Urban Design and Research Section

Bridging Data Silos in Local Government

Examples from Arlington County, Virginia

September 30, 2016

Andrew D'huyvetter AICP, Associate Planner

Department of Community Planning, Housing and Development
Planning Division
Urban Design and Research Section

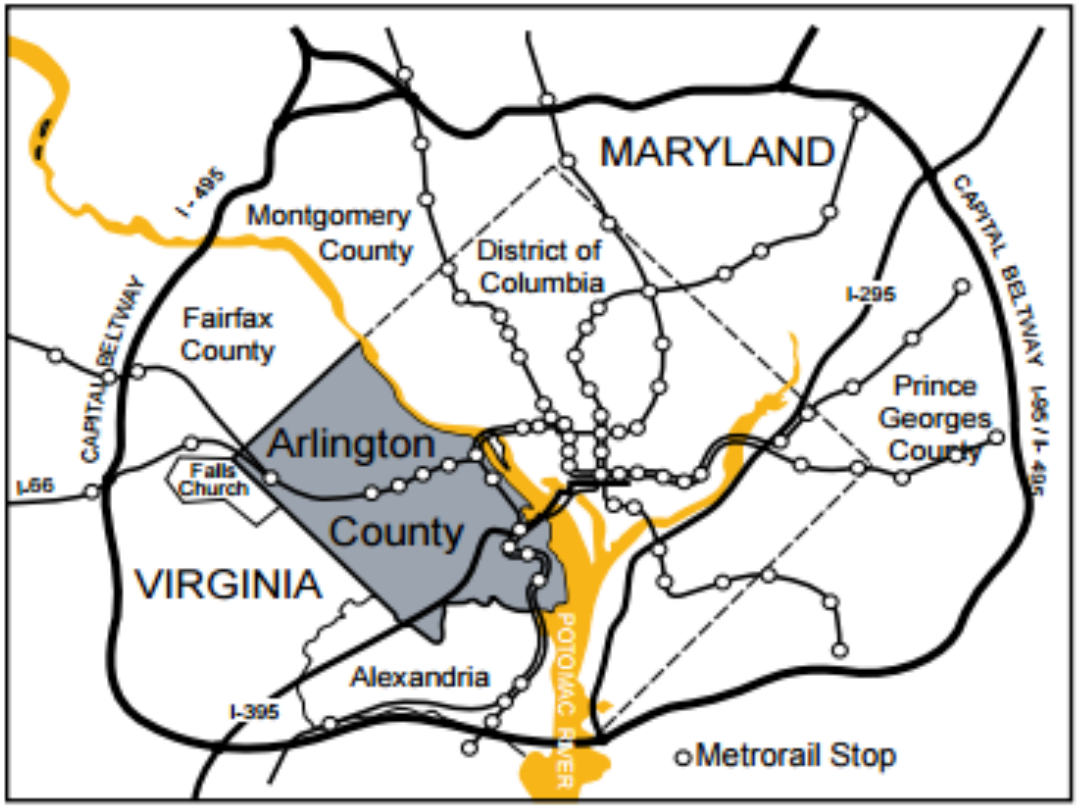
ARLINGTON
VIRGINIA

Agenda

- Local Government Context
- Data Problem #1: Counting Housing Units
- Data Problem # 2: Estimating Employment
- Challenges
- What we've learned
- Next Steps

Local Government Context

ARLINGTON IN THE DC METRO AREA



Local Government Context

Arlington County

- 25.8 square miles
- Population: 220,400 (2016)
- Employment: 211,000 (2016)
- Housing Units: 112,300 (2016)
- Daytime Population: 288,000 (2016)
- Total Office Space: 40 Million SF (2016)
- Total Retail Space: 8.4 Million SF (2016)
- 11 Metrorail Stations

