Introduction to Databases

July 2022

Roadmap

- Conceptual Overview of Databases

- Big picture of database design

- Learn basic SQL queries with Harry Potter Database

- How data analysts use databases: Steven Zhang

UBC
MFRE

Setup

- Install DB Browser for SQLite
    - [Download](#) standard installer for 64-bit Windows

Learning Objectives

- Appreciate the value of databases

- Differentiate the use of a spreadsheet and a database

- Describe the steps involved in database design

- Write SQL queries to utilize data

What is a database?

- A database is a collection of data/information that is organized in a certain structure.

- A database management system (DBMS) is a software designed to manage or access the database.

Types of databases

- **Relational databases**
  - Data is stored in tables (rows/cols) and are related to each other.
  - Popular products: MySQL, PostgreSQL, Oracle, **SQLite,** Airtable
  - Main programming language: SQL

- Non-relational database
  - Data can be stored using key-value pairs, JSON, graphs, etc.
  - Popular products: MongoDB, Redis

UBC
MFRE

Question

- **What are some databases you commonly use or interact with?**

UBC
MFRE

Examples of databases

- Netflix

- Bank

- Amazon

- Legends Haul

Database vs. spreadsheet

- One of the main important technical difference between a database and a spreadsheet is how they store data.

- In a spreadsheet, data is stored in cells. You can format, edit, and manipulate data within that cell.

- In a database, cells contain data that come from external tables. If you change a record in one place, it will make the change in all other places that references that record.

Why use a database?

- Size of data

- Multiple users updating data

- Redundancy of data

- Data accessibility and speed

Limits of Excel

Fetzer and Graeber (2021)

"Conservative estimates suggest that the failure of timely contact tracing due to the data glitch is associated with more than 125,000 additional infections and over 1,500 additional COVID-19-related deaths. Our findings provide strong quasi-experimental evidence for the effectiveness of contact tracing."

How to decide?

- What kind of data do you have?

- How do you want to use that data?

- How important is data integrity?

- Do you need deep search capabilities?

- How many people need access?

UBC
MFRE

Roadmap

- Conceptual Overview of Databases
- Big picture of database design
- Learn basic SQL queries with Harry Potter Database
- How data analysts use databases: Steven Zhang

Relational databases

- Data is organized in a tabular structure
    - Rows are called "**records**" or "**tuples**"
    - Columns are called "**fields**" or "**attributes**"
    - Data is linked or related to each other using **keys**

- Structured Query Language (SQL) is a standard programming language to define, manipulate, and query relational databases
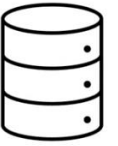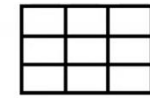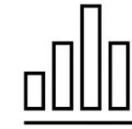
Database design

- Requirement Analysis
  - Discuss with users what they want from the database

- Conceptual Database Design
  - Design a high-level overview and description of your database
  - Can use diagrams such as the Entity-Relationship model

- Logical Database Design
  - Operationalize conceptual design in a database management system

Database design

- **Entities** are real world objects that you want to include in your database.

- **Attributes** are characteristics that describe the entity.

- Designing a database can be an iterative process.

Database design

- A **schema** represents the blueprint/structure/configuration of your database

Customers(*cust_id*: integer, *last_name*: varchar, *first_name*: varchar, *email*: varchar, *address*: varchar, *mobile_num*: integer)

cust_id	first_name	last_name	email	address	mobile_num
1	Olivier	Ntwali	olivier@mfre.com	Vancouver, BC	6041234567
2	Krisha	Lim	krisha@mfre.com	Vancouver, BC	7781234567

Case study: Harry Potter

- How many movies have you watched?

Sample database

ActorName	Character	CharDescription	EpTitle	EpYear	AttributeDescription	AttributeValue	RotTomRate	MetacriticRate	BFCARate
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84

Issues

- Updating data

ActorName	Character	CharDescription	EpTitle	EpYear	AttributeDescription	AttributeValue	RotTomRate	MetacriticRate	BFCARate
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84

Issues

- Deleting data

ActorName	Character	CharDescription	EpTitle	EpYear	AttributeDescription	AttributeValue	RotTomRate	MetacriticRate	BFCARate
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84

actors

characters	
Character	**Description**
Harry Potter	Protagonist of the series
Albus Dumbledore	Headmasters of Hogwarts
Lord Voldemort	Antagonist of the series

actors	
ActorName	**Character**
Daniel Radcliffe	Harry Potter
Richard Harris	Albus Dumbledore
Michael Gambon	Albus Dumbledore
Richard Bremmer	Lord Voldemort
Christian Coulson	Lord Voldemort

What other entities should we have?

