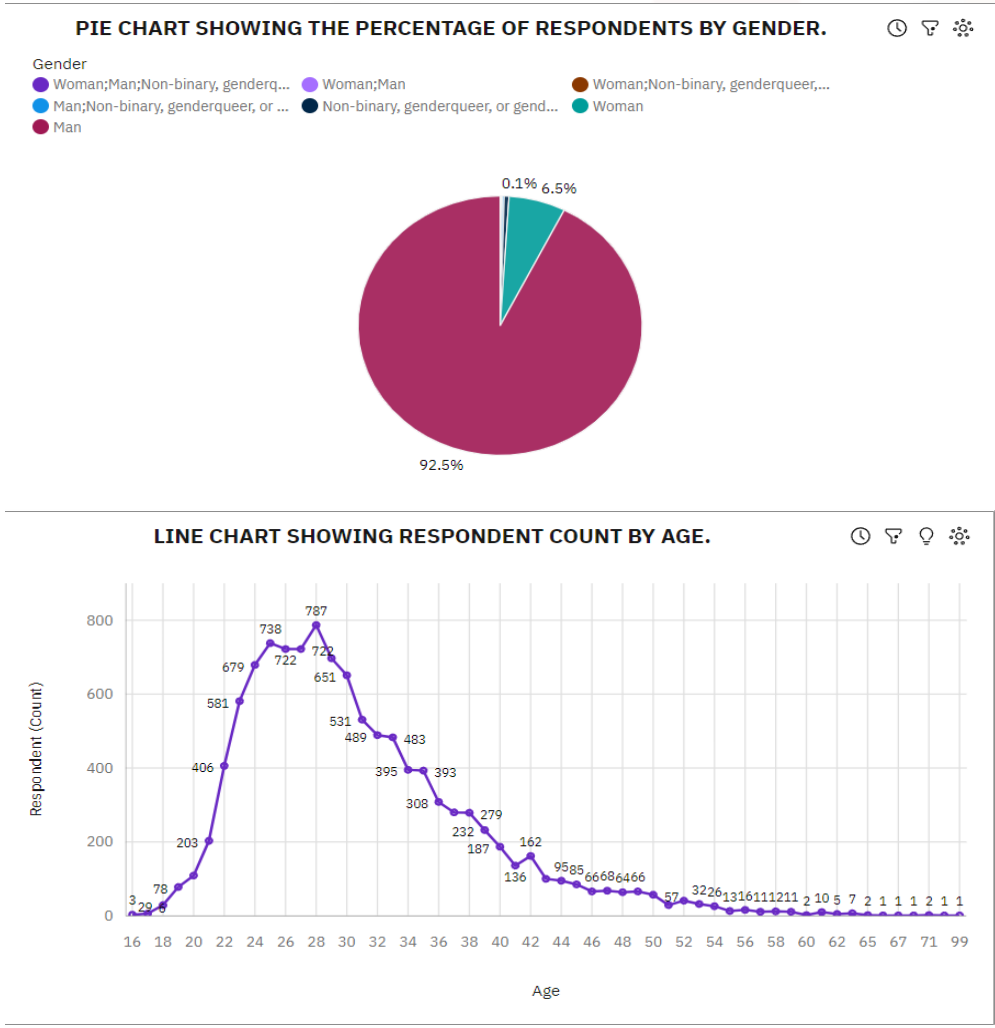
SKILLS NETWORK

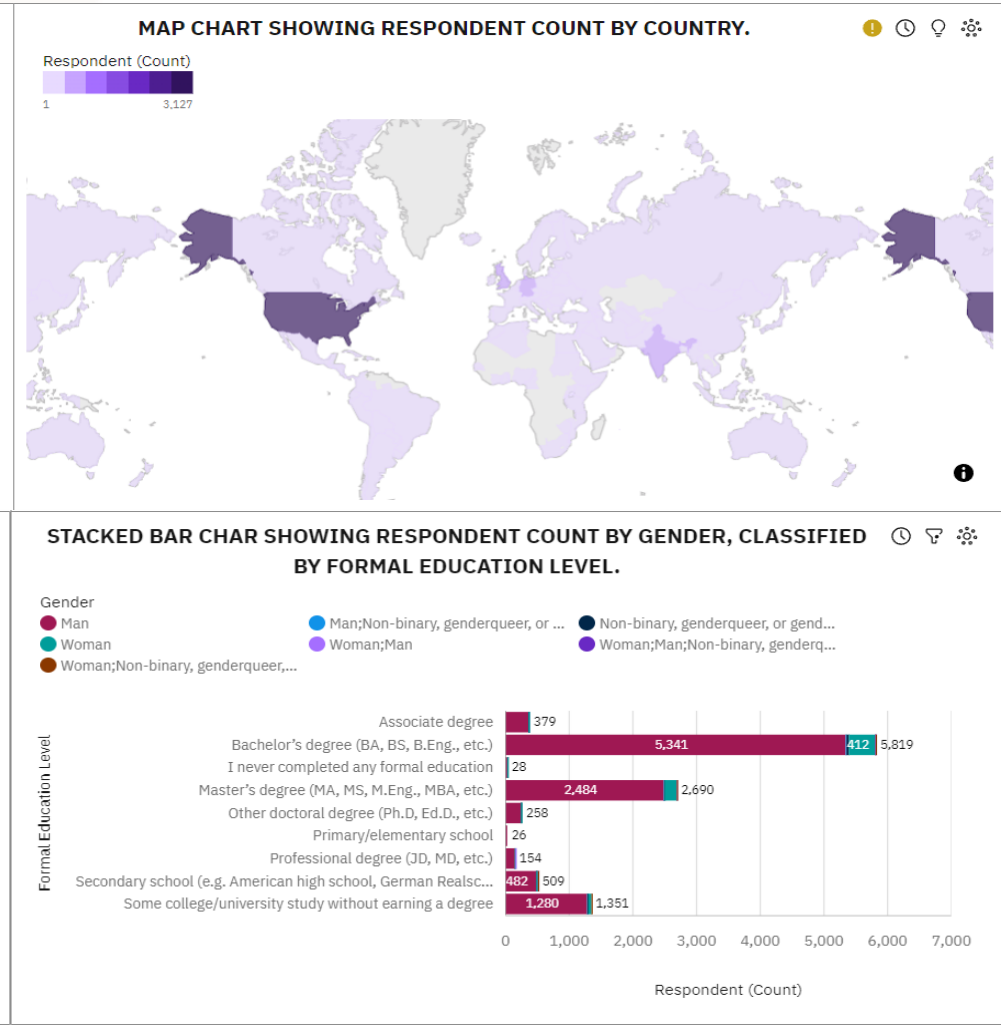
RESULTS

Future Technology Trends

- **Top 10 Desired Languages for Next Year:** The bar chart revealed that programming languages such as Python, TypeScript, and Go were among the TOP 10 desired for learning or adoption in the upcoming year. This suggests a growing interest in these languages, likely due to their increasing relevance in emerging technologies.
- **Top 10 Desired Databases for Next Year:** The column chart indicated that respondents are looking to explore databases like MongoDB, Redis, and PostgreSQL in the future, which may reflect trends toward NoSQL and in-memory databases.
- **Desired Platforms:** The Tree map chart identified the platforms respondents are interested in adopting in the coming year, with Linux, Docker, Windows, and cloud-based platforms seeing significant interest. This is indicative of the broader shift towards cloud computing and open-source environments.
- **Top 10 Desired Web Frameworks for Next Year:** The hierarchy bubble chart showed respondents are keen on exploring frameworks like React.js, Vue.js, Angular/Angular.js, ASP.NET, and jQuery, indicating a trend towards more versatile and performance-optimized web development tools.

RESULTS


PIE CHART SHOWING THE PERCENTAGE OF RESPONDENTS BY GENDER.
