DATA ANALYTICS

Chapter 1: Introduction to Data Analytics				
Teaching Hours: 10	Marks Distribution			
	Remember = 04 M	Understanding= 04M	Applying = 08 M	Total = 16 M

Unit - I Introduction to Data Analytics

1.1 Data Analytics: An Overview, Importance of Data Analytics

1.2 Types of Data Analytics: Descriptive Analysis, Diagnostic Analysis, Predictive Analysis, Prescriptive Analysis, Visual Analytics

1.3 Life cycle of Data Analytics, Quality and Quantity of data, Measurement

1.4 Data Types, Measure of central tendency, Measures of dispersion

1.5 Sampling Funnel, Central Limit Theorem, Confidence Interval, Sampling Variation

1.1 Data Analytics: An Overview, Importance of Data Analytics
Data Analytics :

1.1 Data Analytics: An Overview

Data analytics perform a vital role in every organization and technologies such as health care, financial trading, Internet of Things, Smart Cities or Cyber Physical Systems. However, these diverse applications are driving us to new research challenges. In this context, the book provides a broad picture on the concepts, techniques, applications, and open research directions in data analytics.

Data analytics is a single source of reference for researchers and application developers in the areas of data analytics, data science, and big data. It is also a valuable resource for graduate students, research scholars, data analysts, and technical practitioners who are interested in designing and developing advanced data analytics technologies, systems, and applications.

Data analytics may cover corporate procedures, enhancing decision-making and encouraging business growth. Data analytics is the science of examining raw data to obtain insightful information. It is used in many industries to allow organizations to make better decisions, and verify and disprove existing theories or models. With the help of data analytics, commercial industries can describe, predict, and improve business performance. It can also be used to support smart decisions.

Analyzing data sets to gain information that can the applied to solve problems across various sector is a component of the discipline of data analytics. It uses some disciplines computer programming, statistics, and mathematics, to provide precise dat specialization, and undeck your earning potential in this dynamic Geld. There is huge a Discover the diverse career opportunities data analytics, from entry-levelses In which the broad view of data analytics is shown. Experts and Scientist enlarge structure it quickly and utilize the same in an efficient manner to provide specifically to the needs of every customer

For example, a person who has just come to know about a particular product might a certain advertise that is very different from the one that is shown to the user that known about that same product for months. In order to make this possible, it is businesses to gather consumer data that is unstructured, then structured properly, meaning of the data with their property attribute so that

customer can use the same t achieve this level data scientist plays an important role in transforming raw data manen into actionable data

The data analyst use their skill to Improve businesses with the knowledge needed customize strategies for maximum impact in a highly competitive market. Executing much amount of data requires huge man power with certain skill which makes j opportunities.

The role of a data analyst is to give answers to difficult questions just by observing patter and connecting them to major incidents that took place over a period of time in the r past. So in this book we study about data analyses, type of data, various techniques to organized data, to extract data and etc.

Data analytics involves various techniques to extract meaningful insights from data, and i commanly categorized into four types descriptive, diagnostic, predictive, and prescriptose Data analytics is the procedure of analyzing unstructured data to obtain insightful, relevant information that can be used to guide and support smart company decisions Analyzing dara sets to gain information that can be applied to solve problems across various sectors is a component of the discipline of data analytics. It uses some disciplines, including computer programming, statistics, and mathematics, to provide precise data analysis.

1.1.1 Importance of Data Analytics

Data analytics is crucial for businesses as it provides insights to improve decision-making, streamline operations, and gain a competitive edge.

By analysing data, organizations can identify trends, patterns, and relationships to optimize processes, reduce costs, and personalize customer experiences.

This, in turn, leads to increased efficiency, innovation, and ultimately, higher profitability.

1. Informed Decision-Making

- Data analytics empowers businesses to make data-driven decisions rather than relying on guesswork or intuition. By analysing data, businesses can identify patterns and trends that can help them make predictions, anticipate customer needs, and identify new opportunities
- This leads to more effective strategies, whether in marketing, product development, or operations.

2. Improved Efficiency and Cost Savings

Data analysis helps identify inefficiencies and bottlenecks in business processes, allowing for streamlining and cost reduction. By analyzing resource allocation and process data. organizations can identify areas where they can cut expenses, boost productivity, and save time. This can lead to reduced operational costs and increased profitability.

