



SQL Data Analysis on Stack Overflow Developer Survey 2024

Introduction

Dataset Overview:

The dataset used in this project is the Stack Overflow Developer Survey 2024. It was conducted by Stack Overflow, a well-known platform where developers ask and answer programming-related questions. The survey includes responses from developers all over the world, covering a wide range of roles, experience levels, and backgrounds. The dataset contains over 89,000 responses.

Reason behind this Analysis

One of the major challenges in the tech industry is the lack of clarity about the rapidly evolving needs, behaviors, and preferences of software developers worldwide. Organizations often struggle to answer critical questions like: What educational backgrounds are most common among professional developers? Which programming languages or tools are being adopted by beginners? How do employment trends differ across countries? Without this knowledge, companies risk making uninformed hiring, training, and product development decisions.

This project solves that problem by conducting a comprehensive analysis of the 2024 Stack Overflow Developer Survey using SQL-based data analytics. By leveraging SQL skills such as data filtering, grouping, aggregation, and window functions, we uncovered valuable patterns – such as the dominance of full-time employment in countries like the UK and Germany, the rise of Python and VS Code among learners, and the shift towards remote-friendly tools like WSL and Linux-based OS. These insights provide actionable solutions for tech employers, educators, and policymakers to design better career pathways, align tools with developer needs, and support a global, diverse tech workforce.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Analysis Overview

To perform this analysis, I used MySQL as the primary tool. MySQL helped efficiently handle large volumes of survey data through structured queries. Key skills applied included data cleaning (removing NULL/NA values), grouping and aggregation (GROUP BY, COUNT, AVG), and window functions.

Additionally, I used Common Table Expressions (CTEs) and sorting techniques to organize and rank the results clearly. These SQL techniques enabled me to extract accurate, meaningful insights and summarize them into a comprehensive report.

Here is the github repository link of the analysis:

<https://github.com/Nimra-Iman/sql-projects-and-all-concepts/tree/main/sql%20analysis%20projects/stackOverflow%20analysis%20sql%20Project>



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

-- The top 10 countries by number of respondents?

```
SELECT
    country, COUNT(ResponseId) AS respondents
FROM
    survey_results_2024
WHERE COUNTRY != 'UNKNOWN-COUNTRY'
GROUP BY country
ORDER BY respondents DESC
LIMIT 10;
```

The United States had the highest number of respondents (11,095), followed by unknown countries (6,507), Germany (4,947), and India (4,231). Other major contributors included the UK (3,224), Ukraine (2,672), France (2,110), Canada (2,104), Poland (1,534), and the Netherlands (1,449).



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

-- How many respondents have used
JavaScript in the past year?

```
SELECT
    (SUM(LanguageHaveWorkedWith LIKE '%JavaScript%')
     / COUNT(LanguageHaveWorkedWith) * 100) AS HIGHEST_JS_USERS
FROM
    survey_results_2024;
```

About 57% respondents have used Javascript in the past year

-- How many developers have a Stack Overflow account?

```
SELECT
    (SUM(SOAccount = 'Yes') / COUNT(SOAccount) * 100)
FROM
    survey_results_2024;
```

About 69% respondents have a stackoverflow account in 2024.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What are the most used databases among respondents?

```
WITH RECURSIVE split_cte AS (
    SELECT
        ResponseId,
        SUBSTRING_INDEX(DatabaseHaveWorkedWith, ';', 1) AS db,
        substring(DatabaseHaveWorkedWith, locate(';',DatabaseHaveWorkedWith)+1) AS remaining
    FROM survey_results_2024
    WHERE DatabaseHaveWorkedWith IS NOT NULL

    UNION ALL

    SELECT
        ResponseId,
        TRIM(SUBSTRING_INDEX(remaining, ';', 1)) AS db,
        SUBSTRING(remaining, LOCATE(';', remaining)+1) AS remaining
    FROM split_cte
    WHERE remaining LIKE '%;%'
)
-- Final result
SELECT db, COUNT(*) AS usage_count
FROM split_cte
WHERE db IS NOT NULL AND db != ''
GROUP BY db
ORDER BY usage_count DESC;
```

