Fall 2016 - Week 1 - MIDS - W205

Exercise #1

You are tasked with storing… and retrieving "Knowledge Graph" data for a large search engine. The data stored relates to Actors, Directors and Movies. An Actor, Director or Movie card are presented when the appropriate search query is performed.

Part 1

1. What types of queries would you need to perform on the data?
2. Write the CREATE statements to store the textual info for our Knowledge Graph.
3. Use as many constraints as you can and should.
4. What are your primary keys?
5. Give an example of an operation that would fail due to "schema-on-write"? Constraints?
6. How many queries would you need to perform to display the Actor page?
7. What happens if you are missing some of the information?

Part 2

1. The Product Manager decide to add TvSeries data. How would you represent this data? Seasons? Episodes?
2. How would you find an actor's acting roles in movies & tv series?
3. Could you know if a relation is a TvSeries or Movie (so you can display it differently)?
4. Could you say who acted or directed a specific episode?

Brad Pitt

Actor

William Bradley "Brad" Pitt is an American actor and producer. He has received multiple awards and nominations including an Academy Award as producer under his own company Plan B Entertainment. Wikipedia

Born: December 18, 1963 (age 52), Shawnee, OK

Height: 5' 11"

Spouse: Angelina Jolie (m. 2014), Jennifer Aniston (m. 2000–2005)

Children: Shiloh Nouvel Jolie-Pitt, Zahara Marley Jolie-Pitt, More

Siblings: Doug Pitt, Julie Pitt Neal

Movies   View 45+ more

Fight Club 1999 — Fury 2014 — World War Z 2013 — Inglourious Basterds 2009 — Mr. & Mrs. Smith 2005

More images

Fight Club

R · 1999 · Drama film · 2h 31m

8.8/10	79%	66%
IMDb	Rotten Tomatoes	Metacritic

A depressed man (Edward Norton) suffering from insomnia meets a strange soap salesman named Tyler Durden (Brad Pitt) and soon finds himself living in his squalid house after his perfect apartment is destroyed. The two bored men form an underground club with strict rules and fight other men who are f... More

Release date: September 21, 1999 (USA)

Director: David Fincher

Screenplay: Jim Uhls

Adapted from: Fight Club

Box office: 100.9 million USD

5. The Product Manager decided to add Athletes as well as politicians. How would you store this data?
6. Would Arnold Schwarzenegger present a problem? Cause data duplication?

Part 3

1. Larry, Sergey and David Filo command you to switch to DynamoDB.
2. How would you store Movie, Actor, Director, TvSeries, Politician, Athlete information now?
   1. Define the schema you'd use.
   2. What keys would you use for your tables?
   3. How do the different entities "connect" (reference) each other?
3. How many queries would you need to display the Actor Card? The movie Card? Would you need to scan the whole table?