Optimizing urban mobility

Niva Ran

Data Incubator Fellowship interview, 20200413

Step 1: What is the demand on the current docking stations?

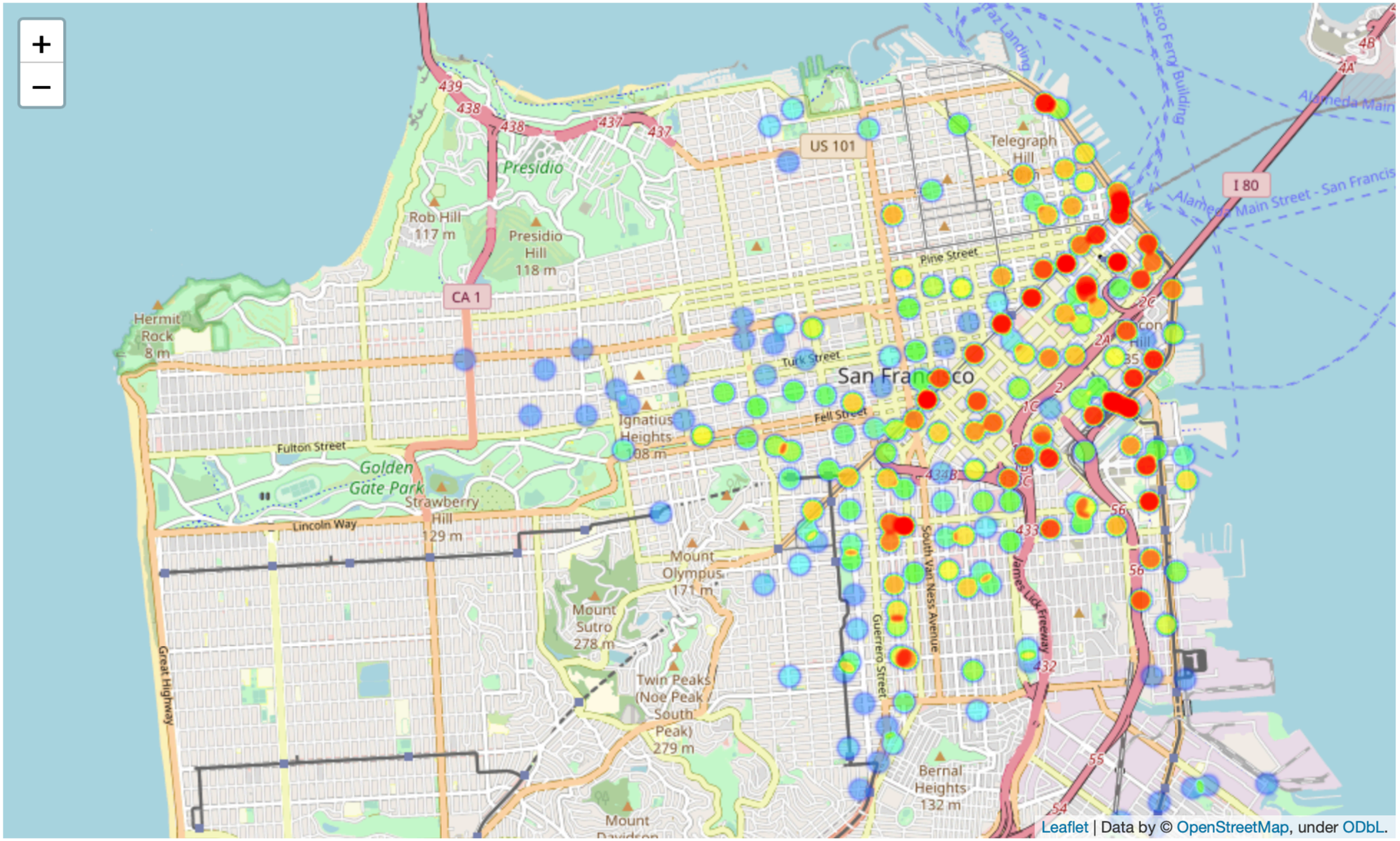
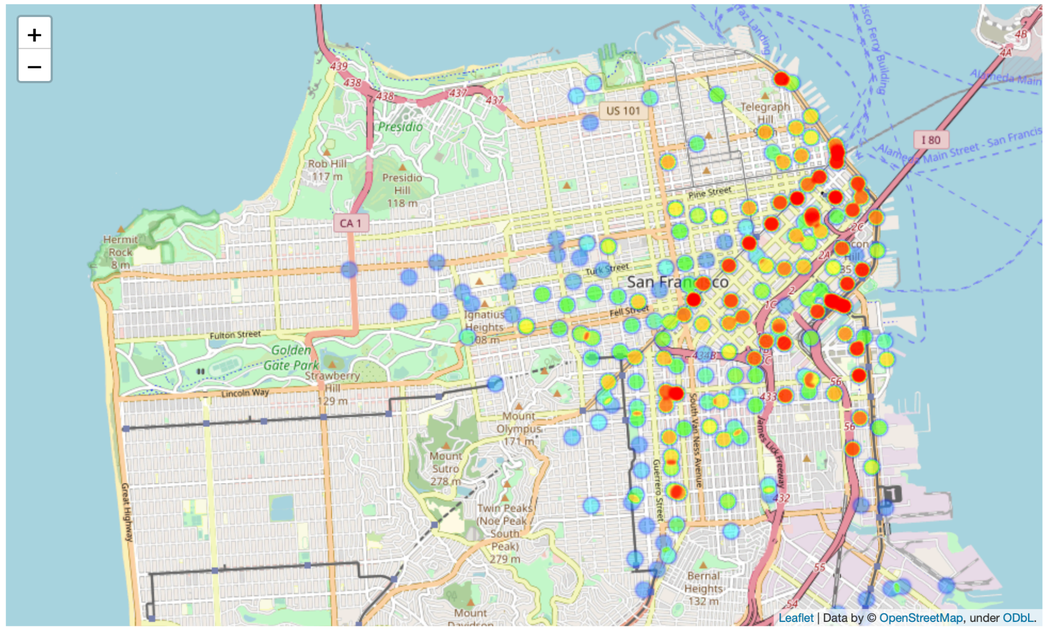
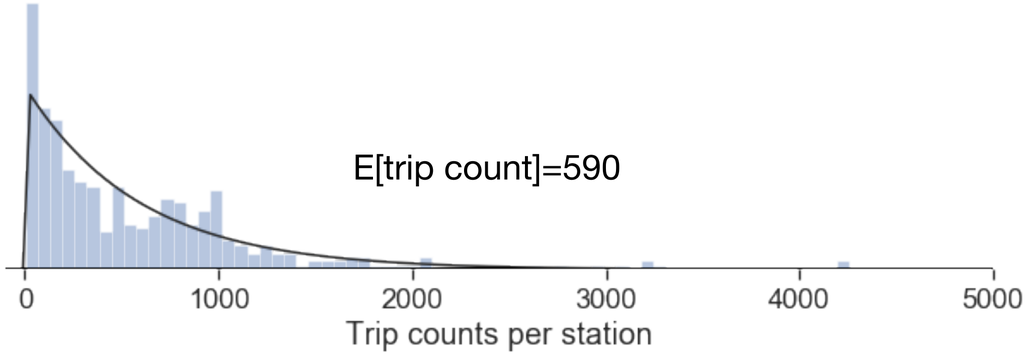


