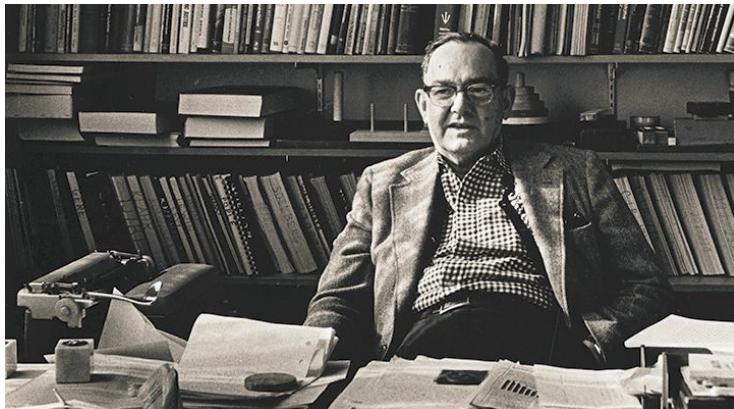


# Comprehensive Planning & Data Science



“All behavior involves conscious or unconscious selection of particular actions out of all those which are physically possible to the actor and to those persons over whom he exercises influence and authority.”  
— Herbert A. Simon, [Administrative Behavior, 4th Edition: A Study of Decision-making Processes in Administrative Organisations](#)

# Agenda

1. Homework questions, Review
2. What is comprehensive planning?
3. How comprehensive planners fail with forecasting
4. From comprehensive to strategic planning
5. Case Studies:
  - A. Bike share planning in Philadelphia
  - B. Lancaster Urban Growth Area

# Homework Questions?

# Review

1. What makes a good indicator?
2. Spatial process is important - what was the spatial process associated with TOD in Philadelphia?
3. Working with census data, longitudinal data

# **What is comprehensive planning?**

# Comprehensive Planning

1. Sometimes required by law
2. Goal-driven
3. Assess current conditions
4. Large geography
5. Lots of public input
6. Comprehensive across domains
7. **Forecast** 10 - 30 years out
8. Data-driven
9. Scenario-driven (ie. evaluating alternatives)
10. Monitor the plan

# PHILADELPHIA 2035

THE PLAN | GET INVOLVED | NEWS



Credit: Photo by Photo by Blended Angels for Visit Philadelphia

# PHILADELPHIA 2035

The Vision for Philadelphia's Growth and Development

<https://www.phila2035.org/>

# **How comprehensive planners fail with forecasting..**

## **An exercise**

# Comprehensive Planning & Forecasting

“Let’s plan for \_\_\_\_\_ in 20 years...”

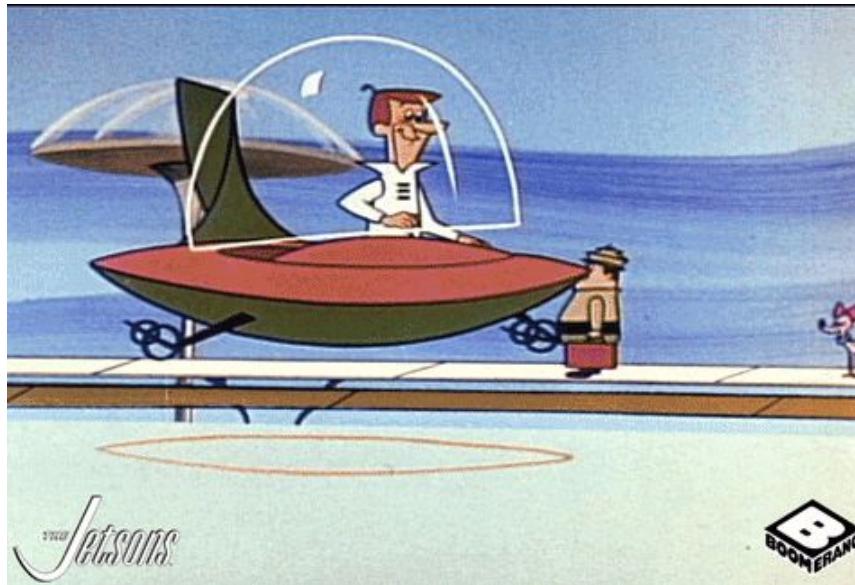
# Comprehensive Planning & Forecasting

## Transportation

What trends today could we never have forecasted in 2001?

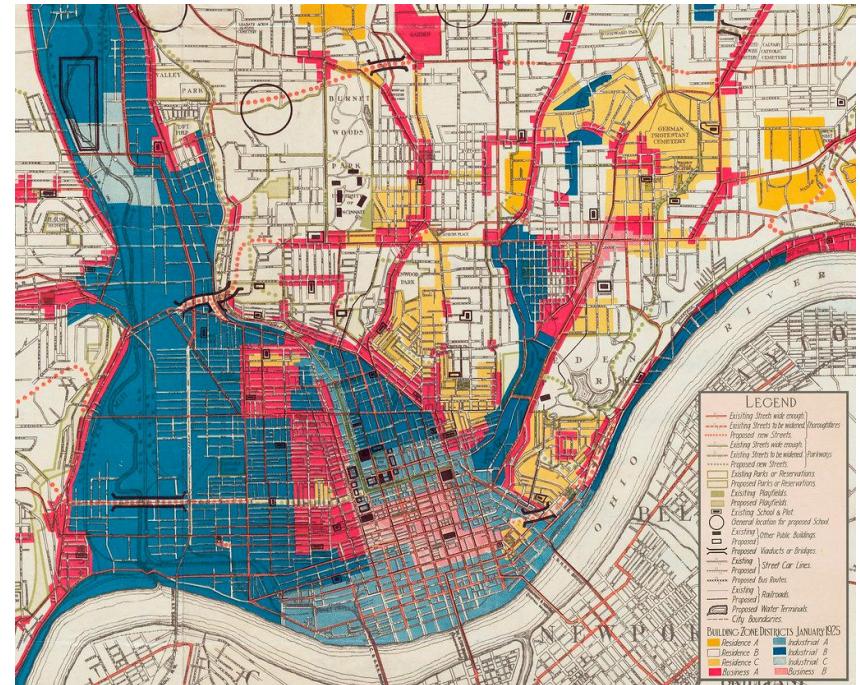
# Comprehensive Planning & Forecasting

What **transportation** trends do you think are going to be important in 2041?



Me: You can't plan for an uncertain future, when future changes will be transformational

Planners: *Hold my beer*



### **Method of Calculating Growth of Population**

The population of a city as at present distributed, and as it will grow and be distributed in the future, is obviously the basis upon which scientific city planning must be formulated. Therefore, careful studies were made of the probable total future population of Cincinnati as a first essential of the Plan.

