Chapter 2.1

Data model

A data model is a notation for describing data information

3 parts:

- Structure of data
- Operations on the data.
- Constraints on the data
  - Limitations on what can be stored (e.g. data types).
  - Domain specific (eg. student id)

Relational model

- Structure.

  Based on relations / tables

title	year	genre
Terminator	1984	Action
Alien	1979	Horror

- Operations
  - Based on Relational Algebra (implemented in SQL)
- Constraints
  - Year between 1984 and 2015
  - Genre in a given set

Much more later.

Basics of Relational Model (2.2)

- Data represented as a 2-dim. table called relation
- A relation represents a set of objects/entities/elements and their relations.

Movies:

Attributes

title	year	genre
Terminator	1984	Action
Alien	1979	Horror

Tuples

No order and No duplicates

Schema

- Represents the characteristics of a relation: name and attributes

  Movies(title, year, genre)

DB Schema

- Set of schemas of all relations in the DB.

Some properties of relations

- every tuple is unique
- every attribute
  - must have a domain
  - must be atomic (no lists, sets, etc).

Domain

- Set of potential values an attribute can take.

Instance

The "current" set of "values" in the relation/db.

Keys

- One of the most important types of constraints.
- A key is a set of attributes of a relation R s.t. two tuples in R cannot have the same values in these attributes.
- Allows to uniquely identify a tuple in a relation
- What is a good key for the Movies relation?
- Underscoring the attr. in Schema:
  ⇒ part of the key

Movies (title, year, genre)

The key documented in the schema is know as the primary key of the relation.

Artificial primary keys

- Attributes created specifically for the purpose of uniquely identifying a tuple in a relation.