Embedding Similarity Distributions by Continent and License Status