ActorName	Character	CharDescription	EpTitle	EpYear	AttributeDescription	AttributeValue	RotTomRate	MetacriticRate	BFCARate
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	WorldWide	974,755,371	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	WorldWide	878,979,634	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	WorldWide	796,688,549	91	82	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Richard Bremmer	Lord Voldemort	Antagonist of the series	Harry Potter and the Philosopher's Stone	2001	Budget	125,000,000	80	64	90
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Richard Harris	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Christian Coulson	Lord Voldemort	Antagonist of the series	Harry Potter and the Chamber of Secrets	2002	Budget	100,000,000	83	63	84
Daniel Radcliffe	Harry Potter	Protagonist of the series	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84
Michael Gambon	Albus Dumbledore	Headmasters of Hogwarts	Harry Potter and the Prisoner of Azkaban	2004	Budget	130,000,000	91	82	84

entities

characters

Character	Description
Harry Potter	Protagonist of the series
Albus Dumbledore	Headmasters of Hogwarts
Lord Voldemort	Antagonist of the series

actors

ActorName	Character
Daniel Radcliffe	Harry Potter
Richard Harris	Albus Dumbledore
Michael Gambon	Albus Dumbledore
Richard Bremmer	Lord Voldemort
Christian Coulson	Lord Voldemort

episodes

EpTitle	EpYear
Philosopher's Stone	2001
Chamber of Secrets	2002
Prisoner of Azkaban	2004

cast

EpTitle	Actor
Philosopher's Stone	Daniel Radcliffe
Philosopher's Stone	Richard Harris
Philosopher's Stone	Richard Bremmer
Chamber of Secrets	Daniel Radcliffe
Chamber of Secrets	Richard Harris
Chamber of Secrets	Christian Coulson
Prisoner of Azkaban	Daniel Radcliffe
Prisoner of Azkaban	Michael Gambon

financials

EpTitle	FinDescrip	FinValue
Philosopher's Stone	Worldwide	974,755,371
Philosopher's Stone	Budget	125,000,000
Chamber of Secrets	Worldwide	878,979,634
Chamber of Secrets	Budget	100,000,000
Prisoner of Azkaban	Worldwide	796,688,549
Prisoner of Azkaban	Budget	130,000,000

rating

EpTitle	RotTomRate
Philosopher's Stone	80
Chamber of Secrets	83
Prisoner of Azkaban	91

entities

episodes		
EpID	**EpTitle**	**EpYear**
1	Philosopher's Stone	2001
2	Chamber of Secrets	2002
3	Prisoner of Azkaban	2004

actors		
ActorID	**ActorName**	**Character**
1	Daniel Radcliffe	Harry Potter
2	Richard Harris	Albus Dumbledore
3	Michael Gambon	Albus Dumbledore
4	Richard Bremmer	Lord Voldemort
5	Christian Coulson	Lord Voldemort

entities

characters

Character	Description
Harry Potter	Protagonist of the series
Albus Dumbledore	Headmasters of Hogwarts
Lord Voldemort	Antagonist of the series

actors

ActorID	ActorName	Character
1	Daniel Radcliffe	Harry Potter
2	Richard Harris	Albus Dumbledore
3	Michael Gambon	Albus Dumbledore
4	Richard Bremmer	Lord Voldemort
5	Christian Coulson	Lord Voldemort

episodes

EpID	EpTitle	EpYear
1	Philosopher's Stone	2001
2	Chamber of Secrets	2002
3	Prisoner of Azkaban	2004

cast

EpID	ActorID
1	1
1	2
1	4
2	1
2	2
2	5
3	1
3	3

financials

EpID	FinDescrip	FinValue
1	Worldwide	974,755,371
1	Budget	125,000,000
2	Worldwide	878,979,634
2	Budget	100,000,000
3	Worldwide	796,688,549
3	Budget	130,000,000

