System Requirements Specification Index

For

Pyspark Usecase

Football player analysis L1

1.0

IIHT Pvt. Ltd.

Step to access the work environment

Step 1 use the URL to login provide the username and password

Step 2 you can see the launch Environment

PySpark-Football Analysis-L1-Solution-Reborn Test

Submit Assessment

1 Question Type: Project Based Question

Marks - 10

PySpark-Football Analysis-L1-Solution-Reborn Test

Provisioning VM ...

Please wait. It will take 5 to 10 minutes to provision your environment. Kindly don't refresh the page
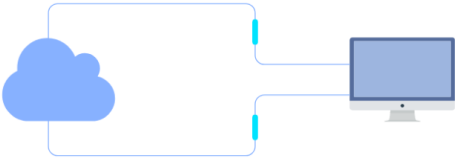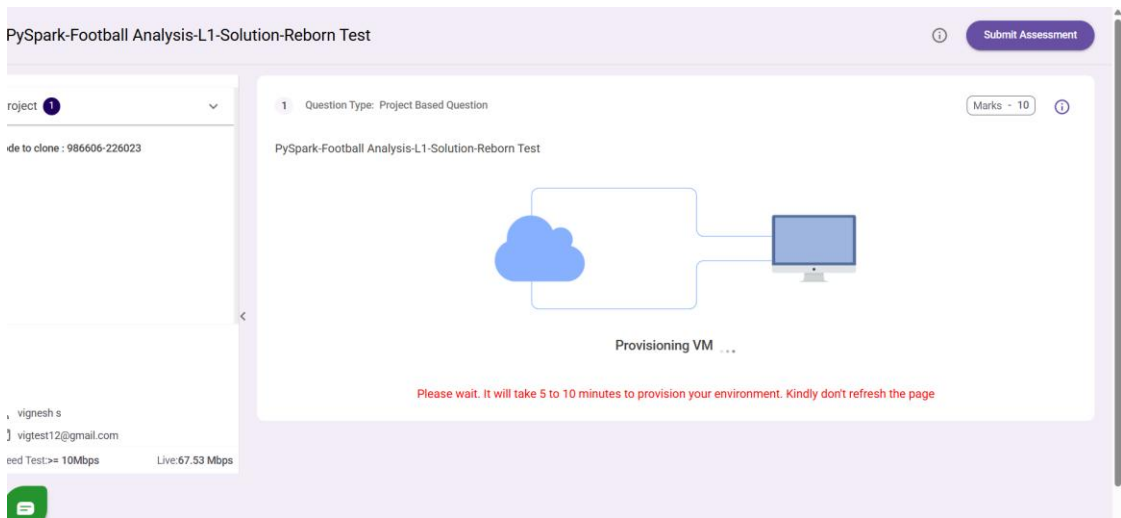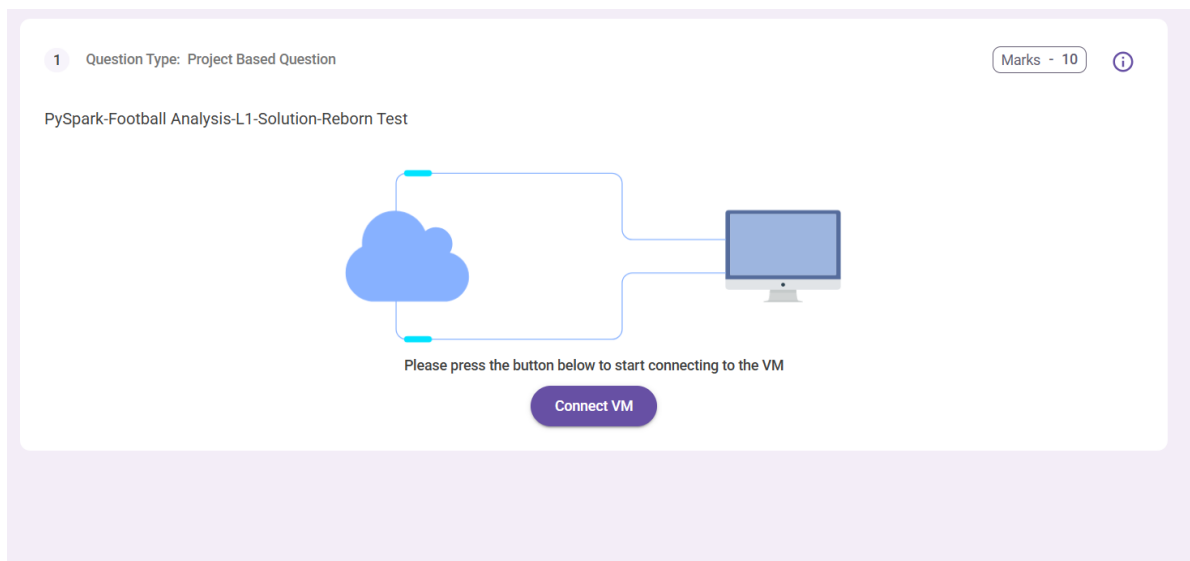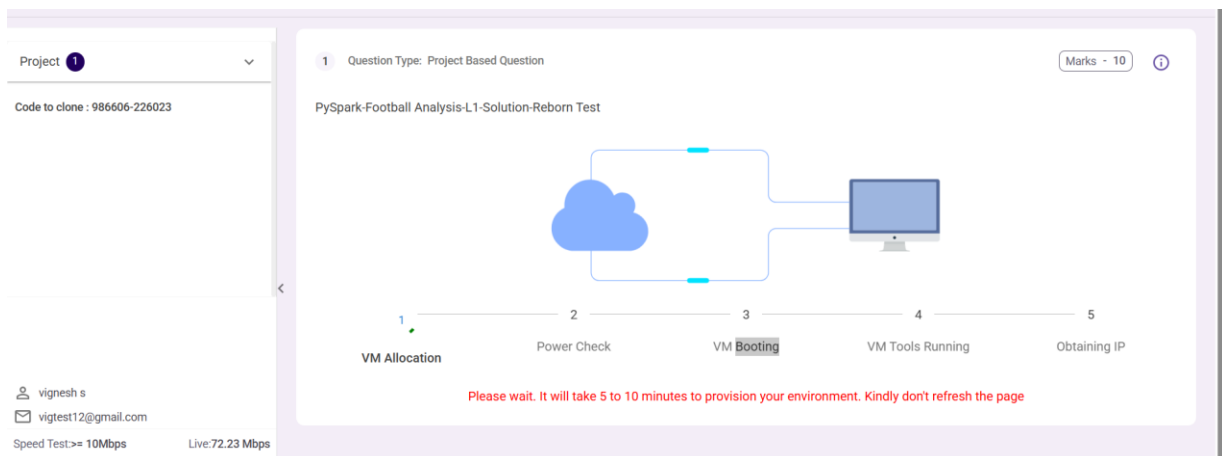
roject 1

