Quarter Long Project

Final Project Report

Michael J. Newman

The Pennsylvania State University

Geography 586, Fall 2 2013

INTRODUCTION

As an adult leader in the Boys Scouts of America, I've been "volunteered" as the chairperson for the "Scouting for Food" (SFF) program in my area. Primarily I received this assignment because others know that I'm "handy with maps". Started in 1985, SFF is a national food drive service project administered at a local level. My local Scout Council, Pennsylvania Dutch Council–BSA, has organized the food drive in partnership with food banks that serve local school districts. The food drive occurs over two weekends, on the first weekend scouts deliver grocery bags and instructions to homes. A week later, scouts pick up grocery bags full of donated items. The donated items are sorted and delivered to a local food bank.

In the Warwick School District (WSD), 10 Scout units collect donated food from approximately 10,000 residences and deliver it to the Lititz/Warwick Community Chest (LWCC). LWCC delivered food to 273 families in 2011 and 523 families in 2012. Historically, the WSD area yields approximately 21,000 pounds of non-perishable food items. Last year, SFF delivered 17,000 pounds of food to LWCC. The 20% decline may have been caused by incomplete collection of donated food. Anecdotal evidence from the LWCC suggests that possibly thousands of homes were missed in the collection effort last year. The number of Scouts has decreased while the number of homes in the collection area continues to rise. The population for WSD has risen from 21,000 in 1990 to 31,000 in 2010 with a proportional increase in housing units and roadways. During that same time period the number of Scout units in WSD has dropped from 18 to 10 and in the next 6 months will drop further to just 8.

The last complete reassignment of SFF collection areas was done sometime in the 1990's by hand with liberal use of photocopied ADC map books and highlighters. The existing areas no longer make sense because they don't reflect the demographic and geographic changes over the past 20 years.

RESEARCH QUESTIONS / OBJECTIVES

Aims (as Stated in Final Project Proposal)

> *My intent is to find the minimum number of areas that will cover the greatest possible number of homes with a minimum of hassle. The newly created areas will be based on: land use, address density, availability of sidewalks, road hierarchy, and Scout unit demographics. I will additionally consider if creation of areas that are different for the distribution of the empty donation bags versus collection of the full bags is worthwhile.*

Re-stated Objective

In the face of rising food demand, with the falling number of scout units and the reduction in collected food, now is the time for creating new SFF collection areas. My objective is to modernize and reorganize the Scouting for Food collection areas within WSD. The new collection areas will: cover 100% of the school district; be sized proportionately based on the number of registered scouts per unit; and be located to place younger scouts in areas with sidewalks or closed neighborhoods.

Deliverable Items

- Presentation-size wall maps for each scout unit depicting their respective collection area in PDF format.
- Letter-size map book in PDF format for field use during SFF bag distribution and collection.
- ESRI file geodatabase containing all project data, suitable for redistricting in following years or generating additional products.
- Google Earth KMZ file of SFF collection areas for desktop mapping without need of ESRI software.
- Form template in Excel and PDF format for LWCC to record pounds of food collected to use for future analysis.

DATA: SOURCES, PREPARATION, AND ANALYSIS

Data Sources

Lancaster County, Pennsylvania GIS Department

- Geocoded Address Points (one per postal address)
- Cultural Features (for final map products)
- Land Parcels with Land Use classification
- Road Centerlines with Classification & Routing attributes
- Warwick School District Boundary

Boy Scouts of America, Pennsylvania Dutch Council, Lancaster, PA

- Scout Unit Demographics Table
    - SFF Participants
    - Unit Type
    - Unit Number

See Appendix A for a complete list of feature classes used.

Preparation

The data received from Lancaster County GIS was in shapefile format. All shapefiles were imported into a file geodatabase for easy access and increased geoprocessing speed during analysis. All data was in the same projection (NAD83, PA State Plane South, US Feet) and thus no further transformations were required. To further shorten geoprocessing and screen redraw times, a study area was established by creating a 1 mile buffer area (to mitigate edge effects) around the Warwick School District. Analysis layers were created by clipping addresses, parcels, and road centerlines to the study area.

