import pandas as pd

import math

import numpy as np

import re

location = pd.read\_csv('alldata\_0222.csv')

location = location.fillna('NA')

location = location[['Summons Number','Street Name']]

def remove\_th(string):

string = string.upper()

pattern = re.compile(r'^(.\*)([1-9]\*)TH(.\*)')

m = pattern.match(string)

if(m):

return (m.group(1) + m.group(2) + m.group(3))

else:

return string

def remove\_co(string):

string = string.upper()

pattern = re.compile(r'^(.\*)C/O(.\*)')

m = pattern.match(string)

if(m):

return (m.group(1)+m.group(2))

else:

return string

def avenue\_street\_(string):

string = string.upper()

string = string.replace("AVENUE", "AVE")

string = string.replace("STREET", "ST")

string = string.replace("\_", " ")

string = ' '.join(string.split())

return string

location['Street Name'] = location['Street Name'].apply(remove\_th)

location['Street Name'] = location['Street Name'].apply(remove\_co)

location['Street Name'] = location['Street Name'].apply(avenue\_street\_)

location.to\_csv('location.csv',encoding = 'utf-8')

BLUE, RED, YELLOW, BLACK, GREEN, GREY, OTHER