Introduction

I was hired by Eduvos to analyze survey data to apprehend the technological tools that their past graduates are presently using in the work environment. I was provided with a comprehensive survey dataset file that presented all there is to know about the popular tools within the technology field. The objective is to identify the most popular programming languages, databases, web frameworks and other tools that past students are currently using. In doing so, Eduvos seek to updates its IT courses to ensure that they align to the industry trends and provide students with relevant skills that suit the job market .

Data Overview

The dataset provided consists of various responses from Eduvos IT graduates, storing all types of information. To ensure I could attain the relevant information to present my dashboard to the stakeholders at Eduvos, I needed to perform all the cleaning and preprocessing steps needed to get the dataset into a more practical format.

Cleaning and Preprocessing steps:

i. The first step was to select only the essential columns mentioned in the scenario; these key elements include:

- **Campus:** The campus at which the past student graduated
- **StudyField:** Field of study (IT, Computer Science or Data Science)
- **Branch:** Graduates' type of work
- **Role:** Graduates' title
- **EduLevel:** Highest level of education
- **ProgLang:** Relevant programming languages the graduate uses
- **Databases:** Relevant databases languages the graduate uses
- **Platform:** Relevant cloud platform the graduate uses
- **WebFramework:** Relevant web framework the graduate uses
- **Industry:** The industry the graduates work in
- **AISearch:** Relevant AI search tool the graduate uses (e.g. ChatGPT)
- **AITool:** Relevant AI developer tool the graduate uses (e.g. GitHub)
- **Employment:** Whether the graduate was employed or not

ii. The next step was to ensure that all missing values within the newly selected columns are dealt with correctly. Removing all items with missing values in any column was the only appropriate course of action. This was accomplished by first converting all blank values to NA and then using the omit function to eliminate the entry entirely. By doing this, it improves dataset integrity, streamlines data handling, and increases performance.

iii. Standardizing the categorical columns was the third phase in the dataset's cleaning and preprocessing process. I standardized it to just "Durban Campus" if the entry contained "Umhlanga Campus" or "Durban Campus" because it was known that Durban/Umhlanga was a problem in this situation under the campus column. Both the Mbombela Campus/Nelspruit Campus and the Nelson Mandela Bay Campus/Port Elizabeth Campus experienced the same problems following some manual searches. Therefore, I also standardized these to the Mbombela Campus and Nelson Mandela Bay Campus, respectively.

iv. The dataset was then limited to only include the three to five campuses with the highest number of responses. The top 5 campuses were chosen since they would provide me with more data and material to work with for the visualization. The top 5 campuses with the most entries were identified using the count function.

v. The next stage was to transform character columns to factors in order to increase the dataset's preparedness. This would enhance data integrity and facilitate data ordering and ggplot2 plotting.

vi. Disregarding any duplicate entries was the last stage in the cleaning and preprocessing procedure. By doing this, it prevents redundancy, improves efficiency, and eliminates misleading insights.

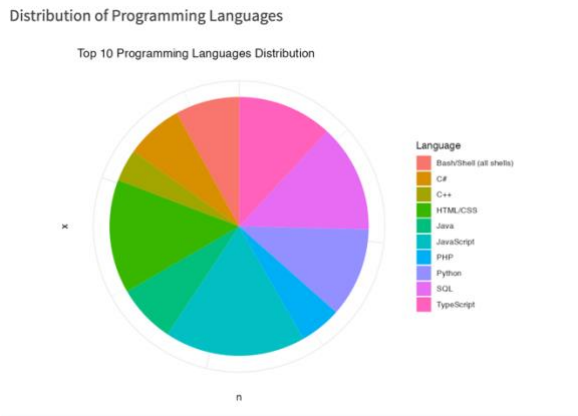
Insights and Analysis

To configure suitable insights and analysis, the best method was to visualize the key correlations and trends within each variable. The major factors included:

