Please complete the assigned problems to the best of your abilities. Ensure that the work you do is entirely your own, external resources are only used as permitted by the instructor, and all allowed sources are given proper credit for non-original content.

1. Practicum Problems

These problems will primarily reference the lecture materials and the examples given in class using Python. It is suggested that a Jupyter/IPython notebook be used for the programmatic components.

1.1 Problem 1

Load the Movielens 100k dataset (ml-100k.zip) into Python using Pandas data frames. Convert the ratings data into a utility matrix representation and find the 10 most similar users for user 1 based on the cosine similarity of the centered user ratings data. Based on the average of the ratings for item 508 from similar users, what is the expected rating for this item for user 1?

```
The 10 most similar users to user 1:
user_id
773    0.204792
868    0.202321
592    0.196592
880    0.195801
429    0.190661
276    0.187476
916    0.186358
222    0.182415
457    0.182253
8      0.180891
Name: 1, dtype: float64

The predicted rating for user 1 on item 508 is: 4.2116
```

1.2 Problem 2

Load the Movielens 100k dataset (ml-100k.zip) into Python using Pandas data frames. Build a user profile on centered data (by user rating) for both users 200 and 15, and calculate the cosine similarity and distance between the user's preferences and the item/movie 95. Which user would a recommender system suggest this movie to?

```
--- Analysis Results ---
Cosine similarity between user 200 and item 95: 0.0890
Euclidean distance between user 200 and item 95: 6.9553
Predicted rating for user 200: 4.1087

Cosine similarity between user 15 and item 95: 0.1305
Euclidean distance between user 15 and item 95: 7.8602
Predicted rating for user 15: 3.0492

The recommendation system should recommend item 95 to user 15
Reason: User 15 has higher cosine similarity (0.1305 > 0.0890)
```

-

END