# POPULATION CURVES

OF THE  
CITY OF CINCINNATI

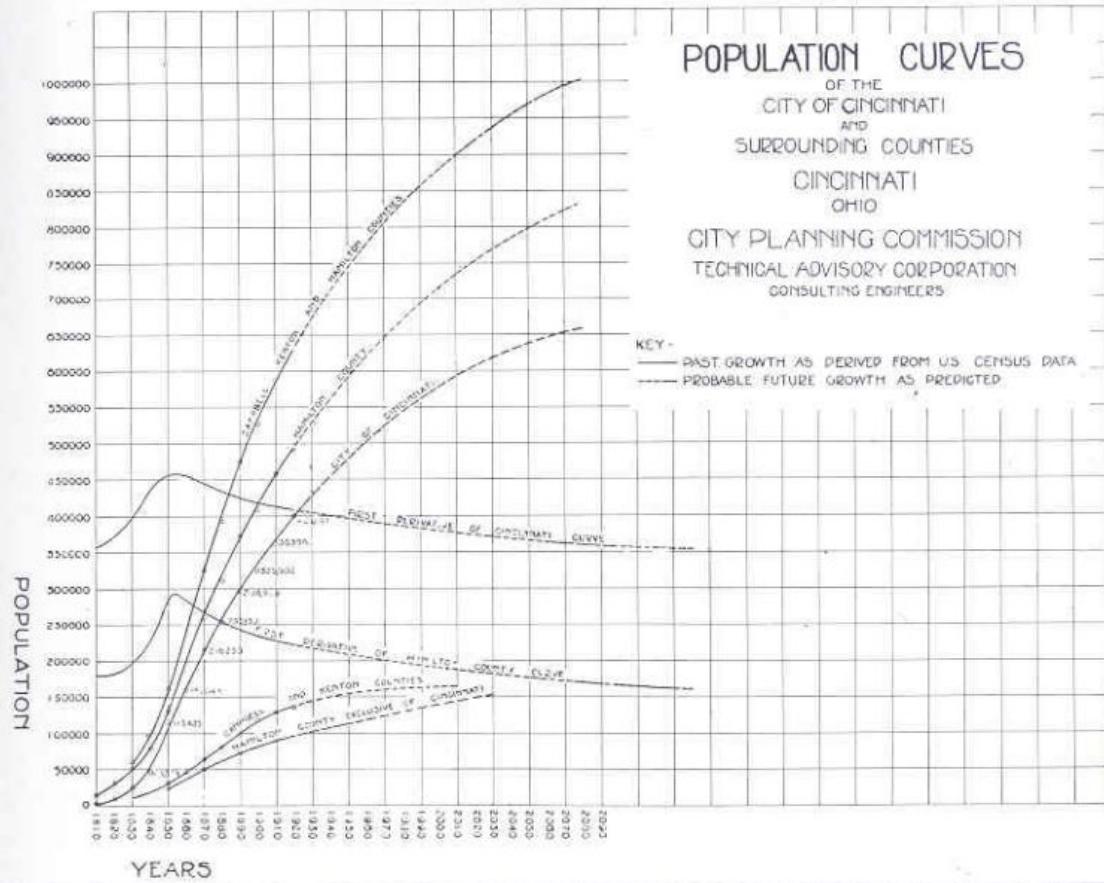
AND  
SURROUNDING COUNTIES

CINCINNATI  
OHIO

CITY PLANNING COMMISSION  
TECHNICAL ADVISORY CORPORATION  
CONSULTING ENGINEERS

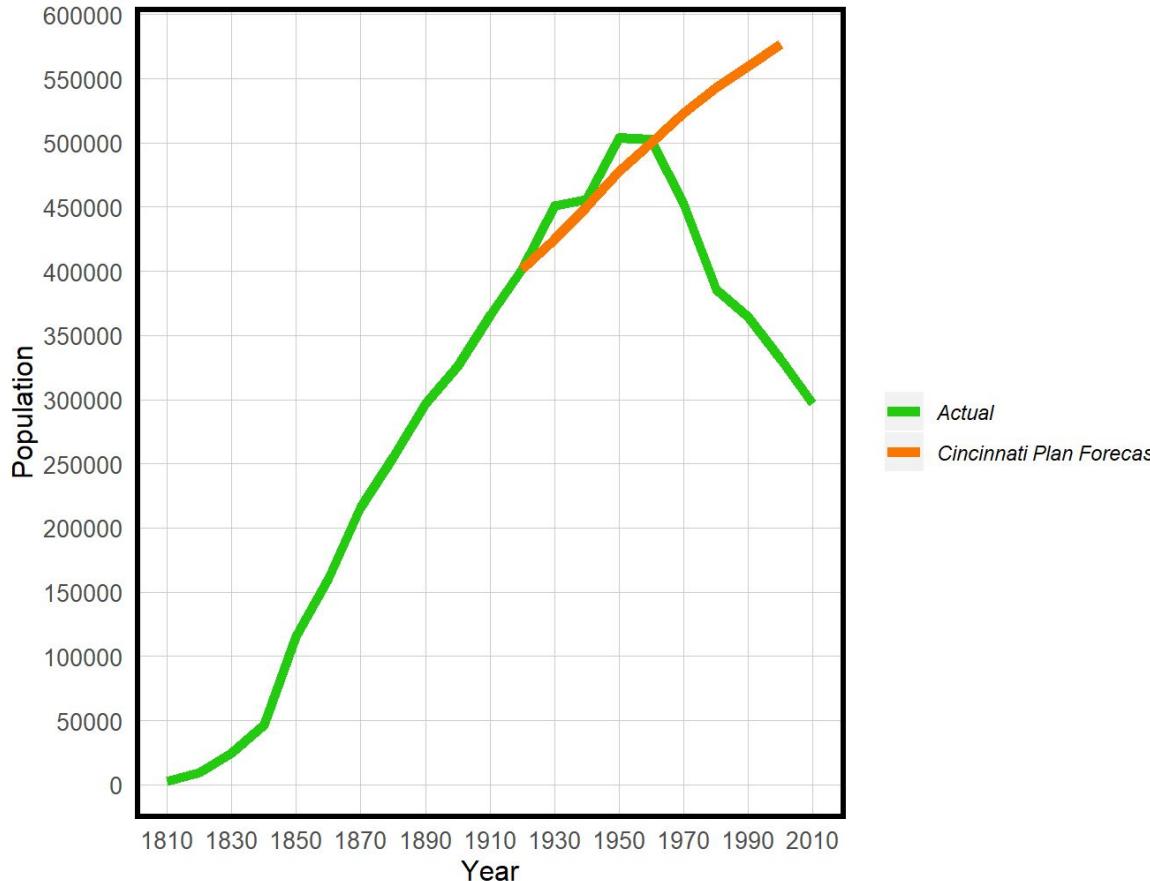
KEY -

- PAST GROWTH AS DERIVED FROM US CENSUS DATA
- PROBABLE FUTURE GROWTH AS PREDICTED

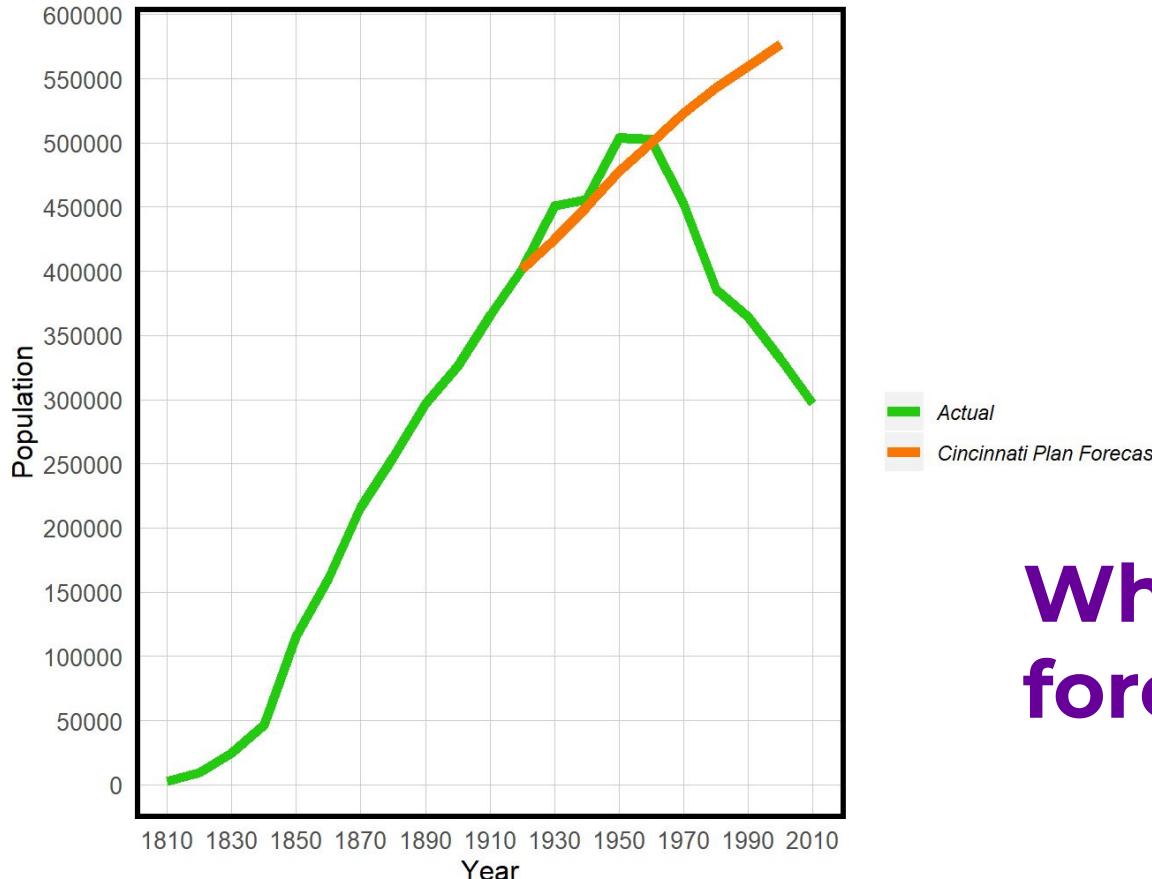


The growth of this city is slowing down.

## 1925 Plan for Cincinnati: Population projections



1925 Plan for Cincinnati:  
Population projections



Why did the  
forecast fail?

**The forecast failed because there is no reason to expect that ‘current conditions’ will generalize to the future.**

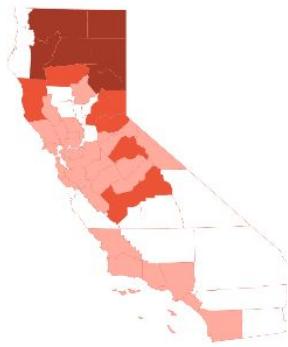
What 'systems' would we need to understand in order to reasonably forecast population into the future?

# HOW CLIMATE MIGRATION WILL RESHAPE AMERICA

Millions will be displaced. Where will they go?



By 2070, some 28 million people across the country could face Manhattan-size megafires. In Northern California, they could become an annual event.



A megafire on average every ...

■	1-2 years
■	2-5 years
■	5+ years
■	No data

\*High Emissions scenario

Sea-level rise could displace as many as 13 million coastal residents by 2060, including 290,000 people in North Carolina.

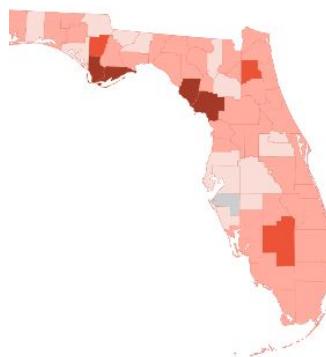


Percent of properties below high tide ...

■	5-25%
■	1-5%
■	Under 1%
■	0

\*High Emissions scenario

By 2060 in Florida and elsewhere, the costs of sea-level rise and hurricanes will be compounded by knock-on economic challenges, from growing crime to falling productivity.

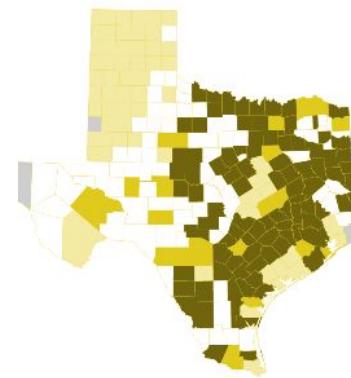


Economic damages as a proportion of G.D.P. ...

■	10%-55%
■	5-10%
■	1-5%
■	0-1%
■	Economic benefits

\*High Emissions scenario

Corn and soy production will decrease with every degree of warming. By 2060, parts of Texas may experience a drop in yields of more than 92 percent.



Crop yield decline by:

■	60-92%
■	30-60%
■	0-30%
■	Yield increases
■	No data

\*High Emissions scenario

# Manage Financial Risk Under Climate Change

risQ leverages economic and physical sciences to drive climate adaptation

[LEARN MORE →](#)

January 12, 2020

New Tool For Measuring Climate Risk In Municipal Bonds Emerges From Stealth

January 12, 2020

Intercontinental Exchange Partnership: What It Means For The Municipal Market

April 18, 2019

Review: BlackRock+Rhodium Climate Risk Analysis

March 20, 2019

Quantifying Wildfire Risk To Municipal Debt In California

## **Comprehensive Planning vs Strategic Planning**

## **Comprehensive Planning**

1. Sometimes required by law
2. Goal-driven
3. Assess current conditions
4. Large geography
5. Lots of public input
6. Comprehensive across domains
7. Forecast 10 - 30 years out
8. Data-driven
9. Scenario-driven (ie. evaluating alternatives)
10. Monitor the plan

## **Strategic Planning**

1. Driven by internal stakeholders
2. Goal-driven
3. Assess current decision-making process
4. Small or large geography
5. May involve no public input
6. Specific to a single domain
7. Near term (or time independent) is better
8. Data-driven
9. Scenario-driven (ie. evaluating alternatives)
10. Monitor the plan with high resolution time/space indicators.



SERVICES ▾

PROGRAMS &amp; INITIATIVES

NEWS &amp; EVENTS

PUBLICATIONS &amp; FORMS

**COVID-19:** Learn how to stay safe, get vaccinated, and find testing. Visit [phila.gov/covid](#).**Election day is Nov. 8, 2022**

Check your voter status, apply for a mail-in ballot, and more.

[MAKE A PLAN TO VOTE](#)

# Philly Tree Plan

Creating a ten-year plan for the growth and care  
of Philadelphia's urban forest.

PHILADELPHIA  
PARKS & RECREATION[Home](#)[Vision](#)[Our team](#)[Timeline](#)[Resources](#)

# **Forecasting - use the right tool for the job**

<https://palakagr.github.io/MUSAPracticum/#>

# Use case 1: Predicting bike share demand in Philadelphia

IndeGo bike share system - what do they need to know to launch their network?

Goals - create an efficient and cost-effective transit alternative, enhance transit equity... what else?

Strategy - put the stations and the bikes in the right places to optimize goal-related outcomes

# Use case 1: Predicting bike share demand in Philadelphia

In cities with bike share systems, are ridership patterns sufficiently general enough that they can be used to predict bike share usage in cities without such systems?

**Take data from  
These bike share  
Systems, and  
predict for...**

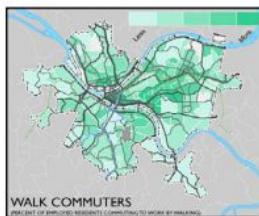
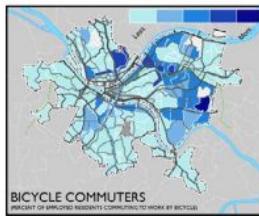
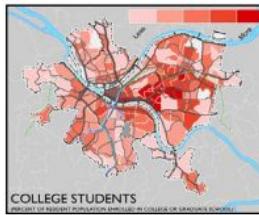


**Philadelphia**



**This was in 2013, nearly a year before the Indego system debuted.**

**What was the ‘business as usual’ approach to bike share planning?**



## LOCAL CONTEXT ANALYSIS

Alta uses GIS and census data to map the demographics of a community, such as household income, vehicle ownership, bicycle and pedestrian commuters, employment centers, parks, greenways, and key destinations. Individual maps are overlaid to form a composite heat map that shows where demand for bike share is likely to be highest.

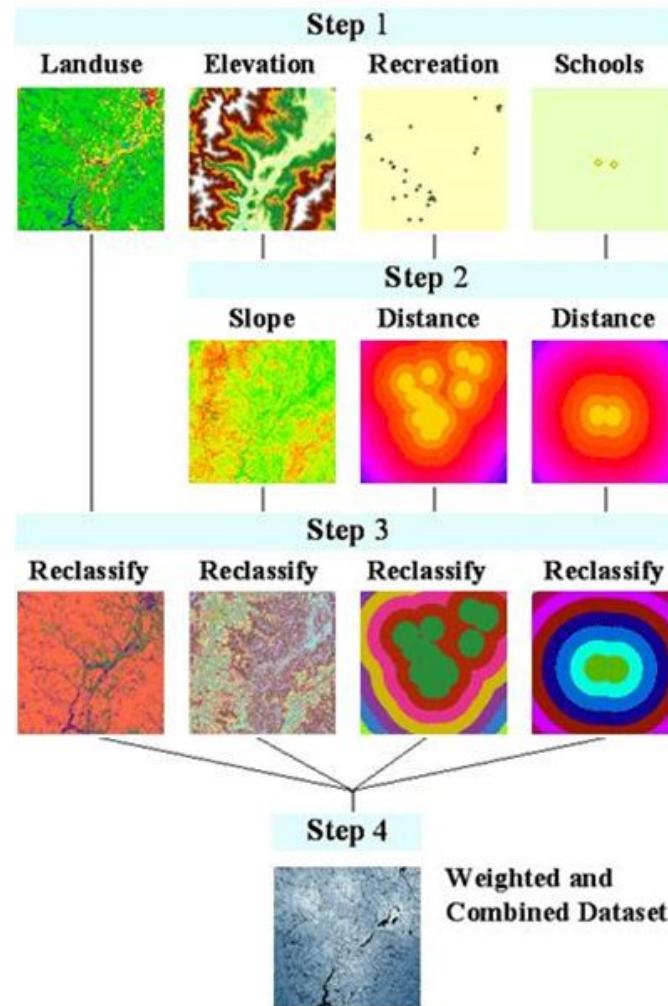
## SYSTEM SIZE, DEMAND, AND STATION LOCATIONS

Alta's heat maps show areas that are well-suited for bike share, and are used to make recommendations for system size and service area. Alta also uses GIS data to create analysis maps focused on bike share system equity.