Analysis

Finding Residential Addresses

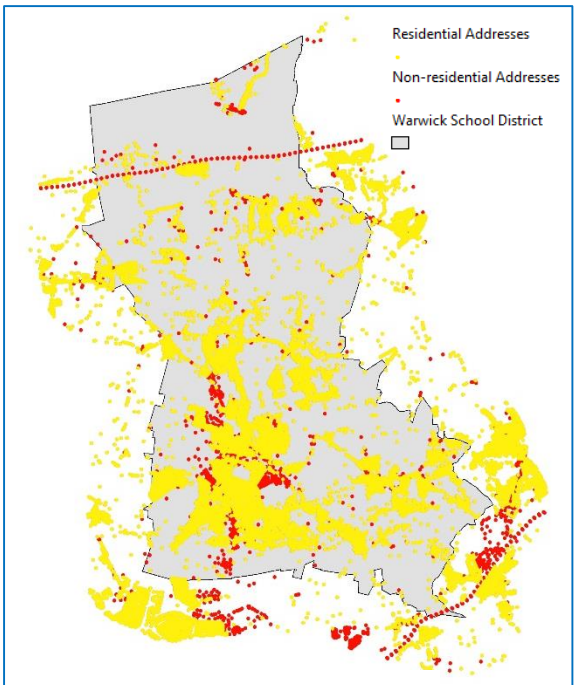

FIGURE 1: Address classification results.

Scouting for Food relies on food donations from private households; scout units are instructed to deliver bags to residential addresses only. People involved in emergency services and postal delivery are usually very familiar with the geography and address distribution within a community, but the average citizen is largely ignorant of how many residences might be on a particular street or in a neighborhood. Children are even less likely to have a grasp of which properties are residential or commercial. In Lancaster County, no agency has categorized lists of addresses.

The first step in this project was to find only the residential addresses within the study area. The parcel coverage contains an enumerated field for land use code. To find residential addresses a spatial join was performed on the parcel and address feature classes: addresses inherited the land use code of their containing parcel. A new feature class containing only residential addresses was then created by querying and extracting points having land use codes indicating the presence of housing units. Figure 1 (above left) shows the distribution of residential and non-residential address points in the study area. The study area contained 19,149 addresses, 16,740 of which are residential. 11,202 residential addresses fall within Warwick School District.

(As an aside, the linear stretches of non-residential address points crossing the northern end and southeast corner of the study area are address points assigned to 1/10 mileposts on the PA Turnpike and US 222 respectively. While not truly "addresses" these exist for emergency service delivery.)

Road Hierarchy & Sidewalk Presence

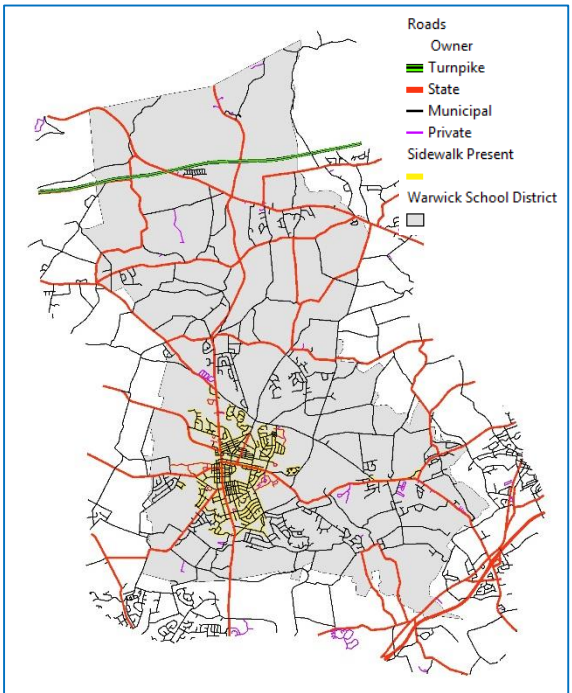

FIGURE 2: Road Hierarchy and sidewalk presence.

To deliver and collect the SFF grocery bags, scouts end up doing a lot of walking or hopping in and out of vehicles. Most drivers don't expect kids walking along roadways on a Saturday morning. Road safety is a primary concern for the boys as drivers speed past unaware of the scouts. In interviews with adult leaders, a high priority was placed on allocating areas so that Cub Scout Packs (boys aged 7 -11) would have access to sidewalks or, in the absence of sidewalks, be assigned less-traveled roads. Figure 2 (left) illustrates road hierarchy: Turnpike and State roads have higher daily volume than municipal or private roads. Sidewalks are concentrated in a relatively small geographic area in and around the central business district of Lititz Borough. The road ownership data was attached to the road features, so analysis was done through symbolization. Sidewalk information was not present within the road data, this information was manually appended to road segments based on local knowledge and reference to 2012 orthophotography. Road hierarchy and sidewalk presence will be factored into which unit gets assigned a certain area.

