Data that Speaks: What the World's Developers Are Teaching Us

Graziolo, Andrea.

23 April 2025.-

OUTLINE

1. **Executive Summary**
   Main conclusions of the analysis.
2. **Introduction**
   Context, objectives and target audience.
3. **Methodology**
   Data sources, processing and tools used.
4. **Results**
   Trends in programming languages
   Trends in databases
5. **Visualisation of Job Offers**
   Graph with labour demand data.
6. **Interactive Dashboard**
   Captures and analysis of the dashboard in Cognos Analytics.
7. **Discussion**
   Critical interpretation of the visualised results.
8. **General Findings and Implications**
   Key conclusions and their impact.
9. **Conclusion**
   Final summary and reflections.
10. **Appendix**
    Complementary visualisations and tables.

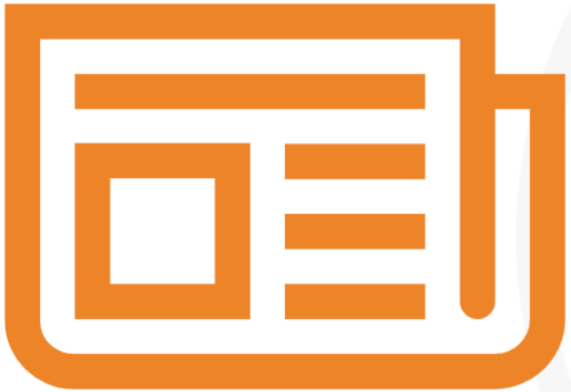
EXECUTIVE SUMMARY

- This analysis draws on information from the Stack Overflow survey, focusing on programmers' salaries, the tools they use, their job types and their level of education.
- Significant patterns were found that connect job type and education level with annual income.
  - Programmers in full-time employment tend to be paid more.
  - Those with higher-level degrees have salaries that can vary widely and are not necessarily higher.
  - Experience and knowledge of technical skills have a greater impact on pay than the level of schooling achieved.
- The most commonly used tools are JavaScript, Python and SQL.
  - The most in-demand databases were MySQL and PostgreSQL..
- Interactive images and panels were created with IBM Cognos to illustrate this clearly and simply.
- These conclusions can be used by recruiters and experts to make informed decisions about hiring and career development.

INTRODUCTION

- The project is based on survey data from Stack Overflow, including responses from developers worldwide.
- The aim is to identify patterns influencing salary, job type and grade level.
- These findings will help us to gain a better understanding of the current state of the technology market, as well as providing better guidance for personal or business decision-making.
- he data were cleaned and processed to remove nulls, missing values, and outliers.
    - Missing values were imputed into the 'ConvertedCompYearly' column.
    - Methods such as one-hot encoding and age ranges were employed to facilitate clearer analysis.

