1. Which advertising medium has the highest impact on sales based on the dataset?

Correlation matrix for Sales expenditure in different media and Sales

	TV	Radio	Newspaper	Sales
TV	1.00	0.05	0.06	0.90
Radio	0.05	1.00	0.35	0.35
Newspaper	0.06	0.35	1.00	0.16
Sales	0.90	0.35	0.16	1.00
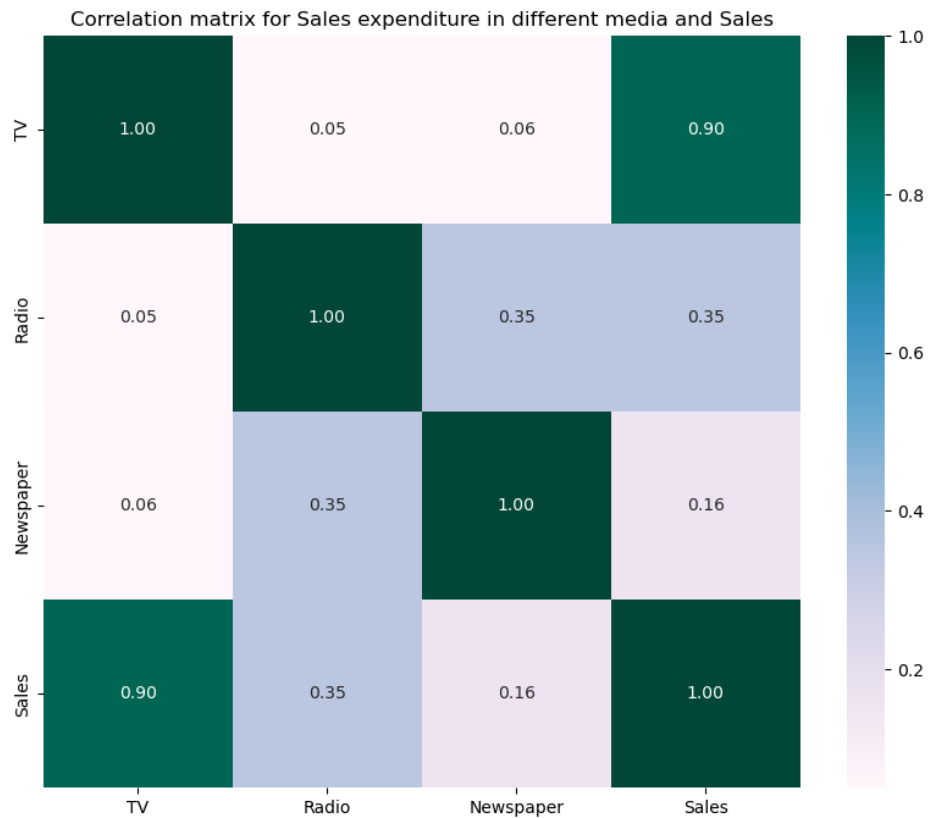
➢ TV medium expenditure has a very strong positive linear relationship with quantity of sales. So, when the expenditure grow, the quantity of sales also grow in equal proportion. So, it has the highest impact on sales.

➢ Radio expenditure has a moderate positive linear relationship with quantity of sales.

➢ Newspaper expenditure has a negligible relationship with quantity of sales.

**Figure 1**: Correlation matrix between advertising media and quantity of sales

2. 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.

