**Case study – Non-linear regression analysis**

**Introduction**

This case study focuses on addressing a non-linear relationship between two variables using non-linear regression analysis. The analysis aims to model the relationship between the independent variable x and the dependent variable y, where the relationship is not linear. The study compares the effectiveness of different polynomial regression models.

**Objective**

To perform non-regression analysis.

**Understanding the data**

The dataset consists of 15 observations with two columns: x (independent variable) and y (dependent variable). The values of x range from 1 to 15, while y shows a non-linear trend, peaking around x = 4 and then decreasing.

**Procedure for coding**

* Import the necessary libraries
* Create the dataset
* Visualize the scattered data
* Define the non-linear regression model
* Calculate MAPE and R-square for the model
* Visualize the model

**Code File Link**

<https://github.com/Ishita2003M/Non--linear-Regression-analysis/blob/main/non-linear-reg.ipynb>

**Interpretation and conclusion**

1 - The goal was to perform non-linear regression analysis on a given dataset to model the relationship between the independent variable (x) and the dependent variable (y).

2 - Dataset Overview 15 observations with columns:

* x (independent variable, range: 1–15)
* y (dependent variable, showing a non-linear trend, peaking around x = 4 and then decreasing).

3 - A scatter plot of the data was created to visualize the relationship between x and y. The plot confirmed a non-linear pattern, suggesting that a simple linear model might not capture the underlying trend effectively.

4 - Key Findings

1. Linear Model (1st Degree) – Poor Fit
   1. MAPE: 71.24% (high error)
   2. R²: 0.363 (only 36.3% variance explained)
   3. Conclusion: Fails to capture non-linearity.
2. Quadratic Model (2nd Degree) – Slight Improvement
   1. MAPE: 70.03% (still high)
   2. R²: 0.587 (better but insufficient)
   3. Conclusion: Better than linear but still not ideal.
3. Cubic Model (3rd Degree) – Strong Fit
   1. MAPE drops to 40.29% (much lower error)
   2. R² jumps to 0.831 (83.1% variance explained)
   3. Conclusion: Captures the peak and decline well.
4. Quartic Model (4th Degree) – Near-Perfect Fit
   1. MAPE: 15.15% (very low error)
   2. R²: 0.971 (97.1% variance explained)
   3. Conclusion: Best model so far, almost fully explains variability.
5. Quintic Model (5th Degree) – Diminishing Returns
   1. MAPE: 15.36% (slightly worse than quartic)
   2. R²: 0.971 (no real improvement)
   3. Conclusion: Risk of overfitting; quartic is preferable.

5 - Best Model: Quartic (4th Degree Polynomial)

* + Lowest MAPE (15.15%) and highest R² (0.971).
  + Effectively captures the non-linear trend without unnecessary complexity.