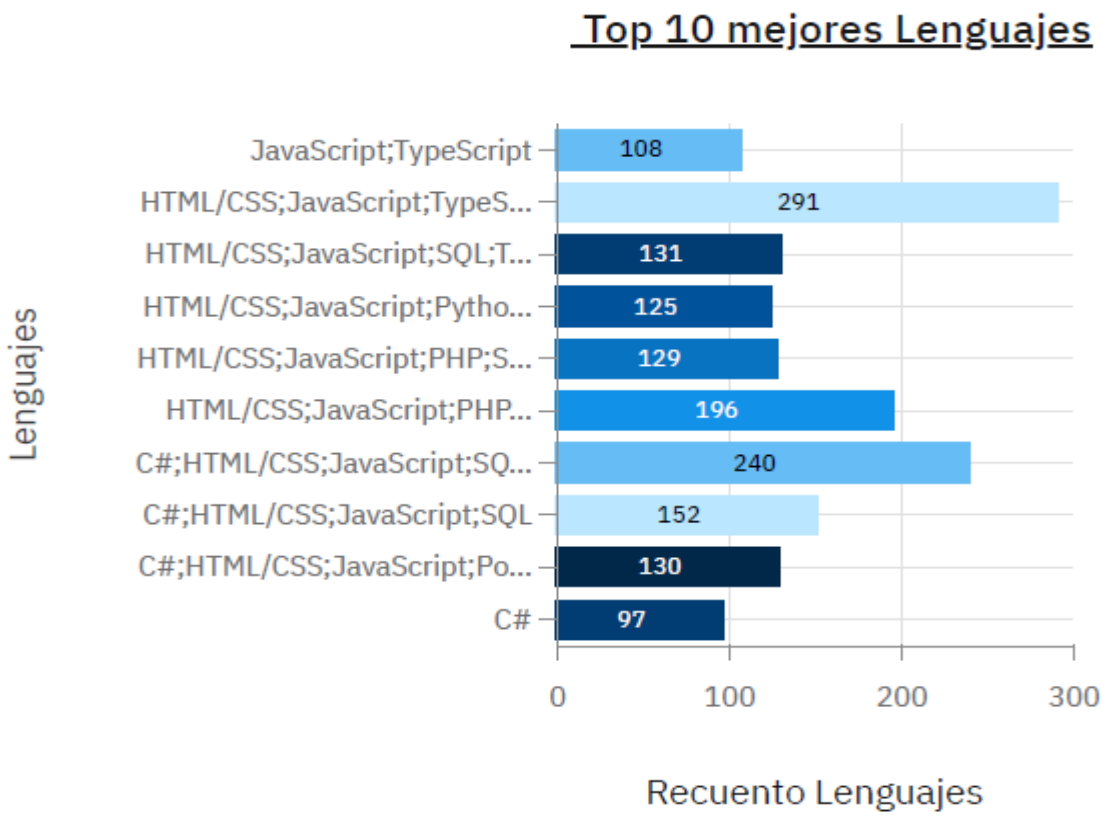
METHODOLOGY



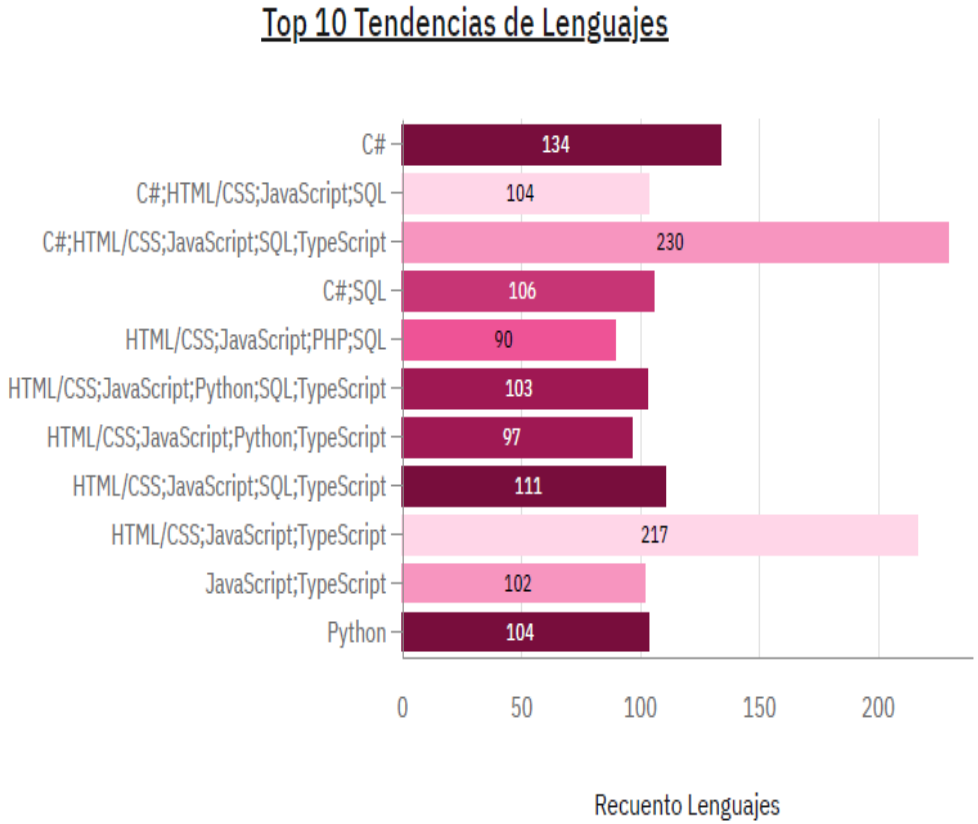
- The survey data were obtained from a reliable and representative sample of the global technology landscape..
- First, we conducted an exploratory analysis and prepared the dataset, ensuring that incomplete entries and outliers, which could have biased the results, were eliminated.
- The analysis focused primarily on annual compensation, the use of technologies, the type of employment, and the educational level of developers.
- The following steps were taken to analyse and visualise the data:
  - Detection of null values and imputation of missing values using the median, mean or frequency, as appropriate.
  - Detection and treatment of outliers using the interquartile range (IQR).
  - Grouping categories into variables such as age and education.
  - One-hot encoding of multi-response variables such as employment and dev type.
- Python (Pandas, Seaborn and Matplotlib) was used for exploratory analysis and graph creation, and IBM Cognos Analytics was used for interactive dashboard design.

PROGRAMMING LANGUAGE TRENDS

Current Year



Next Year



PROGRAMMING LANGUAGE TRENDS - FINDINGS & IMPLICATIONS

Findings

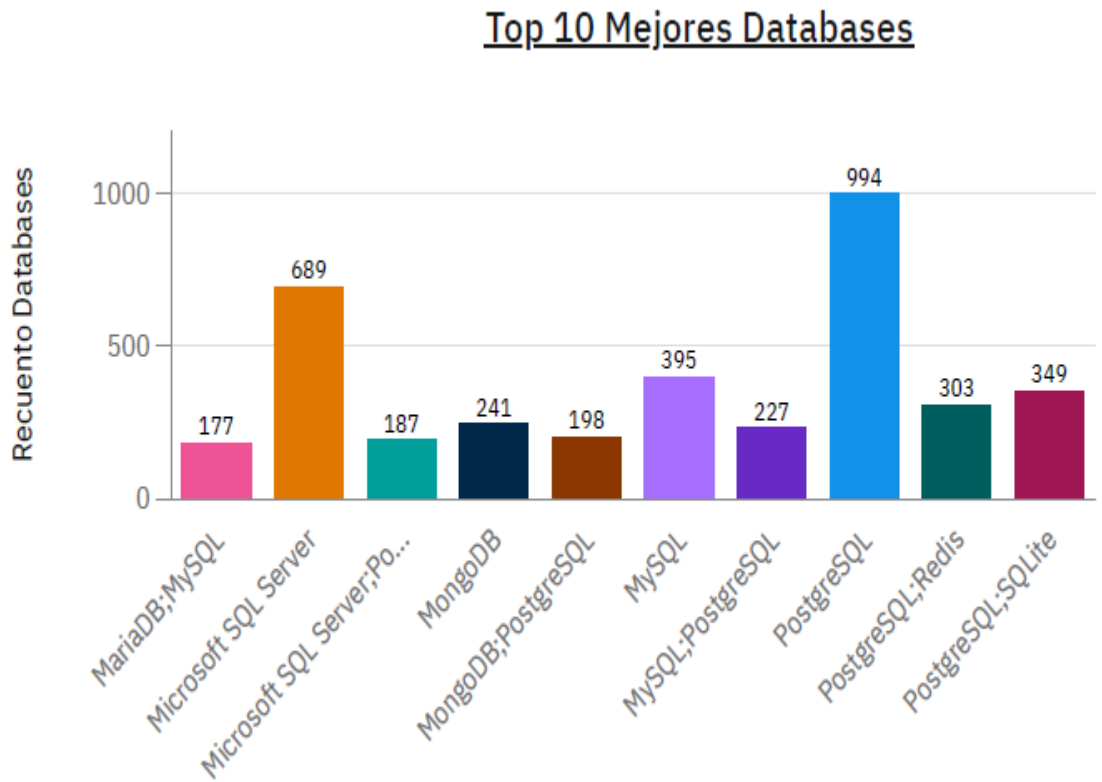
- JavaScript and TypeScript remain the most popular languages among developers.
- he popularity of the combination of HTML/CSS with JavaScript and backend languages such as Python, PHP and SQL indicates that web development remains the main focus.
- C# emerges as a strong contender, particularly in conjunction with web technologies, highlighting its significance in desktop and enterprise environments.
- Looking to the future, Python is emerging as a popular choice, reflecting its increasing use in data science and artificial intelligence.

