Class IX

Lesson Plan

Data Science

1. Introduction

Teaching Objectives

Students will learn about

- Data
- Information
- Indian Information Technology Act, 2000
- Information Concept and Processing
- Information Technology (IT)
- Data, Information, Knowledge and Wisdom (DIKW) Model
- Applications of DIKW Model
- Influence of Data in Your Life
- What are Data Footprints?
- Data Loss and Recovery

Number of Periods	
Theory	Practical
x	x

Teaching Plan

While teaching this chapter, tells the students data is an individual piece of information that contains raw facts and figures about several things around us.

Tell the students Information is processed and interpreted data placed within a meaningful context. The information generated by one system acts as data to other powerful system.

Discuss with the students the qualities of information may be defined in terms of Relevance, Availability, Timeliness, Accuracy, Completeness, Meaningful and Action-Oriented, Conciseness, Commensurate, and Confidence in the Source.

Ask the students to perform **Activity 1** given on Page 157 of the main course book.

Discuss with the students DIKW model describes how the data can be processed and transformed into information, knowledge and wisdom and then finally decisions in pursuit of achieving goals are made.

Tell the students there are various industries, such as healthcare, travel, education are rapidly influenced from data.

Ask the students to perform **Activity 2 and 3** given on Page 171 of the main course book.

Share with the students that digital footprint is a unique data trace of a user's activities, actions, communications or transactions in digital media.

Ask the students digital footprint can be used to track the user's activities and devices.

Tell the students that an active footprint is made when information is collected from the user without the person knowing this is happening.

Share with the students a passive digital footprint is made up of the information that companies are harvesting behind the scenes, such as browsing data, IP addresses and purchasing habits.

Ask the students to perform Activity 4 given on Page 175 of the main course book.

Tell the students data can be lost due to a system crash when the system stops abruptly.

Tell the students that the process of restoring the inaccessible, lost, corrupted, damaged or deleted data is called data recovery.

Ask the students to perform Activity 5 given on Page 176 of the main course book.

Extension

Ask the students some oral questions based on this chapter.

Q. What is the difference between data and information?

Q. Explain the process of transforming data into knowledge and wisdom.

Q. Why is relevance important when considering data and information?

Q. How does availability affect the value of information?

Q. What were the impacts of the introduction of personal computers in the 1980s?

Q. What are the objectives of information technology, and how does it assist management in decision-making?

Q. How does the DIKW Model categorise the processing stages from data to wisdom?

Q. Describe the role of digital footprints in online shopping.

Q. What are the active components of a digital footprint?

Q. What are some reasons for data loss mentioned in the text?

Q. What is the process called for restoring lost or corrupted data?

Q. What is the significance of frequent data backups, especially for large enterprise systems?

Encourage the students to walk through the chapter and ask them to explain any one topic from the chapter.

Evaluation

After explaining the chapter, let the students do the exercises given on pages 176 to 181 in the main course book as **Exercise (Solved and Unsolved Questions)**.

Tell them to solve the critical and thinking skill developing exercises as **Higher Order Thinking Skills** given on pages 181.

Ask the students to practice the project in given in **Applied Project** section given on Page 182 in the main course book.

Suggested Activity

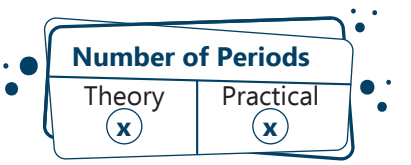
Ask the students suppose you accidentally deleted an important document from your computer. What steps would you take to try to recover the lost data, and how could you prevent similar accidents in the future?

2. Arranging and Collecting Data

Teaching Objectives

Students will learn about

- ☞ Data Collection
- ☞ Data Collection through Smart Gadgets and Devices
- ☞ Types of Data
- ☞ Types of Data for Business and Research
- ☞ Defining Big Data
- ☞ Univariate and Multivariate Data



Teaching Plan

While teaching this chapter, tells the students demand for data insights is increasing exponentially. This data may be structured, unstructured and semistructured.

Tell the students the method of gathering data for calculating and analysing reliable insights is known as data collection.

