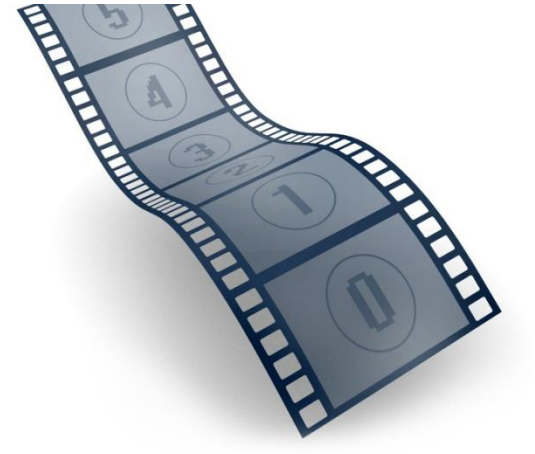
Description of the APIs



Our APIs

We chose to work with APIs from RapidAPI, named "data-imdb" and "Movies/TVShows Data (IMDB)".

We got data about the movies and the actors.

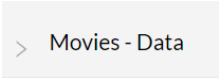
Get genres :

- we needed to get all the genres that are relevant for our movies. There were also genres of series so we filtered them. The data looks like this:

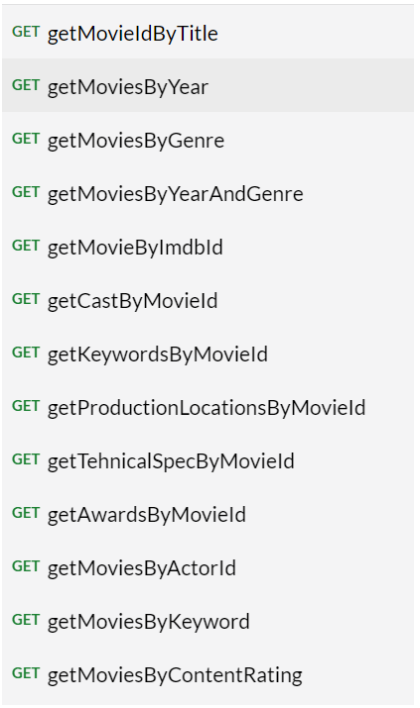

```
{ 1 item
  "results" : [ 27 items
    0 : { 1 item
      "genre" : "Adventure"
    }
    1 : { 1 item
      "genre" : "Family"
    }
    2 : { 1 item
      "genre" : "Fantasy"
    }
    3 : { 1 item
      "genre" : "Crime"
    }
```


GET getGenres

Movies- data :

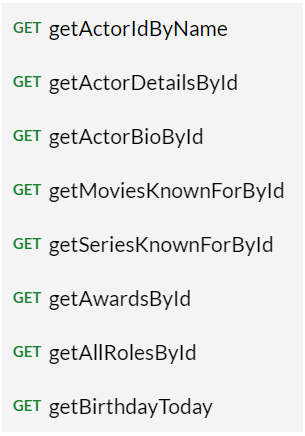


- we used several endpoints from here:
  - getMoviesByYear: It helped us get all the ids of movies from 2000 till now. After we have all the ids of the movies, we can use endpoints that go by MovieId.
  - getMovieByImdbId : From here we got data about the movie itself such as description. This endpoint was really helpful!
  - getAwardsByMovieId: It helped us get all the information we needed about the awards of each movie.
  - getCastByMovieId: It helped us get all the information we needed about the cast of the movies.



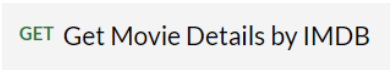
Actors- data :

- we used several endpoints from here:



  - getActorDetailsById: It helped us get all the information we needed about the actor itself. Information like age, height etc.
  - getAwardsById: It helped us get all the information we needed about the awards of each actor and for each director (we did them separately because we have different queries that use different information).

Get Movie Details by IMDB:



  - From here we got more data about the movie itself that we needed.