The most commonly used databases were PostgreSQL (18,171), followed by MySQL (17,818), and Microsoft SQL Server (12,476). MongoDB was also widely used (12,244), along with MariaDB (8,893) and Elasticsearch (6,433). A significant number of responses (15,183) did not specify a database.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Which languages are most admired vs. most desired?

```
WITH RECURSIVE ADMIRE AS (
    SELECT TRIM(SUBSTRING_INDEX(LanguageHaveWorkedWith, ';', 1)) AS AD_LANGUAGE,
        CASE WHEN LOCATE(';', LanguageHaveWorkedWith) > 0 THEN
            SUBSTRING(LanguageHaveWorkedWith, LOCATE(';', LanguageHaveWorkedWith)+1) ELSE ''
        END AS REMAINING FROM survey_results_2024
    WHERE LanguageHaveWorkedWith IS NOT NULL AND LanguageHaveWorkedWith != ''
UNION ALL
    SELECT TRIM(SUBSTRING_INDEX(REMAINING, ';', 1)) AS AD_LANGUAGE,
        CASE WHEN LOCATE(';', REMAINING) > 0 THEN SUBSTRING(REMAINING, LOCATE(';', REMAINING)+1) ELSE '' END
    FROM ADMIRE WHERE REMAINING != ''),
DESIRED AS (
    SELECT TRIM(SUBSTRING_INDEX(LanguageWantToWorkWith, ';', 1)) AS DE_LANGUAGE,
        CASE WHEN LOCATE(';', LanguageWantToWorkWith) > 0 THEN SUBSTRING(LanguageWantToWorkWith,
            LOCATE(';', LanguageWantToWorkWith)+1) ELSE ''
        END AS REMAINING FROM survey_results_2024
    WHERE LanguageWantToWorkWith IS NOT NULL AND LanguageWantToWorkWith != ''
UNION ALL
    SELECT TRIM(SUBSTRING_INDEX(REMAINING, ';', 1)) AS DE_LANGUAGE,
        CASE WHEN LOCATE(';', REMAINING) > 0 THEN SUBSTRING(REMAINING, LOCATE(';', REMAINING)+1) ELSE '' END
    FROM DESIRED WHERE REMAINING != ''),
ADMIRE_COUNT AS (
    SELECT AD_LANGUAGE, COUNT(*) AS COUNT_WORKED FROM ADMIRE GROUP BY AD_LANGUAGE),
DESIRED_COUNT AS (
    SELECT DE_LANGUAGE, COUNT(*) AS COUNT_WANT_WORK FROM DESIRED GROUP BY DE_LANGUAGE)
    SELECT A.AD_LANGUAGE, A.COUNT_WORKED, D.COUNT_WANT_WORK
FROM ADMIRE_COUNT A LEFT JOIN DESIRED_COUNT D ON A.AD_LANGUAGE = D.DE_LANGUAGE UNION
SELECT D.DE_LANGUAGE, D.COUNT_WANT_WORK, A.COUNT_WORKED
FROM DESIRED_COUNT D RIGHT JOIN ADMIRE_COUNT A ON D.DE_LANGUAGE = A.AD_LANGUAGE
WHERE A.AD_LANGUAGE IS NULL
ORDER BY COUNT_WANT_WORK DESC;
```

Popular choices like Python, JavaScript, SQL, and HTML/CSS dominate both categories, showing their continued relevance. However, languages such as Rust, Go, TypeScript, and Zig have significantly more interest compared to their current usage, highlighting trends toward modern and efficient languages. Traditional languages like C, Java, and PHP remain widely used but show less future preference, while newer or niche languages (like Elixir, Dart, and Kotlin) are gaining traction among developers.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What are the average and median salaries for different countries?

```
-- AVERAGE SALARY
select country, round(avg(CompTotal),2) as average_salary from survey_results_2024
group by country ;

-- MEDIAN SALARY
WITH ALL_SALARIES AS (select country,
group_concat( cast(CompTotal as UNSIGNED) order by cast(CompTotal as UNSIGNED))
as salaries, COUNT(*) AS SALARY_COUNT from survey_results_2024
where CompTotal != '' and CompTotal REGEXP '^[0-9]+$'
GROUP BY COUNTRY)
select ALL_SALARIES.country,
SUBSTRING_INDEX(
substring_index(SALARIES, ',', (ALL_SALARIES.SALARY_COUNT/2) +1) , ',', -1)
as
median_salary
from ALL_SALARIES;
```

The average salary figures show that countries like the United States, Switzerland, and Germany report significantly higher average compensation compared to many other countries, reflecting their strong economies and high-paying tech sectors. On the other hand, countries such as India, Pakistan, and Nigeria show relatively lower average salaries, which may correspond to regional economic factors and cost of living differences.

