Assignment 3
Data Visualisation

Olivér Heicz

2025-10-25

Use of AI tools

In this assignment I used AI just for clarification of R syntax and formatting issues.

Task 1

Data Set Description

In this assignment, we work with the `Ecdat` package, more precisely with the dataset called **Wages**. The dataset tracks the wages and jobs of specific workers over a seven-year period, from 1976 to 1982. During this period, the same individuals are analyzed by their wage dynamics, as well as the labour market in the United States.

The formatted dataset contains 4,165 observations across 12 variables, representing 595 individuals, each observed over 7 years. We can analyze both cross-sectional variation (differences between individuals) and the temporal variation (changes over time for the same individual) owing to the data's panel structure.

The dataset includes the following key variables:

- **Experience**: Years of work experience (in years)
- **Gender**: Biological sex of the worker (male/female)
- **School**: Years of formal education completed (in years)
- **Wage**: Monthly wages in US dollars (derived from the logarithmic transformation)
- **Logarithmic Wage**: Natural logarithm of monthly wages
- **Union Membership**: Indicator of whether the worker is a union member
- **Ethnicity**: Categorical variable for racial/ethnic background
- **Occupation**: Type of work or industry sector
- **Residence**: Geographic location or region
- **Marital Status**: Whether the worker is married
- **Health Status**: Health conditions of the individual
- **ID**: Identification number for each individual (1-595)
- **Year**: Years of analysis (1976-1982)

This dataset originates from the Panel Study of Income Dynamics (PSID), a longitudinal household survey conducted in the United States. The PSID is one of the longest-running panel studies in the world and has been widely used in labor economics and social science research.

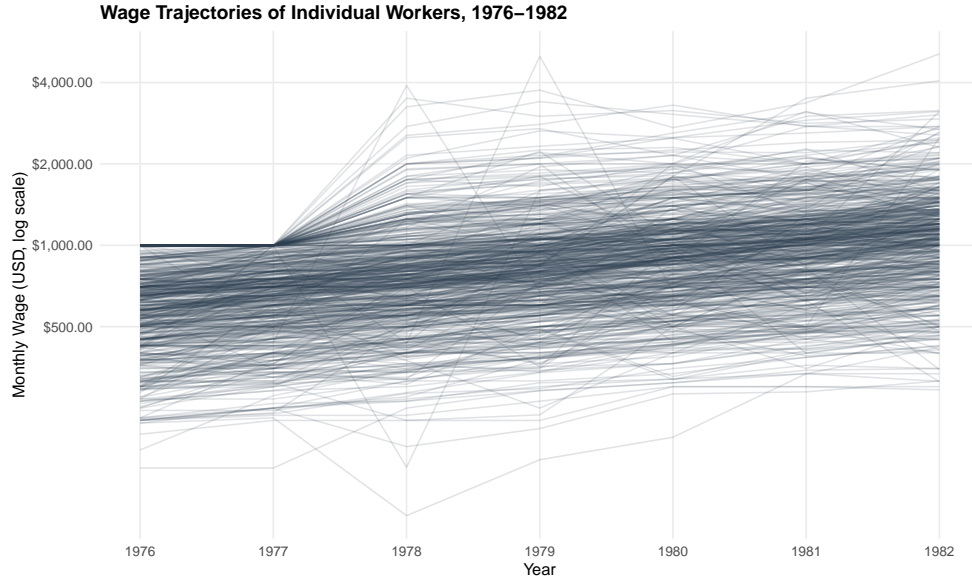**Wage Trajectories of Individual Workers, 1976–1982**

Figure 1: Monthly wages by year for 595 individuals tracked from 1976 to 1982. Each line represents one individual's wage trajectory over the seven-year period. The plot uses a logarithmic scale on the y-axis to better visualize the multiplicative nature of wage growth and reduce the impact of extreme values.

Discussion of Patterns

Looking at the plot above, we can detect a general upward trend in wages between 1976 and 1982 in the USA. The observed individuals, and overall people in the labour market would rightfully earn more with more experience over the years. In addition,it is quite noticeable that some of the individuals have quite different career paths.Some employees earned higher wage increases,while others show stagnation or even occasional decreases. This could be explained by people having different levels of education,work experience and other factors for example social benefits that increase their wages.

Task 2

In task 2, we investigate whether higher education leads to higher salaries. Our observation is set in the year of 1982, so we filter down on that specific year of the years of observations.