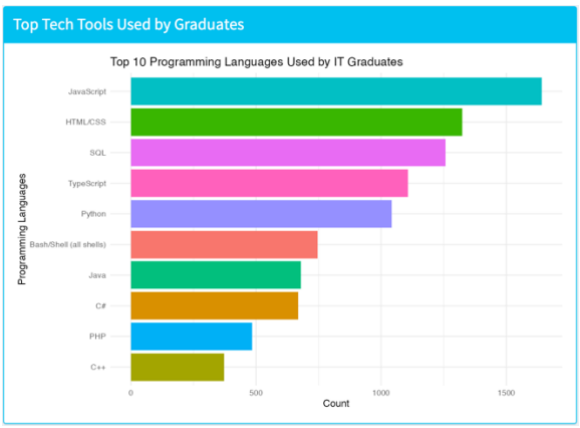
- **Programming Languages:** Discuss popular trends
- **Databases:** Identify trendy databases and usage forms
- **Web Frameworks:** Analyse commonly used frameworks and relationships
- **Cloud Platforms:** Provide updates on prominent cloud platforms and trends
- **AI Tools:** Discuss popular AI tools and their influence

Programming Languages

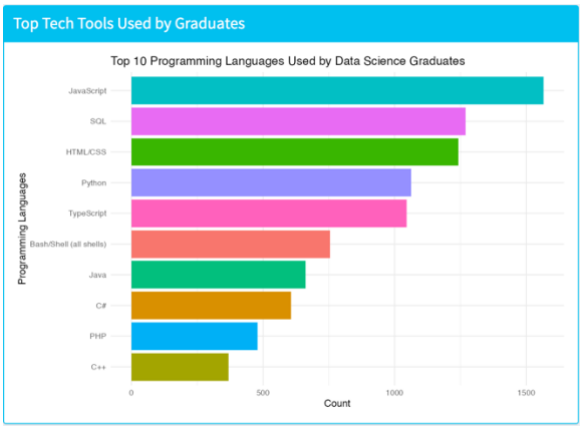
Programming language trends are one of the most significant variables to analyse. The visualization technique was created in such a manner that the user may filter the popular programming languages based on the Study Field. As a result, it is evident which programming languages are most popular in Data Science, Computer Science, and IT. A general comparison was also done to see the top languages across all fields.



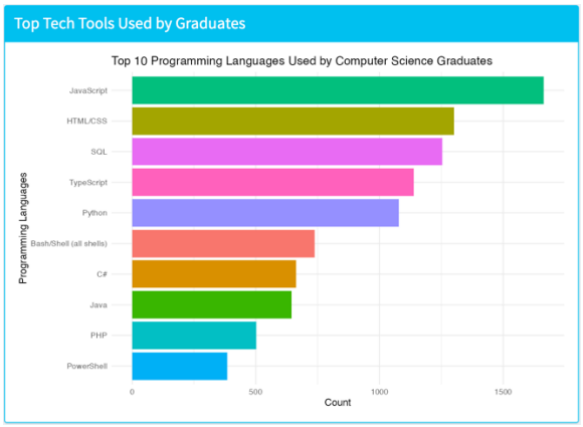
The pie chart shows the most popular programming languages among Eduvos alumni. The top three are JavaScript, HTML/CSS, and SQL. JavaScript is known for its flexibility in both front-end and back-end programming. HTML/CSS are important in web development, as is SQL, which is associated with database administration.



The following chart shows the most common programming languages among IT graduates. The top three languages are JavaScript, HTML/CSS, and SQL. The usefulness of these three languages demonstrates that most IT graduates have been exposed to a wide range of development industries. The top 10 languages are identical to that of all Eduvos graduates.

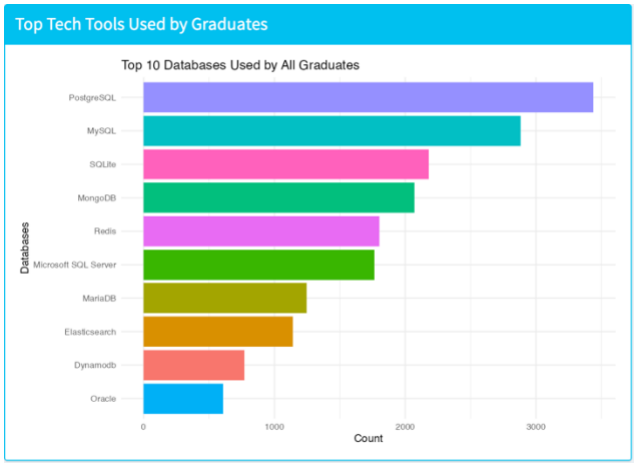


This visualization shows that data science grads choose SQL over HTML/CSS and Python over TypeScript. We can detect these developments since Python and SQL are both more sophisticated in the field of data science.