The U.S. still leads in both average and median salaries, countries with highly skewed distributions (due to a few very high earners) show a more modest median, pointing to greater income disparity. This comparison highlights how average values can be inflated by outliers, whereas the median gives a clearer picture of what a typical developer earns in each country.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Which AI tools are most admired, and how does usage compare -- between professionals and learners?

```
WITH RECURSIVE AI_TOOL AS (
    SELECT MainBranch, CASE
        WHEN MainBranch = 'I am learning to code' THEN 'learner'
        WHEN MainBranch = 'None of these' OR MainBranch = 'NA' THEN ''
        ELSE 'developer' END AS respondent_type,
        SUBSTRING_INDEX(AISearchDevHaveWorkedWith, ';', 1) AS TOOL_USED,
        CASE WHEN (LOCATE(';', AISearchDevHaveWorkedWith)>1)
        THEN SUBSTRING(AISearchDevHaveWorkedWith, LOCATE(';', AISearchDevHaveWorkedWith)+1)
        ELSE '' END AS REMAINAING FROM survey_results_2024
    UNION ALL
    SELECT MainBranch, CASE
        WHEN MainBranch = 'I am learning to code' THEN 'learner'
        WHEN MainBranch = 'None of these' OR MainBranch = 'NA' THEN ''
        ELSE 'developer' END AS respondent_type,
        SUBSTRING_INDEX(REMAINAING, ';', 1) AS TOOL_USED,
        CASE WHEN (LOCATE(';', REMAINAING)>1)
        THEN SUBSTRING(REMAINAING, LOCATE(';', REMAINAING)+1)
        ELSE '' END AS REMAINAING FROM AI_TOOL
    WHERE REMAINAING != '' AND REMAINAING IS NOT NULL
)
SELECT TOOL_USED,
    SUM(CASE WHEN respondent_type = 'developer' THEN 1 ELSE 0 END) AS DEVELOPER_USE,
    SUM(CASE WHEN respondent_type = 'learner' THEN 1 ELSE 0 END) AS LEARNER_USE
FROM AI_TOOL GROUP BY
TOOL USED ORDER BY DEVELOPER USE DESC;
```

The data shows that ChatGPT is the most used AI tool by both developers (35,619) and learners (23,042), making it the top choice for coding assistance. GitHub Copilot is second with high developer use (18,220) but low learner use (812), showing it's more developer-focused.

Google Gemini, Bing AI, and Claude have balanced usage, indicating growing popularity among all users. Tools like Codeium, Perplexity AI, and WolframAlpha also show good engagement. Low-usage tools such as Metaphor, Neeva AI, and AskCodi have minimal adoption, suggesting low awareness or limited use cases.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What is the relationship between developer type and employment status?

```
with recursive emp as (
SELECT ResponseId,
SUBSTRING_INDEX(employment, ',', 1) AS EMPLOYMENT_STATUS,
CASE WHEN LOCATE(';', employment)>1) THEN
SUBSTRING(employment, LOCATE(';', employment)+1) ELSE '' END
AS REMAINING FROM survey_results_2024
UNION ALL
SELECT ResponseId, substring_index(REMAINING , ',', 1) AS EMPLOYMENT_STATUS,
CASE WHEN LOCATE(';',REMAINING)>1) THEN
SUBSTRING(REMAINING, LOCATE(';',REMAINING)+1)
ELSE '' END AS REMAINING FROM emp
WHERE REMAINING IS NOT NULL AND REMAINING != ''
)
SELECT survey_results_2024.devtype,
sum(case when EMP.employment_status = 'Employed, full-time' then 1 else 0 end) as 'Employed, full-time',
sum(case when EMP.employment_status = 'Employed, part-time' then 1 else 0 end) as 'Employed, part-time',
sum(case when EMP.employment_status = 'Independent contractor, freelancer, or self-employed' then 1 else 0 end) as 'Independent contractor, freelancer, or self-employed',
sum(case when EMP.employment_status = 'Not employed, but looking for work' then 1 else 0 end) as 'Not employed, but looking for work',
sum(case when EMP.employment_status = 'Not employed, and not looking for work' then 1 else 0 end) as 'Not employed, and not looking for work',
sum(case when EMP.employment_status = 'Student, full-time' then 1 else 0 end) as 'Student, full-time',
sum(case when EMP.employment_status = 'Student, part-time' then 1 else 0 end) as 'Student, part-time',
sum(case when EMP.employment_status = 'Retired' then 1 else 0 end) as 'Retired',
sum(case when EMP.employment_status = 'I prefer not to say' then 1 else 0 end) as 'I prefer not to say'
from emp join survey_results_2024 on
emp.ResponseId = survey_results_2024.ResponseId
GROUP BY survey_results_2024.devtype;
```

