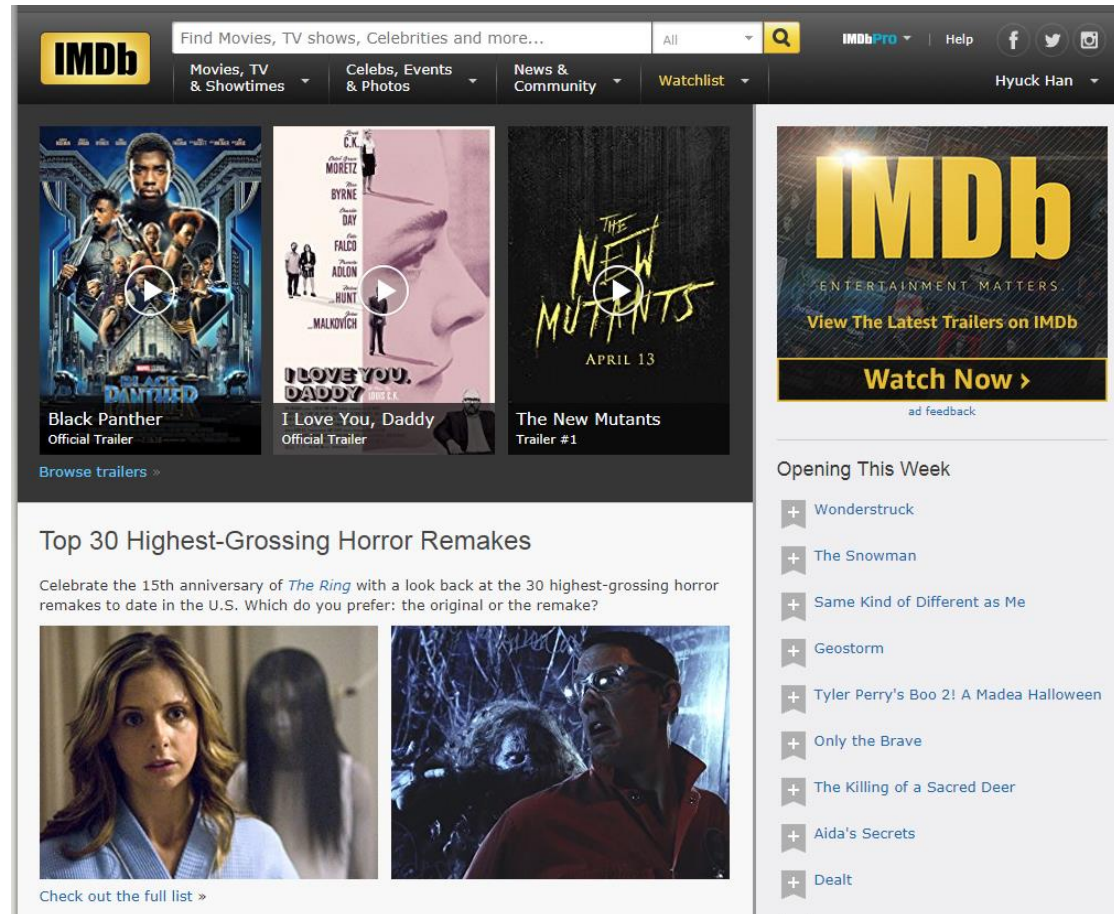# HW

IMDb

www.imdb.com

The world's most popular and authoritative source for movie, TV and celebrity content.

How can we access the IMDb data?

- Please refer to the following url:
  - http://www.imdb.com/interfaces/
- We will use modified data files.
  - The prof. will provide the "movie.dat" file for movie information.

Movie.dat

- Each line has movie id, movie title(year), genre information for a movie.
- Each information in a line is separated by two consecutive colons (i.e., ::).
- Example)
  - 23::Assassins (1995)::Thriller
    - → ID : 23, Title(year) : Assassins (1995), Genre : Thriller
  - 48::Pocahontas (1995)::Animation|Children's|Musical|Romance
    - → ID : 48, Title (year) : Pocahontas (1995), Genre : Animation | Children's | Musical | Romance

Q1. Genre count

- We need know how many movies falls into each genre.
- Input file : movie.dat
- Output : <Genre> <# of movies>
  - Example)
  Thriller 182
  Romance 108
  Children's 58
  ...

- Submission
  - No document, no image
  - Python source code
    - The file name for source code MUST BE **IMDBStudent<Your ID>.py**.
    - Example) If your student id is 20151047, your file for source code MUST BE **IMDBStudent20151047.py**.
    - If you don't follow the rule, your submission may be failed.
- When the due date is passed, the prof. will run your program and check whether your program runs correctly or not.

- Your code MUST process two command line parameters.
  - The first parameter : input file
  - The second parameter : output file
- Your code may be executed through the following command:
  python3 IMDBStudent20151047.py **movie.dat movieoutput.txt**

Uber

www.uber.com

UBER   탑승하기   운전하기   고객지

어디로 갈까요?
Uber로 편하게 이동하세요

Uber develops, markets and operates the Uber car transportation and food delivery mobile apps

uber.dat

- Each line has base number, date, active vehicles, and trips.
- Each information in a line is separated by a single comma.
- Example)
  - B02512,1/1/2015,190,1132
    - → Base number : B02512, Date : 1/1/2015, active vehicles : 190, trips : 1132
    - **Format of date : month/day/year**

Q2. Trips & Vehicles

- We need find the day on which each region has trips and active vehicles.

- Input file : uber.dat

- Output : <region,day> <vehicles,trips>
  - Example)
  B02512,MON 122,1922
  B02512,TUE 451,2200
  B02512,WED   453,3198

  …

  Weekday code : MON/TUE/WED/THU/FRI/SAT/SUN

- Submission
  - No document, no image
  - Python source code
    - The file name for source code MUST BE **UBERStudent<Your ID>.py**.
    - Example) If your student id is 20151047, your file for source code MUST BE **UBERStudent20151047.py** .
    - If you don't follow the rule, your submission may be failed.
- When the due date is passed, the prof. will run your program and check whether your program runs correctly or not.

- Your code MUST process two command line parameters.
  - The first parameter : input file
  - The second parameter : output file
- Your code may be executed through the following command:
  python3 UBERStudent20151047.py **uber.dat uberoutput.txt**

- You should submit your code through github.
  - The name of subdirectory in your project repo must be "**HW2**"

- Deadline : Nov. 2. 11:55AM