

Report on

# JOBS DATA EDA

## Most Popular tools

Prepared by

Laima Lukoseviciute

February 11, 2026

---

## Table of contents

Executive Summary	2
Introduction	2
Selection of Tools . . . . .	2
Data Preparation and Cleaning	3
Observations and Features . . . . .	4
Outliers . . . . .	4
Exploratory Data Analysis	5
Distribution of Features . . . . .	5
Distribution of the Mentions of tools . . . . .	5
Distribution of Job Titles by Job Type . . . . .	7
Distribution of Company Names . . . . .	7
Distribution of Industries . . . . .	7
Distribution of States . . . . .	9
Distribution of Job Posting Times . . . . .	12
Tool Demand by Category . . . . .	12
Tool Demand by Tool Category . . . . .	12
Tool Demand by Job Title . . . . .	13
Tool Demand by Company Industry . . . . .	16
Tools Co-occurrence Frequency . . . . .	16
The Most in Demand Cloud Platforms by State . . . . .	16
Market Leader Deep Dive: FAANG and MANGO . . . . .	18
FAANG . . . . .	18
MANGO . . . . .	19
Key Insights	19
Summary	21
Suggestions for Further Improvements	21
Limitations	22

## Executive Summary

- In 2025 the Modern Enterprise Tech Stack is maturing. While AWS and Git are the bedrock of technical roles, Excel remains the world's most required data tool, proving that human-readable data is still the ultimate business priority. However, the presence of Kubernetes and Power BI shows that businesses are moving toward a future that is both cloud-native and deeply integrated with the Microsoft ecosystem.
- The 2025 job market shows a strong demand for professionals proficient in the following pillars of the Modern Enterprise Tech Stack:
  - Excel is mentioned in 37% of job postings.
  - Followed by the cloud platforms (AWS with 29%, Azure with 24%, and GCL with 14%) for storing, managing, and accessing data.
  - Git(21%) for ensuring the code is versioned.
  - Kubernetes(13%) and Docker(12%) for scalable software.

## Introduction

This report presents an exploratory data analysis (EDA) of job posting data collected throughout 2025. The primary goal is to identify the most required tools designed to collect, store, transform, and analyze data efficiently, or tools that do the supporting role for these operations like git. These will further be referred to as tools. The presence of these tools will be analyzed in the current labor market, in order to understand how these requirements shift across different professional specializations. By analyzing the intersection of job titles and required tools, I will aim to provide some insights into the most popular tools defining the industry in the year 2025.

The initial raw data was gained from a large-scale coresignal jobs data, totaling approximately 60 million global job postings. The analyzed data set used in this report consists of 34 features and around 600 000 observations, ranging from January 1st, 2025, to December 19, 2025. For a detailed breakdown of the features, please refer to the Table 1.

Table 1: The description of variables for data.

Variable Name	Type	Description
title [manager, engineer, analyst, scientist, developer]	STRING BOOL	The professional title of the job listing. A flag indicating if the role is categorized as a management, engineering, analyst, scientist, developer role.
company_name	STRING	The name of the hiring organization.
company_industry	STRING	The sector the company operates in (e.g., Tech, Finance).
broad_industry_group	STRING	The broad sector group that the company operates in (e.g., "Tech, Data and Telecom").
state	STRING	The US state of the job location.
created_at [Excel, AWS, ...]	TIMESTAMP	The date when the job listing was added to the database. Individual flags for 23 tools (e.g., Excel, AWS, ...) indicating if the role listing ask experience in that role.

## Selection of Tools

The tools selected for this analysis represent the Modern Enterprise Tech Stack - a modular, cloud-native ecosystem that has become the standard for high-performing organizations that need to handle a lot of data.

1. Data Storage & Infrastructure - These tools are the “foundations” where massive datasets live.
  - Cloud Data Warehouses: “Snowflake”, “BigQuery”
  - Cloud Platforms (General): “AWS”, “Azure”, “GCP” (Google Cloud Platform)
  - Big Data Storage/Processing: “Databricks”, “Spark”
2. Data Ingestion & Transformation - These tools move data from source to destination and clean it up for use.

- Ingestion (ELT): "Fivetran", "Airbyte"
  - Transformation: "dbt"
  - Streaming: "Kafka"
3. Business Intelligence (BI) & Analytics - The "face" of the data where executives see the results.
- Modern BI: "Power\_BI", "Tableau", "Looker"
  - Spreadsheets: "Excel", "Google\_Sheets"
4. Orchestration & Observability - The "management" layer that ensures everything runs on time and the data is accurate.
- Orchestration: "Airflow", "Prefect"
  - Observability (Quality): "Monte\_Carlo"
5. DevOps & Developer Experience - Tools that ensure code is versioned and software is scalable.
- Version Control: "Git"
  - Containerization: "Docker", "Kubernetes"
  - Infrastructure as Code: "Terraform"

## Data Preparation and Cleaning

For data preparation and cleaning please refer to "Appendix 1: Preparing the data set." It will give the detailed look into how the data frame that I will analyse here was prepared.

A preview of the analysed dataset is presented below in Table 2, Table 3, Table 4.

(596962, 34)

Table 2: Raw data pre-view first 5 rows(part 1)

	title	company_name	company_industry	state	created_at
0	Senior Manager, ...	Amazon	Software Develop...	Tennessee	2025-09-10 20:07:...
1	Business Intelli...	IntelliSavvy	IT Services and ...	Washington	2025-09-22 12:07:...
2	HAZARDOUS SUBSTA...	California Depar...	Environmental Se...	California	2025-06-04 00:17:...
3	PMS 378 Senior T...	CACI Internation...	IT Services and ...	District of Colu...	2025-09-11 23:04:...
4	COMPENSATION ANA...	Boston University	NaN	Massachusetts	2025-09-18 01:04:...

Table 3: Raw data pre-view first 5 rows(part 2)

	title	manager	engineer	analyst	scientist	developer
0	Senior Manager, ...	1	1	0	0	0
1	Business Intelli...	0	1	0	0	0
2	HAZARDOUS SUBSTA...	0	1	0	0	0
3	PMS 378 Senior T...	0	0	1	0	0
4	COMPENSATION ANA...	0	0	1	0	0

Table 4: Raw data pre-view first 5 rows(part 3)

	title	Excel	BigQuery	Airflow	Prefect
0	Senior Manager, ...	1	0	0	0
1	Business Intelli...	0	0	0	0
2	HAZARDOUS SUBSTA...	1	0	0	0
3	PMS 378 Senior T...	1	0	0	0
4	COMPENSATION ANA...	1	0	0	0

## Observations and Features

This section begins the detailed exploration of the dataset's structure. I will examine the characteristics of each column to ensure data integrity and understand the available information.

Specifically, for categorical features (like `title`, `company_name`, and `company_industry`), I'll identify the distinct categories present and count the number of unique observations in each. This step confirms the data types and prepares us for subsequent analysis. Below you can see the breakdown of each type.

Analysed tools:

```
['Excel', 'Google_Sheets', 'Fivetran', 'Airbyte', 'dbt', 'Snowflake', 'BigQuery', 'Airflow']  
['Prefect', 'Power_BI', 'Tableau', 'Looker', 'Git', 'Docker', 'Kubernetes', 'Terraform']  
['AWS', 'Azure', 'GCP', 'Databricks', 'Kafka', 'Spark', 'Monte_Carlo']
```

Categorical features:

```
['title', 'company_name', 'company_industry', 'state', 'broad_industry_group']
```

Date features:

```
['created_at']
```

Job titles:

```
['manager', 'engineer', 'analyst', 'scientist', 'developer']
```

In total 23 tools are present in the data set. The categorical information about the jobs data is `title`, `company_name`, `company_industry`, `state`, `broad_industry_group`. The date when the job posting was created is in the `created_at` column. And if the particular job posting is listed for a manager, engineer, analyst, scientist, or a developer is noted in the corresponding columns.

NOTE: Keep in mind that the same job posting can be in multiple job title categories, like scientist and analyst, as well as single job posting could require experience with multiple tools.

## Outliers

NOTE:

Outliers were mainly handled in Appendix 2. Please refer to the corresponding notebook/report for more details. The function used for verifying the meaning of the words like "excel" or "airflow" utilizes the Ollama Large Language Model (LLM) to identify the meaning of the given word and decide if it is a tool or just a word. This means the categorization is not completely precise. This was necessary because the word "excel" could either be an English verb or the name of a program, also some tools might have some false positives, due to the typos in the job descriptions like the word "perfect" is misspelled to "prefect", thus LLM was employed to avoid these cases.

From the tables Table 5 and Table 6 we see there is not so much additional improvements we could do, we perhaps could fix some typos in the data but that will not be so crucial for this analyses.

Table 5: Description of the categorical data

	title	company_name	company_industry	state	broad_industry_group
count	596962	591502	527733	444562	596962
unique	272068	64081	393	54	11
top	Financial Analyst	Jobs via Dice	Software Develop...	California	Tech, Data & Tel...
freq	3299	27664	94613	63790	226571

From this we can see that there are 54 states in US, which includes states and US territories.

Table 6: Description of the date when job posting was created data.

	created_at
count	596962
mean	2025-07-13 16:26...
min	2025-01-01 00:14...
25%	2025-05-01 00:38...
50%	2025-07-24 07:01...
75%	2025-10-06 09:19...
max	2025-12-19 04:24...

## Exploratory Data Analysis

This section presents the Exploratory Data Analysis (EDA) of the 2025 US job market for tech jobs. The goal is to identify patterns and trends within the job postings. The analysis begins with a general overview of the dataset, including the total number of observations and key features. This initial phase identifies the most common tools, the primary industries hiring for tech roles, and the organizations with the highest volume of postings. We also examine the timing of these posts to identify potential hiring seasons and pinpoint which geographic states are leading in tech employment.

### Distribution of Features

Understanding the foundational characteristics of the data is the first step. This part of the report covers:

- Tool Popularity: An absolute and percentage-based ranking of the 23 identified tools.
- Industry and Company Presence: Identification of the sectors and specific employers driving the most activity.
- Temporal and Geographic Trends: A look at hiring cycles throughout 2025 and the states with the highest density of tech opportunities.

### Distribution of the Mentions of tools

Let's see which tools were the most popular in 2025 (Figure 1). We can see that Excel, AWS, Azure, Git and GCP takes the top 5, for the full list see the Table 7. The comparison of each category can be found below.