Full-stack developers lead the tech scene with the highest number of full-time workers (13,824), followed by back-end (8,137) and front-end developers (2,521). Students are most common in full-stack, back-end, and “Student” categories, showing strong early interest.

Data science, DevOps, desktop apps, and mobile dev roles also have high full-time employment with fewer students. Roles like project managers, engineering managers, and executives are mostly filled by experienced professionals, with minimal student presence.

Emerging fields like AI, blockchain, and cloud show a mix of professionals and students, indicating future growth. Overall, tech is dominated by full-time workers, but students are entering across many roles, preparing for a strong future in the industry.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Which combination of technologies is most common among full-time developers?

```
WITH RECURSIVE TECH AS(
SELECT MainBranch,substring_index(ToolsTechHaveWorkedWith, ';', 1) AS COMMON_TECH,
CASE WHEN (LOCATE(';', ToolsTechHaveWorkedWith)>1) THEN
SUBSTRING(ToolsTechHaveWorkedWith, LOCATE(';', ToolsTechHaveWorkedWith)+1)
ELSE '' END AS REMAINING FROM survey_results_2024

UNION ALL

SELECT MainBranch,substring_index(REMAINING, ';', 1) AS COMMON_TECH,
CASE WHEN (LOCATE(';', REMAINING)>1) THEN
SUBSTRING(REMAINING, LOCATE(';', REMAINING)+1)
ELSE '' END AS REMAINING FROM tech
where REMAINING!='' and REMAINING is not null
)

select COMMON_TECH, count(COMMON_TECH) as count from tech
where MainBranch = 'I am a developer by profession'
group by COMMON_TECH order by count desc;
```

Docker leads all tools with 25K+ users, followed by npm (22K) and Pip (13K). Homebrew, Kubernetes, Yarn, and Vite are also widely used (9K–10K range). Tools like NuGet, Maven, and Gradle fall in the mid-range (6K–7K). Less common tools include Unity 3D, Godot, and Podman. Niche tools like Dagger, Pulumi, and Chef have smaller but active user bases.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What is educational attainment of professional developers?

```
SELECT  
    EdLevel,  
    COUNT(EdLevel) AS Count,  
    COUNT(EdLevel) * 100.0 / SUM(COUNT(EdLevel)) OVER () AS Percentage  
FROM  
    survey_results_2024  
WHERE  
    MainBranch = 'I am a developer by profession'  
    AND EdLevel IS NOT NULL  
    AND EdLevel != 'NA'  
GROUP BY  
    EdLevel  
ORDER BY  
    Count DESC;
```

Among professional developers, nearly 46% hold a bachelor's degree, making it the most common level of education. This is followed by master's degree holders at 28%, showing that a large portion of developers pursue advanced academic qualifications. A smaller group, around 4%, hold professional doctorates such as PhDs or MDs.

