Data Management and Analysis
Assignment 1

NYU, Paris

Due Date: $14-02-2022$

CSV

Write Python codes to do the following. You can only use `csv` package. The file covid.csv is available on Brightspace. The file contains the record of codiv statistics on various days for every location starting from the first case or 31st December 2019 (whichever is earlier) to 29th November 2020.

1. For ever country, output the first day when covid started.
   (Print the output on the terminal)

   Example:
   Algola $2020-03-22$
   Sri Lanka $2019-12-31$
   World $2019-12-31$

2. For each day compute the average new cases of the previous 7 days for every location. (if there are less than 7 days from the current day, calculate the average from the begining of the first case in that location to the current day)

   Print the result on the terminal.

3. A peak for a location is defined as the day such that the 7 day average new cases of the previous day is less than the 7 day average new cases of the current day and also the 7 day average new cases of the next day is less than the 7 day average new cases of the current day.

   Calculate the list of peak days for every location. (Print the result on the terminal)

JSON

Write Python codes to do the following. You can only use `json` package. The file movies.json is available on Brightspace. The file contains some movie information.

1. The Average rating is set to 0 by default for every movie, which is wrong. The ratings are given as a list for every movie. Compute the average rating and update the information for every movie accordingly. Write the updated information (along with old information) to a new file named averageUpdated.json

2. For every genre, compute the movie with maximum ratings. (If there are multiple, then output all of them). Note that some movies are in multiple genres.
   (Print the output on the terminal)

XML

Write Python codes to do the following. The file station.xml is available on Brightspace.

1. Use `xmltodict` package to parse the xml file and create a new json file `stationEmail.json` that contains the information of every station name along with its website.

2. The `Email` tag is empty in all records. Create a new xml file `noEmail.xml` which has the same information as station.xml but the `Email` tag is removed. You can use only `ElementTree` package.