3. Enhanced Customer Experience

Data analytics provides insights into customer behaviour, preferences, and needs, allowing businesses to personalize their offerings and interactions. This leads to improved customer satisfaction, loyalty, and retention. By understanding customers better, businesses can tailor their products, services, and marketing efforts to meet their specific needs.

4. Competitive Advantage

Data analytics helps businesses stay ahead of the competition by identifying emerging trends, predicting future demand, and making informed decisions. Organizations that leverage data analytics can gain a competitive edge by optimizing their operations, improving customer experiences, and developing new products and services. This can lead to increased market share, brand loyalty, and profitability.

5. Innovation and Growth

- Data analytics can uncover new opportunities for growth and innovation by revealing emerging trends, market gaps, and unmet customer needs
- markets, and develop innovative solutions. This can lead to increased revenue, marker share, and long-term growth.

1.2 Types of Data Analytics

Now we can discuss about four main kinds of data analysis now that we have a basic four types of data analytics:

- Descriptive Analytics

  The form of analysis known as descriptive analytics is straightforward and focuses on the surface level of previous events.

- Diagnostic Analytics-

  Diagnostic analytics focuses on the "why" of certain occurrences or patterns rather than just the "what" of descriptive analytics. Further in-depth data analysis is required to identify the variables and reasons that resulted in certain results.

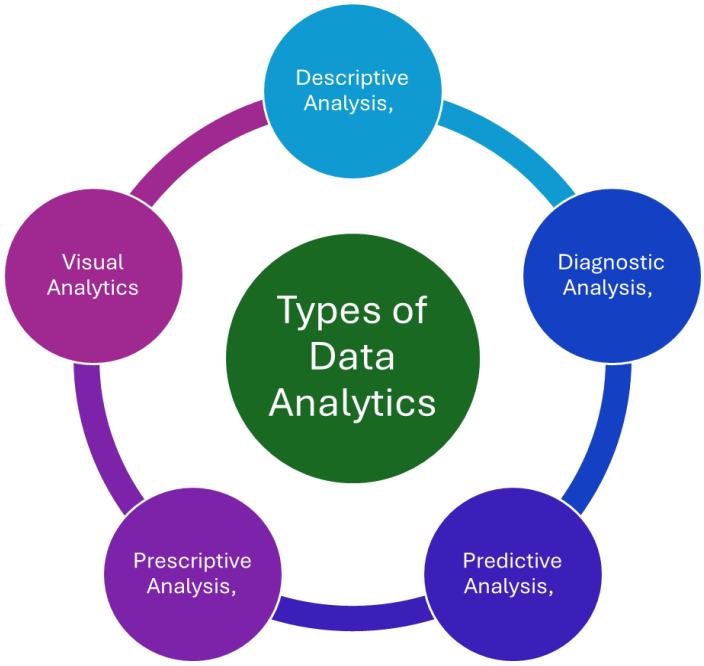
- Prescriptive Analytics

  It makes recommendations for the courses of action and choices to be made.

- Predictive Analytics -

  Predictive analytics seeks to forecast the expected course of events, as

the name indicates.



1.2.1 Descriptive Analytics

- Descriptive analytics is a statistical interpretation used to analyze historical data to identify patterns and relationships. Descriptive analytics seeks to describe an event, phenomenon, ar outcome. Descriptive analytics is about finding meaning within data. Data needs context analytics provide the where and when turning figures into measurable patterns. As a form of data analysis, descriptive analytics is one of the four key types of data analytics.

- Descriptive analytics is the process of using current and historical data to identify trends and relationships. It's sometimes called the simplest form of data analysis because it describes trends and relationships but doesn't dig deeper. Descriptive analytics is relatively accessible and likely something your organization uses daily.

- Following are various tool used to analysis of data like Microsoft Excel or data visualization tools, such as Google Charts and Tableau. These tools can help parse data, identify trends and relationships between variables, and visually display information. Descriptive analytics is especially useful for communicating change over time and uses trends as a springboard for further analysis to drive decision-making.