Other categories, including associate degrees, alternative paths, and even primary-level education, make up the remaining small fraction, highlighting diverse paths into the profession.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

how survey respondents learn to code based on their age?

```
WITH RECURSIVE LEARNING_PLATFORMS AS (
    SELECT
        AGE, SUBSTRING_INDEX(LEARNCODE, ';', 1) AS PLATFORM,
        CASE
            WHEN LOCATE(';', LEARNCODE) > 0 THEN SUBSTRING(LEARNCODE, LOCATE(';', LEARNCODE) + 1)
            ELSE '' END AS REM FROM survey_results_2024
    UNION ALL
    SELECT
        AGE, SUBSTRING_INDEX(REM, ';', 1) AS PLATFORM,
        CASE
            WHEN LOCATE(';', REM) > 0 THEN SUBSTRING(REM, LOCATE(';', REM) + 1)
            ELSE '' END AS REM FROM LEARNING_PLATFORMS WHERE REM IS NOT NULL AND REM != ''
)
SELECT
    PLATFORM, AGE,
    COUNT(PLATFORM) * 100.0 / SUM(COUNT(PLATFORM)) OVER(PARTITION BY AGE) AS total_percentage
FROM LEARNING_PLATFORMS
WHERE PLATFORM IS NOT NULL AND PLATFORM != 'NA'
GROUP BY PLATFORM, AGE
ORDER BY AGE, total_percentage;
```

Younger age groups (especially under 24) mostly use online resources (blogs, videos, etc.)—with up to 36% relying on them. Formal education is common among 18–34-year-olds (16–20%), while online courses and on-the-job training are popular across all adults (14–16%).

Older learners (45+) prefer books (up to 24%) and on-the-job training.

Bootcamps and friends/family are more common for younger users (around 3–7%).

In short:

- Young = Online + School
- Middle age = Courses + Job experience
- Older = Books + Training



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Top 10 online resources for learning code?

```
WITH RECURSIVE ONLINE_RESOURCES AS (
  SELECT
    AGE,
    SUBSTRING_INDEX(LEARNCODEONLINE, ';', 1) AS ONLINE_RESOURCE,
    CASE
      WHEN LOCATE(';', LEARNCODEONLINE) > 0 THEN SUBSTRING(LEARNCODEONLINE, LOCATE(';', LEARNCODEONLINE) + 1)
      ELSE '' END AS REM FROM survey_results_2024
  UNION ALL
  SELECT
    AGE,
    SUBSTRING_INDEX(REM, ';', 1) AS ONLINE_RESOURCE,
    CASE
      WHEN LOCATE(';', REM) > 0 THEN SUBSTRING(REM, LOCATE(';', REM) + 1)
      ELSE '' END AS REM
  FROM ONLINE_RESOURCES
  WHERE REM IS NOT NULL AND REM != ''
)
SELECT
  ONLINE_RESOURCE,
  ROUND(COUNT(ONLINE_RESOURCE)*100/SUM(COUNT(ONLINE_RESOURCE)) OVER(),2) AS RESOURCE_PERCENTAGE FROM ONLINE_RESOURCES
GROUP BY ONLINE_RESOURCE ORDER BY RESOURCE_PERCENTAGE DESC;
```

When it comes to learning programming, most developers still turn to technical documentation (12%) and Stack Overflow (11.5%) for learning. Written tutorials (9.8%), blogs (8.8%), and videos (how-to: 7.8%, courses: 7.2%) are also popular. Books (6.7%), social media (5.9%), and AI tools (5.3%) follow.

Less used are challenges, podcasts, games, and certification videos. Text-based and practical resources lead, but video and AI tools are growing.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

How many years have developers been coding, both overall and professionally?

```
WITH OVERALL_YEARS_OF_CODE AS(
SELECT RESPONSEID,
CASE
WHEN YearsCode = 'Less than 1 year' THEN 'Less than 1 year'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 1 AND 4 THEN '1 to 4 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 5 AND 9 THEN '5 to 9 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 10 AND 14 THEN '10 to 14 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 15 AND 19 THEN '15 to 19 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 20 AND 24 THEN '20 to 24 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 25 AND 29 THEN '25 to 29 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 30 AND 34 THEN '30 to 34 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 35 AND 39 THEN '35 to 39 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 40 AND 44 THEN '40 to 44 years'
WHEN CAST(YearsCode AS UNSIGNED) BETWEEN 45 AND 49 THEN '45 to 49 years'
WHEN YearsCode = 'More than 50 years' THEN 'More than 50 years'
ELSE 'Unknown'
END AS OVERALL_CODING_YEARS FROM survey_results_2024)
SELECT OVERALL_CODING_YEARS AS OVERALL_EXPERIENCE,
ROUND(COUNT(OVERALL_CODING_YEARS)*100/SUM(COUNT(OVERALL_CODING_YEARS)) OVER(),2)
AS OVERALL_CODING_YEARS_PERCENT
FROM OVERALL_YEARS_OF_CODE
WHERE OVERALL_CODING_YEARS IS NOT NULL AND
OVERALL_CODING_YEARS != 'NA' AND OVERALL_CODING_YEARS != 'UNKNOWN'
GROUP BY OVERALL_CODING_YEARS ORDER BY OVERALL_EXPERIENCE;
```

