

## Windows 10

### Python Instructions

#### For running small task

1. Copy both the ratings and tags csv files from the small dataset into an execution folder.
2. Open command prompt in the current directory.
3. Type the following in the command prompt "python Anmol\_Chawla\_task1.py ratings.csv pt1s"
4. Press Ctrl+Z to exit the file and start afresh on the command prompt.
5. Type the following in the command prompt "python Anmol\_Chawla\_task2.py ratings.csv tags.csv pt2s"

#### For running big task

1. Copy both the ratings and tags csv files from the big dataset into an execution folder.
2. Open command prompt in the current directory.
3. Type the following in the command prompt "python Anmol\_Chawla\_task1.py ratings.csv pt1b"
4. Press Ctrl+Z to exit the file and start afresh on the command prompt.
5. Type the following in the command prompt "python Anmol\_Chawla\_task2.py ratings.csv tags.csv pt2b"

### Scala Instructions

#### For running small task

1. Copy both the ratings and tags csv files from the small dataset into an execution folder.
2. Open command prompt in the current directory.
3. Type the following in the command prompt "spark-submit --class task1 datamininghw1.jar ratings.csv st1s"
4. Press Ctrl+Z to exit the file and start afresh on the command prompt.
5. Type the following in the command prompt "spark-submit --class task2 datamininghw1.jar ratings.csv tags.csv st2s"

#### For running big task

1. Copy both the ratings and tags csv files from the small dataset into an execution folder.
2. Open command prompt in the current directory.
3. Type the following in the command prompt "spark-submit --class task1 datamininghw1.jar ratings.csv st1b"
4. Press Ctrl+Z to exit the file and start afresh on the command prompt.
5. Type the following in the command prompt "spark-submit --class task2 datamininghw1.jar ratings.csv tags.csv st2b"