Examples of descriptive analytics:

1) Traffic and Engagement: Reports One example of descriptive analytics is reporting. If your organization tracks engagement in the form of social media analytics or web traffic, you are already using descriptive analytics. These reports are created by taking raw data generated when users interact with your website, advertisements, or social media content and using it to compare current metrics to historical metrics and visualize trends. For example, you may be responsible for reporting on which media channels drive the most traffic to the product page of your company's website. Using descriptive analytics, you can analyze the page's traffic data to determine the number of users from each source. You may decide to take it one step further and compare traffic source data to historical data from the same sources. This can enable you to update your team on movement; for instance, highlighting that traffic from paid advertisements increased 20 percent year over year.

1.2.2 Diagnostic Analytics

Diagnostic analytics is a branch of data analysis that focuses on identifying the reasons behind trends, patterns, and anomalies in data. It helps organizations understand why something happened by drilling deeper into historical data, uncovering correlations, and identifying root causes.

Key features of diagnostic analytics :

1. Root Cause Analysis (RCA)

Determines the underlying causes of business performance issues or successes.

2. Drill-Down and Data Mining

Explores data at multiple levels to pinpoint specific factors influencing outcomes.

3. Correlation and Causation

Uses statistical methods to determine relationships between variables.

4. Comparative Analysis

Examines performance over time, across different segments, or against benchmarks.

5. Advanced Data Visualization

Utilizer charts, heat maps, and dashboards for better insights.

Common techniques used in diagnostic analytics:

- **Regression Analysis** : Identifies relationships between dependent and independen variables.
- **Hypothesis Testing:** Validates assumptions using statistical methods.
- **Machine Learning Models:** Predicts outcomes based on historical patterns.

1.2.3 Prescriptive Analytics:

Prescriptive analytics is the most advanced form of data analytics, providing actionable recommendations based on historical data, machine learning, and optimization algorithms. It goes beyond descriptive (what happened) and diagnostic (why it happened) analytics by suggesting the best course of action to achieve desired outcomes.

Key features of prescriptive analytics:

1. **Decision Optimization:** Suggests the best actions based on available data.
2. **Predictive Modeling:** Uses machine learning to forecast future outcomes.
3. **Scenario Analysis:** Evaluates different strategies to determine optimal results.
4. **Real-Time Recommendations:** Continuously adapts based on live data.
5. **Automated Decision-Making:** integrates with Al-driven systems for efficiency.

Examples of Prescriptive Analytics in Business:

1. **Retail:** Suggesting optimal pricing and promotions to maximize sales
2. **Healthcare :** Recommending personalized treatment plans for patients.
3. **Supply Chain:** Identlying the best logistics route to minimize delays and costs
4. **Finance:** Advising Investment strategies to maximize returns.

1.2.4 Predictive Analytics: Forecasting the Future with Data

Predictive analytics is a data-driven approach that uses statistical techniques, machine learning, and historical data to forecast future trends, behaviors, and outcomes. It helps businesses and organizations anticipate what is likely to happen so they can make informed decisions

Key features of predictive analytics:

1. Forecasting Trends: Uses historical data to predict future patterns.
2. Risk Assessment: identifies potential risks and opportunities
3. Customer Behavior Prediction: Anticipates customer preferences and actions.
4. Fraud Detection: Detects anomalies to prevent fraud
5. Market Demand Prediction: Helps businesses prepare for changes in demand.

1.2.5 Visual Analytics

Visual analytics combines data visualization with analytical processes and interactive tools to explore and understand complex data, enabling users to uncover patterns, make predictions, and support decision-making. It goes beyond simple data visualization by incorporating advanced H] techniques like machine learning and predictive modeling to forecast future trends and outcomes based on historical data.

How visual analytics contributes to predictive data analysis:

Uncovering hidden patterns and insights:

A) Data Exploration

Visual analytics provides a platform for interactive exploration of datasets, allowing users to drill down inte specific areas and identify anomalies, outliers, and relationships that might not be apparent in raw data.

B) Pattern Recognition

Visual representations like charts, graphs, and maps help users spot patterns and trends in data that can be used to build predictive models.

C) Predictive Modelling and Forecasting

Combining Data with Algorithms: Visual analytics integrates predictive modeling techniques with data visualization, enabling users to explore the potential impact of different scenarios and make informed decisions based on future projections.

D] Real-time Insights

Interactive dashboards and visualizations can display real-time data, allowing for immediate insights and proactive adjustments to plans based on predicted outcomes.