1. Data Storage & Infrastructure
  - Cloud Data Warehouses: "Snowflake" is mentioned 2.8 times more than "BigQuery".
  - Cloud Platforms (General): "AWS", is mentioned 1.2 times more than "Azure", and 2.1 times more than "GCP".
  - Big Data Storage/Processing: "Spark" is mentioned 1.5 times more than "Databricks".
2. Data Ingestion & Transformation - These tools move data from source to destination and clean it up for use.
  - Ingestion (ELT): "Fivetran" is mentioned 4.3 times more than "Airbyte"
  - Transformation: "dbt"
  - Streaming: "Kafka"
3. Business Intelligence (BI) & Analytics - The "face" of the data where executives see the results.
  - Modern BI: "Power\_BI", is mentioned 1.2 times more than "Tableau", and 7.7 times more than "Looker"
  - Spreadsheets: "Excel" is mentioned 48 times more than "Google\_Sheets".
4. Orchestration & Observability - The "management" layer that ensures everything runs on time and the data is accurate.
  - Orchestration: "Airflow" is mentioned 14.4 times more than "Prefect"
  - Observability (Quality): "Monte\_Carlo"
5. DevOps & Developer Experience - Tools that ensure code is versioned and software is scalable.
  - Version Control: "Git"
  - Containerization: "Docker" are mentioned approximately equally "Kubernetes"
  - Infrastructure as Code: "Terraform"

Table 7: The most mentioned tools in 2025.

	count	percentage
Excel	218122	36.54
AWS	173477	29.06
Azure	143179	23.98
Git	122983	20.60
GCP	81459	13.65
Kubernetes	77844	13.04
Power_BI	74677	12.51
Docker	74069	12.41

Table 7: The most mentioned tools in 2025.

	count	percentage
Tableau	61319	10.27
Terraform	44385	7.44
Spark	38985	6.53
Snowflake	34179	5.73
Kafka	30023	5.03
Databricks	26292	4.40
Airflow	13499	2.26
BigQuery	12413	2.08
Looker	9767	1.64
dbt	9397	1.57
Google_Sheets	4564	0.76
Fivetran	2032	0.34
Monte_Carlo	1411	0.24
Prefect	939	0.16
Airbyte	476	0.08

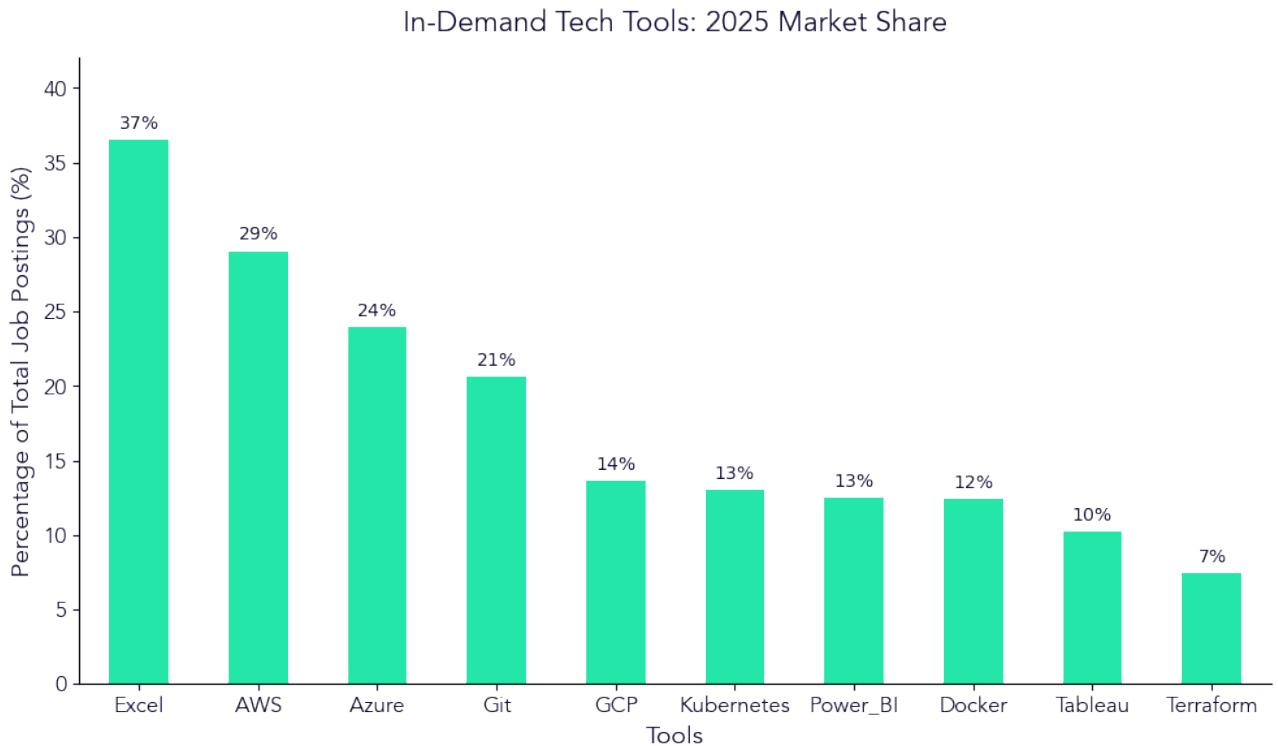


Figure 1: The most mentioned tools in 2025.

## 1. Data Storage & Infrastructure

- AWS is the Uncontested Leader: With nearly 29% of analyzed mentions, AWS is more popular than Azure (24%) and GCP (14%). If a company is hiring for the cloud, they are likely hiring for AWS.
- Big Data Processing vs. Warehousing: Spark (7%) and Snowflake (6%) are neck-and-neck. This suggests that the market is split between companies doing heavy engineering (Spark) and those preferring managed data warehousing (Snowflake).
- BigQuery vs. Snowflake: Within the “Cloud Warehouse” niche, Snowflake is significantly more required (6%) than BigQuery (2%). This highlights Snowflake’s success as a platform-agnostic solution that works across all clouds.

## 2. Business Intelligence (BI) & Analytics

- Excel is Still the Tool of Choice: Despite the rise of “Advanced Analytics,” Excel appears in 37% of analyzed tech-related job posts. It remains the universal language of business, outperforming specialized BI tools like Power BI (13%) and Tableau (10%) by a factor of 3 to 1.

- Power BI Pulls Ahead of Tableau: In the battle of professional visualization, Power BI (13%) has overtaken Tableau (10%). This is likely driven by the deep integration of Power BI into the Microsoft 365 ecosystem, making it a “default” choice for many enterprises.

### 3. DevOps & Engineering

- Git is Mandatory: Mentioned in 21% of jobs, Git is no longer a “bonus” skill; it is a fundamental requirement for anyone working with code or data.
- Kubernetes vs. Docker: Kubernetes (13%) is slightly more popular than Docker (12%). This signals that companies aren’t just looking for people who can build containers, but for those who can manage them at scale in production.
- Infrastructure as Code: Terraform (7%) is the clear leader for managing cloud environments, appearing more frequently than many actual data tools.

### 4. Data Ingestion & Transformation

- The Power of Kafka: For real-time data, Kafka (5%) is the dominant force. It is significantly more popular than specialized ingestion tools like Fivetran or Airbyte.
- The dbt vs. Fivetran Gap: dbt (2%) is more required than Fivetran (0.34%). This suggests that while companies might use various ways to move data, they are increasingly standardizing on dbt to transform it once it arrives in the warehouse.

### 5. Orchestration & Observability

- Airflow Dominates Orchestration: Airflow (2%) is the gold standard for scheduling data jobs, vastly outperforming newer competitors like Prefect (0.16%).
- The Rise of Data Quality: Monte Carlo (0.24%) appearing in the data at all is a significant trend. It shows that “Data Observability” (ensuring data isn’t broken) is transitioning from a luxury to a requirement in 2025.

## Distribution of Job Titles by Job Type

In this section, we examine the distribution of job postings across primary professional categories. Based on the data visualized in Figure 2, there is a clear and substantial demand for technical roles, particularly within the engineering domain.

Engineers represent the majority of the market share at 51%, significantly outpacing other roles. Developers and Analysts follow with 28% and 12% accordingly. Specialized roles such as Scientists and Managers constitute the remaining 4% and 5% of the postings, respectively. These findings suggest that for the 2025 hiring landscape, companies are prioritizing the foundational technical infrastructure provided by engineering talent.

```
engineer      302026
analyst      167770
developer     69526
manager       28194
scientist     21453
dtype: int64
```

## Distribution of Company Names

In this section, we examine the leading organizations driving the demand for technical talent in the United States. While the tech landscape is vast and diverse, a small group of industry giants and specialized platforms accounts for a significant portion of total hiring activity.

From Figure 3, we can observe that Job via Dice maintains a commanding presence, representing 5% of the entire tech job market in the USA. This high volume highlights the platform’s role as a primary aggregator for technical specialized roles.

When analyzing the “Big Tech” sector specifically, Amazon, Microsoft, Google, and Apple all appear within the top 20 hiring entities. This data suggests that while these tech behemoths are influential, the market remains highly fragmented, with a significant amount of hiring distributed across thousands of mid-sized firms and diverse industry sectors.

## Distribution of Industries

In this section, we analyze the distribution of technology-related roles across various economic sectors. Understanding where demand originates provides critical context for the current hiring landscape and identifies the primary

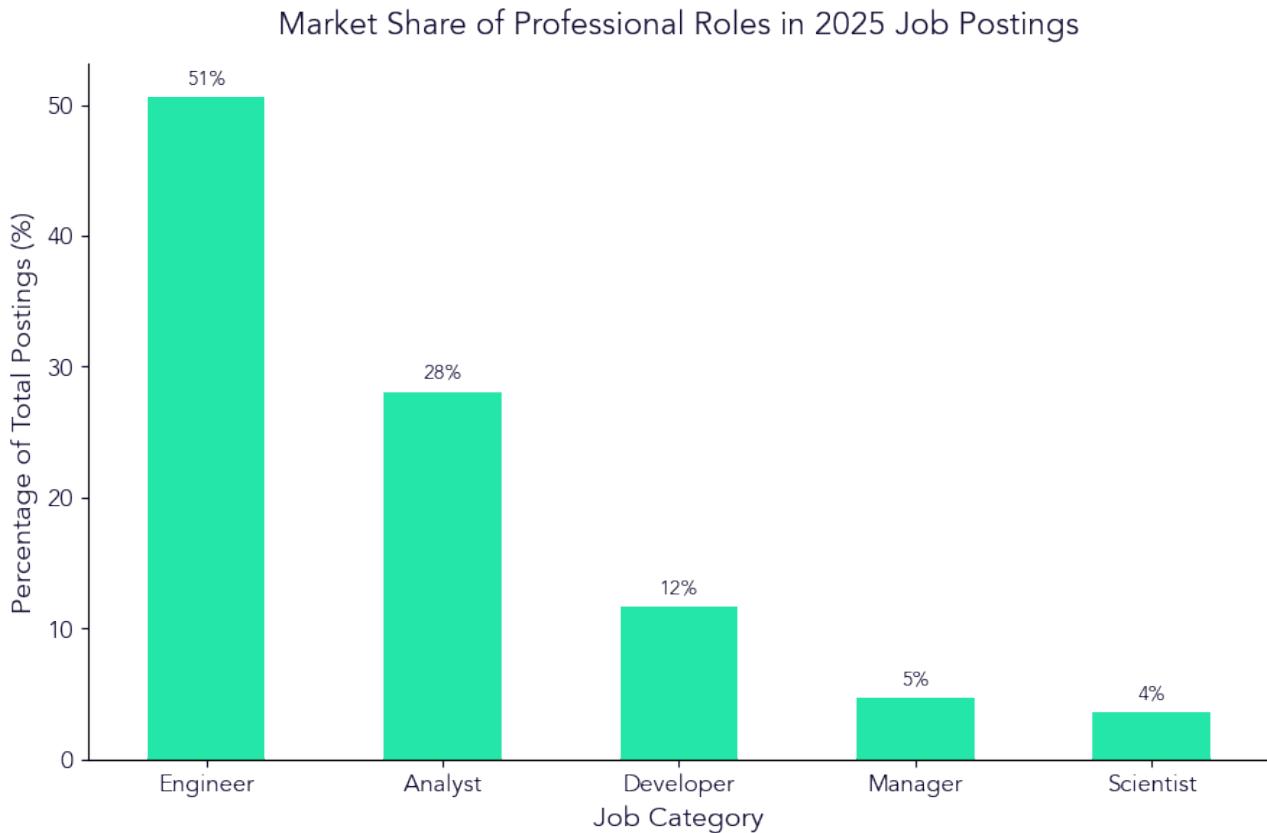


Figure 2: Distribution of job titles in 2025 in job postings.

