CUSTOMER CHURN ANALYSIS

From Data to Strategy

Mentors:

Rashi Sharma

Ritik Shah

Mudit Agrawal

Mridul Gupta

SnT SUMMER PROJECTS - 2024

INDEX

DATA SCIENCE

The field of data science involves using scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data. This multidisciplinary field spans various domains, including statistical analysis, machine learning, and data visualization.

1. Introduction to libraries:

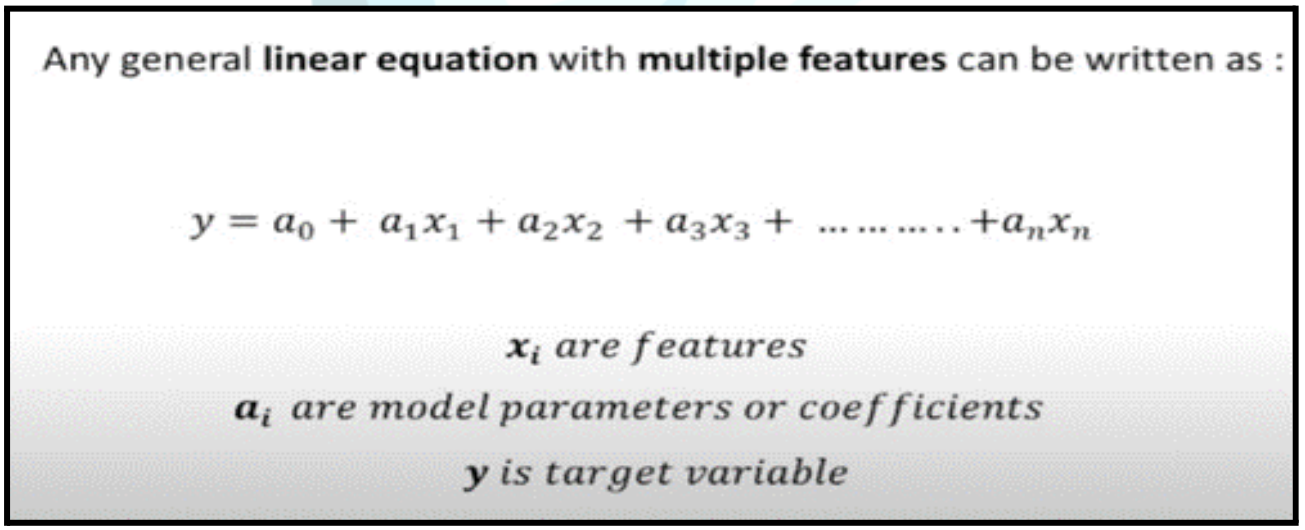
We learnt about Python Libraries, useful for data Analysis - Numpy, Pandas, Matplotlib and Seaborn.

2. Assignment 1:

A basic assignment to understand the fundamentals of numpy and pandas in data analysis.

3. INTRODUCTION TO REGRESSION

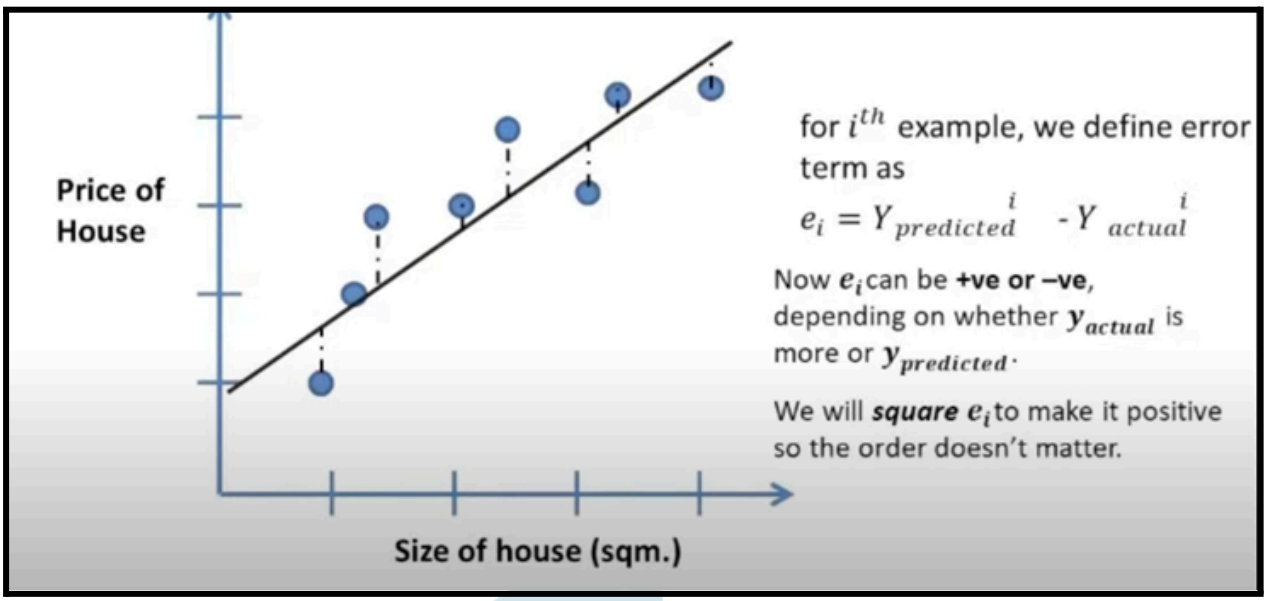
Regression analysis is a statistical method used to understand the relationships between variables. It helps in predicting a dependent variable based on the value of one or more independent variables. This technique is crucial in various fields such as finance, biology, and social sciences for making informed decisions.



LINEAR REGRESSION

Linear regression is a basic and commonly used type of predictive analysis. The overall idea is to examine two things: (1) does a set of predictor variables do a good job in predicting an outcome (dependent) variable? (2) Which variables in particular are significant predictors of the outcome variable, and in what way do they influence it? These regression

estimates are used to explain the relationship between one dependent variable and one or more independent variables.



OLS METHOD

The Ordinary Least Squares (OLS) method is a type of linear least squares method for estimating the unknown parameters in a linear regression model. OLS chooses the parameters of a linear function of a set of explanatory variables by minimizing the sum of the squares of the differences between the observed dependent variable and those predicted by the linear function.

We take the square of each of the errors to -

1. Emphasise Larger Errors
2. Ensure Non-Negative Values
3. Mathematical Differentiability
4. Convexity

Cost function (J) $= \frac{1}{2m}(e_1^2 + e_2^2 + e_3^2 + \ldots\ldots\ldots\ldots + e_m^2)$

$$J(a) = \frac{1}{2m} \sum_{i=1}^{m} (y_{i\,(pre)} - y_{i\,(act)})^2$$

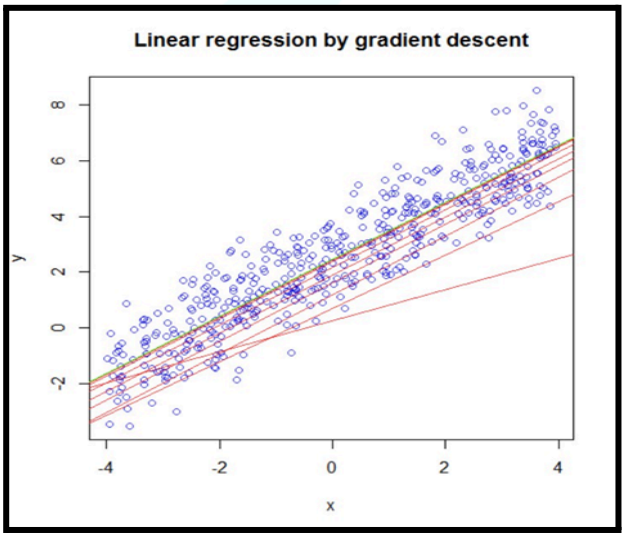
$$J(a) = \frac{1}{2m} \sum_{i=1}^{m} (a_0 + a_1 x_1^{(i)} - y_{i\,(act)})^2$$