## BUSINESS PLAN, EXPECTED COSTS AND REVENUE, AND OPERATING MODEL

Our business plans define expected system costs and develop diversified funding strategies for dock-based, smart-lock, and hybrid bike share systems. We explore all avenues of potential revenue, including user fees, sponsorship opportunities, and public and private funding. Our plans generate interest in bike share with decision makers, potential sponsors, and the general public.

GIS analysis maps for the Bike Share Study, Pittsburgh, PA



## **Weighted raster ‘composite’ approach:**

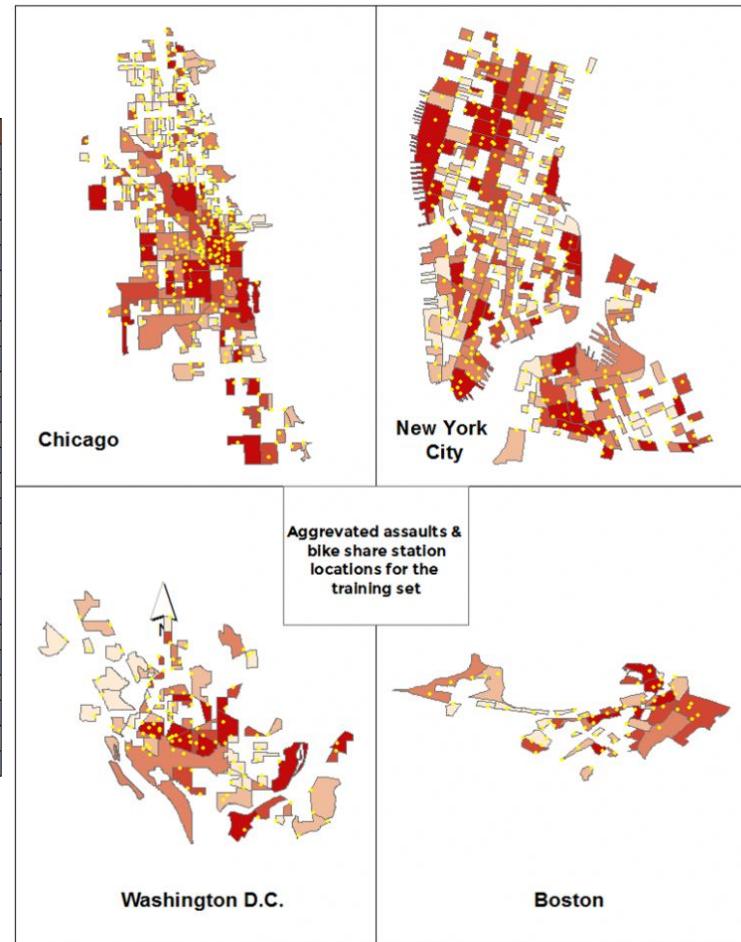
- **Pros:** Empowers experts with domain experience. Let the domain expert ‘set the weights’.
- **Cons:** Experts have been known to be wrong.

## **The machine learning approach**

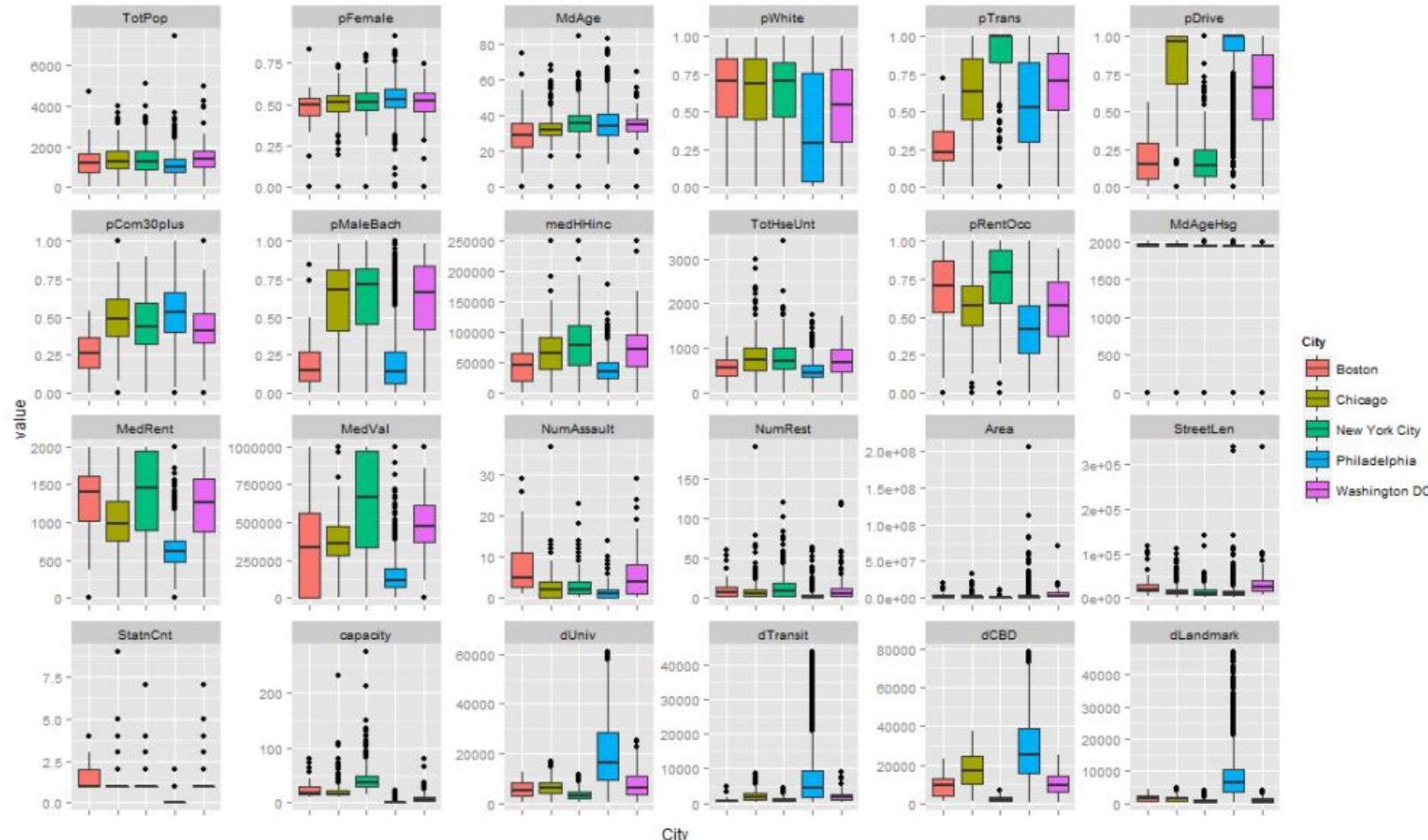
- **Pros:** Gather ridership from a sample of cities similar to the one being planned as well as other exogenous data. In effect, have the weights set by the data or the experience from the sample cities.
- **Cons:** What if there isn’t enough training data?

# Gather data (features) from the training set cities

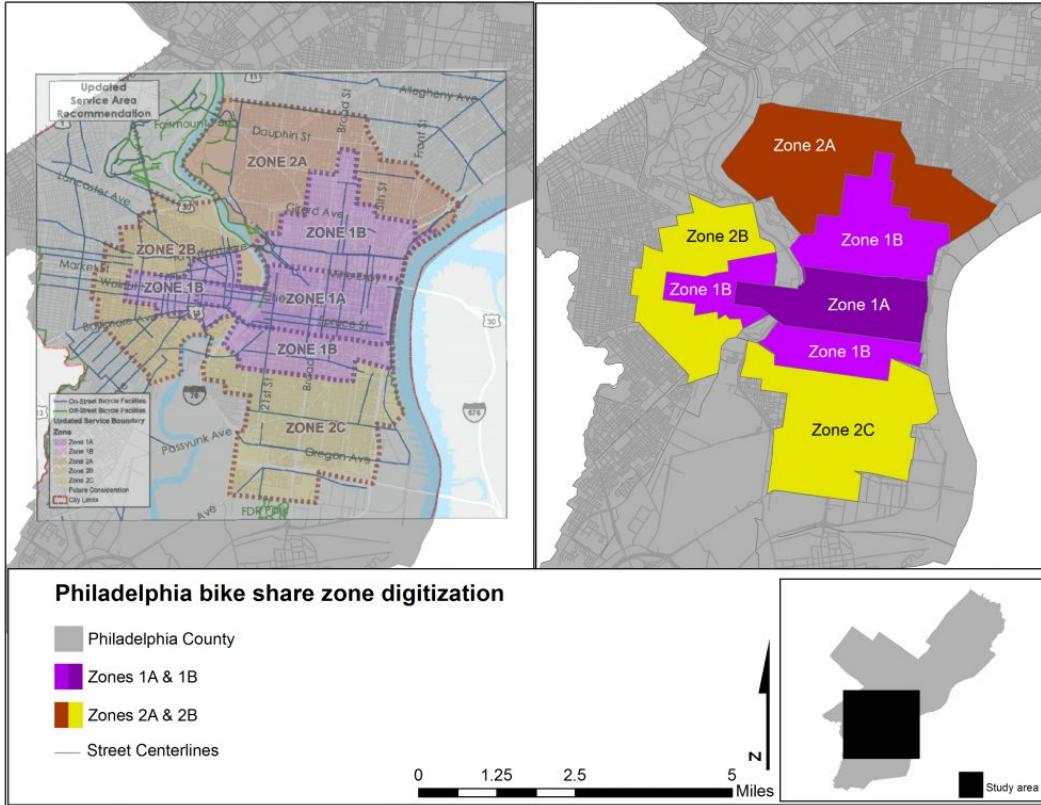
	Variable	Description	Alternative Source*
Demographic (ACS)	TotPop	Total population	
	pFemale	Percent female	
	MdAge	Median age	
	pWhite	Percent white	
	pTrans	Percent who take transit to work	
	pDrive	Percent who drive	
	pCom30plus	Percent who commute 30+ minutes	
	pMaleBach	Percent of males with a bachelor degree	
	medHHinc	Median household income	
	TotHseUnt	Total housing units	
	pRentOcc	Percent Renter Occupied	
	MdAgeHsg	Median Age of Housing	
	MedRent	Median Rent	
	MedValue	Median House Value	
Built Environment	dLandmark	Distance to nearest Landmark	Washington D.C.: Geocommons.com
	dTransit	Distance to nearest transit station	
	dCBD	Distance to CBD	New York City: John Weir, GitHub
	dUniv	Distance to nearest University w/ > 10k student enrollement	Nat. Center Educ. Stat.: Geocommons
	NumAssault	Number of assault crimes	New York City: Thomas Levine
	NumRest	Number of full service restaurants	ESRI Business Analyst Extension
	StreetLen	Total length of streets in feet	
	Area	Total area of block group in sq. feet	
Bike	capacity	Number of bikes/station	
	StatnCnt	Count of bike share stations	
	strtTrip	Number of bike share departures	



