**Assignment 1**

**Instructions**

 This assignment starts with TV shows dataset with around 150k different show entries. First task you have is to explore data set and get understanding of columns and information they hold. Not every column will be needed, so you can minimize dataset only to columns that are needed for an assignment, however in order to accomplish that you will need to deeply understand assignment tasks in front of you. Results of each task should be stored in a separate file as a final output. For the assignments you have to complete following tasks:

1. Find top 50 tv shows based on rating;
2. Find top 50 tv shows based on rating and original language; (each language will be a separate file, e.g. task2.en.csv)
3. Find total number of shows that have vote count greater than 15000;
4. Find top 50 tv shows based on rating where first air data and last air data are between 1st January 2010 and 31st December 2019;
5. Find top 50 tv shows that are still in production;
6. Find top 50 tv shows based on popularity that have a single word in the title (e.g. Riverdale);
7. Find top 50 tv shows based on genre (e.g. Crime, Drama)

Code should be pushed to a **private** repository on your own GitHub profile.

**Assignment Report**

Once you finished with coding part of the assignment, please write one page (500 words) assignment review paper where you should reflect on work you did, with a walkthrough of your process in order to process file and provide reports; finally please provide answers on following questions:  
- What was most challenging?  
- What you expected to be very hard, but it turned to be easy work for you? (if anything)  
- What you expected to be simple, but turned out to be very complex? (if anything)

- If you are starting assignment again, what would you do differently this time?   
  
Report should be pushed to the GitHub repository as well. Report has to be in PDF format;