Most developers have 5 to 9 years of coding experience (27.18%), making it the largest group. This is followed by those with 10 to 14 years (20.16%) and 1 to 4 years (13.65%), showing a strong presence of mid-career developers.

Interestingly, 15 to 19 years (11.78%) and 20 to 24 years (9.21%) also form a solid base of experienced professionals. Newcomers with less than 1 year make up only 0.95%, while veterans with more than 50 years are just 0.42%, highlighting that most of the developer community sits in the early to mid-career phase.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What TYPE OF developers have more than 15 years of coding experience?

```
SELECT
    DEVTYPE, ROUND(AVG(YearsCodePRO), 2) AS EXPERIENCE
FROM
    survey_results_2024
WHERE
    DEVTYPE != 'NA' AND YearsCodePRO != 'NA'
    GROUP BY DEVTYPE
    HAVING EXPERIENCE > 15
ORDER BY EXPERIENCE DESC;
```

Among the most experienced developers, Senior Executives (C-Suite, VPs) lead with 17.3%, showing that many have transitioned into top leadership roles. Engineering Managers follow closely at 15.83%, indicating a strong presence of seasoned professionals in team leadership. Developer Advocates make up 15.08%, reflecting their deep experience in both coding and community engagement. This suggests that high experience often correlates with strategic, managerial, or outreach roles in tech.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What are top 5 popular languages used by people who are learning to code?

```
WITH RECURSIVE LANGUAGE_USED AS(
SELECT MAINBRANCH, SUBSTRING_INDEX(LanguageHaveWorkedWith, ',',1) AS LANG,
CASE WHEN (LOCATE(',', LanguageHaveWorkedWith)>1) THEN
SUBSTRING(LanguageHaveWorkedWith, LOCATE(',',LanguageHaveWorkedWith)+1)
ELSE '' END AS REM FROM survey_results_2024
UNION ALL
SELECT MAINBRANCH, SUBSTRING_INDEX(REM, ',',1) AS LANG,
CASE WHEN (LOCATE(',', REM)>1) THEN
SUBSTRING(REM, LOCATE(',',REM)+1)
ELSE '' END AS REM FROM LANGUAGE_USED
WHERE REM !='NA' AND REM IS NOT NULL AND REM!='')
SELECT LANG, COUNT(LANG)*100/SUM(COUNT(LANG)) OVER() AS COUNT
FROM LANGUAGE_USED WHERE MAINBRANCH
LIKE '%LEARNING TO CODE%' GROUP BY LANG ORDER BY COUNT DESC LIMIT 5;
```

Among individuals who are learning to code, HTML/CSS stands out as the most commonly used language, with 11.27% of learners engaging with it—likely due to its foundational role in web development. JavaScript follows closely at 11.20%, reinforcing the popularity of front-end development as an entry point into programming.

Python, known for its simplicity and versatility, ranks third at 10.97%, making it a top choice for those interested in data science, automation, and general-purpose programming. SQL appears next with 7.40%, reflecting a strong interest in working with databases.

