SAKILA DB NORMALISATION

From the schema provided, we can make some considerations on how to improve the current database to make it more efficient:

PAYMENT and RENTAL table

payment does not have any relation to rental, it's not meeting the referential integrity (2NF)

in this case, **payment id** would be the primary key and **rental id** the foreign key in the related table

STORE and RENTAL table

store and rental are not related, so we can't tell on which store was a given film rented

rental_id would be the primary key and store_id the foreign key in the related table

FILM table special features

on the other hand, in the films table, from the python analysis, there is a violation of the 1NF, where you have more than one value on each **special_features field** (deleted scenes, behind the scenes; trailers, deleted scenes):

special_features
Deleted Scenes,Behind the Scenes
Trailers,Deleted Scenes
Trailers,Deleted Scenes
Commentaries,Behind the Scenes
Deleted Scenes

suggested approach to the previous issue would be creating a separate column, named "special_features_2", and so forth, depending on the number of features that you may have

also, in this database, in the film table, we have no "original_language_id", so we could easily remove this field in order to make our database more efficient:

```
print(data2.isna().sum())
print()
```

```
film_id                    0
title                      0
description                0
release_year               0
language_id                0
original_language_id    1000
rental_duration            0
rental_rate                0
length                     0
replacement_cost           0
rating                     0
special_features           0
last_update                0
dtype: int64
```