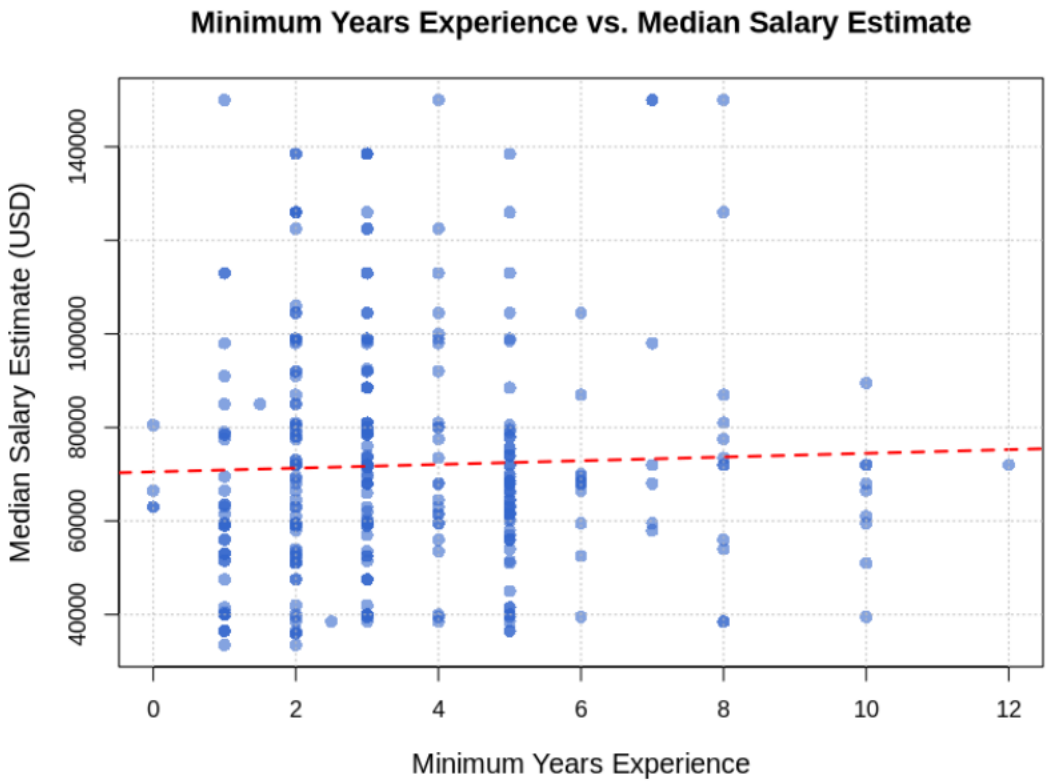
Introduction

This project analyzes a dataset of 400 data analyst job postings to examine the relationship between required experience, programming language skills, and salary expectations. The analysis explores whether years of experience and proficiency in specific data languages (Python, R, both, or neither) are associated with higher median salary estimates in the data analyst job market.

Years of Experience Analysis

Scatter Plot:



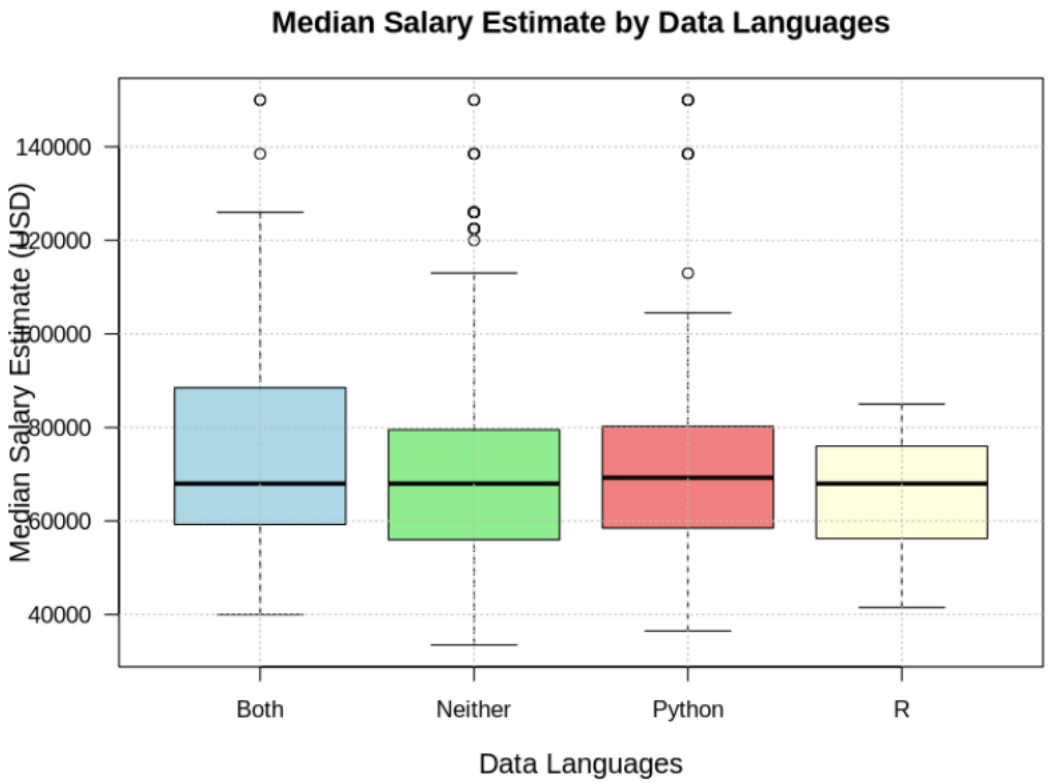
Interpretation:

The scatter plot reveals a surprisingly weak relationship between minimum years of experience and median salary estimates (correlation = 0.034). The data points are widely

dispersed across all experience levels, with salaries ranging from approximately $33,500 to $150,000 regardless of whether positions require 0 years or 12 years of experience. The nearly flat trend line indicates that, on average, additional years of required experience do not substantially increase the median salary estimate. This suggests that other factors—such as industry, geographic location, company size, or specialized technical skills—may play a more significant role in determining salary than years of experience alone.

Programming Language Analysis

Box Plot:



Interpretation:

The box plot displays the salary distributions across four programming language categories. Positions requiring both Python and R show the highest average salary ($76,127), suggesting that versatility in programming languages is valued in the market. Python-only positions follow closely with an average of $74,617, while positions requiring

neither language average $70,366. Interestingly, R-only positions show the lowest average salary ($66,100) and are the least common category with only 15 positions. Despite these differences in mean salaries, the median values across all categories are remarkably similar (around $68,000), indicating that while programming skills may increase upper salary potential, the typical salary remains relatively consistent. The “Neither” category represents the majority of positions (258 out of 400), suggesting that many data analyst roles may focus on other tools such as SQL, Excel, or business intelligence platforms rather than Python or R.

Reflection:

Easy thing was, Julius makes complex data analysis easy and fast**.** The difficult thing was relationship between minimum years of experience and median salary and what surprised me is that there was no sensible correlation or association between experience and median salary in this data. What I learnt is that in the ear of AI data analysis and program code is not challenging task for domain experts and those who have knowledge and skill of AI and also the job market is vary depend on programing skills.