rating

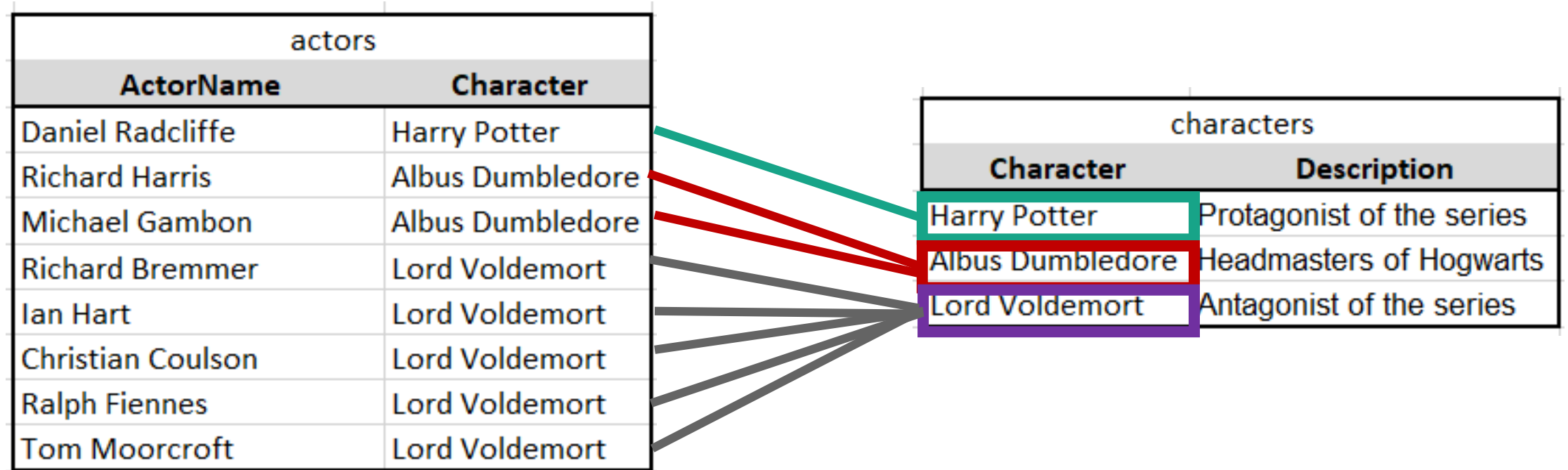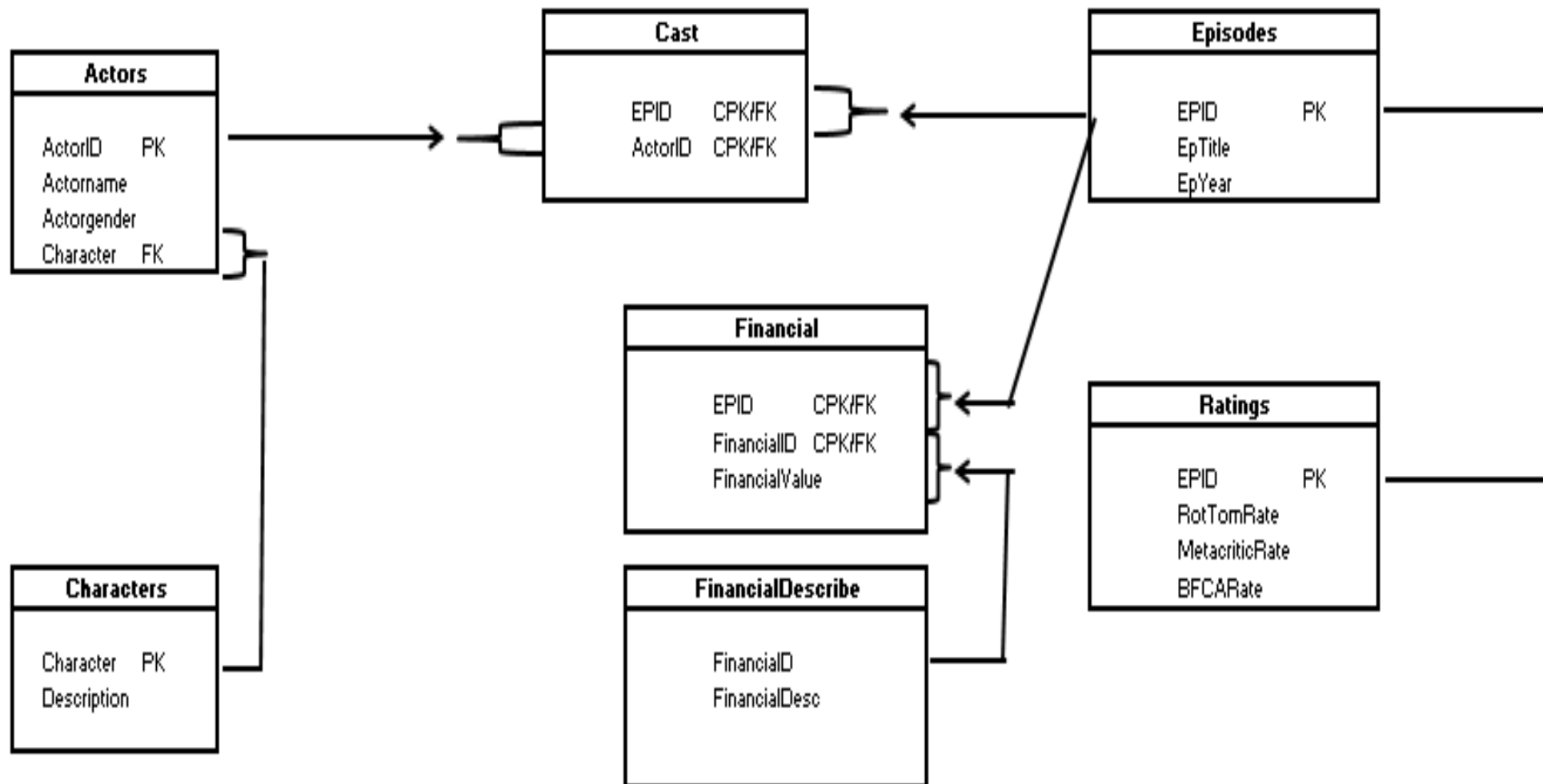| EpID	RotTomRate
1	80
2	83
3	91