de to clone : 986606-226023

, vignesh s
] vigtest12@gmail.com
eed Test:>= 10Mbps   Live:67.53 Mbps

Step 3 Click on the connect vm button

1 Question Type: Project Based Question

Marks - 10

PySpark-Football Analysis-L1-Solution-Reborn Test

Please press the button below to start connecting to the VM

Connect VM

Step 4 you will get a window you need to type the code from that left conner

Project 1

Code to clone : 986606-226023

1 Question Type: Project Based Question

Marks - 10

PySpark-Football Analysis-L1-Solution-Reborn Test

1   2   3   4   5
VM Allocation   Power Check   VM Booting   VM Tools Running   Obtaining IP

Please wait. It will take 5 to 10 minutes to provision your environment. Kindly don't refresh the page

vignesh s
vigtest12@gmail.com
Speed Test:>= 10Mbps   Live:72.23 Mbps
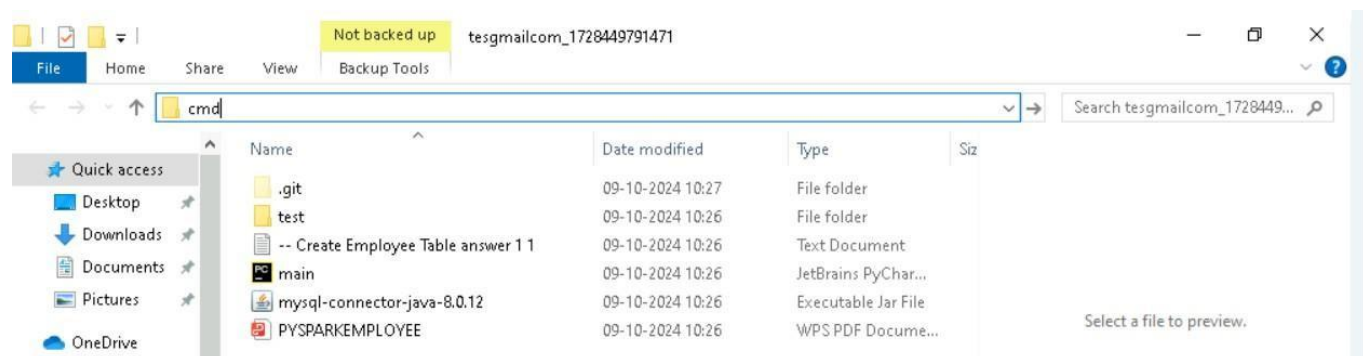
Click on ok

