**RTM/OSS Data Cleaning Functions:**

**Output location: ‘Output/Data\_Cleaning’**

**Outliers**

* **‘outliers’** main function to be used for removing outliers dataset based on z-score.
* The user inputs the variable on which the cleaning is to be performed e.g. Sales value column, and provides threshold z-score value, that is by default 1.96

**SKU Standardization**

* **‘sku\_standardize’** main function for standardization of SKU names/Store names and finding top three matches based on fuzzy score.
* The user has to input two lists of strings – list 1 will have reference names and list 2 will be used to find top three matches for each item in list 1.
* The fuzzy score is out of 100 where 100 means best match. Ranking as 1,2,3 is also given as output.