Analytical Methods for Unsupervised Learning

Modern businesses collect massive amounts of customer data. However, data alone does not automatically create value. Organizations must analyze patterns in the data to understand different groups of customers. One powerful method used by companies worldwide is Cluster Analysis.

In the Philippine business environment—where markets are diverse and consumer behavior varies across income groups, regions, and digital adoption levels—cluster analysis becomes a powerful marketing analytics tool.

For example, a company like Jollibee Foods Corporation may use customer purchase data to determine which groups prefer family meals, value meals, or premium menu offerings.

Through this module, students will learn how cluster analysis works, how clustering algorithms operate, and how these techniques help businesses make data-driven marketing decisions.

Cluster Analytics

Cluster Analytics (also called *Cluster Analysis*) is a statistical technique used in **Marketing Analytics** to group customers into distinct segments based on similarities in their data — such as purchasing behavior, demographics, preferences, or engagement patterns.

Cluster analysis helps businesses identify groups of customers who behave similarly, allowing firms to design more effective marketing strategies, promotions, pricing, and product offerings.

In marketing analytics, cluster analysis is often used for:

- Customer segmentation
- Product grouping
- Market segmentation
- Behavioral analysis
- Personalization strategies

Simply put:

Cluster Analysis = Grouping similar customers together based on their behaviors or characteristics.

Example (Philippine Business Context)

Consider **Shopee**, one of the largest e-commerce platforms in the Philippines.

Shopee collects data such as:

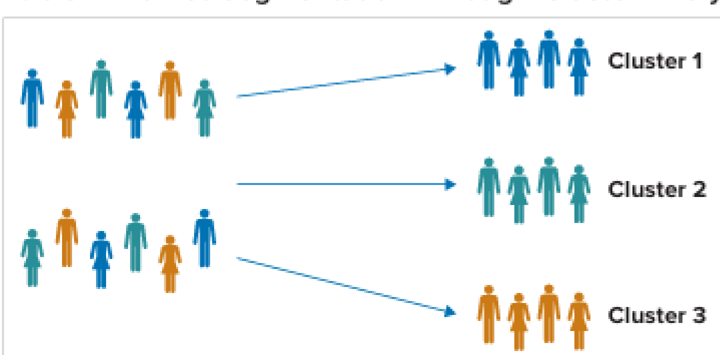
- Purchase frequency
- Amount spent
- Product categories purchased
- Location
- Payment method (GCash, COD, credit card)

Using cluster analysis, Shopee might identify these customer groups:

Cluster	Customer Type	Characteristics
Cluster 1	Bargain Hunters	Buy during flash sales and vouchers
Cluster 2	Loyal Buyers	Regular purchases every month
Cluster 3	Luxury Shoppers	Buy branded electronics and gadgets
Cluster 4	Occasional Buyers	Purchase only during special events

Cluster analysis is an analytical method of segmenting a market based on identifying shared characteristics of groups of individuals. Exhibit 8-1 shows three clusters with different characteristics. On the left are groups with a mix of different types of individuals, and on the right side of the exhibit are the results of cluster analysis in which Cluster 1 is all dark blue, Cluster 2 is light blue, and Cluster 3 is orange. The different colors represent the common characteristics of similar individuals within the clusters and at the same time the differences between the three clusters. Cluster analysis uses algorithms to explore different types of relationships, and then develops smaller groups from larger populations based upon similar characteristics. Marketers can use these insights to improve marketing strategies and better allocate resources when creating messages that resonate with a particular group, gauging new product development, or even selecting test markets that might be most receptive. SPSS Modeler, SAS Enterprise Miner, R Programming, and Python are reliable market segmentation tools to conduct a cluster analysis.

Exhibit 8-1 Market Segmentation Through Cluster Analysis



Why is Cluster Analytics Important in Marketing?

Cluster analytics supports:

1. **Customer Segmentation**
   - Identify high-value vs. low-value customers
   - Separate price-sensitive vs. premium buyers
2. **Targeted Marketing**
   - Personalized promotions
   - Customized email campaigns
3. **Product Strategy**
   - Tailor product bundles per segment
   - Adjust pricing strategies
4. **Customer Retention**
   - Identify churn-risk groups
   - Develop loyalty programs per cluster

How It Works (Conceptual Flow)

Step 1: Collect Customer Data

Examples:

- Age
- Income
- Purchase frequency
- Total spending
- Product category preference

Step 2: Select Clustering Method

Common methods:

- **K-Means Clustering** (most popular in marketing)
- Hierarchical Clustering
- DBSCAN (density-based)

Step 3: Algorithm Groups Similar Customers

Customers inside the same cluster:

- Have **high similarity**
- Are **different from other clusters**

Why Cluster Analysis Matters in Marketing

Cluster analysis provides several strategic advantages:

1. Better Customer Segmentation

Instead of treating all customers the same, businesses identify **distinct groups of buyers**.

2. More Effective Marketing Campaigns

Companies can send **targeted promotions** to specific clusters.

3. Improved Product Development

Understanding clusters helps firms develop **products suited to specific customer groups**.

4. Personalized Marketing

Digital platforms can provide **customized experiences** for different customer segments.

For example, **Globe Telecom may classify customers into clusters such as:

- Heavy data users
- Budget prepaid users
- Postpaid premium subscribers

Each group receives different marketing offers.

How Is Cluster Analysis Used in Practice?

Marketers often try to reach different customer segments with a targeted marketing message or product. Personalizing messages for the different needs and wants of each customer segment is more effective than a single message for a market with diverse characteristics. To design these targeted messages, companies need to be able to identify similar traits within unique customer groups. Clustering algorithms can identify groups with similar traits within the groups and meaningful different traits between the groups. When companies understand these sub-groups, they can then target each cluster using unique strategies that specifically fit their needs and wants. The results of clustering help companies develop targeted marketing campaigns and tactics for each cluster. American Express has experienced much success through market segmentation. The company focuses on demographic characteristics such as income and gender, as well as behavioral characteristics such as spending, shopping preferences, psychographic behaviors, and lifestyles. Using a combination of these characteristics, American Express can identify different clusters of loyal customers. Understanding the different needs and wants of these customers enables them to develop new products and targeted marketing. The result has been a sizable increase in market share for the consumer card market, a decline in customer turnover, deeper customer loyalty, and increased card spending.1 Auto companies often use segmentation strategies. Nissan Motor Group manufactures and sells cars and trucks under the Nissan and Infiniti brand names, among others. Each vehicle produced by the brands is targeted at a specific customer segment. Consider the variety of sedans, sports cars, electric cars, crossovers, SUVs, and trucks offered by the company. It is logical that Nissan customers have substantially different characteristics because vehicles range from starting prices of $15,000 to over $113,000. Another automotive manufacturer, Daimler AG, manufactures and sells cars and trucks under the Mercedes-Benz and Daimler brand names. Daimler Mercedes-Benz developed the CLA model car that started at under $30,000 to appeal to a younger audience of potential customers, including Generations X and Y. The company also created an online community known as Generation Benz to engage likeminded customers and capture insights from social interactions. All automobile manufacturers use segmentation strategies to maximize marketing capabilities, such as new product development, pricing strategy, and advertising. Take, for example, Buzz (young, tech enthusiasts), Jill (suburban moms), and Barry (wealthy, professional men).2 These groups are three of five consumer segments Best Buy, a technology retailer, identified and assigned a category name. These segments of loyal customers were developed from similarities in demographic, lifestyle, and marketplace data. Based on this segmentation, Best Buy spent $50 million remodeling some stores and provided training to serve each of the segments. The company also used these insights to enhance in-store customer experiences and improve sales. Stores that underwent changes and focused on these customer segments reported almost 10 percent in sales growth. Clustering enables marketers to identify hidden patterns and structures in the data. For example, Best Buy didn’t hypothesize or predefine the groups of different customers. Instead, the clustering process generated clusters, and the company provided meaning to the clusters. After clusters are identified, companies typically assign names and definitions to each customer cluster. Distinguishing clusters from the larger population is necessary for understanding and responding to different engagement or buying behaviors. But how does a cluster analysis work?