Gender
Woman;Man;Non-binary, genderq...
Woman;Man
Woman;Non-binary, genderqueer,...
Man;Non-binary, genderqueer, or ...
Non-binary, genderqueer, or gend...
Woman
Man
0.1%
6.5%
92.5%


MAP CHART SHOWING RESPONDENT COUNT BY COUNTRY.
Respondent (Count)
1
3,127

LINE CHART SHOWING RESPONDENT COUNT BY AGE.
Respondent (Count)
800
600
400
200
0
3 29 78 203 406 581 679 738 722 787 722 651 531 489 483 395 393 308 279 232 187 162 136 95 85 66 68 64 66 57 32 26 13 16 11 12 11 2 10 5 7 2 1 1 1 2 1 1
16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56 58 60 62 65 67 71 99
Age

STACKED BAR CHAR SHOWING RESPONDENT COUNT BY GENDER, CLASSIFIED BY FORMAL EDUCATION LEVEL.
Gender
Man
Man;Non-binary, genderqueer, or ...
Non-binary, genderqueer, or gend...
Woman
Woman;Man
Woman;Man;Non-binary, genderq...
Woman;Non-binary, genderqueer,...
Formal Education Level
Associate degree 379
Bachelor's degree (BA, BS, B.Eng., etc.) 5,341 412 5,819
I never completed any formal education 28
Master's degree (MA, MS, M.Eng., MBA, etc.) 2,484 2,690
Other doctoral degree (Ph.D, Ed.D., etc.) 258
Primary/elementary school 26
Professional degree (JD, MD, etc.) 154
Secondary school (e.g. American high school, German Realsc... 482 509
Some college/university study without earning a degree 1,280 1,351
0 1,000 2,000 3,000 4,000 5,000 6,000 7,000
Respondent (Count)

