

Bases de Datos

Clase 2: Modelo Relacional - Álgebra Relacional

Hasta ahora

- Todo el mundo necesita manejar datos
- Salvo que queramos programar cosas que no tienen que ver con nuestro problema, conviene utilizar un DBMS
- Arquitectura de capas:
 - Usuarios finales ven modelo lógico
 - Sistema ejecuta las acciones
 - Usuario no necesita saber cómo el sistema ejecuta las acciones

DBMS

DBMS relacionales comerciales

- IBM DB2
- Microsoft SQL Server
- Oracle

DBMS

DBMS relacionales Open Source

- PostgreSQL - Es la que usaremos nosotros
- MySQL - Usada ampliamente en ambientes de producción
- SQLite - Base de datos pequeña, usada generalmente en contextos de apps móviles

DBMS

Otros DBMS

- Neo4J (Grafos)
- MongoDB (Documentos)
- Cassandra (Key Value - Column Store)
- Apache Jena (RDF)
- Redis (In memory Store)
- Base (Column Store)
- Titan DB (Grafos)
- ...

Modelos de datos

Modelo de Datos

- Un Modelo es una notación para escribir datos
- En este curso se verán en detalles dos modelos:
 - Relacional
 - Semiestructurados (key-value, XML, grafos)
- Modelo relacional es el modelo más usado en ambiente de producción, pero ha existido la necesidad de utilizar cada vez más modelos semiestructurados.

Modelo Relacional

Busca almacenar datos en tablas:

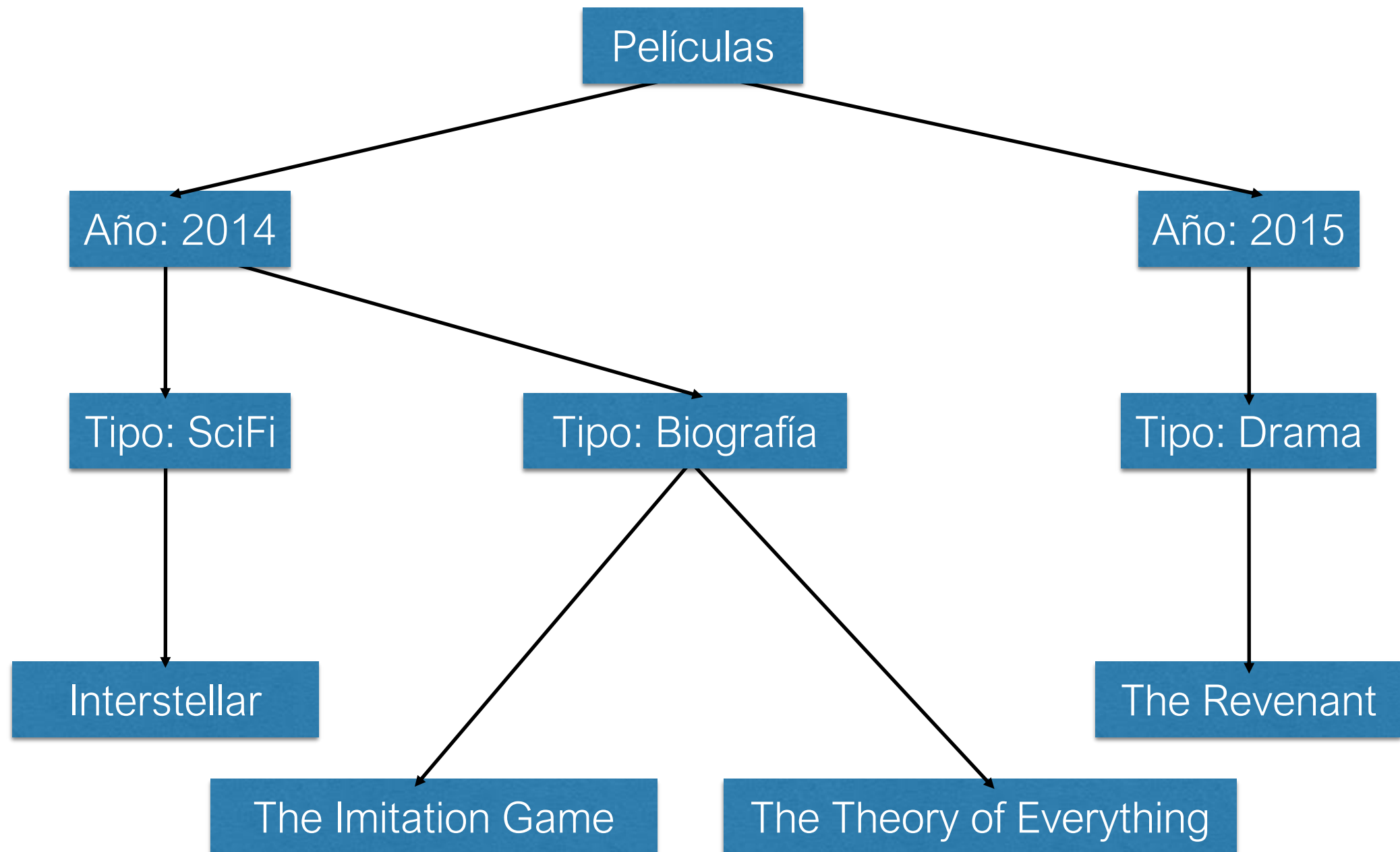
| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|--------------------------|------|-----------|---------------------|
| 1 | Interstellar | 2014 | Fantasía | 8.6 |
| 2 | The Revenant | 2015 | Drama | 8.1 |
| 3 | The Imitation Game | 2014 | Biografía | 8.1 |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 |

