greatlearning
Learning for Life

DSBA CURRICULUM DESIGN

FOUNDATIONS

Data Science Using Python

Statistical Methods for Decision Making

CORE COURSES

Advanced Statistics

Data Mining

Predictive Modelling(Week-1/ 5)

Machine Learning

Time Series Forecasting

Data Visualization

DOMAIN APPLICATIONS

Financial Risk Analytics

Web & Social Media Analytics

Marketing Retail Analytics

greatlearning
Learning for Life

LEARNING OBJECTIVE OF THIS MODULE

- Linear Regression
- Logistics Regression
- LDA

greatlearning
Learning for Life

LEARNING OBJECTIVES OF THIS SESSION

- Multiple Linear Regression
- Concept of R2

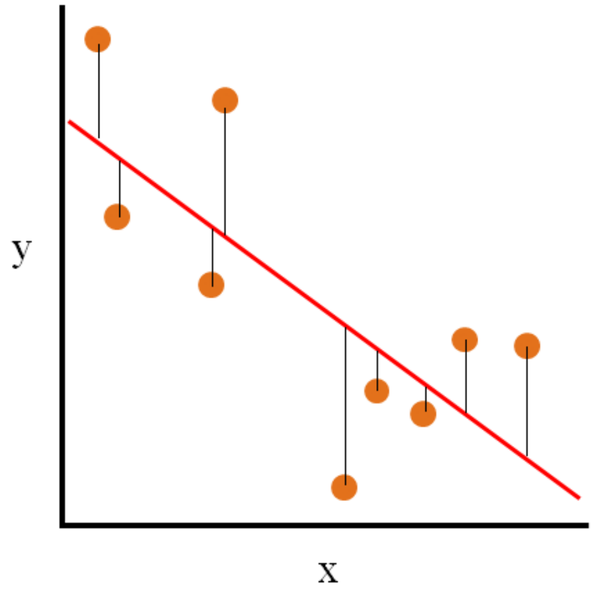
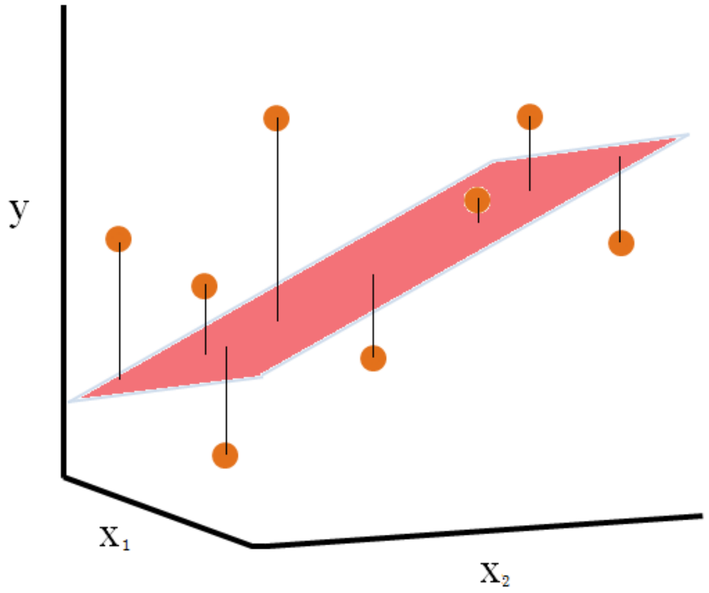
greatlearning
Learning for Life

TRY ANSWERING THE FOLLOWING

- What is the equation of simple linear regression?
- Can we have >2 independent variables in simple linear regression?
- Name few techniques to deal with multicollinearity in a data.

greatlearning
Learning for Life

BROAD OVERVIEW

Simple Linear Regression	Multiple Linear Regression (2 Independent Variables (x_1, x_2))
	
$y = b_0 + b_1 * x_1$	$y = b_0 + b_1 * x_1 + b_2 * x_2 + \ldots + b_n * x_n$

greatlearning
Learning for Life

Industry Application - Linear Regression in Sports Analytics

Sports analytics is a booming field. Owners, coaches, and fans are using statistical measures and models of all kinds to study the performance of players and teams.