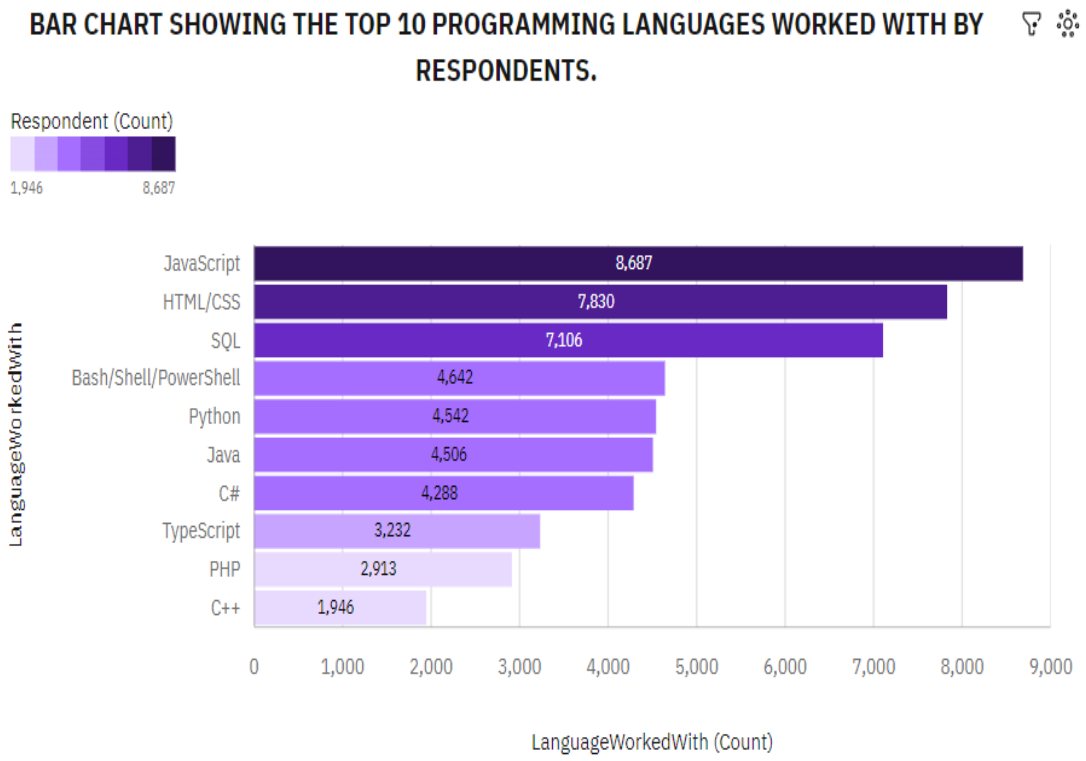
RESULTS

Demographics

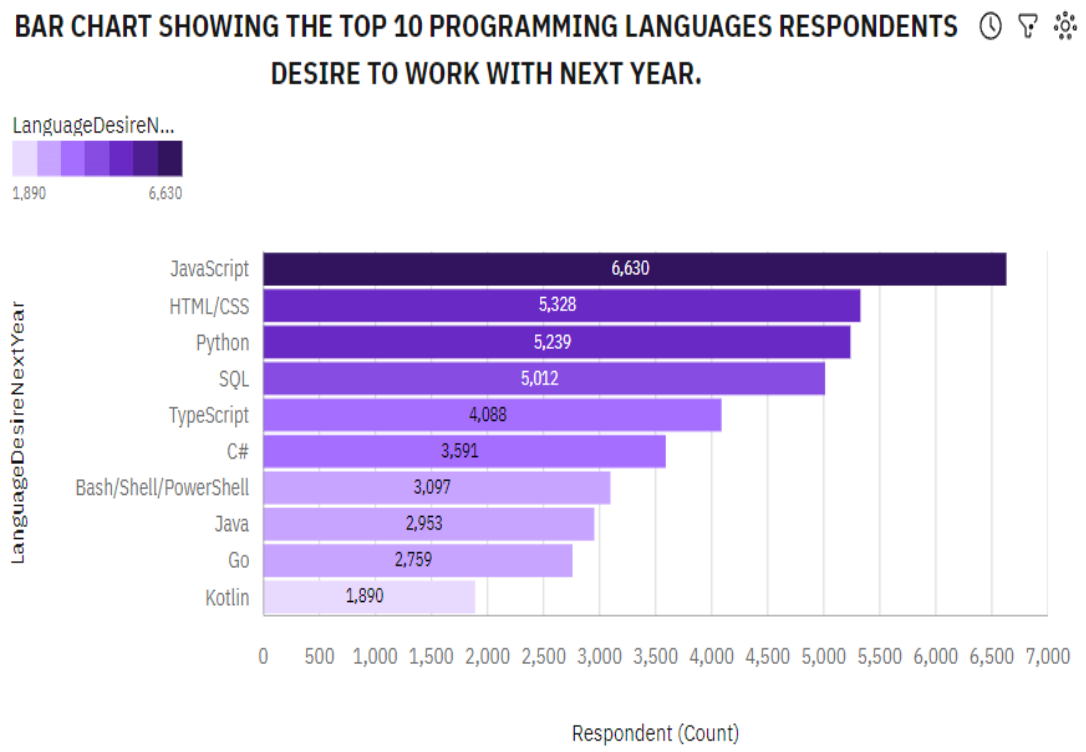
- **Gender Distribution:** The pie chart demonstrated a significant gender gap, with a higher proportion of respondents identifying as male compared to females and other gender categories. This reflects the ongoing gender disparity in the technology sector.
- **Geographic Distribution:** The map chart provided insights into the global distribution of respondents, with significant representation from countries like the United States, India, United Kingdom, and Germany. This indicates the widespread participation and interest in technology from all parts of the world.
- **Age Distribution:** The line chart showed a diverse age range among respondents, with the majority falling between the age bracket of 23 and 35. This suggests that the technology sector is predominantly composed of young professionals.
- **Education Level by Gender:** The stacked bar chart highlighted the educational qualifications of respondents, segmented by gender. A Large proportion of the respondents held at least a bachelor's degree, with a notable portion having completed a postgraduate degree, reflecting the high educational standards in the industry.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

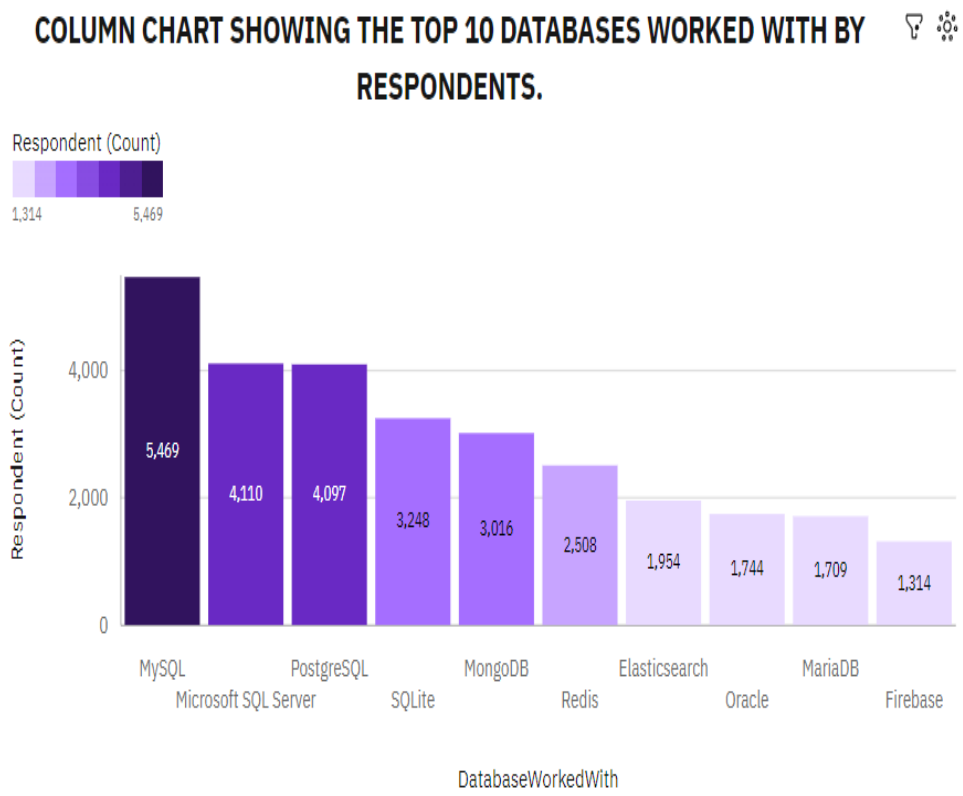
- JavaScript remains the most popular programming language in the current year, with the highest number of respondents (8,687). It continues to be the most desired language for the next year (6,630), indicating its sustained dominance in the industry.
- HTML/CSS and SQL are the second and third most worked-with languages this year, with 7,830 and 7,106 respondents, respectively. They remain highly desired for next year, showing their critical role in web development and data management.
- Python, despite being slightly less used this year compared to HTML/CSS and SQL, shows a significant increase in demand for the next year. This suggests a growing trend toward Python for its versatility and widespread use in fields like data science and machine learning.