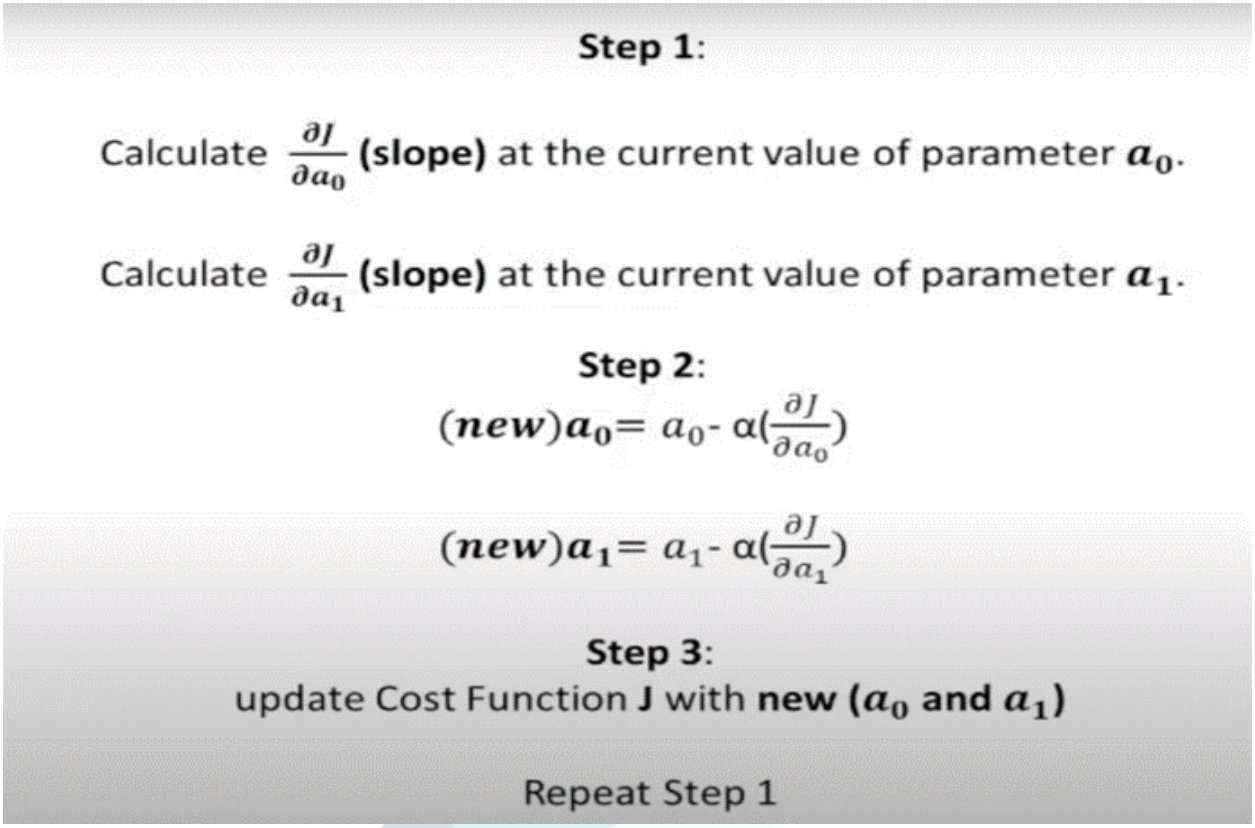
, **Cost function (J)** is a function of parameter space $a = (a_0, a_1)$.

Formulations for performing linear regression using the OLS Method

GRADIENT DESCENT

Gradient Descent is an optimization algorithm used for minimizing the cost function in various machine learning algorithms. It iteratively adjusts the parameters of the model to find the minimum value of the cost function, thereby improving the accuracy of the model. Gradient Descent is an optimisation algorithm used to minimize the cost function in various machine learning algorithms by iteratively adjusting the parameters to find the minimum value, thereby improving model accuracy. It is especially useful for large datasets and complex models.





Formulations for performing linear regression using Gradient Descent

PROBLEMS WITH LINEAR REGRESSION

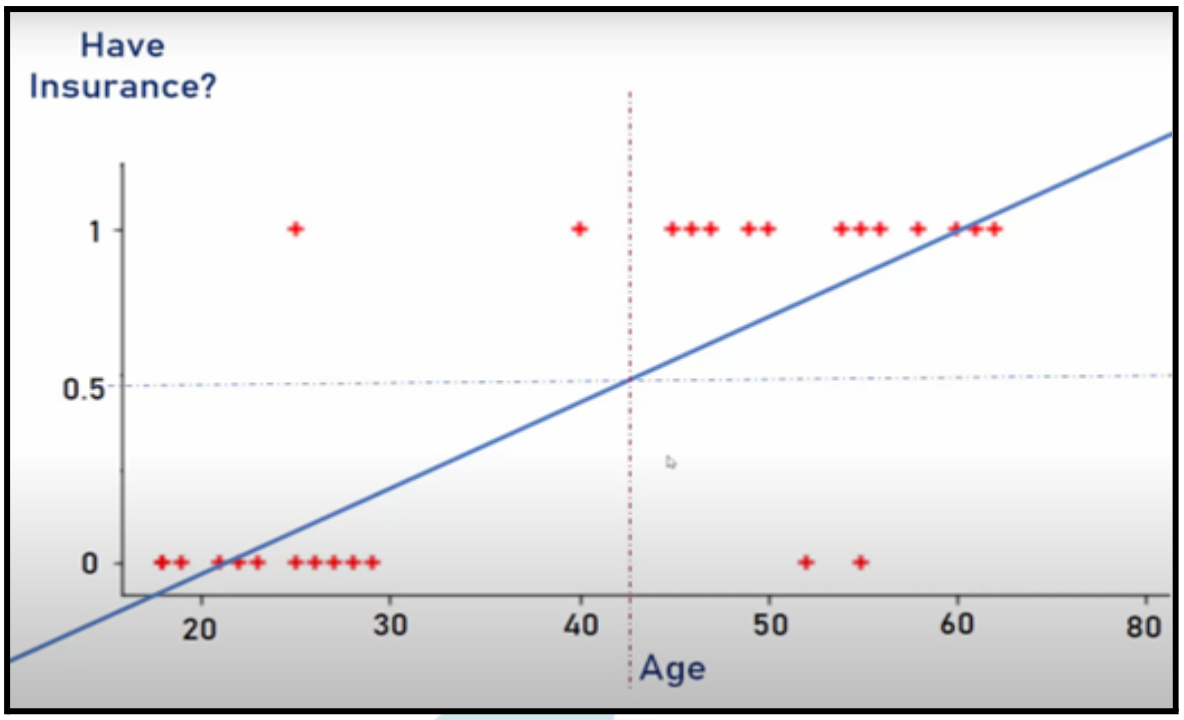
Limitations of using linear regression for classification:

1. Output Interpretation: Linear regression predicts a continuous outcome, which can take any real value. In classification, you need discrete class labels. Interpreting continuous predictions as class labels can be problematic.

2. Decision Boundary: The decision boundary in linear regression is a straight line (in two dimensions), which might not be sufficient to separate classes properly, especially if the data is not linearly separable.

3. Probability Estimates: Linear regression doesn't naturally provide probability estimates for class membership, which are often useful in classification tasks.

4. Model Performance: Linear regression can perform poorly on classification tasks because it's not designed to handle categorical outcomes. Metrics like accuracy, precision, recall, and F1 score might not be optimal.

