New England Blacks in Philanthropy

Financial Support Given to Black Non-Profits

Shunying Chen (U47342767) | Elsa Jaysing (U77504656) | Reha Patel (U17557726)

INTRODUCTION

The idea behind this project was to collect financial information going to predominantly Black non-profits in the Boston area. Predominantly Black non-profits are described as either having a majority Black board of directors, a Black CEO/majority Black leadership, and/or serving a majority Black community.

In order to determine whether a non-profit was predominantly Black, we were tasked with using 990 Forms to cross check with LinkedIn, RocketReach, and GuideStar APIs to determine the race of the non-profit leadership. We specifically wanted to look at what percentage of this charity money is going towards Black non-profits amongst all nonprofits in Boston.

However, as we began exploring these different techniques, we came to realize that most would not work or we did not have the resources to complete parts of the task at hand. As a result, we switched gears to create a proof of concept which could be used by later semesters/teams once resources such as API keys and access to LinkedIn became more readily available.

This document outlines the approaches we took to attempt to answer questions and what approaches we would suggest to future semesters/teams to continue the data exploration for New England Blacks in Philanthropy. It is good to note that this project requires the utmost care and attention to detail as we do not want to misidentify people's identities.

SOURCE DESCRIPTION

GuideStar (Developer Portal)

Guidestar is a database that provides information on non-profit organizations and thus was our main source to look up organizational information from. You are able to filter their database based on geographical location, financial information, and other general information about an organization. A caveat to some of these filters though is that the filters are only useful if each organization is completely filled out to the fullest; however, that is not always the case. We initially thought we could use the "Population Served: Ethnic and Racial Groups: People of African descent" as a way to narrow down organizations serving Black communities, but there were some organizations missing from the list that were included in the Guidestar database.

GuideStar has different API subscriptions that we used to gather our data. The reason we weren't able to obtain all of the data, and instead did a proof of concept, was because of some issues that arose in getting enough API keys from our BU subscription, but nevertheless when filling out this form, a developer got back to us and gave us API keys for their Premier API with 500 calls each so we were able to at least tackle one organization - The Boston Foundation. The information that was useful to us that the key gave back was the organization's financial information and the names, positions, and compensation of the board of directors.

Accessing Guidestar and Obtaining API Keys:

You have to make an account on the Guidestar Developer Portal, and by doing so you will be put in contact with a developer who will help you figure out what type of key you need. Below are two people we were put in contact with who were extremely helpful in giving us keys and resetting the number of API calls we had when we ran out.

Guidestar Developer Contacts:
Tim Reifschneider, Director of Business Development, (he/him/his)
Email: tim.reifschneider@candid.org
Phone Number(s): 516-437-7827, 516-220-6886 (mobile)

Wayne Harris, Business Development Associate
Email: wayne.harris@candid.org
Phone Number(s): 908-803-1111 (mobile)

LinkedIn
LinkedIn provides us with information and connections about the board members allowing us to determine if the board member is connected in any way to Black communities, and to determine if the non-profit they belong to is predominantly Black. More specifically, for each LinkedIn profile it provides a brief summary about the individual and their experiences, interests, skills and endorsements, and education. Although the information provided doesn't allow us to determine right away whether the member is Black or closely connected to the Black community, we can cross reference the results with a specific set of keywords to determine their connection.

However, despite our initial wishes to use the LinkedIn API, we found that LinkedIn limited its API usage leading us to consider different options. This led us to use a web scraper to obtain the same information provided by the LinkedIn API.

PIPELINE
As we moved over to a proof-of-concept format, our team was provided with a general pipeline that consisted of three major steps:

1. Identify grant recipients of the Community Foundations.
2. Extract executives names and compensation.
3. Look up executives names and/or organizations to match with profiles in LinkedIn.

