

# CS506 Bay State Banner Project

Zilin Zhang (U87038789)

Mingyan Yang (U30215865)

Yujing Chen (U70567267)

Yicheng Li (U29503597)

## 1. Motivation/background of the project

### 1.1 Description of organization

The Bay State Banner is an independent newspaper primarily geared toward the readership interests of the African-American community in Boston, Massachusetts. It owns news weekly that reports on the political, economic, social and cultural issues that are of interest to African American and English speaking Latinos in Boston and throughout New England. Started in 1965 by Melvin B. Miller, a graduate of Harvard College and Columbia Law School, the Banner provides its estimated 120,000 weekly print readers and 300,000 online readers with reports on a mix of local, national, international news and cultural events. The Banner is the newspaper of record for this community and has chronicled the struggles and successes of the last 50 years.

### 1.2 Background of the project

To carry out the political plans, politicians need a large amount of money to support the activities. And they need donors. Actually, the real estate industry is the largest donor to Massachusetts politicians. It covers brokers, developers, architects, and anyone with a vested investment in building. They acquire a large number of profits from the real estate industry and become the power of the politicians to support the politicians' activities.

### 1.3 Motivation of the project

The real estate industry covers many people who have a vested investment in building, including brokers, developers, architects and ect.. Our client Yawu Miller, a journalist writing for the Bay State banner would like to examine the lobbying power of the real estate industry in Massachusetts.

## 2. Goals, hypothesis, outcomes

### 2.1 Goals

This project has two primary goals:

- The ultimate output will be a comprehensive dataset of the individuals and businesses affiliated with the real estate industry in Boston/ Massachusetts.
- The ultimate question we are seeking to understand is the influence of the real estate industry in Boston by looking at their prominence as political contributors and city decisionmaking.

## **2.2 Hypothesis of the result**

Naturely, the result may reveal that those individuals and businesses who donate most frequently and have the greatest amount of donation will be those who have the strongest influence. But there are other hypotheses such as those who donate less frequently but have the greatest amount of donation would be the most influential.

## **2.3 Outcomes**

Visualization outcomes(various charts) would be provided to prove our hypothesis. And based on those different charts, certain analysis would be shown straightforwardly to our client to reach our goals.

## **3. Non-goals, out of scope topics**

Our goal is to analyze the donations from the different types of players involved in the real estate industry. It's all about analyzing numbers and presenting the data through charts. We don't need to find out the implicational meaning of this social phenomenon.

In a word, we mainly focus on the data science topic. The society topic is not in our scope.

## **4. End result/product -> definition of done**

### **4.1 End Result**

The end result will be patterns of connections with BPDA decisions and approved projects and involvement in ZBA appeals processes. Through the data visualization, the result of those who are the most influential real estate industry professionals in Boston would be straightforwardly revealed.

### **4.2 Definition of Done**

Acquiring enough useful datasets and merge all datasets. Arrange that datasets and remove duplicates. Match the datasets and analyze the datasets. Presenting them into different charts. All the charts should be straightforward and solve all the problems like who is the most influential among the donors. Most importantly, to reach our goals, the charts should show comprehensive datasets of the individuals and businesses affiliated with the real estate industry in Boston/ Massachusetts and the influence of the real estate industry in Boston by looking at their prominence as

political contributors and city decisionmaking should be easily concluded from the charts.

## **5. Open questions, uncertainties**

### **5.1 Open Questions**

Could our project really benefit society?

How could our project be helpful for city decisionmaking?

According to the results, how can the politicians and society benefit from it?

### **5.2 Uncertainties**

- 1) We have no experience of data science projects, so all the mentioned tools are new to us, such as Scrapy, FuzzyWuzzy and GIS. We might need some time to get familiar with these tools.
- 2) Also, when using OpenCV to extract the text from the PDF, we are not sure how much time it would cost.
- 3) When merging the datasets, we need to cooperate with other teams. During the cooperation, there may be some trouble communicating.
- 4) When we assign the parts, make sure everyone understands the requirements.
- 5) Time arrangement. We may finish the delivery later than our plan.