

2/start-all.sh

Jps

3/hdfs dfs -mkdir /monTravail/test

4/ hdfs dfs -put /home/user1/movies.txt
/monTravail/test/Testven9

5/hdfs dfs -ls /monTravail/test/Testven9

6/

Mapper:

```
#!/usr/bin/env python
```

```
# -*- coding: utf-8 -*-
```

```
import sys
```

```
for line in sys.stdin:
```

```
    line = line.strip().split(',')
```

```
    genre=line[2]
```

```
    print ('%s,%s' % (genre, 1))
```

Reducer:

```
#!/usr/bin/env python
```

```
# -*- coding: utf-8 -*-
```

```
import sys
```

```
genre=None
```

```
count=1
```

```
for line in sys.stdin:
```

```
    line = line.strip().split(',')
```

```
    current_genre=line[0]
```

```
    if genre==current_genre:
```

```
        count+=1
```

```
    else:
```

```
        print('%s,%s' % (genre, count))
```

```
        count=1
```

```
    genre=current_genre
```

```
7/ hs /home/user1/test/mapper.py
```

```
/home/user1/test/reducer.py
```

```
/monTravail/test/testven9/ movies.txt
```

```
/monTravail/test/Resultat
```

