1. For this assignment, I am examining the walkability of my Somerville neighborhood, specifically the area between Powderhouse Circle, Teele and Davis Squares. Using local grocery stores, the Davis Square T station, schools, and libraries as points of reference as well as often-walked-to destinations, I will look at the data quality/accuracy of different data sources (specifically roads in relation to parcels) and how they measure up to the real world. This project will require a certain level of positional accuracy, i.e. the more accurate the data is, particularly inter-data relationships (roads to parcels), the less likely people will seemingly walking into buildings or across yards. However, positional accuracy is not absolutely crucial in a study such as this, mainly because basic inference skills will allow users to realize the gist of the information. If I were making a map for pedestrians to use as far as walking the neighborhood, they would know not to walk directly through a building but to go around it. However, if this information were being used to direct long range missiles, the need for accuracy would increase substantially.

2. The three road center line data sets I am using are the Massachusetts’ Executive Office of Transportation (EOT) data, Census 2000 Tiger data, and Street Map USA (SMUSA) street center line data. The center lines from the EOT are clearly the most appropriate for my specific project, that is, as the streets relate to the parcel environment. The EOT lines are consistently within the parcel streets. The center lines from TIGER and SMUSA vary widely. At certain points they line up nicely with those from the EOT, at other times they are very far off from not only the EOT lines, but each other as well, and often cut right through parcels where clearly buildings exist. In general, the TIGER and SMUSA lines are at least roughly paired together. See visuals.
The three data sets occasionally line up at intersections such as this.
Due to a lack of hydrography in my area of study, I chose to examine land use by parcel and land cover (parcel data from MassGIS and the land cover layer coming from NLCD via Barbara Parmenter). While the two layers do not express exactly the same information, it is interesting to examine how closely the two line up with one another as far as actual land use vs. the listed level of development intensity. One glaring question that is raised when comparing the two layers is how are the levels of developmental intensity measured? The following example compares two areas of the same neighborhood; one is labeled medium intensity development, the other high intensity development. I know from experience that these two areas are more or less identical to each other in building size, density, and use (both heavily or exclusively residential). So why are they classified differently as far as land cover?
Clearly the land cover data is not as detailed as the parcel land use data. The two coincide the best they can much of the time, see next image, but the previously shown image could be misleading if it were simply taken on the land cover layer: the patterns seem erratic.
4. The following are quantitative assessments of each data layer in this project using the land use by parcel data layer as a reference point, as that would be a key component of a walkability study, i.e. the buildings of the neighborhood.
   - The land cover layer cannot be easily quantitatively assessed, due to the fact that it is a more generalized assessment than say the land use by parcel layer. Land cover does not correspond to streets or specific parcels by which to measure it.
   - The EOT street center lines line up almost perfectly with the parcel layer. One exception is in powder house circle. However, the problem actually lays more in the parcel data than the street center lines. The parcel data is off by 50 feet. The island in the rotary (see following example) is only 10 ft. from the parcel at the tufts athletic field. In reality this distance is +/- 60 ft.
The accuracy of the SMUSA street center lines range from 0-110 ft off from the generally accurate EOT center lines. Largely the discrepancy is between 10-35 ft.

The accuracy of the TIGER street center lines is largely similar to the SMUSA lines with some notable variations. The accuracy, in terms of the EOT lines, ranges from 0-110 ft. off. There are also some strange twists and turns in the lines that do not exist in the real world. See images below.
• The school data is accurate in terms of the parcel layer.

