Appendix A

Locater Script

https://github.com/justinmiller33/CampaignFinances/blob/master/Locater/legFinderGlobal.py

Appendix B

Polygon Location Algorithm

```
# Donor Location taken from geolocation
point = donorLocation

# Start with an inconclusive donor district
donorDistrict = inconclusive

# Looping through each district to check if point is in that polygon
foreach polygon in shapefileMap:

    if rayCasting(point,polygon) == inside:

        donorDistrict = polygon

    else:
        continue


# Ray-Casting Algorithm to find if point is in one polygon
function rayCasting(point,polygon):
    count ← 0
     foreach side in polygon:
       if ray_intersects_segment(P,side) then
         count ← count + 1
     if is_odd(count) then
       return inside
     else
       return outside
```

Appendix C

BISG Algorithm

```
# BISG Algorithm

# Looping through each record in the dataset
foreach record in dataset:

        # Get surname proportions
        surnameProps = getSurnameProps(record[fullName])

        # Get racial proportions
        raceProps = getBisgProbs(surnameProps,racord[town]):

# Function to get racial proportions dependent on surname
function getSurnameProps(fullName):

    # Get surname by spliting string at the comma
    surname = splitName(fullName)

    # Clean surname, deleting any affixes (Jr., Dr.)
    surname = cleanSurname(surname)

    # Call census api to get race proportions
    # Returns 4 member array of race distributions for that name [white,black,asian,hispanic]
    surnameProps = getFromCensusApi(PWHITE,PBLACK,PASIAN,PHISPANIC, name=surname)

    return surnameProps

# Function to get racial proportions from surname and geographical distributions
function getBisgProbs(surnameProps,town):

    # Spreadsheet of racial distributions by town or district (number of individuals for each race)
    demographics = load(demographics.xlsx)

    # Load in demographics of the entire US for comparison (number of individuals for each race)
    usDemographics = load(usDemographics.xlsx)

    #  Getting proportion of each race out of the entire united states
    townProps = demographics[town]/usDemographics

    # Getting race props from bisg formula (see METHOD->RACIAL)
    foreach race:
        raceProps[race] = (surnameProps[race]*townProps[race]) /sum(surnameProps*townProps)
```

Appendix D

BISG Scripts

Surname Analysis

https://github.com/justinmiller33/CampaignFinances/blob/master/Race/raceProbs.py

Bayesian + Geocoding:

https://github.com/justinmiller33/CampaignFinances/blob/master/Race/BISG.py

Appendix E

Occupation Groupings

Owner	Arts	Self-Employed	Student	Retired	Legal Services	Executives	Unemployed	Politics
Business Owner	Actor	Self Employed	Student	Retired	Attorney	Business Executive	At Home	Attorney/lobbyist
Co-owner	Artist	Self-employed		retired	attorney	CEO	Home Maker	Campaign Manager
Owner	Author	Self-Employed		RETIRED	ATTORNEY	CFO	Homemaker	Chief of Staff
OWNER	Designer				Attorney At Law	Chairman	Housewife	Director of Government Affairs
Principal Owner	Editor				Clerk	Chief Operating Officer	N/A	Government Affairs
Restaurant Owner	Librarian				District Attorney	COO	None	Government Relations
	Musician				General Counsel	Director	none	Legislative Agent
	Photographer				Lawyer	Executive	Not Employed	Legislative Aide
	Writer				Paralegal	Executive Director	Not employed	Legislator
						Executive Vice President	not employed	Lobbiest
						Founder	Student	Lobbyist
						Managing Partner	Unemployed	lobbyist
						owner		Public Affairs
						Partner		State Representative
						President		
						President & CEO		
						President and CEO		
						President/CEO		
						President/ceo		
						Senior Vice President		
						Vice President		
						Vp		
						VP		