Implications

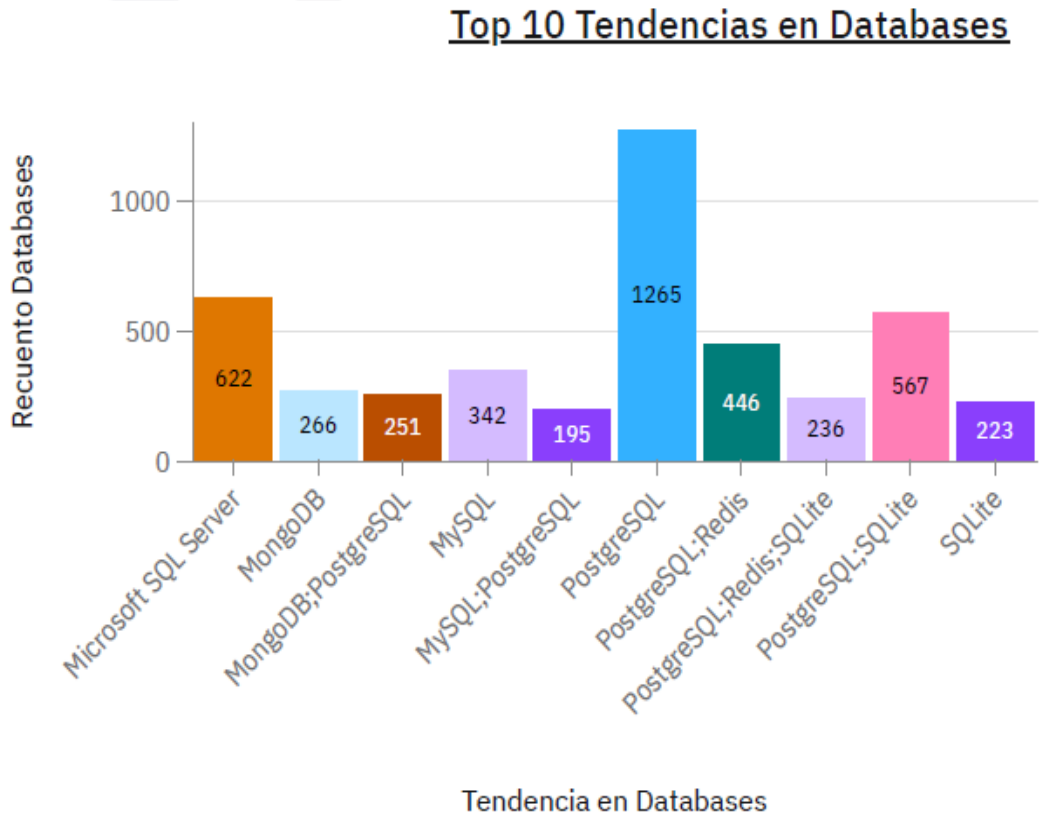
- Demand for full-stack developers continues to rise given the combined use of front-end and back-end languages.
- Languages such as Python and TypeScript provide opportunities to specialise in rapidly expanding fields such as AI, data analytics, and scalable development.
- Companies should continue to train their teams in core technologies such as JavaScript, Python and C#, which remain prevalent across multiple industries.

DATABASE TRENDS

Current Year



Next Year



DATABASE TRENDS - FINDINGS & IMPLICATIONS

Findings

- PostgreSQL is the current leader in terms of both usage and future trends, establishing itself as one of the most versatile and powerful databases.
- Meanwhile, MySQL and Microsoft SQL Server remain popular choices, particularly for enterprise applications.
- MongoDB, a NoSQL database, is increasingly present in future trends, indicating growth in projects with flexible data structures.
- The combination of relational and non-relational databases (e.g. PostgreSQL + Redis or MongoDB + PostgreSQL) is becoming increasingly common.

Implications

- The sustained popularity of relational databases indicates a continued demand for proficiency in SQL.
- Meanwhile, the rise of NoSQL databases such as MongoDB suggests an increased demand for professionals who can handle large volumes of unstructured data.
- The trend of combining multiple databases into a single solution reflects a shift towards more hybrid and specialised architectures.

