Practical Assignment No. 4	
Title:	Linear regression model for sales prediction
Problem Statement:	Build a simple linear regression model to predict sales based on TV marketing expenses.
Objective:	To apply mathematical concepts in computer science for solving the problems.
Outcome:	CO505.2: Use probabilistic models to solve the real-world problems
Software or Hardware Requirements:	Anaconda/Java/GCC
Theory:	**Simple linear regression model:** When modelling between the dependent and one independent variable, if there is only one independent variable in the linear regression model, the model is generally termed as a simple linear regression model. When there are more than one independent variable in the model, then the linear model is termed as the multiple linear regression model. Consider a simple linear regression model $y = \beta_0 + \beta_1 X + \varepsilon$ where y is termed as the **dependent** or study variable and X is termed as the **independent** or explanatory variable. The terms ***β***0 and ***β***1 are the parameters of the model. The parameter ***β***0 is termed as an intercept term, and the parameter ***β***1 is termed as the slope parameter. These parameters are usually called **regression coefficients**. The unobservable error component accounts for the failure of data to lie on a straight line and represents the difference between the true and observed realization of y . ***ε*** is observed as an independent and identically distributed random variable with mean zero and constant variance 2 Given the sample $\{(x_i, y_i);\ i = 1, 2, \ldots, n\}$, the least squares estimates b_0 and b_1 of the regression coefficients β_0 and β_1 are computed from the formulas $b_1 = \dfrac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$ and $b_0 = \dfrac{\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i}{n} = \bar{y} - b_1 \bar{x}.$

	Applications of Linear Regression ● Market analysis by using some marketing strategies and maximising sales ● Financial study through linear models for evaluating an establishment's operational performance ● Sports analysis by predicting game attendance depending on the team's status as well as market size ● Predicts the impact of water and air pollution on the environment ● Recognizes high-risk patients and improves healthy lifestyles
Input/Datasets/Test Cases:	Dataset Link - https://www.kaggle.com/datasets/devzohaib/tvmarketingcsv?select=tvmarketing.csv (Explain Data)
Results:	Write result values in table
Analysis and conclusion:	Write your own analysis of output and conclusion(Minimum 1 statement Analysis, Minimum 1 Statement Conclusion)
References:	Reference Links(Any 2)