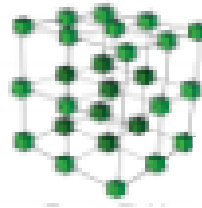


# Data Abstraction

name	rank	gender	year
Jacob	1	boy	2010
Isabella	1	girl	2010
Ethan	2	boy	2010
Sophia	2	girl	2010
Michael	3	boy	2010

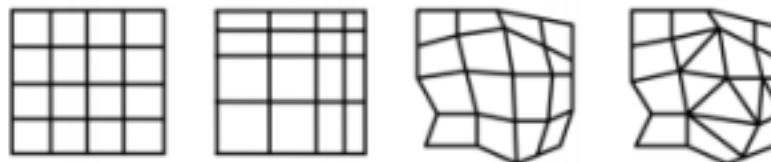
- Four basic dataset types:

- Tables
- Networks
- Fields
- Geometry



- Five basic datatypes

- Items
- Attributes
  - Categorical
  - Ordered
    - Ordinal
    - Quantitative
- Links
- Positions
- Grids



## Slide 1: Tipos e natureza dos dados

### 1. Data representation level (Nível de representação dos dados):

- **Qualitative (ou Categorical):** Dados que descrevem categorias ou qualidades (ex: cor dos olhos, tipo de animal, dias da semana).
- **Quantitative (ou Numeric):** Dados que representam quantidades ou valores numéricos (ex: altura, temperatura, idade).

### 2. Data nature (Natureza dos dados):

- **Continuous (Contínuos):** Podem assumir qualquer valor num intervalo (ex: 3.14, 4.1, -0.1).
- **Discrete (Discretos):** Assumem apenas certos valores distintos (ex: número de pessoas, número de carros).

Frase importante no slide:

"Computer data are discrete but the phenomena may be continuous"

Quer dizer que mesmo que o fenómeno real seja contínuo (ex: temperatura), quando o representamos no computador, temos que discretizá-lo (ex: com casas decimais limitadas).

### 3. Measuring scale (Escala de medição):

As escalas servem para definir como interpretamos os dados:

- **Nominal:** Categorias sem ordem (ex: género, cor dos olhos).
- **Ordinal:** Categorias com ordem (ex: pequeno, médio, grande).
- **Interval:** Dados numéricos com diferença mensurável, mas sem zero absoluto (ex: temperatura em Celsius).
- **Ratio:** Dados numéricos com zero absoluto e com todas as propriedades anteriores (ex: peso, idade, altura).

Nominal e Ordinal são categóricos.

Interval e Ratio são quantitativos.

## Slide 2: Abstração de dados

### 1. Four basic dataset types (Quatro tipos básicos de conjuntos de dados):

- **Tables (Tabelas):** Dados estruturados com linhas e colunas.
- **Networks (Redes):** Representam ligações (ex: redes sociais).
- **Fields (Campos):** Dados distribuídos num espaço contínuo (ex: temperatura num mapa).
- **Geometry (Geometria):** Formas e estruturas espaciais (ex: coordenadas, polígonos).

### 2. Five basic datatypes (Cinco tipos básicos de dados):

- **Items:** Elementos individuais (ex: uma pessoa, um nó na rede).
- **Attributes:** Propriedades associadas aos items (ex: idade, nome, cor).
- **Links:** Relações entre items (ex: amizades, ligações na rede).
- **Positions:** Localizações no espaço (ex: coordenadas GPS).
- **Grids:** Estruturas em malha, úteis para representar campos contínuos (ex: imagem, mapa de calor).

Atributos podem ser:

- **Categóricos:** Ex: "rapaz", "rapariga".
- **Ordenados (Ordered):**
  - **Ordinal:** Ex: classificação (1º, 2º, 3º).
  - **Quantitativo:** Ex: idade, altura.