DASHBOARD

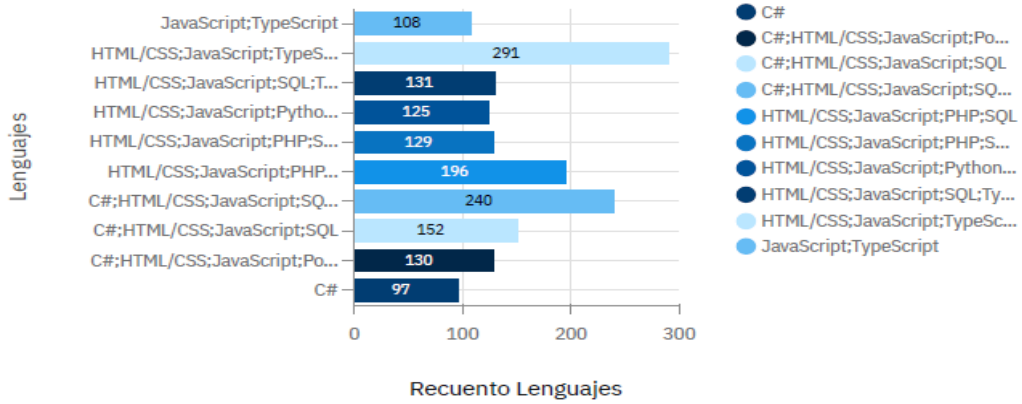
Analysis of technology use and trends.

This interactive dashboard was developed using IBM Cognos Analytics and summarises the current usage of technology, future trends, and the demographic characteristics of the surveyed developers.