Discuss with the students data collection is a methodical process of gathering information in a systematic manner to answer stated questions, test some hypotheses and evaluate the results.

Tell the students Data collection methods can be classified as: Primary data collection and Secondary data collection.

Discuss with the students interviews, questionnaires, surveys, case studies, groups discussions, photography and videography, etc are some ways for data collection.

Tell the students interviews are organised by the data specialist at the cost of organisers or by researcher himself/herself.

Tell the students questionnaire is a set of questions arranged logically and divided into groups with the objective of gathering information for social, political and business research.

Tell the students survey is one of the common methods of diagnosing and solving social problems.

Discuss with the students focus group discussion involves gathering data from people from similar backgrounds or experiences together to discuss a specific topic of interest.

Tell the students data collection apps are mobile applications that make it possible to collect data from a smartphone, tablet or iPad.

Tell the students data collection software is a computerised system for the collection and storage of qualitative and quantitative data in an electronic form.

Ask the students to perform **Activity 1** given on Page 193 of the main course book.

Tell the students data can be quantitative and qualitative. Quantitative data is numerical information (numbers). Qualitative data is descriptive information (it describes something)

Ask the students to perform **Activity 2 and 3** given on Page 196 and 197, respectively, of the main course book.

Share with the students big data can be defined as the technologies and initiatives that involve data that is too diverse, fast-changing or massive for conventional technologies, skills and infrastructure to address efficiently.

Tell the students that big data includes huge volume, high velocity, variety, veracity and value of data. The data in it will be of three types: unstructured data, semi structured data and structured data.

Ask the students unstructured data refers to information that either does not have a predefined data model or is not organised in a predefined manner.

Tell the students that semi-structured data is information that doesn't reside in a relational database but that does have some organisational properties that make it easier to analyse.

Share with the students Structured data concerns all data which can be stored in database like SQL (Structured Query Language) in table with rows and columns.

Ask the students in univariate situation, there is one variable condition. It does not involve multiple parameters or relationships.

Tell the students in multivariate situation, data involves a relationship among multiple variables.

Ask the students to perform **Activity 4** given on Page 216 of the main course book.

Extension

Ask the students some oral questions based on this chapter.

Q. Why is the demand for data insights increasing exponentially?

Q. How can data collection methods be classified?

Q. Who organises interviews for data collection purposes?

Q. Define data collection software and its function.

Q. What are the three types of data found in big data?

Q. Why is it important to gather data systematically rather than haphazardly?

Q. In what ways can big data benefit various industries?

Q. Why is it crucial to consider both quantitative and qualitative data in research?

Q. Can you think of an example where big data has been effectively utilised to drive decision-making?

Q. What steps would you take to analyse data collected from a focus group discussion?

Encourage the students to walk through the chapter and ask them to explain any one topic from the chapter.

Evaluation

After explaining the chapter, let the students do the exercises given on pages 217 to 224 in the main course book as **Exercise (Solved and Unsolved Questions)**.

Tell them to solve the critical and thinking skill developing exercises as **Higher Order Thinking Skills** given on pages 224.

Ask the students to practice the project in given in **Applied Project** section given on Page 224 in the main course book.

Suggested Activity

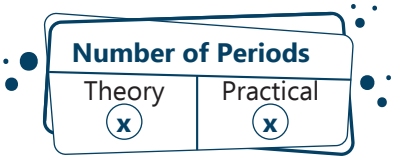
Ask the students suppose you are analysing data collected from various sources using data collection software. What advantages does using such software offer over manual data entry methods?

3. Data Visualisation

Teaching Objectives

Students will learn about

- ☞ Importance of Data Visualisation
- ☞ Introducing OpenOffice
- ☞ Plotting Available Data Histograms
- ☞ Shapes of Histogram and their Meaning
- ☞ General Mistakes done while Plotting Histograms
- ☞ Why Skew Shape Study is Important?
- ☞ Difference between Histogram and Bar Graph
- ☞ When to Use Histogram?
- ☞ Use of Single Variable and Multi Variables Plots
- ☞ Sales Performance



Teaching Plan

While teaching this chapter, tells the students that Data visualisation is a technique to represent data graphically. It helps to identify trends, patterns and even outliers within a large data set.