How Does a Cluster Analysis Function?

Cluster analysis organizes data into two or more, similar groups. A measure of similarity between observations and within groups is computed. Groups exhibit within cluster (intra-cluster) homogeneity or similarities, and dissimilar characteristics (heterogeneity) between groups (inter-cluster). Similarities between groups are calculated using various measures of the distance between groups. The measures of distance between groups are applied to individual members in the groups, and individuals are assigned to the group with which they have the most in common. Smaller distances between individual observations represent greater similarity. Moreover, no previous established dependent variable is specified with cluster analysis, so the technique is considered an unsupervised learning. The goal of cluster analysis, therefore, is to model the underlying structure and distribution of characteristics in the data to separate a dataset into homogenous subgroups. Using cluster analysis, a retailer might be interested in segmenting customers based upon data such as income, distance to the store, and the number of purchases at a particular location. Exhibit 8-2a reflects all customers in the retailer’s database—there is no application of cluster analysis and therefore no subgroups

Exhibit 8-2a All Customers Within a Database

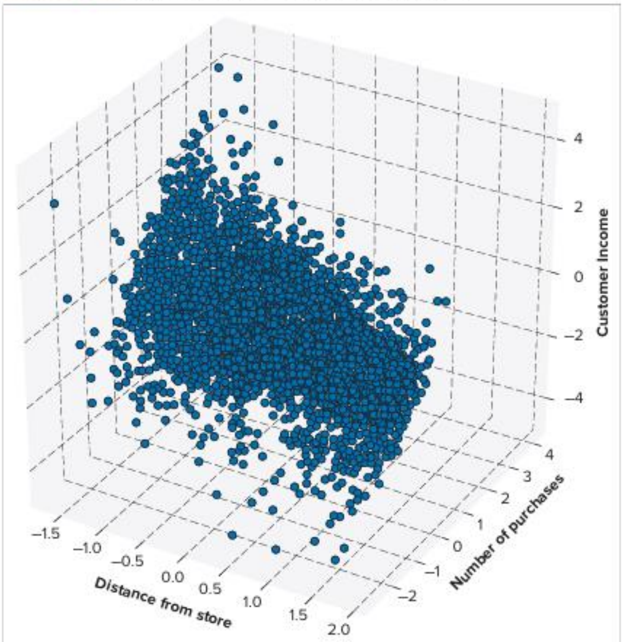
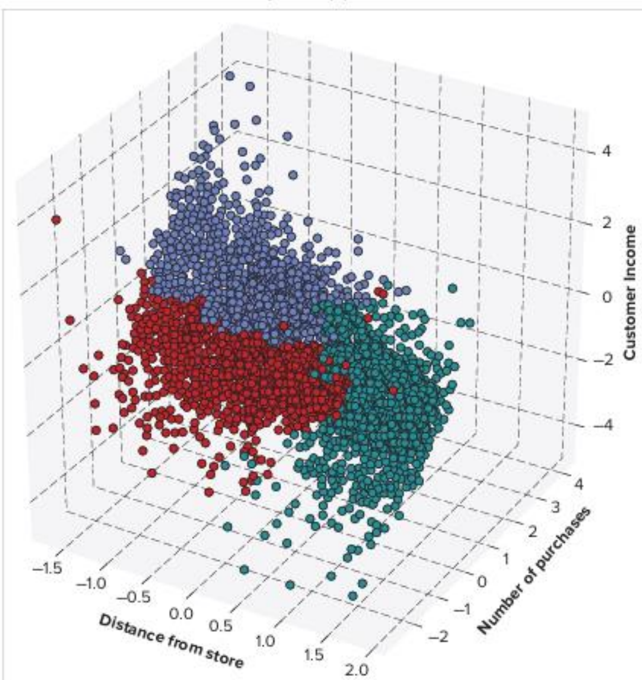


Exhibit 8-2b Cluster Analysis Applied to Customer Database



Using the data for the three specified variables (income, distance, and number of purchases), a cluster analysis can be used to develop separate subgroups. When cluster analysis is applied, Exhibit 8-2b shows the results of a typical market segmentation outcome. What homogenous clusters of customers emerge when considering customer income, their distance to the store, and their number of purchases? In this situation, three clusters emerge from the cluster analysis—

shown as red, purple, and green. Cluster analysis is an efficient and effective method of executing market segmentation to identify subgroups of customers to use in improving business and marketing decisions and therefore performance. To apply cluster analysis, you must know how the technique works. First, there are several different types of cluster analysis. The two most common cluster analysis techniques are k-means clustering and hierarchical clustering.

What Are the Types of Cluster Analysis?

K-Means Clustering

K-Means is one of the **most widely used clustering algorithms** in marketing analytics.

It groups data points into **K number of clusters**, where:

K = number of groups determined by the analyst.

The algorithm assigns each observation to the cluster whose **average value (centroid)** is closest.

How K-Means Works (Simple Explanation)

Step 1 – Choose the number of clusters (K)

Step 2 – Assign customers to the nearest cluster center

Step 3 – Recalculate the cluster center

Step 4 – Repeat until clusters stabilize

Consider a local coffee business similar to **Bo's Coffee** analyzing customer data.

