

## Deliverable I

Our project has sufficient data so we are not planning to collect additional data.

### Perform a preliminary analysis of the data

We started to explore the parcel dataset and we are planning to use that dataset to find out how many potential new houses can be created in district 9 (Counselor Brandon's district). We also explored excel files that contain information about different universities and the number of students who live on and off campus. We also have data in pdf forms. We are also going through this data - using it in combination with the excel files to answer question 1 (see below). Our data is not specific to the district that we need to analyze so we are learning about shape files to filter out the district 9 data from the city dataset.

### 2 Key Questions:

**Question 1 :** How is the population of students staying off campus changing over time?

The Walsh Administration outlined three clear strategic goals regarding student housing and one of them is to reduce the number of undergraduates living off-campus in Boston by 50% by 2030. We want to explore if they are working towards this goal by comparing the 2019-2020 student housing trends with 2018-2019 student housing trends.

District 9 encompasses the Allston and Brighton neighborhoods. The four main universities that likely have students living there are Harvard, MIT, BU, and BC. That is why we will focus on these 4 institutions.

We have found the following:

Undergraduate students:

University	2017	2018	2019	2020	2021
Harvard	0	0	0	0	N/A
MIT	65	66	39	66	75
BU	3065	3221	3314	3211	3789
BC	1057	1082	1111	1334	1443

Graduate students:

University	2017	2018	2019	2020	2021
Harvard	509	464	463	417	606
MIT	497	479	393	399	555
BU	4606	5065	5058	4115	5568
BC	894	878	894	821	772

Total students:

University	2017	2018	2019	2020	2021
Harvard	509	464	463	417	606
MIT	563	545	432	465	630
BU	7671	8286	8372	7326	9357
BC	1951	1960	2005	2155	2215

## Question 2:

What percentage of students lived off campus for each university in 2021?

University	Undergrad Students	Grad Students	Total Students
Harvard	NA	14.25%	NA
MIT	12.6%	9.4%	10.6%
BU	21.4%	42.5%	30.4%
BC	15.1%	13.9%	14.7%

Some of the limitations we have encountered while performing preliminary analysis are explained as follows:

Although the data is mostly complete, we still have missing data (Harvard 2021 undergrad students). Another limitation is that we are only choosing to account for four universities located in proximity to district 9. Some of these students might not be living in district 9 while there may also be a good number of students from other universities living in these neighborhoods. In

addition, covid had disrupted the trends so it is hard to make definitive conclusions. If we can overcome these limitations, our findings can be refined.

**Note:**

On monday(10/17) we had our first meeting with the client. After talking with the client, we realized that the initial project didn't meet client's expectations. So we are going to rescope the project. The client wants us to analyze policy violations data to identify the landlords with policy violations and come up with a metric to classify landlords as good/bad landlords. We will pivot to this direction going forward.