Finally, Java comes in at 6.73%, showing its continued relevance in education, enterprise applications, and Android development. This distribution suggests that beginners are primarily drawn to technologies that offer visible results and wide application.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What are top top 3 most common IDEs used by professional developers?

```
WITH RECURSIVE PLATFORM AS(
SELECT MAINBRANCH, SUBSTRING_INDEX(NEWCollabToolsHaveWorkedWith, ';',1) AS IDE,
CASE WHEN (LOCATE(';', NEWCollabToolsHaveWorkedWith)>1) THEN
SUBSTRING(PlatformHaveWorkedWith, LOCATE(';',NEWCollabToolsHaveWorkedWith)+1)
ELSE '' END AS REM FROM survey_results_2024

UNION ALL

SELECT MAINBRANCH, SUBSTRING_INDEX(REM, ';',1) AS LANG,
CASE WHEN (LOCATE(';', REM)>1) THEN
SUBSTRING(REM, LOCATE(';',REM)+1)
ELSE '' END AS REM FROM PLATFORM
WHERE REM !='NA' AND REM IS NOT NULL AND REM!='')

SELECT IDE, ROUND(COUNT(IDE)*100/SUM(COUNT(IDE)) OVER(),2) AS TOP_IDEs FROM
PLATFORM where IDE is not null and IDE !='NA' GROUP BY IDE ORDER BY TOP_IDEs
desc limit 3 ;
```

The top 3 IDEs used by professional developers are Android Studio (7.25%), IntelliJ IDEA (5.61%), and Visual Studio Code (5.29%), reflecting a strong preference for mobile and general-purpose development environments.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

What are the most commonly used operating systems by developers for personal and professional use?

```
WITH RECURSIVE OSPERSONAL AS(
SELECT SUBSTRING_INDEX(`OpSysPersonal use`, ',',1) AS OS,
CASE WHEN (LOCATE(';', `OpSysPersonal use`)>1) THEN
SUBSTRING(`OpSysPersonal use`, LOCATE(';', `OpSysPersonal use`)+1)
ELSE '' END AS REM FROM survey_results_2024 UNION ALL
SELECT SUBSTRING_INDEX(REM, ',',1) AS OS,
CASE WHEN (LOCATE(';', REM)>1) THEN SUBSTRING(REM, LOCATE(';',REM)+1)
ELSE '' END AS REM FROM OSPERSONAL WHERE REM !='NA' AND REM
IS NOT NULL AND REM!= ''),
OSPROFESSIONAL AS ( SELECT SUBSTRING_INDEX(`OpSysProfessional use`, ',',1) AS OS,
CASE WHEN (LOCATE(';', `OpSysProfessional use`)>1) THEN
SUBSTRING(`OpSysProfessional use`, LOCATE(';', `OpSysProfessional use`)+1)
ELSE '' END AS REM FROM survey_results_2024 UNION ALL SELECT SUBSTRING_INDEX(REM, ',',1) AS OS,
CASE WHEN (LOCATE(';', REM)>1) THEN SUBSTRING(REM, LOCATE(';',REM)+1)
ELSE '' END AS REM FROM OSPROFESSIONAL WHERE REM !='NA' AND REM IS NOT NULL AND REM!= '' ),
OSPERS_COUNT AS (
SELECT OS, ROUND(COUNT(OS)*100/SUM(COUNT(OS)) OVER(),2) AS
PERSONAL_USE_PERCENT FROM OSPERSONAL WHERE OS IS NOT NULL AND OS !='NA' GROUP BY OS ),
OSPROF_COUNT AS (
SELECT OS, ROUND(COUNT(OS)*100/SUM(COUNT(OS)) OVER(),2) AS
PROFESSIONAL_USE_PERCENT FROM OSPROFESSIONAL WHERE OS IS NOT NULL AND OS !='NA'
GROUP BY OS )
SELECT OSPERS_COUNT.OS, OSPERS_COUNT.PERSONAL_USE_PERCENT ,
OSPROF_COUNT.PROFESSIONAL_USE_PERCENT
FROM OSPERS_COUNT LEFT JOIN OSPROF_COUNT ON OSPERS_COUNT.OS = OSPROF_COUNT.OS UNION
SELECT OSPERS_COUNT.PERSONAL_USE_PERCENT, OSPROF_COUNT.OS, OSPROF_COUNT.PROFESSIONAL_USE_PERCENT
FROM OSPROF_COUNT RIGHT JOIN OSPERS_COUNT ON OSPROF_COUNT.OS = OSPERS_COUNT.OS WHERE
OSPROF_COUNT.PROFESSIONAL_USE_PERCENT IS NULL ORDER BY PERSONAL_USE_PERCENT DESC;
```

Among developers, Windows remains the most used OS both personally (27.96%) and professionally (25.42%). MacOS follows with higher professional use (16.96%) than personal (15.02%), while Ubuntu sees a strong presence in both domains (13.06% personal, 14.78% professional).