Figure 1. Relative demand on bike docking station by location in San Francisco. Demand is calculated from cumulative rides departing from a given station, over 6 months (before the COVID19 outbreak).

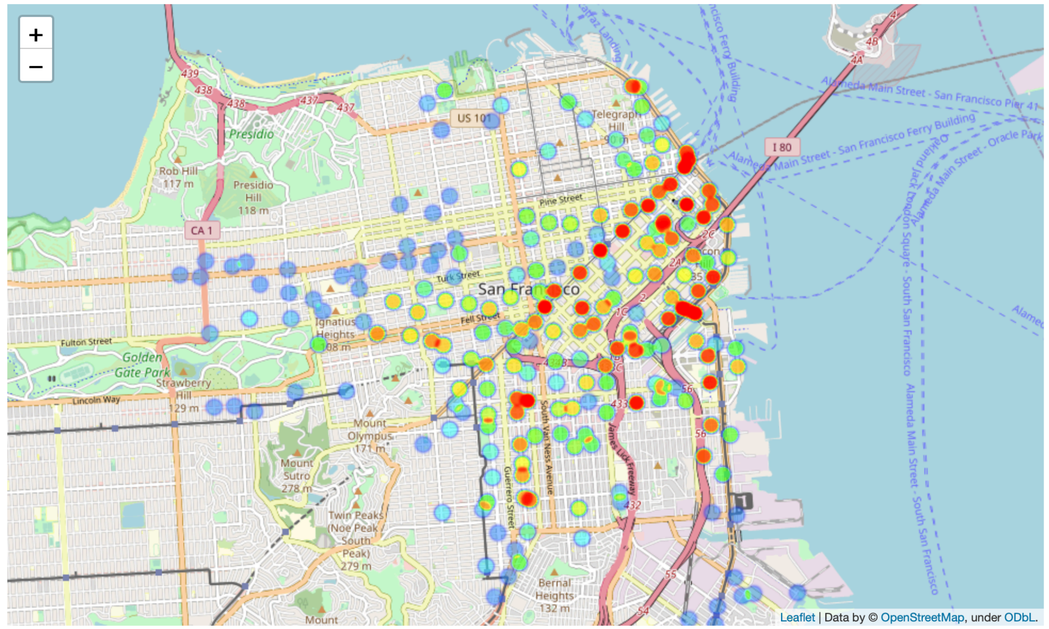
Step 1.A: What data should we use?