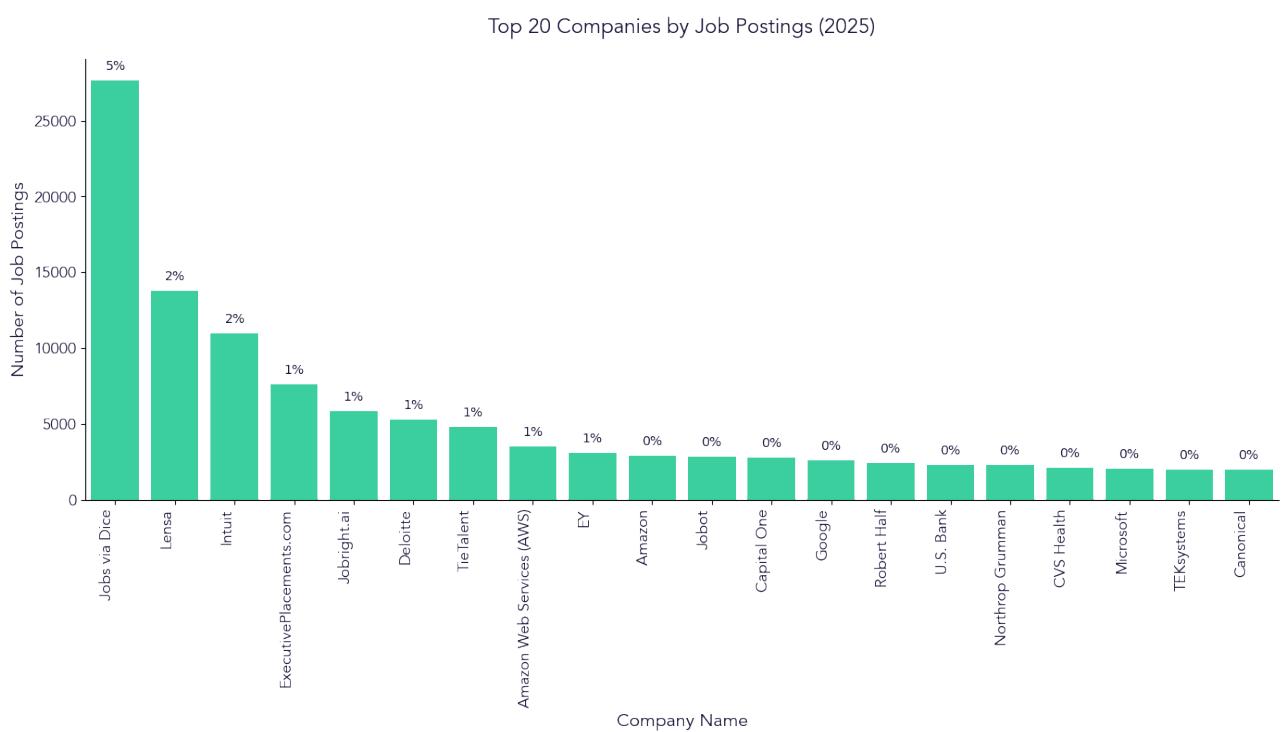


Figure 3: Barplot of the most mentioned companies in 2025 in job postings.

drivers of technical growth.

Based on the data visualized in Figure 4, the Tech, Data & Telecom industry remains the dominant force, accounting for 38% of analyzed professional requirements. This sector continues to outpace all others, reinforcing its role as the primary engine for tech employment.

Outside of the core technology sector, the demand is significantly more distributed. Professional, Legal & Business Services follow with a 15% market share, while Manufacturing, Industrial & Defense represents 9% of the postings. The Finance, Insurance & Real Estate (FinTech) sector accounts for 10%, highlighting a stable need for technical expertise in modernizing financial infrastructure. Notably, all other industries each represent less than 4% of the market share (each), indicating a high concentration of tech talent within the top four sectors.

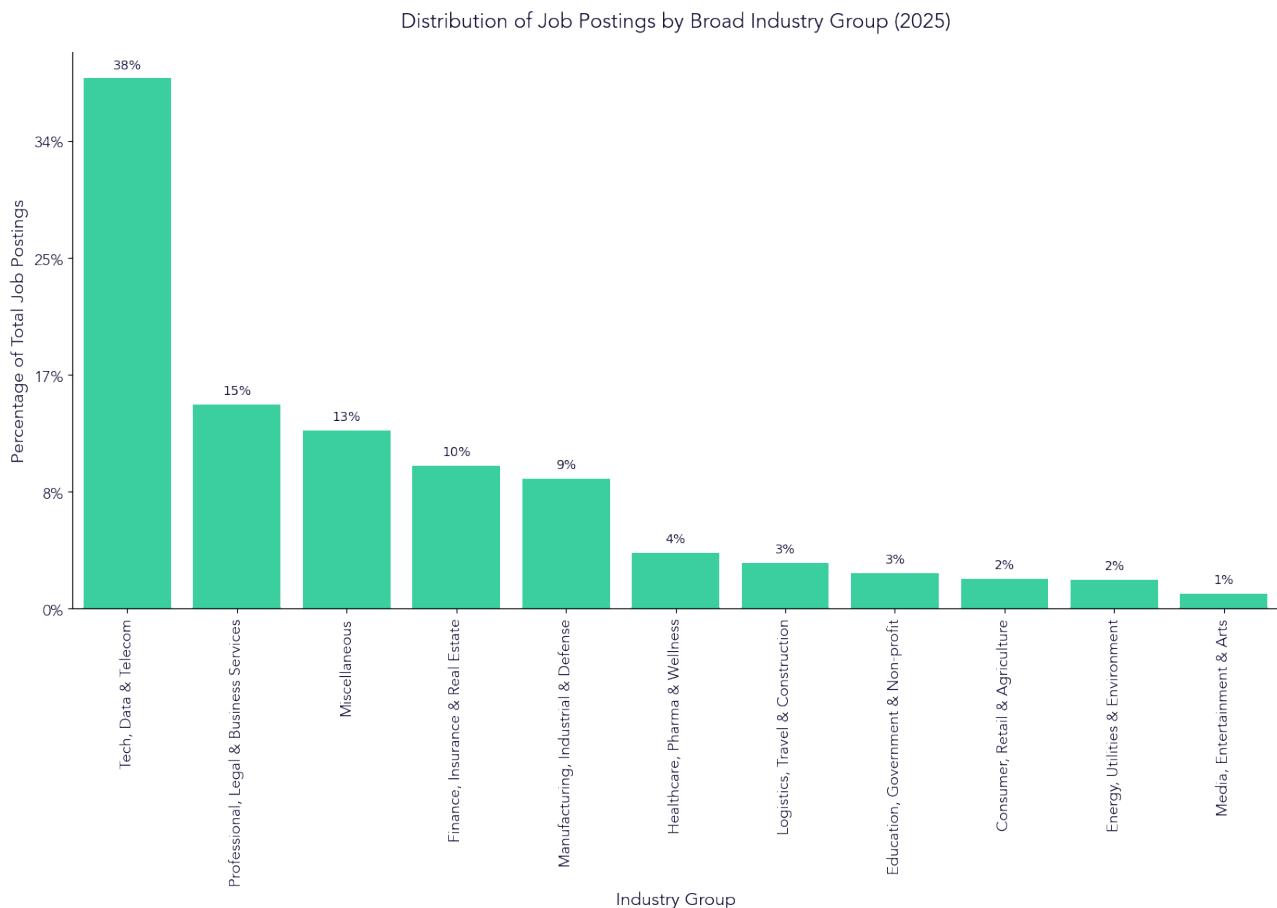


Figure 4: Barplot of the most mentioned broad industries in 2025 in job postings.

Below @dist-company-industries you can see the raw distribution of the top 20 industries that are not cleaned, just for the sake of understanding the raw data. Software development and IT industries are absolute leaders.

### Distribution of States

In this section, we analyze the geographic distribution of tech talent demand across the United States. Identifying these "hiring hubs" allows us to understand the regional concentrations of the digital economy and where companies are focusing their recruitment efforts.

Based on the data visualized in Figure 6 and Figure 8, California remains the primary driver of technical employment, accounting for 11% of analyzed job postings. Texas(8%) follows closely as the second-largest market, reinforcing its status as a significant and growing center for technology and innovation. New York has the third-largest market taking up 5% of analyzed postings. Together, these three states represent a substantial portion of the national demand.

For a more granular view of regional trends, an interactive heatmap is available in the project files at [outputs/figures/tools/us\\_](#) or [html](#) version of this report. This interactive asset allows for state-by-state comparisons and provides specific market share details for each territory.

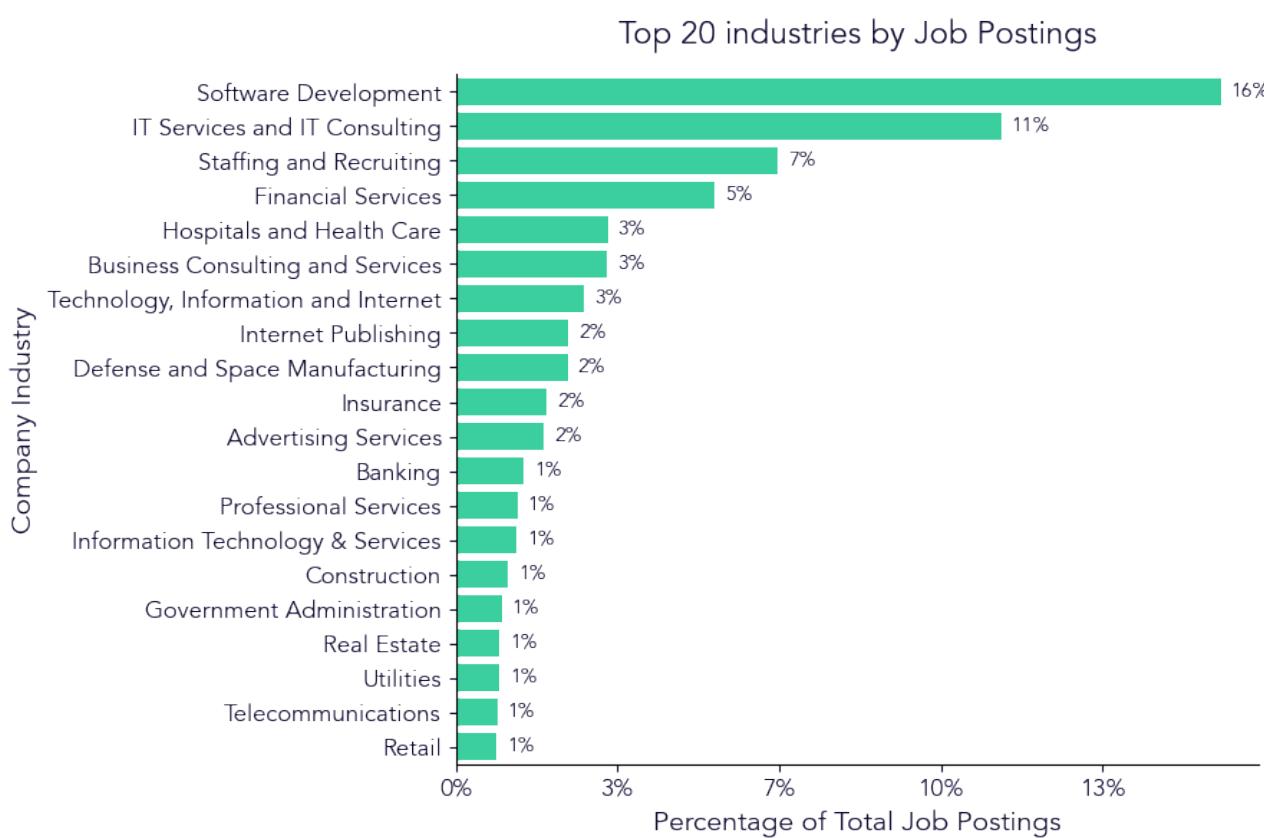


Figure 5: Barplot of the most mentioned raw company industries in 2025 in job postings.

