Abdullah Albijadi, Audrey McMillion, Gordon Ng, Ji Zhang, Thachathum Amornkasemwong

Baystate Banner: Boston voting patterns	
Contact	Yawu Miller Senior Editor yawu@bannerpub.com Baystate Banner / Dorchester Reporter
Organization	Baystate Banner Dorchester Reporter : https://www.dotnews.com/
Organization Description	The Bay State Banner is an African American owned 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. The Dorchester Reporter, founded in 1983, is a community newspaper which is published every Thursday. It serves the Dorchester and adjoining communities in Boston.
Project Type	Data Science
Project Description	The client wants to understand how racial voting patterns in Boston changed over the past decade on the Ward Precinct level. The components to measure voting behavior include ballots cast, blanks cast, registration, eligible voters, and turnout as an auxiliary measure.
Data Sets	The datasets are from independent Spark! Work in compiling and cleaning voter file data from the City of Boston Boston, 2011 - 2019 elections **PULL OUT DEMOGRAPHIC DATA FROM WARD/ PRECINCTS FROM THE DATA ABOVE** Datasets under focus: 2013, 2017 Mayor preliminary and general races

	2016, 2020 Presidential general races 2018 Capuano vs. Pressley Congressional race 2018 DA preliminary and general races 2020 Democratic Senate primary Additional ideas MassGIS (Bureau of Geographic Information) | Mass.gov Electoral Maps | Boston Planning & Development Agency (bostonplans.org) Here is the zoning board of appeal web page. It appears you can search for a project that was reviewed. Getting a database of it is another matter. Zoning Board of Appeal scraped data - In GitHub so will have to download it
Suggested Steps	**Step one:** Finish data collection for preliminary races in the years: 2013 and 2017 Mayoral races, 2016 and 2020 presidential races, 2018's Capuano vs. Pressley race, 2020 Democratic Senate primary. Combine the election data with registration and American housing estimates (2017) data. **Step two:** Find and amend eligible voter data for these years as a new column. **Step three:** Overlay voter trends by race over the years, base geographic unit should be Ward Precincts **Step four:** Create visualizations to summarize trends and patterns in voting by race, geography, and other measures like turnout, blanks cast, etc. **Step five:** Add ZBA approved permits and find the column that might indicate property type being a condo **Step six:** Analyze which Ward Precincts predominantly experience condo conversions and their associated demographics and voting patterns.
Terminology	**Notes on elections** **https://docs.google.com/document/d/1G6I9yaFKCQaJ8Ck2h61LInhM88 4Rz_9GvQwiI-iZYLk/edit?usp=sharing** **Eligible Voters:** size of voting age population in a political unit **Percent registered:** number by political unit registered / total eligible **Turnout:** number of votes cast / registered **Blanks:** submitting a ballot but skipping one or more race

	Nonpartisan: city elections (mayor, city council) in Boston do not separate by party and elections are in odd-numbered years; **Partisan:** state legislature, governor, president, congress, etc, vote by registered party even numbered years. **ZBA**: Zoning Board of Appeal: City agency that approves larger development projects when they need exceptions to the rule, like building higher than the rule for that part of the city.. **Ward**: larger geographic unit for voting purposes. **Precinct**: smaller geographic unit. In db terms, Wards have precincts. Note: wards and precincts do not align to census tract data, zip codes, or neighborhoods. The boundary lines are available in geospatial format. In the voting results, ward and precinct are often presented as four character strings, 0414 would mean, the 14th precinct in Ward 4. **Preliminary or primary election**: first round, narrows down candidates to final two. **Final election**: picks the winner, all just like in the presidential election in the US last year. **Districts**: all voting districts are sets of wards and precincts within wards. A city council district, however, will not align to a state senate or house of representative districts.
Strategic questions	1. How have the demographic changes affected the prospects for candidates of color and white candidates? 2. How do racial voting patterns differ by election type? 3. What are the key predictors in determining support for Black candidates for Boston? 4. What are changes in District 3 particularly (Ward precincts:) 5. How has voter turnout by precinct changed across city council election year?
Additional Information	**Tools & Methods** *Data pre-processing:* Pandas, NumPy for processing and cleaning the data. *Data Visualization:* Matplotlib, Seaborn, Tableau for all kinds of interactive visualizations Steve comment: Need to correlate census tracts to ward and precinct. Source data includes this; do we need to verify it? Need to document procedure for this or redo, sorry.

Risks & Limitations	Potential risks:
	- Most likely all of our tasks are achievable within our given time frame. - There is a potential to go down a rabbit hole with our analysis. - Census data is out of date and does not necessarily line up with wards and precincts. Limitations of the data: - Most of the data sets are fairly complete and reliable. - Currently missing presidential election data and census data. - Presidential election data needs to be converted to csv. - Not all features are clearly identified. - Condo conversion data is difficult to find, may not be publicly available.