Analysis of Job Posting for Data Analyst position

Simon Boakye

3 February 2025

Table of Contents

1.0 Introduction

The analysis examines 400 data analyst job postings to explore the relationship between required experience, programming language skills, and salary expectations. The dataset includes information on median salary estimates, minimum years of experience needed, and programming language requirements (R, Python, both, or neither). Julius AI was used for the analysis.

2.0 Years of experience analysis

2.1 Minimum Years Experience vs. Median Salary Estimate

The analysis reveals a weak relationship between years of experience and median salary for data analyst positions. The correlation coefficient of 0.052 indicates virtually no linear association between these variables. This finding is somewhat surprising, as we might expect higher salaries for positions requiring more experience. The wide distribution of salaries at each experience level suggests that other factors, such as industry, company size, location, or specific technical skills, may play a more significant role in determining compensation than years of experience alone (Figure 1)

Figure 1: Relationship between Minimum Years Experience vs. Median Salary Estimate

3.0 Programming Language Analysis

3.1 Median Salary Distribution by Programming Languages

A notable difference in salary distributions across programming language requirements was observed in the analysis. Positions requiring both R and Python command the highest average salary ($75,762), followed closely by Python-only positions ($74,866). Jobs requiring neither programming language have a mean salary of $70,398, while R-only positions show the lowest average at $64,750. Many roles did not explicitly list programming language skills. The findings suggest that proficiency in multiple programming languages, particularly Python or both Python and R, may be associated with higher compensation in the data analyst job market. The small sample size caution in over-interpreting specific results (Figure 2).

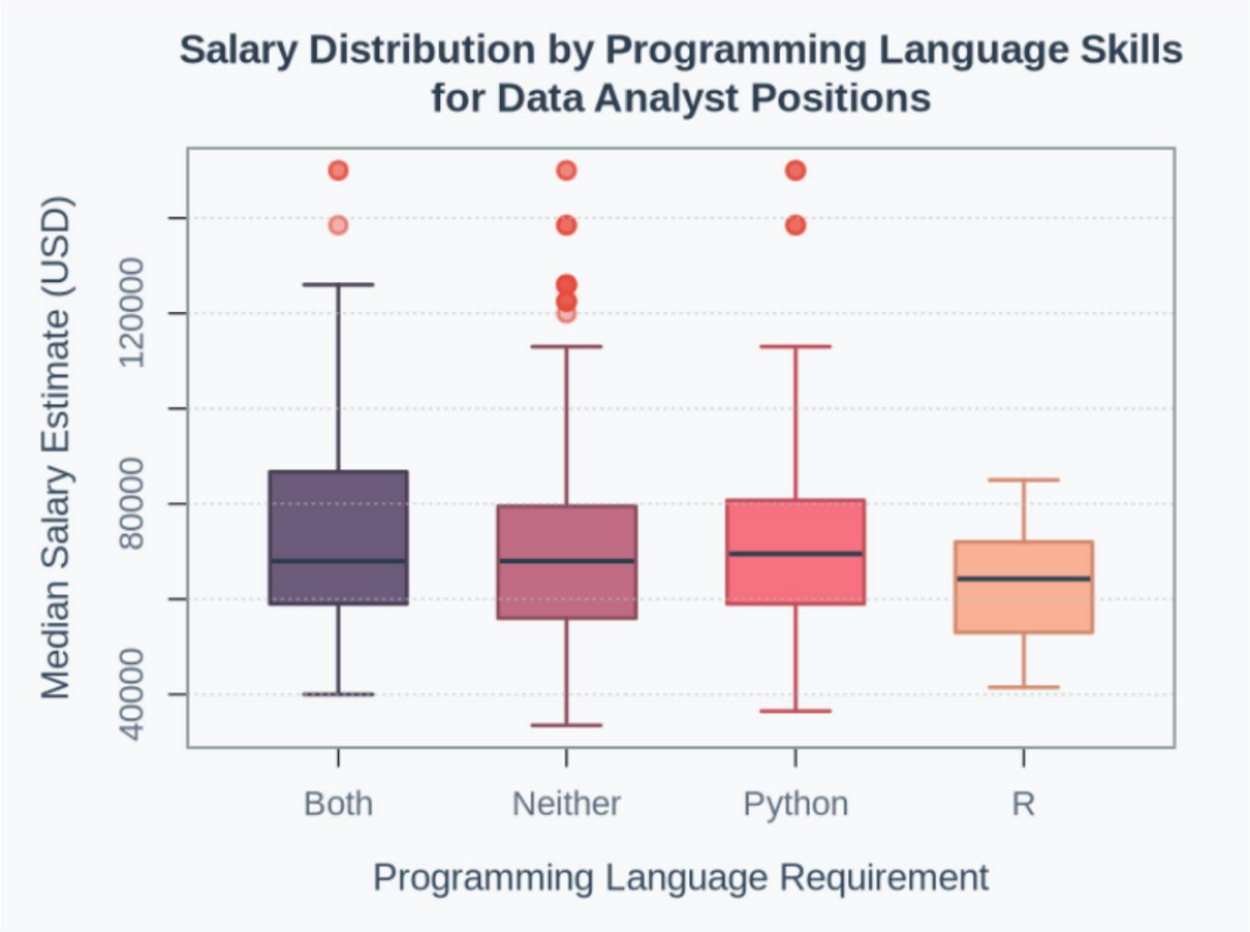

Figure 2: Median Salary Distribution by Programming Languages

4.0 Conclusion

There is weak relationship between years of experience and median salary for data analyst position. Roles asking for both Python and R tend to have a higher average salary than roles that lists neither or R only.

5.0 Reflection on using Julius

While Julius could perform the analysis efficiently, I still needed to think critically about what I wanted to investigate.

One thing that surprised me was how quickly Julius could reveal patterns that might have taken much longer to identify manually. It was also surprising to see how small changes in the way I framed a question could lead to very different insights. This highlighted how important precision and clarity are in data analysis. Overall, I learned that data analysis is not just about using software to produce results, but about critical thinking and decision-making throughout the process.