

Data Analyst Job Postings Analysis Report

Dr Ali Ammar

29 Oct 2025

Prepared with the support of Julius AI

Introduction

This report presents an analysis of 400 data analyst job postings, examining the relationship between experience requirements, programming language skills, and salary expectations. The dataset includes information on job titles, median salary estimates (in USD), minimum years of experience required, and programming language requirements (R, Python, both, or neither).

The analysis focuses on two key research questions:

1. How does the minimum years of experience required relate to median salary estimates?
2. How do programming language requirements affect salary distributions?

Understanding these relationships can provide valuable insights for both job seekers looking to optimize their career development and employers seeking to structure competitive compensation packages.

Key Findings

Dataset Overview

- **Total job postings analyzed:** 400
- **Mean salary:** \$71,793.75
- **Median salary:** \$68,000
- **Salary range:** \$33,500 - \$150,000
- **Postings with experience requirements:** 309 (91 had missing experience data)

Programming Language Requirements Distribution

- **Neither R nor Python:** 255 postings (63.8%)
- **Both R and Python:** 67 postings (16.8%)
- **Python only:** 63 postings (15.8%)
- **R only:** 15 postings (3.8%)

Salary by Programming Language

- **Both R and Python:** Mean = \$76,022, Median = \$68,500
- **Python only:** Mean = \$74,532, Median = \$69,000
- **Neither:** Mean = \$70,433, Median = \$68,000

- **R only:** Mean = \$64,533, Median = \$61,500

Experience and Salary Relationship

- **Correlation coefficient:** 0.053
- The weak positive correlation suggests that years of experience has minimal linear relationship with salary in this dataset, indicating that other factors (such as specific skills, location, or company size) may play more significant roles in determining compensation.

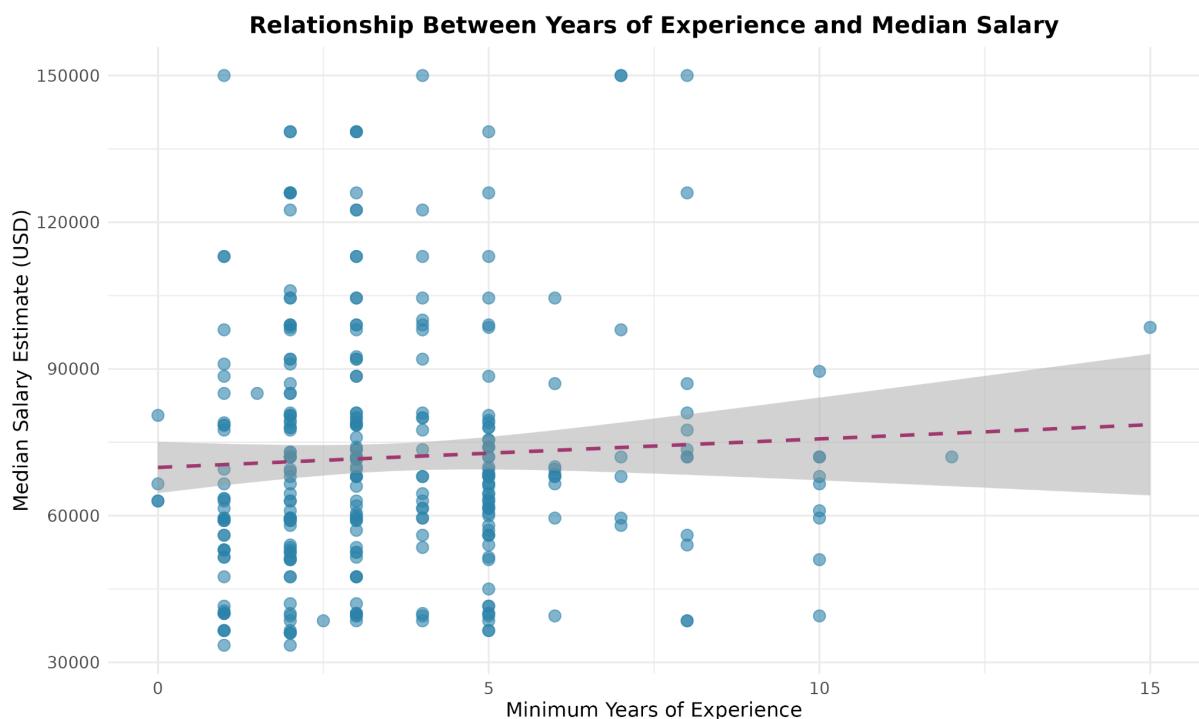


Figure 1: Relationship Between Years of Experience and Median Salary

This scatter plot displays the relationship between minimum years of experience required and median salary estimates. Each point represents a job posting. The dashed trend line shows a slight positive relationship, though the correlation is weak ($r = 0.053$), suggesting that experience alone is not a strong predictor of salary in data analyst positions.

