**2.1.** Data Objects and Attribute Types

- Data sets are made up of data objects.

- A data object represents an entity

- If the data objects are stored in a database, they are *data tuples*

2.1.1. **What Is an Attribute?**

- An attribute is a data field, representing a characteristic or feature of a data object

- The nouns *attribute*, *dimension*, *feature*, and *variable* are often used interchangeably in the literature

- A set of attributes used to describe a given object is called an *attribute vector* (or *feature vector*).

- The distribution of data involving one attribute (or variable) is called *univariate*. A *bivariate* distribution involves two attributes, and so on.

**2.1.6. Discrete versus Continuous Attributes**

- Attributes are organized into nominal, binary, ordinal, and numeric types. The categories are not mutually exclusive.

- There are 2 types of attributes: discrete and continuous

- A discrete attribute has a countable finite or infinite set of values, which can be represented in integer or binary form

+ For example: hair color, number of smokers, number of testers, number of laptops

- The terms numeric attribute and continuous attribute are often used interchangeably in the literature. Continuous properties are often represented as floating-point variables.

+ For example: 3.25kg, 1.32m, 7.25inches

**2.1. Đối tượng dữ liệu và kiểu thuộc tính**

- Tập dữ liệu được tạo thành từ các đối tượng dữ liệu.

- Đối tượng dữ liệu đại diện cho một thực thể

- Nếu các đối tượng dữ liệu được lưu trữ trong cơ sở dữ liệu thì chúng là các bộ dữ liệu

**2.1.1. Thuộc tính là gì?**

- Thuộc tính là một trường dữ liệu, biểu thị một đặc tính, tính năng của đối tượng dữ liệu

- Các danh từ **thuộc tính**, **chiều hướng**, **đặc điểm**, **biến số** thường được sử dụng thay thế cho nhau trong văn học

- Tập hợp các thuộc tính dùng để mô tả một đối tượng nhất định được gọi là vectơ thuộc tính (hoặc vectơ đặc trưng).

- Phân phối dữ liệu liên quan đến một thuộc tính (hoặc biến) được gọi là đơn biến. Phân phối hai biến bao gồm hai thuộc tính, v.v.

**2.1.6. Thuộc tính rời rạc và liên tục**

- Các thuộc tính được tổ chức thành các kiểu **danh nghĩa**, **nhị phân, thứ tự** và **số**. Các loại không loại trừ lẫn nhau.

- Có 2 loại thuộc tính: rời rạc và liên tục

- Một thuộc tính rời rạc có một tập hợp các giá trị hữu hạn hoặc vô hạn đếm được, có thể được biểu diễn dưới dạng số nguyên hoặc dạng nhị phân

+ Ví dụ: màu tóc, số người hút thuốc, số người xét nghiệm, số lượng laptop

- Thuật ngữ thuộc tính số và thuộc tính liên tục thường được sử dụng thay thế cho nhau trong tài liệu. Các thuộc tính liên tục thường được biểu diễn dưới dạng các biến có dấu phẩy động.

+ Ví dụ: 3.25kg, 1.32m, 7.25inches