Course: MMDS
Lab 02
HDFS

Fill answers of the questions below in the given tables.
Your screenshots must contain commands for required operations.

Question 1:

Download and unzip lab02.zip to obtain 365 text files.
Create the folder **/user/<username>/lab02** in your HDFS
Copy the entire **lab02/** from your local filesystem to **/user/<username>/lab02** in HDFS
Take two screenshots to show your results.
For example,

```
Found 1 items
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:27 /user/ntan/lab02/lab02
Found 365 items
-rw-r--r--   1 ntan supergroup          1 2021-05-18 15:27 /user/ntan/lab02/lab02/2021_01_01.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_02.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_03.txt
-rw-r--r--   1 ntan supergroup          1 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_04.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_05.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_06.txt
-rw-r--r--   1 ntan supergroup          1 2021-05-18 15:26 /user/ntan/lab02/lab02/2021_01_07.txt
```

Your screenshot goes here

```
chuong@ubuntu: ~/lab02

File  Edit  View  Search  Terminal  Help
2021_01_21.txt  2021_03_08.txt  2021_04_23.txt  2021_06_08.txt  2021_07_24.txt  2021_09_08.txt  2021_10_24.txt  2021_12_09.txt
2021_01_22.txt  2021_03_09.txt  2021_04_24.txt  2021_06_09.txt  2021_07_25.txt  2021_09_09.txt  2021_10_25.txt  2021_12_10.txt
2021_01_23.txt  2021_03_10.txt  2021_04_25.txt  2021_06_10.txt  2021_07_26.txt  2021_09_10.txt  2021_10_26.txt  2021_12_11.txt
2021_01_24.txt  2021_03_11.txt  2021_04_26.txt  2021_06_11.txt  2021_07_27.txt  2021_09_11.txt  2021_10_27.txt  2021_12_12.txt
2021_01_25.txt  2021_03_12.txt  2021_04_27.txt  2021_06_12.txt  2021_07_28.txt  2021_09_12.txt  2021_10_28.txt  2021_12_13.txt
2021_01_26.txt  2021_03_13.txt  2021_04_28.txt  2021_06_13.txt  2021_07_29.txt  2021_09_13.txt  2021_10_29.txt  2021_12_14.txt
2021_01_27.txt  2021_03_14.txt  2021_04_29.txt  2021_06_14.txt  2021_07_30.txt  2021_09_14.txt  2021_10_30.txt  2021_12_15.txt
2021_01_28.txt  2021_03_15.txt  2021_04_30.txt  2021_06_15.txt  2021_07_31.txt  2021_09_16.txt  2021_10_31.txt  2021_12_16.txt
2021_01_29.txt  2021_03_16.txt  2021_05_01.txt  2021_06_16.txt  2021_08_01.txt  2021_09_16.txt  2021_11_01.txt  2021_12_17.txt
2021_01_30.txt  2021_03_17.txt  2021_05_02.txt  2021_06_17.txt  2021_08_02.txt  2021_09_17.txt  2021_11_02.txt  2021_12_18.txt
2021_01_31.txt  2021_03_18.txt  2021_05_03.txt  2021_06_18.txt  2021_08_03.txt  2021_09_18.txt  2021_11_03.txt  2021_12_19.txt
2021_02_01.txt  2021_03_19.txt  2021_05_04.txt  2021_06_19.txt  2021_08_04.txt  2021_09_19.txt  2021_11_04.txt  2021_12_20.txt
2021_02_02.txt  2021_03_20.txt  2021_05_05.txt  2021_06_20.txt  2021_08_05.txt  2021_09_20.txt  2021_11_05.txt  2021_12_21.txt
2021_02_03.txt  2021_03_21.txt  2021_05_06.txt  2021_06_21.txt  2021_08_06.txt  2021_09_21.txt  2021_11_06.txt  2021_12_22.txt
2021_02_04.txt  2021_03_22.txt  2021_05_07.txt  2021_06_22.txt  2021_08_07.txt  2021_09_22.txt  2021_11_07.txt  2021_12_23.txt
2021_02_05.txt  2021_03_23.txt  2021_05_08.txt  2021_06_23.txt  2021_08_08.txt  2021_09_23.txt  2021_11_08.txt  2021_12_24.txt
2021_02_06.txt  2021_03_24.txt  2021_05_09.txt  2021_06_24.txt  2021_08_09.txt  2021_09_24.txt  2021_11_09.txt  2021_12_25.txt
2021_02_07.txt  2021_03_25.txt  2021_05_10.txt  2021_06_25.txt  2021_08_10.txt  2021_09_25.txt  2021_11_10.txt  2021_12_26.txt
2021_02_08.txt  2021_03_26.txt  2021_05_11.txt  2021_06_26.txt  2021_08_11.txt  2021_09_26.txt  2021_11_11.txt  2021_12_27.txt
2021_02_09.txt  2021_03_27.txt  2021_05_12.txt  2021_06_27.txt  2021_08_12.txt  2021_09_27.txt  2021_11_12.txt  2021_12_28.txt
2021_02_10.txt  2021_03_28.txt  2021_05_13.txt  2021_06_28.txt  2021_08_13.txt  2021_09_28.txt  2021_11_13.txt  2021_12_29.txt
2021_02_11.txt  2021_03_29.txt  2021_05_14.txt  2021_06_29.txt  2021_08_14.txt  2021_09_29.txt  2021_11_14.txt  2021_12_30.txt
2021_02_12.txt  2021_03_30.txt  2021_05_15.txt  2021_06_30.txt  2021_08_15.txt  2021_09_30.txt  2021_11_15.txt  2021_12_31.txt
2021_02_13.txt  2021_03_31.txt  2021_05_16.txt  2021_07_01.txt  2021_08_16.txt  2021_10_01.txt  2021_11_16.txt
2021_02_14.txt  2021_04_01.txt  2021_05_17.txt  2021_07_02.txt  2021_08_17.txt  2021_10_02.txt  2021_11_17.txt
2021_02_15.txt  2021_04_02.txt  2021_05_18.txt  2021_07_03.txt  2021_08_18.txt  2021_10_03.txt  2021_11_18.txt
chuong@ubuntu:~/lab02$ ls -1 | wc -l
365
chuong@ubuntu:~/lab02$
```

