Hadoop - Assignment 2

GUIDO VAN DIJK   
IT2B  
648539

GitHub Link: <https://github.com/guidovdijk/hadoop-assesment/tree/Assignment-2>

17/05/2022

# Steps to take to execute:

The steps assume that the user already has the image shared on Moodle running on his VirtualBox and has Pig running.

Structure  
| …  
| - user/maria\_dev/  
| - diplomacy  
| - orders.csv

**Steps:**

1. Start the Virtual Environment
2. Start up the localhost (<http://127.0.0.1:8888>)
3. Click Launch Dashboard and login
4. Navigate to “File View”
5. Add the “diplomacy” folder in '/user/maria\_dev/ '
6. Navigate to “Pig View”
7. Create a new Pig file
8. Copy the code from the repo’s pig file in the newly created one
9. Check “Execute on Tez” and click “Execute”

# Steps of the code

Steps in the code are shown in the file itself. Can also be read in the repository: <https://github.com/guidovdijk/hadoop-assesment/blob/Assignment-2/Assignment%202.pig>

# Result of the code

*Raw .txt file can be found in the repository to check all the results*

Graphical user interface, text, application

Description automatically generated