Step 5 after few seconds we can see that the your folder is cloned in the desktop .
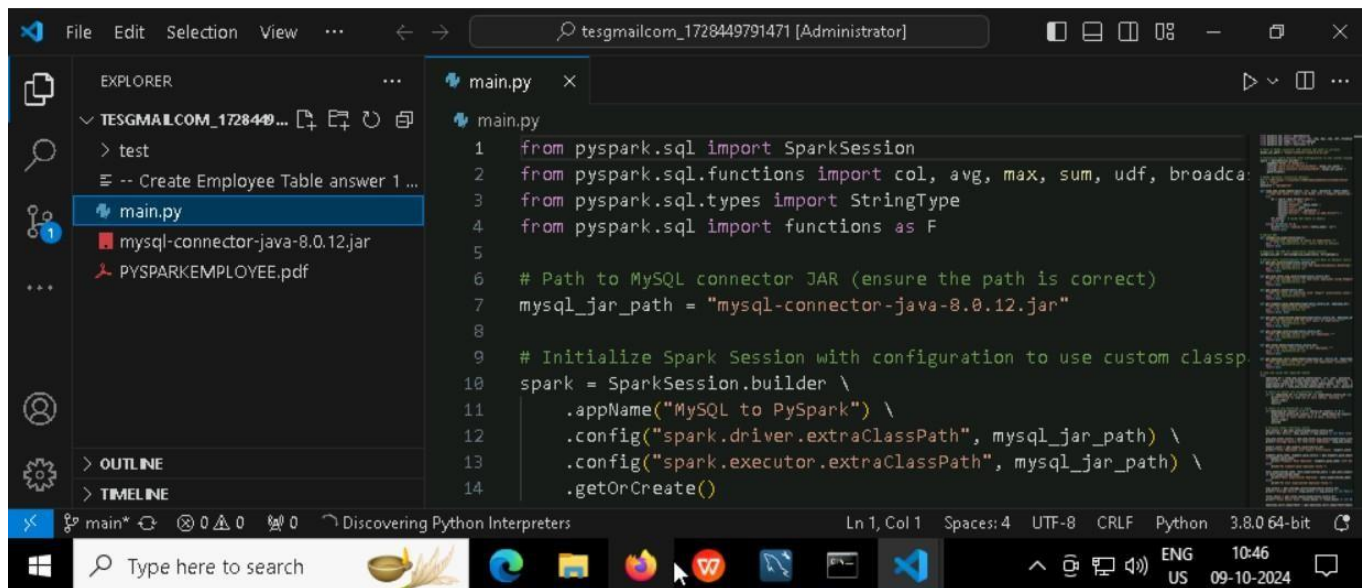
Folder cloned
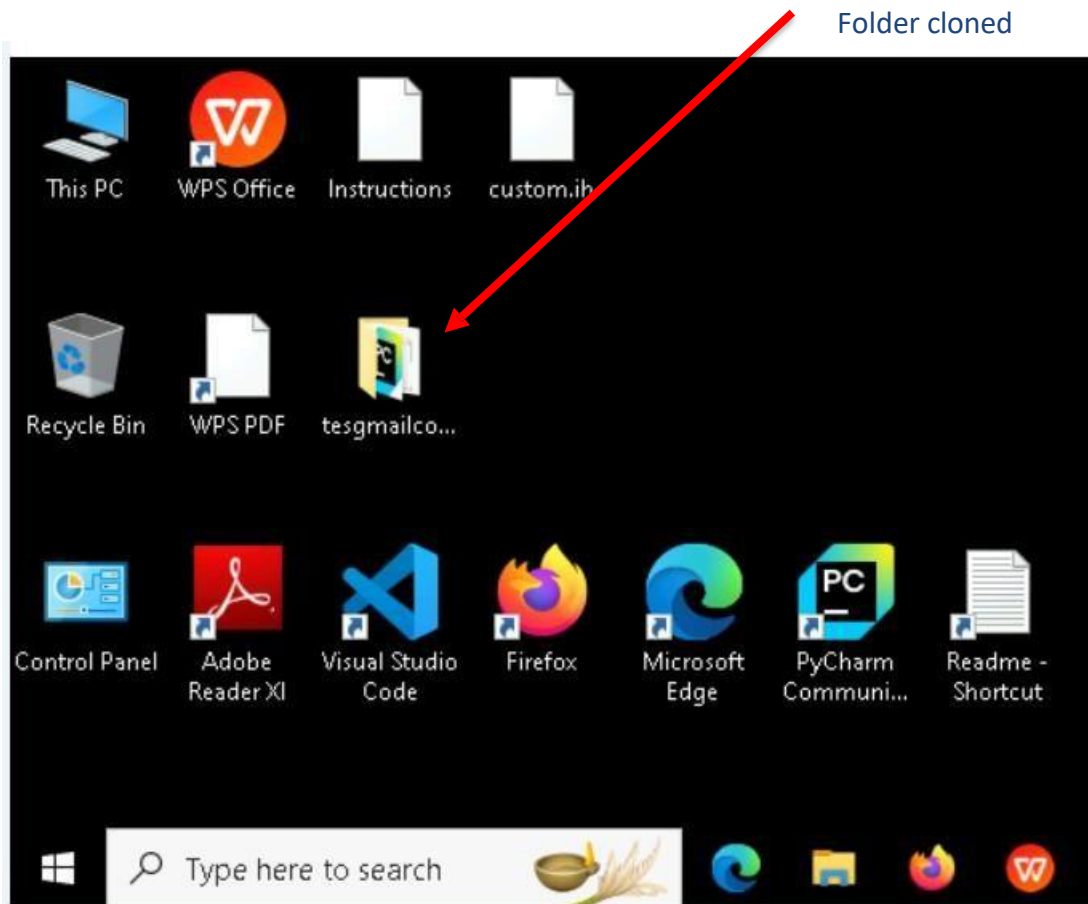
Step 6 go inside the folder type cmd in the top of the file explorer

- Type code. And hit enter you can see that workspace is opened in the visual code

- You can see that workspace is ready to code

> **Note** Please only work with visual code not with any other IDE

- In the folder cloned you will have all the project files needed .

Problem Statement   : **Football analysis with pyspark**

Description   : Use relevant methods operations toperform specified activities which are given in the instructions.

