**Luís Paulo Lima de Brito**

**Exercício HDFS**

1. Iniciar o cluster de Big Data

**$ docker-compose up -d**

2. Acessar o container do namenode.

**$ docker exec -it namenode bash**

3. Baixar os dados dos exercícios do treinamento na pasta home/input

Efetuar o Download do csv em:

**“curl –O** [**https://raw.githubusercontent.com/caiuafranca/dados\_curso/main/alunos.csv**](https://raw.githubusercontent.com/caiuafranca/dados_curso/main/alunos.csv)**”**

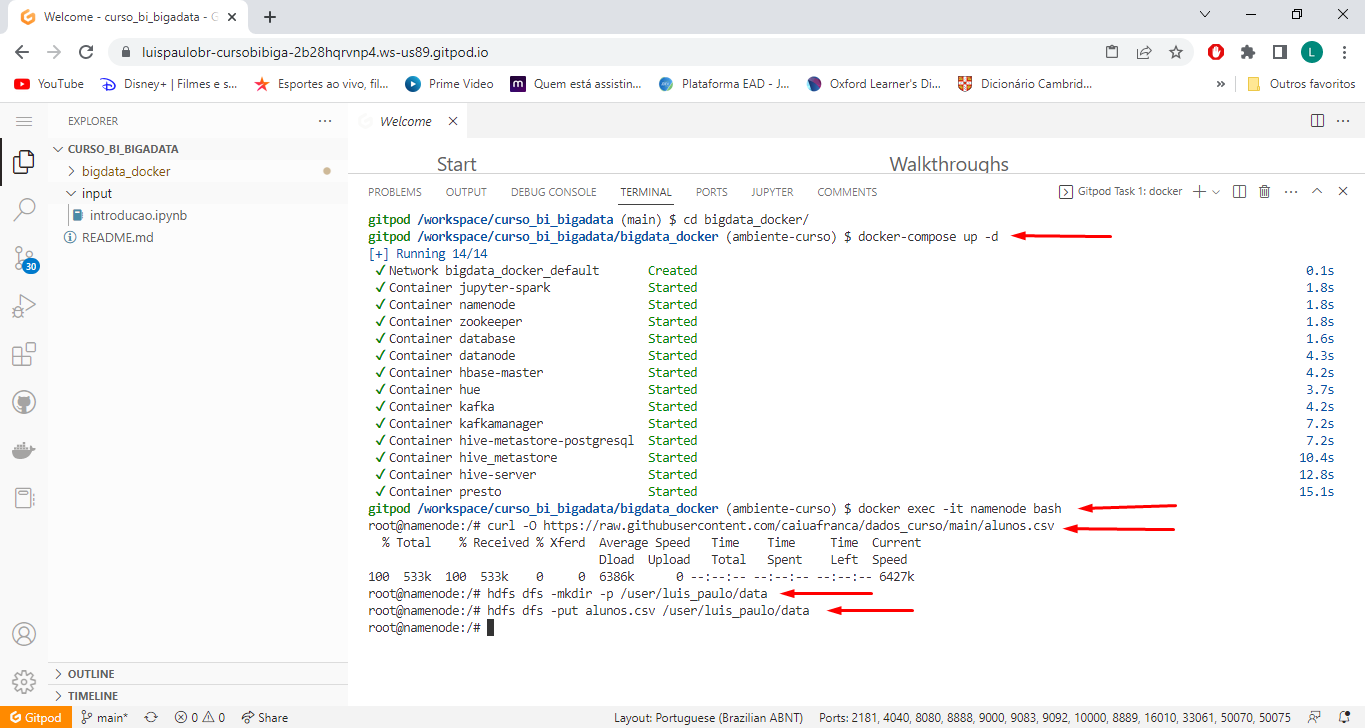
4. Criar a estrutura de pastas no HDFS como apresentada a baixo pelo comando:

