Lab - 4 Assignment (Due on 06/03/2021 23:59:59):

1. Write a spark code for executing the Hash example provided in slide 14 on Hashing from Lab 1 Presentation, on the public file: `gs://bucket_two_2/hash_file.txt.` You would have to find the number of user clicks between 0-6, 6-12, 12-18, and 18-24, as was discussed in the first class.
   a. Submit the python file with your code.
   b. Also, provide the text file containing your output.   [6 marks]

2. Provide a brief description of the functionality of the following services:
   a. HDFS
   b. Hive
   c. Pig
   d. Yarn   [4 marks]

Create a report (as PDF) containing answers to the above questions. Then, zip it along with the Python source code and the text file (for the spark task).

[Please ensure that the name of this zip file should be <yourrollnumber>CS4830Assignment2.zip]

Finally, submit this zip file on moodle.