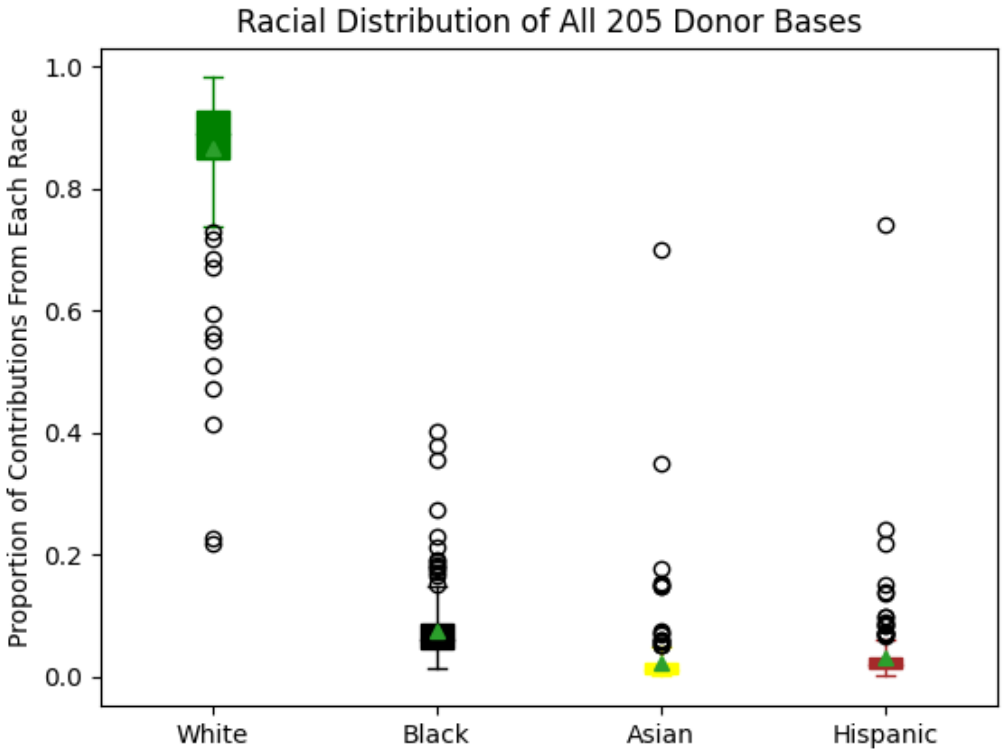
Business	Education	Sciences	Real Estate	Public Services	Blue Collar
Accountant	Adjunct Professor	Chiropractor	Real Estate	Firefighter	Architect
Banker	Educator	Civil Engineer	Real Estate Advisor	Police Officer	Builder
Broker	Professor	Dentist	Real Estate Agent	Social Worker	Carpenter
Business Manager	School Teacher	Doctor	Real Estate Broker	State Police	Construction
Businessman	Teacher	Engineer	Real Estate Developer	State Trooper	Driver
Consultant	teacher	Nurse	Real Estate Development	Trooper	Electrician
consultant		Nurse Practioner	Realtor		Farmer
CPA		Ophthalmologist			General Contractor
Cpa		Optometrist			Mechanic
Finance		Pharmacist			Plumber
Financial Advisor		Physcian			Technician
Financial Analyst		Physical Therapist			
Financial Planner		Physician			
Insurance		Psychologist			
Insurance Agency Owner		Registered Nurse			
Insurance Agent		Rn			
Insurance Broker		Scientist			
Investor		Software Developer			
Marketing		Software Engineer			
Office Manager		Veterinarian			
Product Manager					
Sales					
Sales Manager					
Sales Representative					
Treasurer					

28

[28] The bolded words at the top of each column are the "groupings" names, followed by the self-reported jobs in that group, which are not bolded. The coloring is for visual effects only and does not correlate to the groupings.

Appendix F

In Depth Racial Distributions



Box and Whisker of Racial Distributions

The above chart details the distributions of MA state senator's donor base by race. Each box plot represents 205 representatives' contribution proportion from that race. Note that due to the rare name bias discussed in Methodology: Racial, the minority race proportions are underestimated by an unknown margin. Regardless, there is an interesting variation in the patterns of minority proportions for each candidate as observed in the outliers. As discussed in our racial findings, Liz Miranda, Rithy Uong, William Lantigua have great fundraising support from donors of their own race. Uong and Lantigua are particularly rare outliers, having more than double the support of the runner up Asian and Hispanic donor bases.