Data Analytics / Tarun / 22nd May 2025 Imagine walking into a hospital where doctors already know what your body needs, even before you describe your symptoms. Sounds futuristic? Thanks to the growing benefits of data analytics in healthcare, this future is quickly becoming our reality. From predicting diseases to streamlining hospital operations, data is turning healthcare from reactive to truly proactive. We're [...] 8 Benefits of Data Analytics in Healthcare Read More »Accounting, Data Analytics / Tarun / 20th May 2025 In today's fast-paced digital world, data has become one of the most valuable assets for businesses, governments, and individuals alike. The uses of data analytics have expanded far beyond simple number crunching. It now plays a central role in how decisions are made, problems are solved, and strategies are developed across virtually every industry. Whether 18 Amazing Uses of Data Analytics Read More »General / Quinton Rose / 20th May 2025 If you are a marketer, business owner, or content creator, you understand the importance of videos in today's times. A solid promo video can make your products and services sell quickly. A promo video can grab attention and establish trust among your audience. However, it doesn't need to be difficult to create a solid promo Create a Promo Video Template That Sells – All in CapCut PC Read More »Data Analytics / Tarun / 19th May 2025 If you've ever wondered about the difference between these two roles, you're not alone. The conversation around data analytics vs data engineering is becoming more common, especially as careers in data continue to grow. While the terms might sound similar, they represent two very different parts of the data process. One focuses on analyzing and Data Analytics vs Data Engineering: Understanding the Key Differences Read More »Programming / Quinton Rose / 17th May 2025 In the world of programming, one of the most debated comparisons is between Python and C++. Whether you're building a web application, diving into artificial intelligence, or designing high-performance systems, choosing the right language is crucial. In this blog, we'll explore the difference between Python and C++, their speed, syntax, applications in AI and machine Python vs C++: Key Differences, Speed, Syntax, and AI Applications Read More »General / Quinton Rose / 15th May 2025 These days, if you're not building an email list and using email marketing to stay in touch, you're missing out on one of the most effective ways to connect with your audience. However, with the world's mail getting clogged up, how does one get his messages noticed? That is why the solution is to automate Email Automation Workflows: Best Practices And Examples Read More »Data Analytics / Tarun / 14th May 2025 In today's fast-paced digital world, a career in data analytics has become one of the most exciting and rewarding paths for those who enjoy working with numbers, solving real-world problems, and making sense of information. With the right mix of curiosity, dedication, and skill-building, anyone—from fresh graduates to mid-career professionals—can break into this field and How to Start a Career in Data Analytics Read More »Data Analytics / Tarun / 13th May 2025 In the digital age, data is more than just numbers on a spreadsheet—it's the backbone of decision-making in businesses, healthcare, education, marketing, and almost every industry You can AI-powered healthcare solutions that enhance patient outcomes and clinical efficiency In today's rapidly evolving healthcare landscape, artificial intelligence is no longer a futuristic concept but a transformative force revolutionizing patient care. AI-powered healthcare bots are at the forefront of this revolution, offering unprecedented opportunities to improve diagnosis accuracy, streamline administrative Revolutionizing Patient Care: How to Build a Powerful AI Healthcare Bot System Using Python in 2025 Read More »Data Analytics, Data Science / Tarun / 10th May 2025 When it comes to the world of data, the terms data science vs data analytics often confuse a lot of people, and it's easy to see why. Both fields deal with data, both help businesses make smart decisions, and both are in high demand. But here's the thing: they're not the same. In fact, understanding Data Science Vs Data Analytics: Which One To Choose? Read More » Share — copy and redistribute the material in any medium or format for any purpose, even commercially. Adapt — remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Attribution — You must give appropriate credit , provide a link to the license, and indicate if changes were made . You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original. No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits. You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation . No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material. Linear regression is one of the most commonly used techniques in statistics. It is used to quantify the relationship between one or more predictor variables and a response variable. The most basic form of linear is regression is known as simple linear regression, which is used to quantify the relationship between one predictor variable and one response variable. If we have more than one predictor variable then we can use multiple linear regression, which is used to quantify the relationship between several predictor variables and a response variable. This tutorial shares four different examples of when linear regression is used in real life. Linear Regression Real Life Example #1 Businesses often use linear regression to understand the relationship between advertising spending and revenue. For example, they might fit a simple