Implications

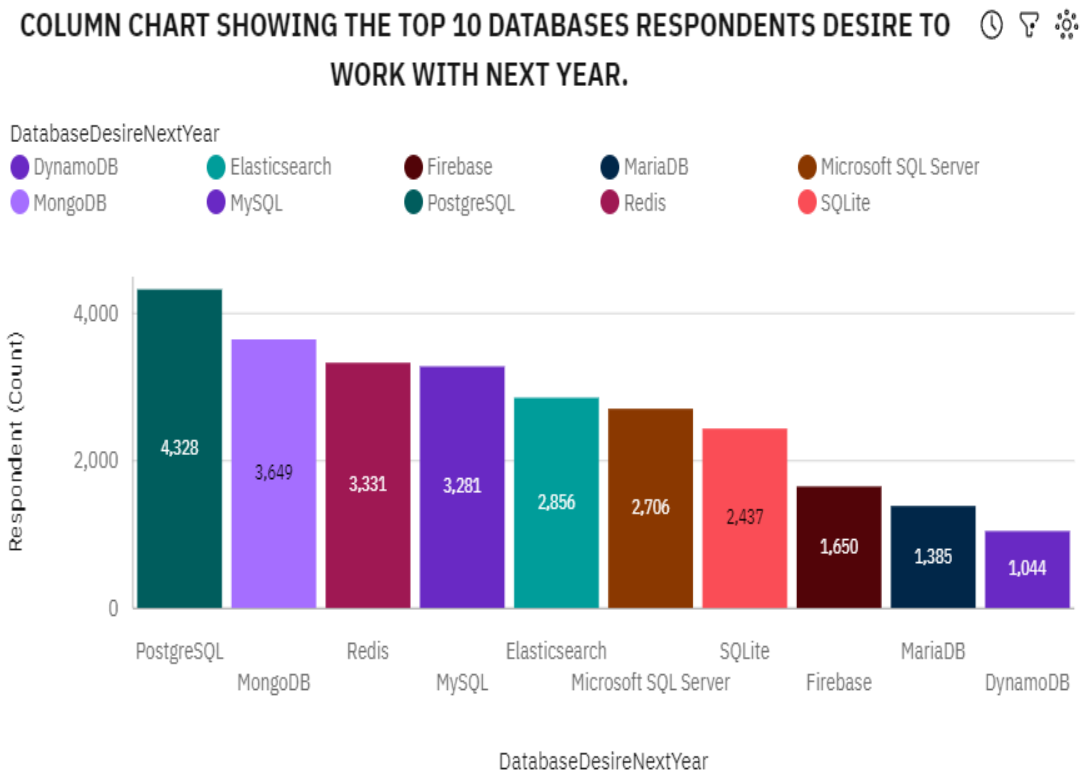
- The consistent popularity of JavaScript implies that developers and companies will likely continue prioritizing JavaScript in their projects, ensuring its relevance in both front-end and full-stack development.
- The strong demand for HTML/CSS and SQL indicates that foundational web development and database management skills will remain crucial for developers. This trend suggests ongoing investment in web technologies and the importance of mastering these languages for career growth.
- The rising interest in Python highlights the increasing importance of this language skill in data analytics, DevOps, data science, machine learning, and automation. As more companies adopt Python, there will likely be more job opportunities and demand for professionals with expertise in this language.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- In the current year, MySQL is the most used database for work with 5,469 respondents. However, the most desired database in the coming year is PostgreSQL with a total of 4,328 respondents showing a slight increase in the number of respondents by 5.64% who desire to work with PostgreSQL database.
- MySQL, PostgreSQL, and Microsoft SQL Server consistently rank among the top six most popular databases for both the current and upcoming year, underscoring their continued dominance in the industry.
- The desire to work with MongoDB is expected to rise by 20.99% in the coming year, with the number of respondents currently using MongoDB increasing from 3,016 to 3,649.

Implications

- **Shifting Preferences:** The increased desire to work with PostgreSQL in the coming year suggests that organizations and developers may start prioritizing PostgreSQL over MySQL for new projects, potentially leading to changes in resource allocation, training, and support for PostgreSQL within the industry.
- **Industry Stability:** The consistent ranking of MySQL, PostgreSQL, and Microsoft SQL Server among the top six databases highlights the stability of these platforms in the industry. Companies can confidently continue investing in these technologies, knowing they will likely remain relevant and well-supported.
- **Growing Adoption of MongoDB:** The significant anticipated increase in the desire to work with MongoDB indicates a growing trend toward NoSQL databases. Organizations may need to consider integrating MongoDB into their database management strategies, providing more support and training for their teams to adapt to this shift.

DASHBOARD

The link to the GitHub page for the project is as follows:

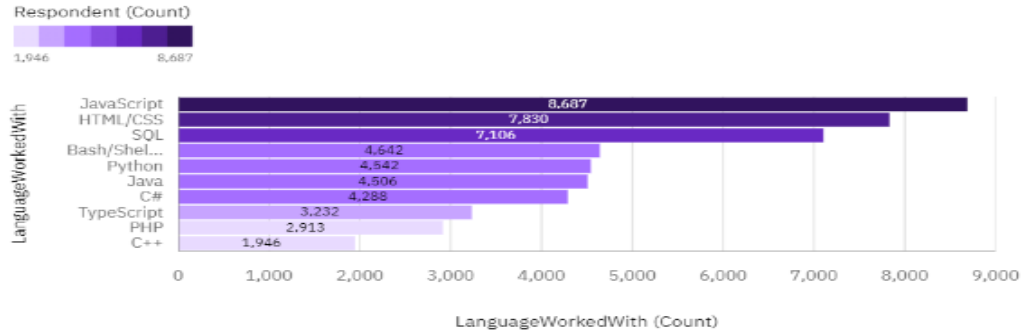
https://github.com/ISakpoba/Final-Coursera-Project/blob/main/FINAL%20COURSERA%20PROJECT%20WITH%20IBM%20COGNOS%20ANALYTICS.pdf

