

Population Model

Can Utility Data predict Neighborhood Instability?

January 19, 2018

Team Members:

Dianna Phillips, Craig Schafer, Mike Dougherty, Michael Radoff, Jackie Stice, Loretta Mills, Tonya Cromwell, Addison Spradlin, John Paganelli, Kevin Gustavson, Trecia Remmert, Breck Buchanan

Tulsa Data Scientist:

Steve Green



CITY OF
Tulsa
A New Kind of Energy.®

Introduction

- Start as Population Growth in the 1st Cohort
- Data Rich – So Many possibilities
 - Blight Team used data
 - First Turn On of Water
 - Consumption
 - Status – Length of Time
 - Neighborhood Stability Rate?
 - Status – Changes
 - Can the increase number of Status Changes (Turnover) predict Neighborhood Instability?
- Tulsa Data Scientists



Questions the team started with:

- What is neighborhood stability?
- What measure(s) can we use?
- What do we expect to see in Tulsa?
- What factors affect stability?
- Are these a positive or a negative effect?
- Do we have data sets we can use for these factors?
- If we don't, who does?



Factors Affecting Neighborhood Stability

A little bit of research led us to the following factors that may have links to utility data:

- Neighborhood upkeep (property abandonment, foreclosure, homeownership rate, property investment by owners)
- Location of neighborhood
- Characteristics of housing (type of housing, value of the homes, etc.)
- Crime

(reference: <http://www.communityprogress.net/increasing-neighborhood-stability-pages-240.php>)



What do we expect from these factors?

The relation of the factors should have a positive or negative affect on neighborhood stability and from our research and discussion this is what we expect to see:

- Discourage neighborhood stability- property abandonment, foreclosure, crime
- Promote neighborhood stability- high home ownership rate, low crime rate, possibly type of housing
- May either promote or discourage- location



Where to get the data?

The following sources were scoped for data:

- Tulsa County
- Census Data
- INCOG
- GIS
- Grants by WIN Department
- Housing and Development Department



Steps to get share data with Tulsa Data Scientists (TDS)

- Who needs to approve – Troy, Kier, Legal
- Legal reviewed sample documents to combine and adjust to one form
- TDS Reviewed and Signed
- RFA for Mayor to sign
- Start to Finish it took 2 months with constant effort.
- Now any City data group can use this Universal Data Authorization form.



Steps to get share data with Tulsa Data Scientists (TDS)

RECEIPT AND RELEASE

THE UNDERSIGNED recipient does hereby acknowledge receipt from the City of Tulsa Engineering Services Department of the DATA listed as follows:

(Example of data requested: Base, Water, Sewer, Storm, Pavement, Aerials, DTM Data, and 3D Contours for Atlas Page 133)

The undersigned acknowledges that the DATA is for information only and is not intended as an accurate representation of Engineering Services Department installations, and does hereby release the Engineering Services Department from any and all claims by the undersigned arising from its use of the DATA.

The undersigned further understands and acknowledges that its use of the DATA is without warranty or representation by the Engineering Services Department of its accuracy. The undersigned recognizes that the DATA was produced at a scale and positional accuracy suitable for planning purposes only, that all installations shown on the DATA must be confirmed by the undersigned in any work that it is doing for itself and/or clients, that additions and deletions are being made regularly, and that no update service is being furnished.

The undersigned further acknowledges and warrants that it will not reproduce any part or all of said DATA other than for its own internal use and that it will not lend, sell, or assign any or all of said DATA to a third



Utility Data

- Carried over from previous Cohort
- Broke down Property Actions by year to fit data model
- For size requirements and data accuracy limited time frame to 2000-present
- Each property takes up 54 rows, totaling a dataset of 8 million rows.
- Joining this amount of data crushed Access/Excel, so we had to switch to [and learn] ACL.