In the final filter, it is worth noting that the top runners for computer science stay the same. However, it is interesting to observe that they prefer C# over Java, and PowerShell has recently entered the top ten.

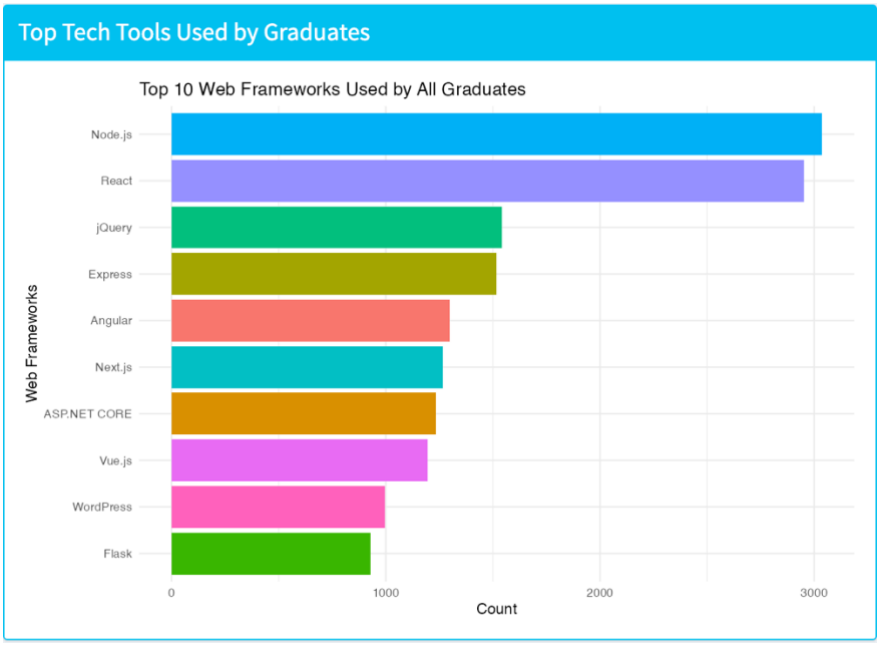
Databases



Although databases may appear to be more significant to people in the data science area, extensive research and visualization reveal that the ten most regularly used databases are the same across the IT, Data Science, and Computer Science sectors. These include MySQL, PostgreSQL, SQLite, MongoDB, and Redis. As seen here, PostgreSQL is highly common among other systems.

Web Frameworks

Web frameworks make application development easier by including components that are reusable along with organized development environments. The bar graph illustrates the top ten web frameworks across all research disciplines. Top five: Node.js, React, jQuery, Express, and Angular.

