@learneverythingai

6 TYPES OF DATA FREQUENTLY USED IN DATA SCIENCE

SHIVAM MODI
@learneverythingai

@learneverythingai

Structured Data

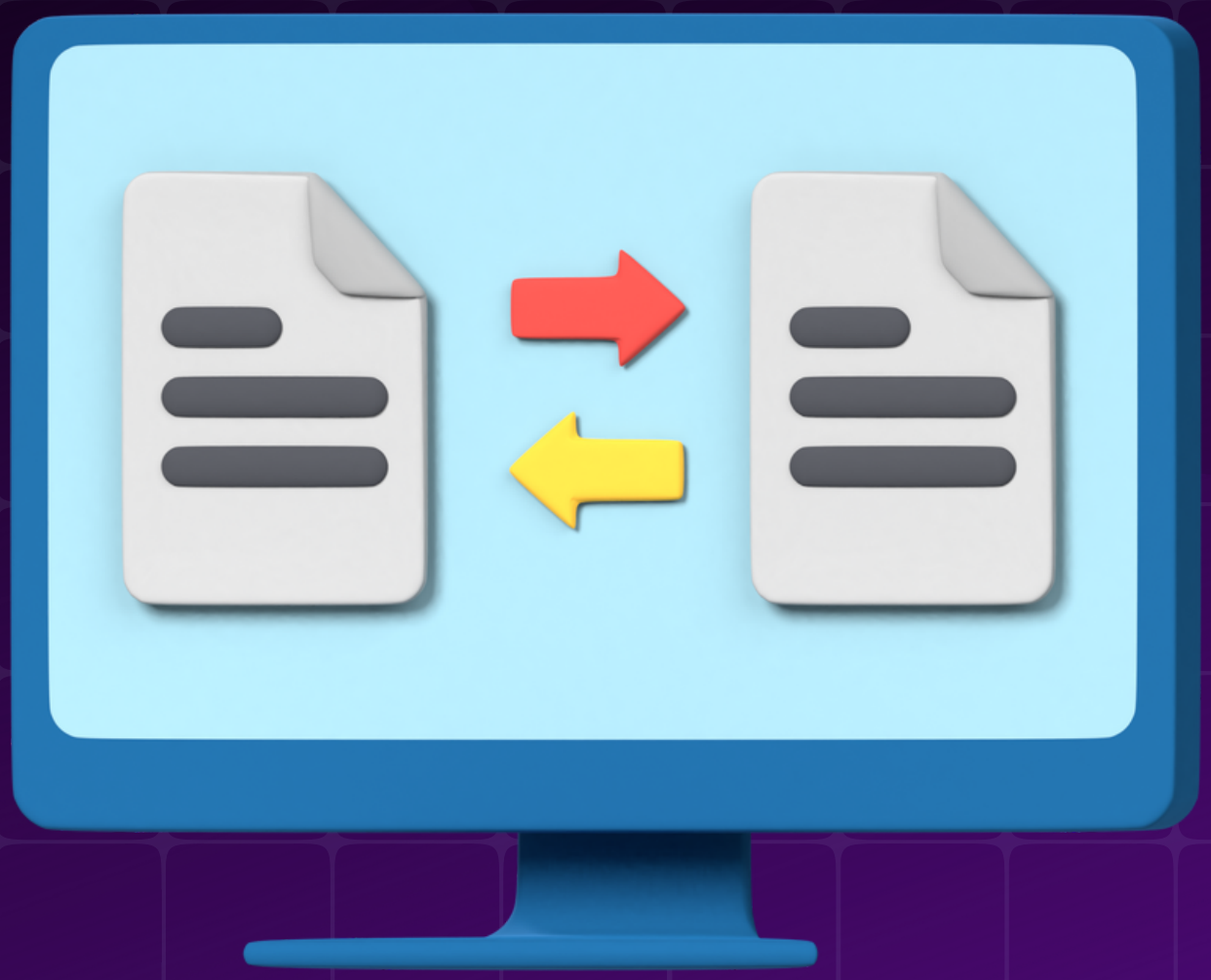
Structured data refers to data that is organized and formatted in a specific way, often stored in databases or spreadsheets. It has a predefined schema and is typically represented in tables with rows and columns. Structured data is easily searchable, analyzable, and can be processed using SQL queries.

@learneverythingai

Unstructured Data

Unstructured data does not have a predefined format and does not fit into traditional databases. It includes text documents, social media posts, images, audio, video, emails, and more. Analyzing unstructured data often requires techniques such as natural language processing (NLP), computer vision, or audio processing.

@learneverythingai



Semi-Structured Data

Semi-structured data lies between structured and unstructured data. It has some organizational properties but does not conform to a rigid schema. Examples include XML files, JSON documents, and log files. Semi-structured data requires specialized techniques for extraction, parsing, and analysis.

@learneverythingai

Time Series Data

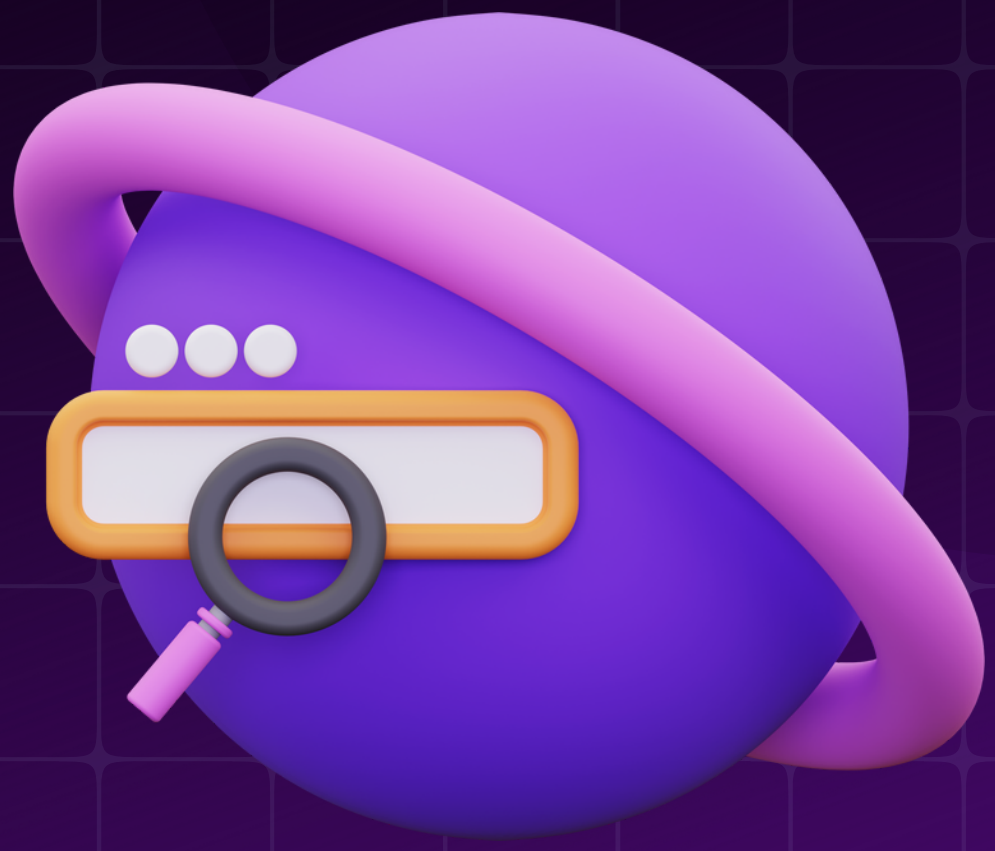
Time series data is a sequence of data points collected over time, typically at regular intervals. It is used to analyze trends, patterns, and seasonality. Time series data is commonly found in areas such as financial markets, weather forecasting, and IoT sensor data.