UBC
MFRE

Relationships

- Relationships determine how entities are connected to each other.

- The cardinality of a relationship determines how many relationships an entity can participate in.
    - One-to-one
    - One-to-many
    - Many-to-many

Relationships

- Relationships determine how entities are connected to each other.

- The cardinality of a relationship determines how many relationships an entity can participate in.
  - One-to-one
  - One-to-many
  - Many-to-many

Assigning Keys

- A primary key (PK) is a unique identifier for an entity.

- The foreign key (FK) is the reference to the primary key of another entity.

- A composite key (CK) is a key that consist of two or more attributes.
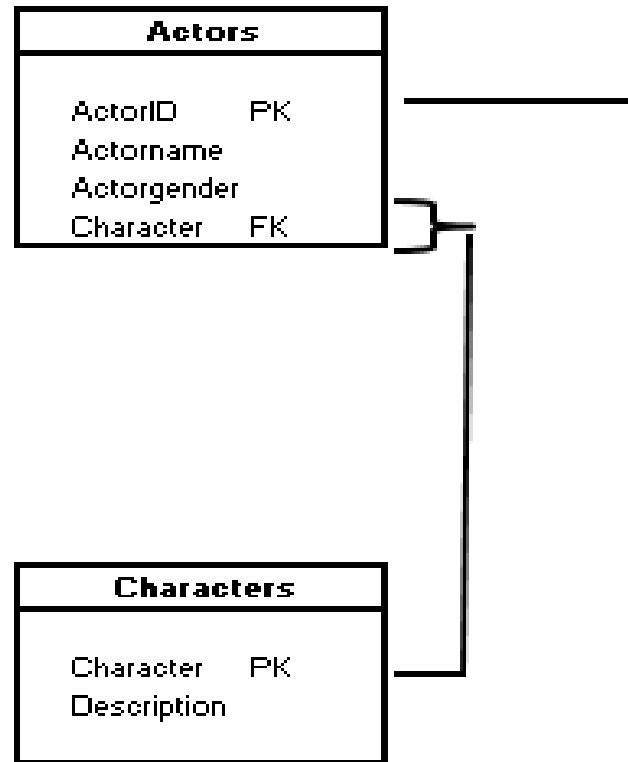
actors - characters

One-to-one

- In a one-to-one relationship, there is a one-to-one mapping from one table to a second table

- In the HP database, the relationship between episode and rating is one-to-one because there is a one-to-one mapping

- In this case, the table's PK also serves as the FK.

UBC
MFRE

One-to-many

- One-to-many relationships are perhaps the most common in a relational database

- In the HP database, there is a one-to-many relationship for the actors and characters entities.