Local Government Context

Local Government Context

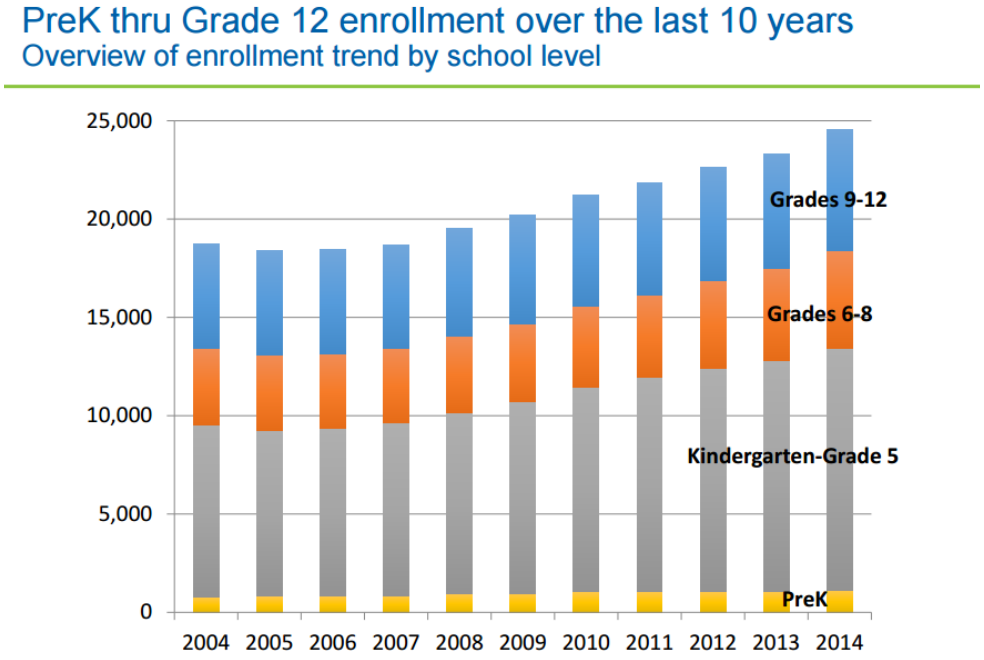
- Demand for high quality service delivery.
- Limited financial and staff resources.
- Arlington public expects you to have the data.
- American Community Survey is unreliable at small geographic scales and difficult for the public to understand.
- Need to supplement federal data with local administrative data.

Local Government Context

- Local governments are a wealth of data
  - Data is generated primarily for financial transactions. (permits, taxes, billing, grant requirements, $$$$, etc.)
  - Local governments should have lots of data…
- We've solved the easy problems.
- Complicated Policy problems require new research approaches
  - How many housing units are in Arlington County?
  - How many jobs are there in Arlington County?

Problem #1: Counting Housing Units

- Problem #1: School overcrowding
  - Significant increase in Public School students over the last 10 years.
  - Demographic shifts move faster than conventional projection techniques.



Problem #1: Counting Housing Units

- Need a comprehensive housing unit data set to create new and improved student generation rates.

All Grade Levels (K-12)					
Housing Type	**APS K- 12 Students**	**% Students by Housing Type**	**Housing Units Countywide**	**% of County Housing Type**	**Student Generation Factor**
Single Family Detached	12,912	52.79%	27,564	25.05%	0.468
Apartment- Garden	5,038	21.76%	15,190	13.81%	0.332
Apartment- Elevator	2,726	11.67%	34,837	31.66%	0.078
Duplex	877	3.78%	2,240	2.04%	0.392
Condo- Garden	1,059	4.53%	12,828	11.66%	0.083
Condo- Elevator	656	2.73%	13,783	12.53%	0.048
Townhouse	654	2.74%	3,582	3.26%	0.183
Total	**23,922**	**100.00%**	**110,024**	**100.00%**	**0.217**

- Census data limited
- So we had to build it.