Step 1: Identifying Grant Recipients of Community Foundations
In the transition to the proof-of-concept project, we were asked to determine the feasibility of identifying the grant recipient of the Community Foundations, whether it be through GuideStar or extracting the information from the 990 forms. In order to actually determine the feasibility, we began with just one Community Foundation, the Boston Foundation. None of the GuideStar API's listed the recipients so we moved on to extracting from the 990 forms. Below is a photo of which part of the 990 form we pulled from.

Additional Data

Software ID:
Software Version:
EIN: 04-2104021
Name: Boston Foundation Inc

Form 990,Schedule I, Part II, Grants and Other Assistance to Domestic Organizations and Domestic Governments.

(a) Name and address of organization or government	**(b)** EIN	**(c)** IRC section if applicable	**(d)** Amount of cash grant	**(e)** Amount of non-cash assistance	**(f)** Method of valuation (book, FMV, appraisal, other)	**(g)** Description of non-cash assistance	**(h)** Purpose of grant or assistance
1 to 4 Foundation 5225 East Camino Cielo Santa Barbara, CA 93105	46-5001370	501 (c) (3)	10,000				Operating Support
350Org 20 Jay Street Suite 732 Brooklyn, NY 11201	26-1150699	501 (c) (3)	12,250				Operating Support

Form 990,Schedule I, Part II, Grants and Other Assistance to Domestic Organizations and Domestic Governments.

(a) Name and address of organization or government	**(b)** EIN	**(c)** IRC section if applicable	**(d)** Amount of cash grant	**(e)** Amount of non-cash assistance	**(f)** Method of valuation (book, FMV, appraisal, other)	**(g)** Description of non-cash assistance	**(h)** Purpose of grant or assistance
5-Cities Community Service Foundation PO Box Zero Grover Beach, CA 93483	77-0437523	501 (c) (3)	10,000				Operating Support
A Far Cry Inc 146A South Street Jamaica Plain, MA 02130	30-0456355	501 (c) (3)	7,500				Arts

Form 990,Schedule I, Part II, Grants and Other Assistance to Domestic Organizations and Domestic Governments.

(a) Name and address of organization or government	**(b)** EIN	**(c)** IRC section if applicable	**(d)** Amount of cash grant	**(e)** Amount of non-cash assistance	**(f)** Method of valuation (book, FMV, appraisal, other)	**(g)** Description of non-cash assistance	**(h)** Purpose of grant or assistance
ABEKAM Inc 279 Pond Street Jamaica Plain, MA 02130	26-4363808	501 (c) (3)	3,000				Operating Support

Because these recipients were stretched across multiple pages, a .PDF to .XLSX converter was used which was then cleaned up. This converter produced the entire PDF on an excel sheet, but only the section titled *Grants and Other Assistance to Domestic Organizations and Domestic Governments* was relevant to the project so the irrelevant rows were deleted and filtered out.