Tell the students Trend analysis is a method in technical analysis that helps in predicting the future movement based on the current trending data.

Ask the students to perform **Activity 1** given on Page 229 of the main course book.

Discuss with the students Sales dashboards are a method of measuring sales performance from a bird's-eye view. They help measure key metrics, individual team members and sales activities.

Ask the students to perform **Activity 2** given on Page 232 of the main course book.

Tell the students Apache OpenOffice is the leading open-source office software suite for word processing, spreadsheets, presentations, graphics, databases and more.

Discuss with the students Apache OpenOffice packages contains Writer, Calc, Impress, Draw, Base and Math applications.

Tell the students A dot plot is a type of simple histogram-like chart used in statistics for relatively small data sets where values fall into a number of discrete bins (categories). It is used for displaying the distribution of quantitative variable where each dot represents a value.

Tell the students bar chart or bar graph is a chart or graph that presents categorical data with rectangular bars with heights or lengths proportional to the values that they represent. These bars can be horizontal or vertical.

Ask the students to perform **Activity 3** given on Page 243, respectively, of the main course book.

Tell the students an outlier is a value that is much larger or smaller than the other values in a data set, or a value that lies outside the given data set.

Discuss with the students number of occurrences of a particular data value in a data set is known as its frequency.

Tell the students a histogram is a popular graphing tool that is used to summarise discrete or continuous data which is measured on an interval scale.

Tell the students Most histograms can be classified into five types: Normal Distribution, Symmetric Distribution, Bimodal Distribution, Right (Positively) Skewed Distribution, and Left (Negatively) Skewed Distribution.

Tell the students uniform shaped histogram indicates data that is very consistent; the frequency of each class is very similar to that of the others.

Share with the students histogram is symmetric if you cut it down the middle and the left-hand and

right-hand sides resemble mirror images of each other.

Tell the students that random distribution lacks an apparent pattern and has various peaks.

Tell the students that skewness does not tell us about the number of outliers. It only tells us the direction.

Tell the students about the certain condition in which the histogram graph is used.

Ask the students to perform **Activity 4** given on Page 259 of the main course book.

Share with the students single variable plot, also known as univariate plot, is used to visualise a continuous variable.

Ask the students multi-variable plots, also known as bivariate plots, are used to display relationship between two variables. They allow you to understand the degree and the pattern of relation between the two variables.

Ask the students to perform **Activity 5** given on Page 268 of the main course book.

Tell the students frequency table is a tabular representation that summarises the raw categorical data.

Ask the students to perform **Activity 6** given on Page 272 of the main course book.

Tell the students Heterogeneous data are any data with high variability of data types and formats.

Tell the students Homogeneous data are made up of data types and formats that are similar to each other.

Extension

Ask the students some oral questions based on this chapter.

Q. What is data visualisation, and how does it assist in data analysis?

Q. Explain the concept of trend analysis.

Q. Describe the components of Apache OpenOffice software suite.

Q. Differentiate between a bar chart and a histogram.

Q. What is frequency and how is it calculated in a dataset?

Q. Explain the purpose of a histogram in data representation.

Q. Enumerate the five types of histograms commonly encountered and their characteristics.

Q. What are multi-variable plots, and why are they useful in data analysis?

Q. What is a frequency table?

Q. Differentiate between heterogeneous and homogeneous data.

Q. How can data visualisation techniques help identify patterns in large datasets?

Q. Can you explain the concept of binning in histograms?

Encourage the students to walk through the chapter and ask them to explain any one topic from the chapter.

Evaluation

After explaining the chapter, let the students do the exercises given on pages 279 to 284 in the main course book as **Exercise (Solved and Unsolved Questions)**.

Tell them to solve the critical and thinking skill developing exercises as **Higher Order Thinking Skills** given on pages 284.

Ask the students to practice the project in given in **Applied Project** section given on Page 284 in the main course book.

