SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution
Coimbatore-35

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT I – INTRODUCTION TO DATA SCIENCE

Empathy:

Understand the Challenge

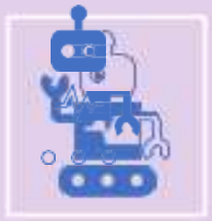
•Companies have huge amounts of raw data but struggle to make sense of it.

•Decisions are often based on **intuition and experience** rather than evidence.

•Users (managers, analysts, customers) need **actionable insights** rather than just numbers.

✓ This step focuses on *why data needs to be understood deeply*.

📌 *Example Insight:* Managers can't spot patterns in big sales data manually.



Data science combines math and statistics, specialized programming,

advanced analytics, Artificial Intelligence(AI) and machine learning with specific subject matter expertise to uncover actionable insights hidden in an organization's data.

Data Science uses advanced analytical theory and various methods such as time series analysis for predicting future. From historical data, Instead of knowing how many products sold in previous quarter, data science helps in forecasting future product sales and revenue more accurately.

Data science is devoted to the extraction of clean information from raw data to form actionable insights.

Define:

📌 ***Frame the Problem Clearly***

Problem Statement:

"How can we turn raw business data into meaningful insights that drive better decisions?"

•Identify key data challenges: incomplete data, inconsistent formats, too much volume, unstructured text, etc.

✓ This defines the real *need for data science* — transforming data into value.

LIFE CYCLE OF DATA SCIENCE:

1

Capture : Data acquisition, data entry, signal reception and data extraction.

2

Maintain : Data warehousing, data cleansing, data staging, data processing and data architecture.

3

Process : Data mining, clustering and classification, data modeling and data summarization.

4

Analyze : Data reporting, data visualization, business intelligence and decision making.

5

Communicate : Exploratory and confirmatory analysis, predictive analysis, regression, text mining and qualitative analysis.

NEED FOR DATA SCIENCE

Data Science is about data gathering, analysis and decision-making.

Data Science is about finding patterns in data, through analysis, and make future predictions.

By using Data Science, companies can make:

- Better decisions (should we choose A or B)
- Predictive analysis (what will happen next?)
- Pattern discoveries (find pattern, or maybe hidden information in the data)

Ideate:

- Generating Data-Driven Solutions Once the problem is defined, multiple ideas and analytical approaches are explored.
- This includes selecting suitable techniques such as data visualization, statistical analysis, machine learning algorithms, and predictive models.
- Data Science enables innovative solutions by combining domain knowledge with algorithms to extract meaningful insights and generate alternative solutions.

- Handling Massive & Complex Data

Traditional data processing and database systems struggle to cope with the volume, variety, and velocity of big data generated today. Data Science provides techniques and tools to manage, process, and extract value from such large and diverse datasets.

- Extracting Meaningful Insights

The primary purpose of data science is to derive actionable insights from raw data. It combines statistics, computing, and domain knowledge to uncover patterns, relationships, and trends that would otherwise remain hidden

- Support for Decision-Making

Insights from data science help organisations make informed decisions, improve operational efficiency, reduce risks, and predict future trends rather than relying on intuition or hindsight.

- Interdisciplinary Approach

Data Science draws from mathematics, statistics, computer science, and domain expertise to address complex real-world problems—making it essential where simple statistics or analytics can’t provide deep solutions.

- Automation & Predictive Ability

By applying machine learning and algorithmic models, data science enables automation of tasks and predictive modelling, helping to forecast outcomes (e.g., customer behaviour, demand trends) and drive strategic actions.

- **Value Creation Across Industries**

  Almost every domain—healthcare, finance, marketing, logistics, government—benefits from data science. It transforms data into **competitive advantage**, personalization (e.g., recommendations), risk detection (e.g., fraud), and optimized operations.

- **Big Data Era Demands**

  Organizations collect data continuously from sensors, transactions, social media, and IoT devices; data science is essential to **process and interpret this continuously expanding data pool**.