PYSPARK TASK -L1

The "Football Player Performance Analysis" use case involves football analysts using a PySpark-based program to extract key insights from a dataset of football players. Analysts load the dataset, which contains player statistics such as goals scored, assists, and matches played, into a Spark DataFrame. The program processes the data to identify top performers in several categories, including the top goal scorer, the player with the most assists, and the player with the best goal-to-match ratio. It also calculates the total number of goals and assists across all players, identifies the team with the most goals, and finds the players who have played the least and most matches. These insights are used by football clubs for tactical decisions, media outlets for reporting, and analysts for scouting. The results are compiled into reports, helping stakeholders make data-driven decisions on team strategies, player transfers, and performance evaluations.

Solve these Questions

     a. What is the name of the player who scored the most goals in the dataset?
     b. Which player provided the most assists in the dataset?
     c. Who has the highest goal-to-match ratio among the players?

     d. What is the total number of goals scored by all players in the dataset?

     e. What is the total number of assists provided by all players in the dataset?

     f. Which team scored the most goals in total based on the dataset?

Execution Steps to Follow:

1. All actions like build, compile, running application,running test

cases will be through Command Terminal.

2. To open the command terminal the test takers, need to go to

Application menu(Three horizontal lines at left top) -> Terminal -> New

Terminal

3. This editor Auto Saves the code

4. If you want to exit(logout) and continue the coding later anytime (using
Save & Exit option on Assessment Landing Page) internal git/repository. Else
the code will not be available in the next login.

5. These are time bound assessments the timer would stop if you logout

and while logging in back using the same credentials the timer would

resume from the sametime it was stopped from the previous logout.

6. To setup environment:

You can run the application without importing any packages

7. To launch application:

   Python football.py

8. To run Test cases:

python -m unittest

Screen shot to run the program
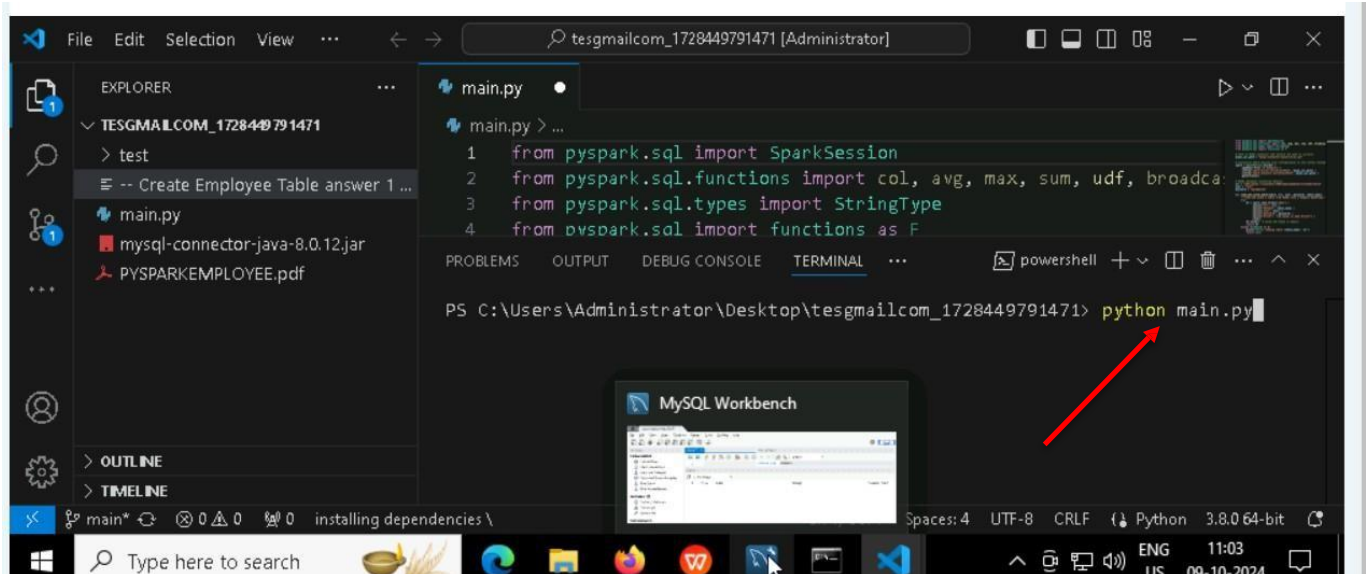
To run the application

