Life Satisfaction Prediction using GDP per Capita

Assignment Note

This assignment is not intended to be a full project. It serves as a follow up to reinforce the concepts we have covered in lecture so far. More comprehensive analyses and detailed projects will be introduced as we progress further into the course.

🧐📚 Learning Objectives

- Load and visualize real-world data.
- Train and compare regression models using Scikit-Learn.
- Interpret model outputs and discuss generalization.

In [8]:

Part 1: Data Exploration and Visualization (20 marks)

Q1 (5 marks) Load the dataset `lifesat.csv` and display the first 5 rows.

In [ ]:

Q2 (5 marks) Print basic `info` and `summary statistics`.

In [ ]:

In [ ]:

Q3 (10 marks) Display a scatter plot for `GDP per capita` vs `Life Satisfaction`. Add labels, title, and discuss the observed relationship.

In [ ]:

Discuss the observed relationship:

In [ ]:

Part 2: Linear Regression Model (30 marks)

Q4 (5 marks) Extract input (X) and target (y). Print their shapes.

In [ ]:

Q5 (10 marks) Train a Linear Regression model & Display coefficient and intercept.

In [ ]:

Q6 (10 marks) Plot the predicted regression line from the model along with a scatter plot of the data.

In [ ]:

Q7 (5 marks) Predict Life Satisfaction for GDP = 37,655.2 USD. Comment on result.

In [ ]:

Part 3: K-Nearest Neighbors Regression (25 marks)

Q8 (5 marks) Train a KNeighborsRegressor (n_neighbors=3).

In [ ]:

Q9 (10 marks)

- Predict Life Satisfaction for GDP = 37,655.2 USD
- and compare with Linear Regression.

In [ ]:

Q10 (10 marks)

- Use n_neighbors 1, 3, 5, and 10 and print the predicted values of life satisfaction in Q10.
- Plot the results using a line plot.

In [ ]:

In [ ]:
