DAILY ONLINE ACTIVITIES SUMMARY

Date:	**01-05-2020**	**Name:**	**PRASANNA**
Sem & Sec	**8th,B**	**USN:**	**4AL16CS068**

Online Test Summary

Subject	**SMS**		
Max. Marks	**60**	**Score**	**45**

Certification Course Summary

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

Coding Challenges

Problem Statement: prob1- ***To print number of vowels in the given string***

Status: Solved

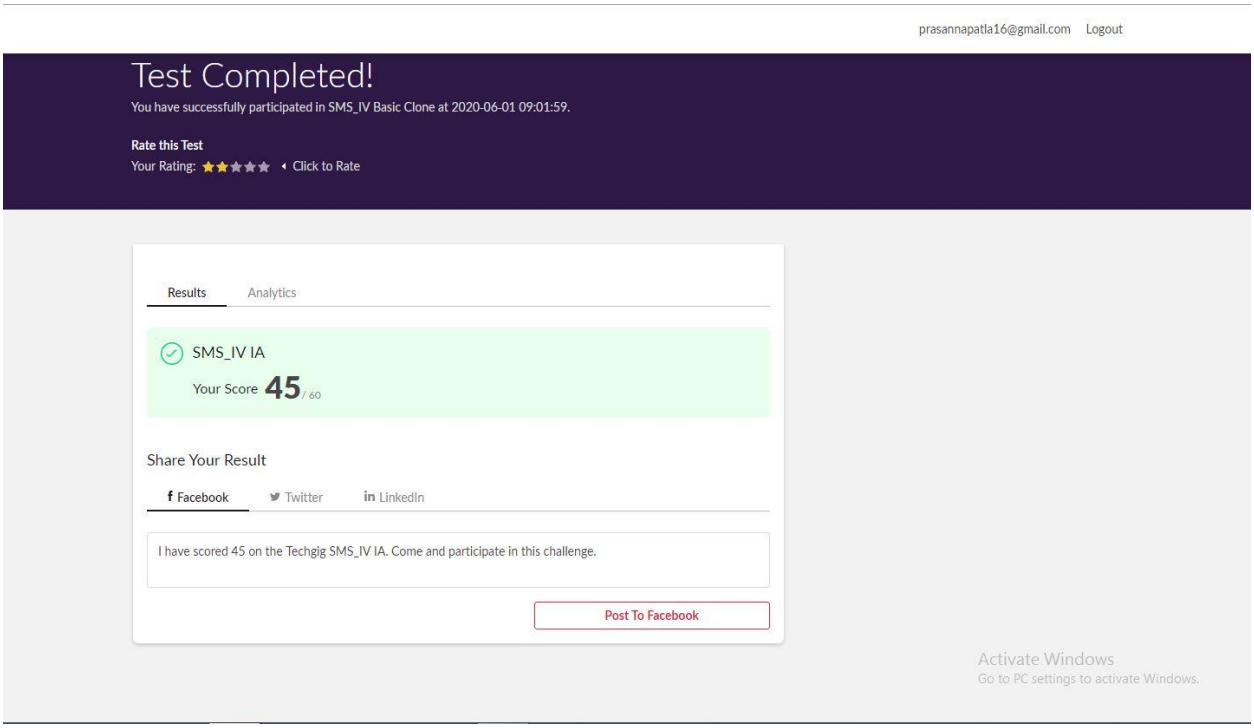
Uploaded the report in Github	**Yes**
If yes Repository name	**prasanna_p**
Uploaded the report in slack	**Yes**

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:

Introduction to Hadoop :

We don't have a straightforward definition to BigData. However, we will try to answer this question in different ways.

In Simple Words, Big Data is a technique to solve data problems that are not solvable using Traditional DataBases and Tools.

In other way, BigData means not just huge amount of Data. BigData means huge amount of data generating at very fast rate in different formats.

Big Data is a Technique to “Store, Process, Manage, Analysis and Report” a huge amount of variety data, at the required speed, and within the required time to allow Real-time Analysis and Reaction.

BigData is Data with has the following three characteristics:

- Extremely Large Volumes of Data
- Extremely High Velocity of Data
- Extremely Wide Variety of Data

the kernel based module This is a fully secured layer or the hardware layer which takes care of the entire Operating system. This also connects with the core OS layer + Linux kernel.

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don't know how to derive value out of it since this data is in its raw form or unstructured format.

Characteristics Of Big Data

(i) Volume – The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data. Hence, **'Volume'** is one characteristic which needs to be considered while dealing with Big Data.

(ii) Variety – The next aspect of Big Data is its **variety**.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

(iii) Velocity – The term **'velocity'** refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.

(iv) Variability – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

Benefits of Big Data Processing

Ability to process Big Data brings in multiple benefits, such as-

- Businesses can utilize outside intelligence while taking decisions

Access to social data from search engines and sites like facebook, twitter are enabling organizations to fine tune their business strategies.

- Improved customer service

Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

- Early identification of risk to the product/services, if any
- Better operational efficiency

Big Data technologies can be used for creating a staging area or landing zone for new data before identifying what data should be moved to the data warehouse. In addition, such integration of Big Data technologies and data warehouse helps an organization to offload infrequently accessed data.

2) Coding Challenges:

1. To display the count of vowels in a sring

Pgrm1:

```
string=input("Enter string:")
vowels=0
for i in string:
    if(i=='a' or i=='e' or i=='i' or i=='o' or i=='u' or i=='A' or i=='E' or i=='I' or i=='O' or i=='U'):
        vowels=vowels+1
print("Number of vowels are:")
print(vowels)
```