Address Density Analysis

In an effort to further define which areas are better-suited for Cub Scout Packs versus Boy Scout Troops several methods of density analysis were performed. The central idea being that areas of highest density should take the least amount of effort for the distribution and collection of SFF grocery bags. Specifically, the densest areas would be better for Cub Scout Packs (younger boys) that have shorter attention spans and less physical stamina.

Kernel Density Estimate (KDE)

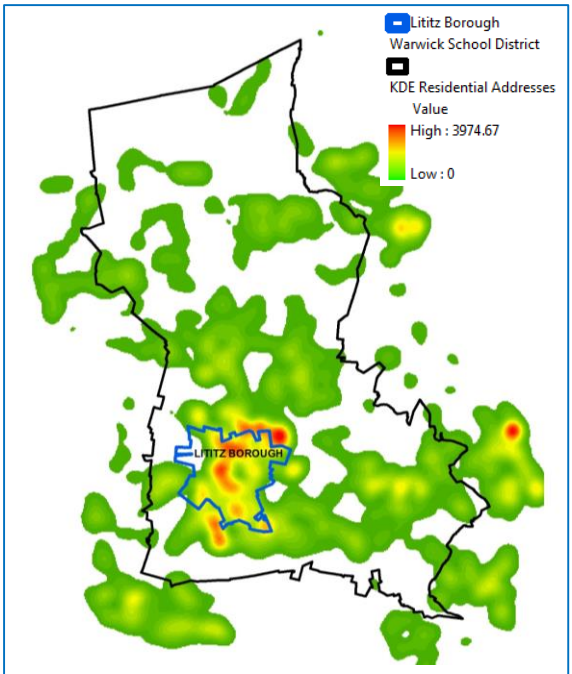

FIGURE 3: Results of Kernel Density Estimate performed on residential addresses.

Figure 3 (left) shows the results of a kernel density estimate performed on the residential address feature class. The densest areas are shown in red and occur within the bounds of Lititz Borough. Note that the areas of densest addressing coincide with the areas of densest road network show previously in Figure 1. The conclusion from this map would be that Cub Scout Packs should be assigned in, or around, Lititz Borough while Boy Scout units be assigned areas outside of the Borough.

Built-Up Areas

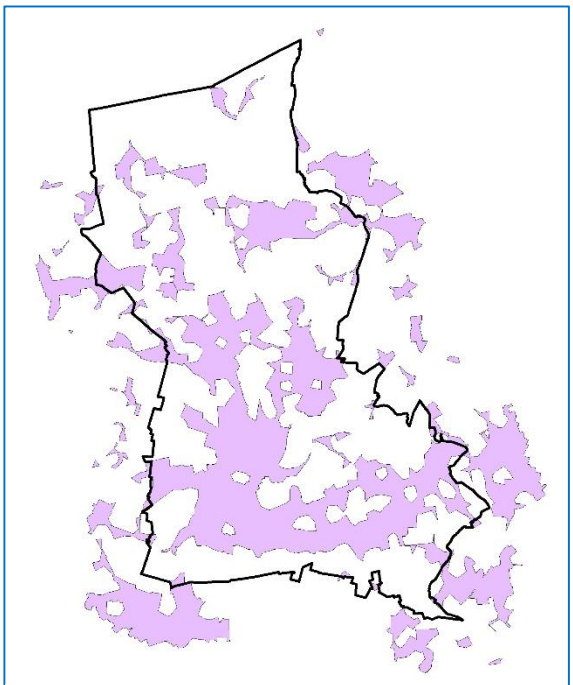

FIGURE 4: Results of Built-Up area Delineation.

Figure 4 (left) shows the results of a tool from the Cartography Toolbox: Delineate Built-Up Areas. According to ArcMap Help 10.2 Online, this tool creates polygons to represent built-up areas by delineating densely clustered arrangements of buildings on small-scale maps. Intended to help cartographers shade "city areas" in map books, this is powerful tool. The results are customizable based on user input of minimum cluster distances and minimum number of cluster members. The resulting polygons can be snapped to user-defined line features such as municipal boundaries or roads. This analysis looked for clusters of at least 4 addresses within 0.2 miles of each other and snapped the results (when possible) to road centerlines. The shape of the output looks surprisingly similar to the densest areas of the Kernel Density Estimate even though it relies on an entirely different set of underlying geoprocessing tools and calculations. This analysis yielded a starting point for collection area outlines.