linear regression model using advertising spending as the predictor variable and revenue as the response variable. The regression model would take the following form: revenue = β0 + β1(ad spending) The coefficient β0 would represent total expected revenue when ad spending is zero. The coefficient β1 would represent the average change in total revenue when ad spending is increased by one unit (e.g. one dollar). If β1 is negative, it would mean that more ad spending is associated with less revenue. If β1 is close to zero, it would mean that ad spending has little effect on revenue. And if β1 is positive, it would mean more ad spending is associated with more revenue. Depending on the value of β1, a company may decide to either decrease or increase their ad spending. Linear Regression Real Life Example #2 Medical researchers often use linear regression to understand the relationship between drug dosage and blood pressure of patients. For example, researchers might administer various dosages of a certain drug to patients and observe how their blood pressure responds. They might fit a simple linear regression model using dosage as the predictor variable and blood pressure as the response variable. The regression model would take the following form: blood pressure = β0 + β1(dosage) The coefficient β0 would represent the expected blood pressure when dosage is zero. The coefficient β1 would represent the average change in blood pressure when dosage is increased by one unit. If β1 is negative, it would mean that an increase in dosage is associated with a decrease in blood pressure. If β1 is close to zero, it would mean that an increase in dosage is associated with no change in blood pressure. If β1 is positive, it would mean that an increase in dosage is associated with an increase in blood pressure. Depending on the value of β1, researchers may decide to change the dosage given to a patient. Linear Regression Real Life Example #3 Agricultural scientists often use linear regression to measure the effect of fertilizer and water on crop yields. For example, scientists might use different amounts of fertilizer and water on different fields and see how it affects crop yield. They might fit a multiple linear regression model using fertilizer and water as the predictor variables and crop yield as the response variable. The regression model would take the following form: crop yield = β0 + β1(amount of fertilizer) + β2(amount of water) The coefficient β0 would represent the expected crop yield with no fertilizer or water. The coefficient β1 would represent the average change in crop yield when fertilizer is increased by one unit, assuming the amount of water remains unchanged. The coefficient β2 would represent the average change in crop yield when water is increased by one unit, assuming the amount of fertilizer remains unchanged. Depending on the values of β1 and β2, the scientists may change the amount of fertilizer and water used to maximize the crop yield. Linear Regression Real Life Example #4 Data scientists for professional sports teams often use linear regression to measure the effect that different training regimens have on player performance. For example, data scientists in the NBA might analyze how different amounts of weekly yoga sessions and weightlifting sessions affect the number of points a player scores. They might fit a multiple linear regression model using yoga sessions and weightlifting sessions as the predictor variables and total points scored as the response variable. The regression model would take the following form: points scored = β0 + β1(yoga sessions) + β2(weightlifting sessions) The coefficient β0 would represent the expected points scored for a player who participates in zero yoga sessions and zero weightlifting sessions. The coefficient β1 would represent the average change in points scored when weekly yoga sessions is increased by one, assuming the number of weekly weightlifting sessions remains unchanged. The coefficient β2 would represent the average change in points scored when weekly weightlifting sessions is increased by one, assuming the number of weekly yoga sessions remains unchanged. Depending on the values of β1 and β2, the data scientists may recommend that a player participates in more or less weekly yoga and weightlifting sessions in order to maximize their points scored. Conclusion Linear regression is used in a wide variety of real-life situations across many different types of industries. Fortunately, statistical software makes it easy to perform linear regression. Feel free to explore the following tutorials to learn how to perform linear regression using different softwares: How to Perform Simple Linear Regression in Excel How to Perform Multiple Linear Regression in Excel How to Perform Multiple Linear Regression in R How to Perform Multiple Linear Regression in Stata How to Perform Linear Regression on a TI-84 Calculator Our platform connects you with top statistical experts from around the globe, ensuring that you receive high-quality guidance for all your projects and tasks. Whether you need help with data analysis, research methodologies, predictive modeling, or any other statistical challenge, our professionals are here to assist you.. Services We Provide Live Tutoring Question & Answer help Programming Help Project based Freelance Work Linear regression is one of the most commonly used techniques in statistics. It is used to quantify the relationship between one or more predictor variables and a response variable. The most basic form of linear is regression is known as simple linear regression, which is used to quantify the relationship between one predictor variable and one response variable. If we