Housing Mortgage Disclosure Act (HMDA)

- The Home Mortgage Disclosure Act (HMDA) data provided information pertaining to mortgage bank loans originating from home purchases, home renovations, and refinancing.
- This data represents where investments or reinvestments are happening in residential areas by home owners.
- Data spans from 2000 to 2016



Geocoding Utility Data

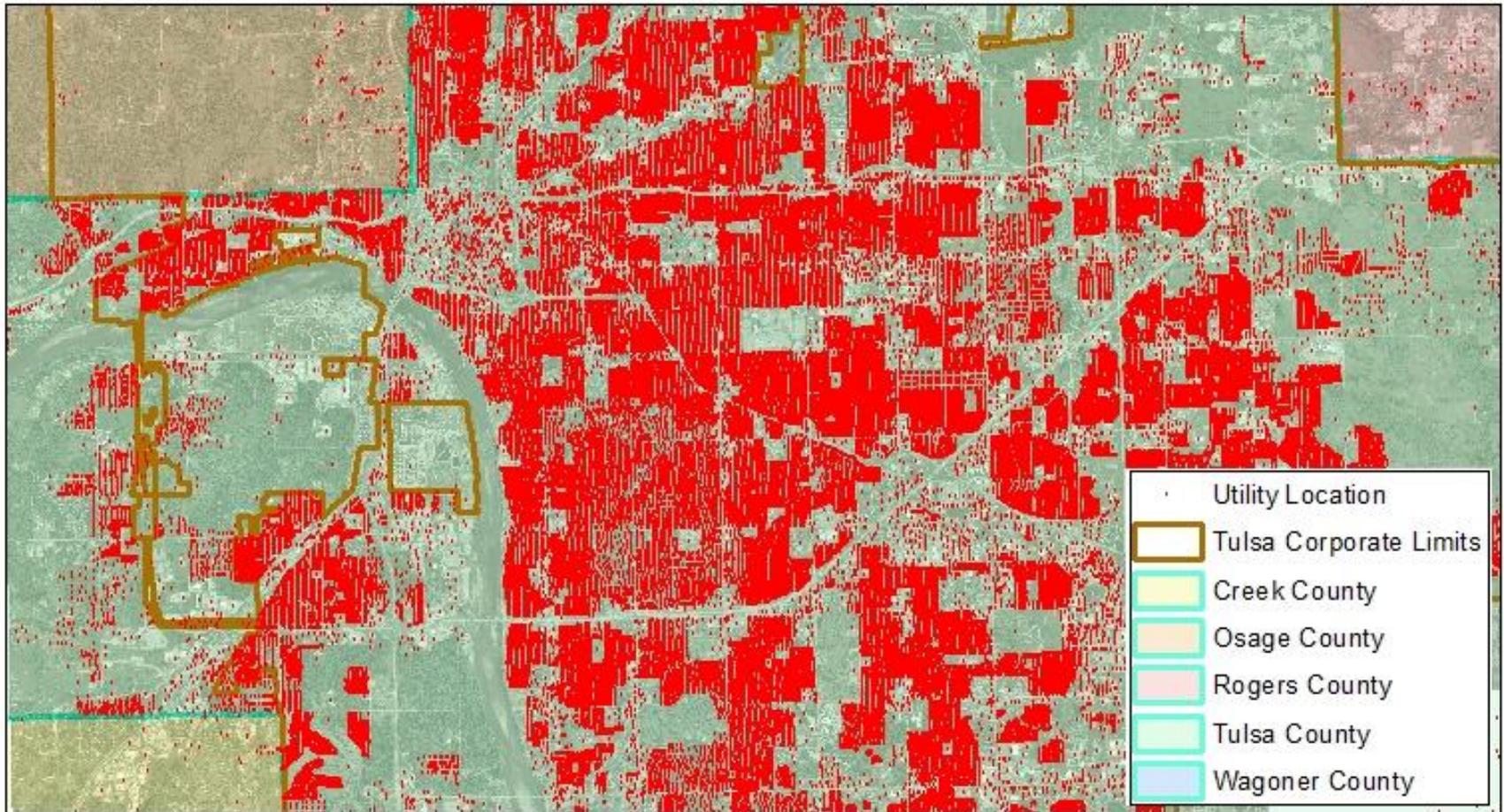
- Addresses from the utility data were used to generate location points associated to Prop_Ref values

Prop_Ref	Prop_Type	.Units	Address	Zip	In-OutCity
008/0905/0681	SF	<Null>	5411 E 121ST PL	74008	IN
008/0905/0682	SF	<Null>	5423 E 121ST PL	740082737	IN
008/0905/0684	IND	<Null>	12236 S SHERIDAN RD	74008	IN
008/0905/0687	SF	<Null>	11814 S SHERIDAN RD	740082035	IN
008/0905/0688	SF	<Null>	11800 S SHERIDAN RD	740082035	IN

- 150,696 total records
 - 133,379 Matched
 - 2,615 Tied
 - 14,702 Unmatched
- 88% of records were located



Geocoding Utility Data (cont.)