E] Improved Decision-Making

Data-Driven Decisions By providing a clear and concise view of data and predicted outcomes, visual analytics enables users to make more informed decisions based on evidence rather than gut feeling

F] Proactive Strategies

Predictive visual analytics helps organizations anticipate potential problems and opportunities, enabling them to develop proactive strategies for risk management, optimization, and growth.

G] Applications Marketing

Predictive visual analytics can help optimize marketing campaigns by identifying high-value leads, predicting customer behavior, and optimizing content and customer journeys.

H] Sales

Visual analytics can help sales teams track performance, identify areas for improvement, and forecast future sales.

I] Supply Chain

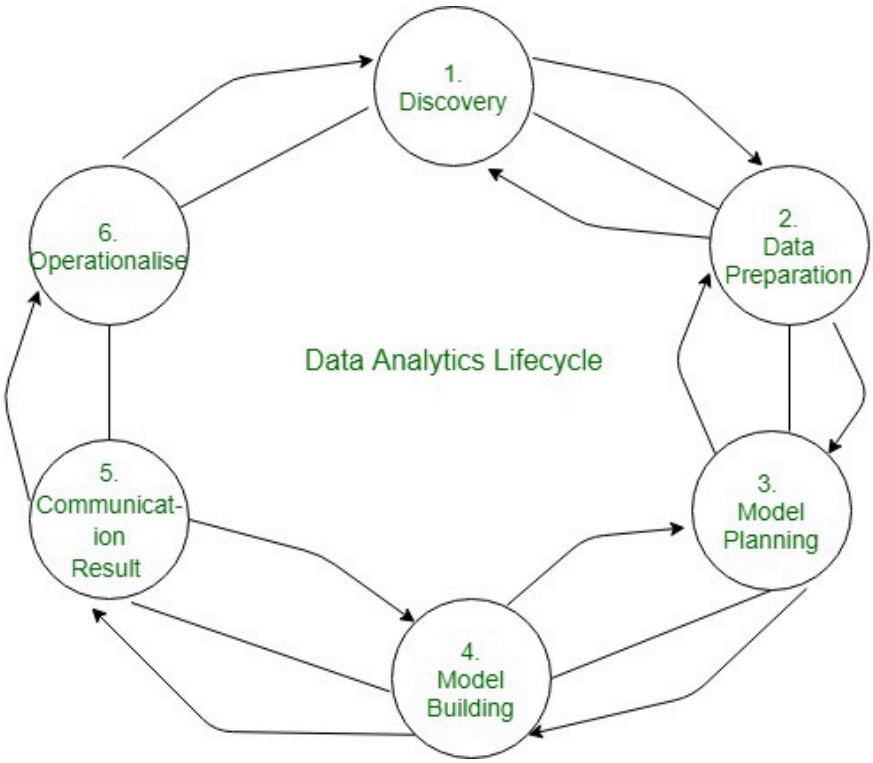
Visual analytics can help optimize supply chain operations by visualizing inventory levels, shipment statuses, and demand fluctuations.

J] HR

Visual analytics can help HR professionals identify patterns in employee data, such as turnover rates and employee engagement, to improve retention and engagement.

1.3 Life cycle of Data Analytics

The data Analytics Life cycle defines analysis process best practices to discover project completion. The Data analytic lifecycle is designed for Big Data problems and data science projects. The cycle is iterative to represent a real project. The synthesis was developed after gathering input from data scientists and consulting established that provide input or piece of the process. Following Fig. 1.3.1 shows the importance of Data Analytics Life cycle.



Phase 1: Discovery

The data science team learns and investigates the problem. Develop context and understanding Come to know about data sources needed and available for the project. The team formulates the initial hypothesis that can be later tested with data. The term learns business domain, including relevant history, such as business unit has attempted similar project in term of people technology, time, and data. Important activities in this phase including framing the business problem as an analytics challenge that can be

addressed in subsequent phase and formulating initial hypothesis to test and begin learning the data.

Phase 2: Data Preparation

This phase requires the presence of sandbox. In which team can be work with data and perform analytics for the duration of the project. Steps to explore preprocess and condition data before modeling and analysis. It requires the presence of an analytic sandbox, the team executes, loads, and transforms, to get data into the sandbox. Data preparation tasks are likely to be performed multiple times and not in predefined order. Several tools commonly used for this phase are-Hadoop, Alpine Miner, Open Refine, etc.

Phase 3: Model Planning

The team explores data to learn about relationships between variables and subsequently, selects key variables and the most suitable models.