Modelo Relacional

- Aunque parecen arreglos o listas (de Python por ejemplo) existen muchas diferencias
- Generalmente, se asume que están en Disco y no en Memoria
- No podemos hacer todo lo que queramos con esto
- Vamos a ver como manejar estos datos:
 - SQL
 - Álgebra Relacional

Datos Semiestructurados

Estructura Jerárquica



Datos Semiestructurados

XML

```
<Películas>
  <Año valor="2014">
    <Tipo valor="Biografía">
      <Película nombre="The Imitation Game" calificación="8.1">
      </Película>
      <Película nombre="The Theory of Everything" calificación="7.7">
      </Película>
    </Tipo>
    <Tipo valor="SciFi">
      <Película nombre="Interstellar" calificación="8.6">
      </Película>
    </Tipo>
  </Año>
  <Año valor="2015">
    <Tipo valor="Drama">
      <Película nombre="The Revenant" calificación="8.1">
      </Película>
    </Tipo>
  </Año>
</Películas>
```

Datos Semiestructurados

Key - Value

```
{
  "2014": {
    "Biografía": [
      { "nombre": "The Imitation Game", "calificación": 8.1 },
      { "nombre": "The Theory of Everything", "calificación": 7.7 }
    ],
    "SciFi": [
      { "nombre": "Interstellar", "calificación": 8.6 }
    ]
  },
  "2015": {
    "Drama": [
      { "nombre": "The Revenant", "calificación": 8.1 }
    ]
  }
}
```

Comparación

Ambos:

- Proveen solución para almacenar datos
- Son versátiles para modelar
- Ambos tienen lenguaje de consultas

Pero:

- Modelo relacional está definido por un **esquema**
- XML es más flexible, no está separado por un esquema

El modelo relacional al ser menos flexible es más simple pero también limitado

Otros Modelos

- Bases de Datos orientados a objetos
- Modelo Objeto - Relacional
- Bases de Datos de Grafos

Tenemos que aprender las diferencias y
cuando usar qué!