1 (a) Name and address of organization or government	(b) EIN	(c) IRC section (if applicable)	(d) Amount of cash grant	(e) Amount of non-cash assistance	(f) Method of valuation (book, FMV, appraisal,	(g) Description of noncash assistance	(h) Purpose of grant or assistance
1 to 4 Foundation 5225 East Camino Cielo Santa Barbara, CA 93105	46-5001370	501 (c) (3)	10,000	0.			Operating Support/Annual Fund
10 Thousand Windows, Inc 348 Canyons Pkwy Livermore, CA 94551	27-2505761	501 (c) (3)	30,000	0.			Operating Support/Annual Fund
350.Org 20 Jay Street, Suite 732 Brooklyn, NY 11201	26-1150699	501 (c) (3)	14,500	0.			Operating Support/Annual Fund
50CAN, Inc. 1625 K Street NW, Suite 400 Washington, DC 20006	27-3069592	501 (c) (3)	15,000	0.			Operating Support/Annual Fund
5-Cities Community Service Foundation - PO Box Zero - Grover Beach, CA 93483	77-0437523	501 (c) (3)	100,000	0.			Operating Support/Annual Fund
826 Boston Inc. 3035 Washington Street Roxbury, MA 02119	20-8065915	501 (c) (3)	10,000	0.			Capital Campaign
826 Boston Inc. 3035 Washington Street Roxbury, MA 02119	20-8065915	501 (c) (3)	2,500	0.			Education
826 Boston Inc. 3035 Washington Street Roxbury, MA 02119	20-8065915	501 (c) (3)	12,800	0.			Operating Support/Annual Fund
9 Dots Community Learning Center 931 North Highland Avenue Los Angeles, CA 90038	45-2834070	501 (c) (3)	100,000	0.			Operating Support/Annual Fund
ABEKAM, Inc. 279 Pond Street Jamaica Plain, MA 02130	26-4363808	501 (c) (3)	25,000	0.			Economic Development
ABEKAM, Inc. 279 Pond Street Jamaica Plain, MA 02130	26-4363808	501 (c) (3)	2,500	0.			Operating Support/Annual Fund
Abraham House, Inc. PO Box 305 Bronx, NY 10454	13-3721924	501 (c) (3)	25,000	0.			Operating Support/Annual Fund
Academy of the Holy Names 1075 New Scotland Road Albany, NY 12208		Religious Org	25,000	0.			Capital Campaign
Academy of the Holy Names 1075 New Scotland Road Albany, NY 12208		Religious Org	5,000	0.			Operating Support/Annual Fund
Access Fund PO Box 17010 Boulder, CO 80308	94-3131165	501 (c) (3)	10,000	0.			Capital Campaign
Access Fund PO Box 17010	94-3131165	501 (c) (3)	1,000	0.			Operating Support/Annual Fund
ACLU Foundation of Massachusetts Inc. - 211 Congress Street, Suite	47-3686152	501 (c) (3)	71,100	0.			Operating Support/Annual Fund
Action for Boston Community Development, Inc. - 178 Tremont Street -	04-2304133	501 (c) (3)	20,000	0.			Emergency Support
Action for Boston Community Development, Inc. - 178 Tremont Street -	04-2304133	501 (c) (3)	1,000	0.			Operating Support/Annual Fund
Actors' Shakespeare Project, Inc. 191 Highland Avenue, Suite 2B	20-0815685	501 (c) (3)	71,500	0.			Operating Support/Annual Fund
Ad Club Foundation, Inc. 22 Batterymarch Street	43-2050395	501 (c) (3)	39,760	0.			Scholarship
Administrators of The Tulane Educational Fund aka Tulane University -	72-0423889	501 (c) (3)	1,000,000	0.			Capital Campaign
Administrators of The Tulane Educational Fund aka Tulane University -	72-0423889	501 (c) (3)	1,500	0.			Operating Support/Annual Fund
Administrators of The Tulane Educational Fund aka Tulane University -	72-0423889	501 (c) (3)	1,500	0.			Scholarship
Adolescent Consultation Services, Inc. - 189 Cambridge Street -	04-3263996	501 (c) (3)	8,500	0.			Operating Support/Annual Fund
Affordable Housing Institute 77 Franklin Street, 7th Floor Boston, MA	92-0185895	501 (c) (3)	25,000	0.			Operating Support/Annual Fund
African Community Economic Development of New England, Inc. - 80	51-0410358	501 (c) (3)	7,000	0.			Education

Step 2: Extracting Executive Names and Compensation (Code: guidestarAPI.py)

Of the data gathered in Step 1, we then went on to use the EINs (Employer Identification Numbers - these numbers are standard across APIs/websites and are unique to each employer) gathered from the 990 forms as a parameter in the GuideStar Premier API. Because each of the three of us only had 500 calls per API key, we had to split the EINs up into 3 groups of 500 by indexing the row which contained the EINs. As a result of this API call, we were able to gather all the information needed, including executives' names, compensation, as well as the financial data of the charities, all in JSON format. Because of the nested nature of the JSON, we ran into issues with cleaning it up and used a JSON to CSV converter to initially validate that the results we were getting would be viable for the client. The results of this are in a file called SampleGuidestarData.csv that will be used in step 3. For future semesters, we also created a starter code file called json_to_csv.py which will read JSON and convert it to a CSV file.