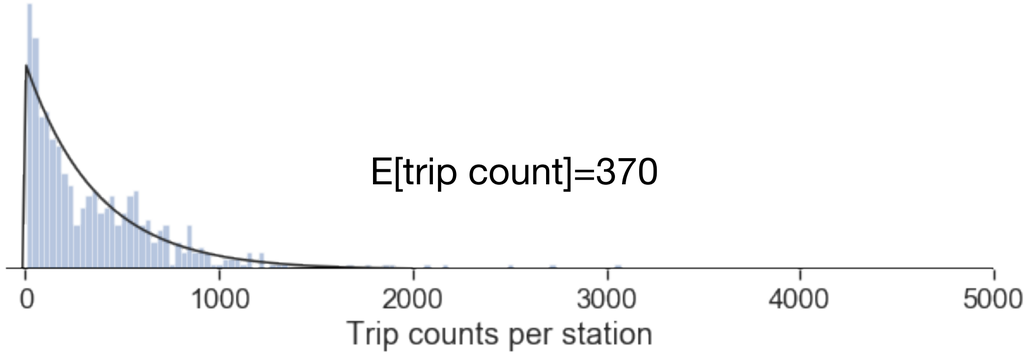
Before Jan 2020





After Jan 2020





Step 2: What is the estimated demand for bikes throughout San Francisco?