The variables used:

Customer	**Monthly Visits**	**Average Spending**
C1	2	150
C2	15	250
C3	20	300
C4	1	120
C5	12	220

K-Means may produce clusters such as:

Cluster	**Description**
Cluster 1	Occasional customers
Cluster 2	Regular customers
Cluster 3	Premium frequent buyers

Marketing action:

- Cluster 1 → discount coupons
- Cluster 2 → loyalty rewards
- Cluster 3 → exclusive membership

K-means clustering uses the mean value for each cluster and minimizes the distance to individual observations. In this type of analysis, the number of clusters (k) is initially specified by the analyst. Different values of k can be examined, for example, ranging from 2 clusters to 12 clusters. The results for the different clusters are then examined with the best number of different homogenous groups chosen, based on what helps the business develop the most effective strategy. Exhibit 8-3 captures a visual representation of the k-means clustering algorithm. The process begins with the analyst deciding on k initial subgroups to experiment with the number of clusters. A good number k to start with is 2 (for two groups). After determining the initial k (2) clusters, the algorithm randomly assigns each observation to one of the k (2) clusters. A cluster seed is randomly selected and designated as the initial cluster centroid. Then, cluster centroids (means) are calculated. Using cluster centroid values, the k-means algorithm continues to reassign observations based upon the proximity of each observation to the cluster centroid. An observation may start out in one cluster it is close to, but when a new cluster is started, the observation may be reassigned to the new cluster if the observation is closer to the centroid of the new cluster. The k-means cluster algorithm reduces the dispersion within a cluster based on the distance or proximity to the centroid (overall group mean) of an observation within a particular cluster. In short, an overall group mean is calculated, and observations are assigned to the group they are closest to in terms of their characteristics. The clustering process continues to evolve until moving observations between clusters can no longer.

Exhibit 8-3 K-Means Clustering

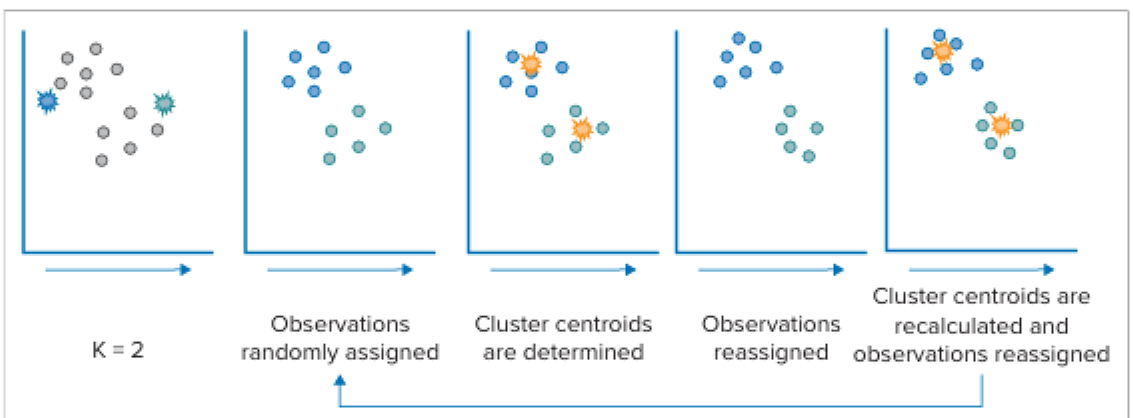
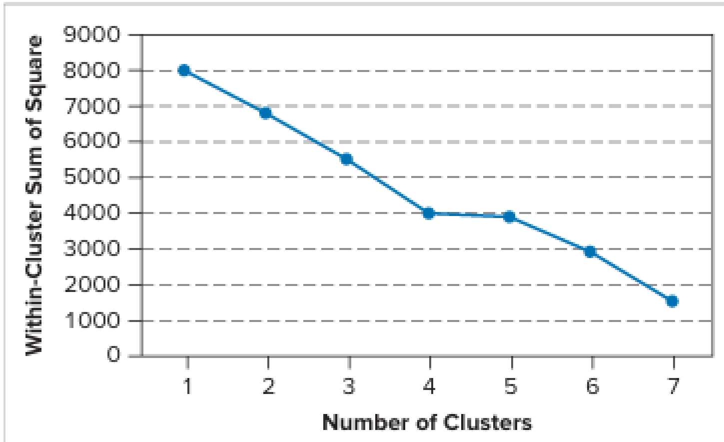


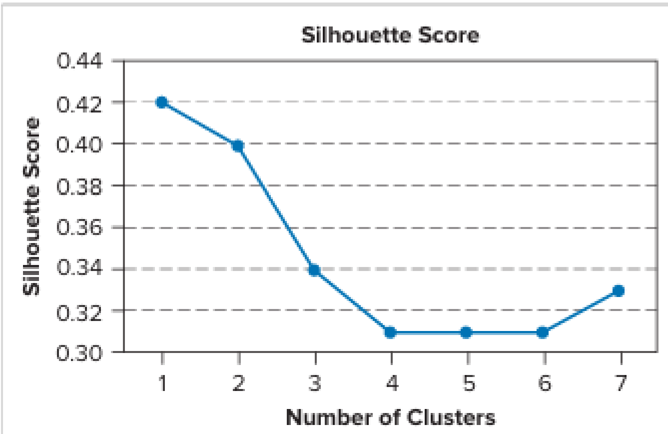
Exhibit 8-4 Elbow Chart



minimize the within-cluster distance. This method is efficient in obtaining solutions and is most often used with large sample sizes because of its simplicity. How is the quality of clusters objectively evaluated? The marketing analyst typically uses a line chart (also referred to as an “elbow chart”) to evaluate the reduction in cluster error (reduction in heterogeneity) as a larger number of clusters is developed. Note in Exhibit 8-4, we see that moving from one to two clusters provides a significant reduction of within-cluster distance (sum of squared error), and so does the move from three to four clusters, but not from four to five. The results in Exhibit 8-4 indicate the optimal number of k (clusters) is four, because the five-group cluster solution does not show a further reduction in the amount of error. Another approach to determining the number of clusters is calculating a silhouette score. The silhouette score is another way to identify the optimal number of clusters for the data. The silhouette score is a measure used in cluster analysis to identify the number of clusters for a given dataset. The silhouette score helps to determine the optimal number of clusters for a given dataset.

The silhouette score is calculated for each data point in the dataset. For each data point, we compute the average distance (a) between the data point and all other points within the same cluster. Then, we compute the average distance (b) between the data point and all other points in the nearest neighboring cluster. Finally, we calculate the silhouette score for the data point as (b - a) / max(a, b). The score ranges from -1 to +1. A score close to +1 indicates a good clustering. The data point is well-matched to its assigned cluster and not to its neighboring clusters. A score close to 0 implies confusion of where the data point belongs. The data point is close to the boundary between its assigned cluster and the nearest neighboring cluster. That means it could belong to either cluster. A score close to -1 suggests that the data point is likely assigned to the wrong cluster, as it is closer to the neighboring cluster than its assigned one. For example, Exhibit 8-5 shows that the optimal number of clusters is four. In addition to examining four clusters, the analyst may decide to also examine solutions for three and even five clusters, and compare them to the four cluster solution. To do so, you would carefully evaluate how meaningful each of the alternative cluster solutions would be in developing a marketing segmentation strategy. The correct number of clusters would be the solution that enables the business to develop the most effective marketing strategy for its customers. In making the decision about the number of clusters, however, there are trade-offs to consider. For example, five clusters will mean five separate marketing strategies must be developed—a unique one for each cluster. In contrast, a solution with three clusters will mean only three unique marketing strategies need to be developed. The more clusters you choose the more

Exhibit 8-5 Silhouette Score

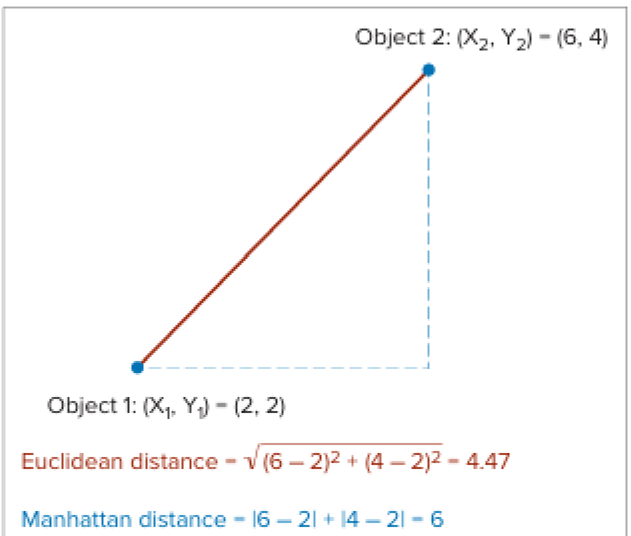


resources the company will need, e.g., qualified marketing strategists and larger budgets to implement those unique strategies, to be successful in attracting different types of customers. Thus, in determining the number of clusters to identify marketers must evaluate the cost of developing and implementing multiple unique marketing strategies as well as identifying customer segments with unique purchasing motives and buying habits. K-Means Issues to Remember When running k-means clustering, it is best to begin with data that has been standardized using z-scores or min-max. Standardization is necessary to obtain the most accurate cluster results—when the clustering variables are measured on very different scales. Recall that in our previous cluster example, we were using the variables income, distance, and number of purchases within a specified time period. Each of the variables is measured differently—income measured in dollars, distance in miles, and the number of purchases measured using a scale of zero to ten purchases. Standardizing converts all three variables to be comparable measures. Another consideration is that the k–means clustering method can only be applied to numerical data. Why is numerical data a requirement? Recall that k-means uses the average datapoint in the cluster as a starting point. Therefore, the mean value is applicable only when the mean is relevant for all of the data used in the clustering process, which requires all data is numerical. If the data is categorical (e.g., Yes or No), the mean would not be meaningful because you cannot average two categories. Thus, if possible, categorical data must be numerically coded before using the k–means algorithm or k-modes algorithm can be used. Hierarchical Clustering A second widely used method for identifying subgroups to use in market segmentation is hierarchical clustering. This method of clustering produces solutions in which the data is grouped into a hierarchy of clusters. Individual observations are combined into subgroups using a measure of distance between observations. There are two commonly used approaches to hierarchical clustering: agglomerative clustering and divisive clustering. With agglomerative clustering (a bottom-up approach), each observation is initially considered to be a separate cluster. That is, if you have 100 observations, you start with 100 separate clusters—one for each observation. Then, in a step-by-step process, each observation is assigned to a cluster that has common characteristics. A linkage method is used to merge smaller clusters into larger clusters. This process of merging observations continues until all observations are included in a single cluster that includes all 100 observations. With divisive clustering (a top-down approach), all records are initially assigned to a single cluster. That is, if you have 100 observations, all 100 observations are specified as being in a

Exhibit 8-6 Ways to Measure Similarity

FUNCTION	DEFINITION
Euclidean	The distance is measured as the true straight line distance between two points.
Manhattan	The distance between two points is not straight—it is a path with right turns as if you are walking a grid in a city. It is also referred to as the "City Block" distance measure.
Matching	Measures the similarity between two observations with values that represent the minimum differences between two points.
Jaccard's	Measures the similarity between two observations based on how dissimilar two observations are from each other.

Exhibit 8-7 Euclidean versus Manhattan



single cluster. Then, a step-by-step process follows in which the most dissimilar observations (records) are sequentially separated from the initial 100-observation cluster. Thus, the process starts with a single cluster of 100 and ends up with 100 different clusters. How is the similarity between observations measured with hierarchical clustering? (See Exhibit 8-6 for a description of the different methods.)

For numerical variables, similarity is most often measured using approaches such as the Euclidean distance or Manhattan distance (see Exhibit 8-7). But if categorical variables are used to cluster, similarity is generally measured using Matching or Jaccard's coefficient. To measure similarity (dissimilarity), a linkage criterion can be used to capture the distance between the resulting clusters. Linkage can be computed using one of the following methods, which are based on linking individual observations both within and between clusters:

• Complete linkage: Similarity is defined by the maximum distance between observations in two different clusters.

• Single linkage: Similarity is defined by the shortest distance from an object in a cluster to an object from another cluster.

• Average linkage: Similarity is defined by the group average of observations from one cluster to all observations from another cluster.

A visual example of the three similarity measures based on linking observations is shown in Exhibit 8-8. A fourth method of measuring distances with hierarchical clustering is Ward's method. Ward's method applies a measure of the sum of squares within the clusters summed