Actual vs Predicted sales
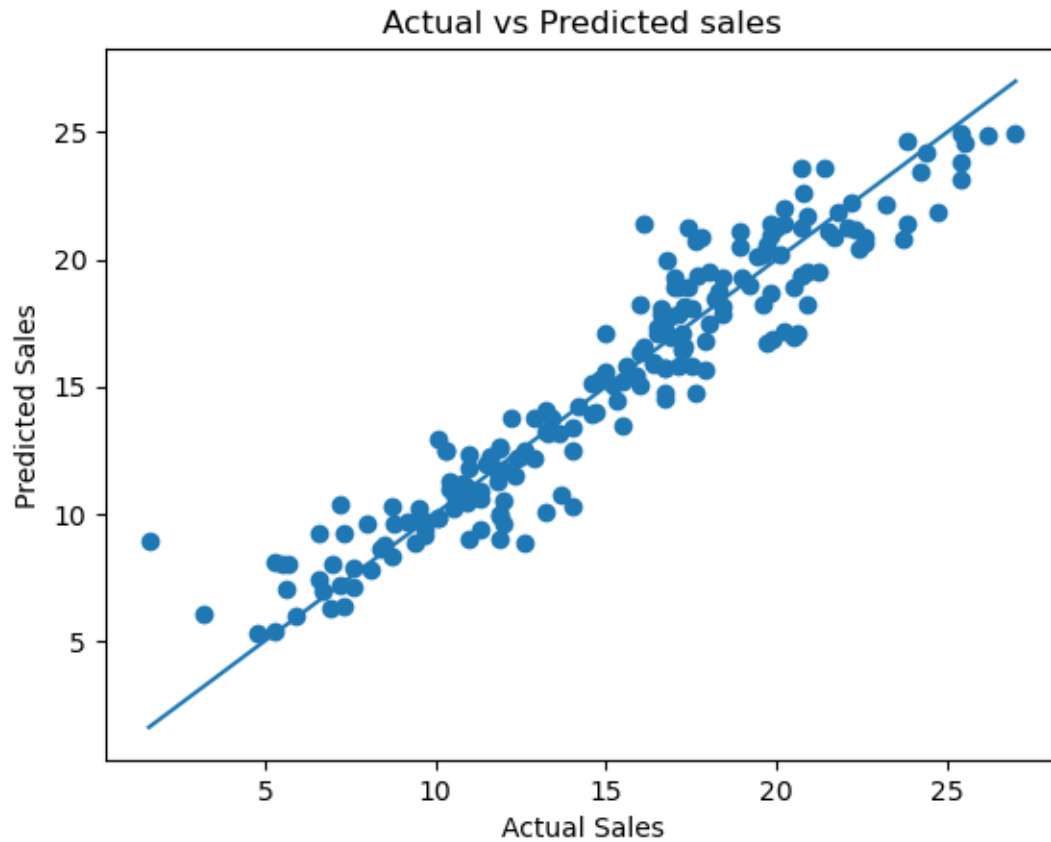
➢ The predicted sales values are near-around the actual sales values which indicate the prediction is good.

Figure 2: Scatter plots showing the predicted sales values and straight line (drawn at 45º) show the actual sales values.

3. How does the performance of the linear regression model change when the dataset is normalized?

Figure 2: Scatter plots showing the predicted sales values (red for unscaled features and green for scaled features) and straight line (drawn at 45º) show the actual sales values.
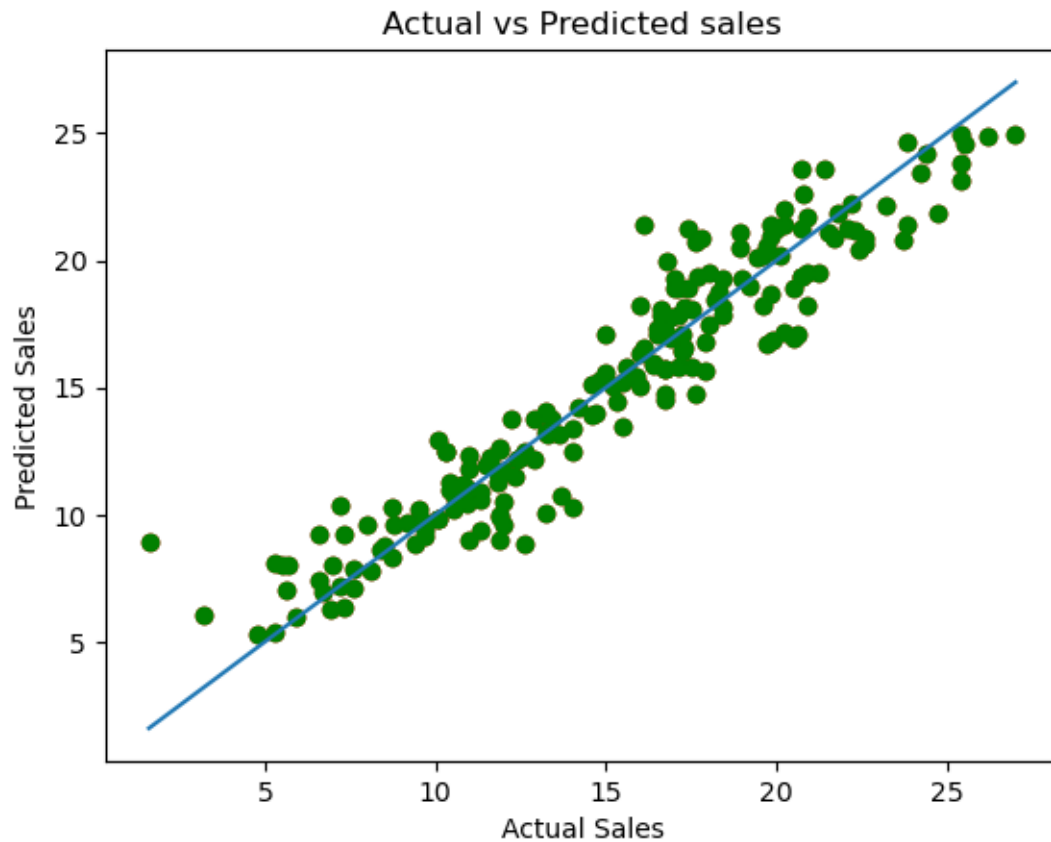
➢ The predicted sales values without unscaling features and with scaling features coincide exactly.