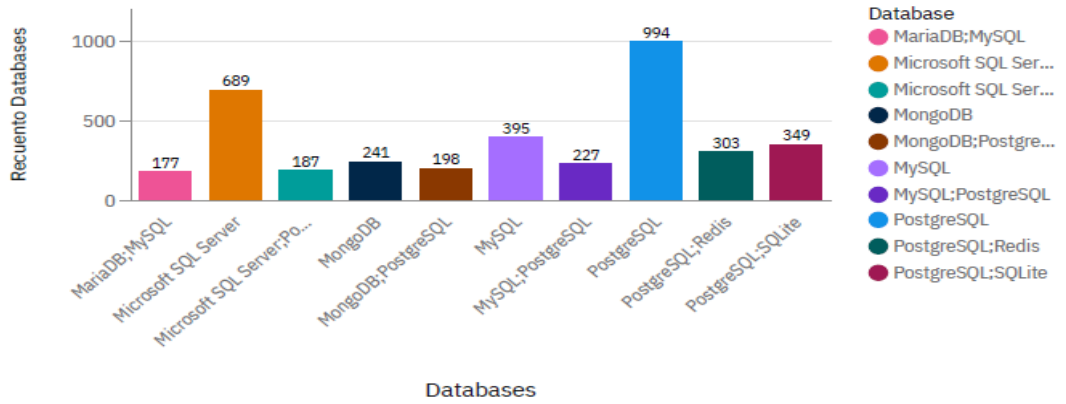
DASHBOARD TAB 1 – CURRENT STATE

Uso actual de tecnología

Top 10 mejores Lenguajes 1

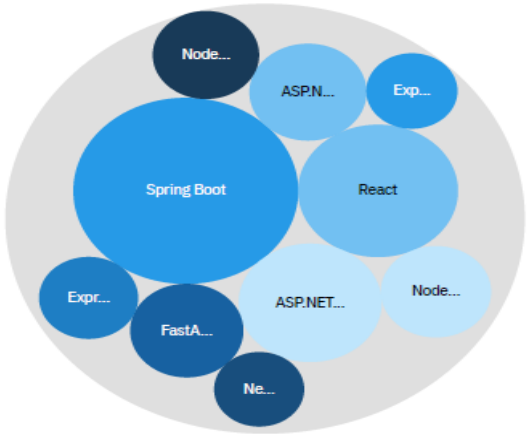


Top 10 Mejores Databases 2



Top 10 Mejores Plataformas 3

Top 10 mejores WebFrame 4

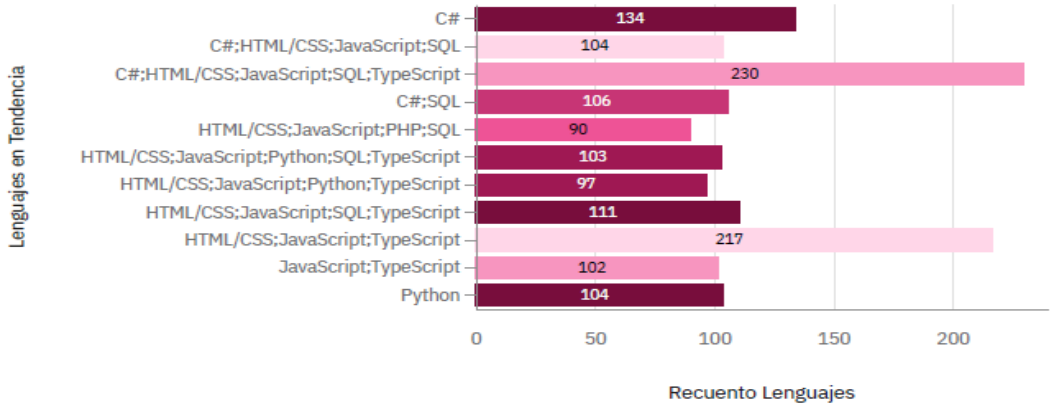


DASHBOARD TAB 2 - TRENDS

Tendencia de tecnología futura

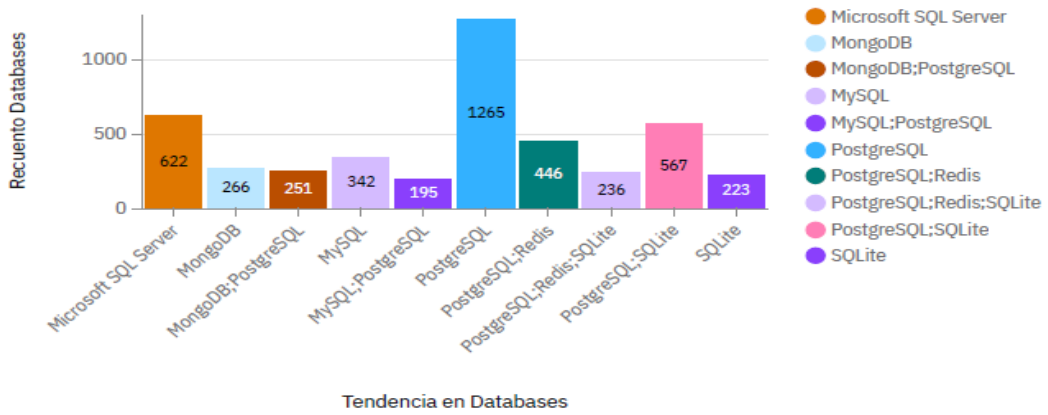
Top 10 Tendencias de Lenguajes

5



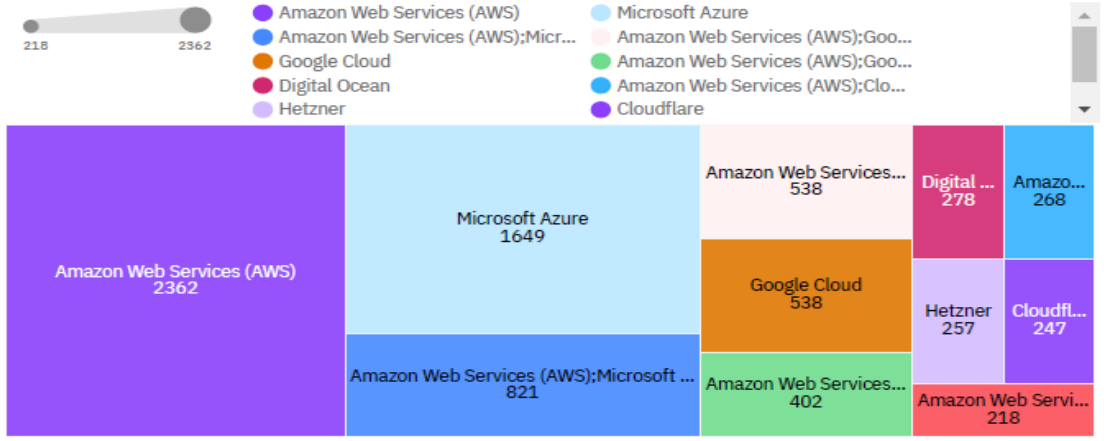
Top 10 Tendencias en Databases

6



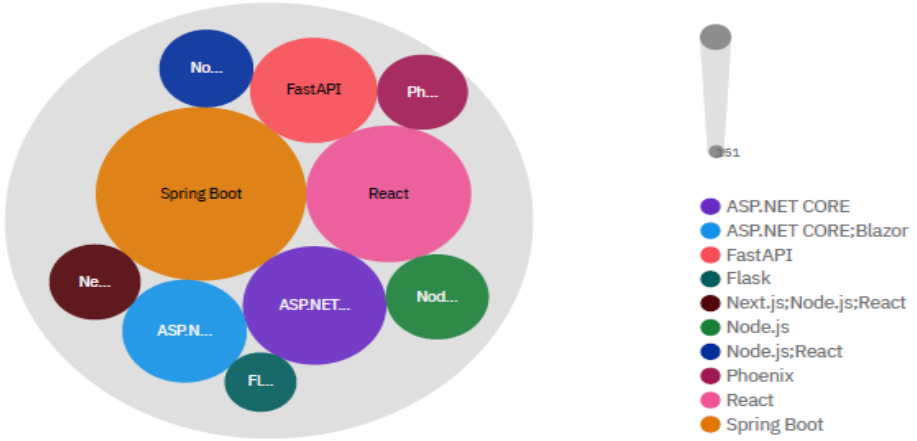
Top 10 Tendencias de Plataformas

7



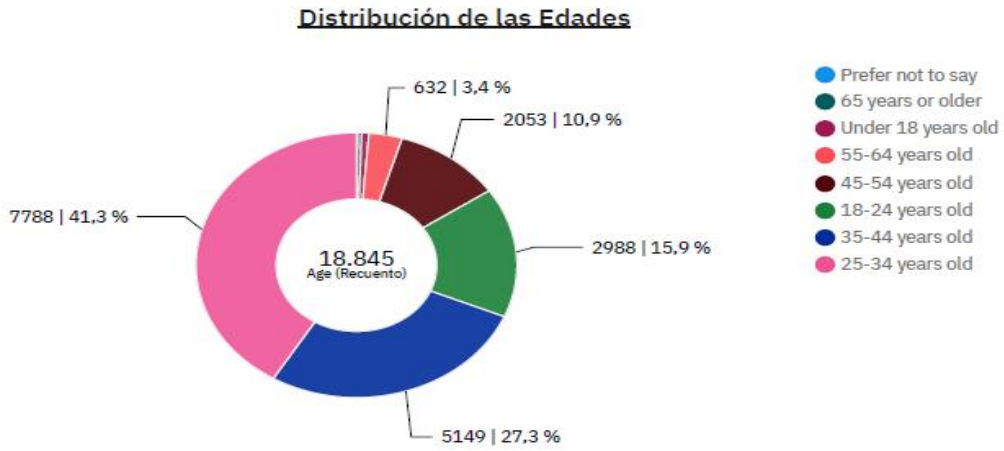
Top 10 Tendencias de Webframe

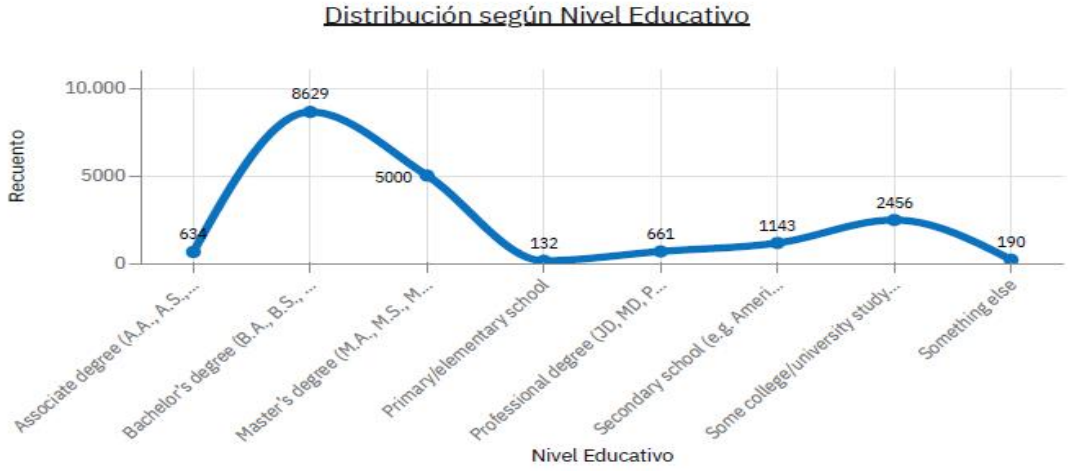
8

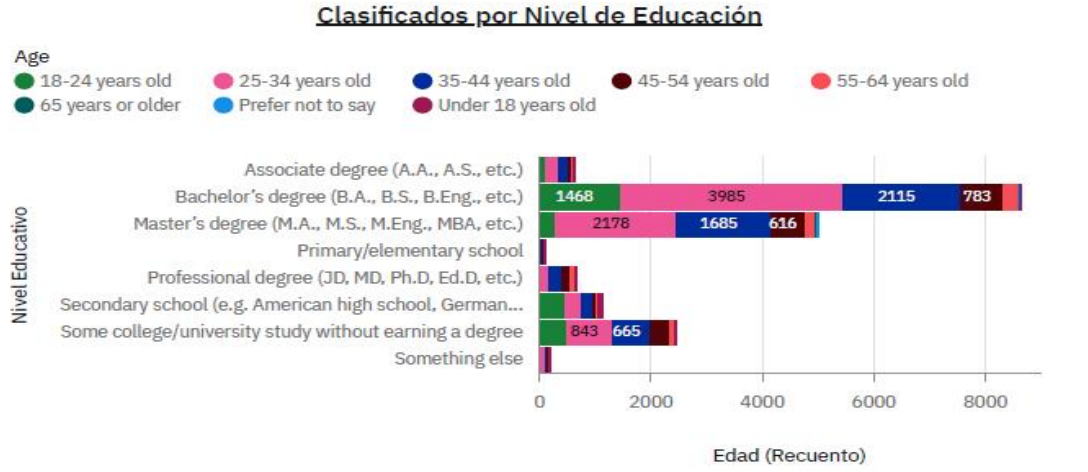


DASHBOARD TAB 3 - DEMOGRAPHY

Demografía







DISCUSSION

- While the majority of the developers surveyed were aged 25–34 and had a university education, this may be due to a participation bias in the survey, whereby the most active users of platforms such as Stack Overflow tend to be young professionals.
- The prevalence of programming languages such as JavaScript and Python may reflect not only their technical value, but also their ease of learning and the wide availability of online resources.
- Although PostgreSQL and MongoDB appear to be the preferred choice, this does not imply that technologies such as Oracle or DB2 are out of favour. Rather, they may simply be more prevalent in closed corporate environments that do not participate in such surveys.
- The popularity of cloud platforms such as AWS and Azure reflects the global digital transformation, but does not imply that all companies have migrated to the cloud. Many are still in the process or operating in hybrid environments.
- As the analysis is based on self-reported data, there may be biases in the responses, particularly with regard to salary, experience and job satisfaction.

OVERALL FINDINGS & IMPLICATIONS

Findings

- The programming languages that are currently in the most widespread use (JavaScript, Python and SQL) are also projected to remain relevant in the future.
- PostgreSQL is the current leader in terms of both usage and future projections, followed by MySQL and MongoDB.
- The majority of developers are aged between 25 and 34, and hold at least a university degree.
- Full-time employment is common, and there is a direct relationship between professional experience and job satisfaction.