• The library layer seems to be quantitatively accurate in terms of location; however, qualitatively it is not. Only one library appears for this data layer and it is the library within the Amnesty International office near Davis.
• The MBTA train station layer is not accurate. It can be assumed that the data is simply marking that the station is in Davis Square, which it is successful in doing. However, for the purposes of a walkability study or simply a neighborhood walking map, the data is misleading. In reality, there are two entrances to the Davis Sq. T station, neither of which is represented in the data. The data shows the T station to be on the wrong side of the street. See below.
- I geo-coded the grocery store layer during the in-class assignment on Tuesday. In short, in relation to all three road center lines and the land use parcel data, the street-level accuracy is not good. The two Teele Sq. stores are located in the middle of the street, as is the Davis Square grocer. Additionally, the data does not include all grocers in the zip-code as previously thought. At least one grocer, located in Davis Sq. has been omitted. See below for quantitative discrepancies.
5. I have already somewhat qualitatively addressed the positional accuracy of the four optional layers (groceries, T station, libraries and schools). As has been noted and evidenced in the previous questions, there are many areas in which the layers do not lineup well with each other. The TIGER and SMUSA road center lines cut through many buildings for example (see above images). In terms of the EOT center lines and the land use parcel later, the schools layer was lined up well, as was the library layer; however, some information was missing. One library and at least one grocery were omitted from the data layer. The MBTA train station in Davis Square was also not accurately placed on the map with regards to the parcel layer. It would be interesting to see the accuracy of these items when compared with orthophotos, but gauging from past assignments, I’m fairly certain that the Somerville land use parcel layer lines up well with the MassGIS orthophotos so I don’t believe it would be much different.

6. The optional layers I have chosen are not appropriately accurate for my chosen project. If I were to give this map to a person who does not know Somerville, there is no guarantee they would be able to locate all of the selected locations on foot. They would get to the school, as well as the Amnesty library, however, they clearly wouldn’t find the actual existing library and grocery store that are not listed on the map. If they were looking for the other grocery stores, they would probably be able to find the real stores based on a reasonable proximity, but some of them, particularly those in Teele Square, are far enough from the actual location that it wouldn’t immediately be obvious to someone standing at the selected locations. See above images for examples of this.
7. As far as completeness for each of the selected data layers, I’ve already mentioned a few incompletes based on my actual knowledge of the selected neighborhood (a missing library and grocery store). Aside from the missing information, and the clearly inaccurate positioning, the data sets, specifically the attribute tables, seem to be nicely complete for each of my selected layers (groceries, schools, libraries, and T station).

8. With regards to currency, my knowledge of the area has told me that the library and grocery store data layers are not up to date. However, I know that both the missing library and grocery store are not new additions to the neighborhood. There must be some explanation as to why they are missing from the data sets that I am unaware of. By examining the metadata for each data set, I saw that the MBTA station layer was last updated in 2006 and I know it hasn’t changed since then (I would classify that as up to date). The schools layer is listed as relevant for December 2007. Again I would classify that as relevant as I do not believe any major changes have occurred with schools in the neighborhood since that time. The library data is listed as relevant in early 2005. This could be questionable as I did not live in the area then. I know the building that currently houses the library that is not listed is a very old building, however, I would be very surprised if the current use as a library branch was any less than 4 years old. As far as the grocery layer is concerned, I pulled the data from Reference USA just a few days ago. As that company makes a living on providing up to date information, I’m sure the grocery address data is current. Given that, I’m not sure why a grocery store is missing unless it is listed as something other than its actual use.

9. The attribute accuracy for my optional layers is excellent, particularly for the needs of my fake-project of walkability in the Davis Area. The grocery, school, and library attribute information thoroughly include address and street info, the names of the institutions, etc. All the fields are filled out and accurate to my knowledge. The attribute information for the MBTA train station is lacking. All the fields are filled out; however, there are very few fields to begin with. It basically only gives the station name and the line of the T which services the station. This could be due to the fact that the data layer I used was simply an MBTA node layer—that is there could be a more detailed layer that exists specifically for train station entrances. I assume that is the reason for so little attribute information. The attribute table for the land use parcel layer is excellent for the needs of my project. There are very few holes in the data, and I checked the accuracy of the listed information for one parcel and it was spot on. The attribute info for the EOT streets is good. The addresses appear accurate, as well as street names, type, and class. There appear to be a few holes in the TIGER street info but nothing too glaring. Mainly, there are lines in the table that do not have any information in them. However, they are consistently sparse; that is, the lines with info are filled out fully and seemingly accurately, and those with no info are fully empty with the exception of the object ID field. The SMUSA attributes seem fine as well.