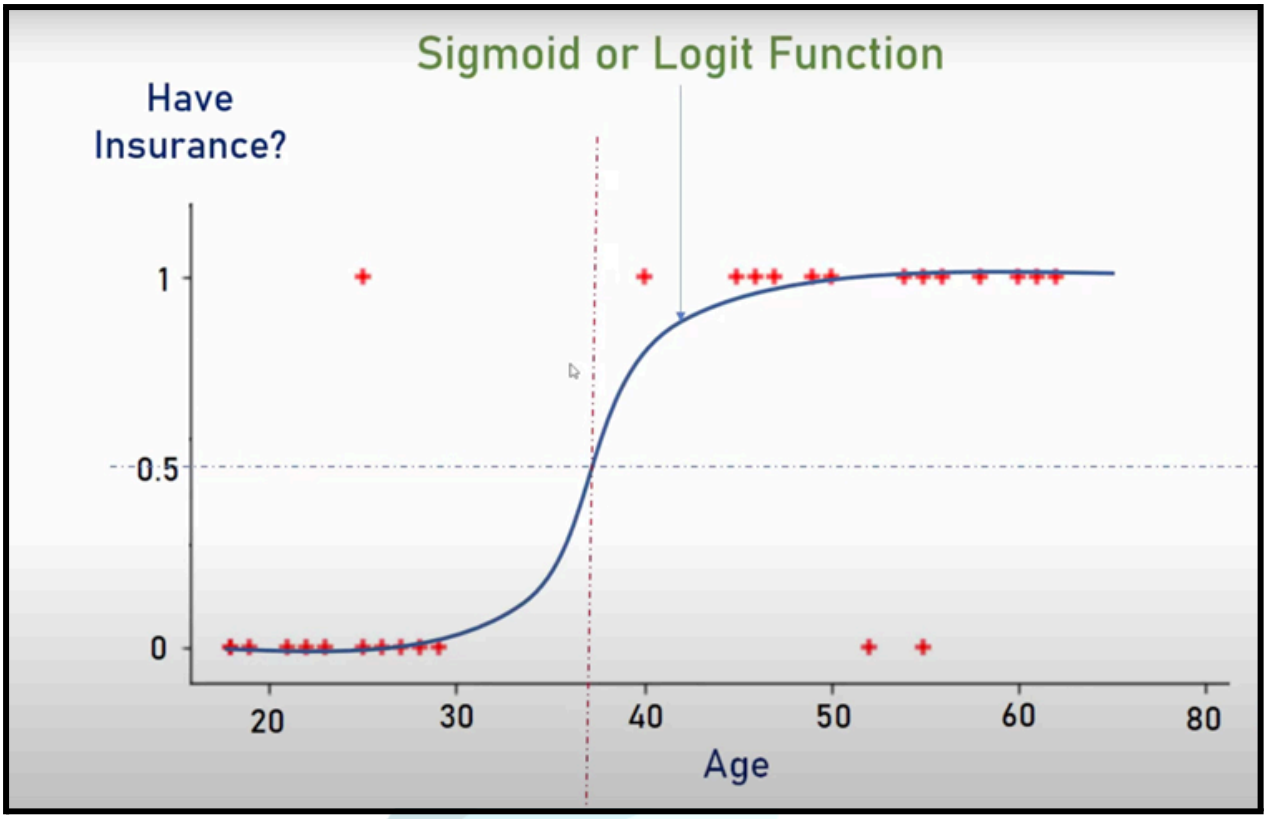


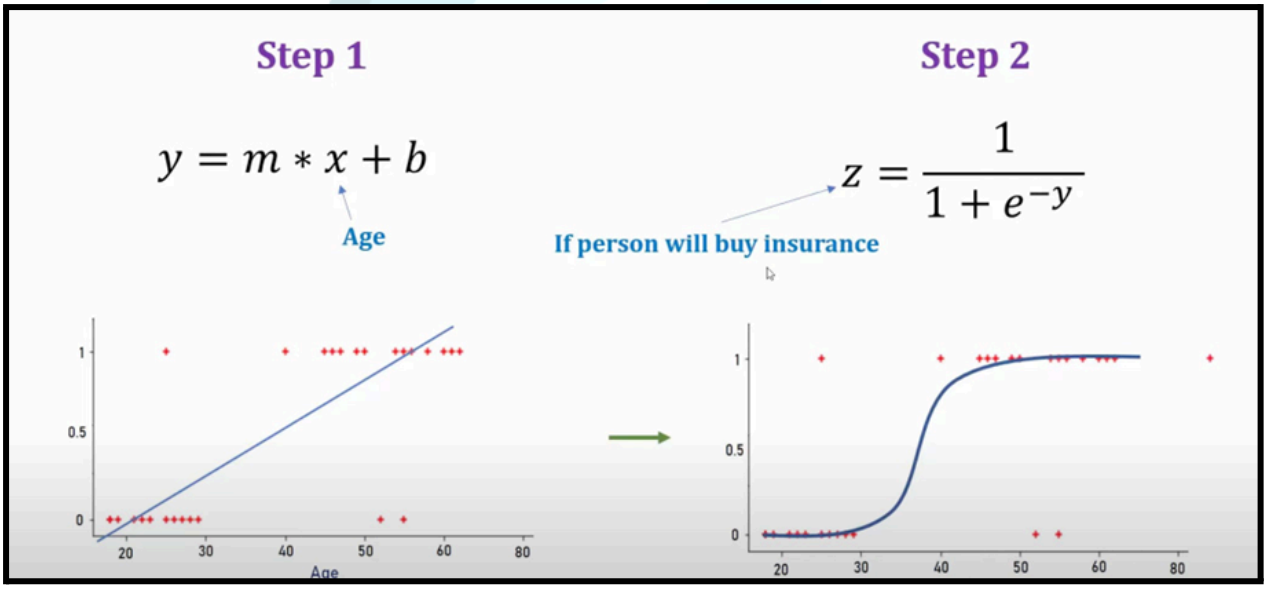
Wrong Classification

LOGISTIC REGRESSION

Logistic regression is a statistical method used for binary classification problems, where the outcome variable is categorical with two possible outcomes (e.g., yes/no, success/failure). Unlike linear regression, which predicts a continuous outcome, logistic regression predicts the probability that a given input point belongs to a specific class.

The logistic regression model is used to model the probability of a binary outcome. The relationship between the dependent variable (Y) and the independent variables (X1, X2, ..., Xp) is modelled using the logistic function (also known as the sigmoid function)



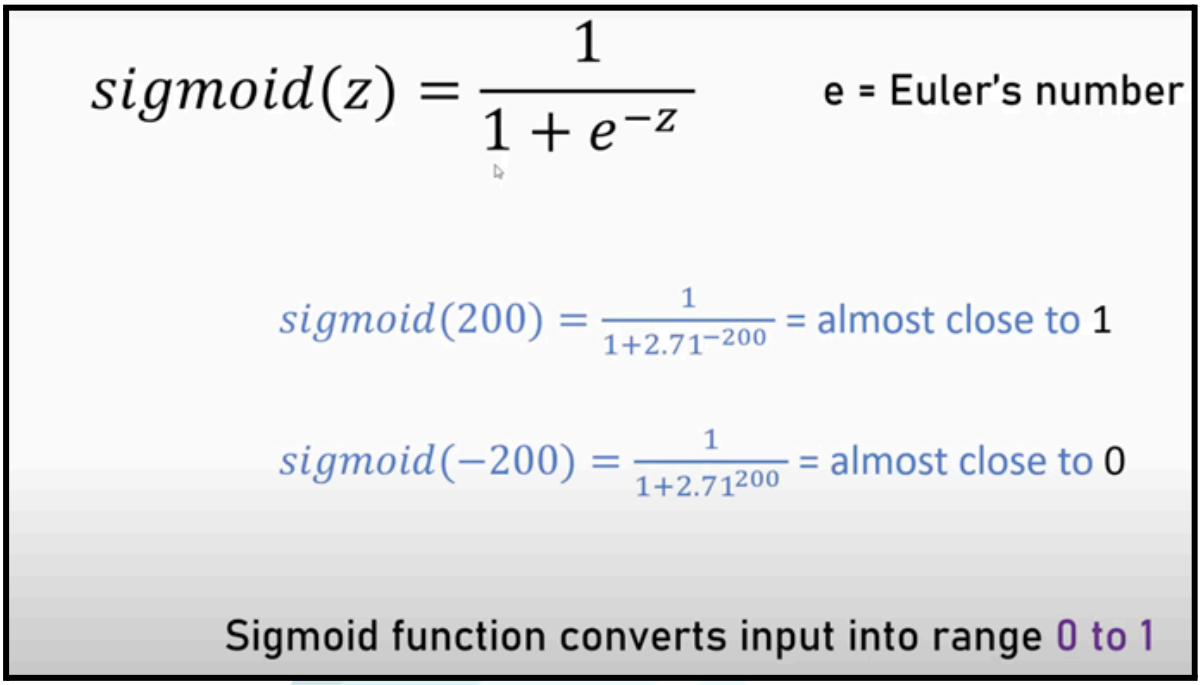


Depiction of how sigmoid differs from linear and in what way is it more suitable for classifying

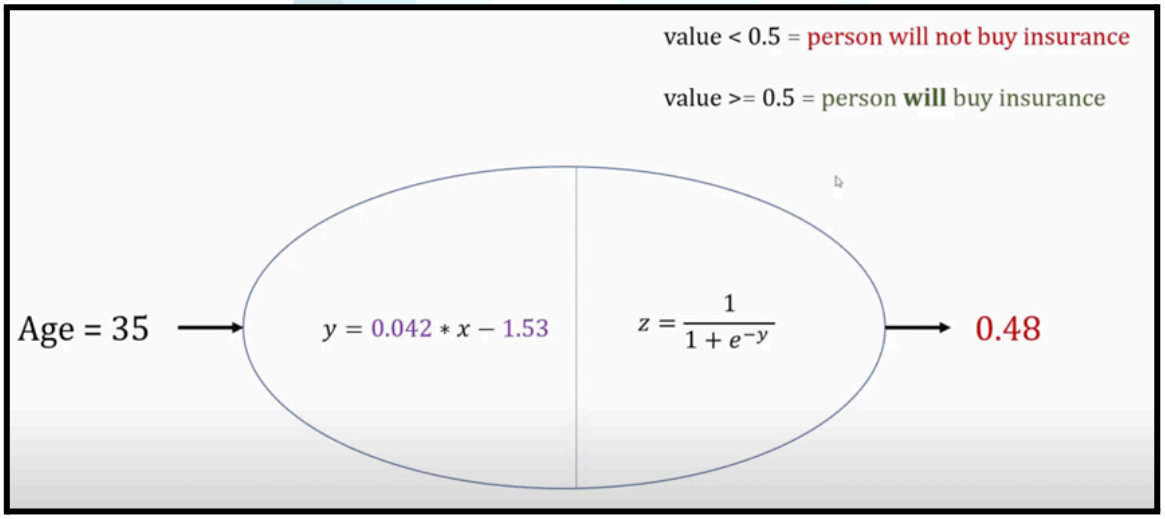
WHY SIGMOID FOR CLASSIFICATION?

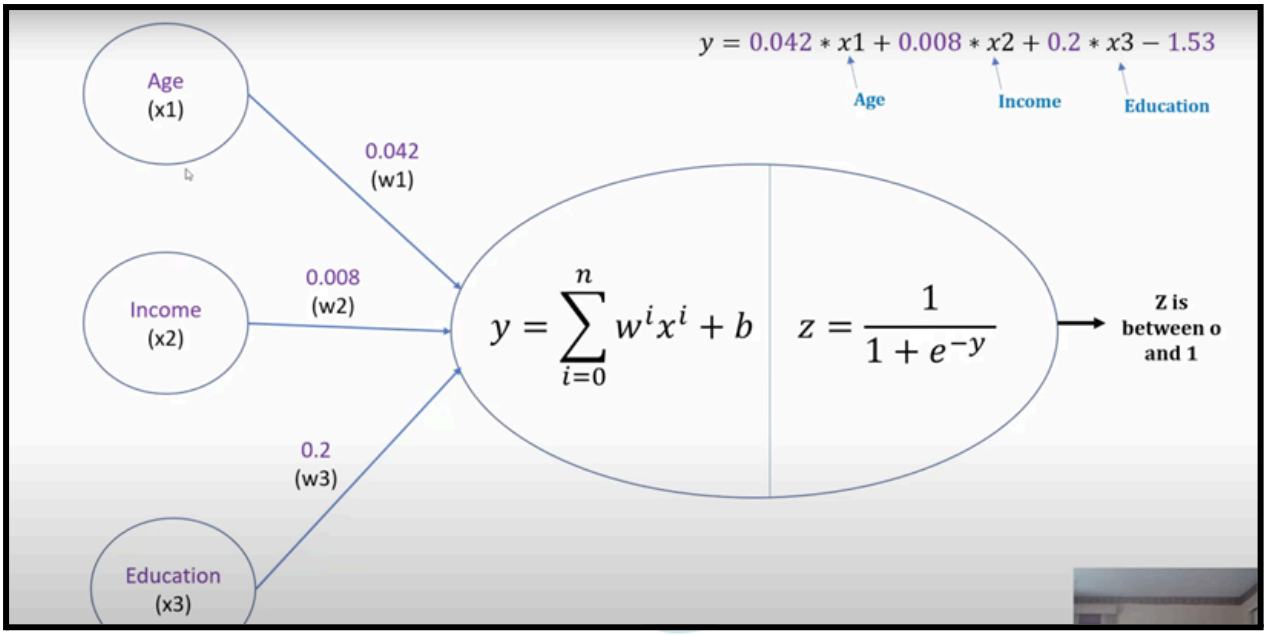
1. Output Range: The output values are in the range (0, 1), making it suitable for representing probabilities.

