Description	Which Columns Changed	Changes
Import dataset	All	376812 lines imported
Convert to date time format then to ISO	REPORTED_DATE	376812 lines changed
Conversions for reported date to ISO	FIRST_OCCURRENCE_DATE	376812 lines changed
Stripping white space from multiple cols	OFFENSE_TYPE_ID, OFFENSE_CATEGORY_ID, INCIDENT_ADDRESS, NEIGHBORHOOD_ID	0 lines changed
Remove special characters from address	INCIDENT_ADDRESS	32305 lines changed
Dropped columns that will not be used or were updated	Previously changed cols that were updated	7 columns dropped
Manually replace abbreviated neighborhood names	NEIGHBORHOOD_ID	1 unique name replaced
Find number of missing neighborhoods	NEIGHBORHOOD_ID	584 missing neighborhoods
Calcualted haversine to find closest neighborhoods	CREATED nearest_neighborhood	582 added nearest neighborhoods
Dropped columns no longer used or not relevant	GEO_X	
	GEO_Y	
	DISTRICT_ID	
	nearest_neighborhood_dist_km	
	OBJECTID	
	OFFENSE_TYPE_ID	
	OFFENSE_CATEGORY_ID	
	IS_CRIME_y	
	IS_TRAFFIC_y	
	first_occurrence_date_dt	
	reported_date_dt	
	OFFENSE_TYPE_ID_trim	
	OFFENSE_CATEGORY_ID_trim	