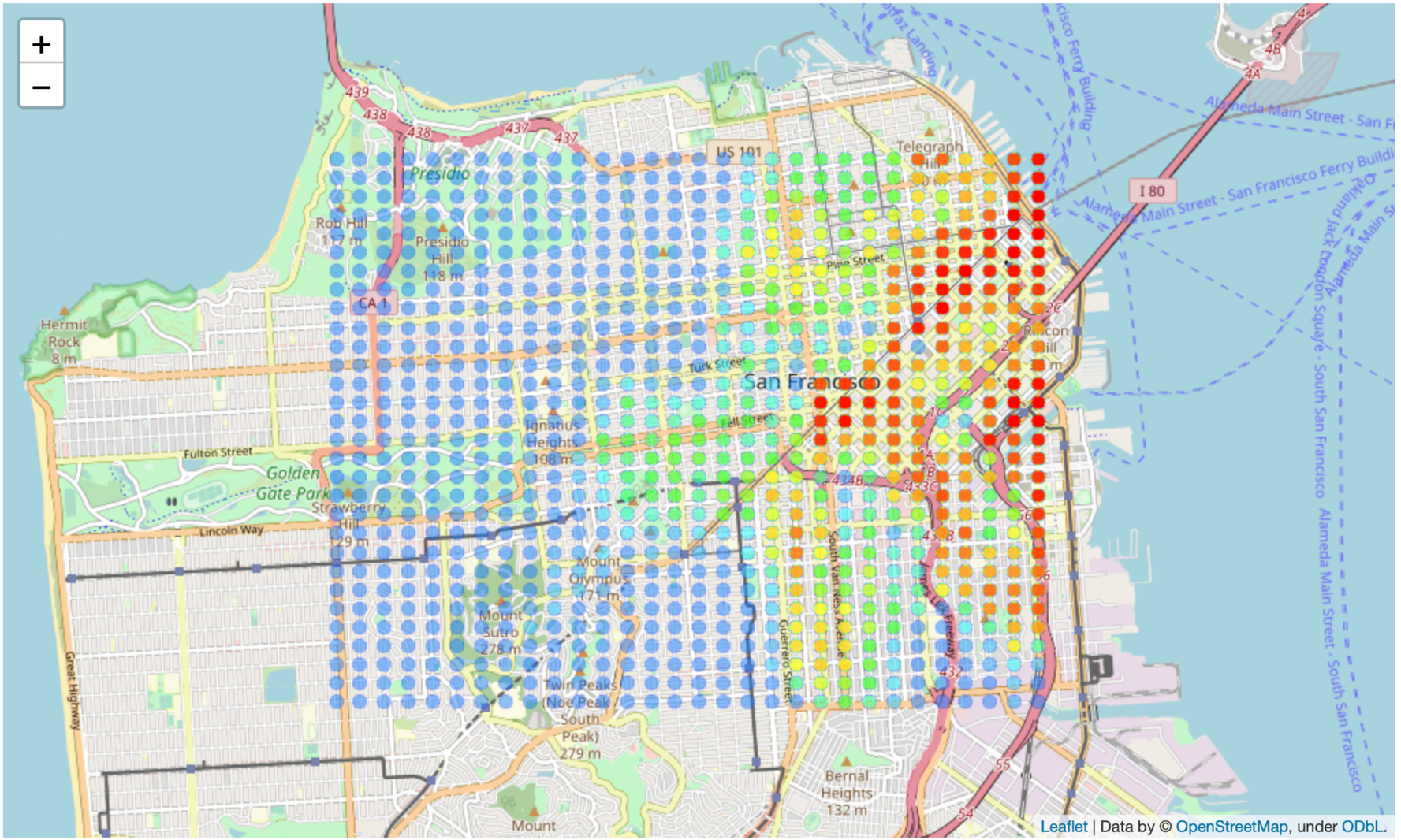
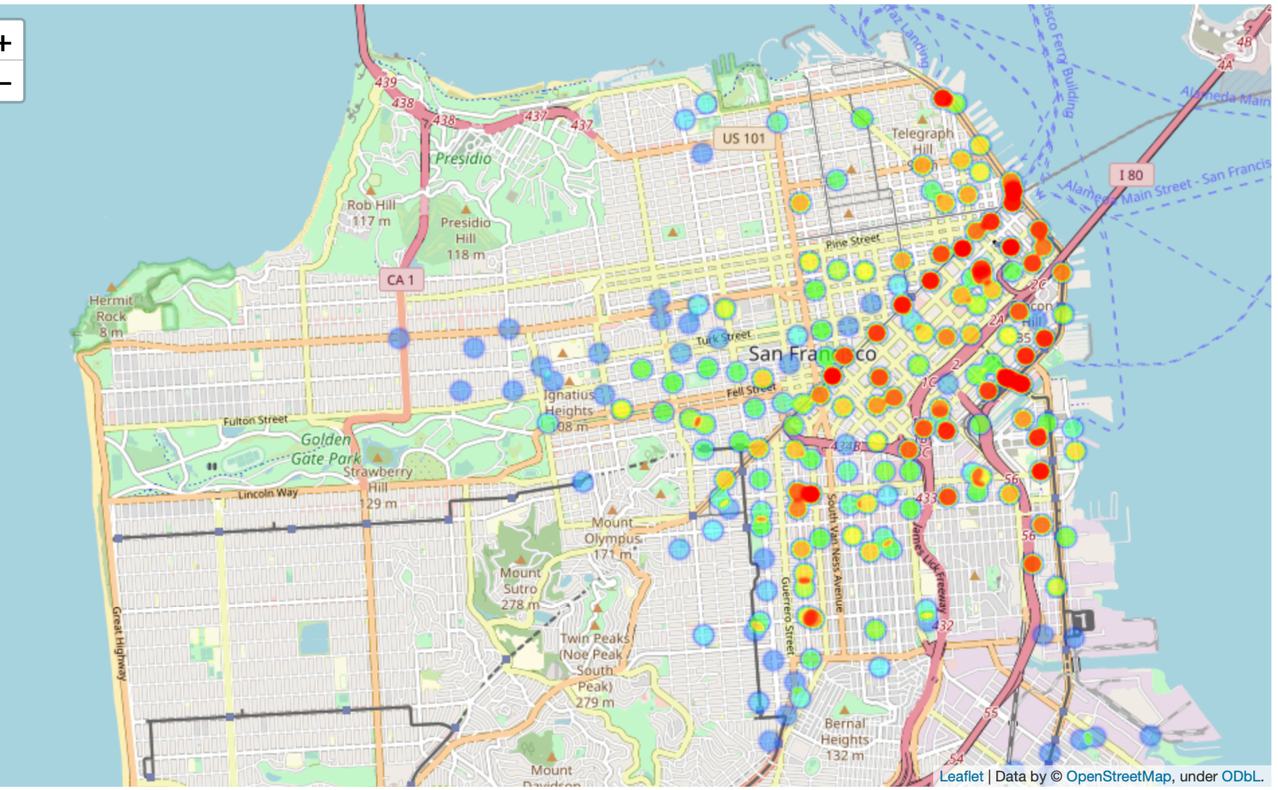