Step 3: Looking up Executives via LinkedIn (Code: linkedin.py)

Once we have obtained all the names and companies of the executives, we can use that information and input it into the web scraper we have built to get the LinkedIn profile information. The LinkedIn web scraper we built logs onto LinkedIn and performs a search by taking the search phrase “name” + “keyword” and constructs a valid url and we can then reach. An example of this:

```
search = "Teri Williams OneUnited Bank"
url = “https://www.linkedin.com/search/results/all/?keywords=Teri%
20Williams%20OneUnited%20Bank&origin=GLOBAL_SEARCH_HEADER”
```

This method was used because LinkedIn hides the search bar such that we could not access it. Once we have reached the search result page, we were able to use BeautifulSoup to access the html of the top result and get the href attribute link that will lead us to the specific profile. Once we are able to get the profile link, we input the link into a LinkedIn web scraper library that will perform the scraping for us and obtain the results about the board member.

```
Teri Williams Cohee

About
["Teri Williams is President and Chief Operating Officer and serves on the Board of Directors of OneUnited Bank. She is
responsible for implementation of the Bank's strategic initiatives, as well as the day to day operations of the bank. Un
der her leadership, OneUnited Bank has consolidated the local names and product offerings of four (4) banks to create a
powerful national brand supported by innovative products and services. She believes the financial services industry has
not connected with urban communities to fully support economic development and wealth building. OneUnited Bank can serve
 as a bridge by offering affordable financial services for all and financial workshops. She brings 30 years of financial
 services expertise from premier institutions such as Bank of America and American Express, where she was one of the you
ngest Vice Presidents. Ms. Williams holds an M.B.A. with honors from Harvard University and a B.A. with distinctions fro
m Brown University. She has served as Chairman of the Black Economic Council of Massachusetts (BECMA) in Boston and is o
n the board of the 79th Street Corridor in Miami. She previously served as Chair of the Urban Initiative Task Force of t
he Miami-Dade Beacon Council. Ms. Williams has received numerous notations and awards for her contribution to urban comm
unities including from the Urban League, NAACP and the National Black MBA Association."]

Experience
[President & COO at None from None to None for None based at None, Vice President at None from None to None for None bas
ed at None]

Education
[None at Harvard Business School from None to None, None at Brown University from None to None]

Interest
[Harvard Business School Alumni (Official), Black Enterprise Networked, Harvard Business School, South Florida's Black P
rofessionals Network, Chief Executive Officer (TGL), National Black MBA Association]

Accomplishments
[]
```

After obtaining the scraped results, we converted it into a string and to prevent case sensitivity when performing the keyword search we converted it into lowercase. We are then able to cross reference the given set of keywords with the LinkedIn profile to determine whether the board member is Black or connected to Black communities.

Discussion Concerning LinkedIn

After some testing, we realized that we cannot search based on "name" + "company". This is because the data we gathered about each organization does not include the company at which each board member works at. The reason we want to include a company in our search is to help narrow down the search, as the way LinkedIn obtains the profile link is by taking the top search result and finding the relevant html class and class name and taking the href. Therefore, if there are multiple results, we are not able to parse to each one and find the person we are looking for. This is caused by how LinkedIn names the class name. Depending on which part we are looking at, the class name could be randomized or they could be exactly the same. This causes a problem as we are not able to iterate through all the results and obtain the profile links without the exact location of the href. At this point, we realized another problem and the slight infeasibility of this portion of the pipeline. This stems from the issue of how can we be certain that whoever we are looking for is actually the correct person. Even if we were able to iterate through all the names in the LinkedIn search, it is sometimes uncertain which individual, out of potentially hundreds and thousands, is the individual associated with the organization. We cannot just cross reference the organization name with the board member's profile because sometimes the name is slightly different or the board member may not have included the organization name in their profile. This is made evident when we actually did a test run of some data we already collected, which is discussed below.