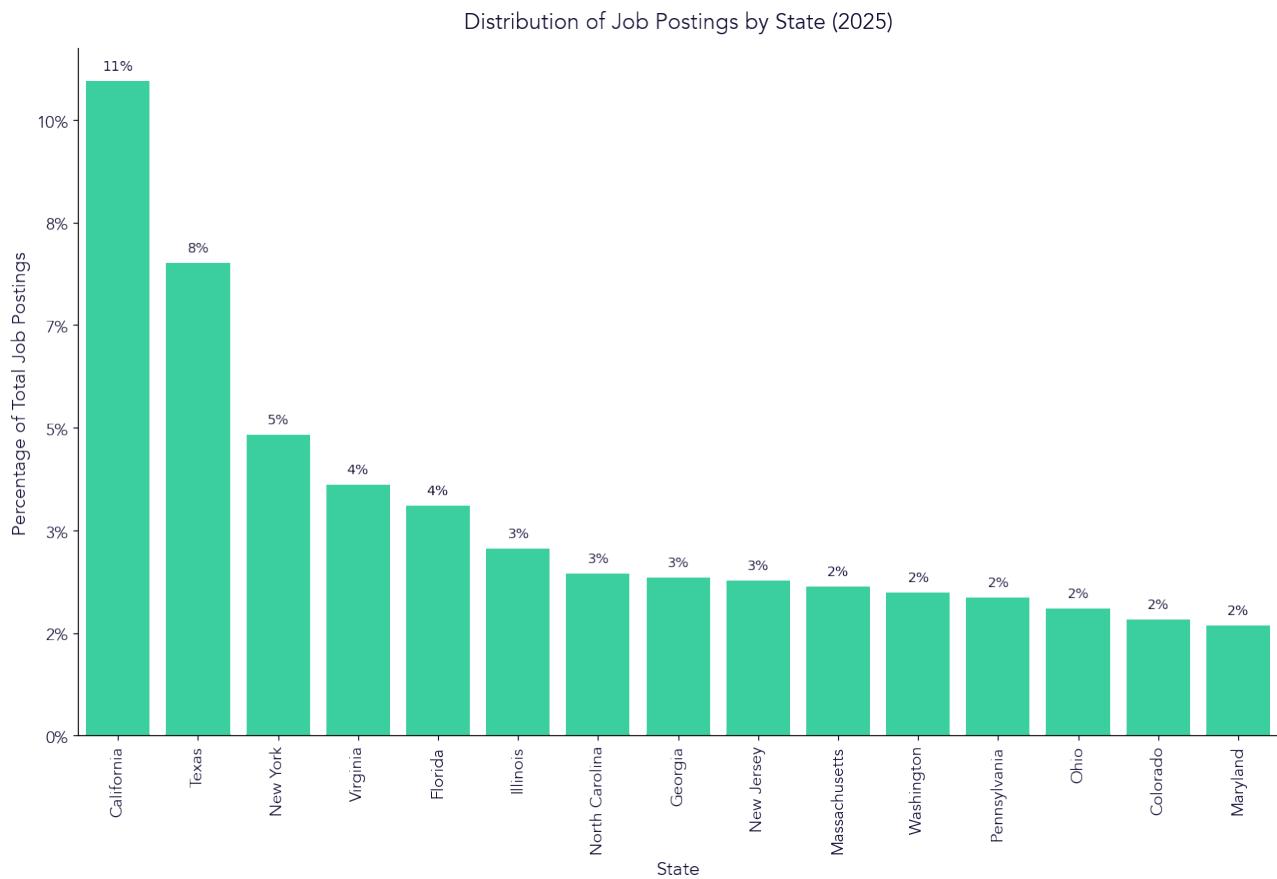


Figure 6: Distribution of job postings by state in 2025.

Unable to display output for mime type(s): text/html

(a) Distribution of job postings by state in 2025.

Unable to display output for mime type(s): text/html

(b)

Unable to display output for mime type(s): text/html

(c)

Figure 7

#### 2025 US Job Market: Regional Demand Distribution

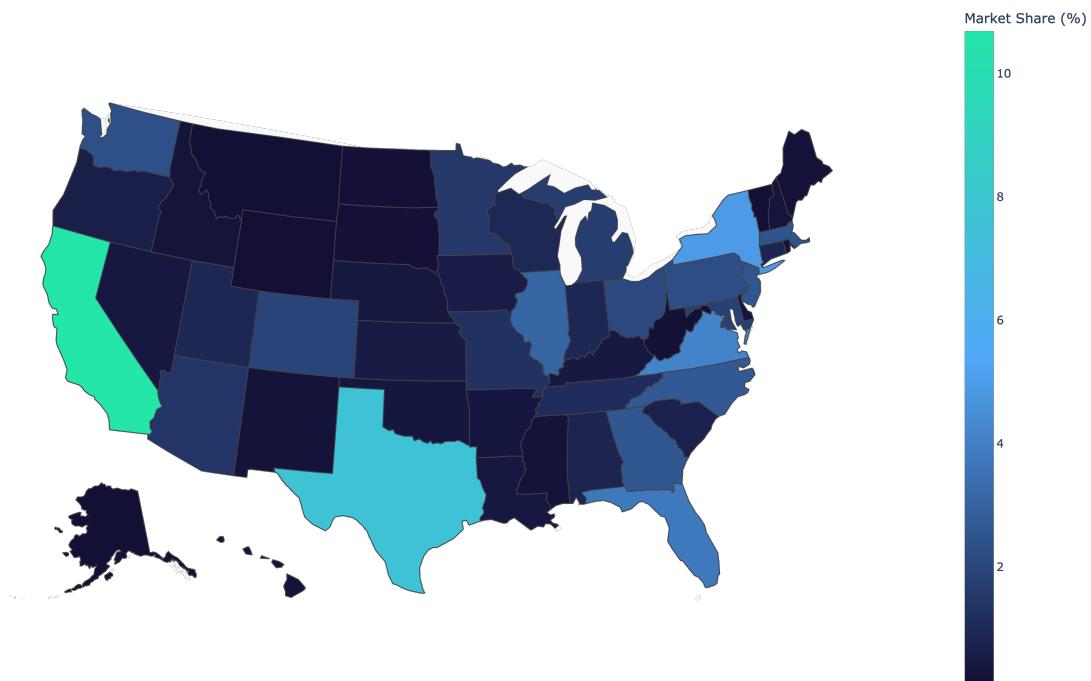


Figure 8: 2025 US Job Market. Regional Demand Distribution

## Distribution of Job Posting Times

In this section, we analyze the seasonal variations in hiring activity throughout 2025. Based on the data visualized in Figure 9, the hiring landscape exhibits a distinct “double peak” pattern. The highest volume of activity occurs in August (12%) and November(13%). These surges likely correspond to the conclusion of the summer period and the finalization of year-end technical projects, respectively.

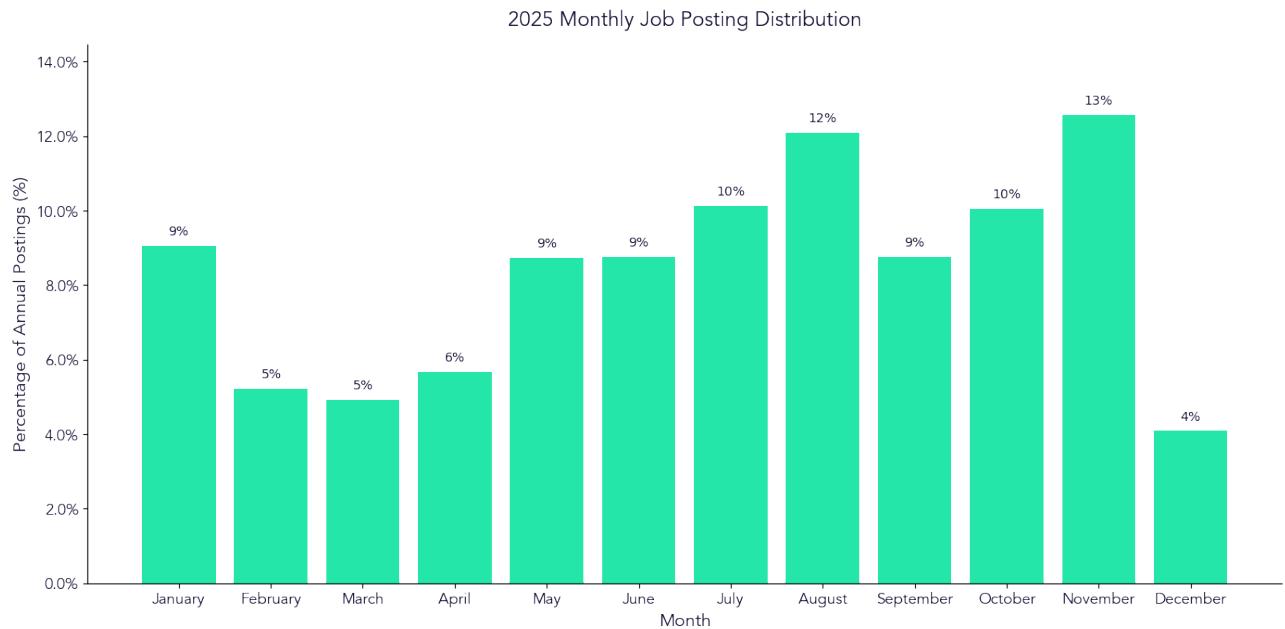


Figure 9: Distribution of job postings by month in 2025.

## Tool Demand by Category

After establishing the baseline, the analysis investigates how tool requirements vary across different contexts. This section explores the relationship between tools and:

- Job Titles: Which tools are synonymous with roles like “Data Analyst” or “Data Engineer.”
- Tools Duos: Which tools are going together in the jobs descriptions.
- Industries: How the tech stack in “Finance” differs from “Healthcare” or “Manufacturing.”
- Geography: Regional preferences for specific technologies across the United States.

