Title:

Advertising Impact on Sales

Dataset: This case study used the data in the file ***advertising_dataset.xlsx***

This file is available in Session Four: Introduction to Modelling.

Expected working time: ~45 minutes.

Context:

You work for a marketing agency that specialises in advertising analytics. Your client is a consumer goods company that wants to understand the impact of their advertising efforts on product sales. They collected data on advertising expenditures through different channels (TV, radio, and newspaper) and corresponding sales figures. Your task is to analyse the dataset and build regression models to predict future sales based on advertising spending.

Tasks:

1. Load the dataset and describe the data (number of points, number missing values, etc.).

2. Create scatter plots to visualize the relationships between each advertising channel (TV, radio, newspaper) and sales.

3. For each of the advertising channels, build a simple linear regression model to predict sales based only on that channel. Interpret the regression coefficients and evaluate the model's performance using the R-squared method.

4. Which of the three advertising channels explains better by itself the level of sales?

5. Build a multivariate linear regression model that includes all three advertising channels (TV, radio, newspaper) as independent variables to predict sales and evaluate the model's performance using the R-squared method. Is multivariate linear regression better than any of the models using only one variable?

6. Use the trained multivariate regression models to predict sales for new data points (predict datasheet). Which of the scenarios proposed in the datasheet forecasts a higher level of sales?