Check data types and nulls

Split multiple columns by delimiter into rows to extract needed data

Merge tables to get certain data

Duplicate table to many copies to create dimensions and factless tables

Select the needed columns and remove the others

Remove all duplicate rows

Also we found that some columns had inappropriate data so we eliminated it from our mode

…………………………………………………………..

In dimension model we use galaxy schema model

Dimensions :

Movie\_dim

Date\_dim

Productn\_comp\_dim

Genres\_dim

Product\_country\_dim

Keywords\_dim

Crew\_dim

Cast\_dim

Spoken\_language\_dim

……………………………………………………

Fact and factless tables

Genres\_factless

Productn\_componies\_factless

Movie\_fact

Product\_countries\_fact

Keywords\_factless

Spoken\_language\_factless

Crew\_factless

Cast\_factless

…………………………………..

In date dimension

We use the formula

= List.Dates(#date(1980,01,01), 365\*37, #duration(1,0,0,0))

To add a date column

Then we change the data type to date

Then we extract some columns needed

(Year , month , quarter , week of year , day , day\_name )