DASHBOARD TAB 1

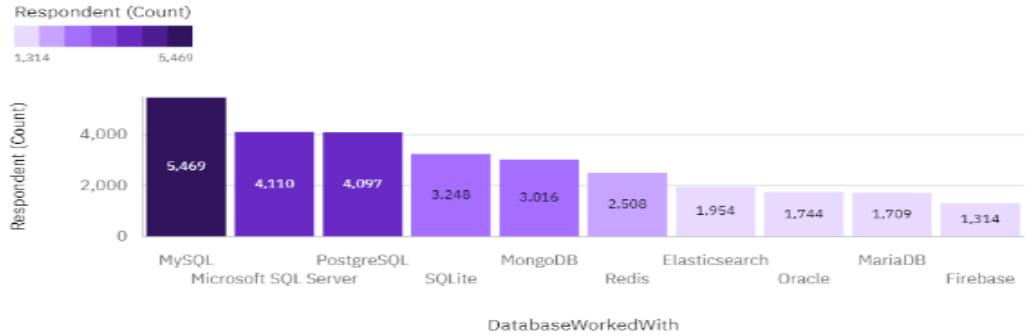
8/9/24, 8:10 PM FINAL COURSERA PROJECT

Current Technology Usage

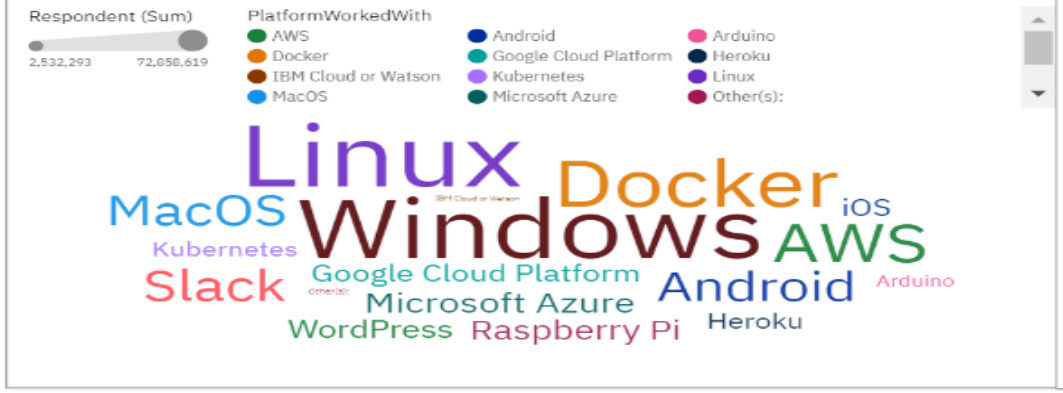
BAR CHART SHOWING THE TOP 10 PROGRAMMING LANGUAGES WORKED WITH BY RESPONDENTS.



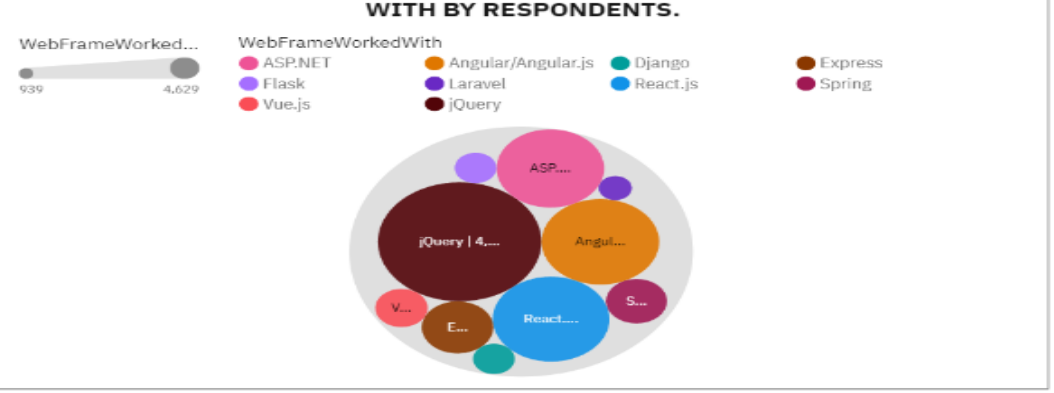
COLUMN CHART SHOWING THE TOP 10 DATABASES WORKED WITH BY RESPONDENTS.



WORD CLOUD SHOWING THE PLATFORMS WORKED WITH BY RESPONDENTS.



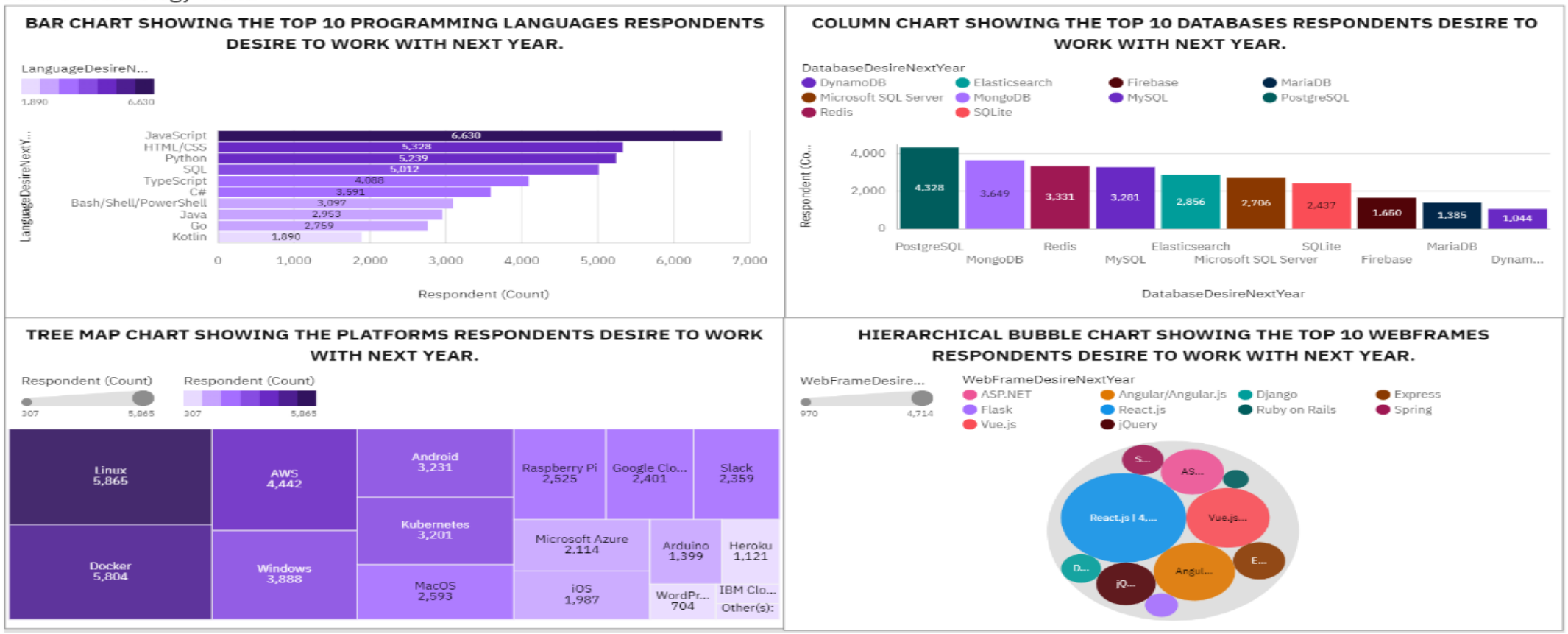
HIERARCHICAL BUBBLE CHART SHOWING THE TOP 10 WEBFRAMES WORKED WITH BY RESPONDENTS.