- **Python football.py**
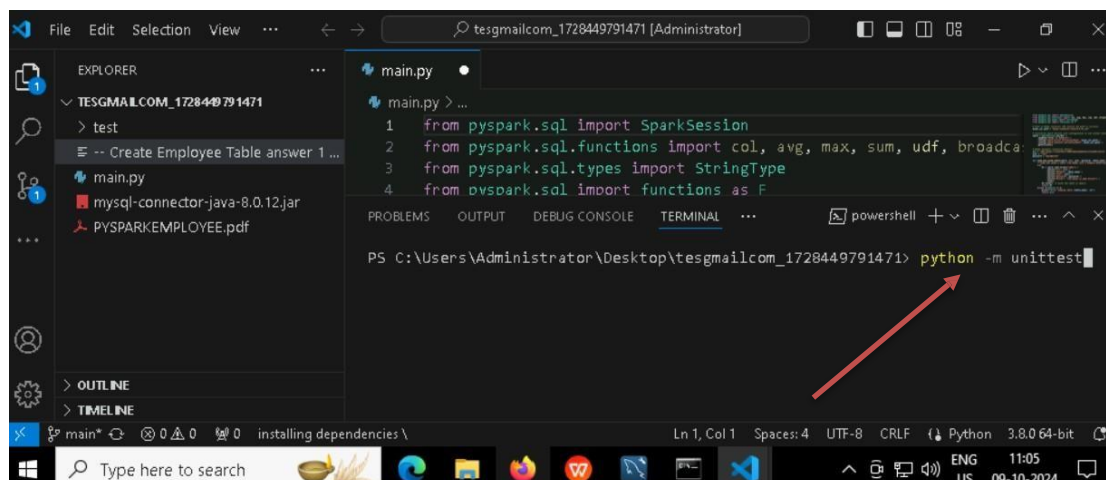
To run the testcase

- **Python -m unittest**

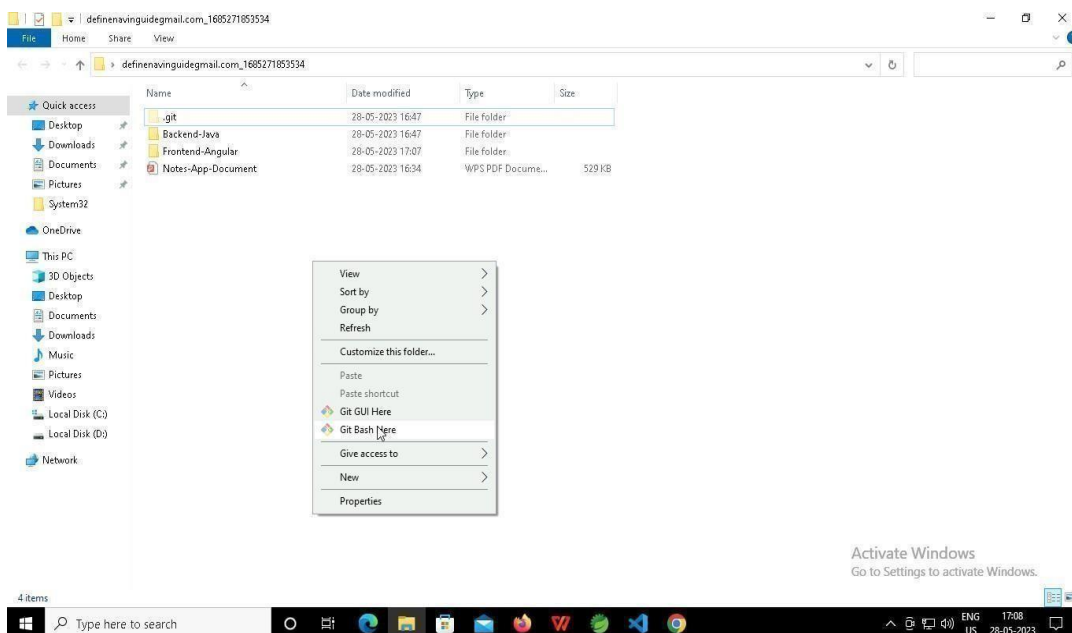## Screenshot to push the application to github

