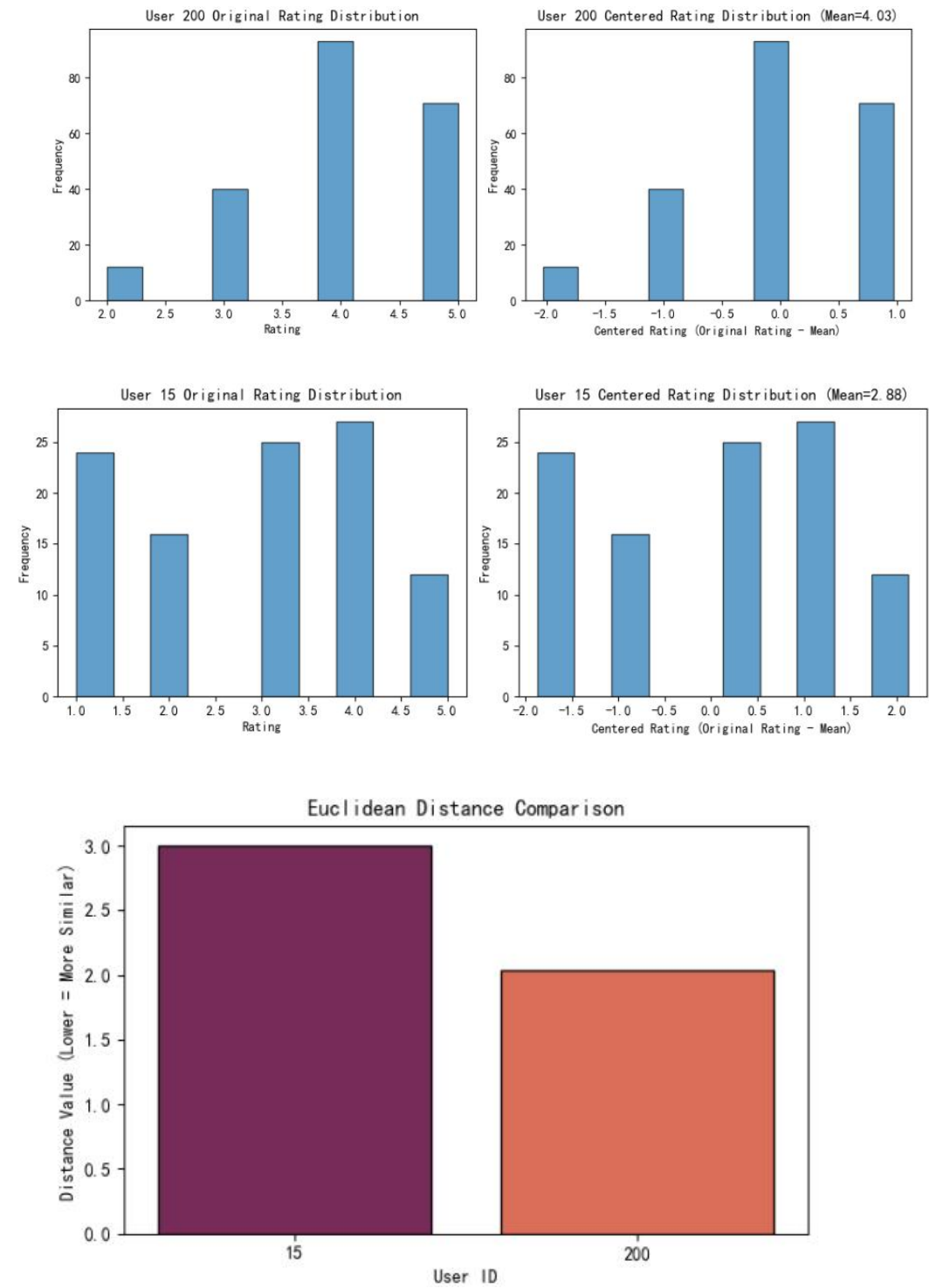


Q1.2

Cosine Similarity between User 200 and Movie 95: 0.00, Euclidean Distance: 2.03
Cosine Similarity between User 15 and Movie 95: 0.00, Euclidean Distance: 3.00



According to the analysis results of the Movielens 100k dataset, the cosine similarity and Euclidean distance between the centered ratings of user 200 and user 15 and movie 95 are as follows:

User 200: The cosine similarity is 0.00, and the Euclidean distance is 2.03.

User 15: The cosine similarity is 0.00, and the Euclidean distance is 3.00.

Interpretation of Similarity and Distance:

The closer the cosine similarity is to 1, the more similar the preferences are. The similarity of both users is 0, indicating that there is no directional consistency between their centered rating vectors and the rating pattern of Movie 95.

The smaller the Euclidean distance, the smaller the absolute difference. The distance of User 200 (2.03) is significantly smaller than that of User 15 (3.00), indicating that the difference between their ratings and the actual ratings of Movie 95 is smaller.

User Rating Distribution:

User 200's original average rating is 4.03 (tending towards high scores). After centering, the ratings are concentrated in the range of $[-1.0, 1.0]$, indicating that their evaluations of most movies are close to their own average level.

User 15's original average rating is 2.88 (tending towards low scores). After centering, the ratings fluctuate more significantly, and the differences in movie preferences are more obvious.

So the decision of the recommendation system to prioritize the recommendation of Movie 95 to User 200.