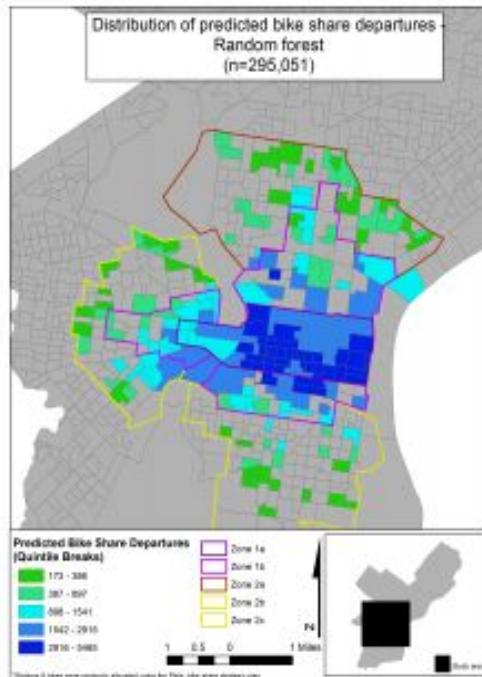
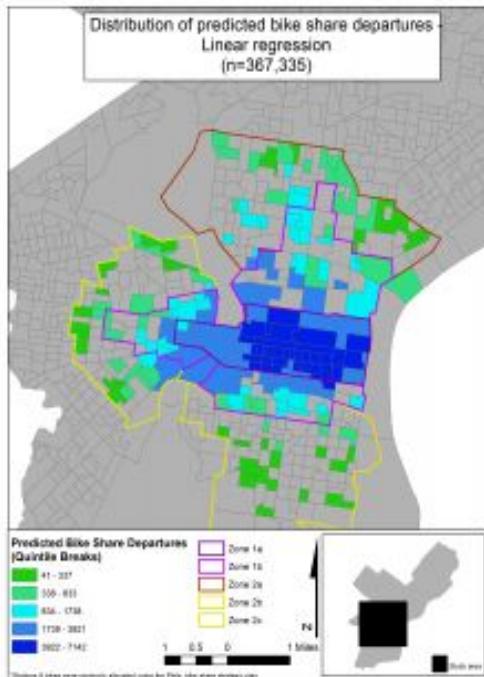
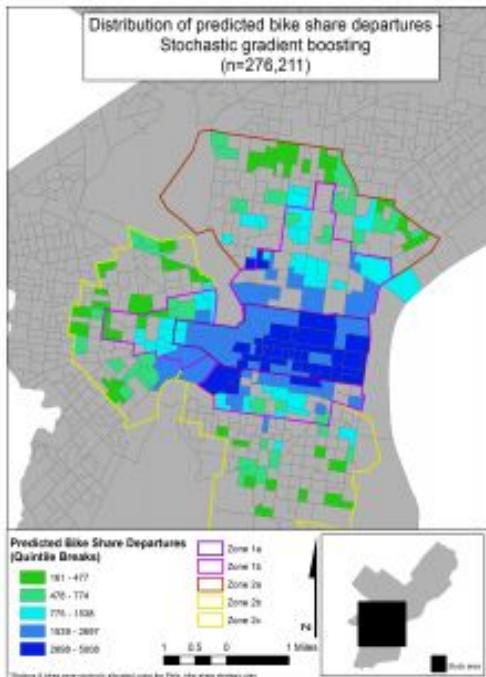
# Do the training set features generalize to Philly?



# How do we figure (supply-side) inputs to the model - namely the location of would-be stations?



Model type	OLS	Random Forest	Gradient Boosting
# of predicted trips	367,355	295,051	267,211



This model predicted 267,211 trips in the first 4 months of Philadelphia's bike share system.

This is **just 17% more** trips than actually occurred in the first four months of service.

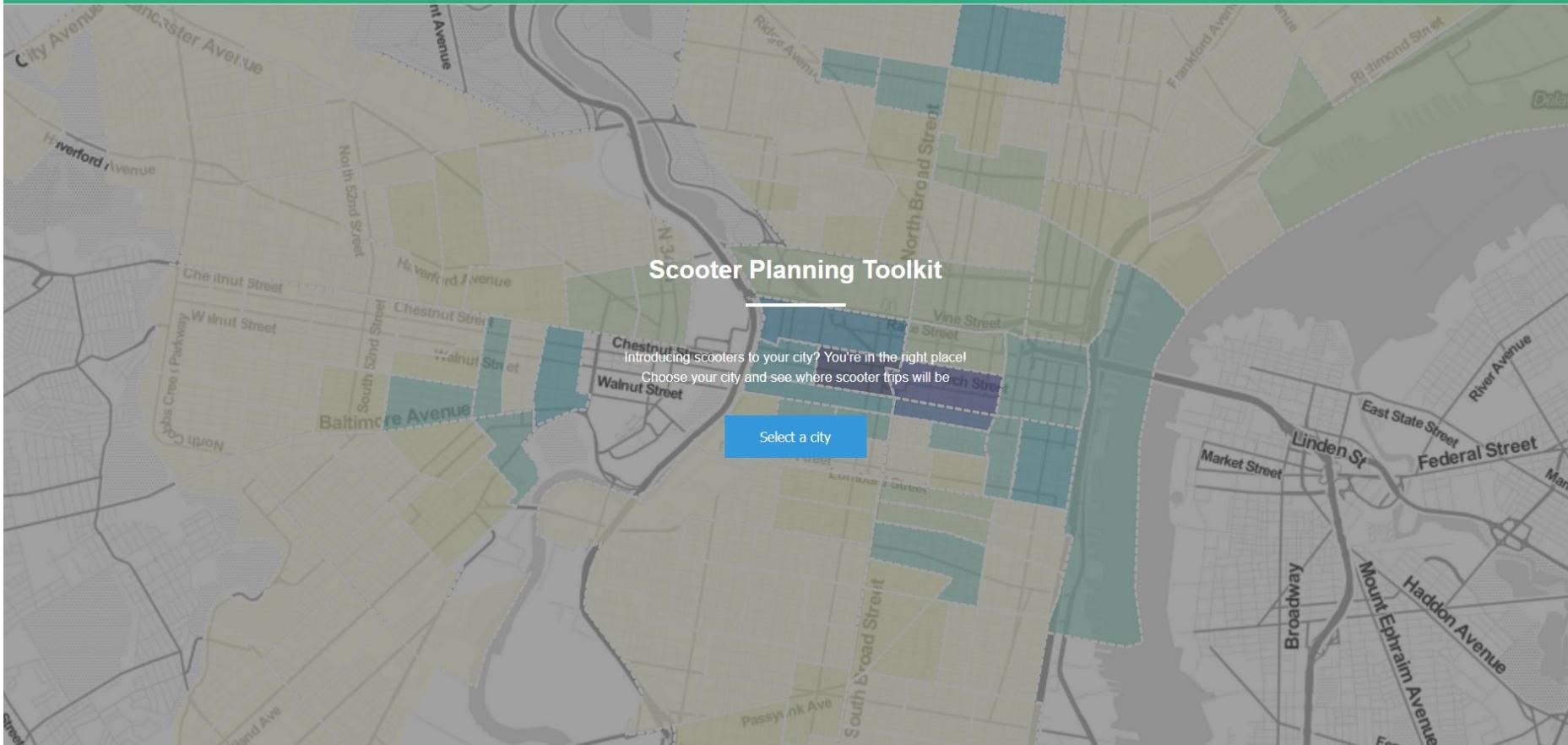
*Can we **democratize** the planning process by placing these empirical tools into the hands of smaller communities for free to help them plan their own systems*

## Scooter Planning Toolkit

Introducing scooters to your city? You're in the right place!

Choose your city and see where scooter trips will be

Select a city

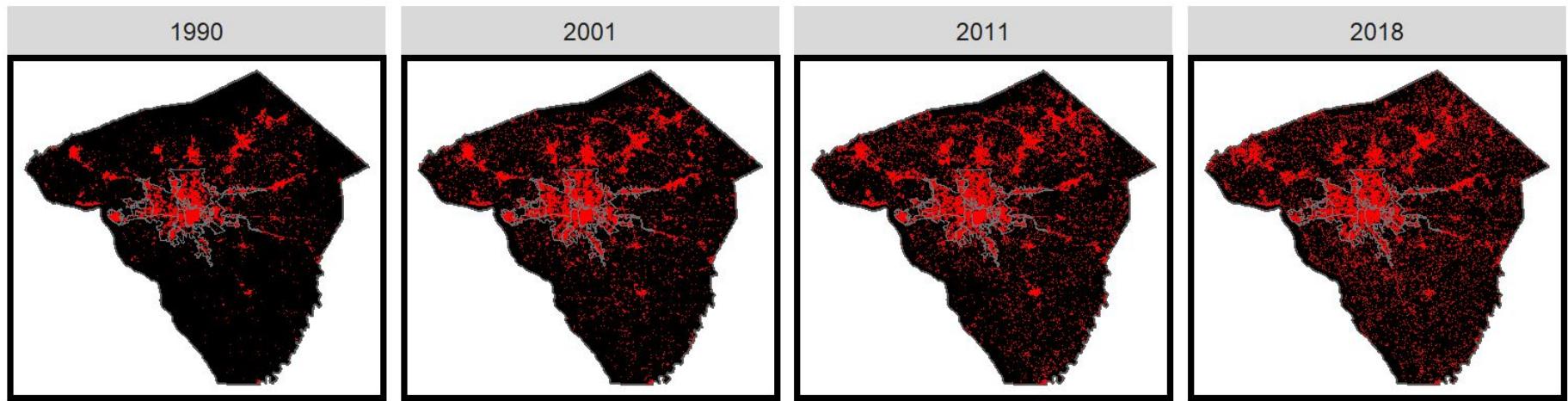


# **Growth Planning Use Case: Where should we extend the Urban Growth Area in Lancaster, PA?**



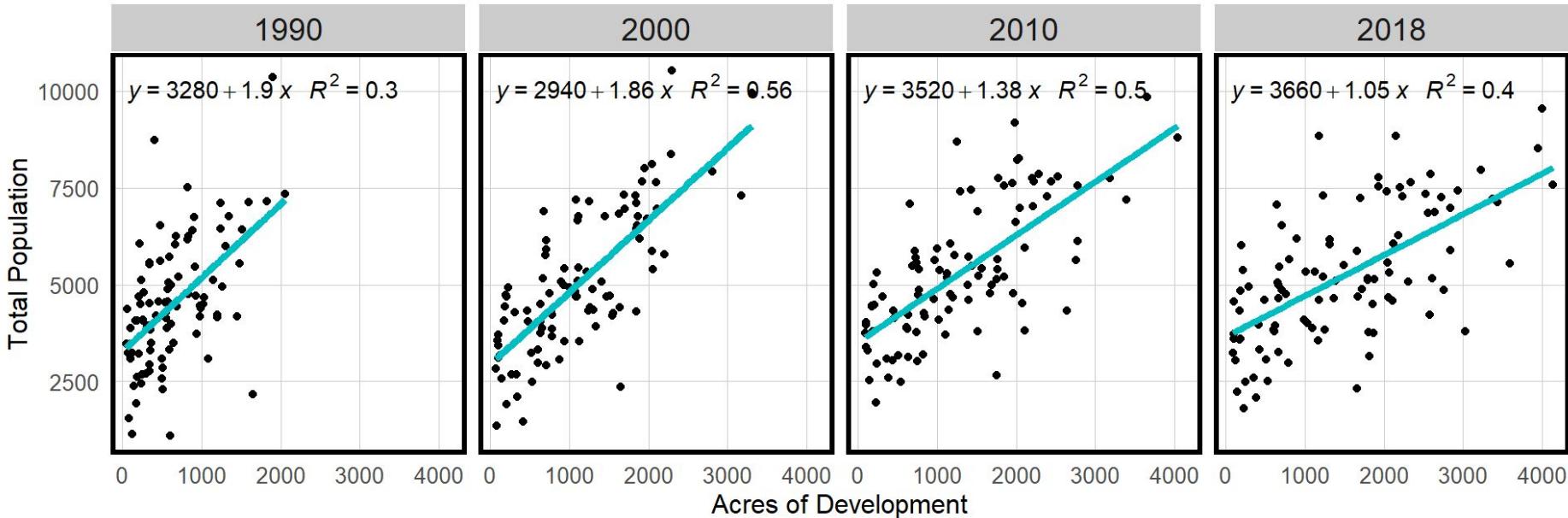
# Land Cover Change - Lancaster, PA - 1990 - 2018