2. Differentiability: The sigmoid function is smooth and differentiable, which is useful for gradient-based optimisation algorithms like gradient descent.
3. Monotonicity: The function is monotonically increasing, meaning it consistently rises as the input increases.



Depiction of how a feature here is transformed to a suitable label ranging between 0 and 1 and thus, by setting a benchmark, we can easily classify the same.





Depiction of how a neuron works in a network, Here we have assigned different weights to each feature randomly just to show the working.

CODE SNIPPET FOR LOGISTIC REGRESSION

```
class LogisticRegression():
    def __init__(self):
        self.w = None
        self.b = None

    # return dot product of weight vector with data point
    def dot_pro(self,x):
        return (np.dot(self.w,x) + self.b)

    # return the probability of the point belonging to a class.
    def predict(self, x):
        return 1.0/(1.0 + np.exp(-self.dot_pro(x)))

    # Calculates gradient w.r.t w
    def gradient_w(self,x,y):
        pred = self.predict(x)
        return (pred - y)*x

    # Calculates gradient w.r.t b
    def gradient_b(self,x,y):
        pred = self.predict(x)
        return (pred - y)

    # Fit method
    def fit(self, x_train, y_train, epochs=100, learning_rate=0.01, refit=True):
        # initializing weights with random values.
        if refit:
            self.w = np.random.randn(x_train.shape[1])
            self.b = 0
        for i in range(epochs):
            grad_w = 0
            grad_b = 0
            for x, y in zip(x_train, y_train):
                grad_w += self.gradient_w(x, y)
                grad_b += self.gradient_b(x, y)

            self.w = self.w - learning_rate*grad_w
            self.b = self.b - learning_rate*grad_b
```

Assignment 2

Linear regression is a fundamental technique in statistical modelling and machine learning used to model the relationship between a dependent variable and one or more independent variables. In this assignment, we performed linear regression using two different methods: Ordinary Least Squares (OLS) and Gradient Descent. We then evaluated and compared the models using three metrics: R-squared (R^2) score, Mean Squared Error (MSE), and Mean Absolute Error (MAE).

```
table.head()
```

✓ 0.0s Open 'table' in Data Wrangler

	TV	Sales
0	230.1	22.1
1	44.5	10.4
2	17.2	9.3
3	151.5	18.5
4	180.8	12.9

The dataset looked like this

```python
class MyLRFromOLS:
    def __init__(self):
        self.m = None
        self.b = None

    def fit(self, X_train, y_train):
        # Convert data to numpy arrays for easier manipulation
        X_train = np.array(X_train).flatten()
        y_train = np.array(y_train)

        # Calculate values of m and b
        n = len(X_train)
        numerator = (n * np.sum(X_train * y_train)) - np.sum(X_train) * np.sum(y_train)
        denominator = (n * np.sum(X_train**2)) - (np.sum(X_train))**2

        self.m = numerator / denominator
        self.b = (np.sum(y_train) - self.m * np.sum(X_train)) / n

    def predict(self, X_test):
        # Convert X_test to numpy array if it's not already
        X_test = np.array(X_test).flatten()
        return self.m * X_test + self.b

    def get_coeff(self):
        return self.m, self.b
```

Writing code for OLS Method

```
X_train = X_train.values.reshape(-1, 1)
X_test = X_test.values.reshape(-1, 1)

from sklearn.preprocessing import StandardScaler
scaler= StandardScaler()
scaler.fit(X_train)
X_train_scaled=scaler.transform(X_train)
X_test_scaled= scaler.transform(X_test)
```

Scaling the data

```
#Get values of m and b
olslr.get_coeff()
```

```
(4.484453920009182, 14.167499999999999)
```

Getting the values of slope and intercept

```python
import numpy as np

class MyLRFromGD:

    def __init__(self, learning_rate, epochs):
        self.m = 0.5  # assign random value
        self.b = 100  # assign random value
        self.lr = learning_rate
        self.epochs = epochs

    def fit(self, X_train, y_train):
        # calculate b and m using GD
        X_train = np.array(X_train).reshape(-1, 1)
        y_train = np.array(y_train).reshape(-1, 1)
        n = len(X_train)
        for i in range(self.epochs):
            y_pred = self.m * X_train + self.b
            # Calculate slope wrt b and m
            loss_slope_b = (-2/n) * np.sum(y_train - y_pred)
            loss_slope_m = (-2/n) * np.sum(X_train * (y_train - y_pred))

            self.b = self.b - (self.lr * loss_slope_b)
            self.m = self.m - (self.lr * loss_slope_m)

    def predict(self, X_test):
        X_test = np.array(X_test).reshape(-1, 1)
        return self.m * X_test + self.b

    def get_coeff(self):
        # Write code to get values of m and b
        return self.m, self.b
```

Writing the code for gradient descent

```python
# Get values of m and b
m, b = gdlr.get_coeff()
print(f"Coefficients: m = {m}, b = {b}")
```

```
Coefficients: m = 4.484453913303477, b = 14.167500144453294
```

Getting values of slope and intercept

Using metrics for accuracy

R2 Score

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

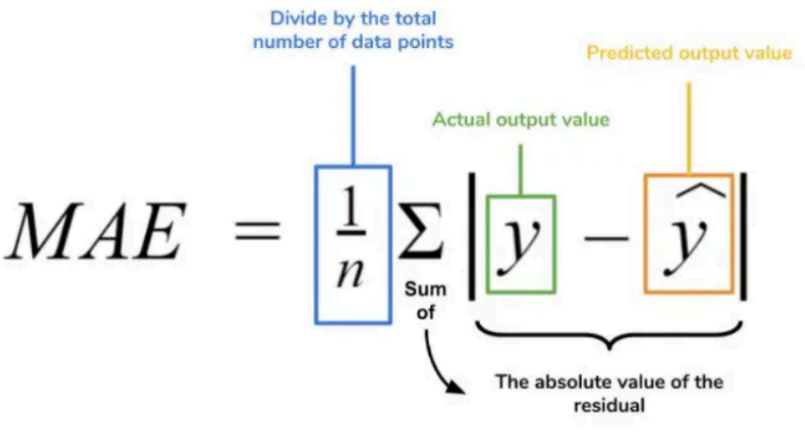
Measures the proportion of variance in the dependent variable explained by the independent variables

ROOT MEAN SQUARED ERROR

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

The average of the squared differences between the predicted and actual values

MEAN ABSOLUTE ERROR



The average of the absolute differences between the predicted and actual values

```python
class evaluate:

    def __init__(self, y_pred, y_test):
        self.y_pred = np.array(y_pred)
        self.y_test = np.array(y_test)

    def r2score(self):

    # Calculate the R^2 score
        res = np.sum((self.y_test - self.y_pred) ** 2)
        total = np.sum((self.y_test - np.mean(self.y_test)) ** 2)

        r2 = 1 - (res / total)
        return r2

    def mae(self):
        # Calculate the Mean Absolute Error
        mean_absolute_error = np.mean(np.abs(self.y_test - self.y_pred))
        return mean_absolute_error

    def rmse(self):
        # Calculate the Root Mean Squared Error
        rmse = np.sqrt(np.mean((self.y_test - self.y_pred) ** 2))
        return rmse
```

Writing the code

```
# First calculate all the metrics for olslr_pred
evaluation = evaluate(y_test,olslr_pred)
r2 = evaluation.r2score()
meanabserror=evaluation.mae()
meansquare=evaluation.rmse()
print("R-squared:", r2)
print("Mean absolute error:",meanabserror)
print("Root mean squared error:",meansquare)
```

```
R-squared: 0.11992385717310161
Mean absolute error: 3.2353682995006396
Root mean squared error: 4.21816233753514
```

```
res_gdlr = np.sum((np.array(y_test) - np.array(gdlr_pred)) ** 2)
res_ols = np.sum((np.array(y_test) - np.array(olslr_pred)) ** 2)

print("Residual sum of squares (gdlr):", res_gdlr)
print("Residual sum of squares (ols):", res_ols)
```

```
Residual sum of squares (gdlr): 64724.977873048716
Residual sum of squares (ols): 711.7157402319965
```

4. Exploratory Data Analysis

Data Types and Correlation EDA involves analysing combinations of categorical and numerical data to find correlations and understand how variables affect each other. Tools like histograms, box plots, heat maps, and scatter plots are commonly used. For example, a heatmap can show the correlation matrix of different variables, indicating how strongly they are related.

