

ST 3082 / DS 3003 Lab worksheet

7<sup>th</sup> November 2025

---

1. ‘**Income**’ dataset is a simulated dataset used in the reference book titled ‘*An Introduction to Statistical Learning*’.
  - a. Draw a suitable graph to explore the relationship between **Income** and years of education, **Education**.
  - b. Fit a simple linear regression model to predict the **Income** using **Education**. Illustrate this fitted line on the graph drawn in a).
  - c. Fit a polynomial regression with order 3 to predict **Income** using **Education**. Illustrate the fitted curve on the graph drawn in a).
  - d. Which model provides a better fit to the data?
  - e. Explore whether the variable **Seniority** is related with Income.
  - f. Draw a 3D plot to visually illustrate the relationship between all three variables.
  - g. Fit a linear regression model to predict **Income** using both **Education** and **Seniority**. Represent the fitted regression plane on the graph drawn in f).