DASHBOARD TAB 2

8/9/24, 8:10 PM FINAL COURSERA PROJECT

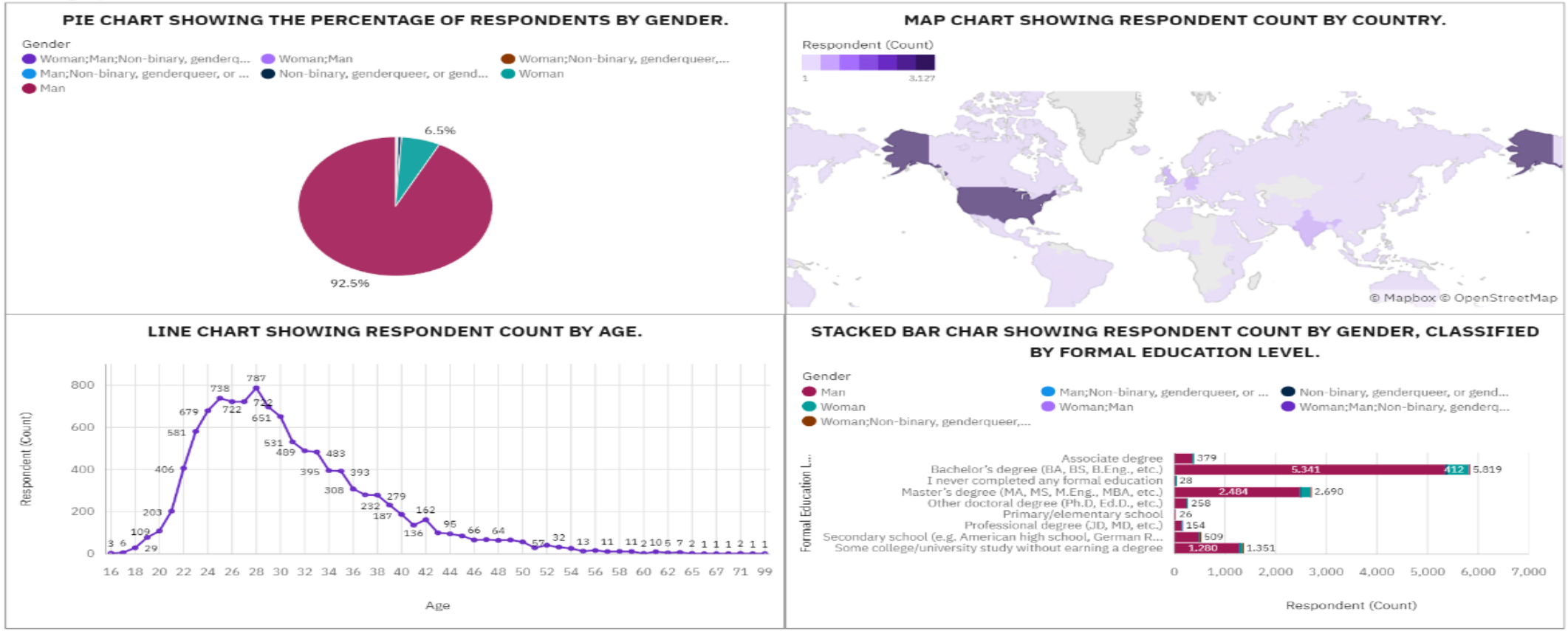
Future Technology Trend



DASHBOARD TAB 3

8/9/24, 8:10 PM FINAL COURSERA PROJECT

Demographics



DISCUSSION

OVERALL FINDINGS & IMPLICATIONS

Findings

- **Finding 1:** The most commonly used programming languages among the survey respondents are JavaScript, HTML/CSS, and SQL, while the predominantly used databases are MySQL, Microsoft SQL Server, and PostgreSQL, indicating their widespread adoption in the industry.
- **Finding 2:** The respondents plan to learn or adopt in the coming year the following top three programming languages JavaScript, HTML/CSS, and Python. They also desire to work with PostgreSQL, MongoDB, and Redis. This highlights the trends towards modern, high-performance technologies.
- **Finding 3:** The demographic analysis reveals a significant gender disparity in the technology sector, with a predominant male representation, and a concentration of respondents in the 22-35 age range, reflecting the youthfulness of the industry.