Here are some more examples of how the data looks like in the API:

In order to get the data from the API to python, we use requests (an elegant and simple HTTP library for Python).

We needed this information to use each endpoint:

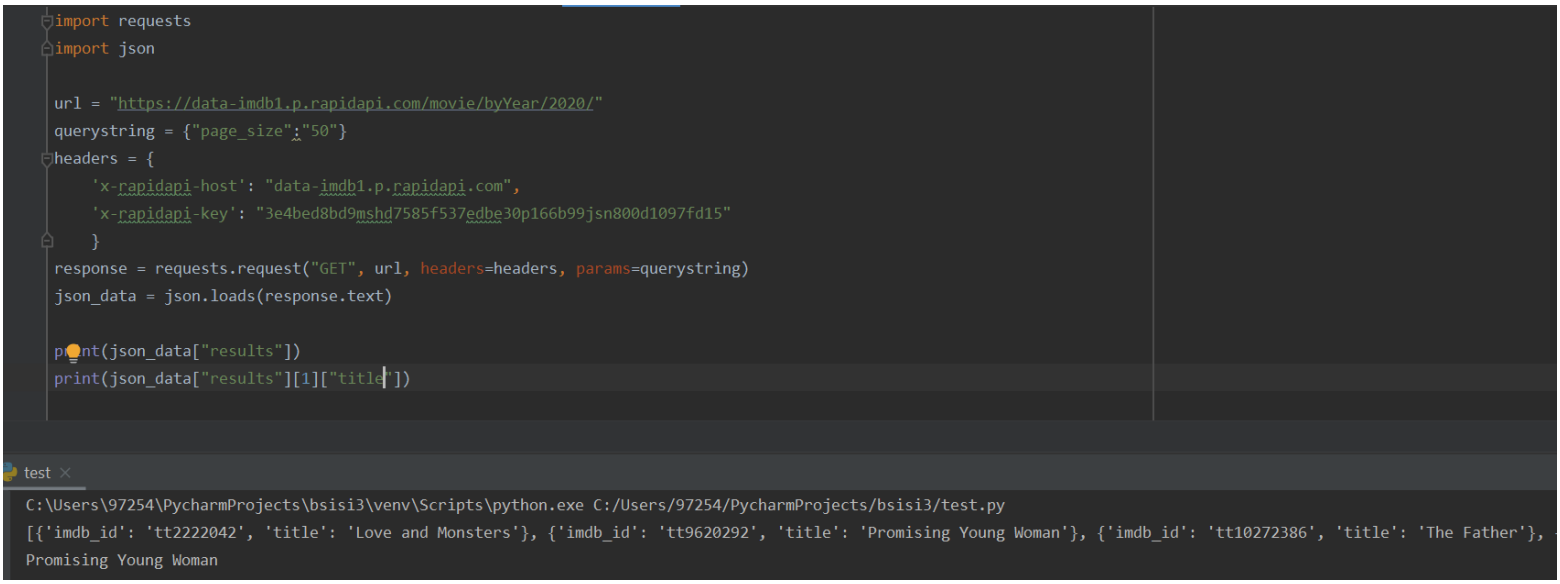
- url- the url of the endpoint.
- Querystring- we used the default of 50 items in each page.
- Headers of the API.

With this information we can use the request function of requests.

```
url = "https://data-imdb1.p.rapidapi.com/movie/byYear/2020/"

querystring = {"page_size":"50"}

headers = {
    'x-rapidapi-host': "data-imdb1.p.rapidapi.com",
    'x-rapidapi-key': "3e4bed8bd9mshd7585f537edbe30p166b99jsn800d1097fd15"
    }

response = requests.request("GET", url, headers=headers, params=querystring)
```

And then we converted it into json, so that it will be easier for us to play with the data.

Here is an example:



As you can see, we asked for movies from the year 2020. Each movie has an imdb_id (movie_id in our DB) and title (the name of the movie). Each movie is a dictionary, and they are all in an array.