

PySpark Assignment:

Task1-

1. Go to command prompt on Windows.
2. In command line change directory to where spark is installed and then run command-
"bin\spark-submit Snehal_Shirgure_task1.py <ratings file> <outputfile>"
And hit enter
3. Output saved to given <outputfile> in ascending order.

Task2-

1. Go to command prompt on Windows.
2. In command line change directory to where spark is installed and then run command-
"bin\spark-submit Snehal_Shirgure_task2.py <ratings file> <tags file> <outputfile>"
And hit enter
3. Output saved to given <outputfile> as a folder and inside it there is a csv file part -xxx which contains final result in in descending order.