Question 2:

Create 12 folders, corresponding to 12 months in a year, in **/user/<username>/lab02** in HDFS.
Notice: folder names must in the format like _01, _02, …, _12
Take a screenshot to show your result.
Hint:

- loop in bash scripts
- $(printf …) for formatting strings in bash scripts.

For example,

```
Found 13 items
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_01
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_02
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_03
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_04
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_05
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_06
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_07
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_08
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_09
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_10
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_11
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:30 lab02/_12
drwxr-xr-x   - ntan supergroup          0 2021-05-18 15:27 lab02/lab02
```

Your screenshot goes here

```
Try: sudo apt install <deb name>

chuong@ubuntu:~/lab02$ ls
01  02  03  04  05  06  07  08  09  10  11  12  lab02-copy
chuong@ubuntu:~/lab02$
```

Question 3:

Text files in **lab02/** have a filename format like **YYYY_MM_DD.txt**

Move each text file in **lab02/** to the folder which is corresponding to the month in its filename. E.g. **20201_01_01.txt** to **_01/**

Take a screenshot to show the content in **lab02/_01/**

Take a screenshot to show the sizes (in bytes) of the 12 folders.
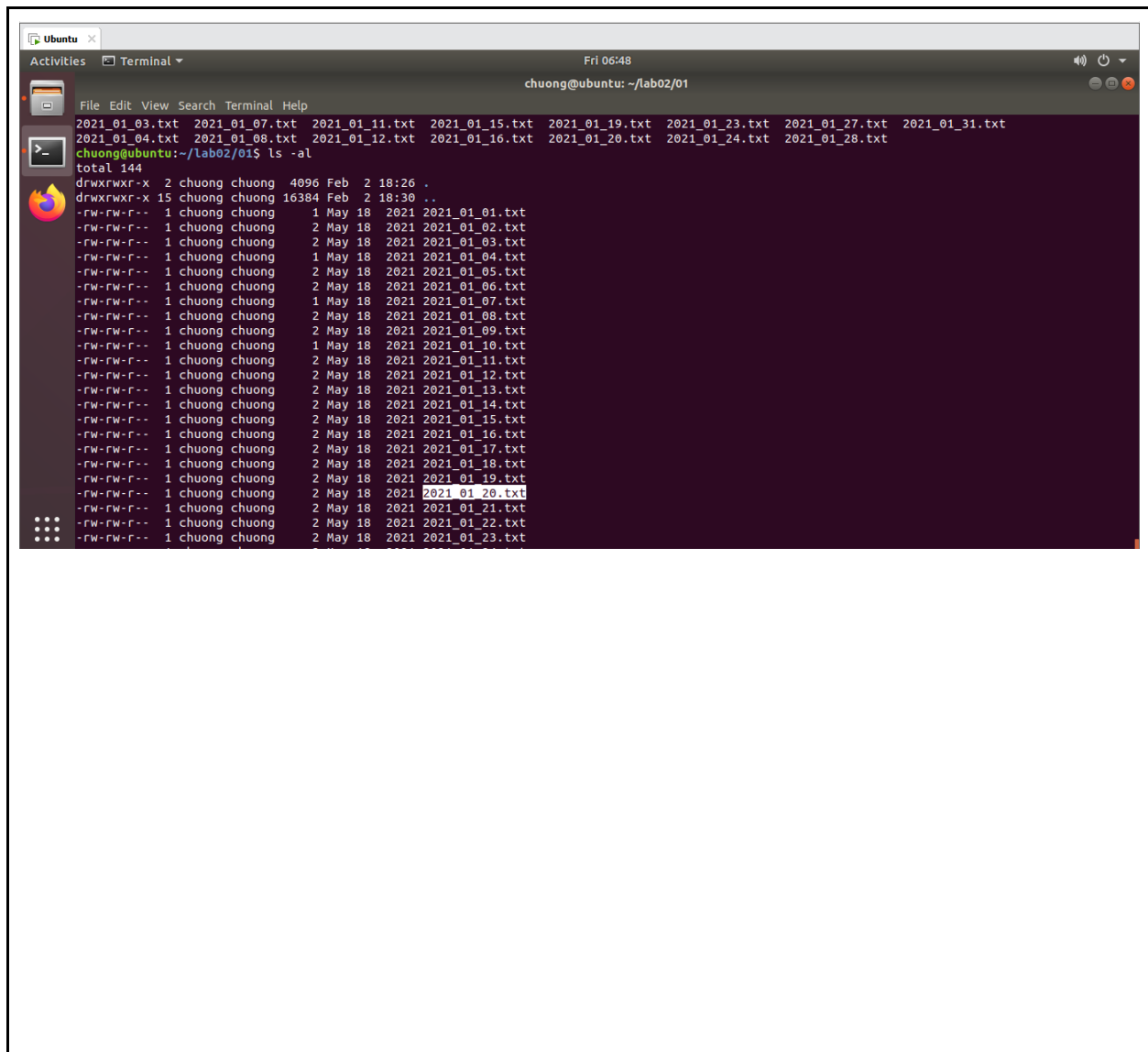
Hint: loop in bash scripts