have more than one predictor variable then we can use multiple linear regression, which is used to quantify the relationship between several predictor variables and a response variable. This tutorial shares four different examples of when linear regression is used in real life. Linear Regression Real Life Example #1 Businesses often use linear regression to understand the relationship between advertising spending and revenue. For example, they might fit a simple linear regression model using advertising spending as the predictor variable and revenue as the response variable. The regression model would take the following form: revenue = β0 + β1(ad spending) The coefficient β0 would represent total expected revenue when ad spending is zero. The coefficient β1 would represent the average change in total revenue when ad spending is increased by one unit (e.g. one dollar). If β1 is negative, it would mean that more ad spending is associated with less revenue. If β1 is close to zero, it would mean that ad spending has little effect on revenue. And if β1 is positive, it would mean more ad spending is associated with more revenue. Depending on the value of β1, a company may decide to either decrease or increase their ad spending. Linear Regression Real Life Example #2 Medical researchers often use linear regression to understand the relationship between drug dosage and blood pressure of patients. For example, researchers might administer various dosages of a certain drug to patients and observe how their blood pressure responds. They might fit a simple linear regression model using dosage as the predictor variable and blood pressure as the response variable. The regression model would take the following form: blood pressure = β0 + β1(dosage) The coefficient β0 would represent the expected blood pressure when dosage is zero. The coefficient β1 would represent the average change in blood pressure when dosage is increased by one unit. If β1 is negative, it would mean that an increase in dosage is associated with a decrease in blood pressure. If β1 is close to zero, it would mean that an increase in dosage is associated with no change in blood pressure. If β1 is positive, it would mean that an increase in dosage is associated with an increase in blood pressure. Depending on the value of β1, researchers may decide to change the dosage given to a patient. Linear Regression Real Life Example #3 Agricultural scientists often use linear regression to measure the effect of fertilizer and water on crop yields. For example, scientists might use different amounts of fertilizer and water on different fields and see how it affects crop yield. They might fit a multiple linear regression model using fertilizer and water as the predictor variables and crop yield as the response variable. The regression model would take the following form: crop yield = β0 + β1(amount of water) + β2(amount of water) The coefficient β0 would represent the expected crop yield with no fertilizer or water. The coefficient β1 would represent the average change in crop yield when fertilizer is increased by one unit, assuming the amount of water remains unchanged. The coefficient β2 would represent the average change in crop yield when water is increased by one unit, assuming the amount of fertilizer remains unchanged. Depending on the values of β1 and β2, the scientists may change the amount of fertilizer and water used to maximize the crop yield. Linear Regression Real Life Example #4 Data scientists for professional sports teams often use linear regression to measure the effect that different training regimens have on player performance. For example, data scientists in the NBA might analyze how different amounts of weekly yoga sessions and weightlifting sessions affect the number of points a player scores. They might fit a multiple linear regression model using yoga sessions and weightlifting sessions as the predictor variables and total points scored as the response variable. The regression model would take the following form: points scored = β0 + β1(yoga sessions) + β2(weightlifting sessions) The coefficient β0 would represent the expected points scored for a player who participates in zero yoga sessions and zero weightlifting sessions. The coefficient β1 would represent the average change in points scored when weekly yoga sessions is increased by one, assuming the number of weekly weightlifting sessions remains unchanged. The coefficient β2 would represent the average change in points scored when weekly weightlifting sessions is increased by one, assuming the number of weekly yoga sessions remains unchanged. Depending on the values of β1 and β2, the data scientists may recommend that a player participates in more or less weekly yoga and weightlifting sessions in order to maximize their points scored. Conclusion Linear regression is used in a wide variety of real-life situations across many different types of industries. Fortunately, statistical software makes it easy to perform linear regression. Feel free to explore the following tutorials to learn how to perform linear regression using different softwares: How to Perform Simple Linear Regression in Excel How to Perform Multiple Linear Regression in Excel How to Perform Multiple Linear Regression in R How to Perform Multiple Linear Regression in Stata How to Perform Linear Regression on a TI-84 Calculator Linear regression is a fundamental statistical and machine-learning technique used to establish relationships between two or more variables. It is especially useful for making predictions and understanding how one variable influences another. In this article, we'll delve into the concept of linear regression and explain it using real-life examples.What is Linear Regression?Linear regression is a statistical method used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. This equation is represented as:[Y = aX + b]Where:(Y) is the dependent variable.