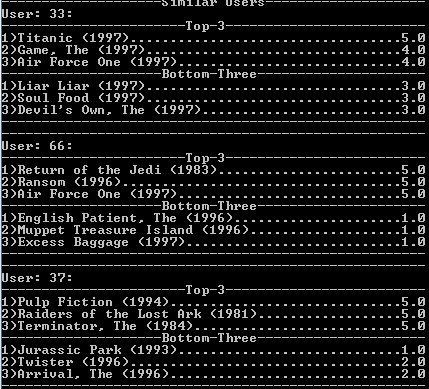
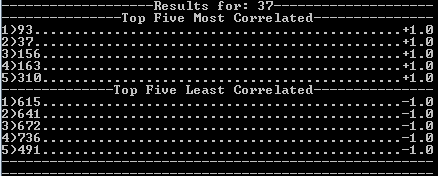
1. For this part of the assignment I used a slightly modified version on a code developed by Kevin Clemmons called data\_extractor.py and substitute\_you.py to create the resulting information below:

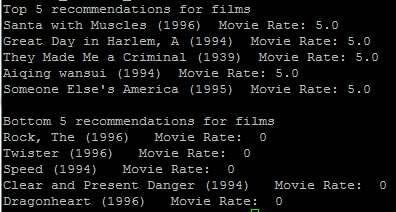


The data I used was my name “Matt”, my age “23”, my gender “M”, and my occupation “student” to get these results. I got a few more results, but chose these three. I chose User 37 as my substitute because I like all 3 of his top 3, though I did kind of like Jurassic Park even though they did not.

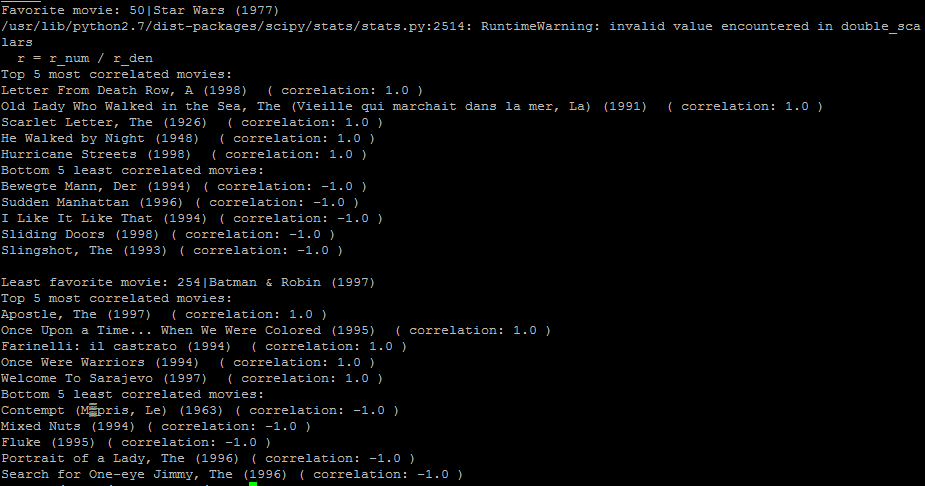
1. Using another slightly modified code developed by Kevin Clemmons called correlation.py I found users correlated most and least with User 37, and these were the results.



1. Using a slightly modified code developed by Chevelle Taylor-Sakyi called rate.py I found the movies that would be best and worst recommended for User 37 based on correlation.



1. Using another slightly modified code developed by Chevelle Taylor-Sakyi called like.py and the data given, I chose “Star Wars” (now known as Star Wars Episode IV) as my favorite from the list, and “Batman and Robin” as my least favorite. The resulting list based on correlation was a bit confusing admittedly.



I have never heard of any of the movies listed as most/least correlated for either film, except the Scarlet Letter and I’ve still never seen it. I can’t really say whether or not I like most of them because I’ve never seen any of them, but I’m pretty sure a good number of them are not like Star Wars.