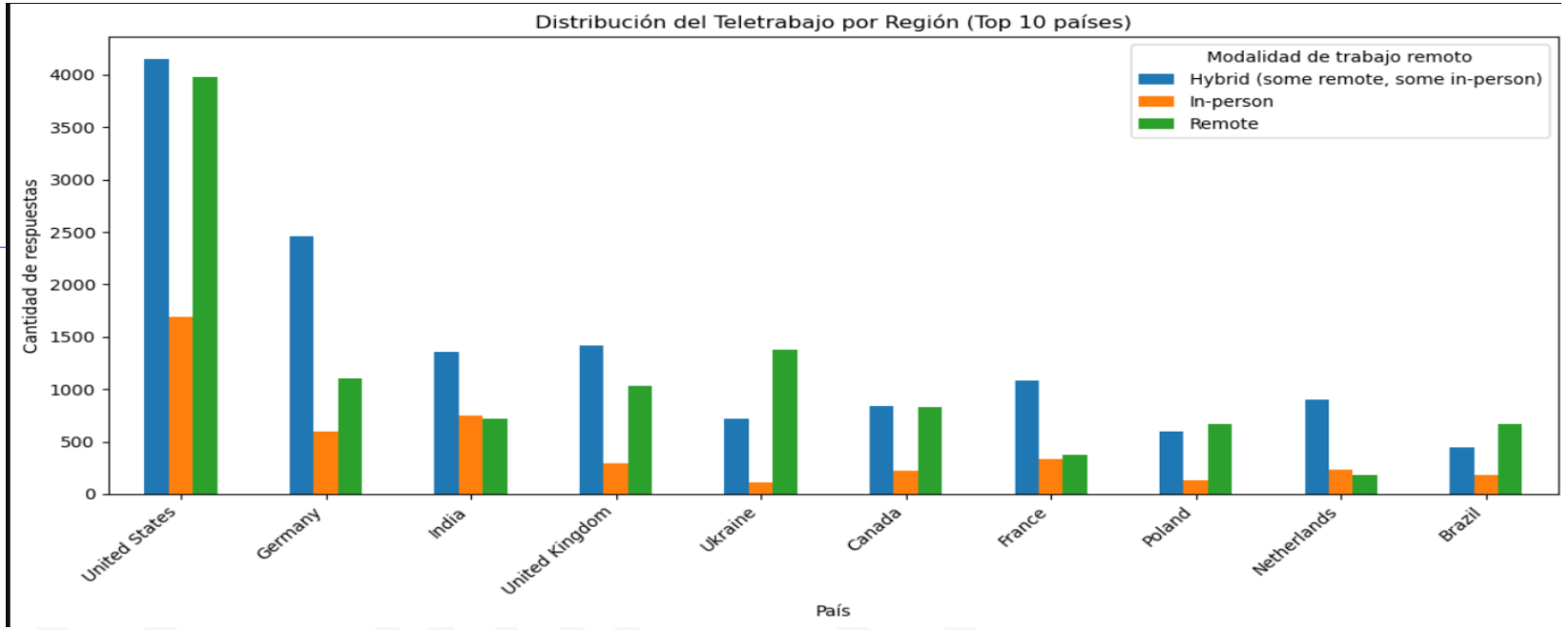
Implications

- Organisations should prioritise the development and recruitment of key technologies, such as JavaScript, Python and PostgreSQL.
- Ongoing technical training, especially in cloud technologies and NoSQL databases, is essential.
- Talent attraction strategies should focus on young professionals while also considering long-term retention through development opportunities.
- The high level of job satisfaction among more experienced employees suggests that investing in professional growth has a positive impact on talent retention.

CONCLUSION

- Analysis of the 2023 Stack Overflow survey revealed significant patterns relating to developer compensation, technology usage, employment type and educational level.
- It found that the most widely used languages and technologies today will also be in high demand in the future, thereby reaffirming their value in the market.
- Combinations of relational and non-relational databases reflect an evolution towards more flexible data architectures that can adapt to different needs.
- The relationship between experience, satisfaction, and educational level provides valuable insights for designing training, recruitment, and retention policies for technological talent.

APPENDIX


Distribución del Teletrabajo por Región (Top 10 países)
Modalidad de trabajo remoto
Hybrid (some remote, some in-person)
In-person
Remote
Cantidad de respuestas
0
500
1000
1500
2000
2500
3000
3500
4000
United States
Germany
India
United Kingdom
Ukraine
Canada
France
Poland
Netherlands
Brazil
País