In this phase, the data science team develops data sets for training, testing, and production purposes. Team builds and executes models based on the work done in the model planning phase Several tools commonly used for this phase are-MATLAB and STASTICA.

Phase 4: Model Building

- Team develops datasets for testing, training, and production purposes.
- Team also considers whether its existing tools will suffice for running the models or if they need more robust environment for executing models.
- Free or open-source tools Rand PL/R, Octave, WEKA, Commercial tools - MATLAB and STASTICA.

Phase 5: Communication Results

- After executing model team need to compare outcomes of modeling to criteria established for success and failure.
- Team considers how best to articulate findings and outcomes to various team members and stakeholders, taking into account warning, assumptions.
- Team should identify key findings, quantify business value, and develop narrative to summarize and convey findings to stakeholders.

Phase 6: Operationalize

- The team communicates benefits of project more broadly and sets up pilot project to deploy work in controlled way before broadening the work to full enterprise of users.
- This approach enables team to learn about performance and related constraints of the model In production environment on small scale which make adjustments before full deployment The team delivers final reports, briefings and codes.

1.3.1 Quality and Quantity of Data in Analytics

Quality of Data

High-quality data ensures accurate, reliable, and actionable insights. Key dimensions include:

- **Accuracy:** Data should be correct and free from errors.
- **Completeness:** No missing or incomplete values.
- **Consistency:** Uniform data across multiple sources.
- **Timeliness:** Data should be up to date.
- **Relevance:** Data should be meaningful to the analysis.

Quantity of Data

The amount of data impacts the depth of analysis:

Small Data: Suitable for simple statistical analysis, small-scale decision-making.

Big Data: Large volumes of structured and unstructured data, requiring specialized tools like Hadoop, Spark.

Balanced Data: A sufficient amount of high-quality data leads to better model performance and insights.

Optimal data : analytics requires both high-quality and sufficient quantity of data to ensure reliable insights and decision-making

1.3.2 Measurement in Data Analytics

Measurement in data analytics refers to the process of collecting, quantifying, and analyzing data to derive meaningful insights. It involves defining key metrics, selecting appropriate methods , and ensuring accuracy for decision making.

Type	Definition	Example
Nominal	Categorical data with no order.	Customer gender (Male, Female, Other).
Ordinal	Ordered categories with ranking	Customer satisfaction (Low, Medium, High)
Interval	Numeric data with equal intervals, no true zero.	Temperature in Celsius.
Ratio	Numeric data with a true zero point.	Revenue, Age, Sales figures.

Importance of accurate measurement:

- Ensures data-driven decision-making.
- Improves model performance in Al and ML
- Enhances business intelligence and forecasting.
- Identifies trends and anomali

1.4 Data Types, Measures of Central Tendency and Measures of Dispersion

1.4.1 Data Types

Data types in analytics determine how data is categorized and analyzed. They are classified Into as follows:

A. Qualitative (Categorical) Data

- **Nominal Data** - Categories without any order.

  Example: Gender (Male, Female, Other), Colors (Red, Blue, Green).

- **Ordinal Data**-Categories with a meaningful order but no fixed interval.

  Example: Satisfaction levels (Low, Medium, High), Education level (Bachelor's, Master's, PhD).

B. Quantitative (Numerical) Data

Discrete Data - Whole numbers/countable values.

Example: Number of students in a class, Number of orders placed.

Continuous Data - Measured values with infinite precision.

Example: Height, Temperature, Weight,

1.4.2 Measures of Central Tendency

Measures of central tendency describe the center or typical value of a dataset.

1. Mean (Average)

Formula:

Mean = ΣX/N

2. Median

The middle value when data is arranged in order.

Example: The median salary in a company gives a better sense of typical earnings when there are extreme values.

3. Mode

The must frequently occurring value in a dataset.

Example: The most popular product sold in a store.

Useful for categorical data.

1.4.3 Measures of Dispersion

Measures of dispersion describe the spread or variability in a dataset.

1. Range

Formula: **Range = Maximum Value - Minimum Value**

Example: If the highest salary in a company is $200.000 and the lowest is $30,000, the range 1. Is $170,000.

2. Variance(σ^2)

Measures how far data points are from the mean.

Formula (for a population): $\sigma^2 = N\Sigma(X - \mu)^2$

Example: Stock price volatility.