*Author's estimations via supervised classification. Source: USGS*

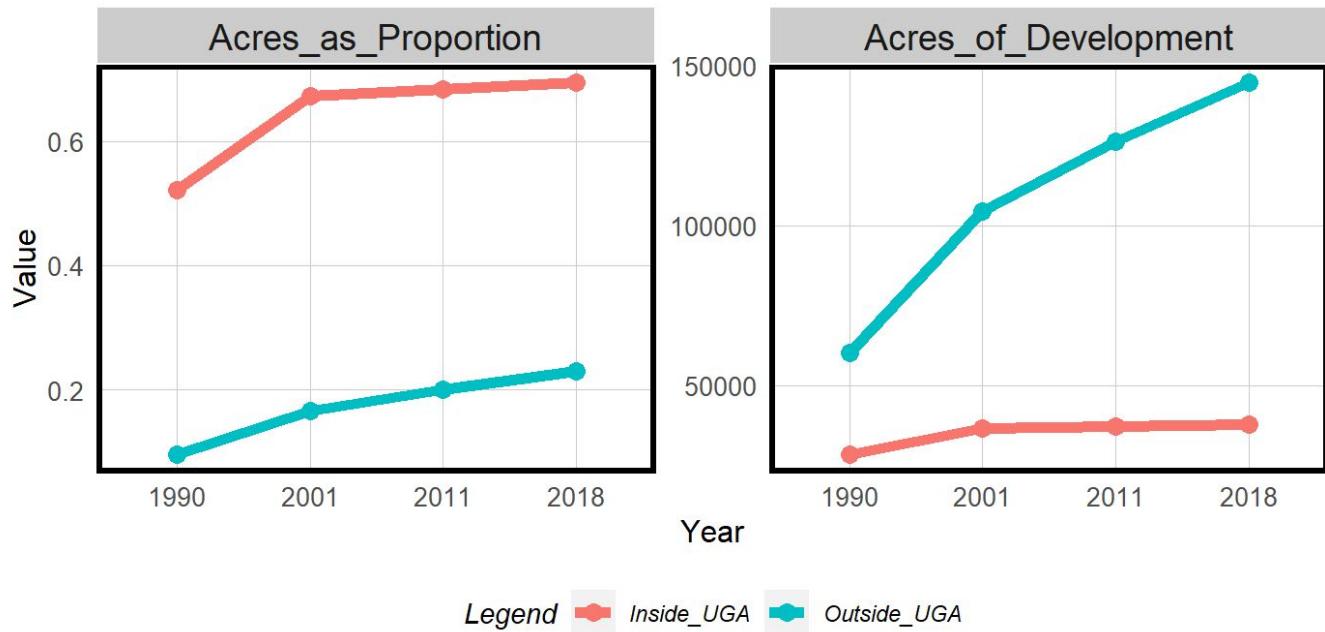
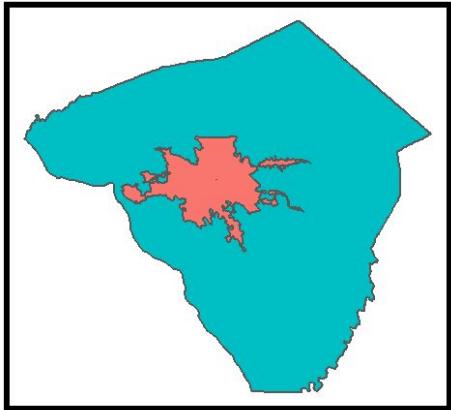


# Total Population as a function of Urban Development

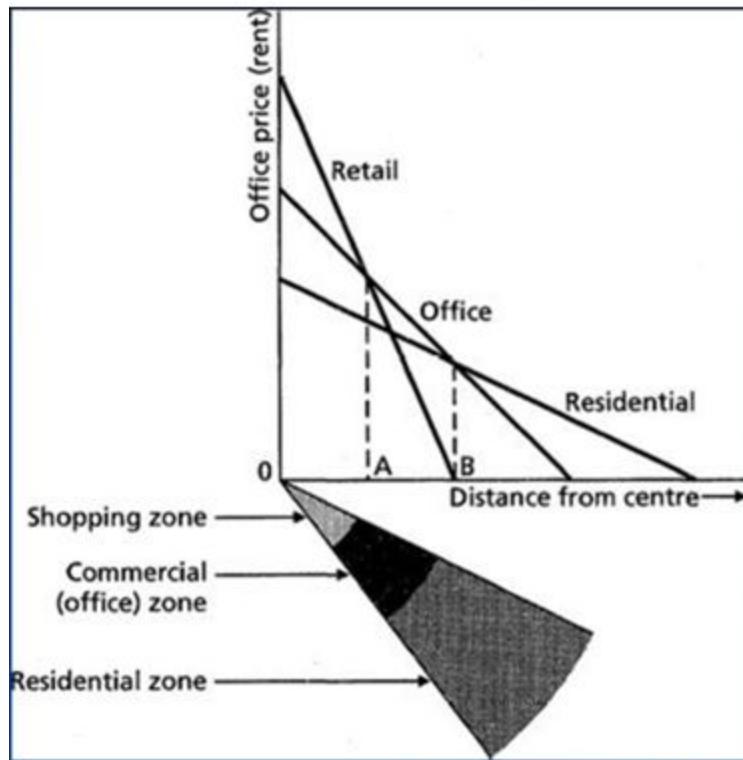
Lancaster, PA; Census Tracts; 1990 - 2018.



## Growth inside & outside of the UGA

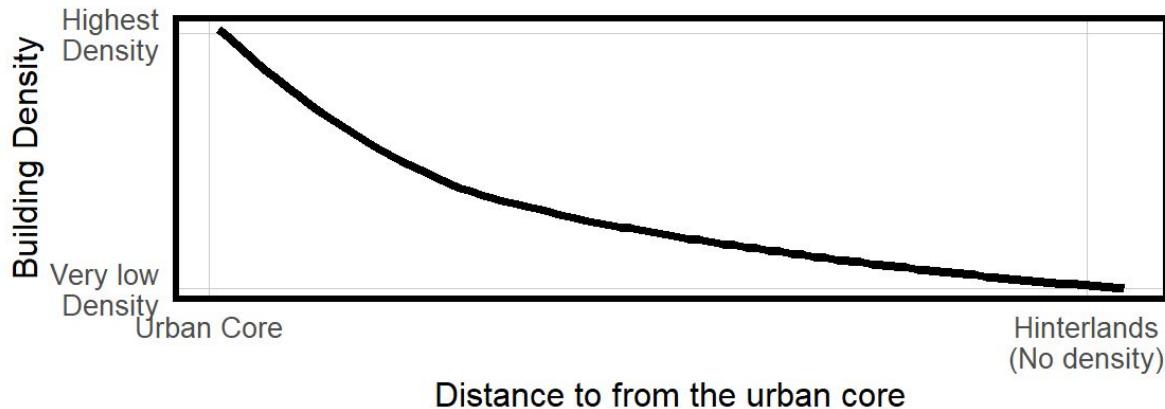


## Bid-Rent Theory

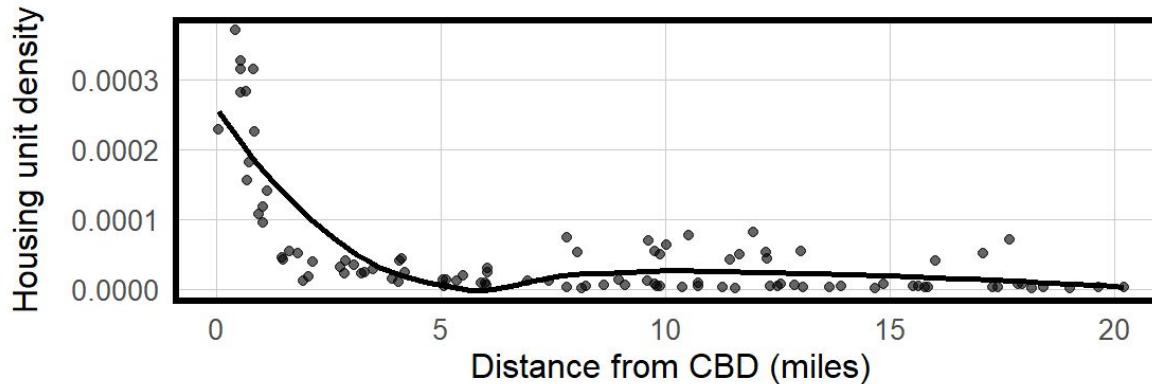


Von Thunen (1826), Alonso (1964)