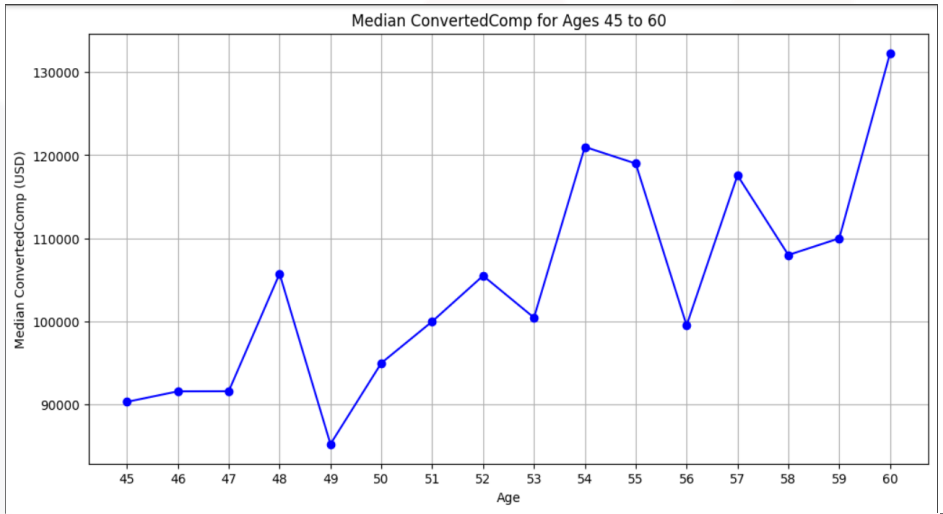
Implications

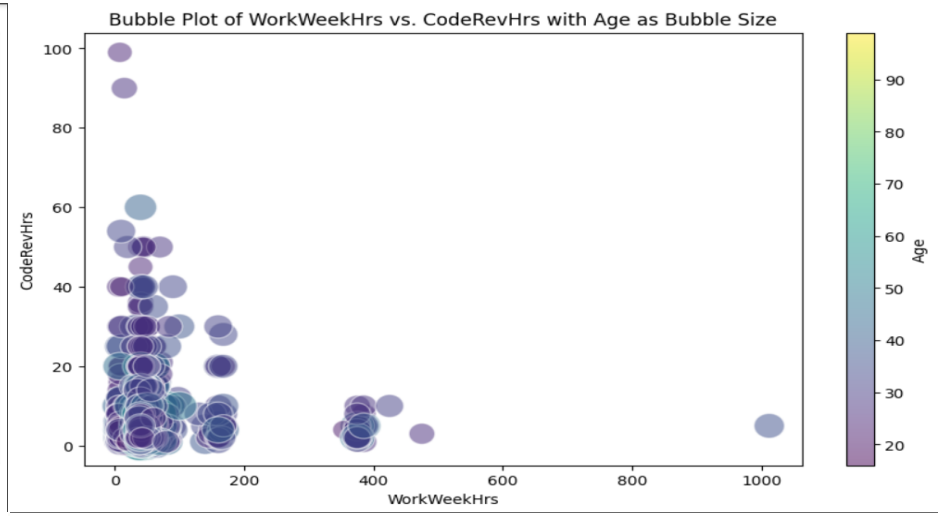
- **Implication 1:** The dominance of JavaScript, HTML/CSS. SQL and Python programming languages suggest that educational institutions and other training programs should focus on these languages to align with industry demands and improve employability.
- **Implication 2:** The growing interest in JavaScript, HTML/CSS, Python, PostgreSQL, MongoDB, and Redis programming languages and databases indicates that companies should invest in these technologies to stay competitive and meet the evolving needs of the tech workforce.
- **Implication 3:** The gender disparity in the technology sector underscores the need for initiatives to promote diversity and inclusion, particularly in encouraging more women to pursue careers in technology.

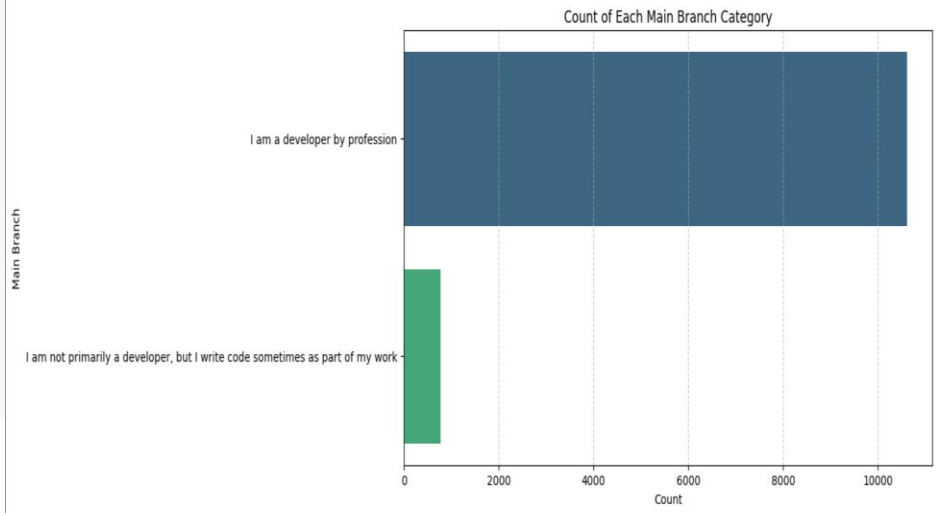
CONCLUSION

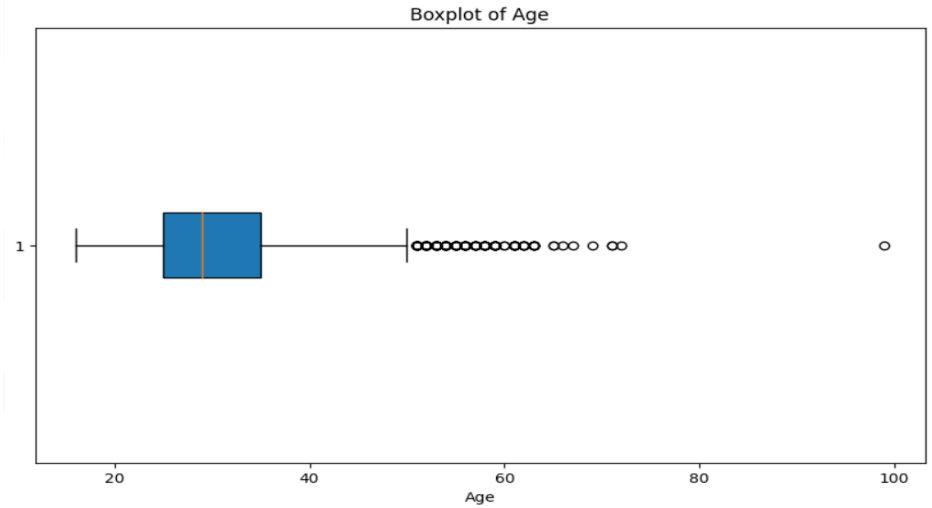
- The project successfully identified the current trends in technology usage, highlighting key programming languages, databases, platforms, and frameworks that are prevalent in the industry.
- The analysis of future technology trends provided insights into the technologies that professionals are most eager to learn, indicating a shift towards more modern and scalable solutions.
- The demographic data revealed significant patterns in age and gender distribution within the tech industry, pointing out areas where diversity and inclusion efforts are needed.
- Overall, the findings of this project offer valuable guidance for educators, employers, and policymakers in shaping the future of technology education, workplace practices, and industry development.