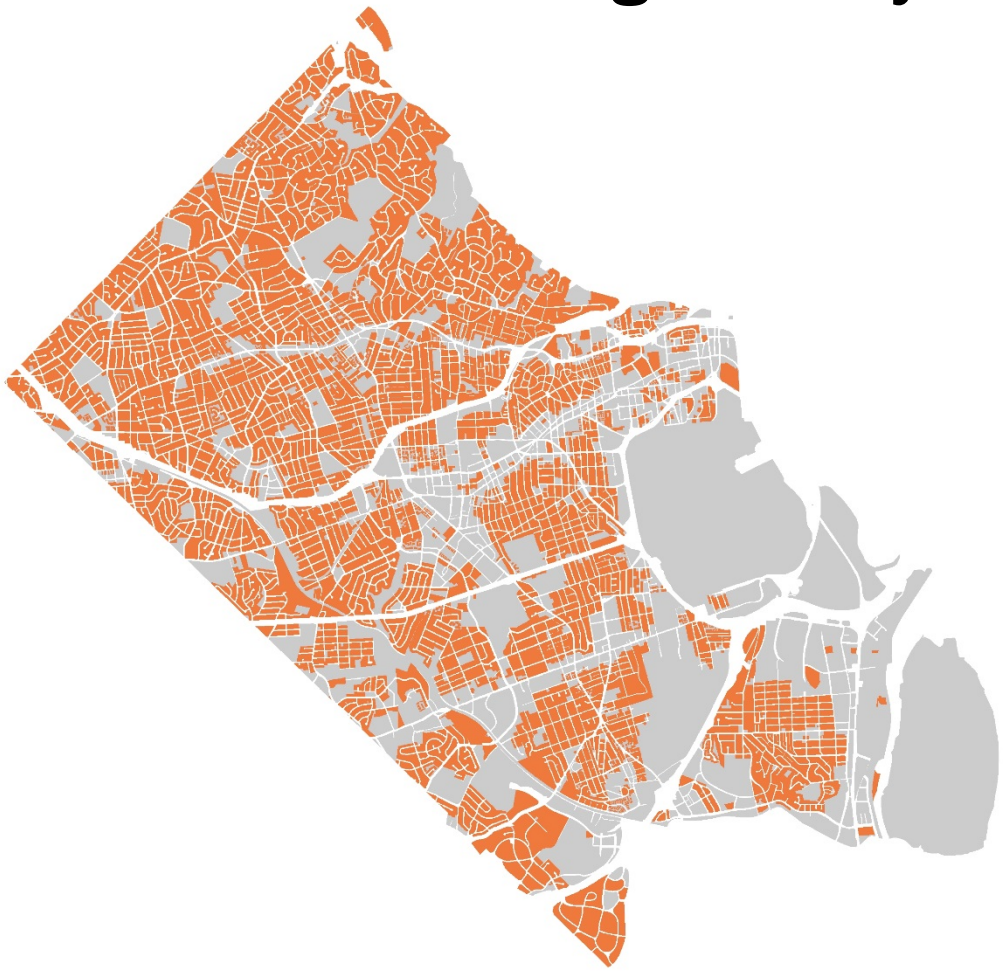
Problem #1: Counting Housing Units

- Arlington County has never before constructed a dataset of absolute housing units defined by housing unit type with detailed characteristics

	Single Family	Multifamily Condo	Apartments
Real Estate Assessments	**X**	**X**	*****
ATRACK (rental apartment survey)			**X**
Census Bureau	**Sample of Housing Units**		
Permitting System	**Not equipped to reliably count housing units.**		

