

# Analyzing In-Office Salary Trends in AI, Data, and Machine Learning Roles

(by André Jardim)

## Problem Statement

Despite the increasing popularity of remote work, a substantial portion of AI, data, and machine learning roles remain in-office. This analysis will focus on understanding salary trends for in-office positions, examining how factors like job title, experience level, company size, company location, and changes over time impact compensation. By highlighting these trends, I aim to provide insights valuable to companies designing competitive compensation packages and individuals seeking clarity on in-office salary expectations.

## Business Impact

By examining in-office salary trends in these fast-growing fields, the analysis will provide companies with benchmarks for competitive compensation strategies, aiding in talent acquisition and retention. This data-driven approach will also benefit industry professionals, offering clear benchmarks that support informed career decisions and negotiations. For policymakers and analysts, the findings can illuminate regional disparities in salary and the relative competitiveness of the U.S. against other regions.

## General Dataset Information

**File Name:** salaries.csv

**Description:** AI, ML, Data Salaries 2020 – 2024

**Dataset Details:** 59,325 Rows / 11 Columns

**Size:** 3,304 KB (3.3 MB)

**Source:** [AIJobs.net](https://ajobs.net)

**Method of Collection:** Direct download from AIJobs.net

## Target Features

The columns that will be used to simplify and analyze the problem statement are:

- ***salary\_in\_usd***: The salary of the individual in U.S. dollars. This is the primary target variable for salary analysis.
- ***work\_year***: The year in which the salary data was recorded. It helps track salary trends over time.
- ***experience\_level***: The level of experience of the individual, categorized as Entry-Level, Mid-Level, Senior, or Executive.
- ***job\_title***: The specific role or job title of the individual (e.g., Data Scientist, AI Engineer). This helps segment salary analysis by job role.
- ***company\_location***: The geographical location of the company (e.g., United States, United Kingdom, Canada), providing insights into location-based salary trends.
- ***company\_size***: The size of the company, categorized as Small, Medium, or Large, which can influence salary variations across different organizations.

## Analysis

### Salary Distribution by Experience Level in 2024:

- In 2024, salaries increase with experience level, with entry-level roles earning an average of \$93,954, mid-level at \$118,841, senior at \$148,740, and executive at \$173,866. The variability in salaries also rises across levels, with senior and executive roles exhibiting a wider range, reflecting diverse roles and responsibilities. Entry-level positions show lower salary variance, while senior and executive roles have significant spreads, indicating that higher positions often have larger discrepancies based on company size, industry, and role specialization. Overall, experience level strongly influences salary, with higher positions commanding substantially higher pay, though with increased variability.

### Salary Distribution by Company Size in 2024:

- In 2024, salaries are highest in large companies (\$158,860), followed by small companies (\$144,377) and medium companies (\$128,079). The small company group is underrepresented in the dataset, with only 4 records remaining after data cleaning and transformation, which makes it harder to draw reliable conclusions from this group. Medium-sized companies show the most salary variability, with salaries ranging from \$35,625 to \$400,000. Large companies also have a broad salary range, from \$54,438 to \$247,065, but with fewer extreme values. Overall, salary tends to increase with company size, although medium-sized companies show the greatest variation in salary.

### Salary Distribution by Company Location in 2024:

- In 2024, salary trends across different company locations exhibit notable variations. The United States stands out with the highest average salary of \$140,832, followed by Canada at \$118,240, and the United Kingdom at \$104,003. The U.S. also shows the greatest salary range, ranging from \$51,050 to \$247,065, indicating significant salary disparities within the region. In contrast, Canada has a narrower range, with salaries between \$58,340 and \$232,750, and a lower standard deviation, suggesting more consistency in salary distribution. The United Kingdom, while having the lowest average salary, displays a broader salary range from \$35,625 to \$400,000, with considerable variation, as reflected by the high standard deviation. These differences underscore the varying economic conditions, cost of living, and job market characteristics across these regions.

### Salary Distribution by Job Title in 2024:

- The analysis of job titles and salary data for 2024 reveals key trends in both the prevalence and compensation of roles across the dataset. The top 10 job titles by count reflect the most commonly held positions in the industry, with Data Analyst, Data Engineer, and Data Scientist being particularly frequent. These roles, though widespread, show notable variability in salary, with positions like Machine Learning Engineer and Software Engineer offering higher average salaries.
- In contrast, the top 10 job titles by average salary are more specialized and managerial, with roles such as ML Infrastructure Engineer, Head of Machine Learning, and Head of AI commanding significantly higher compensation. These positions, however, tend to be less common, reflecting the demand for specialized skills and leadership in cutting-edge fields like machine learning and artificial intelligence. It is important to note that the presence of lower-frequency observations in the dataset—such as roles with only one or two instances—can skew the average salary for these job titles, potentially affecting the ranking of higher-paying positions.

### Salary Distribution by Work Year:

- The salary data across different work years reveals some notable trends. In 2020, the average salary was the highest at \$155,775, but with a very high standard deviation (\$176,924),

indicating a wide range of salaries, including extreme values. In 2021, the average salary dropped to \$139,793, with a more consistent distribution (lower standard deviation of \$70,035). The trend continued in 2022, where the average salary fell further to \$124,198, and the distribution tightened. 2023 saw a slight recovery with an average of \$136,222, while 2024's average salary was \$130,556.

- It's important to note that years with fewer data points, such as 2020 (4 observations) and 2021 (14 observations), may not be fully representative of the broader trends and could have skewed the overall averages. The variability in these years highlights the influence of low sample counts on salary averages.

## Conclusion

In conclusion, the analysis of in-office salary trends in AI, data, and machine learning roles reveals that experience, company size, and geographic location are key factors influencing compensation. Higher experience levels correspond with higher salaries, particularly for senior and executive positions, which exhibit greater variability. Larger companies tend to offer higher salaries, while medium-sized companies show the greatest salary range. The United States leads in average salary, with Canada and the United Kingdom presenting more consistent but lower salaries.

The dataset, focused primarily on 2024, also highlights the diversity of job titles within the industry, with a significant number of different roles contributing to salary variability. Specialized roles, such as ML Infrastructure Engineer and Head of Machine Learning, command higher salaries, while more common roles like Data Analyst and Data Scientist show lower average salaries but are more widely represented. These insights are valuable for companies to optimize their compensation strategies and for individuals seeking clarity on in-office salary expectations in the evolving AI and machine learning fields.

**Access the full Jupyter Notebook for this project [here](#).**