Exam

Exercise 1 – Databases

Given the following table containing data on movies, their actors and the roles the actors play in that movie:

movie:

MID	Title	Prod	AID	ActorName	Role
1	The Matrix	1999	23	Keanu Reeves	Neo
2	Star Wars	1977	45	Mark Hamill	Luke Skywalker
1	The Matrix	1999	98	Carrie-Ann Moss	Trinity
3	Speed	1994	23	Keanu Reeves	Jack Traven
4	The Prestige	2006	28	Hugh Jackman	Robert Angier
4	The Prestige	2006	28	Hugh Jackman	Gerald Root

Here, `MID` is a unique ID for movies, and `AID` is a unique ID for actors, `Title` is the title of the movie, `Prod` is the year the movie was produced, `ActorName` is the name of the actor and `Role` is the name of the character the actor played in that movie.

1. Give examples of data duplication in the table above.
2. Make a database (i.e. a collection of tables) with a better structure containing the same data as the table above.

Solution

1 The table combines movies with actors and the role they play in that movie. This means that data on every movie is repeated once for evey actor and/or role in that movie, so e.g. title (`The Matrix`) and production year `1999` for movie with ID `1` is repeated once for every actor in that movie (twice in the above table). Similarly, the name of the actor with ID `23` is repeated once for every movie that actor plays in.

2 We fix this by splitting the table into three, one for movies, one for actors, and one for roles played by actors in movies:

movie:

MID	Title	Prod
1	The Matrix	1999
2	Star Wars	1977
3	Speed	1994
4	The Prestige	2006

actor:

AID	ActorName
23	Keanu Reeves
28	Hugh Jackman
45	Mark Hamill
98	Carrie-Ann Moss

role:

MID	AID	Role
1	23	Neo
2	45	Luke Skywalker
1	98	Trinity
3	23	Jack Traven
4	28	Robert Angier
4	28	Gerald Root

Exercise 2 – SQL

Given the following table on products from a particular general store:

product:

product_id	name	category	price	vat
1	TV	Electronics	9995	0.25
2	Rice	Food	39	0.1
3	Socks 6pc	Clothing	199	0.2
4	Laptop	Electronics	8599	0.25
5	Blueberry	Food	27	0.1
6	Chocolate	Food	45	0.2
7	Headset	Electronics	899	0.25
8	Sweater	Clothing	849	0.2
9	Pants	Clothing	1099	0.2

where `product_id` is a uique ID for products, `name` is the product's name, `category` is the name of the category of the product, `price` is the base price (before applying VAT) and `vat` is the Value Added Tax for the product. The *total price* for a product is the `price` pluss the `vat` as a percentage of `price`. E.g. the *total price* of `Rice` is `39 + (39 * 0.1) = 42.9`.

1. What would the following query return:

```sql
SELECT category
FROM product
WHERE name = 'TV';
```

2. What would the following query return:

```sql
SELECT name, category
FROM product
WHERE price > 1000 OR vat = 0.25;
```

3. What would the following query return:

```sql
SELECT min(price) AS price_cheapest
FROM product
WHERE category = 'Electronics';
```

4. Write an SQL-query that finds the name and category of all products having a price greater than `1000` or less than `500`.
5. Write an SQL-query that finds the number of products in the `Clothing` category with a price below `1000`.
6. Write an SQL-query that finds the name and *total price* of all products.

Solution

1

category
Electronics

2

name	category
TV	Electronics
Laptop	Electronics
Headset	Electronics
Sweater	Clothing
Pants	Clothing

3

price_cheapest
899

4

```sql
SELECT name, category
FROM product
WHERE price >= 1000 OR price <= 500;
```

5

```sql
SELECT count(*) AS num_clothes
FROM product
WHERE category = 'Clothing';
```

6

```sql
SELECT name, price + (price * vat) AS total_price
FROM product;
```