1. We can analyse combinations of categorical and numerical data find the correlation between them and study how they might affect each other

2. We can analyse a single variable(univariate) using histograms, box plots etc. and multiple variables(multivariate) using heatmaps, scatter plots etc.

3. We can use Pandas Profiling for Automating EDA, which gives concise reports with minimum lines of code.

5.Feature Engineering

Feature transformation includes handling missing values, encoding categorical variables, detecting and treating outliers, and applying feature scaling techniques such as standardization and normalization. For example, transforming a skewed variable using a logarithmic transformation can improve model performance.

Feature Transformation

This includes handling missing values, encoding categorical variables, detecting and treating outliers, and applying feature scaling techniques like standardisation and normalization.,

Feature Scaling

- **Standardisation: (Xi - Xmean)/σ**
- **Normalisation**
  - **Min - Max Scaling:**
  - **(Xi - Xmin)/(Xmax - Xmin)**
  - **Mean Normalisation**
  - **Robust Scaling**
  - **Max Absolute Scaling**

Feature Construction

Creating new features from existing ones can help capture important relationships and patterns in the data.

Feature Extraction

Dimensionality reduction techniques like Principal Component Analysis (PCA) can help extract the most informative features from high-dimensional data.

Data Imputation

Dropping Columns with Missing Values: Simplest option but might lose potentially useful information.

Imputation: Filling in missing values with a specific number, such as the mean value.

Advanced Imputation Techniques: Addressing systematic biases in imputed values or unique characteristics of rows with missing values.

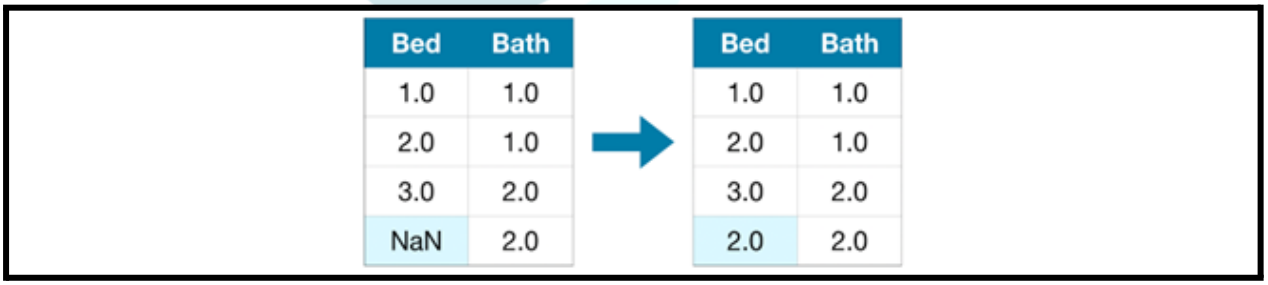
Methods:

1. Drop Columns with Missing Values:

The simplest option is to drop columns with missing values. Unless most values in the dropped columns are missing, the model loses access to a lot of (potentially useful!) information with this approach.

2. Imputation:

Imputation fills in the missing values with some number. For instance, we can fill in the mean value along each column.



3. An Extension To Imputation:

Imputation is the standard approach, and it usually works well. However, imputed values may be systematically above or below their actual values (which weren't collected in the dataset). Or rows with missing values may be unique in some other way.

Bed	Bath
1.0	1.0
2.0	1.0
3.0	2.0
NaN	2.0

Bed	Bath	Bed_was_missing
1.0	1.0	FALSE
2.0	1.0	FALSE
3.0	2.0	FALSE
2.0	2.0	TRUE

Function and Power Transformations.

Handling skewed distributions is essential for improving model accuracy. Techniques include:
1. Power Transformation: Yeo-Johnson, Box-Cox
2.Function Transformation: Logarithmic, Exponential, Square Root, Reciprocal
Many machine learning algorithms assume that the data features are normally distributed, this is why handling skewed distribution becomes an essential task in the data transformation process, as the skewed data might lead to biassed or inaccurate models.

Power Transformation

- Yeo Johnson
- BoxCox

The Yeo-Johnson transform is given by:

$$x_i^{(\lambda)} = \begin{cases} [(x_i+1)^\lambda - 1]/\lambda & \text{if } \lambda \neq 0, x_i \geq 0, \\ \ln(x_i+1) & \text{if } \lambda = 0, x_i \geq 0 \\ -[(-x_i+1)^{2-\lambda} - 1]/(2-\lambda) & \text{if } \lambda \neq 2, x_i < 0, \\ -\ln(-x_i+1) & \text{if } \lambda = 2, x_i < 0 \end{cases}$$

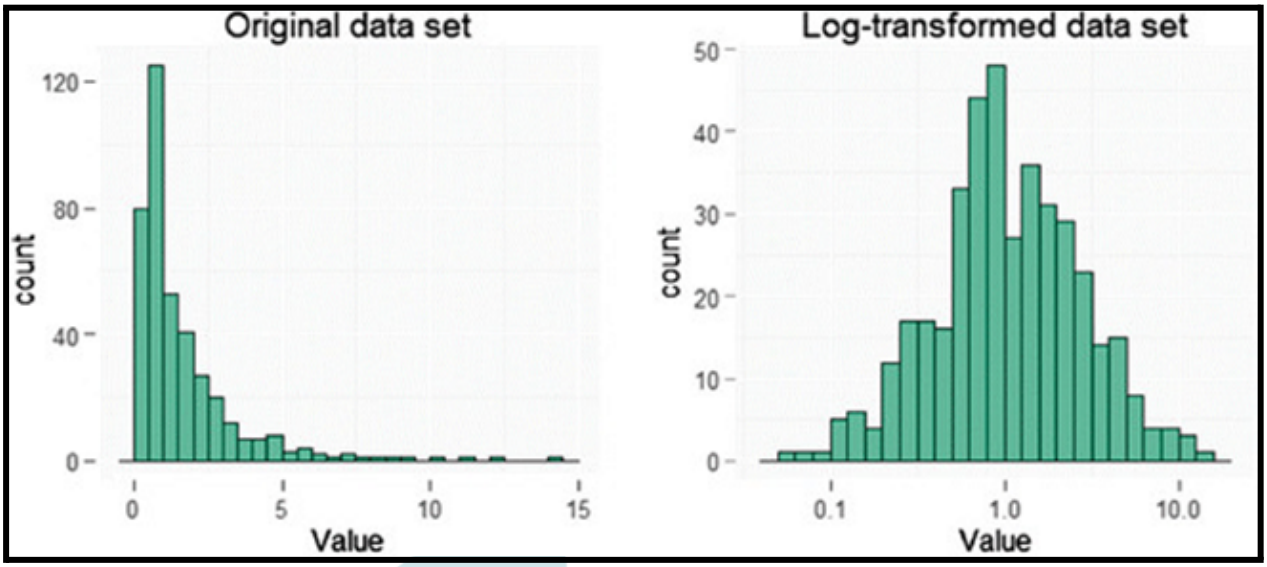
while the Box-Cox transform is given by:

$$x_i^{(\lambda)} = \begin{cases} \dfrac{x_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln(x_i) & \text{if } \lambda = 0, \end{cases}$$

Function Transformation

- Logarithmic
- Exponential

- Square Root
- Reciprocal



Logarithmic Transformation

6. KNN Imputor and Algorithm , MICE algorithm

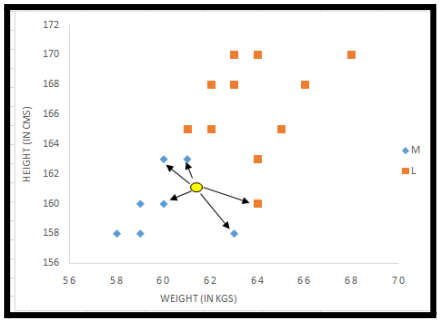
KNN imputation is a more sophisticated approach that considers the proximity of data points. It imputes missing values by averaging or voting based on the values of the k-nearest neighbors in the feature space. This method is effective when there is a strong local relationship between data points.