@learneverythingai

Geospatial Data

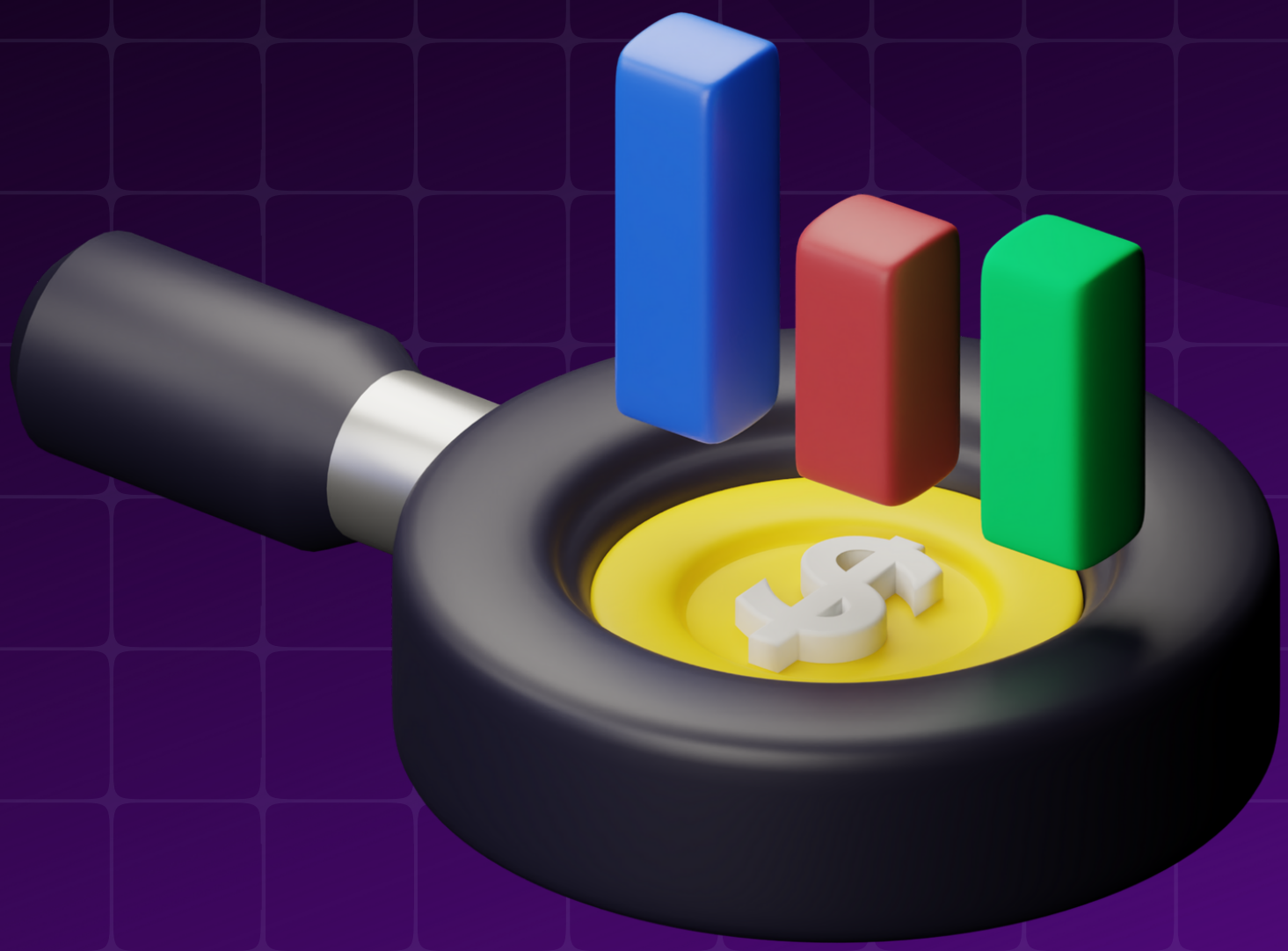
Geospatial data represents information associated with geographic locations. It includes coordinates, addresses, maps, and satellite imagery. Geospatial data is used for spatial analysis, location-based services, urban planning, and environmental monitoring.

@learneverythingai



Big Data

Big data refers to extremely large and complex datasets that cannot be easily managed or processed using traditional methods. It encompasses structured, unstructured, and semi-structured data, often generated at high volumes, velocities, and variety. Big data technologies, such as distributed computing frameworks and NoSQL databases, are used to store, process, and analyze such data.



These different types of data require specific techniques and tools for collection, storage, preprocessing, analysis, and visualization in the field of data science.

Like this Post?

- Follow Me
- Share with your friends
- Check out my previous posts

SHIVAM MODI

@learneverythingai

Share This

FOLLOW

SAVE THIS

www.learneverythingai.com