For example,

```
Found 31 items
-rw-r--r--   1 ntan supergroup          1 2021-05-18 15:27 lab02/_01/2021_01_01.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 lab02/_01/2021_01_02.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 lab02/_01/2021_01_03.txt
-rw-r--r--   1 ntan supergroup          1 2021-05-18 15:26 lab02/_01/2021_01_04.txt
-rw-r--r--   1 ntan supergroup          2 2021-05-18 15:26 lab02/_01/2021_01_05.txt
```

```
58   58   lab02/_01
52   52   lab02/_02
58   58   lab02/_03
56   56   lab02/_04
58   58   lab02/_05
56   56   lab02/_06
58   58   lab02/_07
58   58   lab02/_08
56   56   lab02/_09
58   58   lab02/_10
56   56   lab02/_11
58   58   lab02/_12
```

Your screenshot goes here

Question 4:

Using **cat** to display the content of all files (in month order)
- day 10th
- day 01st
- day 07th
- day 04th

Take 4 screenshots of the corresponding results above.

For example,

```
2021-05-18 15:57:20,510 INFO s
se, remoteHostTrusted = false
e
c
2
1
2
1
a
9
3
7
f
e
```

```
chuong@ubuntu:~/lab02/lab02-copy$ cat thang04
_years__<3_.chuong@ubuntu:~/lab02/lab0thang10
_<3__i__lovechuong@ubuntu:~/lab02/lab0thang01
__you__for__chuong@ubuntu:~/lab02/lab0thang07
a__thousand_chuong@ubuntu:~/lab02/lab02-copy$ cat thang04
_years__<3_.chuong@ubuntu:~/lab02/lab02-copy$
chuong@ubuntu:~/lab02/lab02-copy$
```

Question 5:

Concatenate contents in the 4 screenshots of Question 4 to form the completed message. Finally, write it in the given table.

Your answer goes here

*_<3__i__love__you__for__a__thousand__years__<3_*

```
chuong@ubuntu:~/lab02/lab02-copy$ cat thang04
_years__<3_.chuong@ubuntu:~/lab02/lab0thang10
_<3__i__lovechuong@ubuntu:~/lab02/lab0thang01
__you__for__chuong@ubuntu:~/lab02/lab0thang07
a__thousand_chuong@ubuntu:~/lab02/lab02-copy$ cat thang04
_years__<3_.chuong@ubuntu:~/lab02/lab02-copy$
chuong@ubuntu:~/lab02/lab02-copy$
```

Submission Notice

- Export your answer file as pdf
- Rename the pdf following the format:

<center><student number>_HoTen.pdf</center>

  E.g: **123456_NguyenThanhAn.pdf**
  If you have not been assigned a student number yet, then use 123456 instead.
- Careless mistakes in filename, format, question order, etc. are not accepted (0 pts).