Exhibit 8-8 Similarity Measures

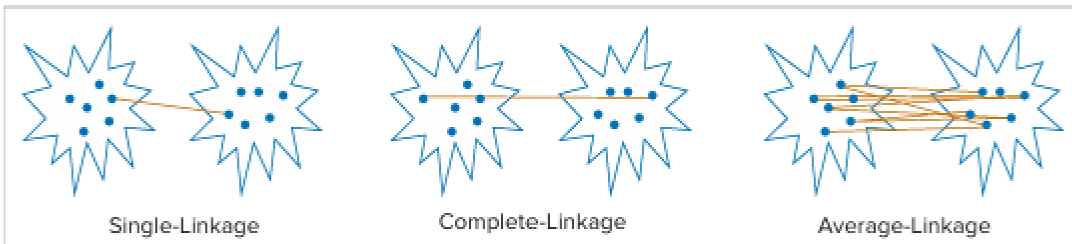
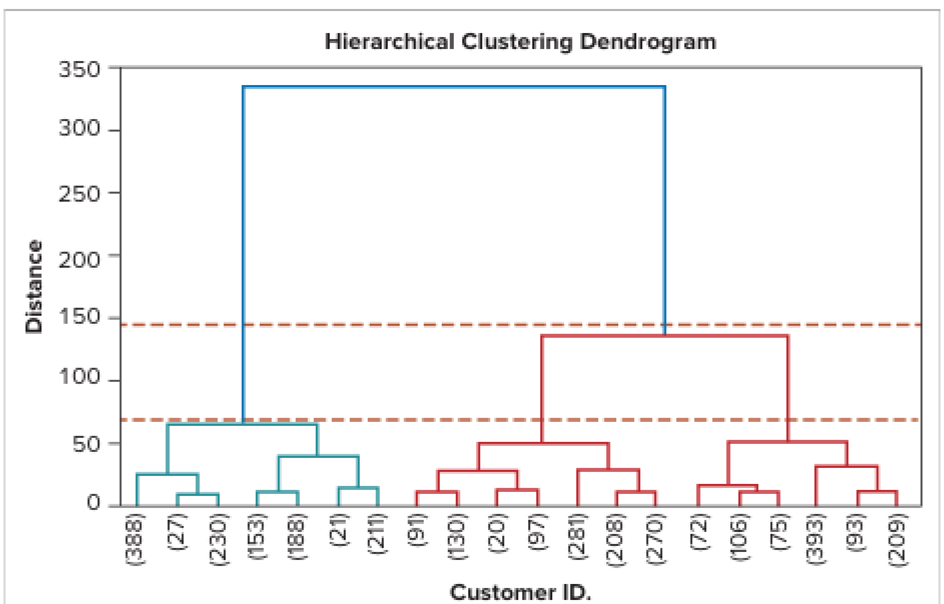


Exhibit 8-9 Hierarchical Clustering Dendrogram



over all variables. The Ward's method process selects two clusters to combine based on which combination of clusters minimizes the within cluster sum of squares for all clusters across all of the separate clusters. All four measurement approaches for determining clusters when applying hierarchical clustering can be illustrated with a dendrogram. Exhibit 8-9 shows a dendrogram that illustrates the process for hierarchical clustering. The treelike graph provides an illustration of the hierarchy of clusters in the dataset. We can see in the dendrogram example in Exhibit 8-9 that there are two main branches. Each of these branches represents a possible cluster. The height of each branch indicates the distance, measuring dissimilarity. In other words, the longer the vertical line is, the more separated the cluster is from the other clusters. Thus, when you look at the graph, you may be able to visually decide how many clusters are needed to represent the correct number of clusters for the data. As an example, if you impose a horizontal line on the dendrogram and count the number of vertical lines it crosses, it will suggest the number of clusters. The two gray dotted lines suggest the possibility of either two clusters (higher line) or three clusters (lower line). Identifying the correct number of clusters using a dendrogram is a subjective process. The analyst takes into account their knowledge of the markets served by the business to assess the practical meaning of the clusters emerging from the dendrogram. To develop an understanding of the markets served, it is helpful to review the descriptive statistics for each cluster to define the characteristics of the clusters and assign them an appropriate name. For example, a grocery delivery service may find there are two clusters. The characteristics of the one cluster include higher income, weekend shoppers, between the ages of 40 and 50, and less price sensitive. The second cluster is made up of weekday shoppers, with medium incomes, between the ages of 25 and 35, and shopping mainly for discounted items. This understanding will enable the marketing analyst to identify the first cluster as a market segment that represents “Higher Income, Weekenders” and the second cluster as representing “Weekday, Price Sensitive” customers. Additionally, the number of clusters is two, so it should be relatively easy to develop a cost-effective strategy to serve both clusters. If you are developing a marketing strategy, segmenting customers into a smaller number of clusters will be easier to manage than a larger number. Hierarchical Clustering Issues to Remember While k-means works only with numerical values, hierarchical clustering can be executed with a mixed set of data that can include both categorical and numerical values. Also, similar to k-means clustering, the hierarchical clustering process should be executed with standardized data developed using either z-scores or min-max.

Hierarchical Clustering

Hierarchical clustering is another clustering method that **builds clusters step by step**.

Instead of choosing K immediately, this method builds a **tree-like structure called a dendrogram**.

There are two types:

1. **Agglomerative clustering (bottom-up)**
2. **Divisive clustering (top-down)**

Most marketing analytics applications use **agglomerative clustering**.

Agglomerative Clustering Process

Step1
Each data point starts as its own cluster.

Step2
The algorithm merges the two most similar clusters.

Step3
Clusters continue merging until only one cluster remains.

Example (Philippine Retail)

Consider **SM Supermalls** analyzing shopper behavior.

Variables collected:

- Shopping frequency
- Spending level
- Preferred product category

Hierarchical clustering might reveal groups such as:

Cluster	Customer Type
Cluster 1	Grocery shoppers
Cluster 2	Fashion shoppers
Cluster 3	Entertainment seekers

Marketing strategies could include:

- Grocery promotions
- Fashion sales events
- Cinema bundle promos

Difference Between K-Means and Hierarchical Clustering

Feature	K-Means Clustering	Hierarchical Clustering
Number of clusters	Must be specified first	Determined later
Speed	Faster	Slower
Data size	Works well with large datasets	Better for small datasets
Output	Cluster groups	Tree diagram (dendrogram)

Cluster analytics is widely used in many industries:

Banking

Banks like **BDO Unibank analyze customers based on:

- loan usage
- savings patterns
- credit card spending

Telecommunications

Companies such as **Smart Communications classify users based on:

- data consumption
- prepaid vs postpaid
- digital service usage

E-commerce

Platforms like **Lazada segment customers by:

- shopping behavior
- preferred brands
- purchase frequency

These insights help businesses create **data-driven marketing strategies**.

Promotion and Advertising Effectiveness in Marketing Analytics

Promotion and advertising are essential elements of marketing strategy because they influence how consumers become aware of products and ultimately decide to purchase them. However, companies today cannot rely solely on intuition or creativity when designing promotional campaigns. Instead, modern businesses use **marketing analytics** to measure whether their advertising investments actually generate results.

Marketing analytics allows organizations to analyze promotional campaigns using **data-driven methods**, including exploratory data analysis (EDA) and regression analysis. These tools help companies identify patterns in marketing data, understand relationships between advertising expenditures and sales performance, and predict the potential impact of future campaigns.

Through **exploratory data analysis**, marketers can examine patterns such as trends in advertising spending, customer engagement, and sales performance. After identifying potential relationships in the data, analysts apply **regression models** to quantify the impact of advertising variables such as promotional budget, social media engagement, or discount offers on sales outcomes.

This module introduces students to how businesses evaluate promotion and advertising effectiveness using analytics. It focuses on the role of EDA in understanding marketing data and regression analysis in measuring the impact of promotional activities.

