{ Idea for scrapping }

Since we have to extract genre wise movies we we should go to imdb genre page

imdb. com /feature /genre /

→ popular movies
   └→ Action, Adventure, Animation ------------ , Western

total total 23 genre

We are using Beautiful soup and for requesting pages we are using urllib. request

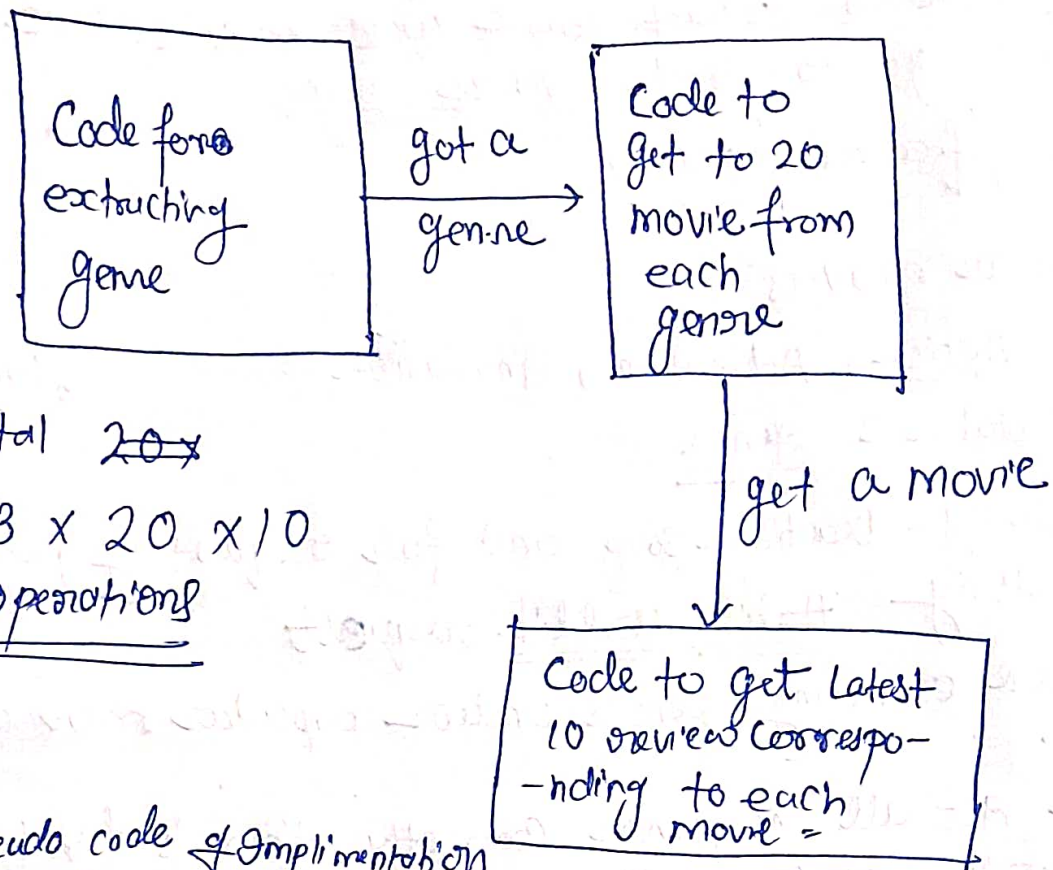then we are extracting genres under popular movies section,

once we get all the genre correctly this is first successful step ☺

Now we Need to run some for loops since we need 20 movie from each genre-

On getting each genre we move to its page & extract 20 movies out of it & each movie we need to get 10 reviews of latest users we will request each time by urllib. request In this sence we need total.

$$23 \times 20 \times 10 \text{ request which is huge}$$

① Definitly this Idea will take some time in execution Because server also taking its time for execution!

```
┌─────────────────┐              ┌─────────────────┐
│ Code for        │   got a      │ Code to         │
│ extracting      │   genre      │ get to 20       │
│ genre           │ ───────────→ │ movie from      │
│                 │              │ each            │
│                 │              │ genre           │
└─────────────────┘              └─────────────────┘
                                          │
                                          │ get a movie
total 20✗                                 │
                                          ▼
23 × 20 × 10                     ┌─────────────────┐
                                 │ Code to get latest
operations                       │ 10 review correspo-
                                 │ -nding to each
                                 │ movie =
                                 └─────────────────┘
```

Pseudo code of Implimentation

for genre in genres :

    get genre_name
    get genre_link (Need to more in corsponding movies)
    requist (genre link) = movies
    count = 0; (to counter for No. of movies)

    for movie in movies-

        # get review of each movie
        review = requst (review link)

        for review in reviews

            ~~Extract (countries)~~;
            Extract (reviews);

Extract all these data in CSV file in each
    iteration, & get the work done !!