## Similarity score distribution on shuffled dataset