Modelo Relacional

Modelo Relacional

Los datos se almacenan como tablas:

Películas

| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|--------------------------|------|-----------|---------------------|
| 1 | Interstellar | 2014 | Fantasía | 8.6 |
| 2 | The Revenant | 2015 | Drama | 8.1 |
| 3 | The Imitation Game | 2014 | Biografía | 8.1 |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 |

Distinguimos:

- **Relaciones:** a cada tabla le llamamos relación
- **Atributos:** son las columnas de la relación
- **Tuplas:** son las filas de la relación

Modelo Relacional

Películas

| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|--------------------------|------|-----------|---------------------|
| 1 | Interstellar | 2014 | Fantasía | 8.6 |
| 2 | The Revenant | 2015 | Drama | 8.1 |
| 3 | The Imitation Game | 2014 | Biografía | 8.1 |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 |

- En este caso tenemos la relación **Películas**
- Los atributos de la relación **Películas** son **ID_Película**, **Nombre_Película**, **Año**, **Categoría** y **Calificación**.

Modelo Relacional

Esquema

Para denominar relaciones escribimos su nombre y luego sus atributos entre paréntesis:

Películas(id, nombre, año, categoría, calificación)

Modelo Relacional

Esquema

Un esquema es un conjunto de relaciones con sus atributos:

Películas(id, nombre, año, categoría, calificación)

Actor(id, nombre, edad)

Actuó_en(id_actor, id_película)

Modelo Relacional

Dominio

En la práctica, asumimos que cada atributo tiene un dominio (float, integer, string, date, ...)

Películas(id: **int**, nombre: **string**, año: **int**,
categoría: **string**, calificación: **float**)

Modelo Relacional

Instancia

Una **instancia** de un **esquema** es un conjunto de tuplas para cada relación del esquema

Esto es un esquema:

| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|-----------------|-----|-----------|---------------------|
|-------------|-----------------|-----|-----------|---------------------|

Modelo Relacional

Instancia

Una **instancia** de un **esquema** es un conjunto de tuplas para cada relación del esquema

Esto es una instancia:

| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|--------------------------|------|-----------|---------------------|
| 1 | Interstellar | 2014 | Fantasía | 8.6 |
| 2 | The Revenant | 2015 | Drama | 8.1 |
| 3 | The Imitation Game | 2014 | Biografía | 8.1 |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 |

Modelo Relacional

Restricciones de integridad

Son restricciones que imponemos a un esquema que todas las instancias deben satisfacer

La restricción más importante son las **llaves**

Un conjunto de atributos forma una **llave** en una **relación** si no permitimos que existan dos tuplas para esa relación con los mismos valores en todos los atributos de la llave, y no hay un subconjunto de esos atributos que cumpla esa condición.

Modelo Relacional

Ejemplo: Llaves

¿Cuál es la llave?

| ID Película | Nombre Película | Año | Categoría | Calificación (IMDB) |
|-------------|--------------------------|------|-----------|---------------------|
| 1 | Interstellar | 2014 | Fantasía | 8.6 |
| 2 | Batman | 2005 | Acción | 8.3 |
| 3 | The Imitation Game | 2014 | Biografía | 8.1 |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 |
| 5 | Batman | 1995 | Acción | 5.4 |

Nota:

- Batman Begins (2005)
- Batman Forever (1995)

Modelo Relacional

Llaves

Cuando escribimos las relaciones subrayamos las llaves

Películas(id, nombre, año, categoría, calificación)

Cómo consultar bases de datos

Ejemplos de Consultas

actores

| id | nombre | edad |
|-----|---------------------|------|
| 1 | Leonardo DiCaprio | 41 |
| 2 | Matthew McConaughey | 46 |
| 3 | Daniel Radcliffe | 27 |
| 4 | Jessica Chastain | 39 |
| ... | ... | ... |

actuo_en

| id_actor | id_pelicula |
|----------|-------------|
| 1 | 2 |
| 2 | 1 |
| 4 | 1 |
| 3 | 3 |
| 1 | 5 |
| ... | ... |

peliculas

| id | nombre | año | categoria | calificacion | director |
|-----|--------------------------|------|-----------|--------------|-------------|
| 1 | Interstellar | 2014 | SciFi | 8.6 | C. Nolan |
| 2 | The Revenant | 2015 | Drama | 8.1 | A. Iñárritu |
| 3 | Harry Potter | 2011 | Fantasía | 8.1 | D. Yates |
| 4 | The Theory of Everything | 2014 | Biografía | 7.7 | J. Marsh |
| 5 | Inception | 2010 | Adventure | 8.8 | C. Nolan |
| ... | ... | ... | ... | ... | ... |