3. Standard Deviation (σ)

Square root of variance, measuring data spread in original units.

Formula: $\sigma = N\Sigma(X - \mu)^2$

Example : A lower standard deviation in exam scores means students performed more consistently

4. Interquartile Range (IQR)

Measures the middle 50% of data.

Formula:

IQR : Q3 - Q1

Example: Used to detect outliers in salary distributions.

1.5 Sampling Funnel: Understanding the Process of Data Selection

A sampling funnel refers to a structured approach for narrowing down a large dataset of population into a smaller, more representative subset for analysis. It ensures that the selected sample maintains the key characteristics of the overall population while reducing bias and improving efficiency.

Why Use a Sampling Funnel?

Efficiency: Reduces the time and cost of analyzing large datasets.

Accuracy: Ensures representativeness while minimizing bias

Scalability: Allows businesses to test hypotheses before full-scale implementation.

1.5.1 Stages of the Sampling Funnel

1. Population Identification

- Define the entire group from which you want to sample
- Example: All customers of an e-commerce platform.

2. Target Population Selection

- Identify a specific segment relevant to the analysis.
- Example: Customers who made a purchase in the last six months.

3. Sampling Frame Creation

- Develop a list of elements that meet selection criteria.
- Example: Customer database containing emails, purchase history and demographics.

4. Sampling Method Selection

Choose an appropriate sampling technique:

Probability Sampling: Random, Stratified, Cluster Sampling

Non-Probability Sampling: Convenience, Judgment, Snowball Sampling

5 . Sample Extraction

Apply the chosen method to select a subset of data.

Example: Randomly selecting 1.000 customers from the database.

6. Data Cleaning and Preparation

Remove inconsistencies, missing values, or irrelevant entries.

Ensure data quality before analysis.

7. Analysis and Interpretation

Conduct statistical or machine learning analysis on the sample.

Use insights to infer conclusions about the broader population.

1.5.2 Central Limit Theorem (CLT)

The Central Limit Theorem (CLT) states that, regardless of the original populations distribution, the sampling distribution of the sample mean will approach a normal (bell-shaped distribution as the sample size increases, provided the samples are independent and randomly selected.

Why is CLT Important?

Justifies Using Normal Distribution: Even if the population is skewed, we can apply normal distribution assumptions for large samples.

Facilitates Hypothesis Testing: Enables techniques like t-tests and confidence intervals in inferential statistics.

Supports Predictive Modeling: Helps in making data-driven decisions based on sample

data. If we take random samples of size n from a population with any distribution that has a finite mean u and finite standard deviation o, then as the sample size n increases, the distribution of the sample mean will approach a normal distribution with:

Mean: **μ**

Standard deviation (Standard Error): σn

Key Principles of CLT

1. Normality of Sampling Distribution:

If you take multiple random samples from any population, their means will form a normal distribution as sample size increases.

2. Mean of Sampling Distribution:

The mean of the sample means ($\mu_{\bar{x}}$) is equal to the population mean (**μ**).

$$\mu_{\bar{x}} = \mu$$

3. Standard Error (SE):

The standard deviation of the sample means ($\sigma\bar{x}$) decreases as sample size increases, following the formula:

$\sigma\bar{x} = \sigma n$

Where:

$\sigma\,\bar{x}$ = Standard error of the mean

σ = Population standard deviation

n = Sample size

Example Scenario

Imagine you are a statistician studying the number of hours people spend on their phones each day in a city. **Let's assume:**

The population mean (u) of phone usage is 4 hours per day.

The population standard deviation (0) is 2 hours.

The population distribution is not normal (maybe it's skewed or has a different shape).

Now, you want to estimate the average number of hours people in the city spend on their phones, but instead of surveying the entire population, you take a sample of 50 people.

Steps to Illustrate the Central Limit Theorem

Step 1: Take multiple random samples from the population. You take several random samples of size 50 (say, 30 or 50 samples), and for each sample, calculate the sample mean of phone usage.

Step 2: Plot the sample means. Even though the original population distribution might be skewed or have some other shape, the distribution of the sample means will approach a normal distribution as the sample size (50) gets large enough.

Step 3: The sample mean's dieribution. According to the Central Limit Theorem, the distribution of the sample means will:

- Have a mean equal to the population mean 4 hours
- Have a standard deviation equal to the population standard deviation divided by the square root of the sample size, le.

