A graph of salary distribution

AI-generated content may be incorrect.Figure1: Boxplot of Salary vs. Gender The boxplot compares salary distributions for male (0) and female (1) employees. The median salary for men is higher than for women, and the presence of several outliers in the male category suggests that some men have exceptionally high salaries. The difference between the quartiles indicates salary disparity by gender

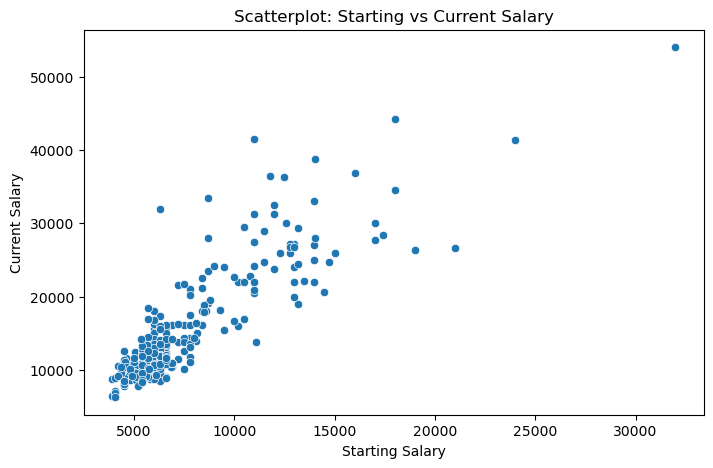


Figure 2: Scatterplot of Work Experience vs. Salary

A graph of a salary by gender

AI-generated content may be incorrect.

Figure 3: Boxplot of Salary vs. Gender. The boxplot compares salary distributions for male (0) and female (1) employees. The median salary for men is higher than for women, and the presence of several outliers in the male category suggests that some men have exceptionally high salaries. The difference between the quartiles indicates salary disparity by gender

A screenshot of a graph

AI-generated content may be incorrect.

Figure 4: Correlation Matrix for Salary Data

The heatmap visualizes the correlation coefficients between different variables. Strong positive correlations, such as between starting salary (salbeg) and current salary (salnow), suggest that an employee's initial salary is a strong predictor of future earnings. High correlations between age and work experience are expected, while multicollinearity in other variables suggests adjustments were necessary for the regression model

A graph of a graph

AI-generated content may be incorrect.

Figure 5: Histogram of Residuals for OLS Regression Model

The histogram of residuals helps evaluate whether the OLS model assumptions hold.

A graph with a red line

AI-generated content may be incorrect.

Figure 6: The Q-Q plot compares the distribution of residuals to a normal distribution. The points mostly align with the red diagonal line, indicating that the residuals follow a normal distribution, with only a few deviations at the extremes. This supports the assumption of normality in regression residuals.

A graph with blue lines and a line

AI-generated content may be incorrect.

Figure C7: Histogram of VAR Residuals

The residuals of the VAR model appear to be somewhat normally distributed, but there are some deviations that could affect forecast accuracy.

A graph with green and purple lines

AI-generated content may be incorrect.

Figure C8: Forecast Comparison – OLS vs VAR vs ARIMA

This graph compares actual GDP (Inv) values with forecasts from three models: OLS, VAR, and ARIMA. The OLS model predictions (purple) appear to be significantly higher than the actual data, suggesting potential limitations in the model when applied to economic forecasting. The ARIMA model (red dashed line) and VAR model (blue dotted line) attempt to follow the actual data trends but show notable deviations.