## Tool Demand by Tool Category

Since the comparison of the tools is not so straight forward and saying that Excel is more popular than AWS (like in Figure 1) seems like comparing apples and bananas, the better approach is to compare the tools with other tools from their category. Thus below you can find two plots ()

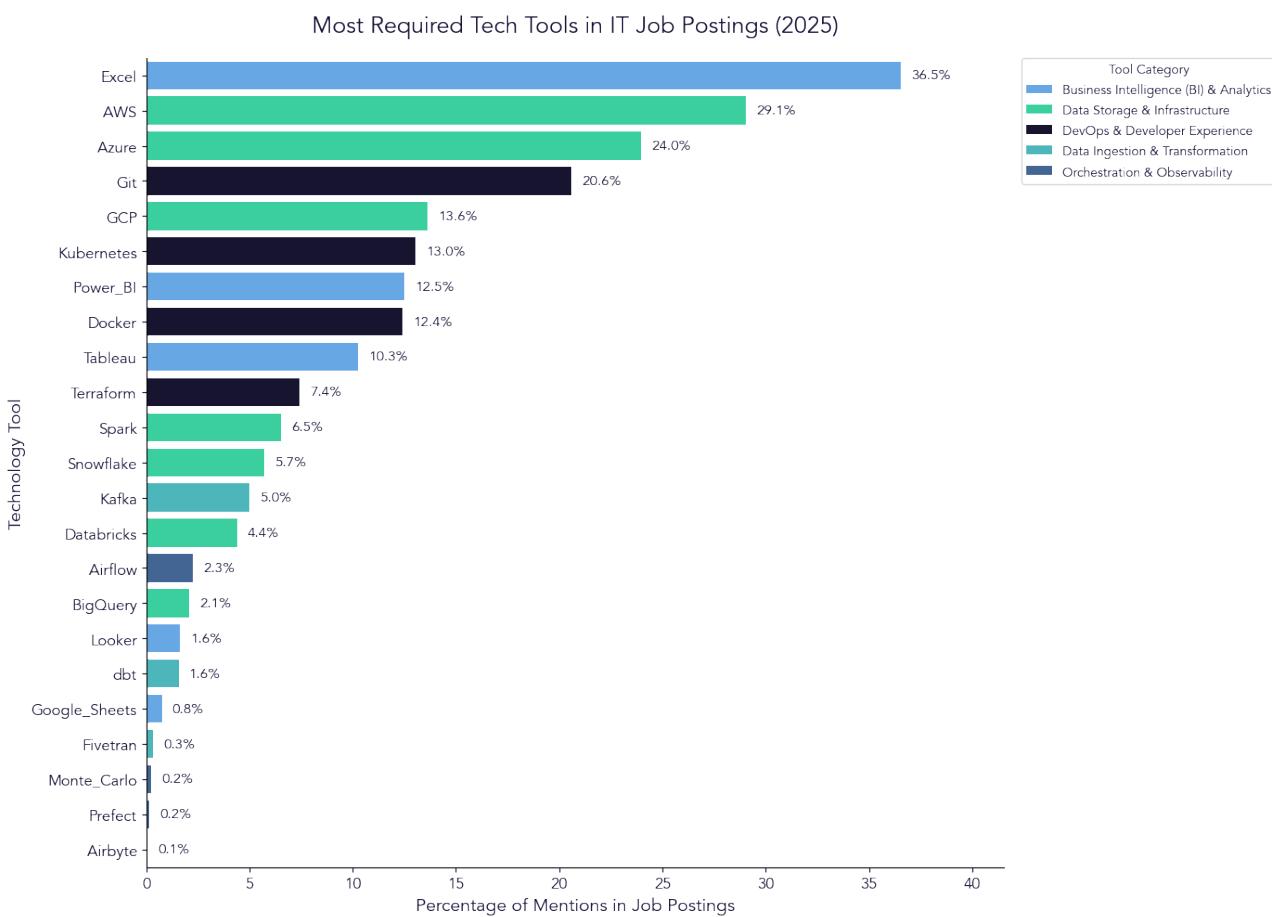
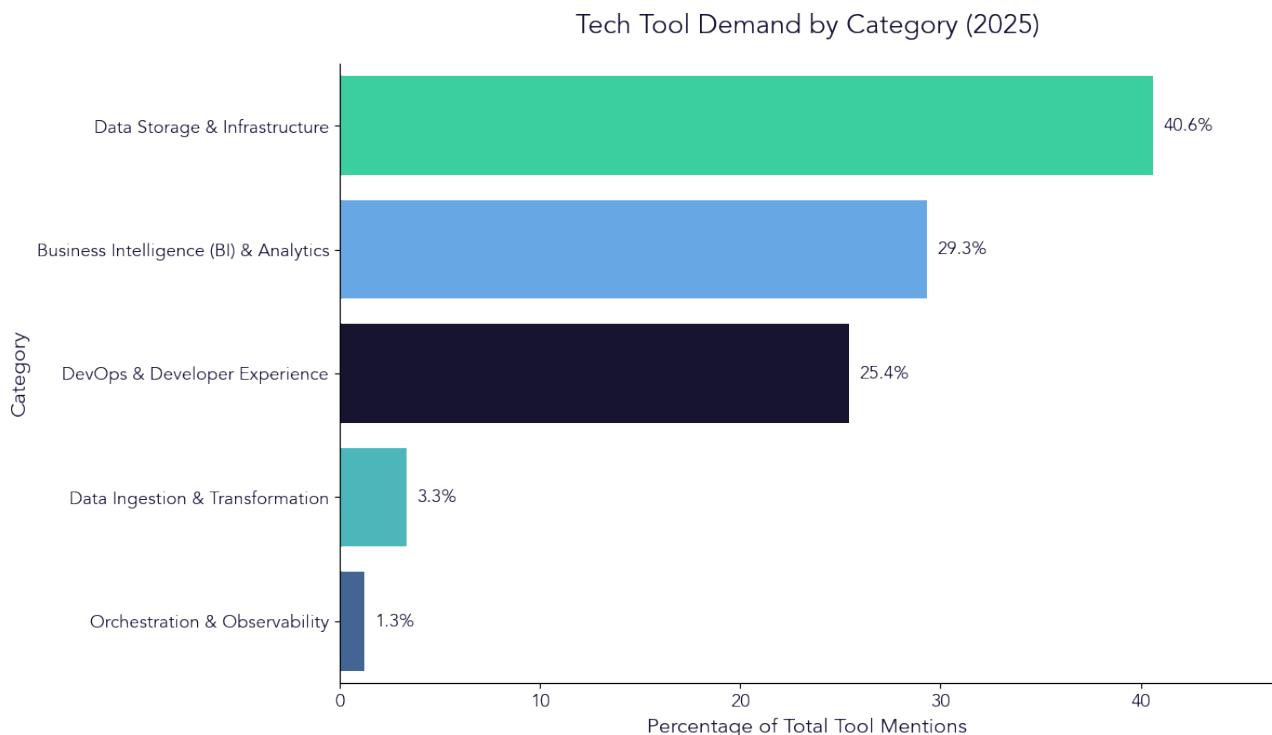


Figure 10: Distribution of job postings by month in 2025.



### Tool Demand by Job Title

In this section, I will analyze the specific technical requirements for selected roles. Understanding this data provides insight into the tools required to oversee modern technical teams and data-driven projects.

- Technical managers in 2025 are no longer just “people managers”; they are Infrastructure Strategists. The high demand for “AWS” and “Azure” indicates that managers must understand cloud unit economics and ar-

chitectural trade-offs. The presence of “Kubernetes” at 14% suggests that leadership roles now require an understanding of how to scale services, while “Excel” remains the primary tool for budget tracking and capacity planning. For a more detailed breakdown of these trends, please see Figure 11.

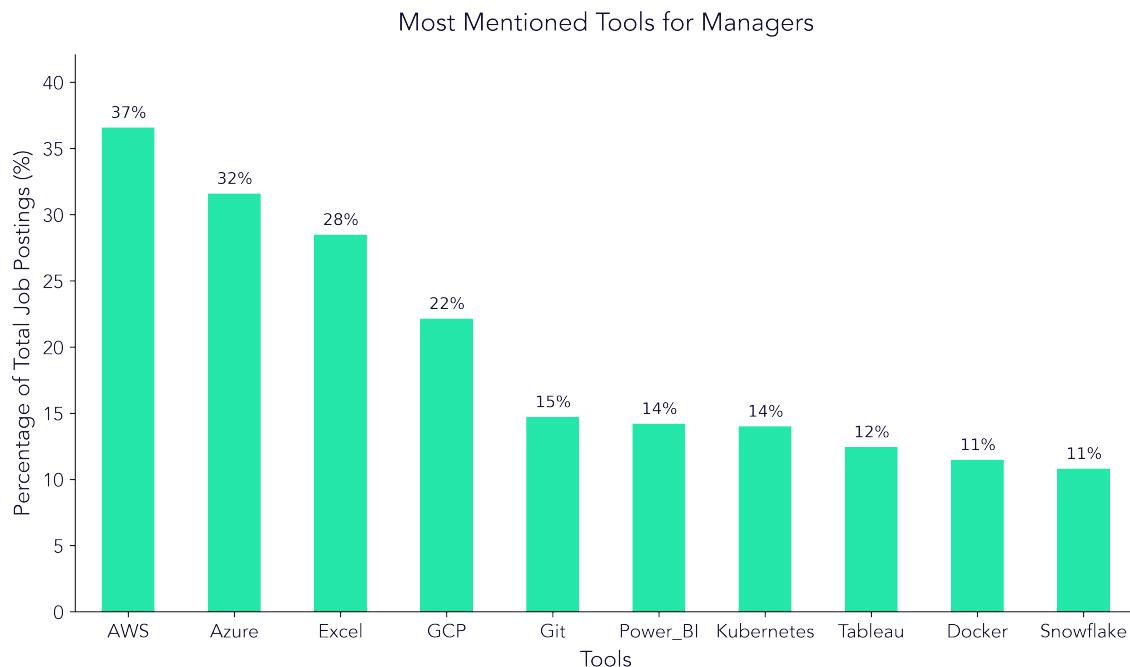


Figure 11: Programming languages demand for managers distribution

- For engineers, the stack is heavily weighted toward Deployment and Automation. With “AWS” and “Azure” leading, followed by a strong showing from “Git” and “Kubernetes”, the data indicates a “DevOps-first” engineering culture. Interestingly, “Terraform” (13%) is a critical differentiator here, showing that “Infrastructure as Code” is a core expectation for modern engineering talent. For a more detailed breakdown of these trends, please see Figure 12.

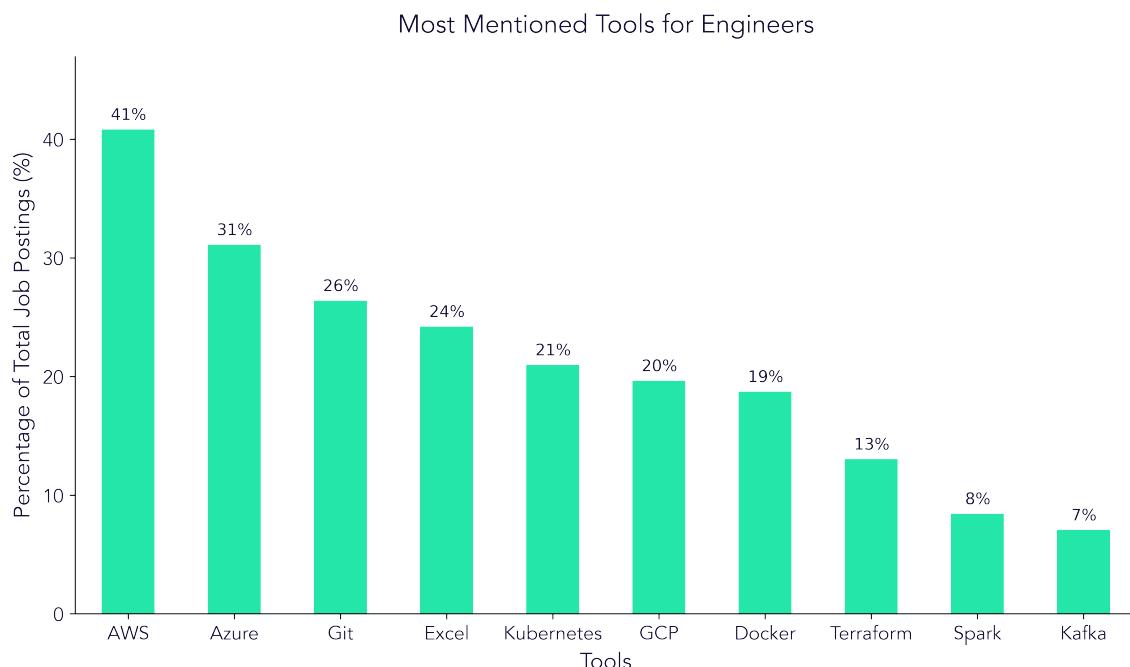


Figure 12: Programming languages demand for engineers distribution

- The analyst profile remains the most specialized and concentrated. A staggering 70% requirement for “Excel” proves it is the undisputed baseline. However, the 24% for “Power\_BI” versus 20% for “Tableau” confirms that Microsoft’s ecosystem dominance is successfully converting “Excel” users into “Power BI” developers. For analysts, the world is almost entirely about visualization and reporting. For a more detailed breakdown of these

trends, please see Figure 13.

