

Week 3 workshop, Tun Linn Thein, 4 Feb 2026

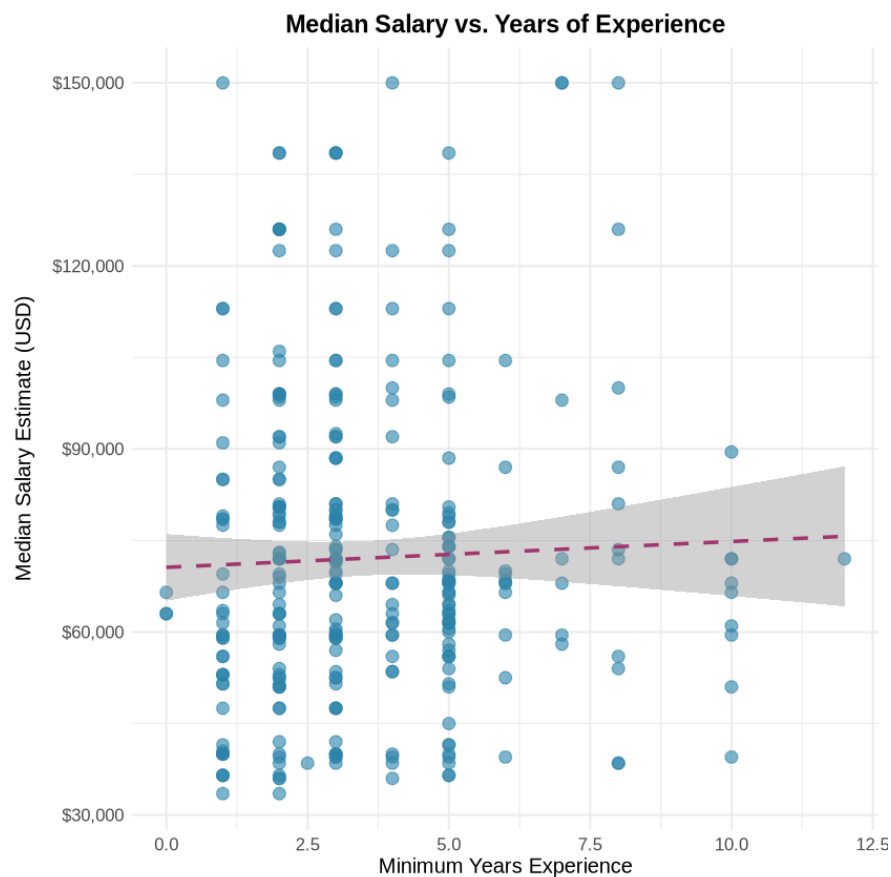
Introduction

I downloaded a dataset from a project which was originally scraped from Glassdoor in June 2020 to understand the Data Science/Engineer/Analyst job market. I used Julius AI for data analysis.

Years of Experience Analysis:

This plot shows the relationship between minimum years of experience and median annual salary estimates. I've included a linear regression line to show the trend. The data includes 308 job postings (92 had missing experience data).

This shows a positive relationship between years of experience and annual salary, with the dashed line indicating the overall trend.



Programming Language Analysis:

This plot compares salary distributions across four categories:

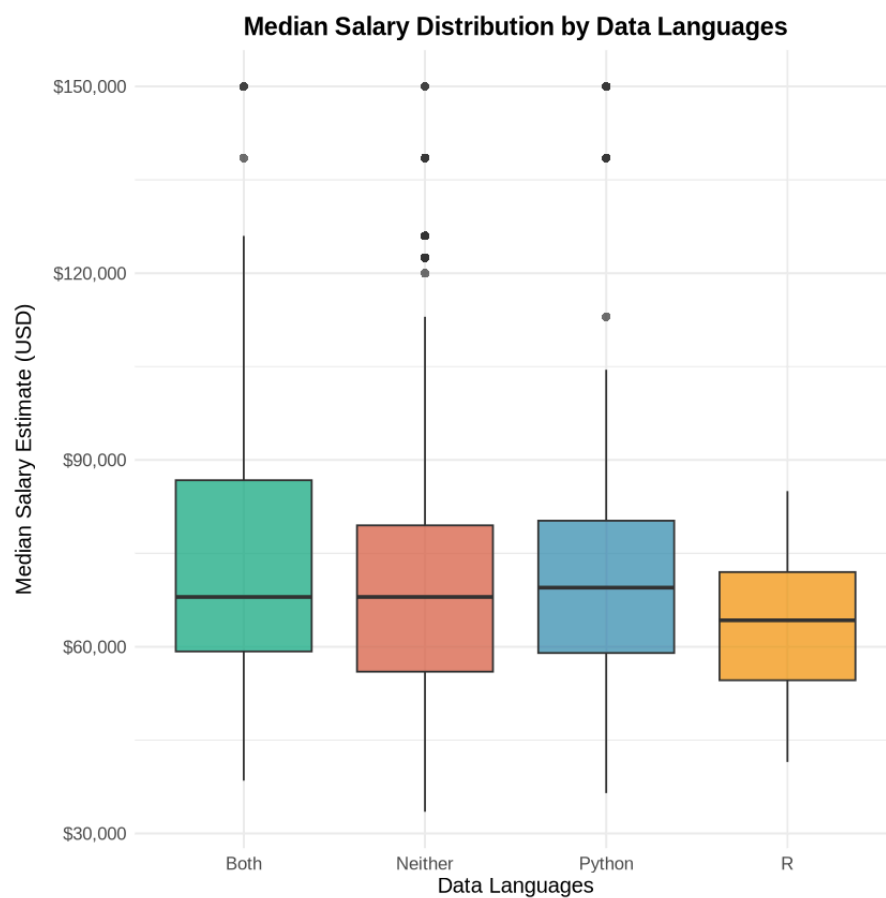
Both (R and Python): 67 postings

Neither: 256 postings

Python: 63 postings

R: 14 postings

This shows that jobs requiring "Both" R and Python tend to have higher median salaries, while "Neither" has the most variation. Jobs requiring Python alone also show competitive salaries compared to those requiring R alone.



Personal Reflection on assignment:

Although generative AI is quick and convenient to analyse data, it is important to repeat the prompts a few times and check/verify the outputs and interpret carefully.