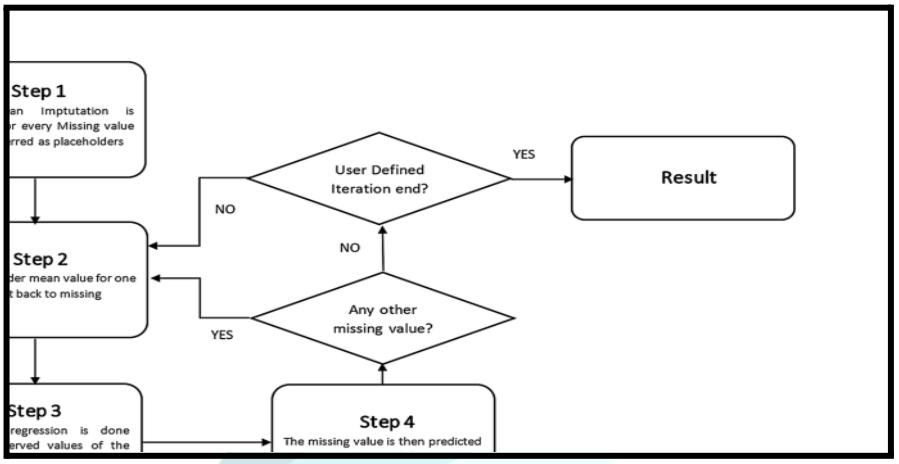
A	B	C	D
45		24	37
62	21	34	
	64	33	57
34	34	56	
	23		48
64			69
56		22	24
86	45	26	
		78	97
21	39	91	

A	B	C	D
45	35	24	37
62	21	34	39
51	64	33	57
34	34	56	41
34	23	36	48
64	65	66	69
56	34	22	24
86	45	26	52
88	87	78	97
21	39	91	50

MICE Algorithm:

The MICE algorithm is a powerful statistical tool that reconstructs missing data by iteratively estimating values based on relationships between variables



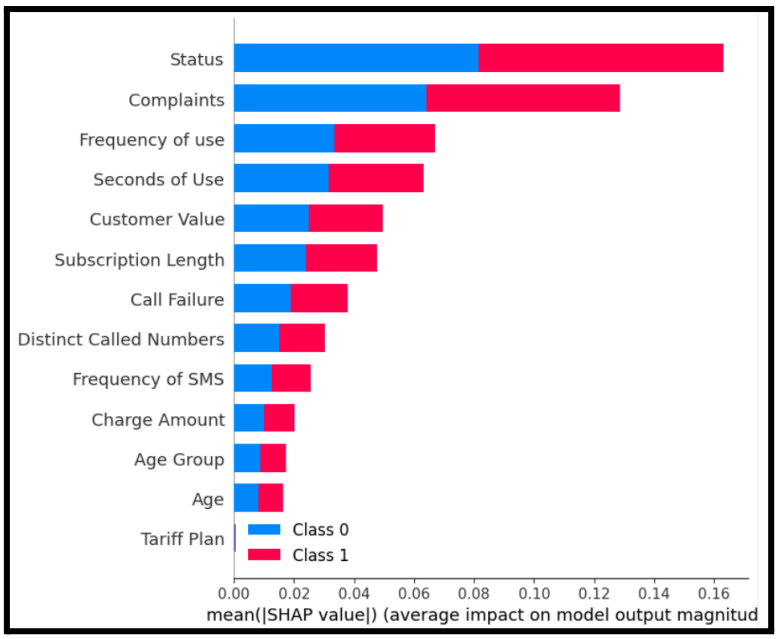


8. Introduction to SHAP and SHAPLEY values

What are SHAP Value?

SHAP (SHapley Additive exPlanations) values are a way to explain the output of any machine learning model. It uses a game theoretic approach that measures each player's contribution to the final outcome. In machine learning, each feature is assigned an importance value representing its contribution to the model's output.SHAP values show how each feature affects each final prediction, the significance of each feature compared to others, and the model's reliance on the interaction between features.

The summary plot shows the feature importance of each feature in the model. The results show that “Status,” “Complaints,” and “Frequency of use” play major roles in determining the results.



Application of SHAP Values:

- Model debugging
- Feature importance
- Anchoring explanations
- Model summaries
- Detecting biases

What is Shapely Value?

Shapley values come from cooperative game theory and provide a way to fairly distribute the payoff (in machine learning, this is typically the prediction) among the players (features) based on their contribution. They ensure that each feature's contribution to the prediction is fairly assessed, taking into account all possible combinations of features.

Using Shapley values in machine learning provides detailed insights into feature contributions for individual predictions and overall feature importance. Outputs include force plots for individual predictions, summary plots for feature distributions, and bar plots for global feature importance.

Why are Shapley Values Important?

Shapley values provide a robust framework for interpreting model predictions by:

- Quantifying Feature Contributions
- Ensuring Fairness
- Model-Agnostic Interpretability

How Shapley Values Work:

1. **Baseline Prediction**: Start with a baseline prediction, which is usually the average prediction over the training dataset.
2. **Marginal Contributions**: For each feature, compute the marginal contribution by adding the feature to all possible subsets of the remaining features and measuring the change in the prediction.
3. **Average Over Permutations**: The Shapley value for a feature is the average of its marginal contributions across all possible permutations of the feature set.

CONSULTING

8. Introduction to Consulting

Consulting is a professional service offered by individuals or firms, and it involves advising an organization on how to achieve better results.

Types of Consulting

- Strategy Consulting
- Management Consulting
- Operational Consulting
- Financial Consulting
- Legal Consulting
- IT Consulting

First Four were Explained Briefly

Strategy Consulting

Strategy consulting is the process of advising companies on important business decisions by providing strategic consulting advice. It generally includes researching and analyzing issues, identifying challenges and opportunities, and making recommendations. (Examples)

Management Consulting

Management consulting is the practice of providing consulting services to organizations to improve their performance or in any how to assist in achieving organizational objectives. (Examples)

Operational Consulting

From a functional perspective, operations consulting aims at boosting the processes, ways of working and underlying systems across a broad range of areas, of which Sales &

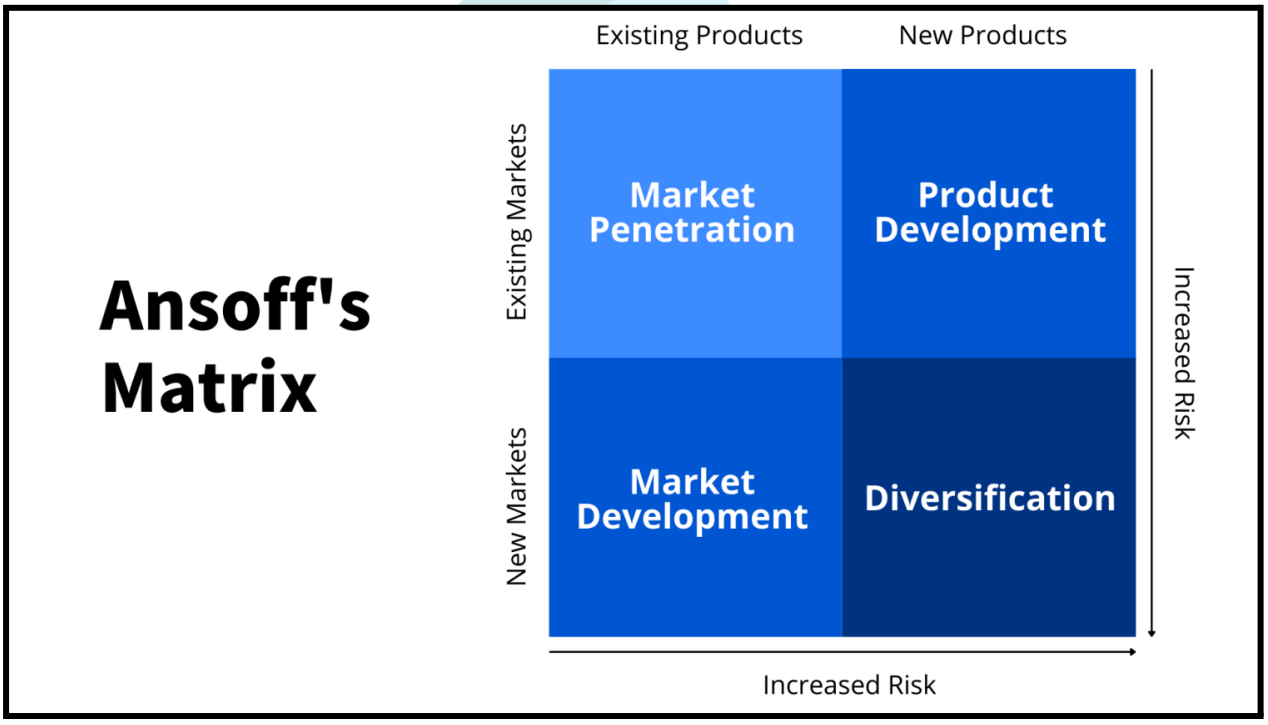
Marketing, Supply Chain, Sourcing & Procurement, Finance and Research & Development are the largest areas in terms of size. (Examples)

Financial Consulting

Financial consultants or advisors offer clients a big picture analysis of their finances. They break down different aspects of a client's financial life, including assets, expenses and income, and help them create a financial plan to reach different types of goals. (Examples)

9. Introduction to Marketing, customer segmentation, market segmentation and strategy

- **Ansoff Matrix :**



The Ansoff Matrix is a marketing tool used to create business growth strategies. It outlines four different strategies businesses can use to achieve growth, each with varying levels of risk.

- Market penetration involves selling existing products to existing markets. This strategy is considered to be the least risky.

- Product development involves creating a new product for an existing market. More risk is involved in this strategy than market penetration because it requires creating a new product.

- Market development involves selling existing products to new markets. This strategy is also riskier than market penetration because it involves entering a new market.

- Diversification involves creating a new product for a new market. This is the riskiest of the four strategies because it requires innovation in both product and market.

- **Relationship b/w Analysis, Strategy, and Programs :**

In market segmentation analysis, strategy and programs work together in a sequential manner to effectively target and reach specific customer groups. They relate as follows :

1. Analysis: This is the foundation. You use market research and customer data to understand the overall market. This analysis involves:

    - Identifying customer needs and wants: What are potential customers' different motivations, pain points, and preferences?

    - Segmenting the market: By analyzing demographics, behavior, and psychographics, you group customers with similar characteristics into distinct segments.

    - Evaluating segment attractiveness: Analyze each segment's size, growth potential, and profitability to determine which ones are most attractive for your business.

