Average Similarity Score Comparison across Models