Suggested Activity

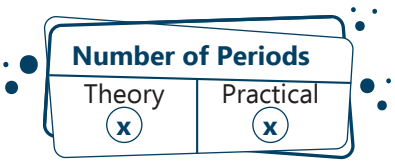
Ask the students plot a histogram to visualise the distribution of sales across different months from the given dataset that contains the monthly sales figures for a retail store chain.

4. Ethics in Data Science

Teaching Objectives

Students will learn about

- ☞ Data Ethics
- ☞ Ethical Guidelines around Data Analysis
- ☞ Common Causes of Data Quality Problems
- ☞ Need for Ethical Guidelines in Data Science
- ☞ Goals of Ethical Guidelines
- ☞ Data Governance Framework
- ☞ Goals of Data Governance



Teaching Plan

While teaching this chapter, tells the students that Ethics can be defined as a system of principles that help in identifying right or wrong and good or bad.

Tell the students Data ethics can be defined as a set of moral responsibilities for collecting, generating, analysing and disseminating data and is concerned with the impacts of this data on society.

Discuss with the students the primary ethical and moral question around the sharing of data is whether your practices respect a person's or group's privacy.

Tell the students Data governance refers to a set of principles and practices that ensures formal and consistent management of important data assets or information.

Discuss with the students Veracity is the obligation to tell the truth and is more specific than other principles such as beneficence or fidelity with which it is commonly associated.

Tell the students improper data analysis is an ethical issue because it can result in publishing false or misleading conclusions.

Tell the students data quality meets six dimensions: accuracy, completeness, uniqueness, validity, consistency and timeliness.

Tell the students data accuracy means to make the data error-free, so that it can be used as a reliable source of information.

Discuss with the students completeness of data quality that ensures there is no data missing from your data set and all possible data that is required is present.

Tell the students uniqueness of data quality framework which specifies that there is no duplication of data. It is measured against all records within a data set.

Tell the students validity specifies that the values are according to the requirement.

Tell the students timeliness refers to the availability of data for analysis and for making decisions.

Share with the students Private data of the customers and their identity must remain private and confidential.

Tell the students Some of the causes of data quality problems: Manual data entry errors, OCR errors, Lack of complete information, Ambiguous data, Duplicate data, Data transformation errors, Poor organisation, Too much data, Inconsistent data.

Tell the students that goal of ethical guidelines is to help data analysts make decisions ethically.

Tell the students professional integrity means to behave in accordance with ethical principles and act in good faith, intellectual honesty and fairness.

Share with the students data integrity can be defined as the reliability and trustworthiness of data. It specifies the state of your data that is valid or invalid or the process of ensuring and preserving the validity and accuracy of data.

Ask the students data governance framework can be defined as the process of building a model for managing enterprise data. It empowers an organisation to define guidelines and rules on data management.

Ask the students about the goals of data governance.

Extension

Ask the students some oral questions based on this chapter.

Q. How does data ethics help prevent mishandling and unethical use of data?
Q. Discuss the primary ethical concern regarding data sharing and privacy.
Q. Why is uniqueness important in ensuring data quality?
Q. Define validity in the context of data quality and provide examples.
Q. Explain the ethical issues raised by big data analytics, particularly regarding biases.
Q. Discuss the goals of ethical guidelines in data analysis.
Q. Define professional integrity and accountability in the context of data analysis.
Q. Why is it crucial for organisations to establish data governance policies and rules?

Q. Discuss the importance of governing data for organisations.

Q. How does data governance help minimise business risks and reduce operational costs?

Encourage the students to walk through the chapter and ask them to explain any one topic from the chapter.

Evaluation

AAfter explaining the chapter, let the students do the exercises given on pages 294 to 298 in the main course book as **Exercise (Solved and Unsolved Questions)**.

Tell them to solve the critical and thinking skill developing exercises as **Higher Order Thinking Skills** given on pages 298.

Ask the students to practice the project in given in **Applied Project** section given on Page 298 in the main course book.

Suggested Activity

"There is a students' data list kept on Computer Lab's system, which is accessible by some students without authorisation."

Ask the student how would you ensure data privacy and integrity?