Problem #1: Counting Housing Units

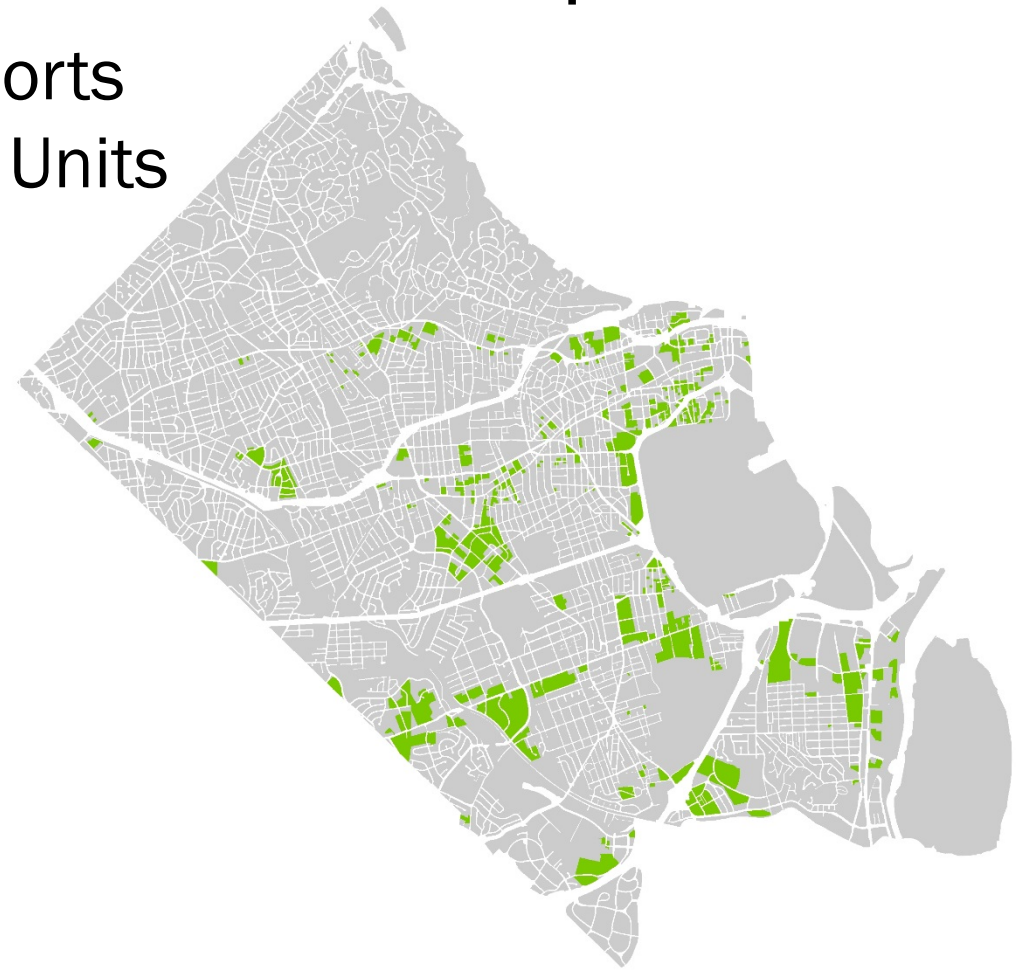
Real Estate Assessments: Single Family and Condo