"Blocks": Address/Acre

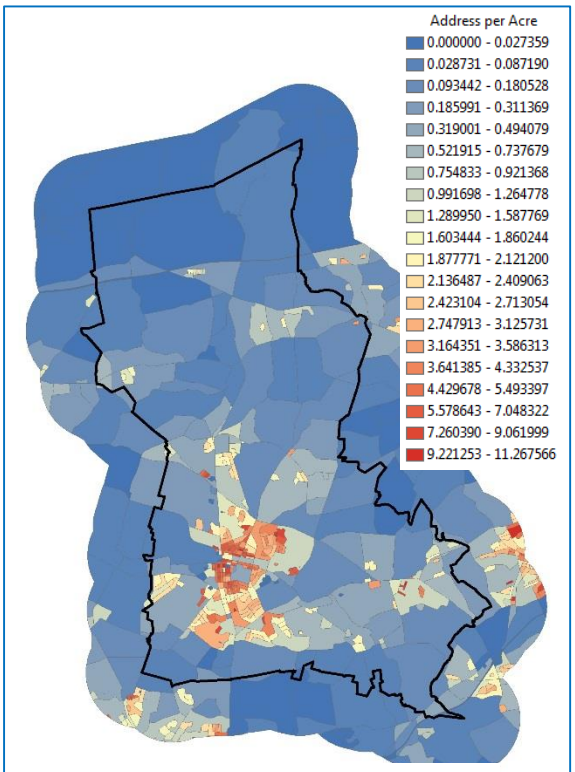

FIGURE 5: Addresses per Acre assigned by "block". Categorized by Jenks Natural Breaks

Figure 5 (left) shows the density of addresses / per acre within a given "block". The block areas are polygons defined by the intersection of roadways. The blocks were built by generating polygons from road centerlines. Address points were then spatially joined to the polygons to obtain a count of addresses within each block. A geographic calculation was then applied to each polygon to compute the number of acres in the polygon divided by the number of addresses.

By categorizing the output, a density pattern similar to that of the KDE is revealed. The benefit of this analysis is that it produced a series of vector polygons with attached address counts. Aggregate block groupings (with associated address counts) can be created by selecting and merging individual blocks. This analysis was the result of failed parcel aggregation attempts documented later.

Address Apportionment

Unit Name	# Scouts in Unit	% of Total Scouts	Addresses Apportioned
Pack 27	3	1.02%	114
Pack 44	33	11.19%	1,253
Pack 100	35	11.86%	1,329
Pack 142	36	12.20%	1,367
Pack 154	31	10.51%	1,177
Troop 27	8	2.71%	304
Troop 44	22	7.46%	835
Troop 100	10	3.39%	380
Troop 142	80	27.12%	3,038
Troop 154	37	12.54%	1,405
TOTALS	295	100%	11,202

TABLE 1: Address apportionment based on scouts in unit as a percent of total scouts.

During interviews with current scout leaders, no one is sure of how areas were previously assigned. However the common understanding is that areas were assigned based on land area. Someone tried to make things "fair" by visually balancing polygon areas by hand. It seems that most people were happy with this arrangement when each scout unit had roughly the same number of scouts and no one was interested in "wasting time" on figuring out new collection areas.

The past five years have seen some significant shifts in the number of scouts in each unit, but the areas remained the same. Units with a lower number of scouts were unable to complete their assigned areas in a timely fashion, leaving hundreds (thousands?) of homes without SFF grocery bags. Every house that did not receive a bag was one less opportunity for community participation in the food drive, lowering the yield to the LWCC. To prevent another occurrence of "missed houses" future areas will be created with an eye towards available manpower: each unit will be assigned the number of addresses that is proportional with their number of scouts compared to all scouts within the WSD area. Table 1 (above left) shows the ideal allocation of addresses based on this approach. This can be expressed by the following equation:

$$\text{Assigned Addresses} = (\text{Total Addresses}) \frac{\text{Scouts in Unit}}{\text{Scouts in WSD}}$$

