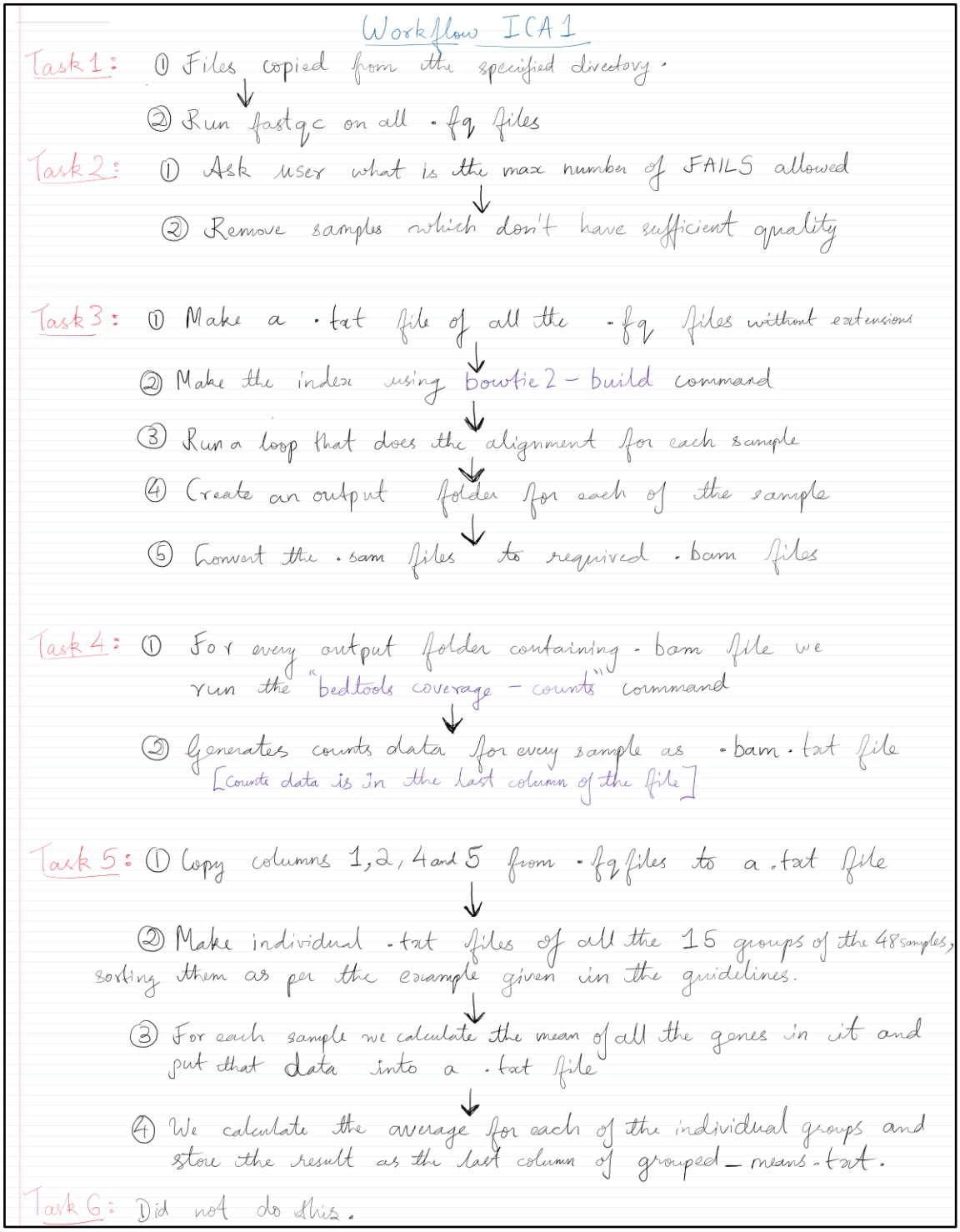
Information required as per guidelines:

1. The link for my PUBLIC GitHub ICA1 repository is as given:
   https://github.com/B236494-2023/ICA1
2. The ccrypt encryption key is: **123456**
3. The basic overview/flowchart of the pipeline is as depicted below:



4. ***Programme Parameters:*** In the code comments have been added where a description of the parameters is useful. Specific important parameters are as given below:
    - **Task 4:** We use the coverage -counts in the command in

```
bedtools coverage -counts -a ../TriTrypDB-46_TcongolenseIL3000_2019.bed -b ${file} > ${file}.txt
```

because we want to only count the number of overlaps and don't require any computation of fractions.

5. ***Useful Information For User:***
    - You will have to run the code in the homespace so that it runs without any errors. I should have corrected for it but I realised it too late.
    - A directory will be created called "folder_to_run_ICA" into which all the data is first copied and then sub-directories are created according to the tasks to be performed.
    - The last task was not completed and the option to give the user input for deciding what is the acceptable minimum quality of raw sequence data was not implemented into the code.
    - Didn't address the eventuality that there might be new samples with time 72h or 96h.
6. ***Difficulties Faced:***
    - I faced difficulty in making the workflow for this assignment as there were a lot of new tools I have never used so I wasn't aware of what the output was going to be for each step. It was quite hard to plan ahead without actually doing it first.
    - Handling this size of the data was also difficult as loops needed to be written which were efficient at getting the job done accurately.
    - Found it a bit harder to find solutions in different forums/webpages. Felt like the best way to go is to read the manual using '–help' as you have always suggested. This is mainly regarding bowtie2 and bedtools. For the actually coding there's a lot of help available. Youtube was also really helpful.
    - Time taken to complete the ICA was much more than I expected but it truly challenged me to work hard on solving the problem. I realised I need to start much earlier than I think.
    - Proper understanding of which parameter to use was confusing but got better with practice. Need more practice.