(X) is the independent variable.(a) is the slope (coefficient) of the line.(b) is the y-intercept.Simple Linear RegressionSimple linear regression is the case when there is only one independent variable. Let's explore a simple example to understand this concept.Predicting House PricesImagine you want to predict the selling price of a house based on its square footage. In this scenario:Y is the selling price.X is the square footage.A and b are the regression coefficients that we will determine from the data.We collect data on various houses, noting their square footage and selling prices. Using simple linear regression, we can determine the best-fit line that represents this relationship, allowing us to predict house prices based on square footage.Read Blog: Difference B/w Regression and RegressivenessMultiple Linear RegressionIn many real-world situations, multiple independent variables affect the dependent variable. Multiple linear regression allows us to model these more complex relationships.Predicting GDPSuppose you want to predict a country's GDP based on several factors, such as population, inflation rate, and foreign direct investment. Here:Y is the GDP.X_1, X_2, and X_3 are the population, inflation rate, and foreign direct investment, respectively.a_1, a_2, a_3, and b are the regression coefficients.Multiple linear regression helps us understand how these factors collectively influence a country's GDP.Interpretation and LimitationsLinear regression provides valuable insights, but it has its limitations. It assumes a linear relationship between variables, and real-life data often doesn't fit this assumption perfectly. Additionally, it's sensitive to outliers, multicollinearity, and other statistical issues.Read Blog: What is Logistic Regression in Machine LearningConclusionLinear regression is a powerful tool for understanding and predicting relationships between variables in various fields, from economics to medicine. By using real-life examples, we've demonstrated how simple and multiple linear regression can be applied to make informed decisions and predictions. Keep in mind that while linear regression is a valuable technique, it's essential to consider its assumptions and limitations when applying it to real-world problems. A simple Linear Regression model is one of the most fundamental techniques in machine learning and statistics. Whether you are a data science newbie or just brushing up on the basics, understanding linear regression is essential. Frequently, we measure two or more variables on each individual and try to express the nature of the relationship between these variables (for example, in the simple linear regression model and correlation analysis). Using the regression technique, we estimate the relationship of one variable with another by expressing the one in terms of a linear (or more complex) function of another. We also predict the values of one variable in terms of the other. The variables involved in regression and correlation analysis are continuous. In this post, we will learn about the Simple Linear Regression Model. We are interested in establishing significant functional relationships between two (or more) variables. For example, the function $Y=f(X)=a+bx$ (read as Y is a function of X) establishes a relationship to predict the values of variable Y for the given values of variable X. In statistics (biostatistics), the function is a simple linear regression model or the regression equation. The variable Y is called the dependent (response) variable, and X is called the independent (regressor or explanatory) variable. In biology, many relationships can be appropriate over only a limited range of values of X. Negative values are meaningless in many cases, such as age, height, weight, and body temperature. The method of linear regression is used to estimate the best-fitting straight line to describe the relationship between variables. The linear regression gives the equation of the straight line that best describes how the outcome of Y increases/decreases with an increase/decrease in the explanatory variable X. The equation of the regression line is$$Y=\beta_0 + \beta_1 X,$$where β_0 is the intercept (value of Y when $X=0$) and β_1 is the slope of the line. Both β_0 and β_1 are the parameters (or regression coefficients) of the linear equation. The best-fitting line is derived using the method of the \textit{Least Squares} by finding the values of the parameters β_0 and β_1 that minimize the sum of the squared vertical distances of the points from the regression line, The dotted-line (best-fit) line passes through the point ($\overline{X}, \overline{Y}$). The regression line $Y=\beta_0+\beta_1X$ is fit by the least-squares methods. The regression coefficients β_0 and β_1 are both calculated to minimize the sum of squares of the vertical deviations of the points about the regression line. Each deviation equals the difference between the observed value of Y and the estimated value of Y (the corresponding point on the regression. The following table shows the \textit{body weight} and \textit{plasma volume} of eight healthy men. SubjectBody Weight (KG)Plasma Volume (liters)158.02.75270.02.86374.03.37463.52.76562.02.62670.53.49771.03.05866.03.12 The parameters β_0 and β_1 are estimated using the following formula (for simple linear regression model): \begin{align} \beta_1 &= \frac{n\sum\limits_{i=1}^{n} x_iy_i -\sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i} {n \sum\limits_{i=1}^{n} x_i^2 – \left(\sum\limits_{i=1}^{n} x_i \right)^2}\\\beta_0 &= \overline{Y} – \beta_1 \overline{X} \end{align} Regression coefficients are sometimes known as "beta-coefficients". When the slope ($\beta_1=0$, then there is no relationship between X and Y variables. For the data above, the best-fitting straight line describing the relationship between plasma volume with body weight is$$Plasma\, Volume = 0.0857 +0.0436\times Weight$$Note that the calculated values for β_0 and β_1 are estimates of the population values and, therefore, subject to sampling variations. Real Estate: Predicting House Prices (Estimate home prices based on size to guide buyers and sellers.)Independent Variable (X): Size of the house (sq ft) Dependent Variable (Y): Price of the house Education: Predicting Student Scores (Teachers or students can predict likely outcomes based on study habits.)X: Hours studied Y: Exam scores Healthcare: Predicting Blood Pressure (Understand how blood pressure tends to rise with age, aiding diagnosis.)X: Age of patient Y: Systolic blood pressure Energy: Predicting Electricity Usage (Power companies use this to forecast demand and manage resources) X: Temperature (°C or °F)Y: Electricity consumption (kWh) Manufacturing: Predicting Machine Failures X: Hours a machine has been in use (Predict maintenance schedules and avoid production delays.)Y: Number of breakdowns or wear percentage Business: Predicting Sales Based on Advertising Spend (Helps businesses decide how much to invest in advertising.)X: Advertising expenditure (in $\$$) Y: Product sales (in units) Agriculture: Predicting Crop Yield (Estimate yield based on expected rainfall to plan for food production.)X: Amount of rainfall (mm) Y: Crop yield (kg per acre) Finance: Predicting Stock Prices (Although basic, it helps in forecasting trends over time (note: simple linear regression has limits in volatile markets))X: Time (days or months) Y: Stock closing price Transportation: Estimating Fuel Consumption (Predict fuel needs and optimize transportation costs.)X: Distance traveled (km) Y: Fuel used (liters) E-commerce: Predicting Customer Spending (Analyze user behavior and optimize website experience for better conversion.)X: Time spent on the website Y: Amount spent on a purchase , This dataset of size n = 51 are for the 50 states and the District of Columbia in the United States (poverty.txt). The variables are y = year 2002 birth rate per 1000 females 15 to 17 years old and x = poverty rate, which is the percent of the state's population living in households with incomes below the federally defined poverty level. (Data source: Mind On Statistics, 3rd edition, Utts and Heckard). The plot of the data below (birth rate on the vertical) shows a generally linear relationship, on average, with a positive slope. As the poverty level increases, the birth rate for 15 to 17 year old females tends to increase as well. The following plot shows a regression line superimposed on the data. The equation of the fitted regression line is given near the top of the plot. The equation should really state that it is for the "average" birth rate (or "predicted" birth rate would be okay too) because a regression equation describes the average value of y as a function of one or more x-variables. In statistical notation, the equation could be written \(\hat{y} = 4.267 + 1.373x \). The interpretation of the slope (value = 1.373) is that the 15 to 17 year old birth rate increases 1.373 units, on average, for each one unit (one percent) increase in the poverty rate. The interpretation of the intercept (value=4.267) is that if there were states with poverty rate = 0, the predicted average for the 15 to 17 year old birth rate would be 4.267 for those states. Since there are no states with poverty rate = 0 this interpretation of the intercept is not practically meaningful for this example. In the graph with a regression line present, we also see the information that s = 5.55057 and r2 = 53.3%. The value of s tells us roughly the standard deviation of the differences between the y-values of individual observations and predictions of y based on the regression line. The value of r2 can be interpreted to mean that poverty rates "explain" 53.3% of the observed variation in the 15 to 17 year old average birth rates of the states. The R2 (adj) value (52.4%) is an adjustment to R2 based on the number of x-variables in the model (only one here) and the sample size. With only one x-variable, the adjusted R2 is not important. Example 2: Lung Function in 6 to 10 Year Old Children The data are from n = 345 children between 6 and 10 years old. The variables are y = forced exhalation volume (FEV), a measure of how much air somebody can forcibly exhale from their lungs, and x = age in years. (Data source: The data here are a part of dataset given in Kahn, Michael (2005). "An Exhalent Problem for Teaching Statistics", The Journal of Statistical Education, 13(2). Below is a plot of the data with a simple linear regression line superimposed. The estimated regression equation is that average FEV = 0.01165 + 0.26721 × age. For instance, for an 8 year old we can use the equation to estimate that the average FEV = 0.01165 + 0.26721 × (8) = 2.15. The interpretation of the slope is that the average FEV increases 0.26721 for each one year increase in age (in the observed age range). An interesting and possibly important feature of these data is that the variance of individual y-values from the regression line increases as age increases. This feature of data is called non-constant variance. For example, the FEV values of 10 year olds are more variable than FEV value of 6 year olds. This is seen by looking at the vertical ranges of the data in the