The Windows Subsystem for Linux (WSL) is notably more used professionally (8.98%) than personally (8.09%), showing its growing role in development workflows. Android is used more personally (8.44%) than professionally (4.49%), whereas niche systems like Red Hat, AIX, and Solaris show minimal but specialized professional usage.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

EMPLOYMENT STATUS BY GEOGRAPHY?

```
WITH RECURSIVE emp AS(
SELECT COUNTRY, SUBSTRING_INDEX(EMPLOYMENT, ',',1) AS employ,
CASE WHEN (LOCATE(',', EMPLOYMENT)>1) THEN
SUBSTRING(EMPLOYMENT, LOCATE(',',EMPLOYMENT)+1)
ELSE '' END AS REM FROM survey_results_2024
UNION ALL
SELECT COUNTRY, SUBSTRING_INDEX(REM, ',',1) AS EMPLOY,
CASE WHEN (LOCATE(',', REM)>1) THEN
SUBSTRING(REM, LOCATE(',',REM)+1)
ELSE '' END AS REM FROM emp
WHERE REM !='NA' AND REM IS NOT NULL AND REM!='')
SELECT COUNTRY, EMPLOY,
ROUND(COUNT(EMPLOY)*100/SUM(COUNT(EMPLOY)) OVER(PARTITION BY COUNTRY),2)
AS DEVELOPER_PERCENTAGE FROM emp
GROUP BY COUNTRY,EMPLOY ORDER BY COUNTRY,DEVELOPER_PERCENTAGE DESC ;
```

In different parts of the world, developers follow unique work paths. In the US, a fair number work full-time, but many are freelancers (11.1%) or full-time students (9.1%). Germany sees a majority (56.3%) in full-time jobs, with students and part-time workers making up a notable portion.

Over in India, most developers are employed full-time (55.9%), yet it also has a large student community (19.5%) and a visible freelance and job-seeking presence. The UK stands out with the highest full-time employment (66.4%) and strong freelance activity (12.3%), while in Ukraine, full-time roles (60.9%) and freelancing (14.6%) dominate, showing a vibrant mix of experience and independence.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

NUMBER OF EMPLOYES IN ORGANIZATION IN WHICH PEOPLE WORK?

```
SELECT
    OrgSize,
    COUNT(OrgSize) * 100.0 / SUM(COUNT(OrgSize)) OVER () AS PERCENTAGE_COUNT
FROM
    survey_results_2024
WHERE
    OrgSize IS NOT NULL
    AND OrgSize != 'NA'
GROUP BY
    OrgSize
ORDER BY
    PERCENTAGE_COUNT DESC;
```

Most developers surveyed work in mid-sized companies. The largest share, 20.5%, are part of organizations with 20 to 99 employees, closely followed by 18.3% in firms with 100 to 499 employees. Large enterprises with 10,000 or more employees employ 11.7%, while 11.3% work at companies with 1,000 to 4,999 staff.

Smaller setups are also common, with 10.2% in 2 to 9 person teams and 8.6% in 10 to 19 employee organizations. Around 6.7% are in 500–999 employee firms, and 6.5% are solo professionals or freelancers. Bigger firms with 5,000 to 9,999 employees account for 3.9%, and 2.2% respondents were unsure of their organization size.



Analysis by Nimra Iman



SQL Data Analysis on Stack Overflow Developer Survey 2024

Conclusion

This project provides a comprehensive analysis of the 2024 Stack Overflow Developer Survey data, offering deep insights into the global developer community. By exploring various aspects such as education levels, years of coding experience, preferred programming languages, tools, operating systems, and employment types across different countries, the project uncovers key trends and patterns shaping the tech industry today.

Using SQL for data extraction and analysis, this project demonstrates essential skills like querying, aggregation, filtering, and data interpretation. The results not only highlight the dominant technologies and platforms but also reflect the diverse backgrounds and career paths of developers worldwide. Overall, this study equips stakeholders with valuable knowledge to understand the current state of the developer ecosystem and make data-informed decisions in tech-related domains.



Analysis by Nimra Iman