Understanding Promotion and Advertising

Promotion in Marketing

Promotion refers to the communication strategies used by businesses to inform, persuade, and remind customers about their products or services.

Promotional tools include:

- Advertising
- Sales promotions
- Social media marketing
- Influencer marketing
- Public relations
- Email marketing

The main goal of promotion is to **increase awareness, stimulate demand, and influence purchasing behavior**.

Advertising

Advertising is a **paid form of communication used to promote products, services, or brands** through various media channels.

Common advertising channels include:

- Television
- Social media
- Online ads
- Radio
- Billboards
- Search engines

Advertising effectiveness refers to the **extent to which advertising influences consumer behavior and improves sales performance**.

Why Businesses Measure Advertising Effectiveness

Companies must ensure that marketing investments produce measurable returns.

Key questions businesses want to answer include:

- Does advertising increase sales?
- Which advertising channels are most effective?
- How much should the company spend on advertising?
- What type of promotion generates the highest customer response?

Marketing analytics provides tools to answer these questions using **data-driven methods**.

1. Promotion and Advertising in Marketing

Promotion refers to the marketing activities that communicate the value of a product or service to potential customers. Advertising, on the other hand, is a paid form of promotion used to reach a wider audience.

Common promotional tools include:

- Television advertisements
- Social media campaigns
- Online advertisements
- Promotional discounts
- Influencer marketing
- Sales promotions

Companies invest significant financial resources in promotional campaigns. Because advertising costs can be high, organizations must evaluate whether these campaigns generate sufficient returns.

Marketing analytics helps businesses answer questions such as:

- Does advertising increase sales?
- Which marketing channels generate the best results?
- How much advertising investment is optimal?

2. Exploratory Data Analysis in Marketing

What is Exploratory Data Analysis (EDA)?

Exploratory Data Analysis (EDA) is the process of **examining and summarizing data to identify patterns, trends, and relationships before applying statistical models**.

EDA is often the first step in marketing analytics because it helps analysts understand the structure of the dataset and detect possible relationships between variables.

Common EDA techniques include:

- Data visualization
- Descriptive statistics
- Trend analysis
- Correlation analysis
- Outlier detection

Through EDA, marketers gain insights that guide further analysis and modeling.

Advertising Data Exploration

Suppose a company collected the following monthly marketing data:

Month	Advertising Budget	Social Media Engagement	Sales
Jan	₱100,000	5,000	₱850,000
Feb	₱120,000	6,200	₱920,000
Mar	₱150,000	7,500	₱1,050,000

Common EDA Techniques in Marketing

1. Descriptive Statistics

Descriptive statistics summarize marketing data using metrics such as:

- Mean
- Median
- Minimum
- Maximum
- Standard deviation

Example:

Average advertising spending = ₱65,000
Average monthly sales = ₱467,500

---

2. Data Visualization

Charts help analysts quickly observe relationships in data.

Examples include:

- Scatter plots
- Bar charts
- Trend lines
- Histograms

A **scatter plot of advertising spending vs. sales** often reveals whether a relationship exists between the variables.

Using EDA, analysts may observe that **higher advertising spending appears to correspond with increased sales and engagement**. However, this observation alone does not confirm causation. To measure the strength of the relationship, regression analysis is required.

Example:

Fast Food Advertising Campaign

One of the most recognizable advertising brands in the Philippines is **Jollibee Foods Corporation**.

The company frequently launches emotional television commercials and social media campaigns to promote products such as Chickenjoy and value meals.

Marketing analysts in the company evaluate whether advertising campaigns increase:

- Restaurant traffic
- Sales revenue
- Online engagement
- Brand awareness

By analyzing advertising expenditure and sales data, the company can determine whether their promotional strategy is effective.

4. Regression Analysis for Advertising Effectiveness

After conducting exploratory data analysis, the next step is to quantify the relationship between promotional activities and sales outcomes using regression models.

Regression analysis helps answer the question:

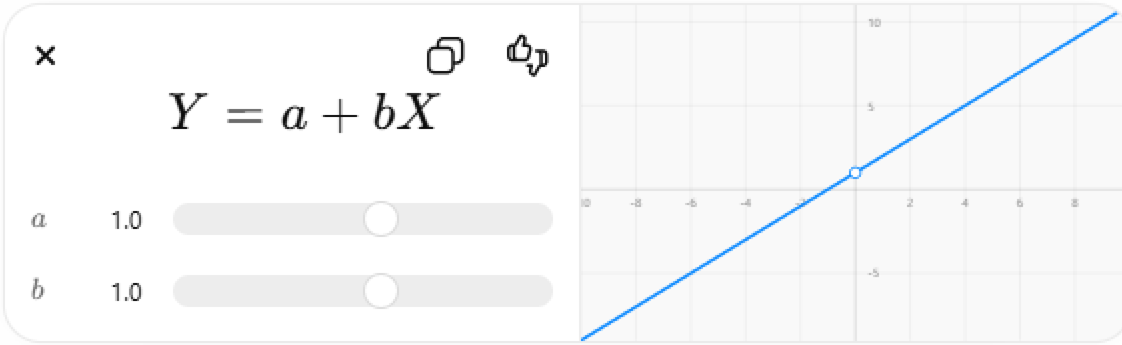
How much do advertising efforts influence sales performance?

Simple Linear Regression in Advertising

Simple regression examines the relationship between **one independent variable and one dependent variable**.

Example:

Independent Variable (X) = Advertising Budget
Dependent Variable (Y) = Sales Revenue



Where:

- **Y** = predicted sales
- **a** = intercept (baseline sales without advertising)
- **b** = advertising effect
- **X** = advertising expenditure

Example Regression Model

Suppose a regression analysis produces the equation:

Sales = 500,000 + 4(Advertising Budget)

Interpretation:

- Even without advertising, baseline sales are **₱500,000**.
- Every additional **₱1 spent on advertising increases sales by ₱4**.

This result suggests that advertising has a **positive impact on sales performance**.

Multiple Regression for Marketing Campaign Analysis

In real business environments, sales are rarely influenced by advertising alone. Multiple marketing factors may simultaneously affect consumer behavior.

Multiple regression allows analysts to examine the effect of **several independent variables**.

Example model:

Sales = a + b_1(Advertising) + b_2(Promotion Discount) + b_3(Social Media Engagement)

Interpretation:

- Advertising increases brand awareness
- Discounts stimulate purchase decisions
- Social media engagement improves customer interaction

Using this model, marketers can determine **which promotional variable contributes the most to sales growth**.

Example:

E-Commerce Advertising Analytics

E-commerce platforms such as **Shopee** rely heavily on digital advertising analytics.

