

Data Market Intelligence

What 2023 job-posting data shows about compensation, skills, and market activity

Market-intelligence briefing based on a public 2023 U.S. job-posting dataset.

CONTENTS

Contents

01	Executive Summary	03
02	Market Scope & Analytical Approach	04
03	Compensation Signal	05
04	Skill Signals	06
05	Market Activity	07
06	Management Implications & Monitoring Framework	08
A	Query Definitions & Filters	09
B	Metric Definitions & Interpretation Limits	10

Executive Summary

The analysis used six SQL queries to examine compensation, skills, employer activity, and monthly posting volume in a public 2023 data-jobs dataset.

EXHIBIT 01

Four signals for market monitoring

01

Returned salaries span a wide range, but cohort definition matters

The top-ten salary query returned standardized annual salaries from \$232,423 to \$375,000. The highest result was a `Sr Data Analyst` listing. The query excluded titles containing `Senior`, but the abbreviation `Sr` was not included in the filter. The result therefore demonstrates why salary benchmarks require standardized title and seniority definitions before they are used for comparison.



02

SQL and Python show broad cross-role presence

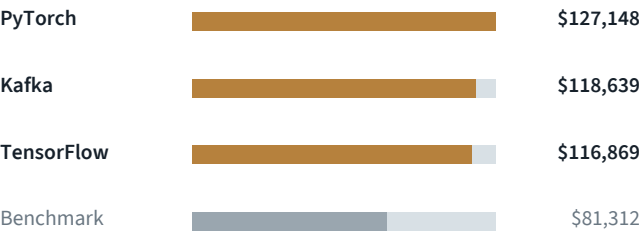
Across Data Analyst, Data Scientist, and Data Engineer postings: SQL appeared in 8,289 associated postings, Python appeared in 7,708. Both appeared across all three selected role categories.



03

Several less-common skills were associated with higher average listed salaries

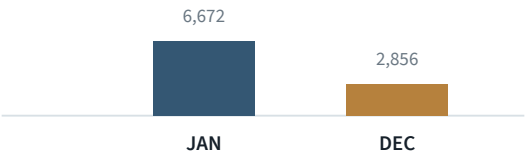
Within the project's under-5%-of-postings screen: PyTorch · \$127,148, Kafka · \$118,639, TensorFlow · \$116,869. The salary benchmark used in this analysis was \$81,312. These figures describe posting prevalence and average listed salary. They do not measure applicant competition or prove that possession of a skill causes higher compensation.



04

Posting activity changed substantially during 2023

January postings: 6,672. December postings: 2,856. That is approximately a 2.3x difference. The pattern supports monitoring posting activity over time. A single year does not establish a permanent hiring season.



MANAGEMENT IMPLICATION

A useful market-intelligence process should rely on repeatable definitions and repeated measurement, rather than isolated salary rankings or one-time posting counts.

Market Scope & Analytical Approach

Job-posting data provides useful market signals when the population behind each metric is explicit.

The project examined six questions covering: listed compensation; skills associated with high-returned salary listings; less-common skills and average listed salary; monthly posting activity; employer posting volume and average listed salary; skill overlap across three data-role categories.

Most analyses focus on U.S. full-time Data Analyst postings. The cross-role skill analysis expands the population to Data Analyst, Data Scientist, Data Engineer.

01 Data structure The PostgreSQL database uses four related tables: JOB_POSTINGS_FACT posting, title, salary, location, date, and role attributes COMPANY_DIM company information SKILLS_DIM skill names and categories SKILLS_JOB_DIM bridge between postings and skills	02 Salary treatment Where annual salary was available, the analysis used that value. Where only hourly salary was available: hourly salary x 2,080 was used to create a standardized annual value.
--	---

ANALYTICAL PRINCIPLE

Different queries use different populations. The report therefore keeps three questions explicit: Who is included? What is the denominator? What does the resulting metric actually measure?

EXHIBIT 1 Information flow and query populations



Compensation Signal

Salary rankings are sensitive to how the comparison group is defined.

The top-ten salary query returned standardized annual salaries ranging from \$232,423 to \$375,000. The highest returned listing was Sr Data Analyst.

The SQL excluded several leadership and seniority terms, including Senior, Director, Principal, Lead, Manager, Head, VP, Chief. It did not exclude the abbreviation Sr. The row therefore remained in the result.

01 What the result supports

The query accurately returned the highest salaries within the written filter.

02 Management implication

Compensation benchmarking requires a standardized cohort before the result is used for salary planning. Role title, seniority, geography, employment type, and salary treatment should remain consistent across measurement periods.

EXHIBIT 2 A filtered salary range, not a market ceiling



Highest-returned listing uses the abbreviated seniority label `Sr`, which was not included in the written exclusion filter.

Skill Signals

Broad skills and specialized skills answer different workforce questions.