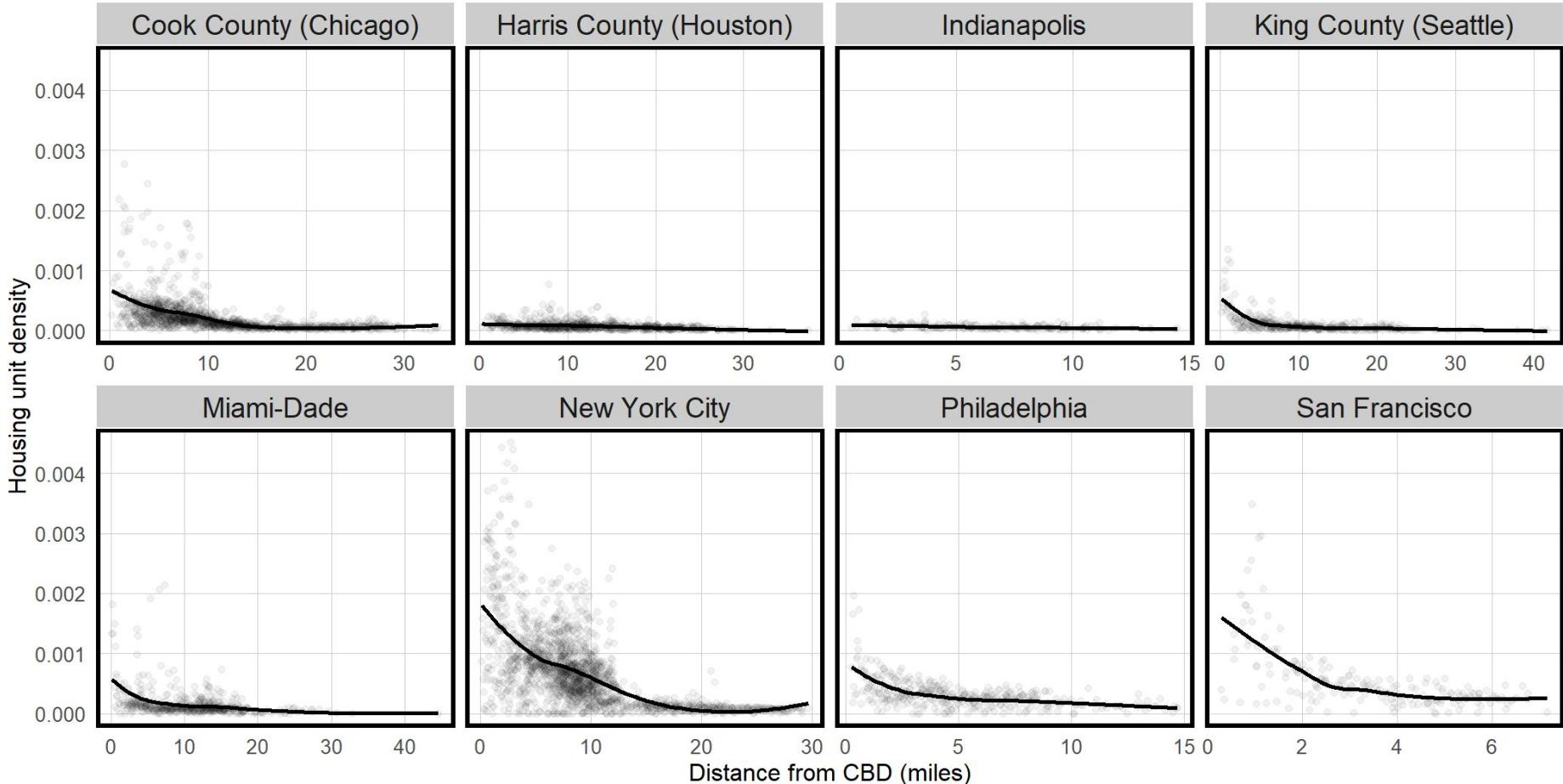
# Bid Rent Curve



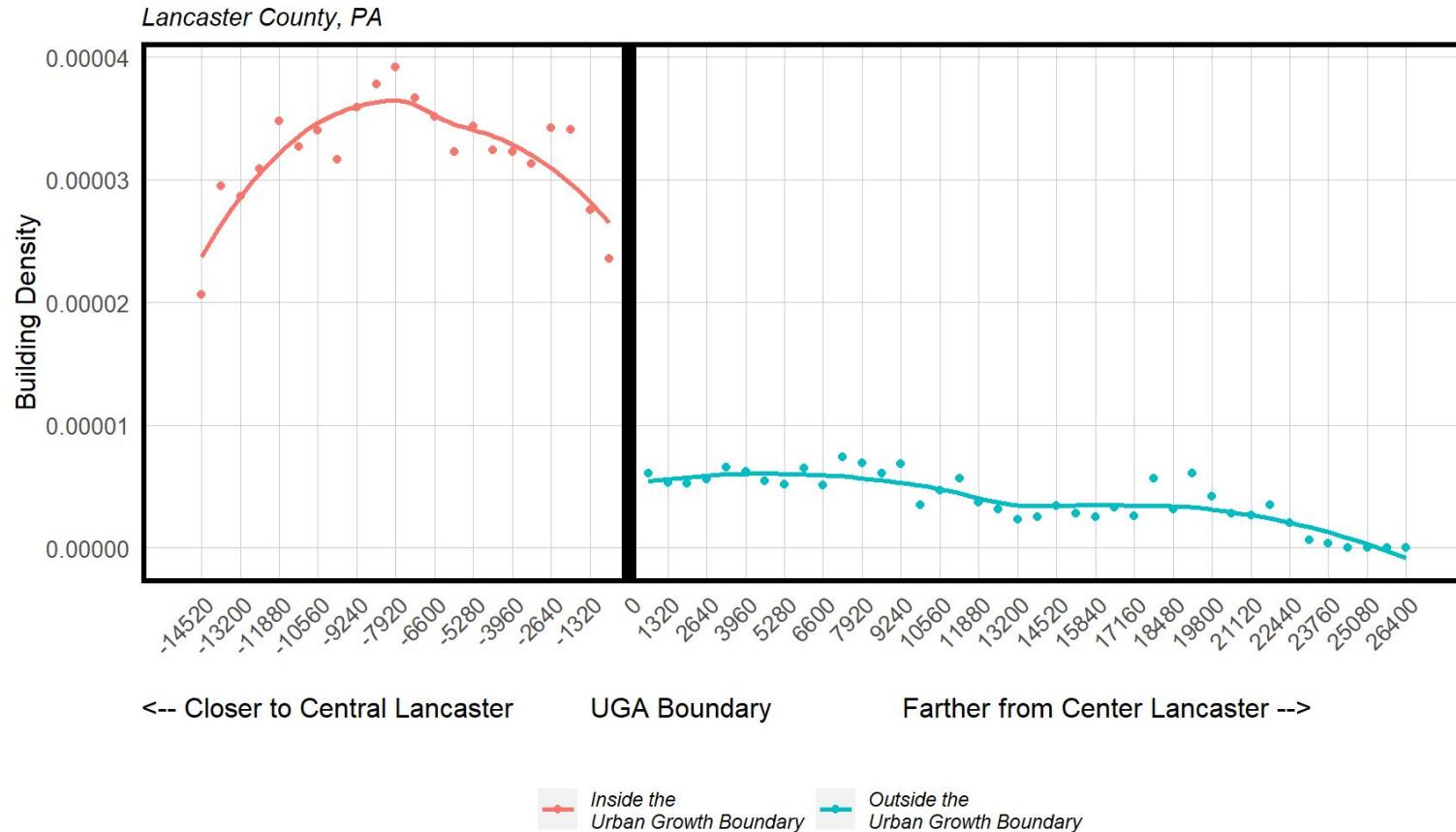
## Lancaster bid-rent



# Bid-rent curves for select U.S. Cities

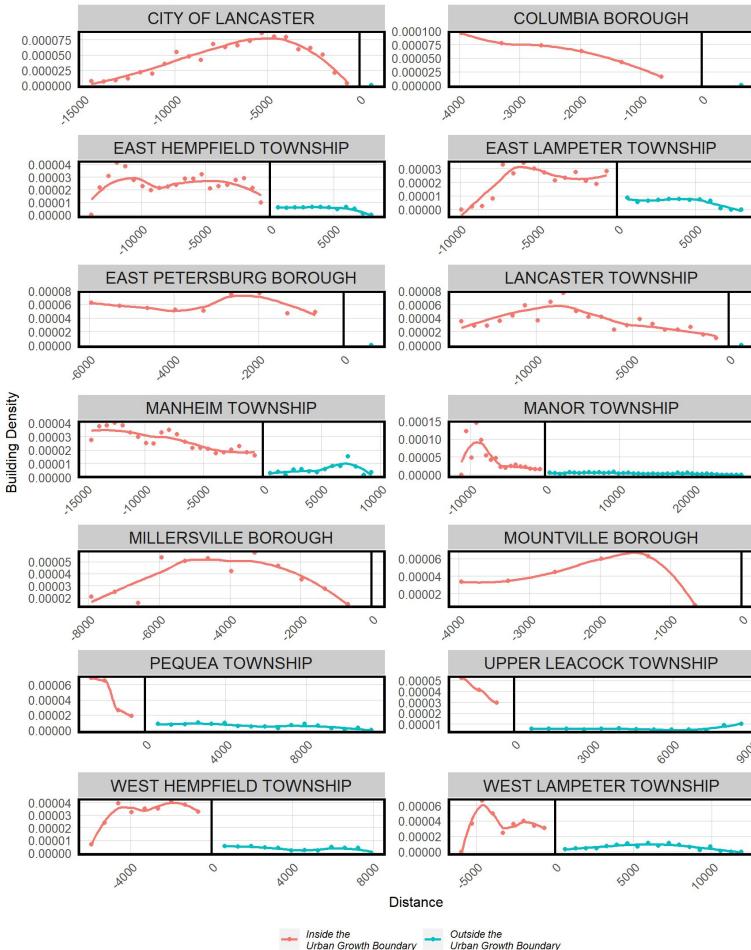


# Bid Rent: Building Count as a function of distance to the Urban Growth Area



## Bid Rent by town

Lancaster County, PA



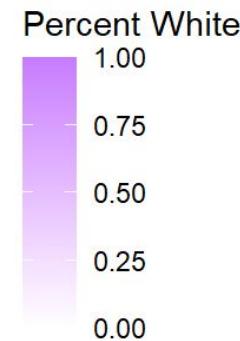
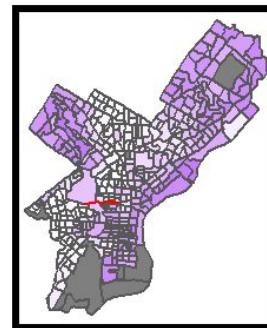
Which town would you extend the boundary?

**Boundaries as a natural  
economics experiment -  
assumptions and design**

**Boundaries in your community**

# Percent White

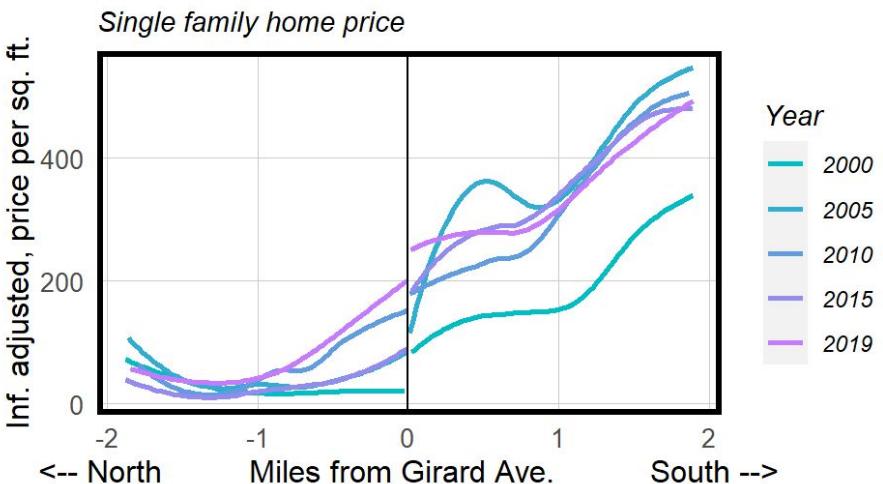
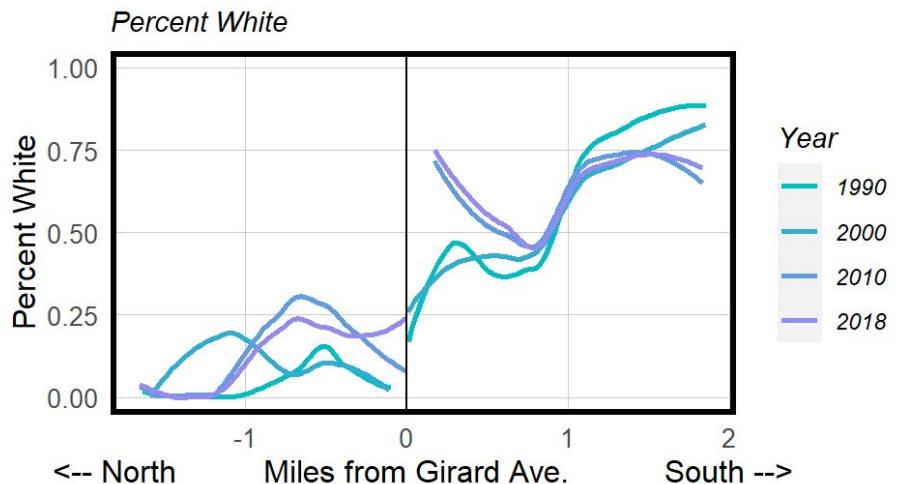
*Girard Avenue in red*



# Distance to Girard



## How Girard Avenue partitions two Philadelphia neighborhoods



Girard Avenue running from Broad St. west to the river. Includes Census tracts and real estate transactions within 2 miles of the Avenue. X-axis units in miles.

# Lab objectives

1. Assignment 1 post-game show
2. Introduce lab geo-processing and visualization objectives
3. Chapter 2 code exercise (Lancaster County) in groups

**Fin**

## **Use case 2:** Building a better Social Service Delivery System

Governments at all levels spend millions on delivering social, health and economic services to disadvantaged populations.

Governments lack the capacity to deliver these services themselves, so they contract with non-profits.

*This is a terrific example of how a budgetary decision-making, which places huge sums of money at the disposal of a single agency, must then be converted to operational decision making. Here, a single **agency** must figure out how to allocate millions to **non-profits** and eventually individual **clients**.*

**What was the ‘business as usual’ approach to service delivery?**

?

**Can we take the client by service level data and use it to forecast local demand?**

If so, we can do a better job ensuring the demand for services are properly aligned with the supply of community services.

Such alignment will optimize the cost/benefit ratio for the agency.



## Aging and Disability Services - Client Level

Data 2016

[View Data](#)

Visualize

Export

API

...

The City's Aging and Disability Services (ADS) addresses the environmental, economic, and social factors that influence the health and well-being of older adults. In an effort to ensure all older adults experience stable health and can age in place, the Human Services Department invests in a combination of direct services and in funding agencies that serve our older adults. ADS supports older individuals, those living with a disability, and their families.

This dataset is taken from an ADS database that tracks services to clients. That is, each record in the dataset holds information from a specific service provided to an individual client. Services are provided both by direct City services and by those the City contracts with. Due to the sensitivity of this data, efforts have been made to remove any personally identified information from the records. As different service providers are required to collect different kinds of information, some fields appear NULL or blank.