Geocoding Utility Data (cont.)



Joining Census Data to Utility Data

- Geocoded Prop_Ref point values were spatially joined to US Census identifiers

Prop_Ref	Prop_Type	.Units	Address	Zip	In-OutCity	Block	BlockGroup	Tract
008/0905/0680	COM	<Null>	109 N ARMSTRONG ST	740084449	OT	401430078011023	401430078011	40143007801
008/0905/0681	SF	<Null>	5411 E 121ST PL	74008	IN	401430076381035	401430076381	40143007638
008/0905/0682	SF	<Null>	5423 E 121ST PL	740082737	IN	401430076381035	401430076381	40143007638
008/0905/0684	IND	<Null>	12236 S SHERIDAN RD	74008	IN	401430076381036	401430076381	40143007638
008/0905/0687	SF	<Null>	11814 S SHERIDAN RD	740082035	IN	401430076372024	401430076372	40143007637
008/0905/0688	SF	<Null>	11800 S SHERIDAN RD	740082035	IN	401430076372026	401430076372	40143007637
008/0905/0689	SF	<Null>	11720 S SHERIDAN RD	740082049	IN	401430076372026	401430076372	40143007637
008/0905/0690	SF	<Null>	11840 S SHERIDAN RD	740082035	IN	401430076372024	401430076372	40143007637
008/0905/0691	SF	<Null>	11818 S SHERIDAN RD	740082035	IN	401430076372026	401430076372	40143007637
008/0905/0692	SF	<Null>	5404 E 121ST PL	74008	IN	401430076381035	401430076381	40143007638
008/0905/0693	SF	<Null>	5311 E 122ND PL	74008	IN	401430076381035	401430076381	40143007638

- Each record was associated to a Block, Block Group, and Tract
 - 7,458 Blocks
 - 409 Block Groups
 - 180 Tracts



