Report

By: Tyler McElroy

In banking, the Office of Foreign Asset Control (OFAC) requires that banks keep track of Specially Designated Nationals (SDNs). SDN Lists contain entities that banks are not allowed to have as customers. Violating this requirement results in severe fines imposed against non-compliant banks.

The notebook associated with this report uses the OFAC SDN list as of 7/13/2025 and compares it against a synthetic list of bank customers. This mimics processes that real banks perform to ensure compliance with OFAC SDN lists.

To start, I imported libraries and data needed to perform this analysis. After that, I obtained summary statistics on both sets of data. Next, I identified rows with null values and remediated the missing values as needed. I identified 1 row in the OFAC SDN list that was null. This null row was a result of a CSV formatting error; as such, the row was removed.

To ensure names can be compared between datasets, preprocessing the names column was needed. I created a new column, Clean_Name, that ensured no special characters existed in either datasets name column. This allowed for easier comparison between datasets.

Comparing strings using "==" will only return true is the strings are exactly the same. This poses a challenge if the names are formatted differently, as is the case with the two datasets used in this analysis. To address the name columns having different naming formats, I used the library "rapidfuzz." This library performs optimized fuzzy matching to compare similar names. For instance, the names "John Smith" and "Smith, John" would return as exact matches using the token sort ratio method from the rapidfuzz library.

The token sort ratio method ignores word order to assess similarity between two strings. It does this by assigning a numeric value for the comparison between two strings. A 0 indicates no relation and a 100 indicates an exact match. The results of my analysis can be found below using 90 as the cutoff threshold value – that is, only comparisons with a score of 90 or above were returned for further investigation.

As for the results, Milan Ivanovic does match a value in the OFAC SDN list and would need to be addressed by the non-compliant bank. Further investigation would prove Jim Newton to be a false positive; as such, no action would be needed for him.

	Bank_Name	OFAC_Name	Score
0	Jim Newton	NEWTON ITM	90.0
1	Milan Ivanovic	IVANOVIC, Milan	100.0