Problem #1: Counting Housing Units

Real Estate Assessments: Apartments

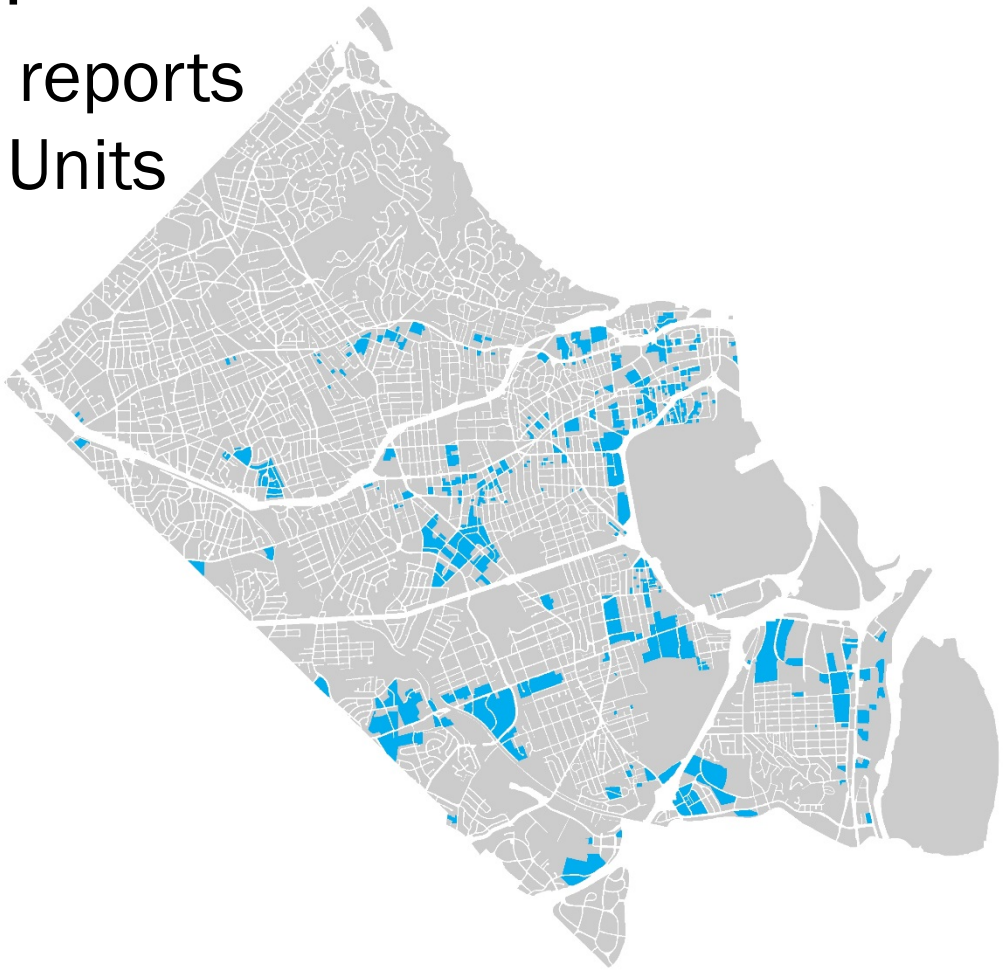
- REA reports 42,948 Units



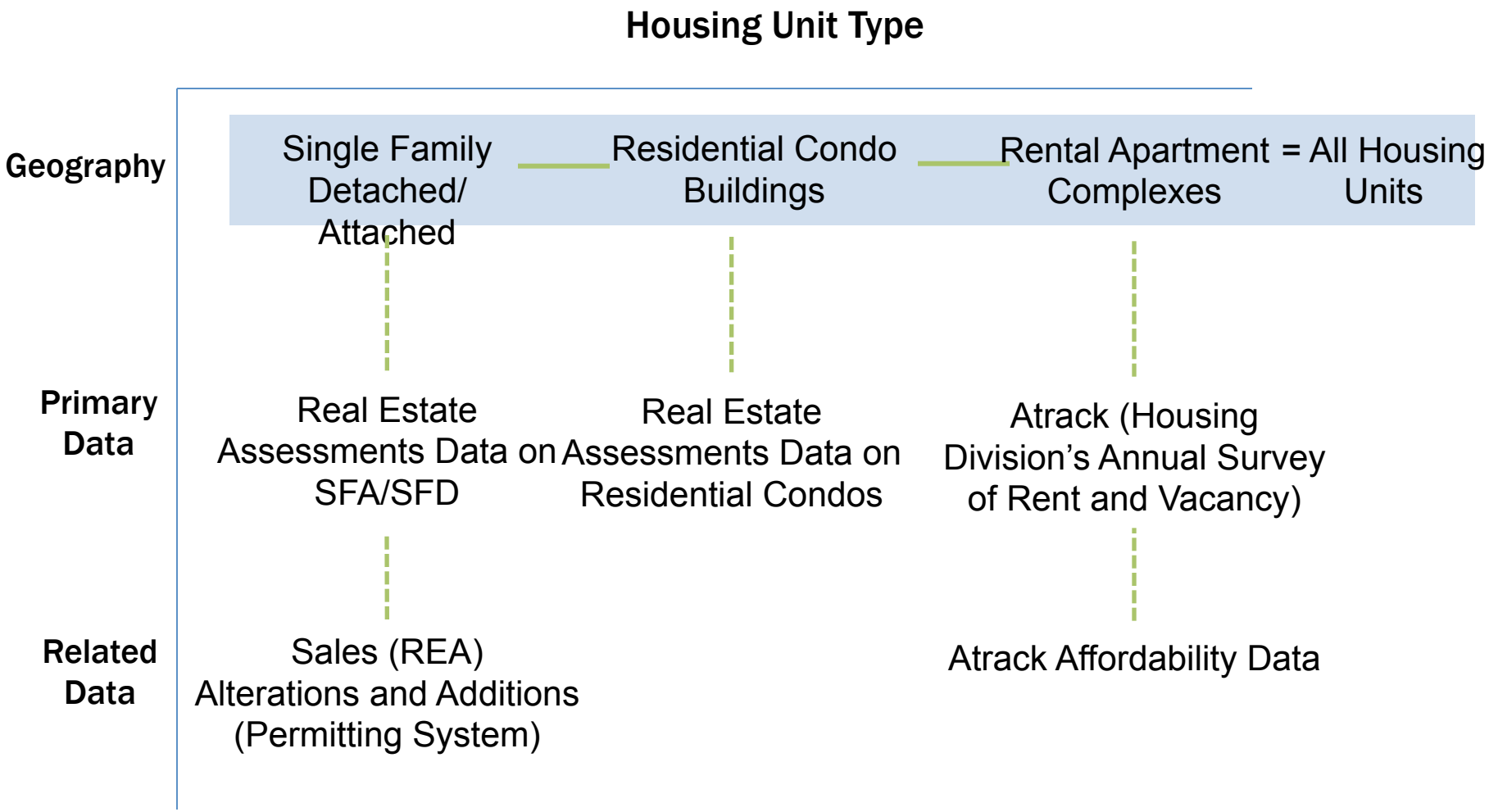
Problem #1: Counting Housing Units

ATRACK: Apartments

- ATRACK reports 50,312 Units



Problem #1: Counting Housing Units



Problem #1: Counting Housing Units

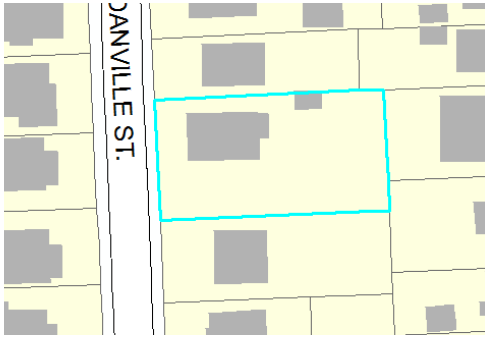
- Almost a year to completion.
