

**Sources and nature of data**

**classification of data (structured, semi-structured, unstructured)**

## Sources and nature of data

### classification of data (structured, semi-structured, unstructured)

#### **Structured data:**

- A phone directory with names, addresses, and phone numbers organized in a table
- A spreadsheet with sales data, including product names, prices, and dates
- A database of student information, including name, ID number, and grades

## Sources and nature of data

### classification of data (structured, semi-structured, unstructured)

#### **Semi-structured data:**

- An email message with a sender, recipient, subject, and body
- An XML file with data organised using tags, but the data within the tags may be unstructured
- A customer review on an e-commerce site, which has a structure (username, date, rating, and review text), but the review text may be unstructured

## Sources and nature of data

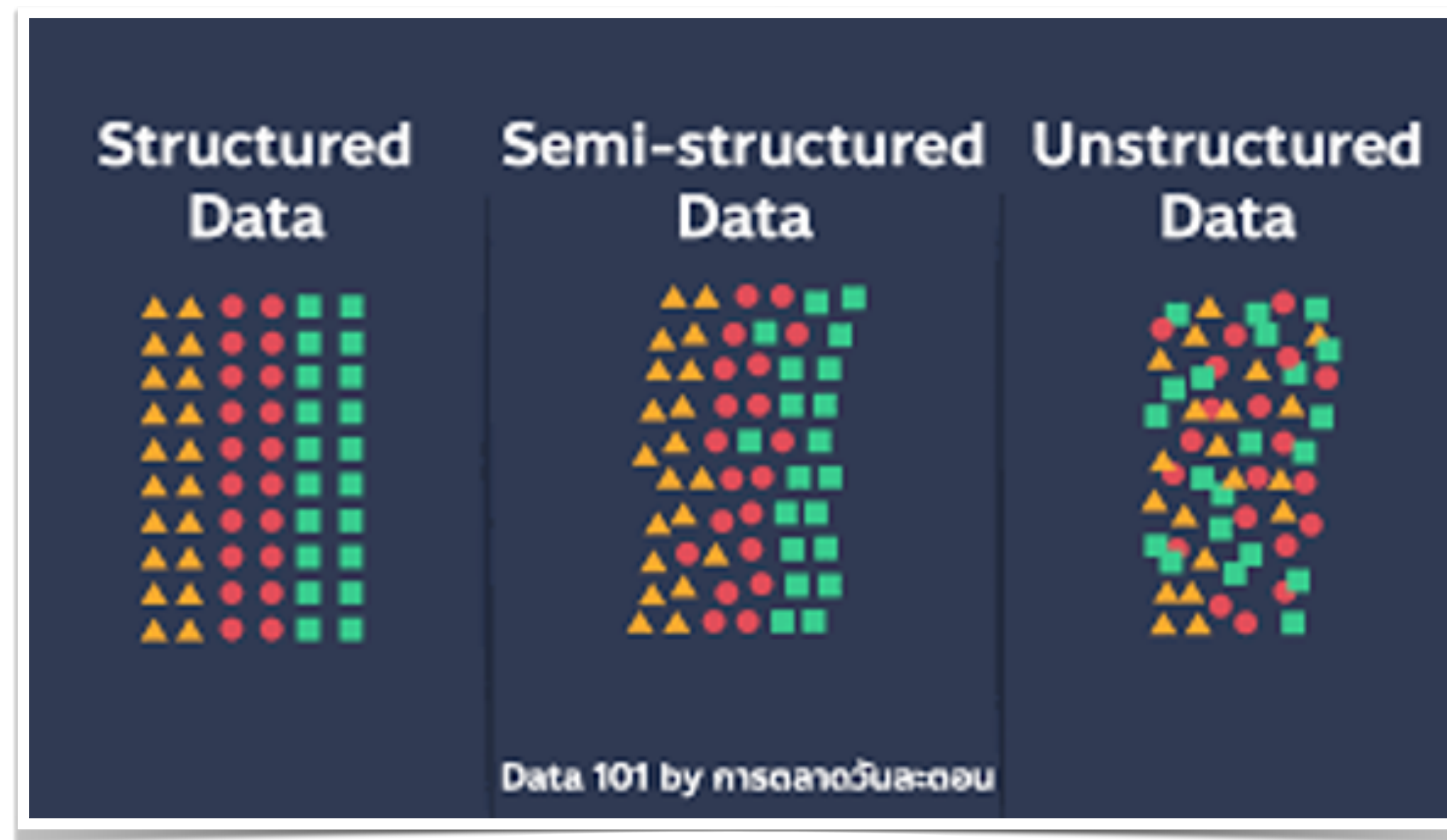
### classification of data (structured, semi-structured, unstructured)

#### **Unstructured data:**

- A handwritten note
- An audio recording of a conversation
- A photo or image
- A free-form text entry, such as a blog post or social media update

## Sources and nature of data

classification of data (structured, semi-structured, unstructured)



# Characteristics of Data

## Characteristics of Data

- Volume
- Velocity
- Variety
- Veracity
- Value
- Variability
- Complexity

## Types of data attributes

- Nominal
- Ordinal
- Interval
- Ratio



# Big Data

# Big Data

Big Data refers to the large and complex data sets that cannot be processed or analyzed using traditional data processing and storage technologies



# Big Data

It is characterized by three main attributes:

- Volume
- Velocity
- Variety

# Big Data

## Example

- Social media platforms, such as Facebook, Twitter, and Instagram.
- This includes user-generated content, such as posts, comments, photos, and videos, as well as metadata, such as location, time, and device information

# Big Data platforms

## Big Data platforms