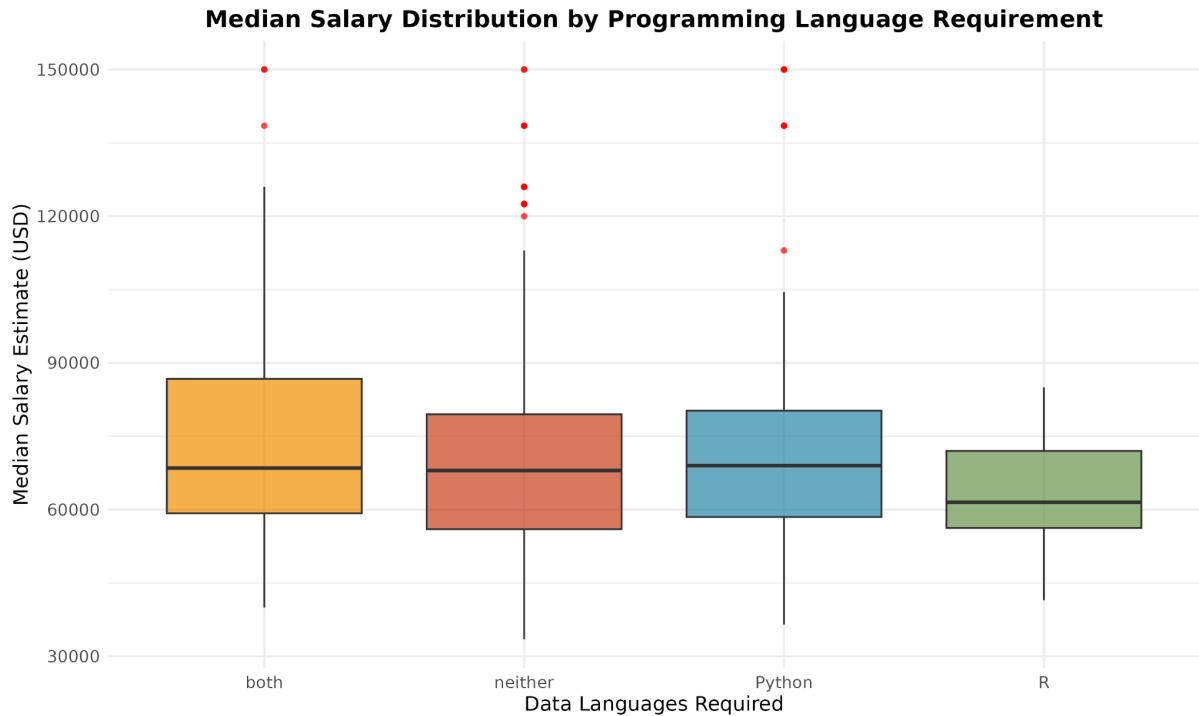


Figure 2: Median Salary Distribution by Programming Language Requirement

This box plot compares salary distributions across different programming language requirements. Positions requiring both R and Python show the highest mean salary (\$76,022), followed by Python-only positions (\$74,532). Interestingly, positions requiring R only show the lowest mean salary (\$64,533), while positions requiring neither language fall in the middle (\$70,433).

Discussion

The analysis reveals several interesting patterns in the data analyst job market:

- Programming Skills Premium:** Jobs requiring both R and Python command higher salaries on average, suggesting that versatility in programming languages is valued by employers.
- Python Dominance:** Python-only positions offer competitive salaries comparable to those requiring both languages, reflecting Python's widespread adoption in data analytics.
- Experience Paradox:** The weak correlation between years of experience and salary ($r = 0.053$) is surprising and suggests that other factors—such as specific technical skills, industry domain, company size, or geographic location—may be more influential in determining compensation.

4. **Market Segmentation:** The large number of positions (63.8%) not requiring R or Python suggests a significant segment of data analyst roles that may rely on other tools (such as SQL, Excel, Tableau, or other business intelligence platforms).
-

Personal Opinion

As my coding experience is totally absent, and my exposure to these tools started really with this course, today's task, and the interaction with both Google Sheets and Julius AI for the second time was pretty easy and without any serious challenges.

One of the pain points was the output in google sheets when extracting the required programming languages, despite being very strict in the prompt and giving only 4 options of response, the tool gave some responses outside this ranges, so I had to clean and harmonize the responses (Python/R or using Python and R instead of R and Python).

Julius AI made the report totally ready for my review and to add this section.
Highly satisfied and excited to use these tools further in my day to day professional activities.