**Assessment:**

**Data Manipulation and Data Cleaning**

| **Section B (80 points total)** |
| --- |

**Clean the book information from British Library**

You are given a CSV file - [Book.csv](https://drive.google.com/file/d/1_zSPomZ1RpXus9hADilFaAfIM2ikoTPR/view?usp=sharing), which contains the information about books from the British Library.

You are then required to complete the following tasks:

1. Read the csv in a data\_frame. (*10 points*)
2. Remove the following columns from the data\_frame (*20 points*)
   * 'Edition Statement',
   * 'Corporate Author',
   * 'Corporate Contributors',
   * 'Former owner',
   * 'Engraver',
   * 'Contributors',
   * 'Issuance type',
   * 'Shelfmarks'
3. Check the content of the column- 'Date of Publication’ and define a function to clean the value (*20 points*) Examples:
   * 1879 [1878] to: 1879
   * [1858.] to 1858
4. Check the content of the column- ' Author’ and define a function to clean the value. And split the name to first name and last name. (*10 points*)
5. Check the content of the column- ‘Title’ and define a function to clean the value (*10 points*). Examples:
   * Walter Forbes. [A novel.] By A. A to: Walter Forbes
   * Love the Avenger. By the author of “All for Gr.. to Love The Avenger
6. Check the content of the column- ‘Place of Publication’ and define a function to clean the value. E.g., the original value: London; Virtue & Yorston to: London. (*10 points*)