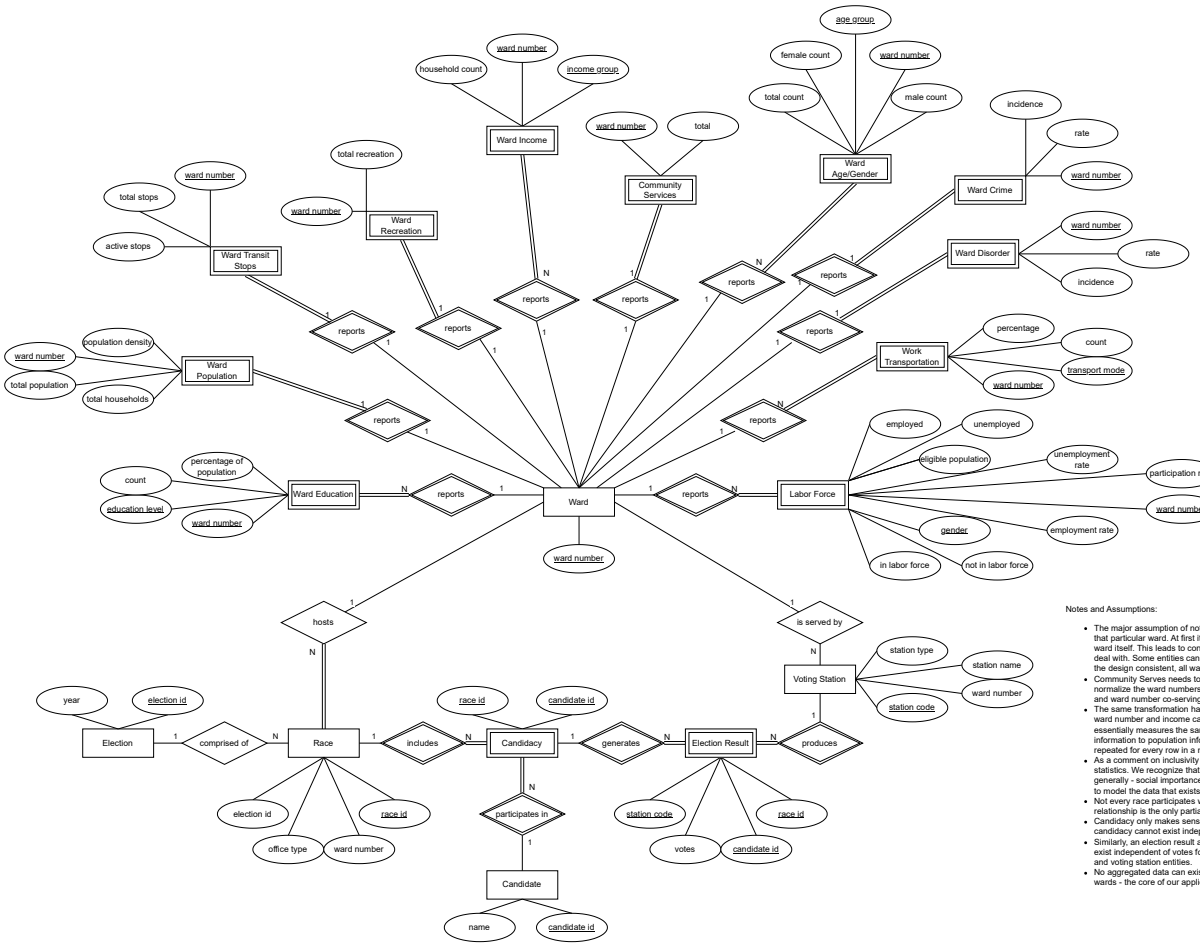

Notes and Assumptions:

- The major assumption of note here is the existence of Ward as a central entity connected to aggregated data about that particular ward. At first it seems like the most natural way to conceptually design this is through attributes of ward itself. This leads to conflicts of repeating groups and sets, which would make the system far more difficult to deal with. Some entities can be connected as attributes because of a 1:1 relation (like income), but in order to keep the design consistent, all ward facts are modelled as entities instead.
- Community Serves needs to be transformed in order to fit the data model. The base csv transformation is to normalize the ward numbers and enlist service types for each ward as a column instead. This results in service type and ward number co-serving as composite keys.
- The same transformation has to take place in the income group - we need the long-form version of the data where ward number and income category server as composite keys, since each separate column in the raw data essentially measures the same thing. As a consequence however, total households is shifted from the income information to population information for the sake of convenience. If we do not do this, then this information will be repeated for every row in a normalized ward income level.
- As a comment on inclusivity - ward information obtained from sources do not consistently include non-binary gender statistics. We recognize that these statistics contain important information regarding political leanings and - more generally - social importance. Because we feel we cannot accurate convey this information, the decision was made to model the data that exists for the sake of the project as is.
- Not every race participates with a ward directly - mayoral candidates are not representatives of wards - thus this relationship is the only partial relationship in the diagram.
- Candidacy only makes sense in the context of a candidate running in a specific race. Without these conditions, candidacy cannot exist independently.
- Similarly, an election result as modelled by our data (not necessarily winners, but vote counts per candidacy) cannot exist independent of votes for candidacy in a station. Thus, election results themselves are dependent of candidacy and voting station entities.
- No aggregated data can exist about a ward without the existence of that ward - thus, all aggregated data regarding wards - the core of our application - is entirely comprised of weak entities.