Project 5.27. HBSC Recruitment Information Web Scraping

Requirement:

1. Guide the user to customize the HSBC job search by requesting the status and desired business area of the user.
2. From <http://www.hsbc.com/careers/students-and-graduates/programmes>, scrape all relevant job entries and organize the information into a DataFrame, which presents the following information of each job where available:
   1. Programme
   2. Areas
   3. Duration
   4. Description
   5. Link
3. Scrape the corresponding web page of each job entry and include the following application details in the DataFrame as well:
   1. Requirements
   2. Applications Open
   3. Applications Close
   4. Additional Languages
4. Display the DataFrame. Ask the user if he or she wants the search result exported to a .csv file. If so, export the DataFrame.

Codes will be graded based on:

1. The format of the codes
2. The organization of the csv file
3. The speed of the codes
4. The extendibility of the codes

Estimated completion time: 20 hours. Time consumption depends on the codes’ quality.

Suggestive points:

Deadline:

Bonus points:

1. Work completion before the deadline will generate an extra 20% bonus.

Penalty:

1. Late delivery/missing deadline will incur -10% of finally determined points of the whole project per missing day.
2. Serious mistakes that made the submitted results not useful will incur -120% of the whole project.