- The table on the "many" side of the equation has the FK, and the table on the "one" side has the PK.
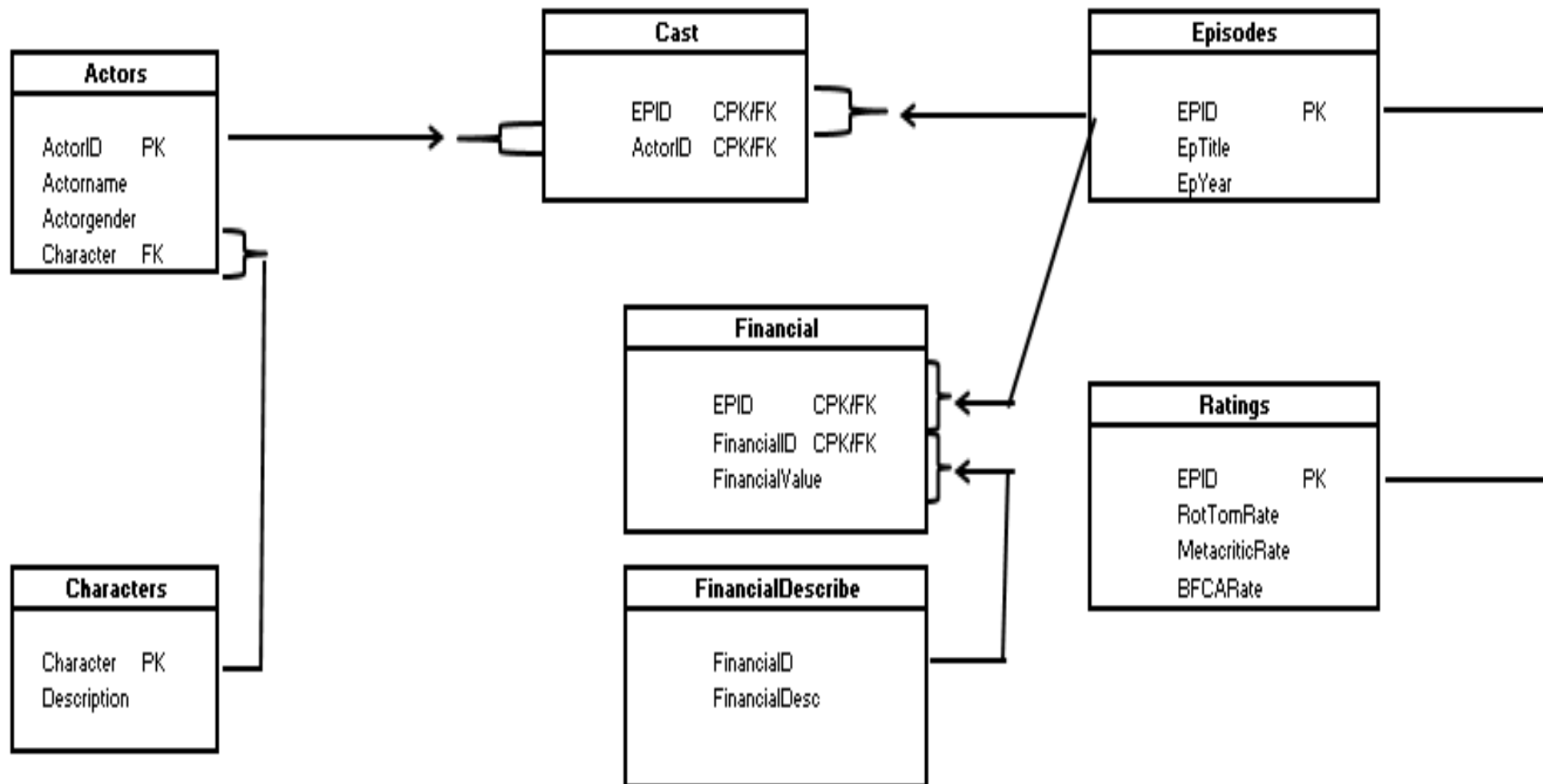
One-to-many

- Across the different episodes Lord Voldemort is played by:
  - Ian Hart, Richard Bremmer, Christian Coulson, Ralph Fiennes and Tom Moorcroft

- Each of these actors only plays Lord Voldemort

Many-to-many

- A many-to-many relationship requires a link/junction table

- In the HP database, *Episodes* and *Actors* have a many-to-many relationship.
  - One actor can play in more than one episode. Each episode have more than one actor.

- The PK is implicit since it is implied by the unique combination of the two FK.

Recap

- Understood the value of a database over a spreadsheet

- Learned about basic database concepts including database design, entities, and relationships