Figure 2. Predicted bike use demand across a specified area in San Francisco. Using the the currently known demand as a training set, the predicted demand is achieved with a k-neirest neighbor (kNN) algorithm.

Step 3: What would be the optimal bike station distribution

Existing bike distribution



Optimized bike distribution

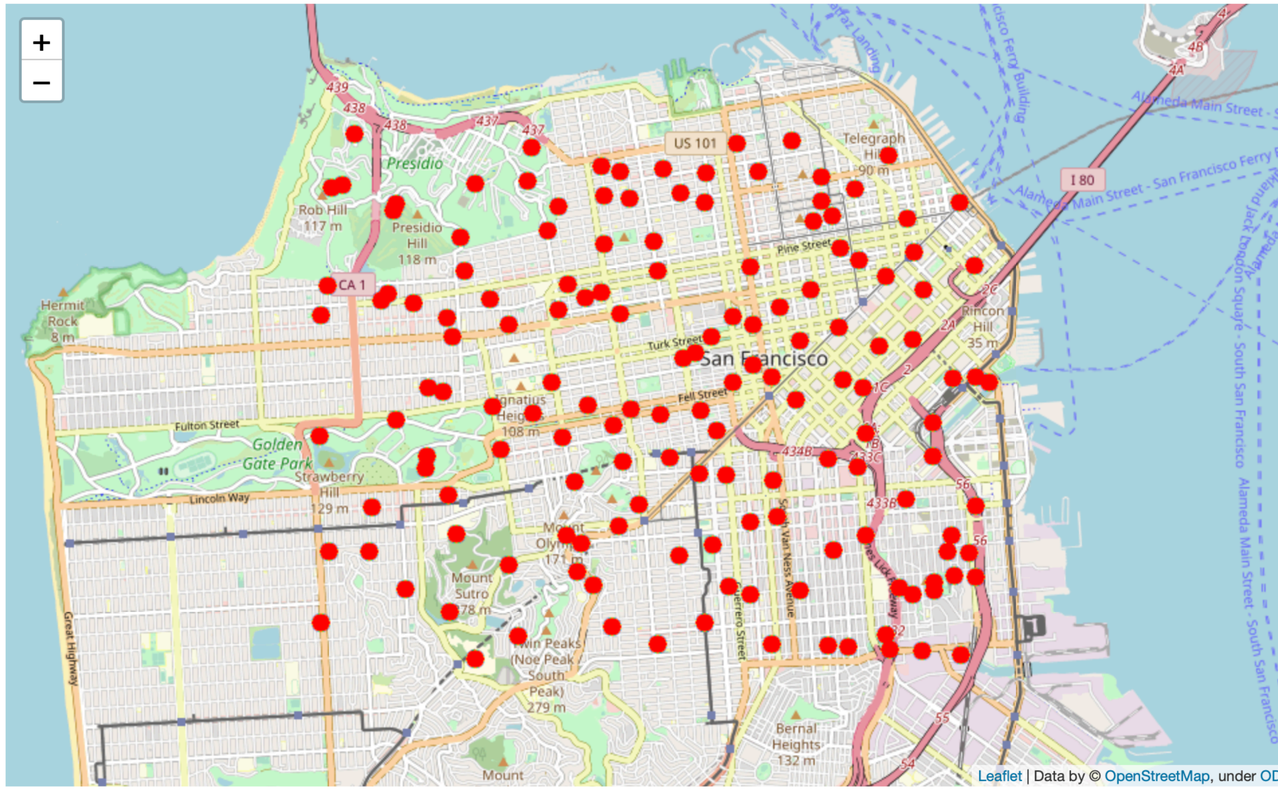


Figure 3. Current bike distribution vs. optimized bike distribution. Using the demand predicted by the kNN algorithm, for each unit of demand, cost is defined to be the Manhattan distance to the nearest station. Gradient descent was used to minis the cost by redistributing the bikes around San Francisco.