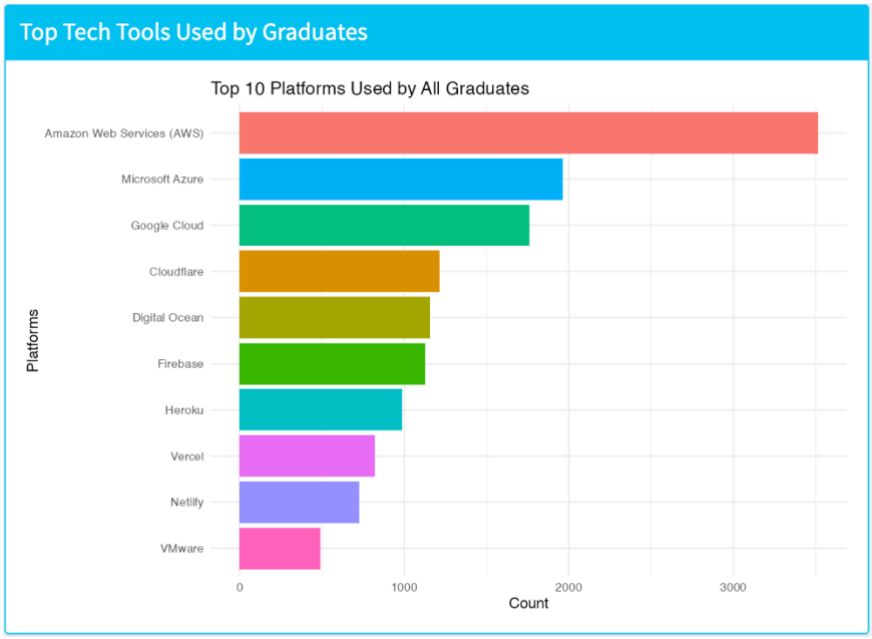


However, after using the filtering technique to determine the most popular web frameworks within each category, the top ten rankings varied somewhat:

These lists are in order from 1 to 10.

IT : Node.js, React, jQuery, Express, Next.js, ASP.NET CORE, Angular, Vue.js, ASP.NET and WordPress
Data Science: Node.js, React, Express, jQuery, Angular, Next.js, ASP.NET CORE, Vue.js, Django and Flask
Computer Science: Node.js, React, jQuery, Express, Angular, Vue.js, Next.js, ASP.NET CORE, WordPress and Flask

Cloud Platforms

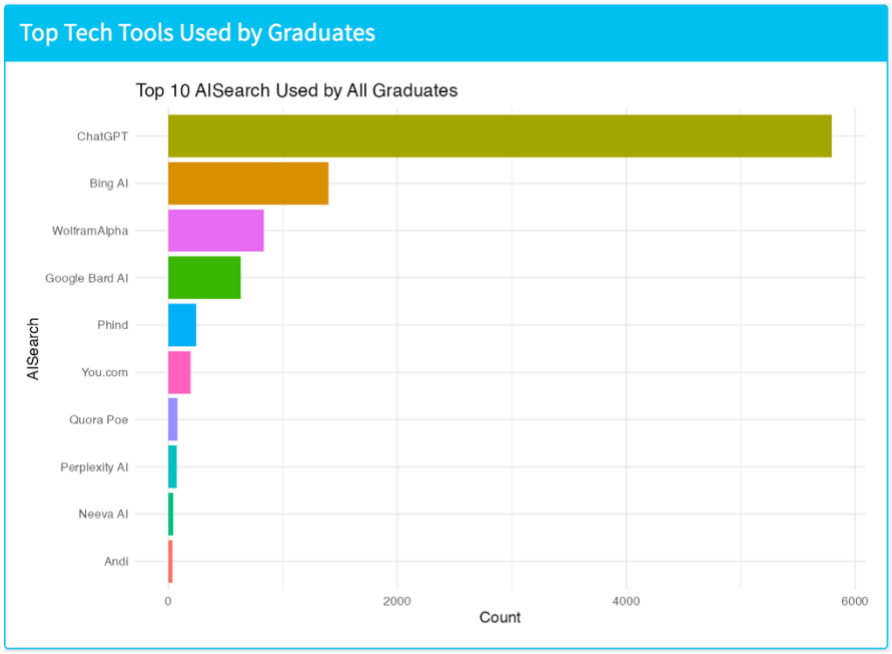


Cloud computing has become a vital part of modern IT infrastructure, allowing organizations to implement scalable and cost-effective solutions. Eduvos alumni regularly use cloud platforms like AWS, Microsoft Azure, and Google Cloud. As demonstrated in the visual, Amazon Web Services (AWS) is the most popular cloud platform among graduates.