Marketing analysts examine how promotional strategies such as:

- Flash sales
- Online advertisements
- Influencer promotions
- Discount vouchers

affect product demand.

For example, during a **9.9 Mega Sale Campaign**, analysts may observe that:

- Advertising spending increased by **40%**
- Website traffic increased by **70%**
- Sales increased by **60%**

Regression analysis helps determine whether these increases are statistically related to promotional spending.

7. Evaluating Advertising Effectiveness

Marketing analysts use several metrics to determine whether promotional campaigns are effective.

Return on Advertising Spend (ROAS)

ROAS measures how much revenue is generated for every peso spent on advertising.

Example:

Advertising cost = ₱100,000
Revenue generated = ₱500,000

ROAS = 5

This means every **₱1 spent on advertising generates ₱5 in revenue**.

Customer Acquisition Cost (CAC)

CAC measures the cost of acquiring a new customer.

Lower CAC indicates a more efficient marketing campaign.

Conversion Rate

Conversion rate measures how many potential customers actually make a purchase.

Example:

10,000 visitors → 500 purchases

Conversion rate = 5%

8. Using Excel for Promotion Analytics

Students can analyze advertising effectiveness using **Microsoft Excel**.

Steps include:

1. Organizing advertising and sales data.
2. Conducting **descriptive statistics** to understand trends.
3. Creating **scatter plots** to visualize relationships.
4. Running **regression analysis** using the Data Analysis Toolpak.
5. Interpreting regression coefficients and R-squared values.

These techniques allow marketers to transform raw marketing data into actionable insights.

Pricing Analytics and Price Elasticity Using Regression Analysis

Pricing is one of the most powerful decisions a company makes because it directly influences both **customer demand and company profitability**. In today's data-driven business environment, organizations do not rely solely on intuition to determine prices. Instead, they apply **pricing analytics**, which uses statistical models and data analysis to understand how customers respond to different price levels.

One of the most widely used tools for pricing analytics is **regression analysis**, a statistical method used to examine relationships between variables. In pricing studies, regression helps businesses estimate how changes in price affect **sales volume, revenue, and market demand**.

This module introduces students to the use of **simple and multiple regression models** for analyzing price-demand relationships and evaluating the **price elasticity of demand**. Students will also learn how to validate regression models and assess predictive performance to ensure that pricing decisions are reliable and data-driven.

By the end of this module, learners will understand how companies can use regression analytics to set optimal prices, forecast demand, and improve revenue management.

1. Pricing Analytics in Marketing

What is Pricing Analytics?

Pricing analytics refers to the **use of data, statistical methods, and analytical models to determine the best price for products or services**.

Businesses analyze historical sales data, consumer behavior, competitor prices, and economic conditions to identify the price that maximizes profit while maintaining competitiveness.

Pricing analytics answers questions such as:

- How will demand change if the price increases?
- What price maximizes revenue?
- How sensitive are customers to price changes?

These questions are often answered using **regression models and elasticity analysis**.

A company like **Jollibee Foods Corporation** regularly evaluates the pricing of its value meals.

Suppose Jollibee wants to study the effect of increasing the price of a **Chickenjoy meal**.

Historical data may look like this:

Price (₱)	Weekly Sales
85	7,200
90	6,900

Price (₱)	Weekly Sales
95	6,500
100	6,100

Using regression analysis, analysts can estimate how price increases affect demand and determine whether raising the price will increase or decrease total revenue.

Price Elasticity of Demand

Definition

Price elasticity of demand measures how **responsive customer demand is to changes in price**.

Elasticity tells managers whether a small change in price will significantly affect sales volume.

Elasticity Formula

Price Elasticity of Demand is calculated as:

$$Elasticity = \frac{\%\ Change\ in\ Quantity}{\%\ Change\ in\ Price}$$

Where:

- **E_n (Elasticity of Demand)** = responsiveness of demand
- **$\%\Delta Qd$** = percentage change in quantity demanded
- **$\%\Delta P$** = percentage change in price

Interpretation of Elasticity Values

Elasticity Value	Meaning	Business Interpretation
E > 1	Elastic Demand	Customers are sensitive to price changes
E = 1	Unit Elastic	Price change proportionally affects demand
E < 1	Inelastic Demand	Customers are less sensitive to price

Example 1: Milk Tea Shop Pricing (Elastic Demand)

Price	**Daily Sales**
₱120	200 cups
₱140	150 cups

Step 1: Calculate Percentage Change in Quantity

$$\frac{150 - 200}{200} \times 100 = -25\%$$

Step 2: Calculate Percentage Change in Price

$$\frac{140 - 120}{120} \times 100 = 16.67\%$$

Step 3: Calculate Elasticity

$$Elasticity = \frac{-25\%}{16.67\%} = -1.50$$

Interpretation

Elasticity = **1.5 (Elastic Demand)**

This means:

- Customers are **very sensitive to price increases**
- When price increased by **16.67%**, sales dropped **25%**

Marketing Insight

For milk tea businesses:

- Increasing price may reduce sales significantly
- Promotions and discounts may attract more buyers

This insight helps marketing managers decide whether to **increase price or focus on promotional strategies**.

3. Example 2: Gasoline Pricing (Inelastic Demand)

Companies like **Petron Corporation** often face **inelastic demand**.

Data Example

Price	Weekly Sales
₱65 per liter	50,000 liters
₱70 per liter	48,000 liters

Step 1: Percentage Change in Quantity

$$\frac{48,000 - 50,000}{50,000} \times 100 = -4\%$$

Step 2: Percentage Change in Price

$$\frac{70 - 65}{65} \times 100 = 7.69\%$$

Step 3: Elasticity

$$Elasticity = \frac{-4}{7.69} = -0.52$$

Interpretation

Elasticity = **0.52 (Inelastic Demand)**

Interpretation

Elasticity = **0.52 (Inelastic Demand)**

This means:

- Demand is **not highly affected by price changes**
- Customers still need fuel even if price increases

Marketing Insight

Fuel companies can increase prices with **less reduction in demand**, allowing them to maintain revenue.

How Price Elasticity Works

Price elasticity becomes powerful when combined with **marketing analytics techniques like regression analysis and predictive modeling**.

Marketing analysts use data to estimate elasticity from **real customer purchasing behavior**.

Role of Regression Analysis

Regression models estimate the relationship between **price and sales**.

Example regression model:

$$Sales = 8000 - 20(Price)$$

Interpretation:

- Every **₱1 increase in price reduces demand by 20 units**

Using this model, analysts can simulate different price scenarios.

Example Pricing Simulation

Price	Predicted Demand	Revenue
₱90	6200	₱558,000
₱95	6000	₱570,000
₱100	5800	₱580,000

Why Price Elasticity is Important

Price elasticity allows companies to make **data-driven marketing decisions**.