- Coordination with the Mapping Center, IT staff, and CPHD's Housing Division.
- Significant data cleaning, creation of new geographic data, and an overhaul of Atrack.
- GIS (Geographic Information Systems) was the way to link databases together.

Problem #1: Counting Housing Units

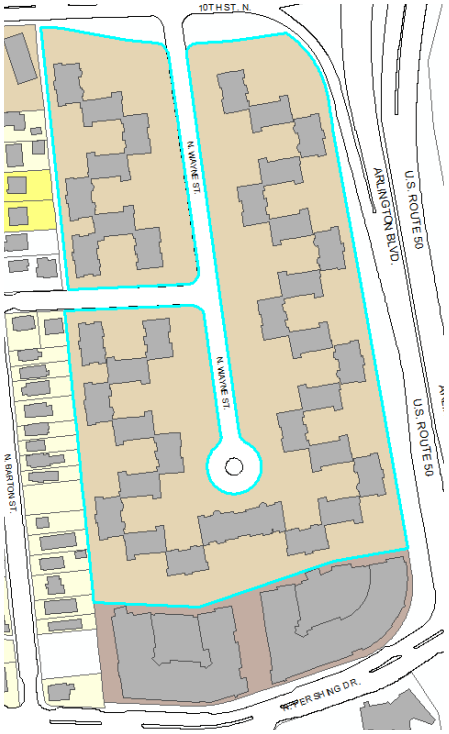
Results: Housing Unit Type Dataset

- GIS layer that can provide analysis at the smallest level of geography.
  - parcel and
  - complex level.