A more concrete example of this problem is if we happen to have a board member with the name John Smith, which is a popular first and last name. If we search on LinkedIn with this name, then we will get a lot of results. If we narrowed down that search to Boston, we would still potentially get a couple hundred or thousand results. At this point, we could use the organization name, but what if the board member didn't include the name? Or, what if there are multiple results that have some association with the organization and included it in their profile, even though they are not necessarily a board member. This complicates how we want to process with this portion of the pipeline.

Despite these concerns, we went ahead and tested our LinkedIn web scraper on some gathered data. When we first constructed our search phrase, we used "name" + "organization name" to build the query. We immediately encountered problems as many queries could not find a result. At this point, we decided to just test it on the name, and for the majority of the tests we were able to get results. We cross referenced those results with the following keywords:

```
keywords = ["alpha phi alpha", "alpha kappa alpha", "kappa alpha psi", "omega psi phi",
            "delta sigma theta", "phi beta sigma", "zeta phi beta", "sigma gamma rho",
            "iota phi theta", "the links", "the links incorporated", "the boulé", "boulé",
            "jack and jill of america", "n.a.a.c.p", "naacp", "the urban league", "urban league",
            "national association of black accountants, inc.", "national association of black accountants",
            "national association of black accountants inc", "national association of black accountants, inc",
            "national association of black accountants inc.", "naba", "national black mba association",
            "nbmbaa", "hbcu", "boston, young black professionals", "boston young black professionals",
            "young black professionals", "ybp", "black networking groups", "black women", "black woman",
            "black men", "black man", "black enterprise", "national association of african americans in human resources",
            "naaahr", "new england blacks in philantropy", "black educators alliance of massachusetts",
            "national association of black social workers", "people of color in independent schools",
            "national society of black engineers", "national black nurses association",
            "student national medical association", "blacks in government", "black lawyers association",
            "national black law students association", "national forum for black public administrators",
            "national association of black journalists", "alabama a&m university", "alabama am university",
            "alabama state university", "bishop state community college", "gadsden state community college",
            "shelton state community college", "concordia college", "miles college", " oakwood university",
            "selma university", "stillman college","talladega college", "tuskegee university",
            "university of arkansas at pine bluff", "arkansas baptist college", "philander smith college"
            "shorter college", "charles drew university of medicine and science", "delaware state university",
            "university of the district of columbia", "howard university", "florida a&m university", "florida am university",
            "bethune-cookman university", "bethune cookman university", "edward waters college", "florida memorial university",
            "albany state university", "fort valley state university", "savannah state university", 'clark atlanta university',
            'interdenominational theological center', 'morehouse college', 'morris brown college', 'paine college', 'spelman college',
            'kentucky state university', 'simmons college of kentucky', 'grambling state university', 'southern university and a&m college',
            'southern university law center', 'southern university at new orleans', 'southern university at shreveport', 'dillard university',
            'xavier university', 'bowie state university', 'coppin state university', 'morgan state university',
            'university of maryland, eastern shore', 'alcorn state university', 'jackson state university', 'mississippi valley state university',
            'coahoma community college', 'hinds community college-utica', 'rust college', 'tougaloo college', 'harris-stowe state university',
            'lincoln university of missouri', 'elizabeth city state university', 'fayetteville state university', 'north carolina a&t state university',
            'north carolina central university', 'winston-salem state university', 'barber-scotia college', 'bennett college',
            'johnson c. smith university', 'livingstone college', 'st. augustine's college', 'shaw university', 'central state university',
            'wilberforce university', 'langston university', 'cheyney university of pennsylvania', 'lincoln university', 'south carolina state university',
            'denmark technical college', 'allen university', 'benedict college', 'claflin university', 'morris college', 'voorhees college',
            'clinton junior college', 'tennessee state university', 'american baptist college', 'fisk university', 'knoxville college',
            'lane college', 'lemoyne-owen college', 'meharry medical college', 'prairie view a&m university', 'texas southern university',
            'st. philip's college', 'huston-tillotson university', 'jarvis christian college', 'paul quinn college', 'southwestern christian college',
            'texas college', 'wiley college', 'norfolk state university', 'virginia state university', 'hampton university', 'virginia union university',
            'virginia university of lynchburg', 'bluefield state college', 'west virginia state university', 'university of the virgin islands']
```

