DAILY ONLINE ACTIVITIES SUMMARY

Date:	15-06-2020	Name:	PRASANNA
Sem & Sec	8th,B	USN:	4AL16CS068

Online Test Summary			
Subject	SMS		
Max. Marks	60	Score	No mail recieved

Certification Course Summary			
Course	Introduction to Hadoop		
Certificate Provider	Great learner academy	Duration	6 Hrs

Coding Challenges
Problem Statement: prob1- *To find number which is divisible by another number*
Status: Solved

Uploaded the report in Github	Yes
If yes Repository name	**prasanna_p**
Uploaded the report in slack	Yes

Online Test Details: (Attach the snapshot and briefly write the report for the same)

Certification Course Details: (Attach the snapshot and briefly write the report for the same)

Coding Challenges Details: (Attach the snapshot and briefly write the report for the same)

1. Online test Details

2. **Certification Course Details**

greatlearning
Learning for Life

Certificate of completion

Presented to

Prasanna

For successfully completing a free online course

Introduction to Hadoop

Provided by
Great Learning Academy
(On June 2020)

To verify this certificate visit verify.greatlearning.in/SARAGTXX

Introduction to Hadoop :

Hadoop is an Apache open source framework written in java that allows distributed processing of large datasets across clusters of computers using simple programming models. The Hadoop framework application works in an environment that provides

distributed *storage* and *computation* across clusters of computers. Hadoop is designed to scale up from single server to thousands of machines, each offering local computation and storage.

.

How Does Hadoop Work?

It is quite expensive to build bigger servers with heavy configurations that handle large scale processing, but as an alternative, you can tie together many commodity computers with single-CPU, as a single functional distributed system and practically, the clustered machines can read the dataset in parallel and provide a much higher throughput. Moreover, it is cheaper than one high-end server. So this is the first motivational factor behind using Hadoop that it runs across clustered and low-cost machines.

Hadoop runs code across a cluster of computers. This process includes the following core tasks that Hadoop performs −

- Data is initially divided into directories and files. Files are divided into uniform sized blocks of 128M and 64M (preferably 128M).

- These files are then distributed across various cluster nodes for further processing.

- HDFS, being on top of the local file system, supervises the processing.

- Blocks are replicated for handling hardware failure.

- Checking that the code was executed successfully.

- Performing the sort that takes place between the map and reduce stages.

- Sending the sorted data to a certain computer.

- Writing the debugging logs for each job.

2) Coding Challenges:

1. *To find numbers which is divisible by another number*

Pgrm1:

```
# Take a list of numbers
my_list = [12, 65, 54, 39, 102, 339, 221,]


# use anonymous function to filter
result = list(filter(lambda x: (x % 13 == 0), my_list))


# display the result
print("Numbers divisible by 13 are",result)
```