

Assignment 1: Analysing the Relationship Between Education and Economic Output

Objective

Investigate the relationship between the level of education and its economic output at country level using simple linear regression analysis.

Research Question

How does the average years of schooling in a country correlate with its GDP per capita?

Data Source

World Bank's World Development Indicators (WDI) database:
<https://data.worldbank.org/indicator>

Assignment Steps

1. Data Collection (15% of grade)

- Access the World Bank data website
- Select the appropriate indicators
- Choose appropriate data period window
- Download the data

2. Data Preparation (20% of grade)

- Clean the dataset, (e.g. removing any countries with missing data)
- Organize the data into a structured format (e.g. Excel)

3. Exploratory Data Analysis (20% of grade)

- Calculate descriptive statistics (mean, median, standard deviation) for relevant variables
- Create a scatter plots
- Briefly describe any patterns or relationships you observe in the data

4. Simple Linear Regression Analysis (25% of grade)

- Conduct a simple linear regression using one independent variable (X) and GDP per capita as the dependent variable (Y)
- Calculate and report the regression equation: $Y = \beta_0 + \beta_1 X$
- Interpret the slope (β_1) and intercept (β_0)
- Calculate and report the coefficient of determination (R-squared)

- Perform and briefly discuss residual analysis to check the assumptions of linear regression

5. Interpretation and Discussion (20% of grade)

- Interpret the results in the context of the research question, based on data selected, and mention trade-offs and calls made (e.g. related to data available)
- Discuss the strength and direction of the relationship found
- Address at least three limitations of the analysis
- Suggest at least two potential confounding factors not captured in this simple model
- Propose at least two policy implications based on your findings

Deliverables

1. A report (1000-1500 words) detailing your methodology, findings, and discussion
2. An appendix with your data and any additional visualizations not included in the main report

Evaluation Criteria

- Accuracy and completeness of data collection and preparation
- Appropriateness and execution of statistical analyses
- Depth and insight of interpretation and discussion
- Clarity and professionalism of the report and Excel document

Due Date

Wednesday 13th November, 11:59pm

Good luck with your analysis!