Short Note

It is not possible to upload all the data here. Because we used all the crime records for several years in the UK as data, it was 7.7gb. Instead, we are uploading our "combined_crimes.csv" file, which we created after extracting London data from all this data. But this way it doesn't work from the beginning to the end, so if you are going to run it, you need to skip the section starting from # Combining Crime Outcome Files for Crime Probability Analysis to Calculating Crime Counts Around Houses Using Haversine Distance. So the first code after the jump is import pandas as pd
from sklearn.neighbors import BallTree
import numpy as np

#1) Upload crime and household data
df = pd.read_csv("combined_crimes.csv") # Crime data
houses = pd.read_csv("data.csv") # Must start as house data

I hope I have made my point.