Putting it All Together: Overlay

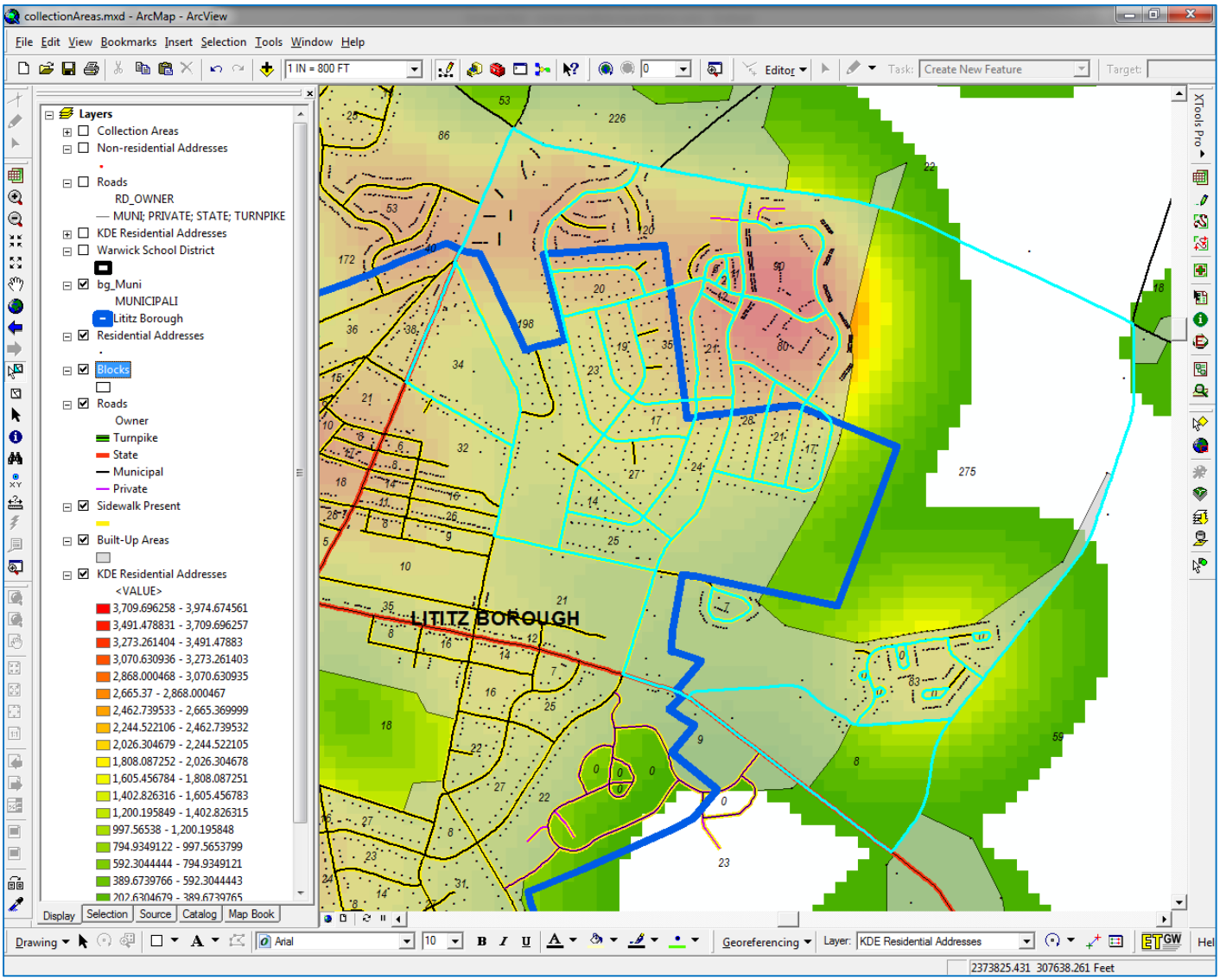

FIGURE 6: Screen capture from ArcMap Desktop during creation of collection areas. Highlighted blocks will be merged to become Area E for Cub Pack 154.

Figure 6 (above) is an in-process screen capture from the final overlay analysis. Creating the final collection area polygons seemed best accomplished through overlay of all preceding analysis output. Attempts at raster-based categorization and map algebra did not yield satisfactory results (see the next section for discussion) and nearest neighbor vector aggregation of blocks failed as well. The best results were obtained by manually selecting and merging blocks, this allowed for decision making “logic” on-the-fly, in particular related to exception situations. During this process it also became clear that using only density measures would “leave out” a substantial number of rural and farm addresses.