The follow is the result returned after our run:

{'JOHN AYLIFFE': [], 'CATHERINE AYLIFFE': [], 'RONALD BOEHM': [], 'NICHOLAS HIRST': [], 'JEREMY FLOYD': [], 'ANDREW MEANS': [], 'ANNE SWARTZ': [], 'JULIE BRYCE': [], 'KRISTEN FOOT': [], 'MIKE GOGIS': [], 'LANCE ROBINSON': [], 'HUMPHREYS MUNAI': [], 'KC GOLDEN': [], 'CHINA BROTSKY': [], 'JESSY TOLKAN': [], 'BILL MCKIBBEN': [], 'JAY HALFON': [], 'LIDY NACPIL': [], 'TASNEEM ESSOP': [], 'MELINA LABOUCAN-MASSIMO TO 319': 'None', 'MARC MAGEE': [], 'VALLAY-LATH VARRO': 'None', 'CHRIS TESSONE': [], 'DERRELL BRADFORD': [], 'MICHAEL PHILLIPS': [], 'ANN BOROWIEC': [], 'ROLAND MARTIN': [], 'JONATHAN SACKLER': [], 'DACIA TOLL': [], 'DAVID WICK': [], 'CAMPBELL BROWN': [], 'DEEPA JAVERI': [], 'ANDREW SCHWEDEL': [], 'MIKE MINER': [], 'ALLAN BUCK': [], 'J JOHNSON': [], 'JESSICA DRENCH': [], 'GILLIAN KOHLI': [], 'MARC FOSTER': [], 'MIMI CURRAN': [], "EMILY D'AMOUR PARDO": 'None', 'LINDA BUTTON': [], 'ANDREW COHN': [], 'DONNA COWAN': [], 'JON FULLERTON': [], 'CAROL GREENWALD': [], 'REVEREND TIM HOUSE': [], 'CHARISSE HOWSE': ['morgan state university'], 'LISA KRAKOFF': [], 'JEFF MAYERSOHN': [], 'PAM ROSENBERG': [], 'JANET TIAMPO': [], 'KEVIN WHALEN': []}

Since we did not want to definitely decide if a board member is Black or associated with the Black community, we elected to keep the list of keywords found in the profile for each individual. In the sample data we tested, we found that the majority did not contain those keywords, and only one person had results. Additionally, while observing the profile pictures of these individuals during the web scraping, we found that most are not Black. In fact, the majority are White. These results are interesting as even in a small subset of data we are already able to gain some insight.

NOTES FOR FUTURE SEMESTERS

Reverse Pipeline

Before obtaining and analyzing all of the data, we need to make sure that the accuracy of our approach is high enough so that we are not misidentifying individuals. In order to test our accuracy a good starting point would be the following:

1. Create a dataset comprised of board members from Black-led organizations whose races we already know.
2. Split the dataset into test and train sets.
3. Run the train dataset on an SVM algorithm. You can use the initial key words as a starting point to identify an individual's race, weighting each association based on how frequently it is found in the profiles of Black individuals.
4. Run the trained SVM algorithm on the test dataset to determine its accuracy. Tweak the keywords and their weights until you have a high accuracy.
5. Once the accuracy is high enough, go through the original pipeline to obtain all of the data and predict those individuals' races.