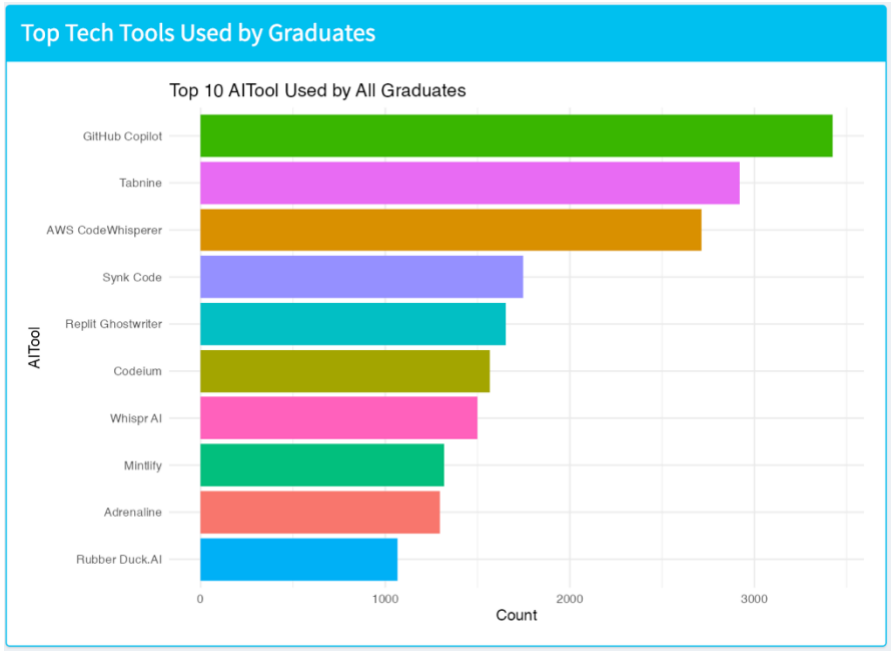
AI Tools

Artificial intelligence is becoming more incorporated into current applications. With the ever-changing technological landscape, it is safe to say that artificial intelligence is here to stay. Organisations utilize these platforms to boost automation and deep learning. The AI tools were divided into two categories: AI search tools and AI development tools.

In terms of AI search tools, the graphic shows that ChatGPT dominates this field. BingAI comes in second position, with three times less usage than ChatGPT. WolframAlpha, Google Bard AI, and Phind round out the top five results.



The bar graph depicts the top ten AI development tools utilized by Eduvos alumni. GitHub Copilot is the most widely used, followed by Tabnine, AWS CodeWhisperer, Synk Code, and Replit Ghostwriter. After using a screening method, it was determined that the top 10 most often used AI Developer tools are consistent throughout the IT, Data Science, and Computer Science sectors.



Recommendations

Curriculum Development

Eduvos' curriculum should be expanded to include practical instruction in JavaScript, HTML/CSS, and SQL, which are often utilized in the IT business. Courses should also include PostgreSQL, MySQL, and SQLite, which are widely used databases in the workplace. There are several web frameworks used in the market; Eduvos should consider including the most regularly used frameworks (Node.js, React, and jQuery) for each profession within their degree courses.

Focus Areas

As previously said, artificial intelligence (AI) plays a significant role in business and cloud computing. It would be advantageous for Eduvos to delve more into AI Developer tools, AI Search tools, and Cloud Platforms. In terms of cloud computing, Eduvos should provide practical training for platforms such as AWS, Azure, and Google Cloud. In terms of AI development tools, GitHub Copilot is an important variable that should be included in Eduvos courses. It would be advisable for Eduvos to adequately instruct students on how to utilize AI Search tools such as ChatGPT to guarantee that they make advantage of this powerful tool while reducing plagiarism.

Industry Alignment

To line with the standards of the industry, Eduvos should seek collaboration with industry organizations that can provide crucial insights into the growing sector. Providing students the opportunity to become certified professionals for platforms would be tremendously beneficial. Platforms like AWS provide courses that any individual may take to become a certified professional on the platform. Eduvos should work to establish connections with important industry players to ensure they are brought up to date on current developments.

Conclusion

The analysis focused on significant technologies utilized by Eduvos IT graduates, such as prominent programming languages, databases, web frameworks, cloud platforms, and artificial intelligence tools. The information acquired will help Eduvos improve its curriculum and better equip students for the changing employment market. Through prioritizing skills relevant to the industry and emerging technologies, Eduvos can improve employability and keep its graduates competitive in the technology industry.