RESULTS

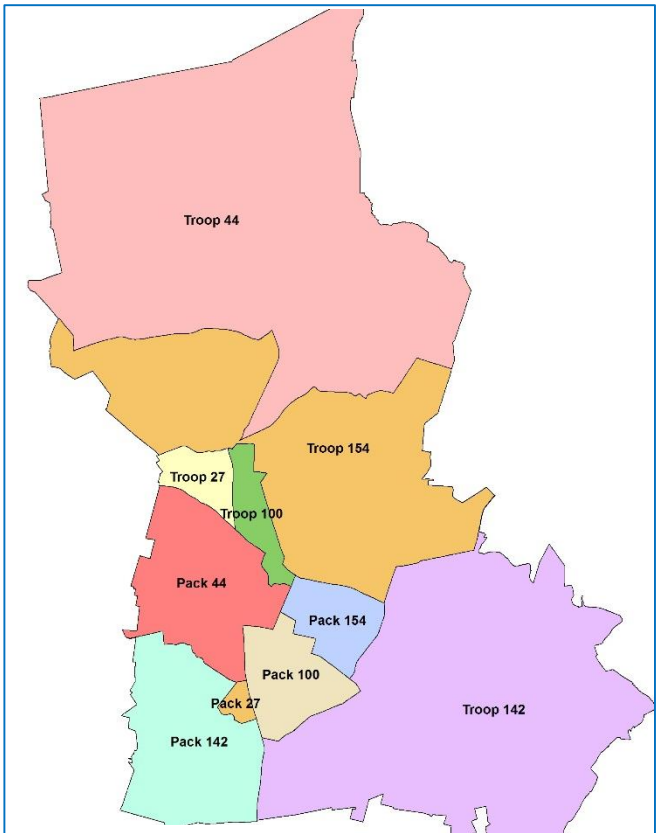

FIGURE 7: New collection areas.

Figure 7 (left) shows the final collection areas resulting from the overlay analysis. Table 2 (below left) shows the number of addresses in each collection area compared to the ideal apportionment described earlier. Based on these two graphics, the objectives for this project have been met: 100% of residential addresses are covered; Cub Scout Packs are clustered in the area of highest address density (with sidewalks); and the quantity of addresses assigned reflects the work capacity of each unit based on registered scouts. The deliverable items resulting from this project are available for review at http://www.thenewmans.us/geog-586-report/ .

Unit Name	Addresses Apportioned	Actual Addresses	% Deviation
Pack 27	114	159	40%
Pack 44	1,253	1,416	13%
Pack 100	1,329	1,519	14%
Pack 142	1,367	1,391	2%
Pack 154	1,177	1,217	3%
Troop 27	304	254	-16%
Troop 44	835	969	16%
Troop 100	380	427	12%
Troop 142	3,038	2,576	-15%
Troop 154	1,405	1,274	-9%
TOTALS	11,202	11,202	0%

TABLE 2: Comparison of ideal apportioned addresses and actual collection areas.

Collection areas based on Density

In the initial conception of this project, the collection areas were to be heavily based on the density of addresses in an attempt to make distribution and collection efforts easier on the volunteer participants. Using a density-based approach would come at the expense of visiting less densely settled areas in rural or farm areas. Upon interviewing the LWCC director, it became clear that visiting every available address was imperative.

Based on LWCC estimates, one out of four households makes a food donation and the average donation is approximately 6.8 pounds. By extrapolation, for every 100 homes “left out” of bag distribution there will be a loss of 170 pounds in the food pantry. The first density analysis (not shown in this report) resulted in 9,512 addresses: the more efficient collection areas would leave out 1,690 addresses of the actual 11,202. This could amount to a potential loss of 2,873 pounds of food. After seeing the results of this analysis it is clear that effectiveness (getting as much food as possible) is linked to visiting every address possible. As a result, the density-based analysis was replaced in favor of an overlay based approach. The question became not which addresses to collect from, but who would do the collection.

Other Abandoned Analyses

Raster Allocation

Several attempts at area allocation using raster processes were attempted. In the end this just didn’t work because there is only one drop-off location (LWCC food bank) and the scout units are hosted at churches very close to the drop-off location. The allocated areas just didn’t make any real-world sense and there was no way to balance areas based on scouts per unit.