Single Family Parcel

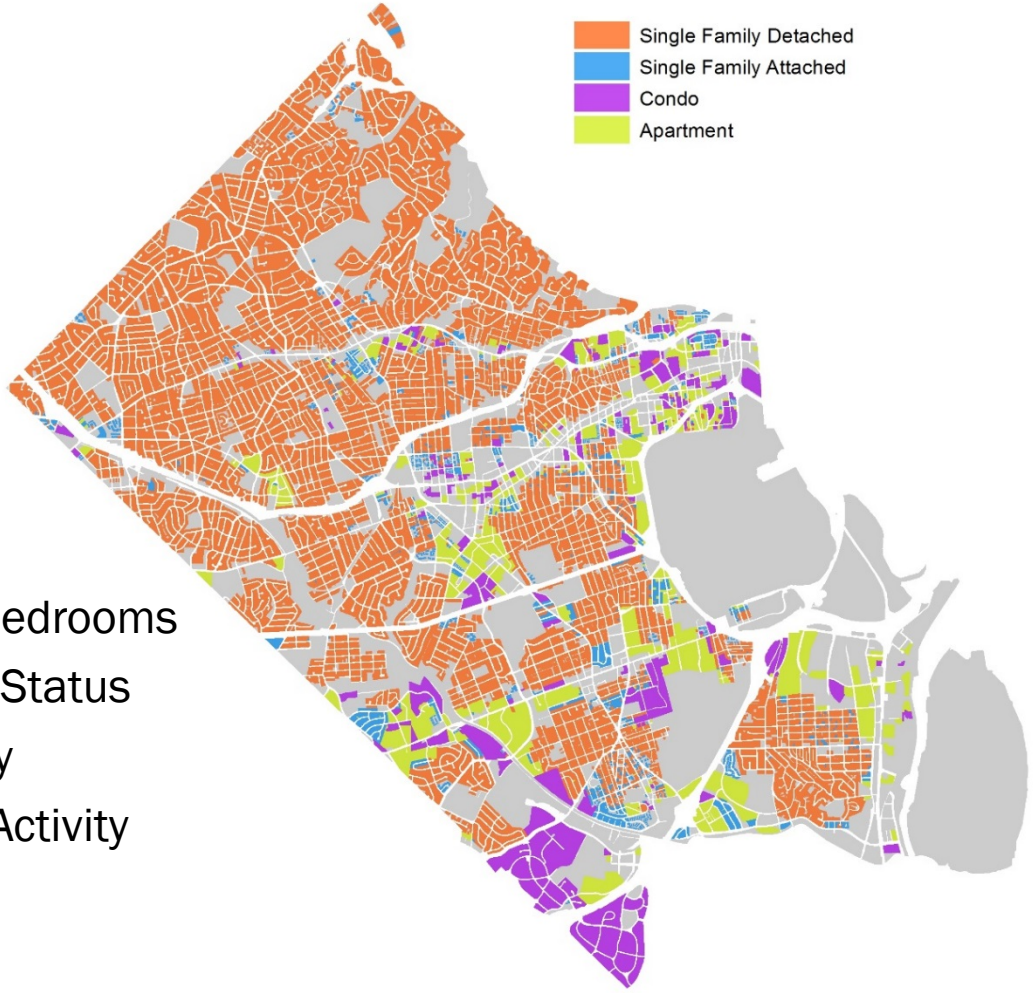


Rental Apartment Complex



Problem #1: Counting Housing Units

Results: Housing Unit Type Dataset



- New Features
  - Number of bedrooms
  - Affordability Status
  - Sales Activity
  - Renovation Activity
  - Year Built.

Problem #1: Counting Housing Units

Results: Housing Unit Type Dataset

- When joined with student data enables exploration of highly detailed student generation rates
  - Improves student projection accuracy
  - Allows for study of the housing supply and student growth over time
- Measure neighborhood change and turnover
- Combine with commission of revenue data on vehicle registrations to explore parking demand

Problem #2: Estimating Employment

- Problem #2: How many jobs are in Arlington?
  - Impacts transportation funding sources
  - Conventional Employment Data Sources cover only some types of employees and have wide disparities in their estimates.
    - 20,000 job difference between county estimates and conventional estimates. $$$$$
  - Arlington is unique: It has a large federal presence.

Problem #2: Estimating Employment

- Arlington employment primarily comes from office buildings.
- Significant federal presence in leased or federally owned sites.
- Numerous military facilities and contractors.
- Problems:
  - Inaccurate reporting of employees to worksites.
  - Inaccurate reporting of the number of employees.
  - Inaccurate or missing address information.
  - Inaccurate reporting of federal employees.

Problem #2: Estimating Employment

- Ongoing efforts with Virginia Tech to:
  - **Analyze** Quarterly Census for Employment and Wages (QCEW) data.
    - Document and try to resolve inaccuracies.
  - **Supplement** the employment data with other County, Private, or Federal data sources.
  - **Find a surrogate** for employment data. Evaluate whether building measurements such as water usage data or cell phone data can be used as proxies for employees in occupied buildings.