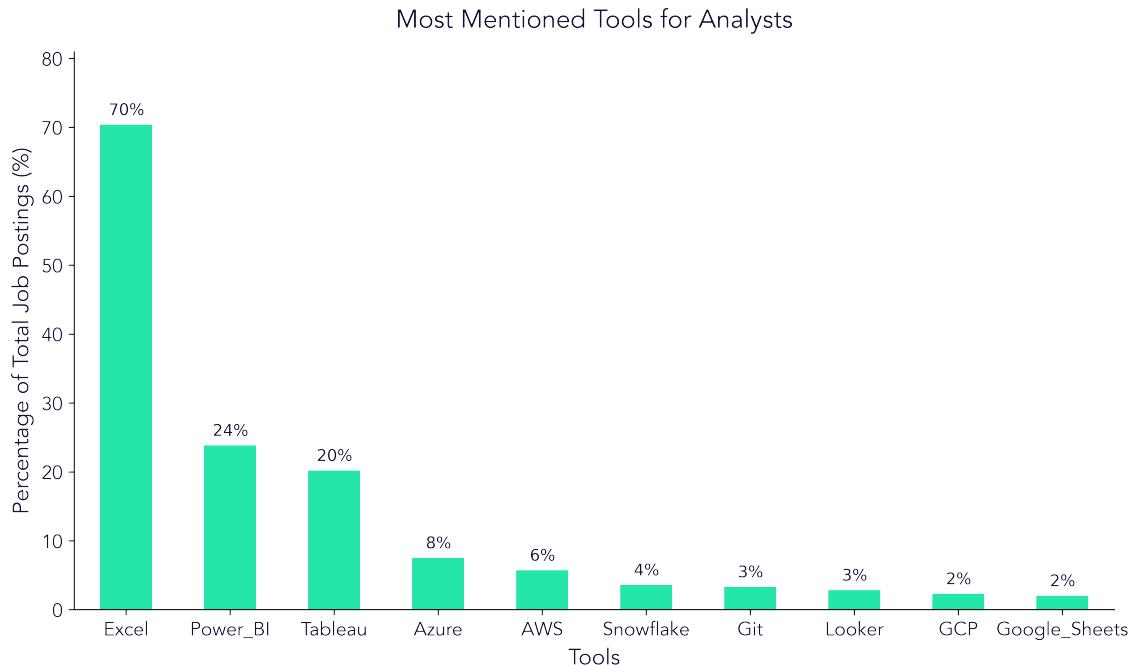


Figure 13: Programming languages demand for analysts distribution

- Data Scientists have the most diverse toolset, bridging the gap between analytics and engineering. The high demand for "Spark" (21%) and "Databricks" (10%) indicates a focus on "Big Data" processing that Analysts typically don't touch. The nearly equal split between "Tableau" (25%) and "Power\_BI" (14%) shows that Scientists still lean toward "Tableau" for complex, research-oriented visualizations. For a more detailed breakdown of these trends, please see Figure 14.

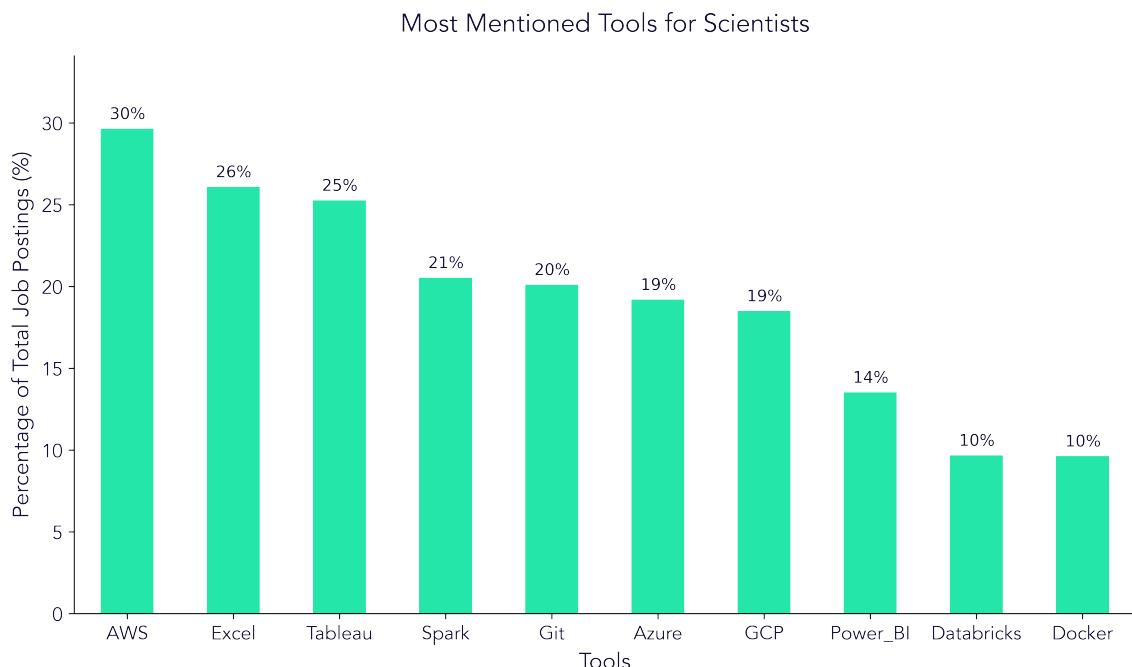


Figure 14: Programming languages demand for scientists distribution

- For Developers, Code Integrity and Real-time Systems are paramount. "Git" is a requirement in nearly half of analyzed postings, the highest for any tool in this category. The mention of "Kafka" (9%) is unique to this group, indicating that developers are increasingly responsible for building the real-time data "pipelines" that the rest of the organization relies upon. For a more detailed breakdown of these trends, please see Figure 15.

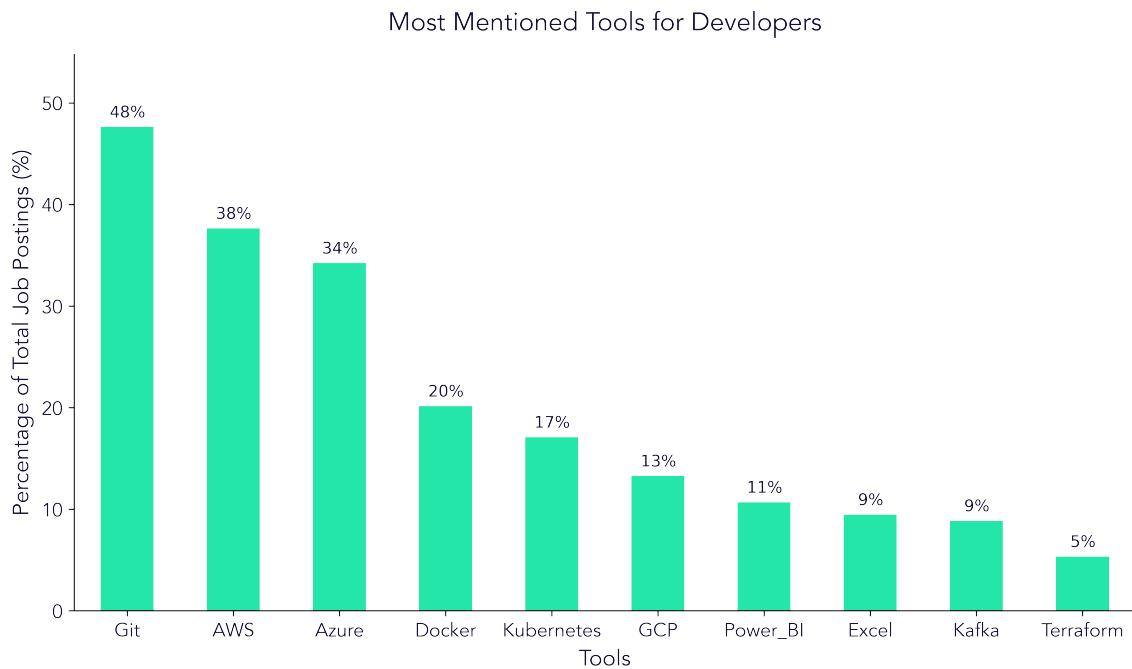


Figure 15: Programming languages demand for developers distribution

#### Tool Demand by Company Industry

In this section, I analyze whether specific patterns of tool usage emerge across different industries. Understanding these trends helps identify which technical ecosystems are maturing within specific sectors and where traditional tools like Excel still maintain a strong foothold.

From the data, it is evident that Excel remains the dominant tool across almost all industries. Its versatility for data entry, quick calculations, and reporting makes it a staple in sectors ranging from Finance to Healthcare. However, a distinct shift occurs within the Tech, Data & Telecom industry. In this sector, the demand for specialized infrastructure and collaboration tools surpasses traditional office software. The most frequently required tools in this group are AWS, Azure, and Git, reflecting a heavy organizational focus on cloud-native development and version-controlled environments.

This shift suggests that while “General Business” industries still rely on Excel as their primary analytical interface, the “Tech” sector has moved toward integrated cloud platforms as their foundational environment.

#### Tools Co-occurrence Frequency

In this section, I analyze the relationships between different technical tools by examining how often they appear together within the same job description. Identifying these pairings reveals the “tech stacks” or integrated ecosystems that employers currently favor in 2025.

Based on the data visualized in Figure 16, there is a strong overlap within cloud infrastructure providers. AWS and Azure are mentioned together in 15% of postings, suggesting a high demand for professionals capable of managing multi-cloud environments. Similarly, Google Cloud Platform and AWS appear together in 11% of jobs, while Google Cloud Platform and Azure share a 10% co-occurrence rate.

These figures indicate that while many companies choose a primary cloud provider, a significant portion of the market requires familiarity with multiple platforms to support redundant or hybrid cloud strategies.

#### The Most in Demand Cloud Platforms by State

In this section, I explore which cloud platforms hold the dominant position within individual states. By identifying the most frequently requested cloud platform in each region, we can map out the technical “territories” of the 2025 US job market.

According to the data visualized in Figure 18, AWS is the most required cloud platform in most of the states. Specific regions with high concentrations of infrastructure and federal operations show a pivot toward specialized cloud tools.

