

Film-Noir|10
Horror|11
Musical|12
Mystery|13
Romance|14
Sci-Fi|15
Thriller|16
War|17
Western|18

Source Data: <http://files.grouplens.org/datasets/movielens/ml-100k.zip>

Requirement

Use the movielens dataset as input and perform the following analysis:

Find the preference in movie genre based on profession:

Output required:

Profession, Genre, wa_rating

$Wa_rating = (\text{no ratings} > 3) / \text{total ratings}$

Output should be sorted by Profession and Genre in ascending order and wa_rating in descending order

The above analysis should be performed using Spark SQL Data Frame API.