Looking at the plot, there does seem to be a positive relationship between education and wages. People with more years of schooling generally earn higher wages, and
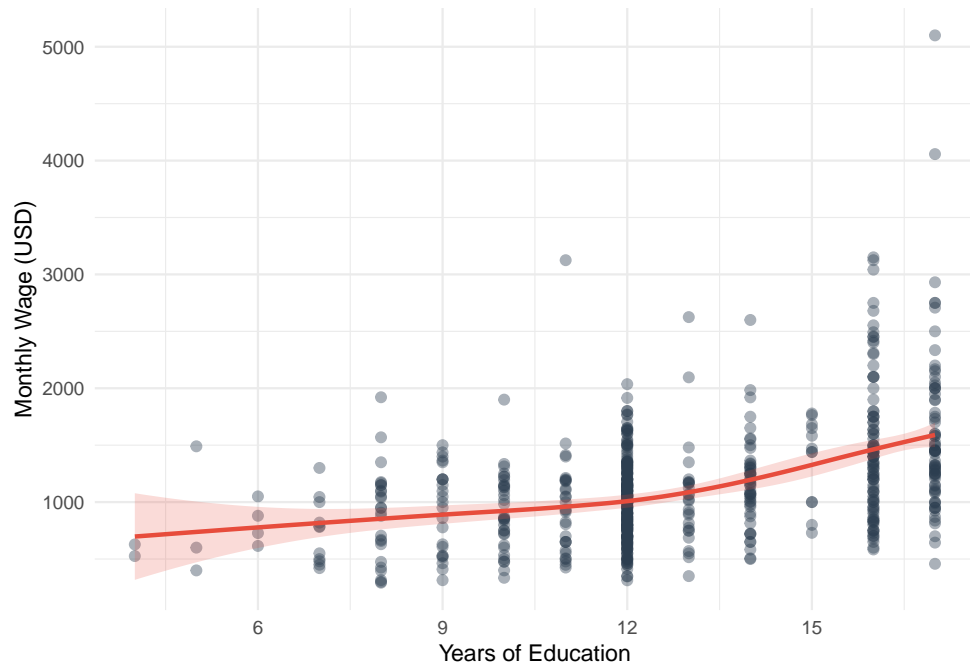
Figure 2: The relationship between years of education and monthly wages in the USA in 1982.

the trend line shows this upward pattern fairly clearly. Someone with 16 years of education which is typically a bachelor's degree appears to earn noticeably more on average than someone with just 12 years.

Task 3

In the final task, we examine how wage levels effect other variables in the dataset. We choose to illustrate the wage levels between males and females.For even better data telling, as the task encourages, we cut the wages variables into low, mid, and high levels.

From the plot it can be seen that there is a quite noticable discrepancy between the sexes. Females only fall into the low and mid levels, while a great percentage of men fall into the mid and even the high category.The plot states very clearly the distribution of wages of both genders and makes it easy to compare the results. Therefore, the bar chart is a good choice to illustrate the disparity of wages.
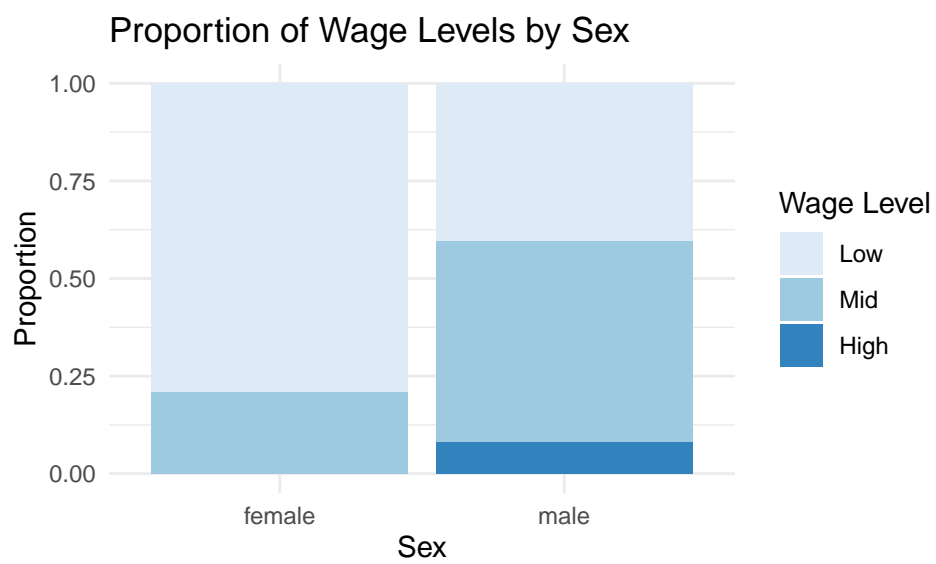
Proportion of Wage Levels by Sex

Figure 3: The wage level differences between genders in the USA.