[Less](#)

Updated  
May 15, 2019

Data Provided by  
City of Seattle

### About this Dataset

Updated  
**May 15, 2019**

Data Last Updated September 8, 2017  
Metadata Last Updated May 15, 2019

Date Created  
September 8, 2017

#### Refresh Frequency

Frequency One Time

#### Department

Department Human Services Department

# Indicators

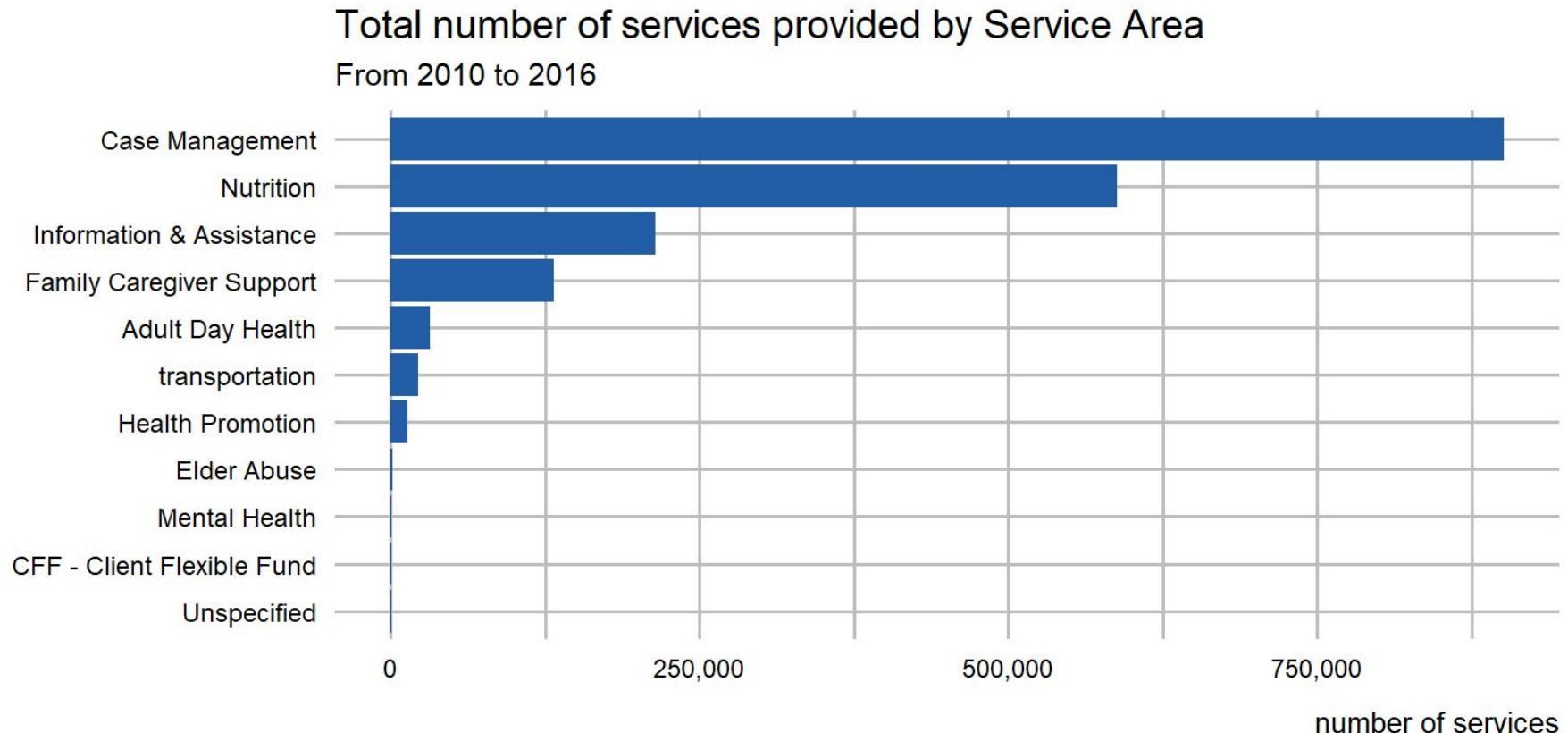


Figure 8

# Indicators

Number of clients over the last seven years

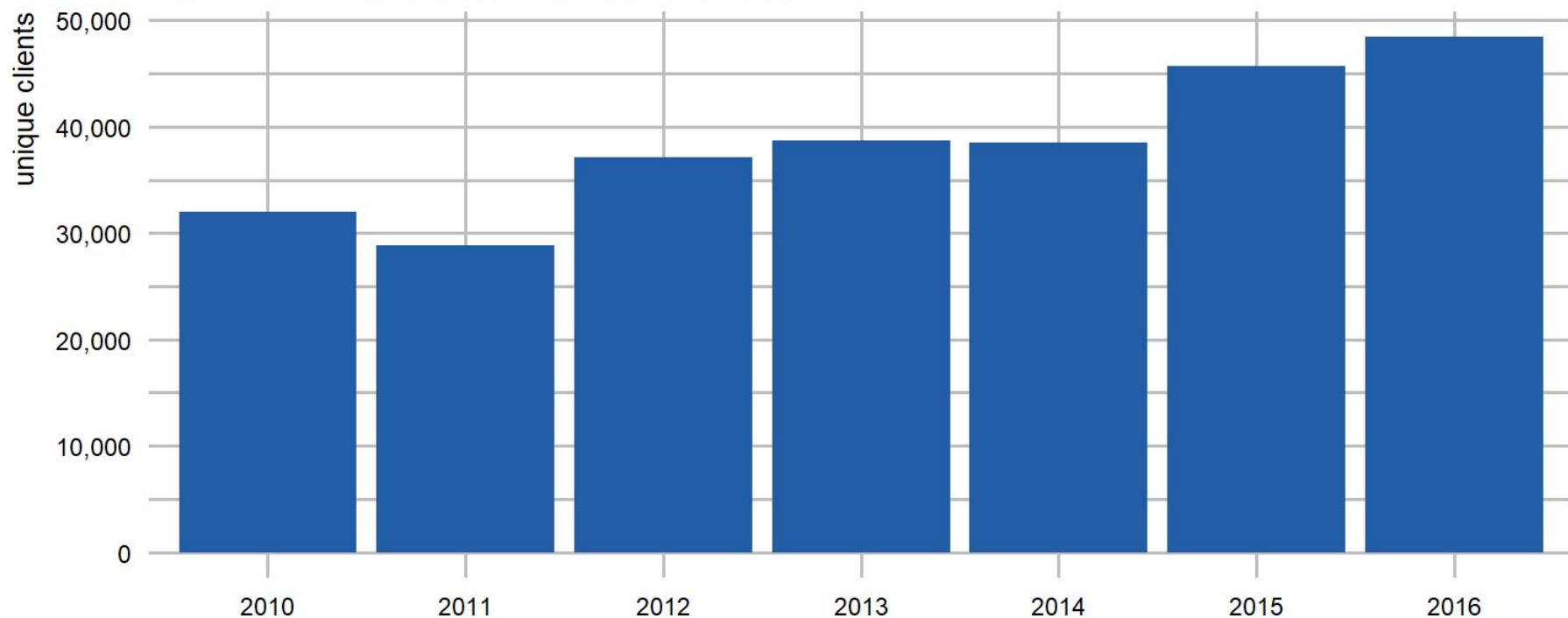


Figure 1

# Indicators

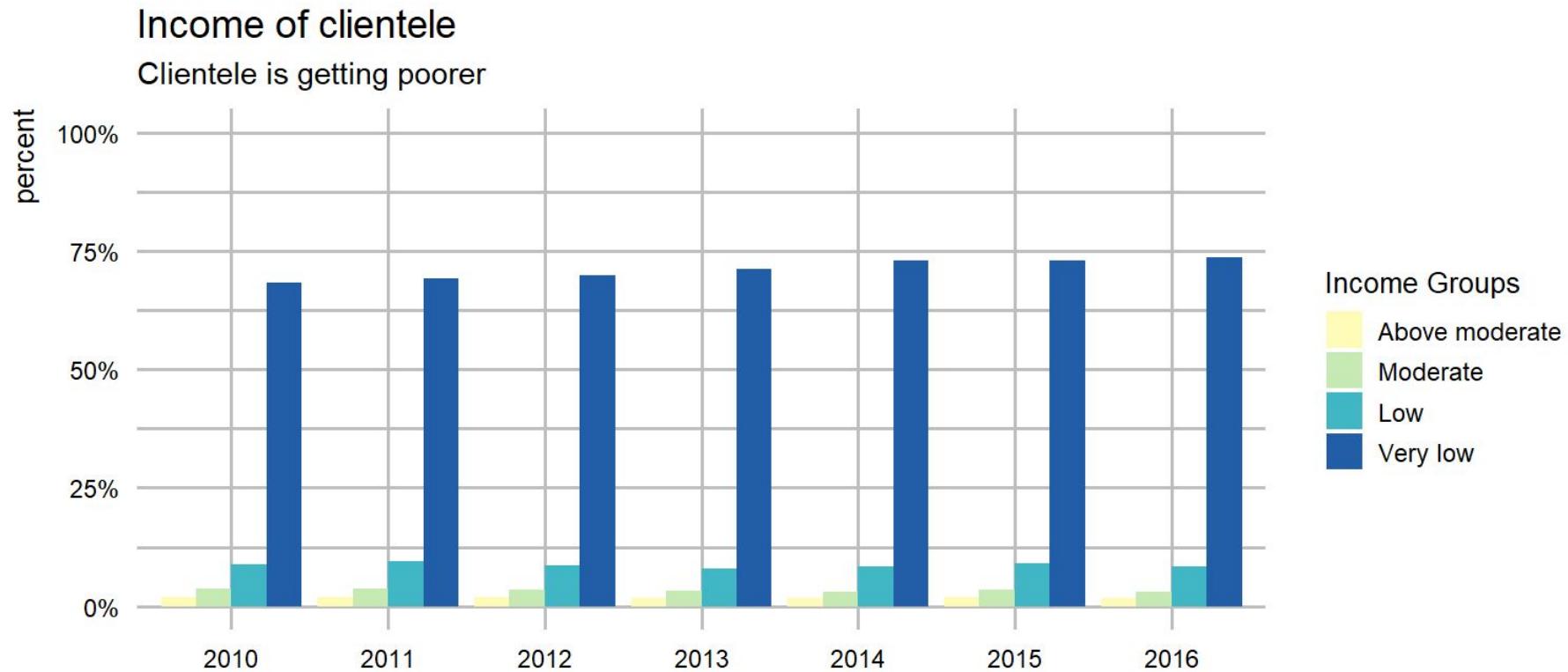


Figure 5

# Indicators

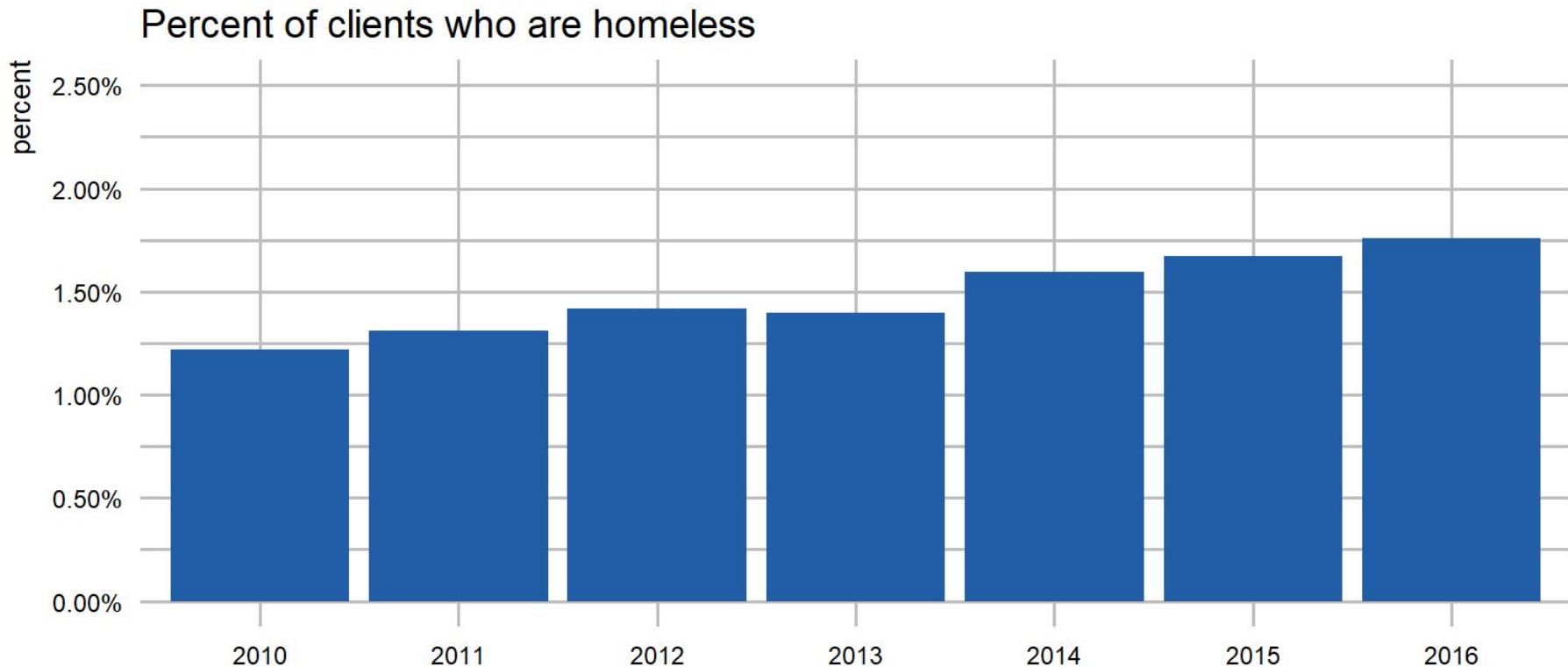


Figure 6

# Indicators

## Activities of Daily Living

Percent of clients who need help with these activities

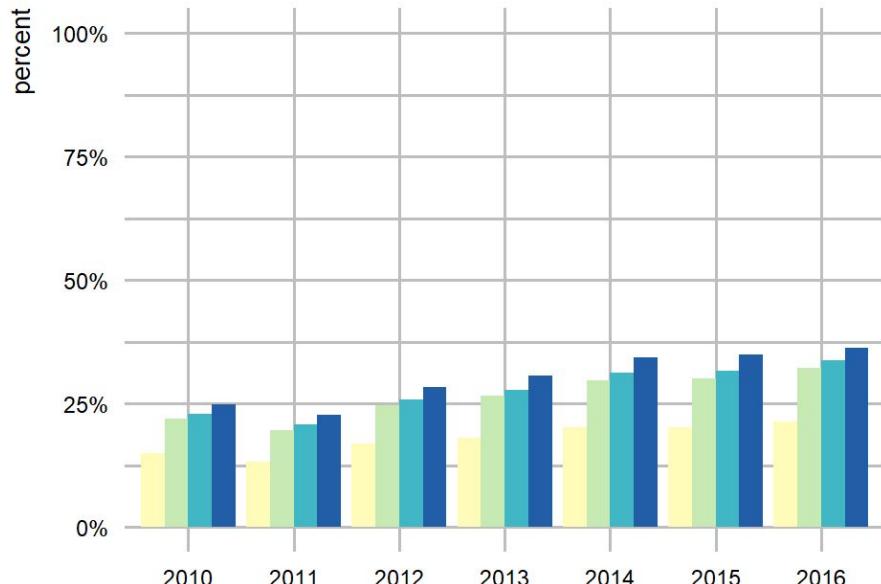
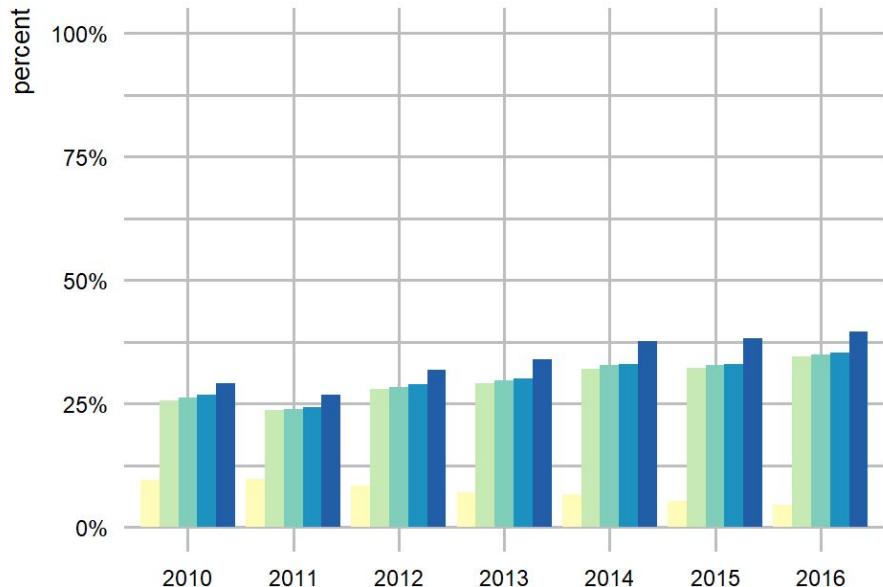


Figure 10

## Instrumental Activities of Daily Living

Percent of clients who need help with these activities