Analyzing the Data

Once the data is collected, we've brainstormed some approaches and steps to go through to answer the questions NEBiP wanted to analyze which are the following:

1. What percent of giving goes to Black-led organizations vs. all donations?
2. What are the average revenues of Black-led organizations compared to average revenues of all organizations?
3. What are the unrestricted net assets of Black-led organizations compared to average unrestricted net assets of all organizations?
4. How does gender impact numbers above?

5. What is the total volume of Black philanthropic dollars?
6. What is the total number of Blacks in the nonprofit sector?

Step 1: Determine which organizations are Black-led organizations
After scraping LinkedIn and using key words to identify associations affiliated with Black communities you can try to figure out whether the individual is Black or not. This of course should not be determined by just one affiliation and should be done responsibly. Once the ethnicity of each individual on a board is determined, you can tally it up to see if the board of directors is led by a majority, which would be greater than 50% of the board. Once that is done for each organization, you would be able to have two sets of organizations, one that contains Black-led organizations and one that contains the rest.

Step 2: Create datasets that isolate gender
Once you have split up the datasets by whether they are Black-led or not, you can also break them up further to have it broken down by gender. This step will require more discussions on the feasibility and best approach to identify one's gender as we want to make sure we're being inclusive to all genders and not misidentifying anyone.

Step 3: Analyze financial data
Once you have the different data sets for each group it will be fairly easy to total up the numbers for the financial data to answer the questions above as donations, revenue, and net assets are all given by the GuideStar API. You can then use graphs and other data visualization techniques to represent the data better as you see fit.

Other Potential Areas to Analyze
One aspect of the grantee information that we noticed for The Boston Foundation is that there were a lot of grantees located outside of Massachusetts. We still think it's important to analyze them because that's where the money from these Massachusetts organizations are going to, but it would be interesting to divide the datasets again based on location (Massachusetts and not Massachusetts) to see exactly how much money is leaving the state and not helping the Black communities here as well as whether they are helping Black communities elsewhere instead or not.

Already Explored Paths for Data Collection
The following is a list of APIs and databases that we already explored to try to collect data. This information is provided in case there ends up being a kink in the process forcing you to explore some other paths.

- Charity Navigator API
  - The Charity Navigator API has the ability to provide the name of the CEO and trustee using the EIN, but does not include compensation for either. However, the

API does provide a general overview of the 990 financial information based on parameters such as city, state, name, etc. We have explored methods to extract just the names of the grantees from our initial spreadsheet because they are joined with the addresses. The code still needs slight tweaking, but what we have gotten up until a certain point has been provided in *get_addresses.py*. Given the limitations of the number of calls we can make with GuideStar, Charity Navigator may be a viable option for 990 financial information.

- RocketReach API
  - RocketReach is an API that gives information about an individual when you use their name, and company name if needed, which we initially tried using to find associated organizations to help determine the demographics of the board of directors based on their affiliation with other Black communities. The cons to using this API, and the reason we didn't use it in our final pipeline, was that there was a mismatch between the names provided by the 990 forms and the names in the RocketReach system, and so there was an inaccuracy in the results.
- RefUSA Database
  - RefUSA is a database that provides demographic information for the board of directors of an organization. The cons to this approach is that the database isn't mainly for non-profit organizations and so it doesn't have a lot of the organizations we were looking for anyways, and even if it did there were sometimes names missing from the board or no demographic information actually listed.
- LinkedIn API
  - Obtaining API keys for LinkedIn would either take too much time to obtain or we wouldn't even be granted keys in the end, so it was an approach that was vetoed based on the time constraint we had to figure out something that would be viable. We then turned to scraping LinkedIn instead since the website still had useful information.