Parcels, Neighbor Aggregation

Several attempts were also made at aggregating collection areas by merging parcels through nearest neighbor proximity and adjoining sides. These attempts failed due to the topology of the parcel feature class. Where parcels touch parcels the boundaries are coincident, but roadways in Lancaster County, PA are not designated as parcels. Basically every roadway in the county creates a corridor of null space in the parcel coverage: the resulting discontinuities made it impractical to work with parcels for analysis. While this analysis failed, it did spur the idea of creating 'blocks' extracted from the road centerline coverage, which ultimately became a crucial part of the final overlay analysis.

Separate Areas for Distribution and Collection

Again, while this might be a more efficient model it was at odds with the capability of scout units. Each unit follows the same pattern: scouts distribute bags in week one and then collect the bags from the same area in week two. The overriding logic is that the scouts remember details about where they dropped off bags (use the side door of the blue house, here is the house with the angry dog, there are two houses here with a common driveway, etc.) that make for faster and easier collection. While logistics companies like UPS and FedEx might use separate delivery and collection routes, the scouts in this area will continue with consolidated areas as in past years.

Road Network Routing

This was left as a last analysis step and some of the service area tools do look promising. The road centerline data from Lancaster County does include attributes to facilitate road network creation (one-way directionality, lane counts, overpasses, impediments) but a lack of familiarity with the Network Analyst extension posed too big of a challenge to complete within the available time for this project.

REFLECTION

THIs has been a very satisfying project: allowing me blend two passions (GIS and Scouting) to accomplish something substantial in both areas simultaneously. The deliverable items from this project will be used to improve a worthwhile service project that impacts hundreds of needy families in my community. At the same time I was able to apply previously theoretical analysis methods to a “real world” problem. While I am not completely facile with all of the analysis methods in this project I believe that the results are accurate and that given more time using these tools I will be able to produce more effective results in less time.

I regret not having more time to find a final allocation method that is more “automated”. I would like to be able to apply the same methods of analysis to other areas within Lancaster County (I have all the required data) but the manual overlay process is too time consuming for me to pursue. I think that a natural avenue for further exploration of this project would be to engage in automation via Model Builder or Python. The crucial hurdle is in generating in logical flow of decisions. After deciding on the number of collection areas (equal to number of scout units) there must be a way to determine starting points for the analysis.

Finally, I’ve documented some of my most significant failures earlier in this report, but have not pointed out the work that is most pleasing to me: the deliverable map products. As I wrote in the introductory forum thread for this course, I want to use GIS to create things that people can use effectively. The wall maps and map book created for this project are truly useful products for the Scouts. Compared to the degraded photocopies they have been using for years the new maps will be a welcome sight. I’ve already delivered draft copies of the documents to area scout leaders and they are very excited to have a clear and easy reference that can be distributed to all of their members. Who knows, some scout looking at the map book might be inspired to pursue GIS on his own someday.

APPENDIX A: Lancaster County GIS Feature Classes Used

Feature Class	Feature Type	Content	Use
bg_AddrPnts	Point	Address Points	Analysis
bg_bldgs	Polygon	Building Footprints	Map Background
bg_brdgcov	Polygon	Covered Bridge Locations	Map Background
bg_bridges	Polygon	Roadway Bridge locations	Map Background
bg_CADRoads	Line	Road Centerlines	Analysis & Map Background
bg_cmtry	Polygon	Cemetery Extents	Map Background
bg_drives	Polygon	Paved Driveway Surfaces	Map Background
bg_EMS	Points	EMS Station Locations	Map Background
bg_Fire	Points	Fire Station Locations	Map Background
bg_hydrol	Line	Stream Centerlines	Map Background
bg_hydrop	Polygon	Water Body Locations	Map Background
bg_Muni	Polygon	Municipal Boundaries	Map Background
bg_parcel	Polygon	Parcel Boundaries	Analysis
bg_parks	Polygon	Park Locations	Map Background
bg_pklots	Polygon	Paved Parking Lot Surfaces	Map Background
bg_police	Points	Police Station Locations	Map Background
bg_rdedge_p	Polygon	Paved Roadway Surfaces	Map Background
bg_rrlines	Line	Railroad Track Centerlines	Map Background
bg_schooldist	Polygon	School District Boundaries	Analysis
bg_specneed	Point	Public & Private Facilities of Note	Map Background
a_LandUseCodes	Table	Enumerated Land Use Codes	Analysis