- **Bridging the Gap Between Data and Action**

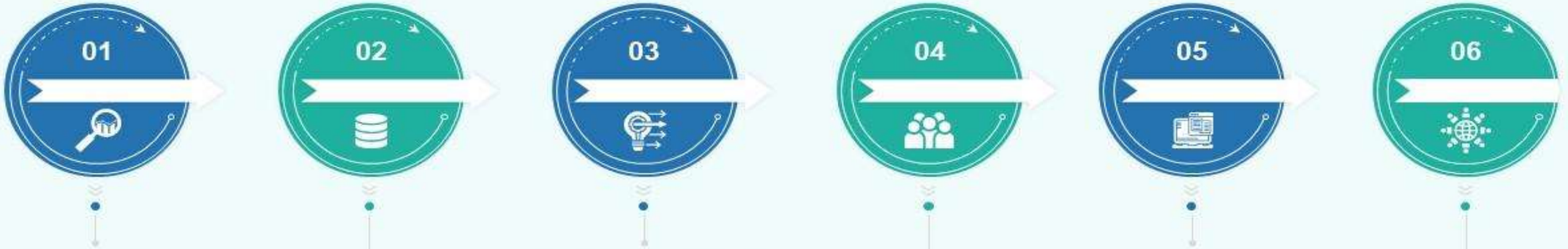
  It is not enough to have data; data science offers the **process and tools to clean, integrate, analyze, model, and present data** effectively so it can be turned into knowledge and action.

Prototyping:

- In this stage, ideas are converted into working models or prototypes.
- Data scientists build predictive models, dashboards, or recommendation systems using tools like Python, R, and ML libraries.
- Prototyping allows organizations to quickly test solutions on real data, reduce uncertainty, and refine models before full-scale deployment.

Data Science Depends on Six Critical Elements

This template depicts that data science depends on six critical elements, including analytic opportunities, data, analytic techniques, people, technology, and culture, also known as ADAPT+C.



Analytic Opportunities

Considers new and existing use cases to apply data science to improve organization mission and operations

Data

Use new and existing data sets and better manage to govern data in support of data science projects

Analytic Techniques

Considers the analytic tradecraft and techniques to be applied to generate insights from data

People

Develop the set of human capital programs, required to set a talented and capable team of data science practitioners

Technology

Find the optimal ways to use existing and new technologies including applications, data platforms, and infrastructure to perform data science projects