Key Applications

1. Pricing Strategy

Businesses determine whether they should **increase or decrease prices**.

2. Promotion Planning

If demand is elastic, **discounts and promotions can significantly increase sales**.

3. Revenue Optimization

Companies can find the **price that maximizes revenue**.

4. Market Segmentation

Different customer segments may have different elasticity levels.

Example:

Students → highly price sensitive
Professionals → less price sensitive

Real Marketing Analytics Example in E-Commerce

Online platforms like **Shopee** analyze price elasticity using large datasets.

Marketing analysts examine:

- Product price
- Discounts
- Flash sales
- Advertising

For example:

When Shopee introduces a **20% flash sale**, demand may increase by **60–80%**.

This indicates **high price elasticity**, which explains why e-commerce platforms frequently use promotions.

Price elasticity of demand is a critical concept in marketing analytics because it allows businesses to **quantify how price changes affect customer demand**.

Through techniques such as **regression analysis, predictive modeling, and demand forecasting**, companies can identify the most effective pricing strategies, optimize revenue, and remain competitive in the market.

Organizations that effectively use pricing analytics gain a strong advantage because they make **pricing decisions based on data rather than intuition**.

Predictive Marketing Analytics

Module Overview

Marketing has evolved significantly in the digital age. Instead of simply analyzing past performance, modern organizations use advanced analytics to **predict future customer behavior, sales trends, and marketing outcomes**. This approach is known as **predictive marketing analytics**.

Predictive marketing analytics uses statistical models, machine learning algorithms, and artificial intelligence to forecast customer actions and support strategic marketing decisions. Businesses apply predictive analytics to anticipate future demand, identify high-value customers, personalize marketing campaigns, and optimize marketing budgets.

Several analytical methods support predictive marketing analytics. **Regression analysis** is one of the most widely used techniques for predicting relationships between variables. More advanced models such as **neural networks** allow companies to identify complex patterns in large datasets. Meanwhile, **Automated Machine Learning (AutoML)** simplifies predictive modeling by automatically selecting the most effective algorithms and parameters.

This module introduces students to the foundations of predictive marketing analytics and demonstrates how businesses apply predictive models to improve marketing performance and decision-making.

What is Predictive Marketing Analytics?

Predictive marketing analytics refers to the use of **historical data, statistical models, and machine learning techniques to forecast future marketing outcomes**.

Rather than asking:

“What happened in the past?”

Predictive analytics asks:

“What is likely to happen in the future?”

Businesses use predictive analytics to forecast:

- Future sales
- Customer purchasing behavior
- Customer churn
- Marketing campaign performance
- Demand for products and services

These insights help marketing managers make **proactive and data-driven decisions**.

Importance of Predictive Marketing Analytics

Predictive analytics provides several strategic advantages for organizations.

Improved Decision-Making

Marketing managers can forecast market trends and prepare strategies before changes occur.

Personalized Marketing

Companies can predict customer preferences and deliver personalized marketing messages.

Customer Retention

Predictive models can identify customers who are likely to stop purchasing, allowing companies to intervene with retention strategies.

Budget Optimization

Marketing teams can allocate advertising budgets to campaigns with the highest predicted returns.

Example: Streaming Platforms

Companies such as **Netflix** rely heavily on predictive analytics to recommend content to users.

By analyzing viewing history, search behavior, and user preferences, predictive algorithms forecast which movies or shows a user is most likely to watch.

This personalized recommendation system significantly improves user engagement and customer retention.

Regression Analysis in Predictive Marketing

Regression analysis is one of the most fundamental predictive modeling techniques used in marketing analytics.

It estimates the relationship between a **dependent variable** and one or more **independent variables**.

Example:

Predicting **sales based on advertising spending**.

Sales = a + bX

Where:

- **Sales** = predicted sales revenue
- **X** = advertising expenditure
- **a** = intercept
- **b** = coefficient measuring advertising impact

Advertising Prediction

Suppose a company develops the regression model:

Sales = 200,000 + 3(Advertising)

Interpretation:

If advertising spending increases by **₱1**, sales are expected to increase by **₱3**.

Marketing managers can use this model to estimate future sales under different advertising budgets.

Example: Online Advertising

E-commerce platforms like **Shopee** use regression models to predict the impact of digital advertising campaigns.

Marketing analysts examine variables such as:

- Advertising spending
- Discount percentage
- Website traffic
- Promotional events

Regression models allow the company to forecast how marketing investments will affect future sales.

Neural Networks in Marketing Analytics

Neural networks are advanced machine learning models inspired by the structure of the human brain.

They consist of interconnected layers of nodes (neurons) that process data and learn patterns from large datasets.

Neural networks are particularly useful when relationships between variables are **complex and nonlinear**, which traditional regression models may struggle to capture.

Applications of Neural Networks in Marketing

Neural networks are used in marketing analytics to:

- Predict customer purchasing behavior
- Detect fraudulent transactions
- Recommend products
- Forecast demand
- Personalize marketing messages

These models can process massive datasets and discover hidden relationships that traditional statistical methods might miss.

Real Business Example: Online Retail Recommendation Systems

Companies like **Amazon** use neural networks to power their recommendation systems.

The system analyzes customer behavior, including:

- Products viewed
- Items purchased
- Browsing patterns
- Customer ratings

Using neural networks, the platform predicts which products customers are most likely to buy next.

This technology significantly increases cross-selling and upselling opportunities.

Automated Machine Learning (AutoML)

Automated Machine Learning simplifies the process of building predictive models.

Traditionally, developing machine learning models required extensive expertise in data science and programming. AutoML platforms automate several steps in the modeling process, including:

- Data preprocessing
- Feature selection
- Model selection
- Hyperparameter tuning
- Model evaluation

This allows analysts and marketers with limited programming experience to build powerful predictive models.

Benefits of AutoML

Faster Model Development

Automated systems test multiple algorithms quickly.

Improved Model Accuracy

AutoML tools identify the best-performing models for a given dataset.

Accessibility

Non-technical users can develop predictive models without advanced programming skills.

Example: Marketing Automation Platforms

Platforms such as **Google** integrate automated machine learning into their advertising tools.

For example, **automated bidding strategies** in Google Ads use machine learning to predict which ad placements are most likely to generate conversions.

The system continuously analyzes data such as:

- User search behavior

- Device type
- Time of day
- Previous engagement

Based on these predictions, the system automatically adjusts advertising bids to maximize campaign performance.

Applications of Predictive Marketing Analytics

Organizations use predictive marketing analytics in many strategic areas.

Customer Segmentation

Predictive models identify high-value customer segments.

Sales Forecasting

Businesses estimate future product demand.

Churn Prediction

Companies identify customers who are likely to stop purchasing.

Marketing Campaign Optimization

Predictive models estimate which campaigns will generate the highest returns.