Jaccard Similarity Distribution using 100 signatures