## Tech Tool Co-occurrence in 2025 Job Postings

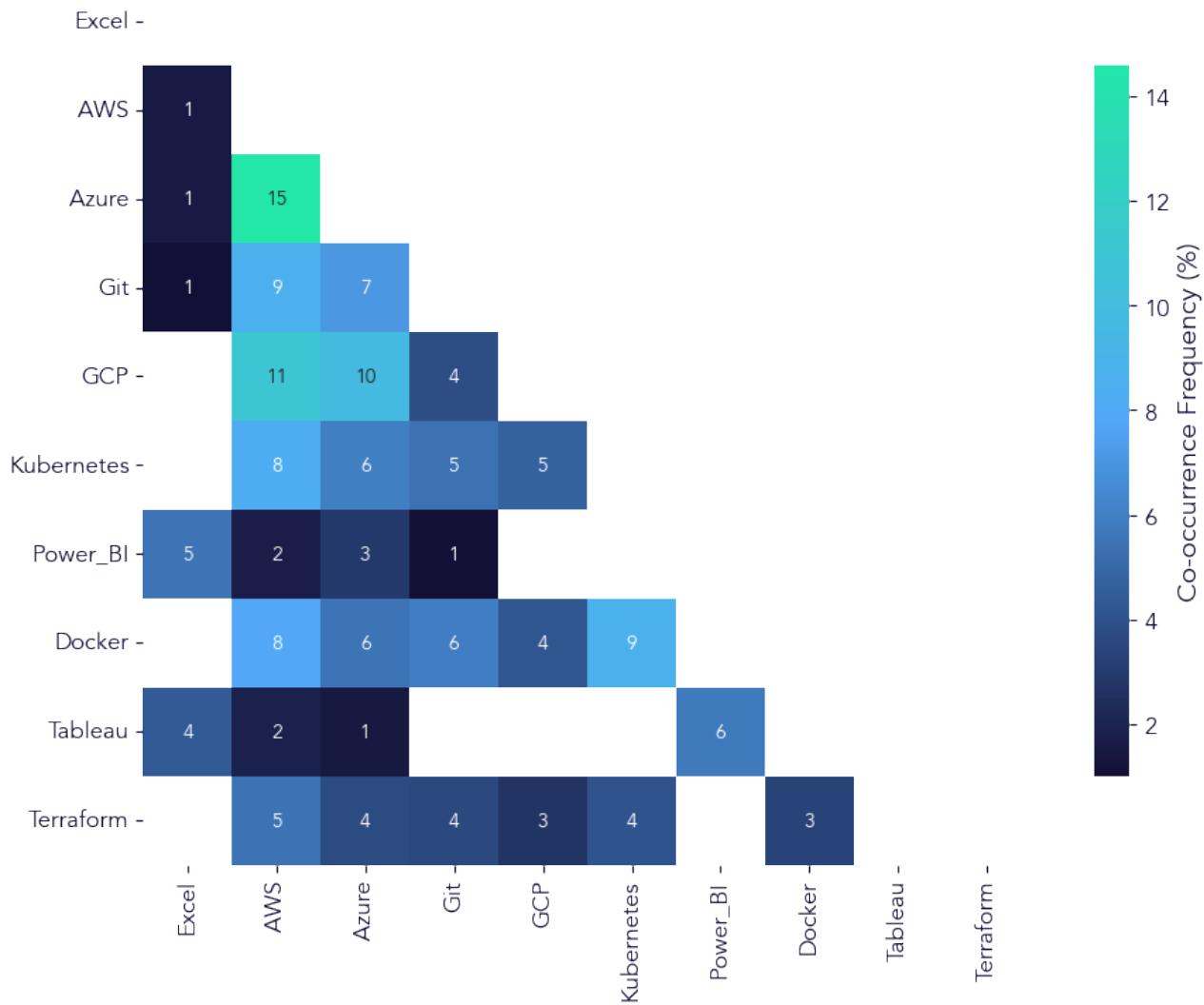


Figure 16: Tool co-occurrence heatmap for 2025.

Unable to display output for mime type(s): text/html

Figure 17: Most popular tools by state in 2025.

## 2025 Most Mentioned Cloud Platforms by State

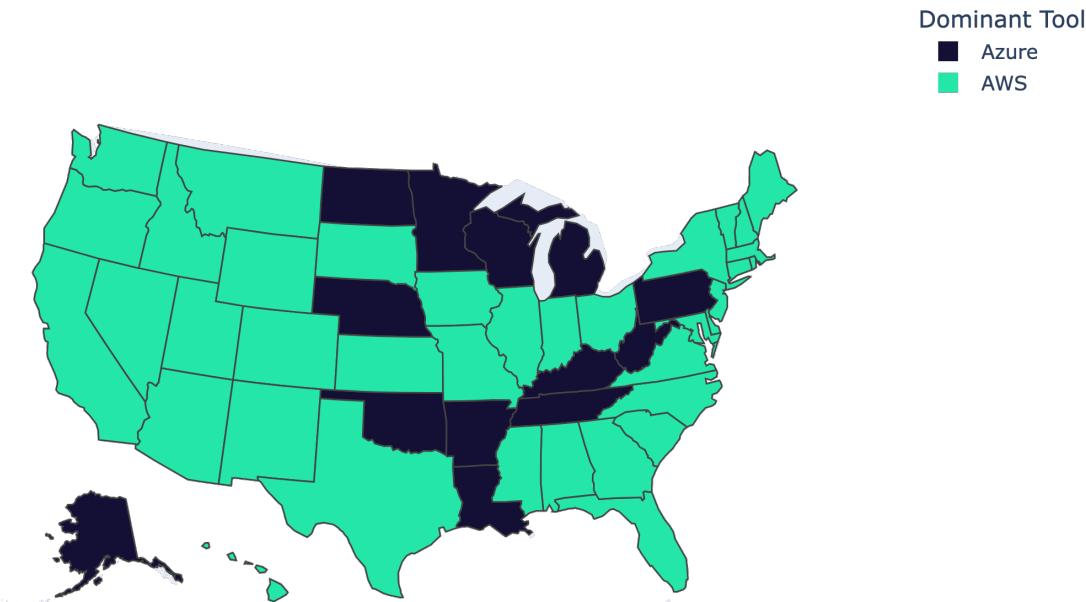


Figure 18: Most popular tools by state.

### Market Leader Deep Dive: FAANG and MANGO

The final part of the analysis focuses on the industry giants that often set the trend for the rest of the market. We compare two distinct groups:

**FAANG** (Meta, Apple, Amazon, Netflix, Google) These established leaders provide a benchmark for high-scale, mature tech environments. We analyze their core tool requirements to see which traditional technologies remain dominant.

**MANGO** (Microsoft, Apple, Nvidia, Google, OpenAI/Anthropic) The MANGO index represents the architects of the current AI era. This section highlights the tool preferred by companies at the forefront of hardware innovation and LLM development. Comparing these groups reveals whether AI-centric firms are pivoting toward newer tools faster than the broader market.

#### Key Differences

**Focus:** FAANG was about digital advertising, e-commerce, and content streaming. MANGO is about generative AI, cloud infrastructure, and AI hardware.

**Context:** While FAANG focused on user attention, MANGO focuses on AI reasoning and intelligence.

#### FAANG

In this section, we examine the hiring patterns of the world's most influential technology giants. These organizations often set the standard for technical requirements and recruitment volume across the global IT sector.

Based on the data in Figure 19, Amazon emerges as the clear leader in recruitment volume among the FAANG group, accounting for 62% of analyzed job postings within this elite cohort. Google follows with 24%, while Apple represents 8% of the demand. Meta and Netflix round out the group with 5% and 2% of the market share, respectively. This distribution highlights Amazon's expansion and its significant role as a primary employer for technical talent.

The technical stack required by these organizations (see Figure 20) reflects their heavy investment in cloud infrastructure and data-driven decision-making. The high demand for AWS (53%) and Google Cloud (24%) is directly linked to the fact that Amazon and Google are the dominant employers within this dataset, essentially hiring to build and maintain their own proprietary ecosystems.

While cloud platforms lead the requirements, the data also shows a strong reliance on traditional analytical tools. Excel (15%) and Tableau (12%) remain vital for business intelligence within these firms, proving that even at the highest level of tech innovation, foundational data tools are still used for daily operations. More specialized infrastructure

tools like Spark, Kubernetes, and Terraform show lower total percentages, but they represent the critical “engine room” skills required for high-level engineering roles.

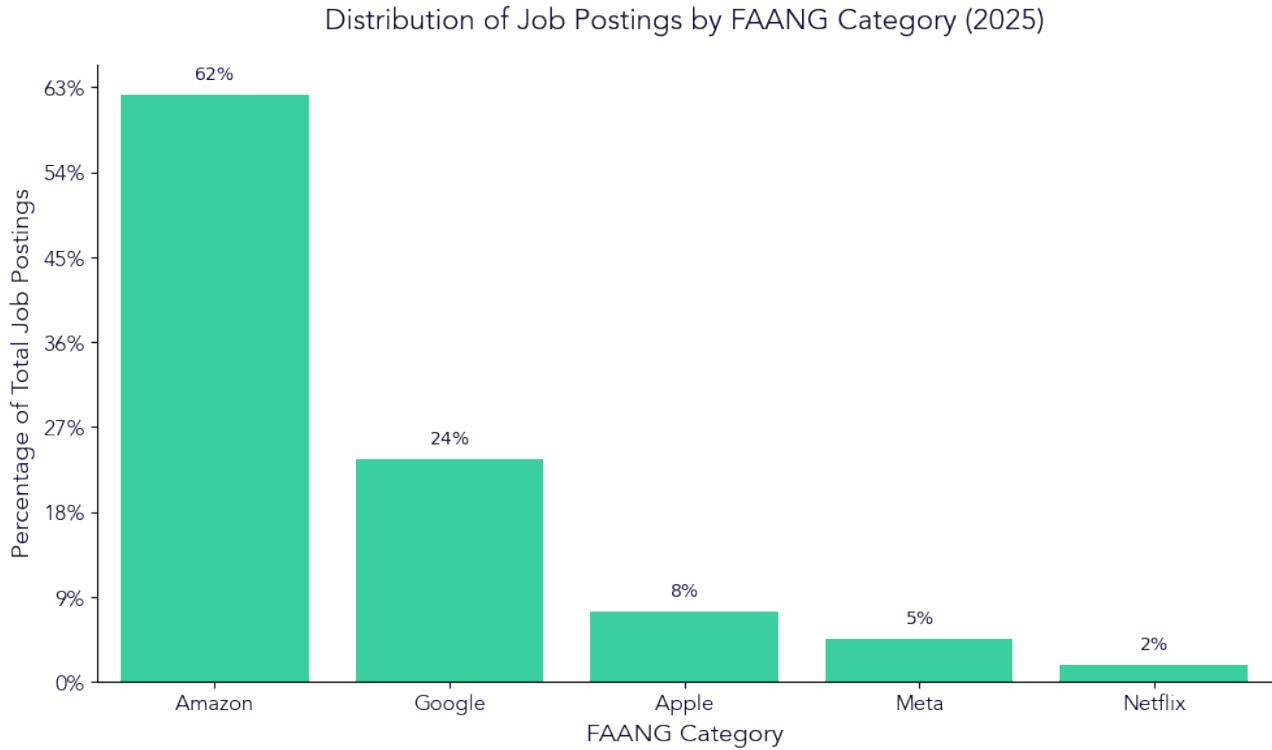


Figure 19: Distribution of job postings within the FAANG companies.

## MANGO

In this section, we pivot our focus to the “MANGO” cohort—Meta, Apple, NVIDIA, Google, and Oracle—representing a more hardware and AI-centric evolution of the traditional tech giants. Analyzing this group allows us to see how the demand for talent shifts when emphasizing cutting-edge infrastructure and silicon innovation.

According to Figure 21, Google stands as the dominant hiring force in this group, responsible for 34% of the job postings. Microsoft and Apple maintain a strong presence with 19% and 17% respectively, while NVIDIA—fueled by the ongoing AI hardware boom—accounts for 15% of the requirements. Meta occupies a smaller portion of the current hiring landscape at 3%. This distribution underscores a market where established cloud and hardware leaders are currently driving the bulk of new technical opportunities.

The tool requirements within the MANGO group (see Figure 22) show a distinct shift compared to the FAANG cohort. Most notably, GCP (41%) and Azure (30%) are the leading technologies. This is due to the fact that Microsoft and Google are central to the MANGO category, while Amazon is excluded, leading to a significantly lower presence for AWS (11%).

Beyond cloud providers, the MANGO group places a much higher premium on deployment and containerization technologies. Kubernetes (19%), Git (16%), and Docker (11%) appear with higher frequency here than in other cohorts. This suggests that the work within these companies is heavily focused on scalable infrastructure and complex DevOps pipelines, likely supporting the massive compute requirements of AI development and global cloud services.

## Key Insights

- Excel and AWS have solidified their roles as the essential tools of the 2025 technical landscape. While Excel remains the near-universal baseline for business data management (required in 37% of postings), AWS has become the “infrastructure lingua franca,” appearing in 29% of analyzed technical roles—more than Azure and GCP combined.
- The demand for Engineers is overwhelming, accounting for 51% of the market share among technical titles. This highlights a critical industry shift: companies are focusing more on building and architecting robust systems than on pure management or stand-alone data analysis.