Joining Parcel Data to Utility Data

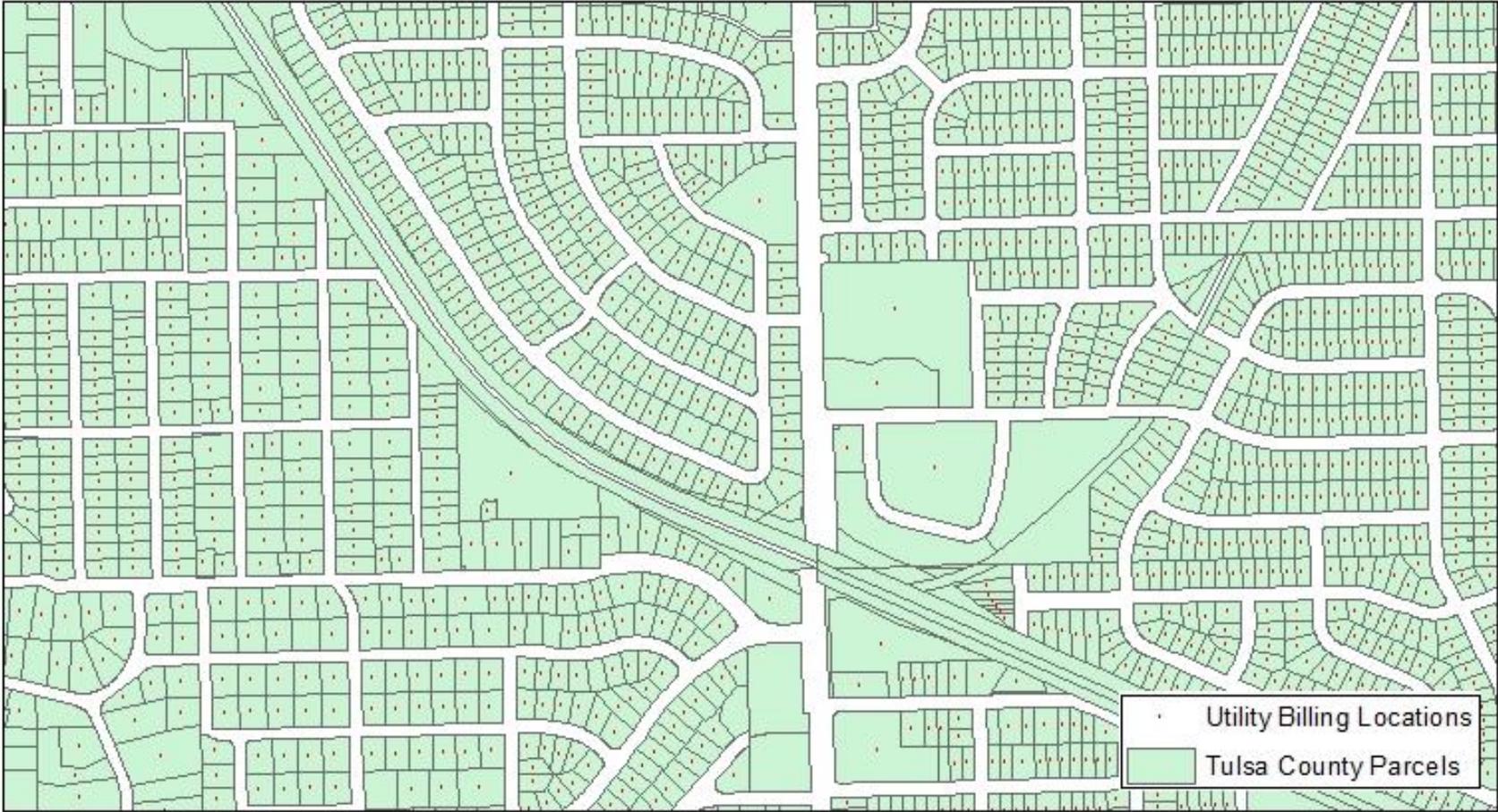
- Geocoded Prop_Ref point values were spatially joined to 2017 Tulsa County Parcels

Prop_Ref	Prop_Type	.Units	Address	Zip	In-OutCity	ACCOUNTNO
130//2710/3865	COM	<Null>	111 N MAIN ST C	74103	IN	R00500920129760
103//2710/3630	COM	<Null>	110 N BOSTON AVE	741032020	IN	R00500920129720
103//2710/3691	COM	<Null>	114 N BOSTON AVE	741032020	IN	R00500920129700
103//2710/3620	COM	<Null>	124 N BOSTON AVE	741032020	IN	R00500920129690
103//2710/3870	COM	<Null>	17 E MATHEW BRADY ST	741032015	IN	R00500920127740
103//2710/4290	IND	<Null>	220 N BOSTON AVE	741032018	IN	R00500920127660

- Each record was associated to 2017 Tulsa County Parcel Account Number
- 2017 Parcels had right-of-way and other non-applicable types removed



Joining Parcel Data to Utility Data (cont.)



Joining Parcel Data to Utility Data (cont.)

- Previous year parcels were joined using the Tulsa County AccountNo value
- Future Improvements
 - Refine removal of non-applicable parcels
 - Explore the benefits of spatially joining all prior years of Tulsa County Parcels instead of joining via the 2017 AccountNo value
 - That process risks associating values from one large parcel to many smaller parcels



Learned this Cohort

- Learning to work with large Data Sets
 - ACL, SQL, etc.
 - Why do this twice – City could create a way to maintain a database or data warehouse so data can be saved and used again.
- Working with outside organizations
 - Transferring Data Files
 - Data Dictionary Essential
 - Legal Agreement
 - Universal so any group can use it if sharing data
- Data Rich – Still could look at Consumption
 - University of Chicago submission



Questions