01 **Broad cross-role presence**

The cross-role query compared three categories: Data Analyst, Data Scientist, Data Engineer. Two skills appeared particularly broadly: 8,289 SQL-associated postings, 7,708 Python-associated postings. Both appeared across all three role groups. For workforce planning, these provide broad skill signals within the analyzed market.

02 **Specialized skills showed a different pattern**

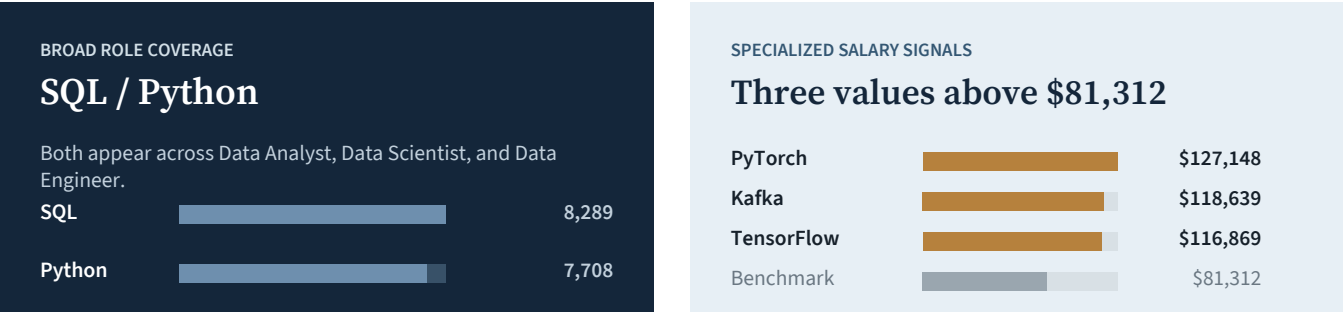
A separate query screened for skills appearing in more than 10 matching postings and below a project-specific 5% prevalence threshold. Three examples had average standardized salaries above the \$81,312 benchmark:

SKILL	MEAN LISTED SALARY
PyTorch	\$127,148
Kafka	\$118,639
TensorFlow	\$116,869
Salary benchmark	\$81,312

03 **Management implication**

These two views should not be treated as the same signal. SQL and Python indicate broader cross-role presence. PyTorch, Kafka, and TensorFlow are narrower posting signals associated with higher mean listed salaries in this analysis. Neither result determines which skill an organization should hire for without first defining the capability the role requires.

EXHIBIT 3 **Broad role coverage and specialized salary signals**

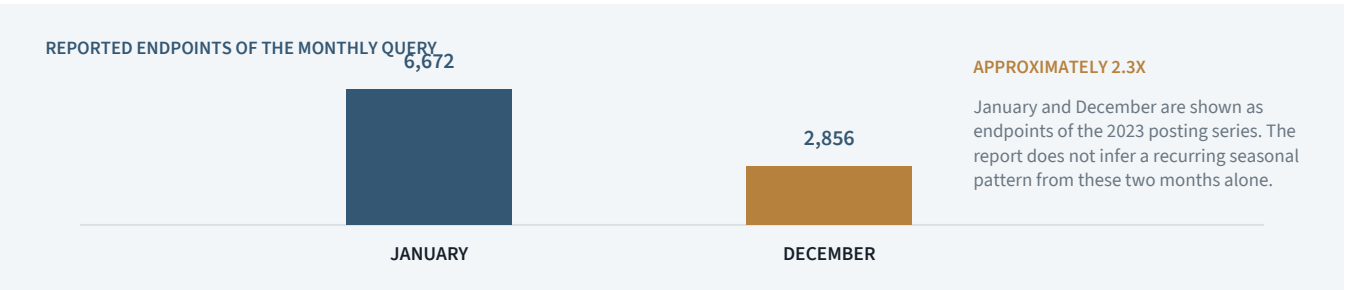


Market Activity

Posting volume changed materially over the course of 2023.

The monthly posting query found 6,672 January postings versus 2,856 December postings, approximately 2.3x difference. The result shows substantial variation in posting activity within the dataset. The evidence does not establish that January is always the strongest hiring month or that December is always the weakest. A repeatable monthly series would be needed to determine whether the pattern persists across years.

EXHIBIT 4 Posting activity varied substantially across 2023



01 Employer activity provides a second market signal

A separate query compares employers on matching posting volume and average standardized listed salary, and includes employers with more than 10 matching postings. The measure is useful for identifying where posting activity is concentrated and how listed compensation differs across employers. It does not measure employer quality.

02 Management implication

Treat posting activity as a market-monitoring measure. Repeated observation is more useful than turning one year's pattern or one employer's posting count into a permanent rule.