- /user/<SEU NOME>/data

**hdfs dfs -mkdir -p /user/luis\_paulo/data**

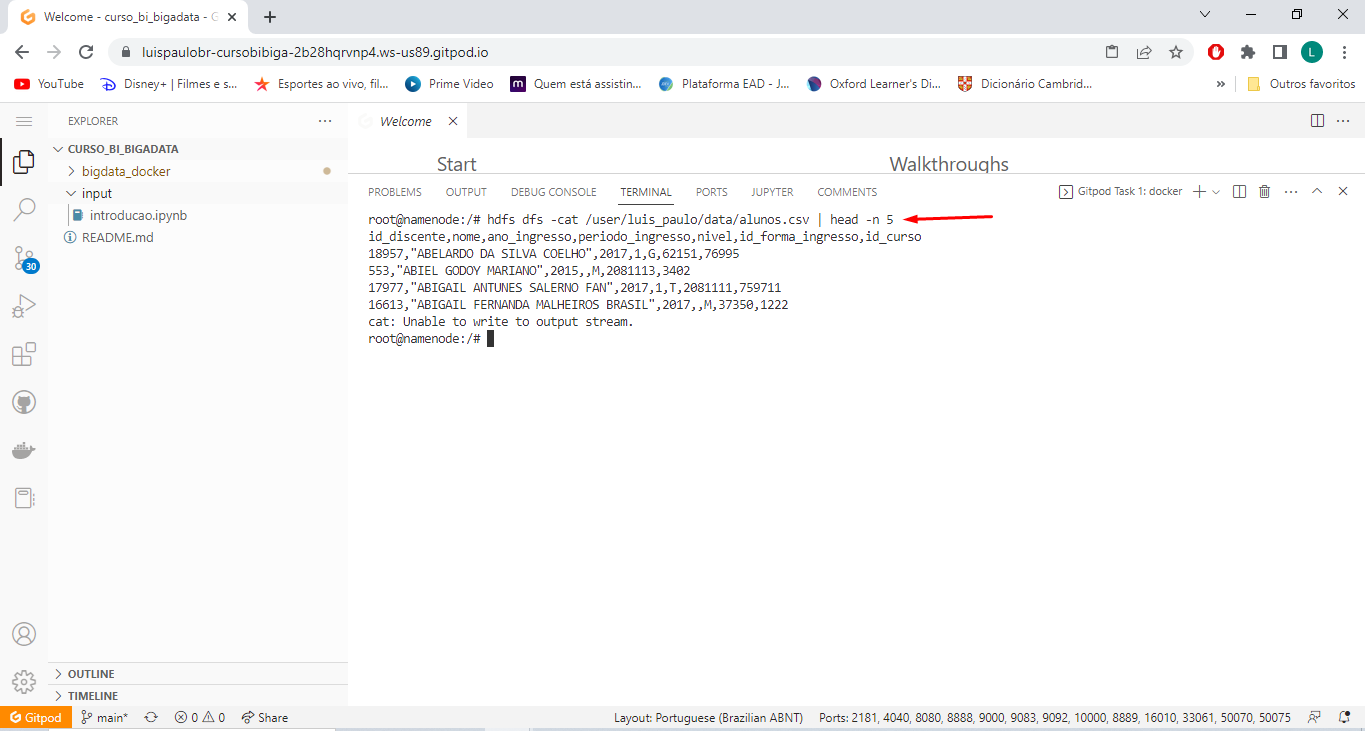
5. Enviar arquivo “/input/alunos.csv” para o diretório criado

