Deliverable 1

State Surplus Project

Actions Completed:

1. Separating the Original Dataset

The problem asks us to decide which land owned by municipal housing authorities and the Mass Department of Transportation (MASSDOT) is most attractive to develop affordable housing, and which land is most attractive for sale. Therefore our first step was to divide the dataset by owners' names and select parcels owned by housing authorities, cities and towns, MASSDOT, and related departments. However, many of the names listed in the dataset are inconsistent. For example, MBTA, MASS BAY TRANSPORT, M.B.T.A all refer to the same land owner. So, before dividing the dataset, we reformatted the owners' names to make them consistent, then we split the dataset into housing.csv, DOT.csv and city_and_town.csv.

2. Analyse Subsidized Housing Inventory by District

To establish the percentage of affordable housing of each district, we extracted data from the Department of Housing and Community Development Chapter 40B Subsidized Housing Inventory (SHI) and applied the data to final.data.csv to find out the percentage of affordable housing of each district. Here we each list five districts which respectively have the highest and lowest percentage of affordable housing. The whole dataset is in the file SHI_District.csv.

Table 1. The five districts with the lowest percentage affordability housing, which may have the greatest need for development

District	**Amount of Affordable Housing** *(%)*
1st Plymouth District	3.19
1st Franklin District	3.47
7th Worcester District	3.79
1st Worcester District	4.05
3rd Hampden District	4.34

Table 2. The top 5 districts which have the largest percentage of affordable housing

District	Amount of Affordable Housing *(%)*
5th Hampden District	19.88
2nd Suffolk District	19.28
11th Hampden District	16.56
10th Hampden District	16.56
21st Middlesex District	15.1

Looking at the file all_district_data.csv, we found that those which have the smallest percentages also have the largest amount of vacant land areas to develop, which means they can benefit more from the project.

3. Analyse Subsidized Housing Inventory by City

The data was also analysed to determine the amount of affordable housing in each city. The entire breakdown by city can be found in the project storage drive, but all cities with less than 2% affordable housing can be found below in Table 3.

Table 3. Bottom Two Percent of SHI

City Name	Amount of SHI *(%)*	Number of Total Units in Town*	Total Number of SHI Units
Becket	0.00%	838	0
Cheshire	0.00%	1481	0
Colrain	0.00%	731	0
Conway	0.00%	803	0
Dunstable	0.00%	1085	0
East Brookfield	0.00%	888	0
Egremont	0.00%	596	0
Erving	0.00%	778	0
Florida	0.00%	335	0
Gosnold	0.00%	41	0
Granville	0.00%	630	0
Hancock	0.00%	326	0
Hawley	0.00%	137	0
Hinsdale	0.00%	918	0
Leyden	0.00%	300	0

Monroe	0.00%	64	0
Monterey	0.00%	465	0
Montgomery	0.00%	337	0
Mount Washington	0.00%	80	0
New Braintree	0.00%	386	0
New Marlborough	0.00%	692	0
New Salem	0.00%	433	0
Oakham	0.00%	702	0
Otis	0.00%	763	0
Pelham	0.00%	564	0
Peru	0.00%	354	0
Petersham	0.00%	525	0
Plainfield	0.00%	283	0
Rowe	0.00%	177	0
Sandisfield	0.00%	401	0
Savoy	0.00%	318	0
Sunderland	0.00%	1718	0
Tyringham	0.00%	149	0
Warwick	0.00%	363	0
Washington	0.00%	235	0
West Stockbridge	0.00%	645	0
Windsor	0.00%	387	0
Blandford	0.19%	516	1
Ashfield	0.25%	793	2
Leverett	0.25%	792	2
Shutesbury	0.26%	758	2
Whately	0.31%	654	2
Buckland	0.35%	866	3
Rochester	0.43%	1865	8
Charlemont	0.49%	615	3
Richmond	0.57%	706	4
Royalston	0.57%	523	3
Rehoboth	0.63%	4252	27
Hardwick	0.68%	1185	8
Middlefield	0.87%	230	2

Dover	0.92%	1950	18
Halifax	0.94%	2971	28
Boxford	1.14%	2730	31
Goshen	1.14%	440	5
Russell	1.16%	687	8
Wendell	1.19%	419	5
Phillipston	1.22%	658	8
Holland	1.24%	1051	13
Clarksburg	1.27%	706	9
Ashburnham	1.28%	2272	29
Brookfield	1.31%	1452	19
Boylston	1.47%	1765	26
Sutton	1.50%	3324	50
Deerfield	1.53%	2154	33
Hull	1.69%	4964	84
Charlton	1.74%	4774	83
West Tisbury	1.84%	1253	23
Southampton	1.90%	2310	44
Mendon	1.93%	2072	40
Wellfleet	1.94%	1550	30
Princeton	1.96%	1324	26
Sheffield	1.99%	1507	30

* Number taken from 2010 census

4. Available Land Area and Parcels

We looked at all_district_data.csv which was constructed last semester. It grouped data by districts and calculated the area of vacant land each district has and how many parcels each district contains. The top five districts with the most available land are shown in Table 4 below.

Table 4. Top five districts with most available land

District	Available Land *(ft^2)*
4th Berkshire District	1659781447
1st Franklin District	1296162864
5th Worcester District	958143775
2nd Franklin District	937041688
1st Worcester District	932572695

The five districts with the most number of parcels are shown in Table 5 below.

Table 5. Top five districts with most available parcels

District	Parcels
1st Worcester District	527
1st Franklin District	454
3rd Bristol District	416
5th Worcester District	392
2nd Franklin District	388

However, some of those land parcels are not completely vacant, either having existing affordable housing or being developed for other purposes. Parcels that are marked as vacant were looked at and the districts were ranked. The five districts with the greatest vacant area are below in Table 6.

Table 6. Top five districts with most available completely vacant land

District	Vacant Land *(ft^2)*
26th Middlesex District	13320447.36
24th Middlesex District	11036420.77

15th Essex District	10094848.64
18th Essex District	9005328.8
34th Middlesex District	7528881.18

The five districts with the most vacant land parcels are below in Table 7

Table 7. Top five districts with most available completely vacant parcels.

District	Parcels
13th Middlesex District	79
12th Middlesex District	55
26th Middlesex District	45
34th Middlesex District	37
11th Plymouth District	30

Those districts have better opportunities either to develop affordable housing or to sell the land to support affordable housing. While there is not a one-to-one correlation between the districts that have low percentages of affordable housing units, there is certainly evident room for them to increase their percentages, based on this dataset.

Based on discussions with the client, it's become apparent that there is some murkiness present around the current state-wide legislation on the required percentage of affordable housing. This is an important factor in determining the urgency and priority level placed on districts with low percentages. Understanding the landscape of this legislation, and what requirements/incentives have been currently placed on districts (if any at all) to develop affordable housing will help us determine which districts we should be focusing our time and resources on investigating.

5. Attempting to secure access to an API

A large component of the proposed questions revolves around analyzing the land based on current market value. There are several APIs that offer property data such as assessed values and recent sales. We have reached out to several APIs including ATTOM and Zillow. ATTOM charges for use above what their free trial API allows. The free ATTOM API has a time limitation on pulling data (once every 10 minutes) and can only bring back 10data points each time. Also, ATTOM does

not have estimated property values for every property tested with. The Zillow API which would have their estimated property values is not available to education research projects. We are still reaching out to Zillow to see if we can get access to API or the data.