Management Implications

- 01

Standardize the market cohort
Fix the definitions used for: role category; seniority; geography; employment type; salary treatment. This makes compensation and market-activity comparisons meaningful across reporting periods.
- 02

Separate broad skills from specialized capability signals
Track widely distributed skills such as SQL and Python separately from narrower technologies associated with particular roles or capabilities. Use role requirements, rather than salary association alone, to decide which skills matter to a hiring plan.
- 03

Monitor the market as a time series
Re-run the same measures on a regular schedule. A recurring view can show whether changes in posting volume, salary benchmarks, skill prevalence persist or reverse over time.
- 04

Add hiring-process data before making talent-supply conclusions
Job postings alone cannot determine: applicant supply; candidate competition; time-to-fill; offer acceptance; completed hires. Those measures require additional recruiting or labor-market data.

Monitoring Framework

INDICATOR	WHAT IT MEASURES	WHAT IT DOES NOT MEASURE
Posting count	Advertisement activity	Completed hiring
Listed salary	Advertised compensation	Accepted compensation
Skill prevalence	Frequency in postings	Applicant supply
Average salary by skill	Salary association	Causal skill premium
Employer posting volume	Employer advertisement activity	Employer quality
Cross-role skill count	Breadth across selected roles	Career resilience

MANAGEMENT CONCLUSION

Use posting data as a repeatable market signal, with the population and denominator attached to every metric.

APPENDIX A

Appendix A - Query Definitions & Filters

Query 1 - Top-returned Data Analyst salaries

Population: `job\_title\_short` = 'Data Analyst', United States, full-time, annual or hourly salary available. Excluded title strings: `Senior` · `Director` · `Principal` · `Lead` · `Manager` · `Head` · `VP` · `Chief`. Salary standardization: annual salary if available, otherwise hourly salary x 2,080. Output: top 10 standardized yearly salary rows. Known title-definition limitation: `Sr` is not included in the seniority exclusions, which allowed the `Sr Data Analyst` listing to remain.

Query 2 - Skills associated with the top-returned listings

The top-returned salary cohort is joined to `skills\_job\_dim` and `skills\_dim` to identify the skills associated with those listings. This query describes the skills attached to that specific returned cohort.

Query 3 - Less-common skills and average listed salary

Population: U.S. full-time Data Analyst postings with salary data. Selected skills must appear in more than 10 matching postings and fall below 5% of the calculated posting denominator. The 5% denominator is broader than the later seniority-filtered result rows because the `job\_total\_count` calculation does not apply those title exclusions. The threshold is therefore a project-specific prevalence screen, not a universal definition of a niche skill.

Query 4 - Monthly posting activity

Population: U.S. full-time Data Analyst postings with the same selected seniority-title exclusions. Postings are grouped by calendar month and ordered chronologically. The query counts posting activity; salary availability is not required for this analysis.

Query 5 - Employer posting volume and salary

Population: U.S. full-time Data Analyst postings with salary data and the selected seniority-title exclusions. Employers must have more than 10 matching postings. Measures: posting count, average standardized salary. The analysis describes employer posting activity and listed compensation.

Query 6 - Cross-role skill presence

Role categories: Data Analyst, Data Scientist, Data Engineer. Population: U.S. full-time postings with salary data. Unlike several Data Analyst queries, this query applies no seniority-title exclusions. Skills must appear across at least two of the three role categories. Measures: role-category count; associated posting count; average standardized salary.

APPENDIX B

Appendix B - Metric Definitions & Interpretation Limits

Standardized yearly salary

Where annual listed salary exists, that value is used. Where only hourly listed salary exists, hourly salary x 2,080 is used as the annualized value.

\$81,312 benchmark

The analysis uses a median standardized salary benchmark for the filtered U.S. full-time Data Analyst population with salary data and the specified seniority-title exclusions. It is a benchmark for the query population, not a universal U.S. Data Analyst salary.

Posting prevalence

The number of analyzed postings associated with a skill. Posting prevalence measures how often a skill appears in the defined posting sample. It does not measure the number of applicants or candidate competition.

Skill salary association

The average standardized listed salary associated with postings containing a skill. The result can identify compensation differences in the sample. It does not establish that adding the skill to a role causes compensation to increase.

Posting volume

The number of postings meeting the query definition. Posting volume measures advertised opportunities in the source data. It does not measure completed hires.

Employer posting volume

The number of matching postings associated with an employer. It indicates employer activity within the filtered dataset. It is not a measure of employer quality.

Cross-role presence

The number of selected role categories in which a skill appears. The analysis uses three categories: Data Analyst, Data Scientist, Data Engineer. Cross-role presence describes breadth within those categories. It does not measure career resilience, lifetime earnings, individual career progression, or future demand.

Time period

The analysis describes a 2023 job-posting dataset. Patterns observed across its months should be treated as historical observations from that year. Repeated data across additional years would be required before describing a persistent seasonal pattern.

Source and method note: public 2023 job-posting dataset, PostgreSQL query outputs, and Tableau views. Metric labels in the main report should be read with the definitions above.