-----X-----

> **You can run test cases as many numbers of times and at any stage of Development, to check howmany test cases are passed/failed and accordingly refactor your code.**
> 1. **Make sure before final submission you commit all changes to git**. For that open theproject folder available on desktop
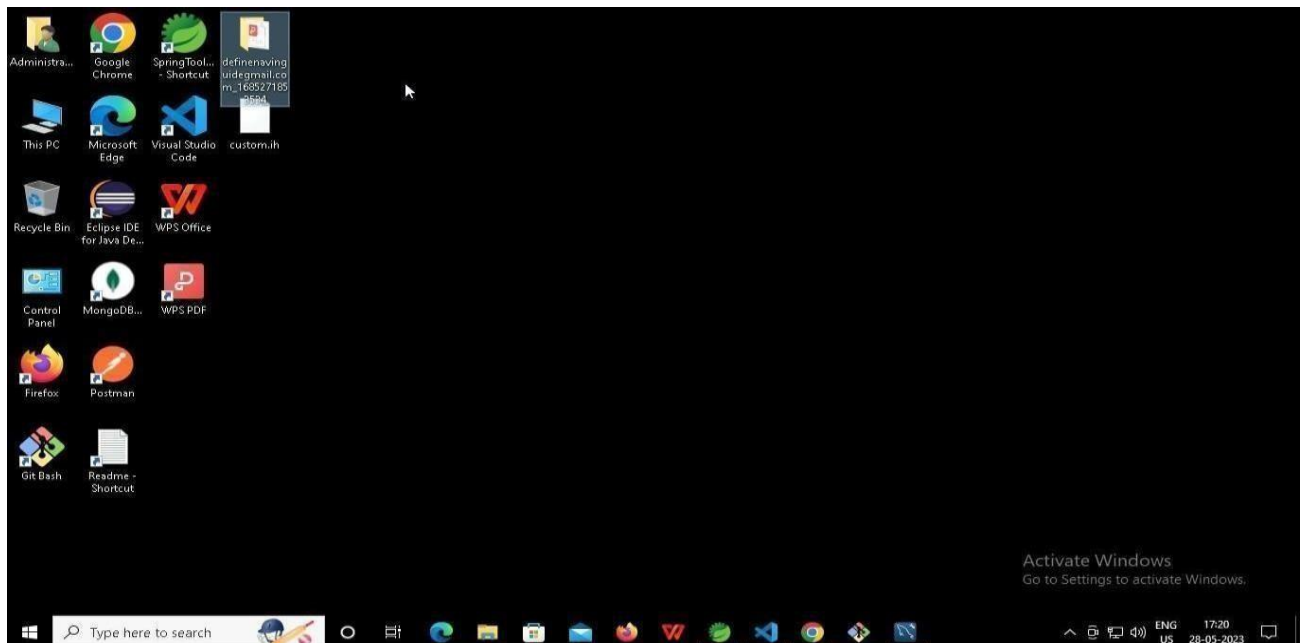
a. Right click in folder and open Git Bash

b. In Git bash terminal, run following commands

c. git status

d. **git add .**

e. git commit -m "First commit"
(You can provide any message every time you commit)

f. **git push**

After you have pushed your code Finally click on the final submission button

You should see a screen like this you will have to wait for the results . after getting this page you can leave the system

-----X-----