4. What is the impact on the sales prediction when only radio and newspaper advertising expenditures are used as predictors?

Figure 2: Scatter plots showing the predicted sales values (red is for all features and green is for only radio and newspaper advertising expenditures features) and straight line (drawn at 45°) show the actual sales values.
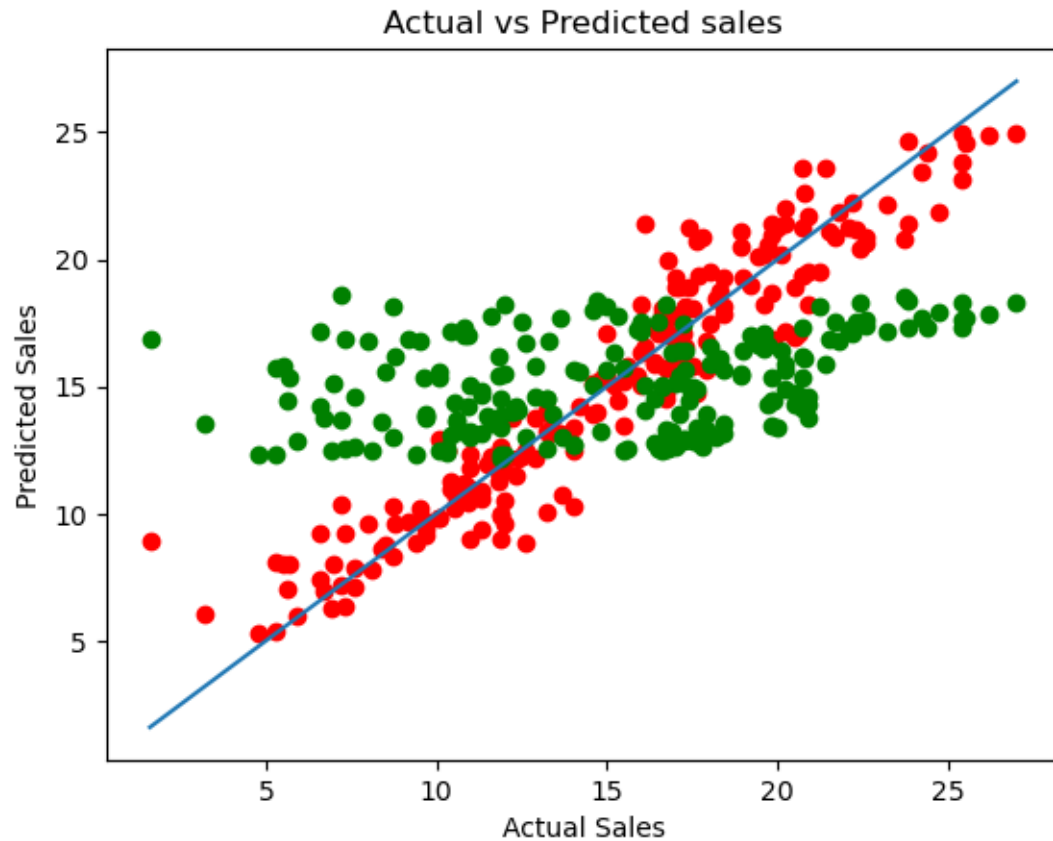
➢ The predicted sales values obtained by considering all the features are near-around the actual sales values while those predicted by considering only radio and newspaper advertising expenditures features are far away from actual sales values which indicate prediction become more inaccurate.