APPENDIX

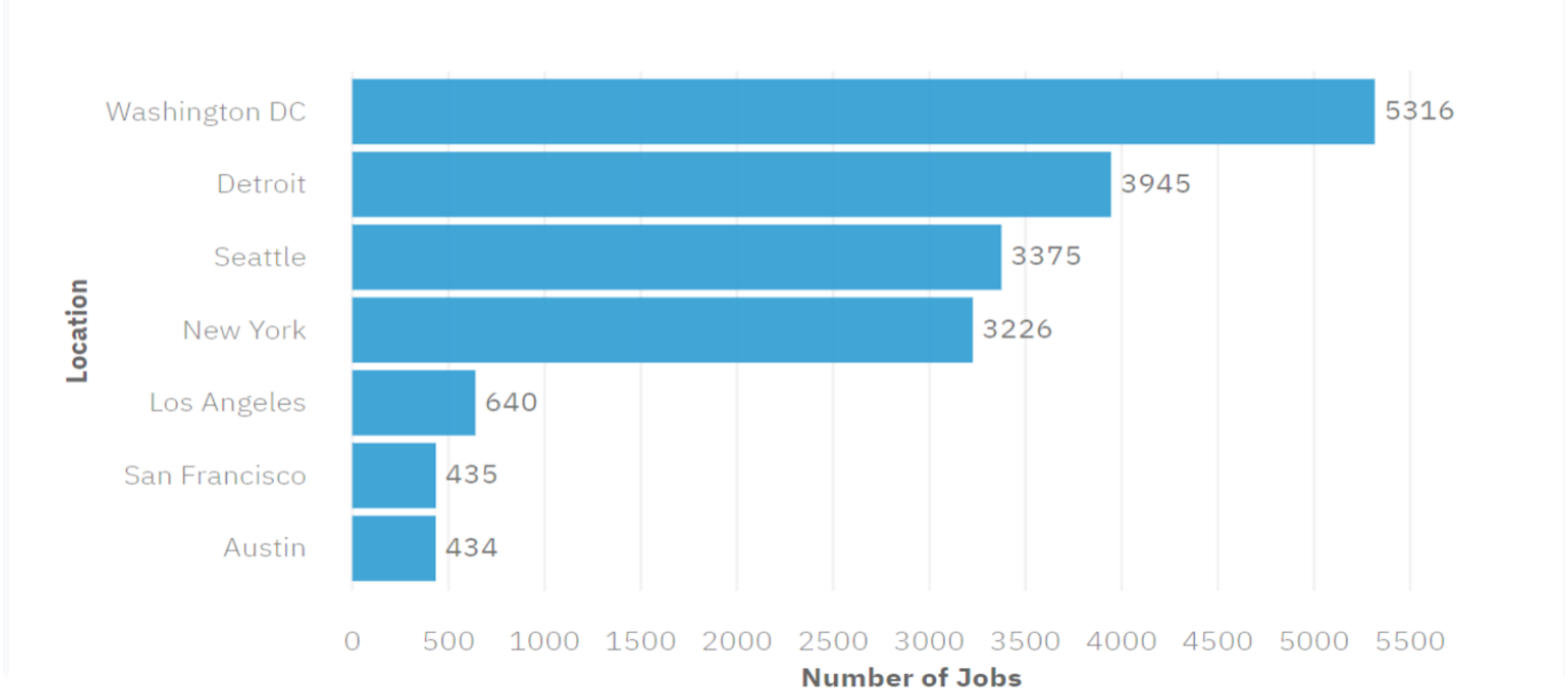

Median ConvertedComp for Ages 45 to 60
Median ConvertedComp (USD)
Age


Bubble Plot of WorkWeekHrs vs. CodeRevHrs with Age as Bubble Size
CodeRevHrs
WorkWeekHrs
Age

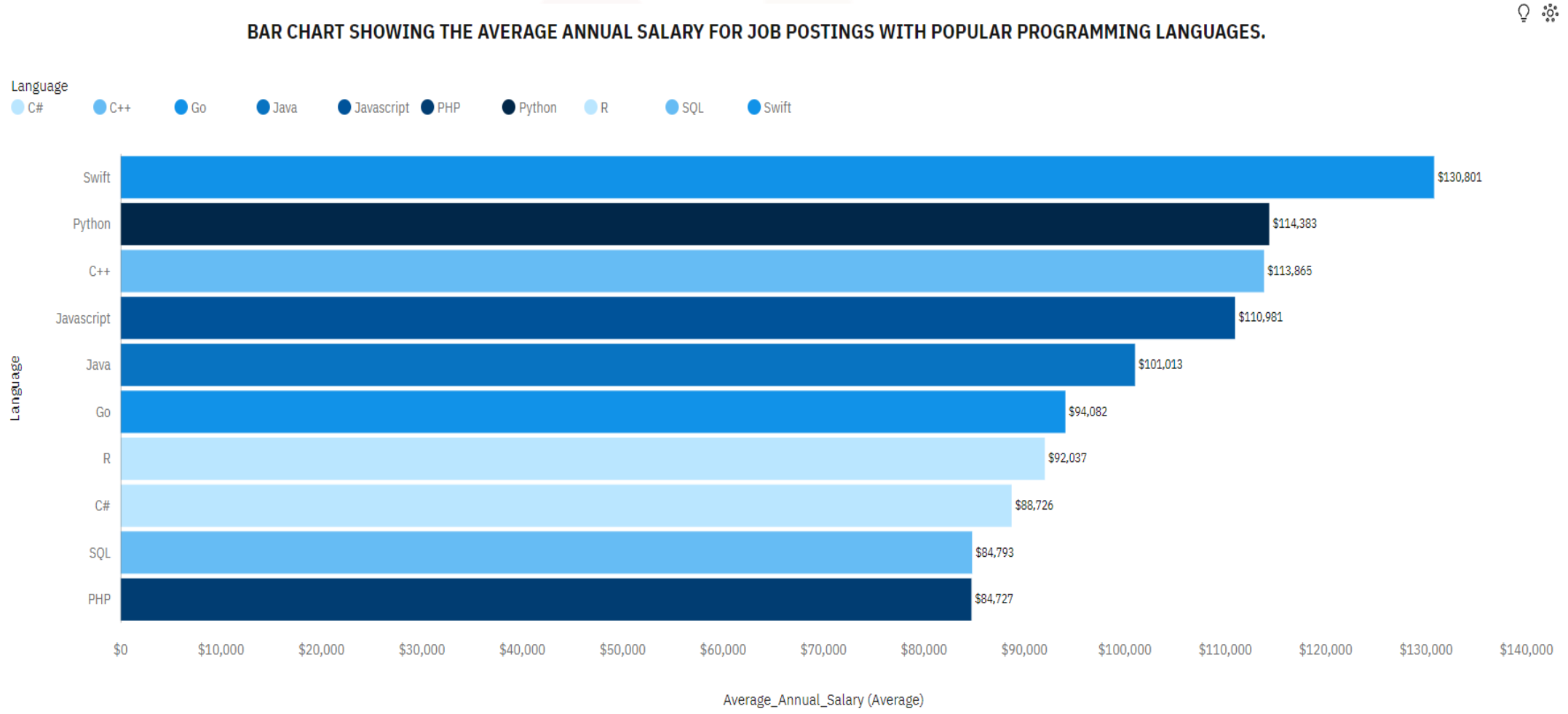

Count of Each Main Branch Category
Main Branch
I am a developer by profession
I am not primarily a developer, but I write code sometimes as part of my work
Count


Boxplot of Age
Age

JOB POSTINGS



POPULAR LANGUAGES




THANK YOU