Ejemplos de Consultas

Ejemplo 1

Liste el nombre de todos los actores

| nombre |
|---------------------|
| Leonardo DiCaprio |
| Matthew McConaughey |
| Daniel Radcliffe |
| Jessica Chastain |

Ejemplos de Consultas

Ejemplo 2

Liste el nombre y la calificación de todas las películas

| nombre | calificacion |
|--------------------------|--------------|
| Interstellar | 8.6 |
| The Revenant | 8.1 |
| Harry Potter | 8.1 |
| The Theory of Everything | 7.7 |
| Inception | 8.8 |

Ejemplos de Consultas

Ejemplo 3

Liste el nombre y la calificación de todas las películas con calificación inferior a 8.5

| nombre | calificacion |
|--------------------------|--------------|
| The Revenant | 8.1 |
| Harry Potter | 8.1 |
| The Theory of Everything | 7.7 |

Ejemplos de Consultas

Ejemplo 4

Liste el todas las películas de Nolan

| id | nombre | año | categoria | calificacion | director |
|----|--------------|------|-----------|--------------|----------|
| 1 | Interstellar | 2014 | SciFi | 8.6 | C. Nolan |
| 5 | Inception | 2010 | Adventure | 8.8 | C. Nolan |

Ejemplos de Consultas

Ejemplo 5

Liste todos los id de los actores de la película “Interstellar”

| id |
|----|
| 2 |
| 4 |

Ejemplos de Consultas

Ejemplo 6

Liste cada actor junto a todas las películas en las que ha actuado

| id | nombre | nombre_pelicula |
|----|---------------------|-----------------|
| 1 | Leonardo DiCaprio | The Revenant |
| 1 | Leonardo DiCaprio | Inception |
| 2 | Matthew McConaughey | Interstellar |
| 3 | Daniel Radcliffe | Harry Potter |
| 4 | Jessica Chastain | Interstellar |

Ejemplos de Consultas

Ejemplo 7

Liste todas las películas en que actúe Leonardo DiCaprio y que sean dirigidas por C. Nolan

| nombre |
|--------|
|--------|

| |
|-----------|
| Inception |
|-----------|

Ejemplos de Consultas

Ejemplo 8

Liste todas las películas y la calificación en que actúe Leonardo DiCaprio o que sean dirigidas por C. Nolan

| nombre | calificacion |
|--------------|--------------|
| Interstellar | 8.6 |
| The Revenant | 8.1 |
| Inception | 8.8 |

Ejemplos de Consultas

Ejemplo 9

Liste el nombre de todos los actores y directores

| nombre |
|---------------------|
| Leonardo DiCaprio |
| Matthew McConaughey |
| Daniel Radcliffe |
| Jessica Chastain |
| C. Nolan |
| A. Iñárritu |
| D. Yates |
| J. Marsh |

Ejemplos de Consultas

Ejemplo 10

Liste el nombre de todos los actores dirigidos por C. Nolan y A. Iñárritu

| id | nombre | edad |
|----|-------------------|------|
| 1 | Leonardo DiCaprio | 41 |

Ejemplos de Consultas

¿Qué podemos concluir?

- Los resultados de las consultas también son tablas
- Parecen haber operaciones en común

Álgebra relacional de selección, proyección y unión

- Lenguaje teórico
- Posee un conjunto de operadores que como input toman tablas, y como output devuelven tablas

$$\pi, \sigma, \cup, \times$$

- Durante el curso veremos cómo este lenguaje forma los cimientos de todos los lenguajes de consulta