2. Strategy: Based on the analysis, you develop a market segmentation strategy that defines how you will target and reach each customer segment. This strategy should consider:

    - Value proposition: How will you tailor your product or service to address each segment's specific needs and wants?

    - Marketing mix: How will you adjust your pricing, promotion, distribution channels, and product features to resonate with each segment?

    - Competitive advantage: How will you position yourself to stand out from competitors within each segment?

3. Programs: Finally, you translate your strategy into actionable programs that reach your target segments. These programs may include:

- Targeted marketing campaigns: Develop specific marketing messages and channels to reach each segment effectively.
- Product development: Adapt or create products/services that cater to the unique needs of each segment.
- Pricing strategies: Implement different pricing models for each segment based on their price sensitivity and willingness to pay.

10. Marketing analysis, segmentation, and strategy development for Udaan and Marico(Consult assignment-1)

- We formed a group of four and were assigned to introduce a Marico product into the Japanese market, using the knowledge of resources provided by mentors.
- Each team had divided the task of formulating the below for Marico:

Target Market Segmentation, Product Strategy, Pricing Strategy, Promotional Strategy, Risk Analysis and Distribution Channel.

The flow of the project is as follows :

I. Executive Summary

- Briefly introduce Marico and the chosen product for Japanese market entry.
- Highlight the potential of the Japanese market and the product's value proposition.
- Summarize the key elements of the proposed market entry strategy.

II. Market Analysis

- Japanese Consumer Landscape:
- Demographics and psychographics of target audience.
- Hair care and personal care product trends in Japan.
- Consumer preferences for natural and healthy ingredients (relevant to Marico's strengths).
- Competitive Landscape:
    - Identify major players in the Japanese hair care/personal care market.
    - Analyze competitor strengths, weaknesses, opportunities, and threats (SWOT).
    - Highlight differentiation opportunities for Marico's product.

III. Market Entry Strategy

- Product Adaptation:
    - Assess the need for product modifications to suit Japanese preferences (e.g., fragrance, packaging).
    - Consider compliance with Japanese labelling and regulatory requirements.
- Pricing Strategy:
    - Analyze competitor pricing and consumer price sensitivity.
    - Determine an optimal pricing strategy that balances profitability and market competitiveness.
- Distribution Strategy:
    - Evaluate potential distribution channels (e.g., online marketplaces, traditional retail stores).
    - Consider partnering with a local distributor or establishing a subsidiary.
- Marketing and Promotion:
    - Develop a targeted marketing campaign emphasizing the product's unique benefits.
    - Utilize appropriate marketing channels (e.g., social media, influencer marketing) to reach the target audience.

IV. Implementation and Risk Management

- Outline a timeline for market entry, including product launch and marketing activities.
- Identify potential risks of entering the Japanese market (e.g., cultural differences, regulatory hurdles).
- Develop mitigation strategies to address these risks.

V. Conclusion

- Recap the key aspects of the market entry strategy.
- Reiterate the potential for success in the Japanese market.
- Briefly mention future considerations for expanding Marico's presence in Japan.

11. Introduction to clustering

Learnt about Hierarchical Clustering:

- Agglomerative
- Divisive

12. Assignment 3(Bank customer churn analysis)

Achieved an accuracy of 97%

Introduction to various frameworks

1. SWOT Analysis

A SWOT analysis is a strategic planning tool that helps an organization identify its strengths, weaknesses, opportunities, and threats. This framework aids in understanding the internal and external factors that can impact the organization's objectives, enabling effective strategy formulation.

During this project, we gained an in-depth understanding of the SWOT analysis framework, which is pivotal in strategic planning and decision-making. SWOT stands for Strengths, Weaknesses, Opportunities, and Threats, and it provides a comprehensive approach to evaluating an organisation's internal and external environments.

1. **Strengths** refer to the internal attributes and resources that support a company's success. These can include brand reputation, financial performance, market share, and operational efficiencies. Identifying and leveraging strengths is crucial for maintaining a competitive edge.
2. **Weaknesses** are internal factors that may hinder a company's performance. Recognizing these weaknesses, such as operational inefficiencies or gaps in product offerings, helps in developing strategies to mitigate them and improve overall effectiveness.
3. **Opportunities** represent external chances for growth or improvement. These could arise from market trends, technological advancements, or shifts in consumer behaviour. Seizing opportunities can lead to significant business expansion and innovation.
4. **Threats** are external factors that could cause trouble for the business. These might include competitive pressures, regulatory changes, or economic fluctuations. Understanding threats is essential for developing contingency plans and risk management strategies.

2.PESTLE Analysis

The PESTLE analysis examines the macro-environmental factors affecting an organization. It stands for Political, Economic, Sociological, Technological, Legal, and Environmental factors. This analysis helps organizations anticipate market trends, identify potential threats, and align their strategies with the external environment.

The PESTLE analysis examines the macro-environmental factors affecting an organization. PESTLE stands for Political, Economic, Sociological, Technological, Legal, and Environmental factors.

1) **Political factors** involve government policies, trade restrictions, and political stability. These elements can significantly influence business operations and strategic planning, especially in terms of compliance and market entry strategies.

    1) Government Policies: Regulatory policies, tax policies, trade restrictions, and tariffs.

    2) Political Stability: The overall political climate and stability of the region or country.

    3) Corruption Levels: The prevalence of corruption and its impact on business operations.

2) **Economic factors** include economic growth, exchange rates, and inflation rates. Understanding these factors helps in forecasting market trends, setting pricing strategies, and making investment decisions.

    1) Economic Growth: GDP growth rates, inflation rates, and overall economic health.

    2)Exchange Rates: Fluctuations in currency exchange rates affecting international trade.

    3)Interest Rates: Central bank policies and their influence on borrowing and investment.

3) **Sociological factors** encompass demographic changes, social trends, and cultural shifts. These insights are crucial for aligning products and marketing strategies with consumer preferences and behaviours.

4) **Technological factors** consider advancements and innovations that could impact the industry. Staying ahead in technology adoption can provide a competitive advantage and improve operational efficiencies.

5) **Legal factors** involve regulations and laws that a company must comply with. This includes labour laws, consumer protection laws, and environmental regulations, which can affect operational practices and product development.

6) **Environmental factors** address ecological and environmental aspects such as climate change, sustainability, and waste management. These are increasingly important for corporate social responsibility and long-term viability, influencing both operational practices and brand reputation.

3. The ABELL model

The Abell Model, also known as the Abell Framework, is a strategic tool used to define the business scope and identify growth opportunities. It focuses on three dimensions:

1. Customer Groups (Who):
    - Identifying and categorizing different customer groups based on demographics, geography, and behavior.
    - Understanding the needs and preferences of these groups to tailor products and services accordingly.
2. Customer Needs (What):
    - Determining the specific needs and wants of the identified customer groups.
    - Analyzing how well current products or services meet these needs and identifying gaps for new offerings.
3. Technologies (How):
    - Assessing the technologies used to satisfy customer needs.
    - Exploring new technologies that can enhance product offerings or create new market opportunities.

The Abell Model helps businesses focus on their core competencies and align their strategies with market demands and technological advancements.

4. TOWS Matrix

The TOWS Matrix is an extension of the SWOT analysis. It provides a framework to develop strategic options by matching internal strengths and weaknesses with external opportunities and threats. The matrix offers four strategic options: Strengths-Opportunities (SO), Weaknesses-Opportunities (WO), Strengths-Threats (ST), and Weaknesses-Threats (WT).

The TOWS Matrix is an advanced version of the SWOT Analysis, providing a framework to develop strategic options by matching internal strengths and weaknesses with external opportunities and threats. It consists of four strategic options:

1. Strengths-Opportunities (SO) Strategies:
    - Leverage internal strengths to capitalize on external opportunities.
    - Example: A company with strong R&D capabilities (strength) can exploit a growing market for innovative products (opportunity).
2. Weaknesses-Opportunities (WO) Strategies:
    - Address internal weaknesses to take advantage of external opportunities.
    - Example: A company with a weak distribution network (weakness) can form alliances with established distributors to enter a new market (opportunity).
3. Strengths-Threats (ST) Strategies:
    - Use internal strengths to mitigate external threats.
    - Example: A company with a strong brand reputation (strength) can counteract increased competition (threat) by emphasizing brand loyalty.
4. Weaknesses-Threats (WT) Strategies:
    - Minimize internal weaknesses to avoid external threats.
    - Example: A company with outdated technology (weakness) can invest in upgrading its technology to withstand competitive pressures (threat).

The TOWS Matrix helps organizations develop actionable strategies by providing a structured approach to analyze their internal and external environments.

Consult Task - SWOT analysis on HUL

Introduction

The report aims to provide a comprehensive SWOT analysis for Hindustan Unilever Limited (HUL), the largest FMCG player in India, with operations extending to Western countries. References to precise documents, articles, and reports support the analysis. The process map provided has been utilised to granularise various aspects of an FMCG company, leading to detailed insights into strengths, weaknesses, opportunities, and threats. The analysis also includes comparisons with competitors to provide a holistic view.

Step 1: Framing Questions for SWOT Analysis

Importance of Preparing Questions and Process Mapping

Advantages of Preparing Questions

Preparing a set of questions before conducting a SWOT analysis offers several advantages:

1. Structured Approach: It provides a systematic framework to ensure that all relevant aspects of the business are examined thoroughly.
2. Comprehensive Coverage: Questions help in covering all dimensions of strengths, weaknesses, opportunities, and threats, ensuring no critical area is overlooked.
3. Focused Research: Questions guide the research process, making it more focused and efficient by identifying specific areas that need in-depth investigation.
4. Consistency: It ensures consistency in the analysis by maintaining a clear and uniform approach throughout the research process.
5. Enhanced Clarity: Clear questions help in obtaining precise and relevant information, leading to more accurate and actionable insights.

