# WEB SCRAPING – ASSIGNMENT 4

- **Read all the problem statements, notes carefully and scrape the required data using any web scraping tool of your choice.**
- **You have to handle commonly occurring EXCEPTIONS by using exception handling programing. To get information about selenium Exceptions. You may visit following links:**
  1. **https://selenium-python.readthedocs.io/api.html**
  2. **https://www.guru99.com/exception-handling-selenium.html**
  3. **https://stackoverflow.com/questions/38022658/selenium-python-handling-no-such-element-exception/38023345**

1. Scrape the details of most viewed videos on YouTube from Wikipedia:
   Url = https://en.wikipedia.org/wiki/List_of_most-viewed_YouTube_videos/
   You need to find following details:
   A) Rank
   B) Name
   C) Artist
   D) Upload date
   E) Views

2. Scrape the details team India's international fixtures from bcci.tv.
   Url = https://www.bcci.tv/.
   You need to find following details:
   A) Match title (I.e. 1st ODI)
   B) Series
   C) Place
   D) Date
   E) Time
   Note: - From bcci.tv home page you have reach to the international fixture page through code.



3. Scrape the details of selenium exception from guru99.com.
   Url = https://www.guru99.com/
   You need to find following details:
   A) Name
   B) Description
   Note: - From guru99 home page you have to reach to selenium exception handling page through code.

4. Scrape the details of State-wise GDP of India from statisticstime.com.
   Url = http://statisticstimes.com/
   You have to find following details:
   A) Rank
   B) State
   C) GSDP at current price (19-20)
   D) GSDP at current price (18-19)
   E) Share(18-19)
   F) GDP($ billion)
   Note: - From statisticstimes home page you have to reach to economy page through code.

5. Scrape the details of trending repositories on Github.com.
   Url = https://github.com/
   You have to find the following details:
   A) Repository title
   B) Repository description
   C) Contributors count
   D) Language used

Note: - From the home page you have to click on the trending option from Explore menu through code.

6. Scrape the details of top 100 songs on billboard.com.
   Url = https://www.billboard.com/
   You have to find the following details:
   A) Song name
   B) Artist name
   C) Last week rank
   D) Peak rank
   E) Weeks on board
   Note: - From the home page you have to click on the charts option then hot 100-page link through code.

7. Scrape the details of Data science recruiters from naukri.com.
   Url = https://www.naukri.com/
   You have to find the following details:
   A) Name
   B) Designation
   C) Company
   D) Skills they hire for
   E) Location
   Note: - From naukri.com homepage click on the recruiters option and the on the search pane type Data science and click on search. All this should be done through code

8. Scrape the details of Highest selling novels.
   Url = https://www.theguardian.com/news/datablog/2012/aug/09/best-selling-books-all-time-fifty-shades-grey-compare/
   You have to find the following details:
   A) Book name
   B) Author name
   C) Volumes sold
   D) Publisher
   E) Genre



9. Scrape the details most watched tv series of all time from imdb.com.
   Url = https://www.imdb.com/list/ls095964455/
   You have to find the following details:
   A) Name
   B) Year span
   C) Genre
   D) Run time
   E) Ratings
   F) Votes

10. Details of Datasets from UCI machine learning repositories.
    Url = https://archive.ics.uci.edu/
    You have to find the following details:
    A) Dataset name
    B) Data type
    C) Task
    D) Attribute type
    E) No of instances
    F) No of attribute
    G) Year
    Note: - from the home page you have to go to the Show All Dataset page through code.