FAANG most wanted tools for 2025

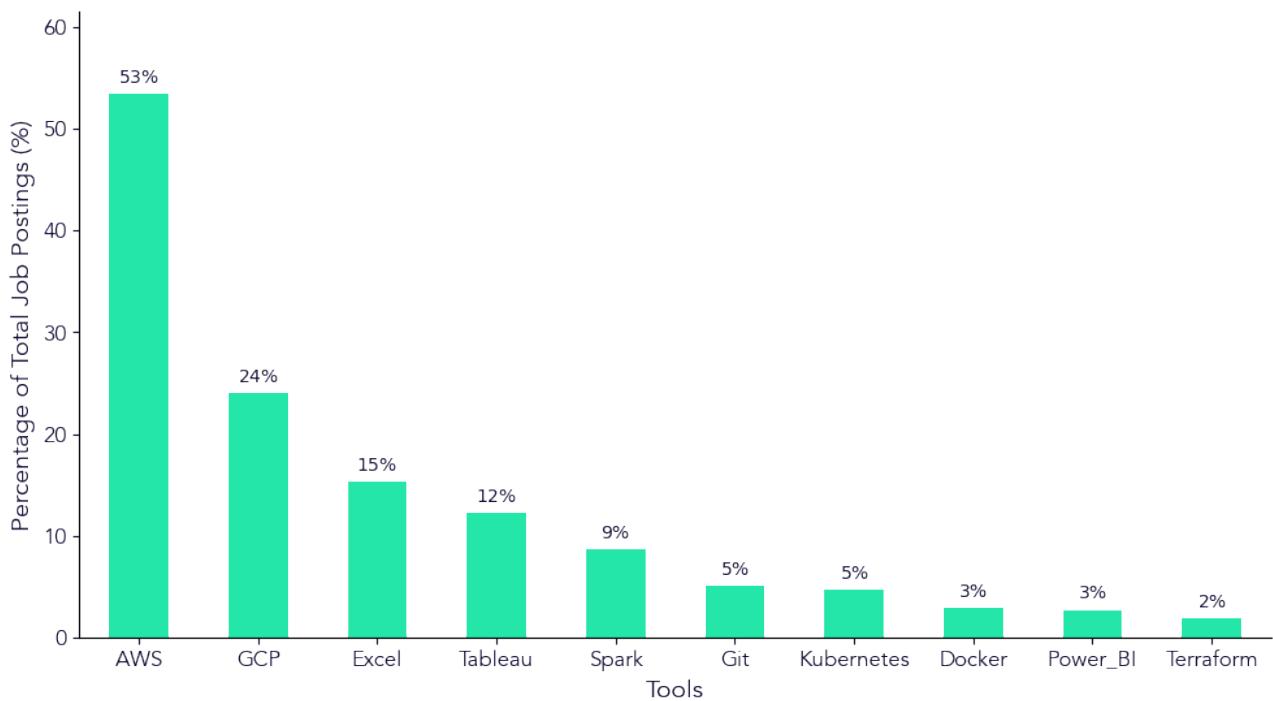


Figure 20: Distribution of demanded tools within the FAANG companies.

Distribution of Job Postings by MANGO Category (2025)

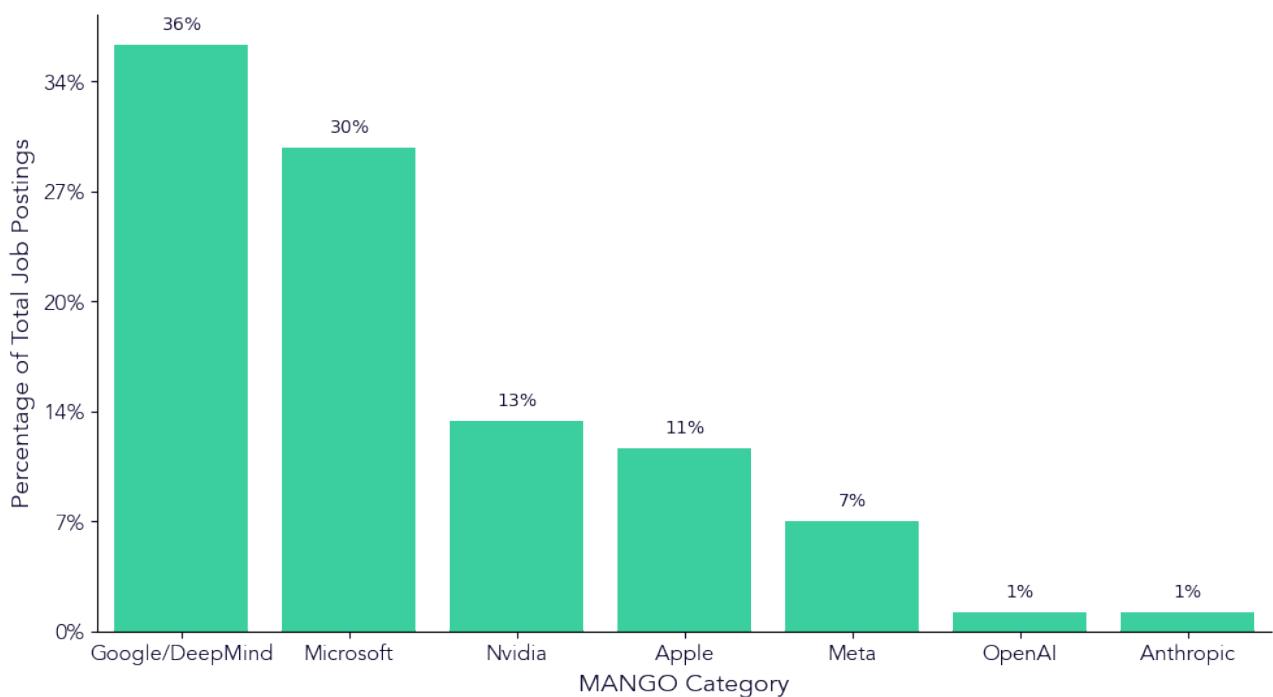


Figure 21: Distribution of job postings within the MANGO companies.

## MANGO Most Wanted Tools for 2025

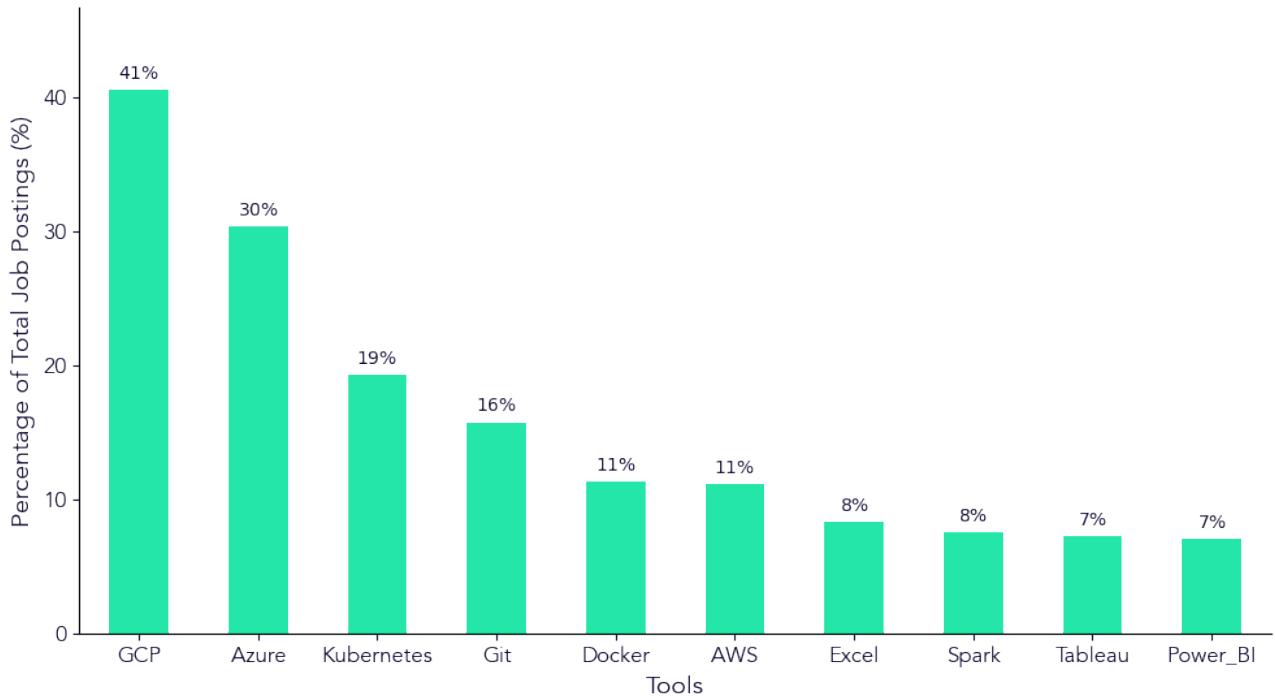


Figure 22: Distribution of demanded tools within the MANGO companies.

- As AI integration matures, the “Modern Data Stack” is shifting towards performance. Tools that facilitate real-time data streaming and high-speed processing, such as Kafka and Spark, are seeing increased co-occurrence with cloud platforms, indicating that “AI-ready” infrastructure is a top hiring priority.
- In the battle for Business Intelligence, Power BI (13%) has gained a clear lead over Tableau (10%). This suggests that organizations are favoring the deep integration of the Microsoft ecosystem to streamline workflows from the desktop (Excel) to the cloud (Azure).
- Tech hiring remains centered in established hubs, with California, Texas, and New York leading national demand. Notably, AWS dominance is strongest in these high-tech corridors, while Excel remains the leader in states with more traditional industry bases.
- Proficiency in version control (Git, 21%) and containerization (Kubernetes, 13%) is no longer an “extra” skill; these have become fundamental requirements across almost all technical job categories.

## Summary

The 2025 technical job market is defined by a massive reliance on cloud infrastructure, data automation, and the maturation of the “Modern Data Stack”. Through an analysis of nearly 600,000 job postings, it is evident that while new specialized tools emerge, the industry remains anchored by foundational technologies.

AWS and Git have become the bedrock of technical execution, but Excel remains the world’s most frequently required tool, proving that human-readable data remains the ultimate priority for business stakeholders. However, the rise of Kubernetes and the shift toward Power BI indicate that organizations are aggressively moving toward a future that is simultaneously cloud-native and deeply integrated into existing enterprise ecosystems.

Hiring activity remains regionally concentrated, with California and Texas leading the demand. The dominance of the “Engineer” role (51% market share) confirms that for the 2025 calendar year, companies are prioritizing the foundational technical infrastructure and the construction of scalable systems over purely analytical or management-heavy positions.

## Suggestions for Further Improvements

**Salary and Tool Correlation:** Integrating salary data would allow for the creation of a “Value Matrix” to identify which toolsets (e.g., dbt + Snowflake) command the highest financial premium compared to more common tools like Excel.

---

Growth Velocity Tracking: Comparing this 2025 data against 2024 benchmarks would help identify “rising stars”—tools like dbt or Monte Carlo that may have lower total volume but high growth velocity.

We can use the tool categories from [stackoverflow](#) and compare the tendencies with our data.

## Limitations

The limitations of this analyses is the data quality, we cannot be sure that we have ALL the job postings from 2025 in USA, yet I believe that the general tendencies would still remain in the data. For being comepletely precice the error rates could be calculated, yet they would bring some confusion for non technical readers, thus the errors remain unclear. As I have mentioned this would not change the general tendencies just could make some difference for the postings where the percentages are very close to eachother.