Importance of Process Map

A process map is crucial for several reasons:

1. Visual Representation: It provides a visual representation of the entire supply chain and business processes, making it easier to understand and analyse each component.
2. Identifying Bottlenecks: Process mapping helps in identifying bottlenecks and inefficiencies in the supply chain, which can then be addressed to improve overall performance.
3. Standardisation: It promotes the standardisation of processes, ensuring consistency and quality in operations.
4. Optimization: By mapping out the processes, companies can identify areas for optimization, leading to cost savings and improved efficiency.
5. Alignment: It ensures alignment of various departments and functions within the organisation, facilitating better coordination and communication.

Some questions we asked

STRENGTHS

1. How well-defined inventory systems and distribution centres do you possess?
2. How great is your brand as an employer, what kind of talent do you attract?

WEAKNESS

1. How are HUL's metrical ratios featured in the market compared to its competitors?
2. How does HUL plan to comply when it's so sensitive to changing laws and regulations of FMCG products?

OPPORTUNITIES

1. What are some of the underlapped areas of growth present currently for HUL?
2. What are new, presently rising, fields of FMCG products that HUL can capitalise on?

THREATS

1. Will any change in regulations of FMCG goods cause serious contingencies in HUL?
2. Are there strong Indian brands that pose a threat to your brand value, does your being the subsidiary of a brand with a history of Indian colonisation discredit your value.

Step 2: Research on HUL

To answer these questions, we will gather information from various sources such as industry reports, company financial statements, news articles, and market analysis reports.

Step 3: Answering the Questions

This led us to a comprehensive SWOT Analysis, we present some points from the same

STRENGTHS

1. HUL products have good visibility owing to strong advertising via TV ads, print ads, online ads, social media engagement, signing popular celebrities and digital marketing.
2. The manufacturing plants are equipped with state-of-the-art technology to ensure high-quality production standards.HUL continuously invests in upgrading its technology to enhance productivity and maintain quality.
3. HUL has a diverse product portfolio in personal care, home care, and foods, similar to ITC and Godrej Consumer, but with a broader range than most competitors

WEAKNESS

1. HUL also has a higher debt-to-equity ratio than rival ITC, HUL's debt-to-equity ratio fluctuated, while ITC consistently maintained a low ratio, signifying a more conservative approach to debt management
2. Decreasing Market Share: HUL has experienced a decrease in market share in some product categories. Competitors focusing on specific products, such as Ghadi and Nirma detergents, have gained market share from HUL, impacting its position in the market. Its profit before tax ratio is also lower than ITC (22.11 vs 37.41)

OPPORTUNITIES

1. HUL can further expand its reach by penetrating rural markets. Initiatives like Project Shakti AMMA have already enabled HUL to tap into rural areas, but there is still untapped potential for growth in these markets. By understanding the unique needs and preferences of rural consumers, HUL can tailor its offerings to cater to this segment effectively.

2. India has a significant number of unorganised businesses in the consumer goods sector. HUL can seize the opportunity to transition these businesses to organised ones, thereby expanding the consumer goods market and strengthening its market presence.

THREATS

1. HUL operates in a highly regulated environment, facing stringent laws related to consumer safety, environmental impact, and labour practices. Compliance with these regulations can be costly and complex, and any lapses can lead to legal issues and reputational damage
2. Local brands like Patanjali are giving special impetus to their make-in-India products and highlighting the fact that HUL is the subsidiary of a foreign company with a history of colonisation, this has a significant negative impact on its brand value

13.Conclusion

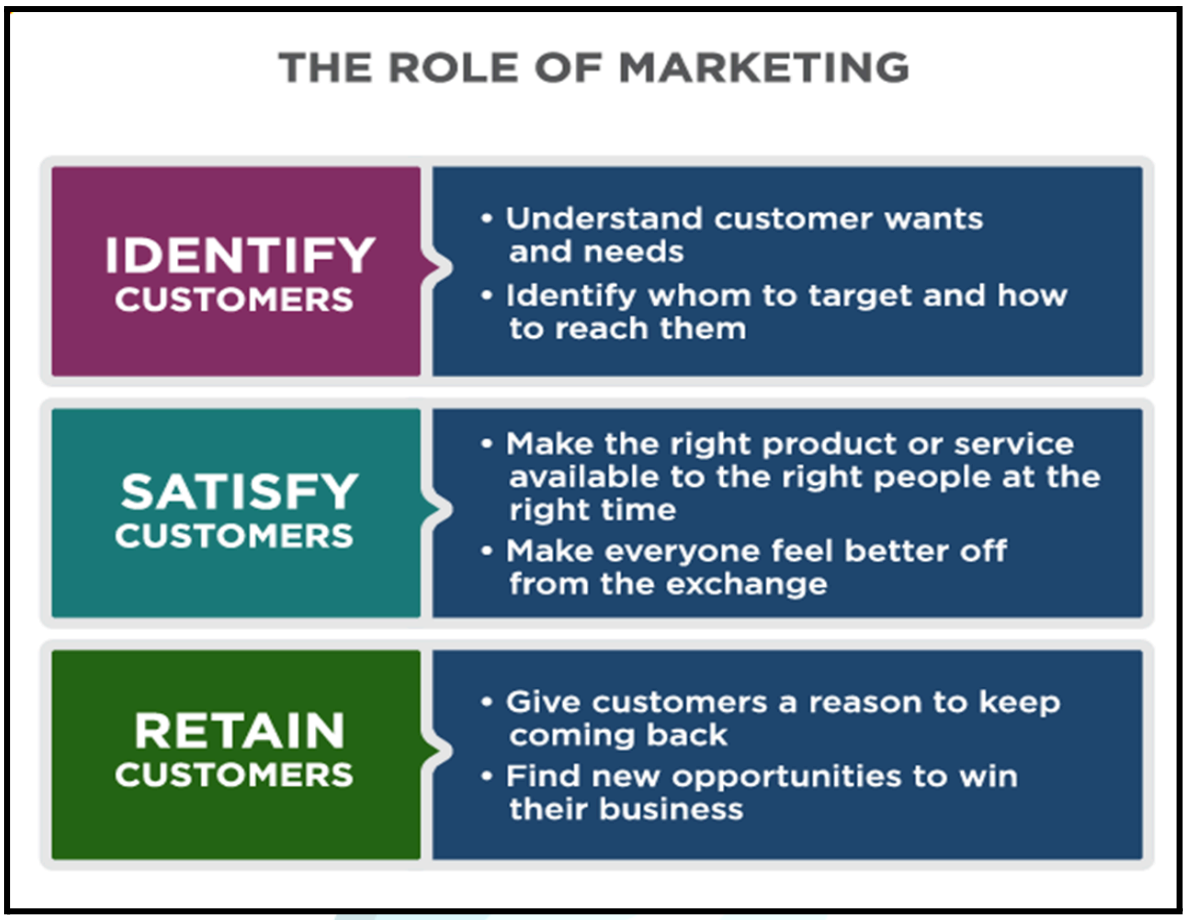
The SWOT analysis of Hindustan Unilever Limited (HUL) reveals a comprehensive view of the company's strategic position in the FMCG sector in India. HUL stands out as a market leader due to its strong brand portfolio, extensive distribution network, continuous innovation, and robust financial performance. The company's commitment to sustainability and corporate social responsibility further strengthens its market position and brand reputation.

However, HUL faces several internal and external challenges. Operational inefficiencies in the supply chain and a high dependency on the Indian market for revenue are notable weaknesses. Additionally, managing a vast brand portfolio and maintaining cost competitiveness require ongoing strategic attention.

HUL is well-positioned to capitalise on significant growth opportunities in the rural FMCG market and the burgeoning e-commerce sector. The increasing consumer demand for sustainable and eco-friendly products presents another avenue for expansion. Moreover, emerging markets in Asia and Africa offer potential for international growth, leveraging HUL's extensive distribution capabilities and parent company Unilever's global network.

Conversely, HUL must navigate intense competition from major players like ITC, Nestlé, and Procter & Gamble. Regulatory changes and economic fluctuations pose significant threats that could impact HUL's operations and profitability. Furthermore, the rapid pace of technological advancements and shifts in consumer preferences necessitate continuous adaptation and innovation.

In conclusion, HUL's strengths provide a solid foundation for sustained growth, while its proactive approach to addressing weaknesses and leveraging opportunities can help mitigate potential threats. The company's ability to innovate, adapt to changing market dynamics, and expand its presence in emerging markets will be crucial in maintaining its leadership position in the highly competitive FMCG sector in India.


THE ROLE OF MARKETING
IDENTIFY CUSTOMERS
• Understand customer wants and needs
• Identify whom to target and how to reach them
SATISFY CUSTOMERS
• Make the right product or service available to the right people at the right time
• Make everyone feel better off from the exchange
RETAIN CUSTOMERS
• Give customers a reason to keep coming back
• Find new opportunities to win their business

Prepared By

Manit Kapoor
Ayansh Shankar Yadav
Aditya Singh
Raihaan Kakkat
Aakriti Prajapati
Chirag Goyal
Arnav Gupta