Task Order: 1

Date: 19 March 2018

Due Date: 26 March 2018

Subtask Description:

* Find all duplicates in ‘FullStreetAddress’
* Filter duplicates where ‘CloseDate’ is within 5 years or 1825 days of each other
* Change Boolean variables to TRUE = 1 and FALSE = 0
* For duplicates (matching ‘FullStreetAddress’) subtract numerical values (of all columns) between the older and newer date
* Return or print a dataframe with the original values of each row for the earliest duplicate and difference or delta for each following duplicate (see example\_df)

Additional Notes:

Visualization work to follow in task order 2.

Data File:

1\_jan\_17-7\_mar\_18.csv located at https://drive.google.com/open?id=1fmvEbfctjIzBF83OPO37SQG89Ms7Htw2