Problem #2: Estimating Employment

Proof of Concept: Building Water Usage

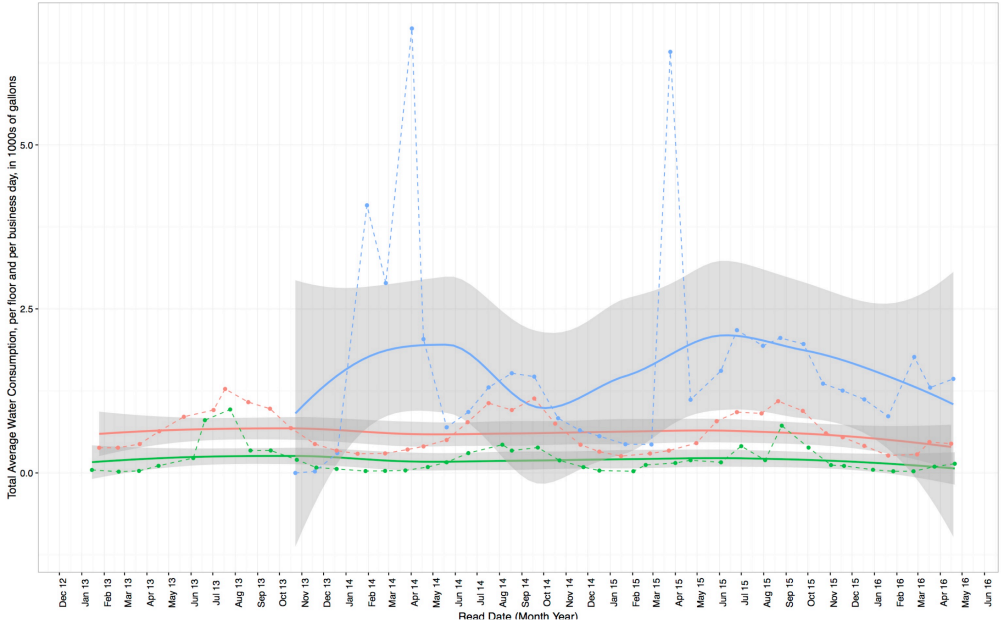
- Building water usage has many components.
- **Assumptions**: office only, no laundry; no landscaping; constant cooling
- **Estimation Steps:**
  - 1. Model water usage as a function of cooling degree days.
  - 2. Using the model to calculate the min/max water usage due to cooling and subtract from total water usage.

Problem #2: Estimating Employment

Proof of Concept: Building Water Usage

- 3. Using water usage tables which estimate the water usage per male/female employee and visitor, estimate the mean, min., and max. occupancy.

Example:
QCEW=135,
Mean Water
Estimate=124 (30,
218).



Problem #2: Estimating Employment

Next Steps

- Refine the water usage model to account for additional sources of variability and extend to different commercial building types.
  - Evaluate other proxies for building occupancy such as cell phone usage.
  - Build a data set with the building as the experimental unit and for each building collect three variables, QCEW employment number, water usage, and date.

Challenges

- Data, like departments, can operate in silos.
- Financially related datasets were set up for a specific purpose. It takes staff and political capital to use it for a secondary purpose.
- Finding the cross-connection
- Inconsistent address formatting.
- Cleaning data required staff time.
- Level of GIS skill varies, many databases not set up for GIS.

What We've Learned

- GIS is the key platform for linking data across departments.
- Most of the work is data cleaning.
- Start with the small victories:
  - Address formatting
- Build relationships
  - Address privacy concerns
  - Find automated solutions
  - Have a data sharing agreement.
- You won't know it works until you try
- Testing the limits of public information

Next Steps

- Review and inventory county data sources.
  - Identify potential uses, cross connections, and relationships between users.
- Connect data analysts (end users) with data input staff.
- Figure out who's in charge.
- County Data Steering Committee to engage in this kind of data driven approach to policymaking and find sample projects for VT SDAL Community Learning Initiative.

Thank you!

Bridging Data Silos in Local Government

Examples from Arlington County, Virginia

September 30, 2016

Andrew D'huyvetter AICP, Associate Planner
adhuyvetter@arlingtonva.us

ARLINGTON
VIRGINIA