**hdfs dfs -put alunos.csv /user/luis\_paulo/data**



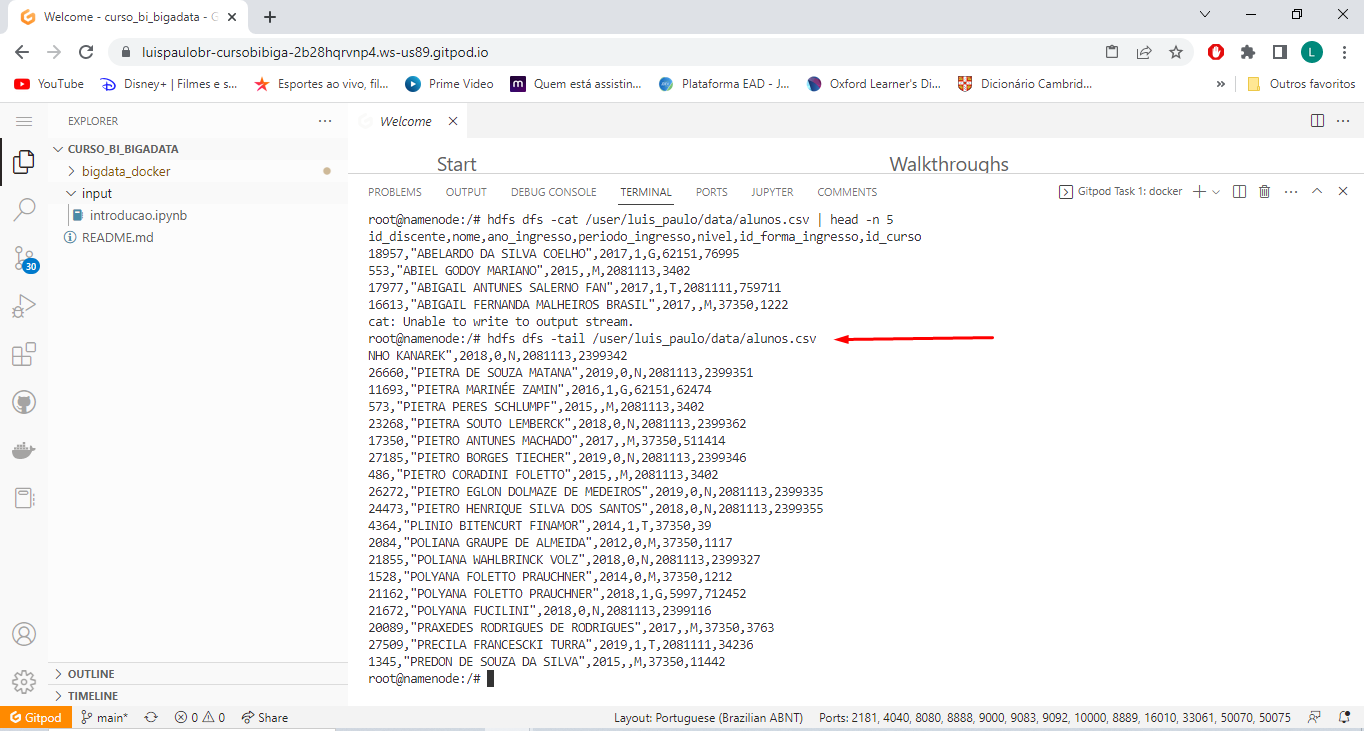
6. Verificar as 5 primeiras linhas do arquivo “alunos.csv”

**hdfs dfs -cat /user/luis\_paulo/data/alunos.csv | head -n 5**



7. Mostrar ultimas linhas do arquivo “alunos.csv”

**hdfs dfs -tail /user/luis\_paulo/data/alunos.csv**



8. Contar linhas do arquivo “alunos.csv”

**hdfs dfs -cat /user/luis\_paulo/data/alunos.csv | wc -l**

9. Criar um arquivo em branco com o nome de “teste.txt” em /user/<SEU NOME>/data

**hdfs dfs -touchz /user/luis\_paulo/data/teste.txt**

10. Apagar arquivo “teste.txt”

**hdfs dfs -rm /user/luis\_paulo/data/teste.txt**

11. Exibir o espaço livre e o uso do disco

**hdfs dfs -df -h /**