Billy Beane the manager of Oakland As met Peter Brand, a young Yale economics graduate with radical ideas about how to assess player value. Rather than relying on selector's experience and intuition, Brand uses Linear Regression, selecting players based on their on-base percentage (OBP) while ignoring their perceived weaknesses. Brand and Beane use this methodology to hire undervalued players.

With one third the budget Billy Beans' team performed at par with the highest paid teams. This led to evolution of a new stream of Analytics known as Sabermetics.
You may watch the Brad Pitt starrer Moneyball for Reference.

Reference:https://en.wikipedia.org/wiki/Sabermetrics

greatlearning
Learning for Life

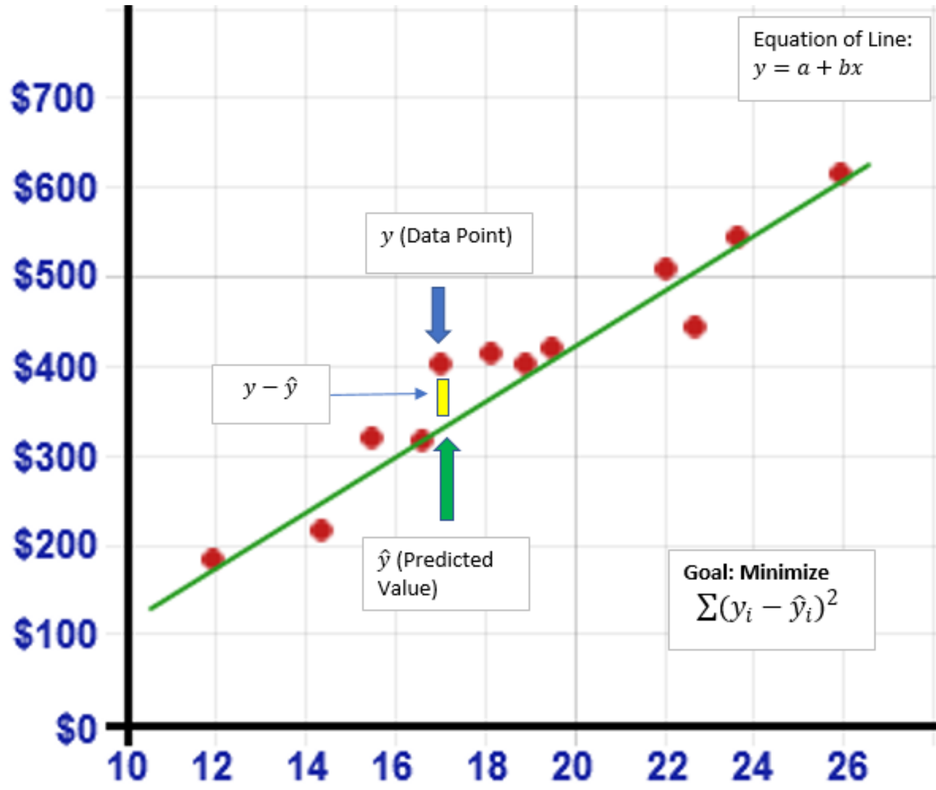
Industry Application - Linear Regression in Financial Markets

The financial performance of a company is a primary concern for every stakeholder especially for investors. The measurement of the financial health of a company through the reported financial statements gives a qualitative analysis of the company's position as well as an account of how the company has utilised its capital in production.

These Financial Statements can be used to derive critical ratios that are broadly divided into following categories:

- Liquidity ratios
- Leverage ratios
- Efficiency ratios
- Profitability ratios
- Market value ratios

We can obtain these ratios and the corresponding stock prices of various listed companies to create a Linear Regression model which predicts the stock price given we know the ratios.



Reference: https://pdfs.semanticscholar.org/0a82/305452d2fa7cdca91bdbbe6f409905ace42a.pdf

greatlearning
Learning for Life

CASE STUDY - Price Prediction(Airbnb)

Airbnb, Inc is an online marketplace for arranging or offering lodging, primarily homestays, or tourism experiences. Airbnb has close to 150 million customers across the world. Price is the most important factor considered by the customer while making booking into a property. Strategic pricing of the properties is important to avoid losing customers to the competitors.

We have a data of 74111 Airbnb properties across the nations. Based on this data build a simple and multiple linear regression model to predict the strategic pricing of a new listed property on Airbnb.

greatlearning
Learning for Life

ANY QUESTIONS

greatlearning
Learning for Life

HAPPY LEARNING