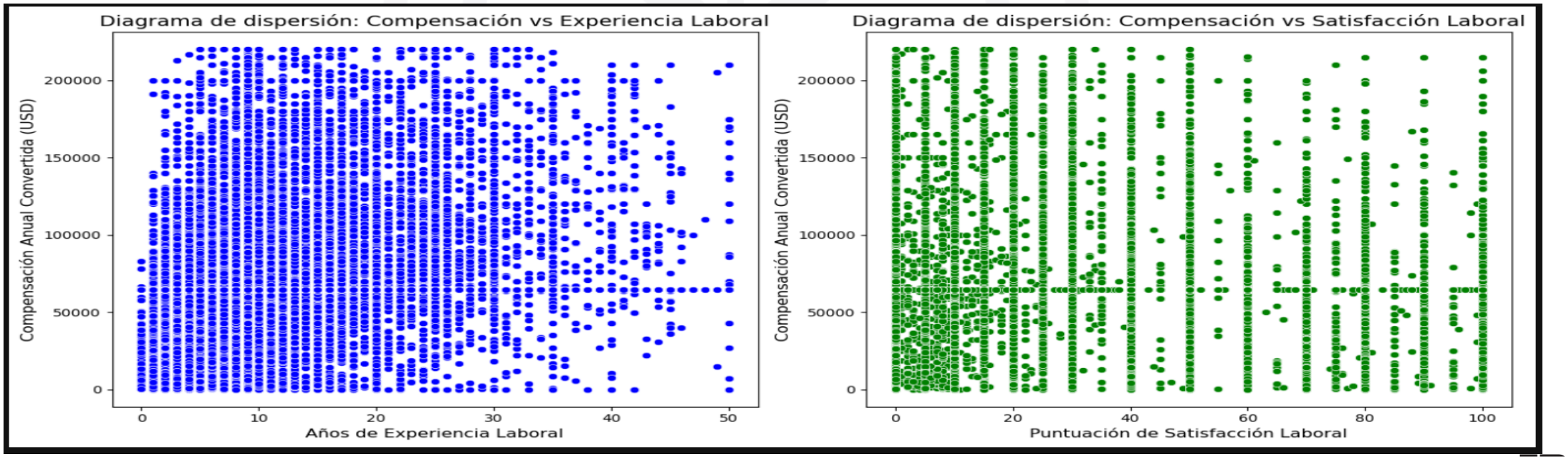
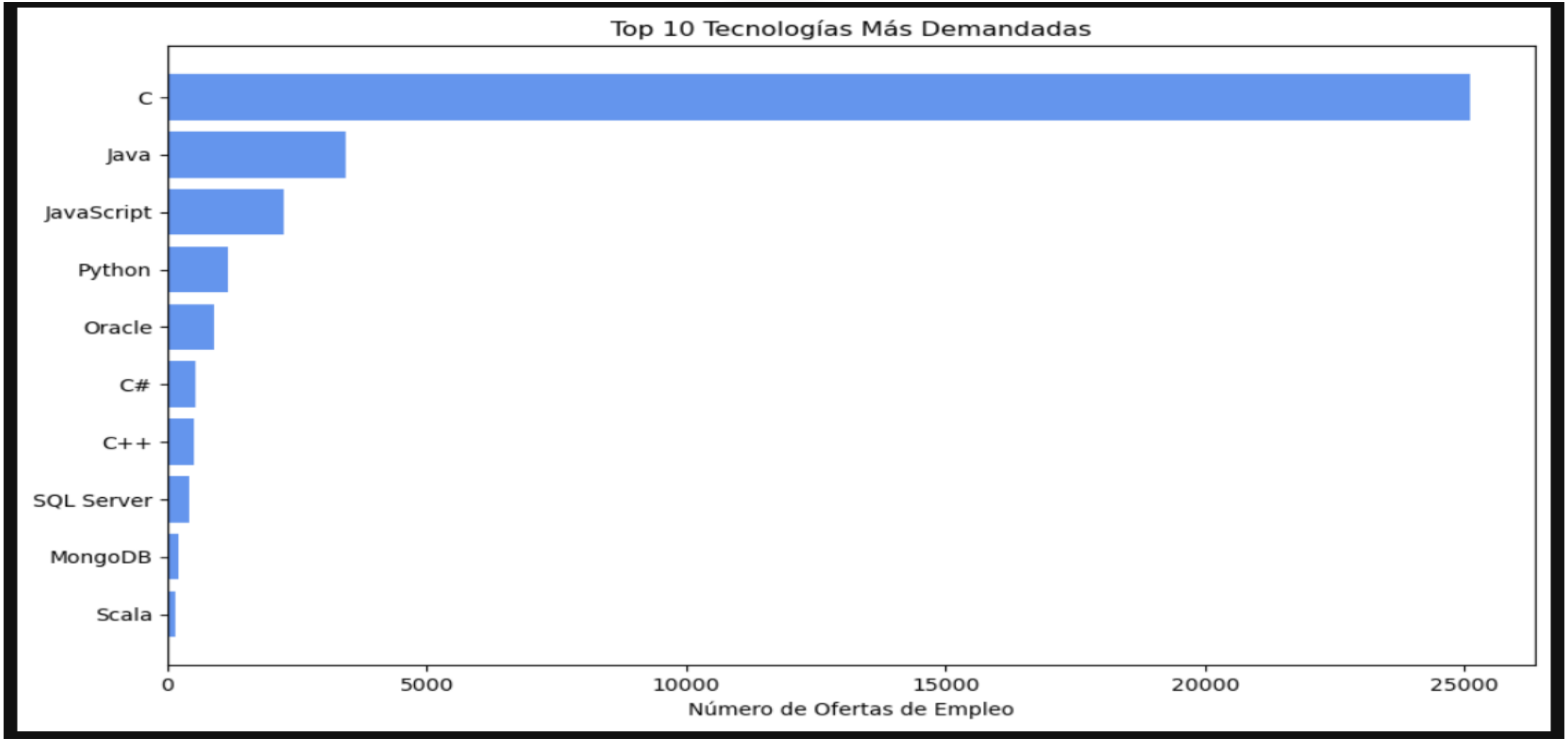


Diagrama de dispersión: Compensación vs Experiencia Laboral
Compensación Anual Convertida (USD)
0
50000
100000
150000
200000
0
10
20
30
40
50
Años de Experiencia Laboral
Diagrama de dispersión: Compensación vs Satisfacción Laboral
Compensación Anual Convertida (USD)
0
50000
100000
150000
200000
0
20
40
60
80
100
Puntuación de Satisfacción Laboral

JOB POSTINGS



POPULAR LANGUAGES