  Standard Error (SE) = σn 250-0.28 hours
- So, even if the original distribution of phone usage is not normal, the sampling distribution the sample means will be approximately normal with:

  Mean = 4 hours.

  Standard Error = 0.28 hours.

Step 4: Apply the CLT. As you continue to collect more samples of size 50, the distribution t the sample means will get closer to a normal distribution. The Central Limit Theorem ensures that for large enough sample sizes, the distribution of the sample means will be approximately normal, regardless of the shape of the population distribution.

Step 5: Inference. Now, you can use the normal distribution to make inferences about the population mean. For example, you can calculate the probability that the average number of hours spent on the phone in a sample of 50 people is between 3.5 t 4.5 hours using the normal distribution.

Summary of the CLT

In this example:

Population Mean (μ): 4 hours

Population Standard Deviation (σ): 2 hours.

Sample Size (n): 50 people

Standard Error: 0.28 hours

Sampling Distribution: The distribution of sample means (from samples of 50 people) will be normal with mean 4 and standard deviation 0.28.

1.5.3 Confidence Interval

Computing a confidence interval

- For a normal population with known Standard Deviation &, we can sample and compute a Confidence Interval for the Population Mean **μ**
- Suppose we want to estimate the true yield difference for 2 corn hybrids. From data comparing many pairs of corn hybrids, we know that the differences are normally distributed, and we also have very good knowledge of the standard deviation of these differences. We have a sample of 16 locations where the 2 corn hybrids are both planted and take the difference in hybrid yields. Each difference is an individual from a population of all such differences. A 95% confidence interval for the true average difference in hybrid yields is centered on the average difference (2) plus and minus 2 standard errors of the mean ($\sigma/\sqrt{\mathbf{n}}$).

Estimating a true mean

We use properties of the normal distribution and sample averages to get this confidence Interval. The reasoning is as follows.

Suppose we have a normal distribution with known standard deviation s not known, and we wish to estimate it. e, but the true mean u

1. We know from the properties of the normal distribution that 95% of the individuals differ from the true population mean by no more than 1.96 (about 2) standard deviations.
2. By the central limit theorem, 95% of sample averages will differ from the population mean by less than about two standard errors ($\sigma/\sqrt{\mathbf{n}}$).
3. Therefore, we can estimate u using the sample average. Y, and an interval of 1.96 ($\sigma/\sqrt{\mathbf{n}}$) above and below it
4. We have 95% confidence that this procedure will give correct results, provided our assumptions are met.

1.5.4 Sampling Variation

Also known as sampling error or sampling distribution, refers to the differences in statistics (like the mean) that occur when taking multiple samples from the same population. This natural variation happens because each sample is a subset of the larger population and will likely have slightly different characteristics.

Sampling distribution

A sampling distribution is the theoretical distribution of all possible sample statistics (eg. sample means) that could be obtained from a population. It helps in understanding the probability of obtaining a certain sample mean given the true population mean.

Applications of data analytics

Applications for data analytics have advanced throughout time due to developments in the industry. Here are some crucial areas where data analytics excels:

1. Education: Officials may utilize data analytics to improve management and education decisions. These programs would improve administrative control and learning. You may gather preference information from each student and utilize it to develop a curriculum to enhance the existing curriculum.

2. Delivery and logistics: Data analytics are used in logistics to optimize delivery process and streamline operations. As a result, the sector has fared better, which has increased th number of customers. It boosts productivity by facilitating real-time data exchanges and business insights across partners.

3. Digital marketing and advertising: Marketers utilize data analytics to understand the audience and achieve high conversion rates. Digital advertising professionals emplo analytics to learn about the target market's demographics, including their age, race, gender and other characteristics. Additionally, they filter their audience using this technology accordance with their tastes and behaviors.

4. Transportation: The transport sector might be completely transformed thanks to dat analytics. It is especially helpful when moving a sizable group to a site demanding smooth movement. By enhancing transportation systems and intelligence, data analytics may be use to reduce traffic congestion and enhance travel.

5. Security: Data analytics, particularly predictive analytics, is used by security staff to forecast future instances of crime or security breaches. Additionally, they can investigate recent assaults. Analytics enables the investigation of additional potential holes, the action of end users or devices involved in a security breach, and the manner in which IT system were penetrated during an assault.