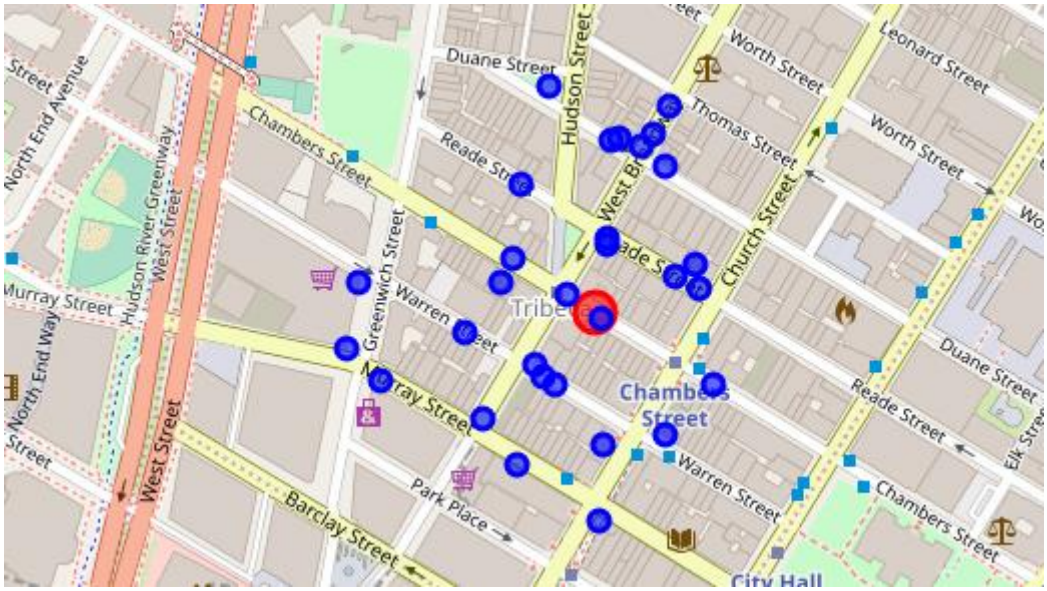
1. Background

We all have different choices when selecting an apartment. Price can be a one obvious measure that we consider when selecting an apartment. Also, some people have different neighborhood choices. This analysis trying to help to a person who considers the distance to the nearest GYM and well as the price of the apartment. We are trying to cluster the data to help our "Gym Goer" to find the best apartment which fits to him. "Gym Goer" plans to buy the apartment in the **next month**. We will be using data preprocessing, augmentation, analyzing techniques with cluster the neighborhood for out "Gym Goer" and finally visualize the clusters ranking the clusters to live. The data analyzing process will be a trial and error process considering the state-of-art clustering techniques (prioritizing the techniques followed throughout the specialization) applying to this specific problem that we have in the hand.



2. Data description

This link has New York NYC Property Sales data released by New York City department of finance. The Department of Finance's Rolling Sales files lists properties that sold in the last twelve-month period in New York City for all tax classes. These files include:

- the neighborhood;
- building type;
- square footage;
- other data.

Also, we use the places data from the foursquare API. Here we assume that the current places will be there when he buys the apartment in the next month.