Course: MMDS

*Lab 03*

**MapReduce**

***Fill answers of the questions below in the given tables.***

***Your screenshots must contain commands for required operations.***

## Question 1:

Given a tsv file [WHO-COVID-19-20210601-213841.tsv](https://drive.google.com/file/d/1TG6orBmU74s1_Z3NDsyntRb9-OAHIuy_/view?usp=sharing) which is corresponding to the [WHO Coronavirus (COVID-19) Dashboard](https://covid19.who.int/table).

Students are required to create a folder, named **lab03**, in HDFS and then copy the tsv to **lab03/input/**

Take a screenshot to show the content of **lab03/input/** in HDFS

|  |
| --- |
| *My screenshot* |

## 

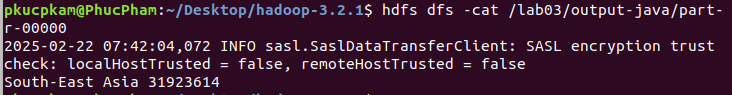
## Question 2:

Create one and only one java file, named **ASEANCaseCount.java**, to run a MapReduce job that counts the number of cumulative total cases among ASEAN countries (*South-East Asia Region in the given data table*).

The output of the MapReduce job is located in **lab03/output-java/**.

Submit the source code file following the instructions in Submission Notice.

Result screen:



## Question 3 (*optional*):

Create a pair of Python files, named **ASEANDeathCountMapper.py** and **ASEANDeathCountReducer.py**, to run a MapReduce job that counts the number of cumulative total deaths among ASEAN countries (*South-East Asia Region in the given data table*).

The output of the MapReduce job is located in **lab03/output-python/**.

Submit the source code files following the instructions in Submission Notice.

Result screen:

