Sales Prediction Dataset

The dataset provided contains information about the advertising expenditures of a company on various platforms (TV, Radio, newspapers) and the corresponding sales of a product. Here's an explanation of the dataset:

1. TV: This column represents the amount of money spent on advertising the product on television. TV advertising is a traditional and widely used medium for reaching a broad audience.

2. Radio: This column indicates the advertising expenditure on radio. Radio advertising is known for its ability to target specific demographics and local audiences.

3. Newspaper: This column shows the advertising cost spent on newspaper advertising. Newspaper advertising is often used for targeting specific geographic regions or demographics.

4. Sales: This column represents the number of units sold corresponding to the advertising expenditures on TV, Radio, and newspapers.

Questions:

1. What is the average amount spent on TV advertising in the dataset?
2. What is the correlation between radio advertising expenditure and product sales?
3. Which advertising medium has the highest impact on sales based on the dataset?
4. Plot a linear regression line that includes all variables (TV, Radio, Newspaper) to predict Sales, and visualize the model's predictions against the actual sales values.
5. How would sales be predicted for a new set of advertising expenditures: $200 on TV, $40 on Radio, and $50 on Newspaper?
6. How does the performance of the linear regression model change when the dataset is normalized?
7. What is the impact on the sales prediction when only radio and newspaper advertising expenditures are used as predictors?