**Homework 1**

**Problem #1 (of 2): Monsoon account creation and workshop**

* Completed the Self-Paced Workshop
* Obtain and submit the validation codes to self-validate your account

The obtained codes are:

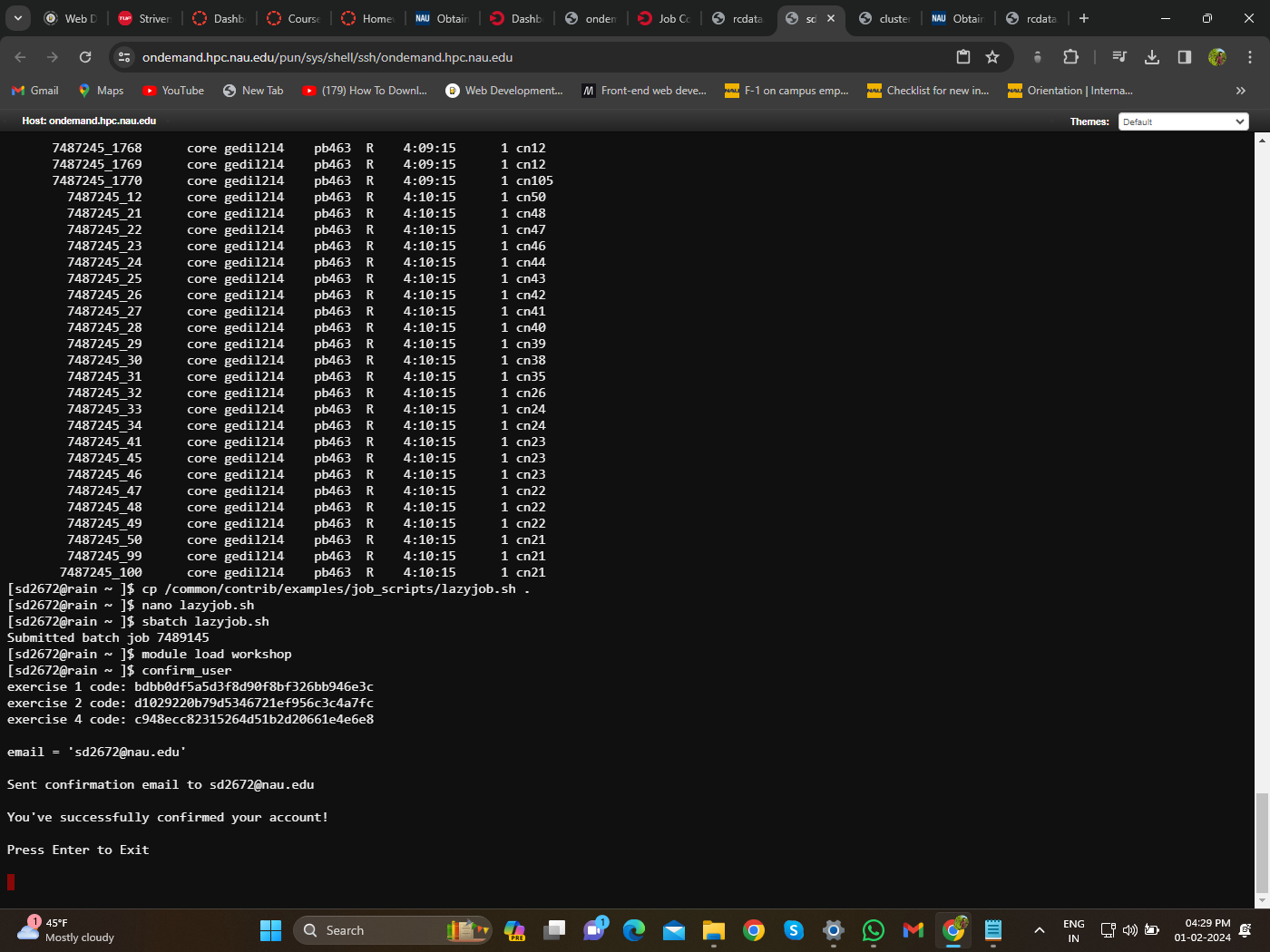
The secret code for exercise1 is: bdbb0df5a5d3f8d90f8bf326bb946e3c

The secret code for exercise2 is: d1029220b79d5346721ef956c3c4a7fc

The secret code for exercise4 is: c948ecc82315264d51b2d20661e4e6e8

* Take a screenshot of the successful ‘confirm user’ command (see example below) and submit it as part of your writeup to complete problem #1 of the assignment.

The captured screenshot is:



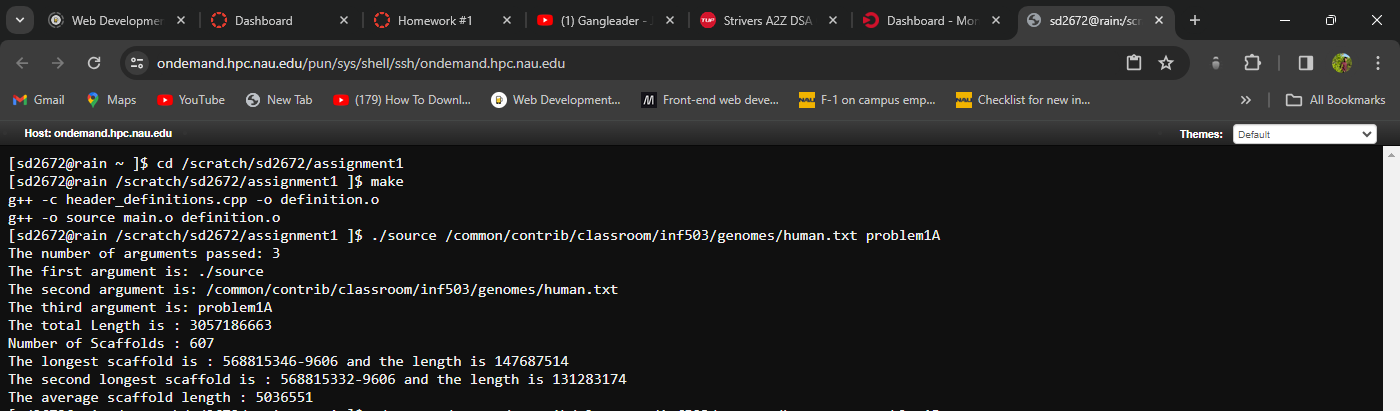
# Problem #2 (of 2): basic text processing

**Part A:**

* How many scaffolds were there?

**Ans:**607

* What was the longest and shortest scaffold? Provide names of scaffolds and lengths.
  + Longest scaffold: 147687514 and Scaffold name: 568815346-9606
  + Second longest scaffold: 131283174 and Scaffold name: 568815332-9606
* What was the average scaffold length?
  + Average scaffold length: 5036551



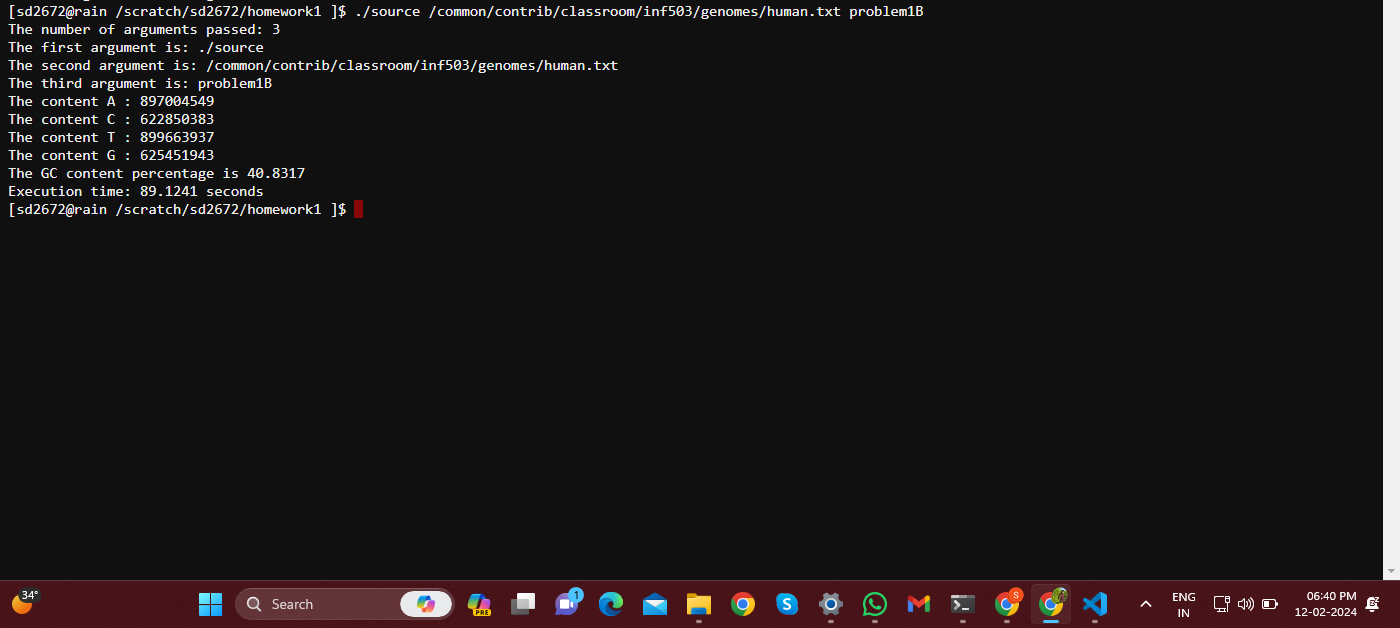
**Part B:**

* How long does it take (in seconds) to execute this function? Hint: You will need to use system time within your code to get accurate time estimates.

**Ans:** 89.1241 seconds

* What was the GC content of the human genome (percent of C’s and G’s in the genome)?

**Ans:** GC content: 40.8317%



**Steps of execution:**

* Created total of three files main.cpp, header\_definitions.cpp and header.h
* Header.h file contains all the header files that are used in the program
* The main.cpp contains the main function and all the function calls required to get desired output
* Definition.cpp file contains all the function definitions which are declared in the header file
* Created a make file to run the code
* Uploaded all the above files to a directory on monsoon
* There I have opened terminal and entered the command “make” then source executable file is generated next we need to run the source file with file path and the part of execution.
* The command for execution of part A is
* ./source /common/contrib/classroom/inf503/genomes/human.txt problem1A
* The command for execution of part B is
* ./source /common/contrib/classroom/inf503/genomes/human.txt problem1B