Culture

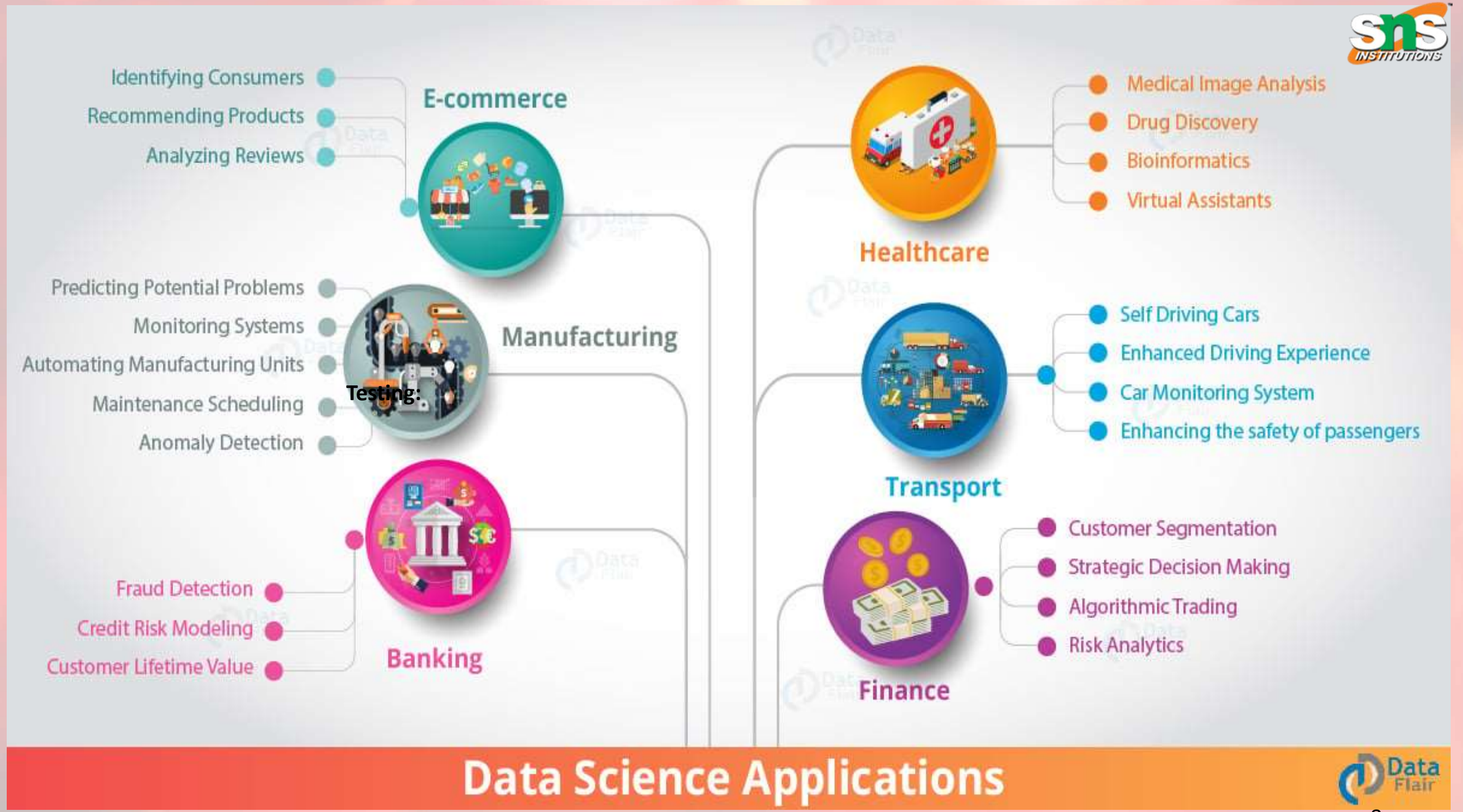
Find the optimal ways to use existing and new technologies including applications, data platforms, and infrastructure to perform data science projects

This slide is 100% editable. Adapt it to your needs and capture your audience's attention.

Where is Data Science Needed?

Testing:

- For route planning: To discover the best routes to ship
- To foresee delays for flight/ship/train etc. (through predictive analysis)
- To create promotional offers
- To find the best-suited time to deliver goods
- To forecast the next year’s revenue for a company
- To analyse the health benefits of training
- To predict who will win elections


Identifying Consumers
Recommending Products
Analyzing Reviews
E-commerce
Predicting Potential Problems
Monitoring Systems
Automating Manufacturing Units
Maintenance Scheduling
Anomaly Detection
Manufacturing
Fraud Detection
Credit Risk Modeling
Customer Lifetime Value
Banking
Medical Image Analysis
Drug Discovery
Bioinformatics
Virtual Assistants
Healthcare
Self Driving Cars
Enhanced Driving Experience
Car Monitoring System
Enhancing the safety of passengers
Transport
Customer Segmentation
Strategic Decision Making
Algorithmic Trading
Risk Analytics
Finance
Data Science Applications

Testing:

What To Expect ? When Pursuing Data Science as a Career?

Some potential jobs for you, once you gain data science skills, include:

- Analytics Manager
- Business Intelligence Analyst
- Clinical Data Manager
- Data Analyst
- Data Warehousing Specialist
- Database Administrator

Some of the top industries that employed data scientists in 2021 were:

- Scientific research and development services
- Computer systems design and related services
- Management of companies and enterprises
- Management, scientific, and technical consulting services
- Insurance carriers and related activities