plot. This may lead to problems using a simple linear regression model for these data, which is an issue we'll explore in more detail in Lesson 4. Above, we only analyzed a subset of the entire dataset. The full dataset (fev_dat.txt) is shown in the plot below: As we can see, the range of ages now spans 3 to 19 years old and the estimated regression equation is FEV = 0.43165 + 0.22204 × age. Both the slope and intercept have noticeably changed, but the variance still appears to be non-constant. This illustrates that it is important to be aware of how you are analyzing your data. If you only use a subset of your data that spans a shorter range of predictor values, then you could obtain noticeably different results than if you had used the full dataset. You might have read lot of tutorials on Linear Regression and already have the assumption - Linear Regression is not easy to Understand. We will make Linear Regression very easy for you. Let's boil down each concept and learn with help of Examples. If you have no idea what Linear regression is, this tutorial will be help you understand the basics. Linear regression might sound like a complex term, but it's actually a very simple concept. Linear Regression is all about finding patterns in data. When two things are connected, (like - hours of study and test scores, OR temperature and ice cream sales) linear regression helps us understand and predict how one affects the other.Basically, Linear Regression is asking if Thing-1 will change, how Thing-2 will respond? Answer of this question is often found by drawing a straight line through data points on a graph.How Does Linear Regression Work?Linear regression helps us answer questions about relationships in data. For example:Is there a consistent connection between the amount of time you spend studying and your test scores?Can we predict future trends based on past data?This is done by identifying two types of variables:Independent Variable: The thing we control or know (e.g., hours studied).Dependent Variable: The thing we want to predict (e.g., test scores).Linear regression tries to find the best-fit line through the data. This line is like a rule or formula that tells us:When the independent variable (e.g., hours studied) increases, how much does the dependent variable (e.g., test scores) increase or decrease?If we know the independent variable's value, what's the most likely value for the dependent variable?What is Best Fit Line?Among all possible lines you could draw through the data, linear regression finds the one that minimizes the errors (the gaps between the line and the points). This is called the line of best fit.Understand Linear Regression with Help of Example The more time you spend in studying, the better your test scores. Linear regression helps us find the relationship between these two things and use that relationship to make predictions.Think of it like this:You collect some data: how many hours you studied and the scores you got on tests.You plot this data on a graph.Then, you draw a straight line through the points in such a way that it's as close as possible to all the points. This line shows the trend.Once you have this line, you can use it to make predictions. For example, if you studied 5 hours for a test, the line can help you estimate what score you're likely to get.Intuition Behind Linear RegressionLet's say you're tracking how much time you spend studying and the test scores you get. You gather the following data:Hours of Study Test Score 2 50 4 70 6 90If you plot this data on a graph:The x-axis represents the hours of study.The y-axis represents the test score.You will see the points roughly form a straight-line pattern. Linear regression helps us draw the best possible straight line through these points. Once we have the line, we can use it to predict scores for other study hours, like 3, 5, or even 8 hours.Math Behind Linear RegressionThe equation of a straight line is: where,y: The value we want to predict (your test score).x: The value we know (hours of study).m: The slope of the line (how much yy changes when xx changes by 1 unit).c: The y-intercept (the value of yy when x=0x=0). Step 1: Find the Slope (m)The slope shows how much y (test score) changes for every 1 unit increase in x (study hours). From the data:Change in x Change in yFrom 2 to 4 From 50 to 70From 4 to 6 From 70 to 90The slope is:m = Change in y / Change in x = 20 / 2 =10So, for every extra hour of studying, your test score increases by 10 points.Step 2: Find the Y-Intercept (c)The y-intercept is the value of y when x=0 (what your score would be if you didn't study at all). Let's use one of the data points, say (2, 50) to find c :y = mx + c50 = 10(2) + c50 = 20+ cc = 30So, the equation of the line is y = 10x + 30Step 3: Using the Equation to Make PredictionsNow that we have the equation y=10x+30, we can use it to predict test scores for any amount of study time.If you study for 3 hours: y = 10(3) + 30 = 60If you study for 5 hours: y = 10(5) + 30 = 80So, you can expect a score of 80, if you study for 5 Hours.If you want to score at least 90, how much should you study?90 = 10x + 3090 - 30 = 10xx = 60/10 = 6 hours. In above example:The slope (m=10) tells us how much the score improves for each extra hour of study.The y-intercept (c=30) tells us the starting score when no studying is done.Goal of Linear RegressionThe main goal of linear regression is to find the values of m (slope) and c (y-intercept) that define the best-fit line. Once we have these values, we can:Understand the relationship between x and y.Make predictions about y for any given value of x.For example:If m>0, it means there's a positive relationship (as x increases, y also increases).If m