Heavy Housework   Driving   Cooking   Chores   Shopping

# Indicators

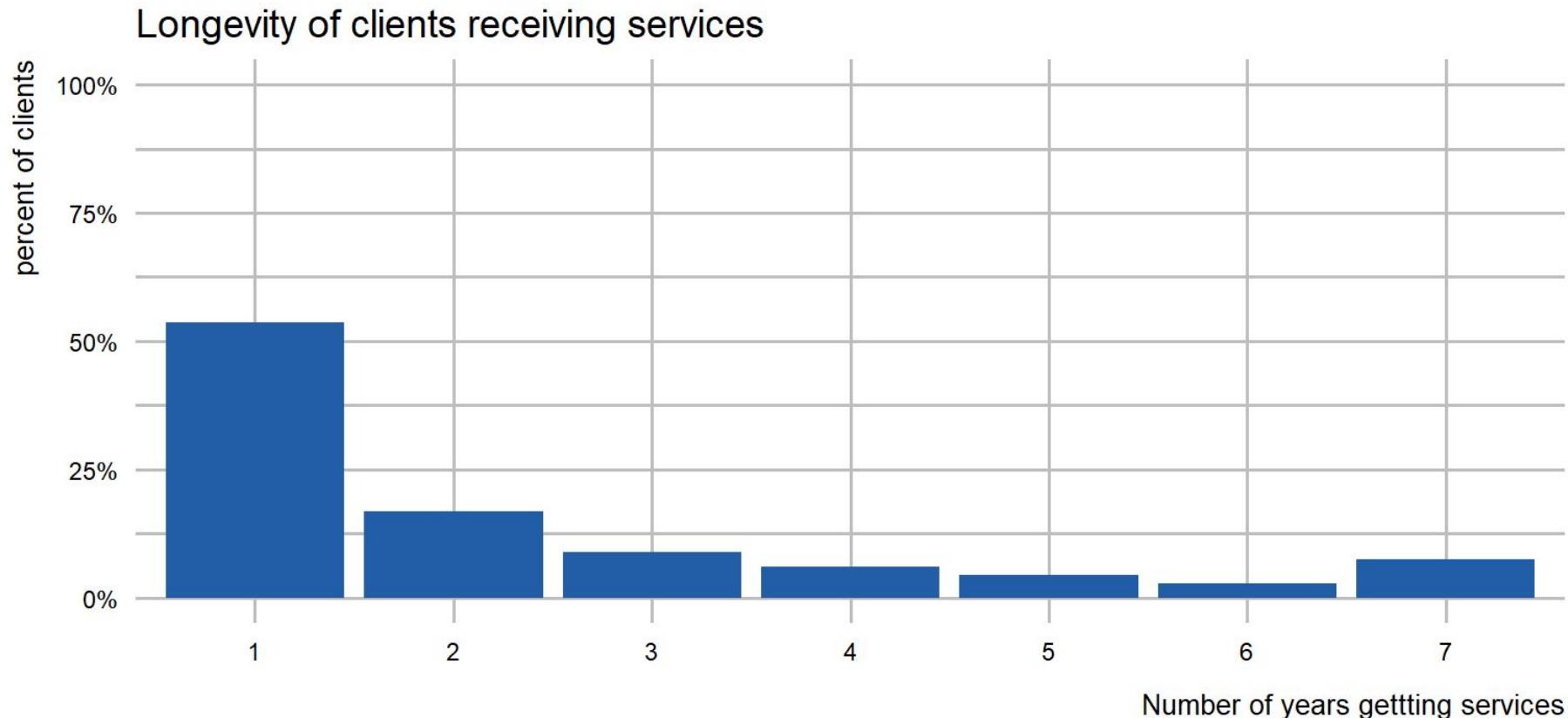


Figure 11

# Indicators

Average length of receiving services by demographic groups



Figure 13

# Forecast - Healthcare analytics

The models were trained on a sample of 2010 - 2015 data and then tested on clients who received services in 2016

The data are split into five different 'stories':

- Demographics
- Location
- Activities clients need help with
- Patient history
- Service Providers for that service

Three separate models are estimated for each story - a Negative Binomial model, a Random Forest model, and an XGBoost model. We then 'ensemble' the predictions from these sub-models into a final model.

# Forecast

## Predicted count of Nutrition services for ensemble models

2016 test set

Vertical line represents observed count

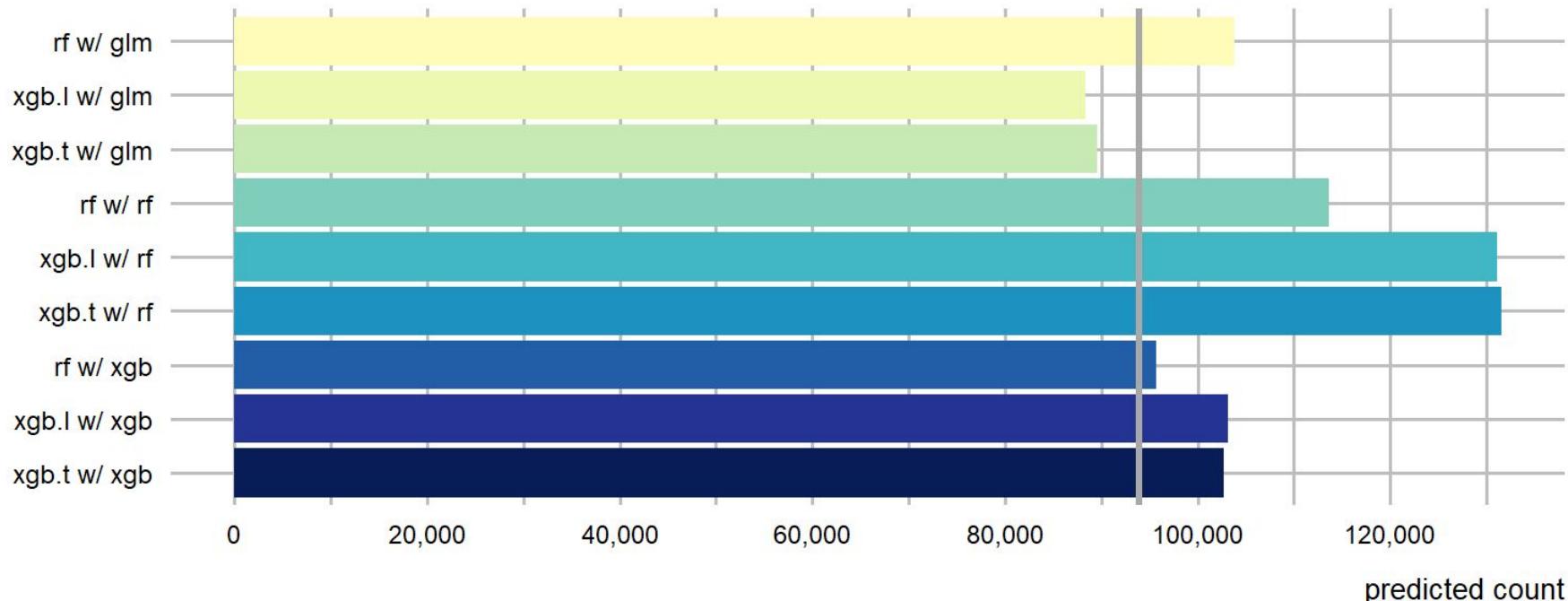


Figure 14

# Forecast

## Goodness of fit metrics for ensemble models

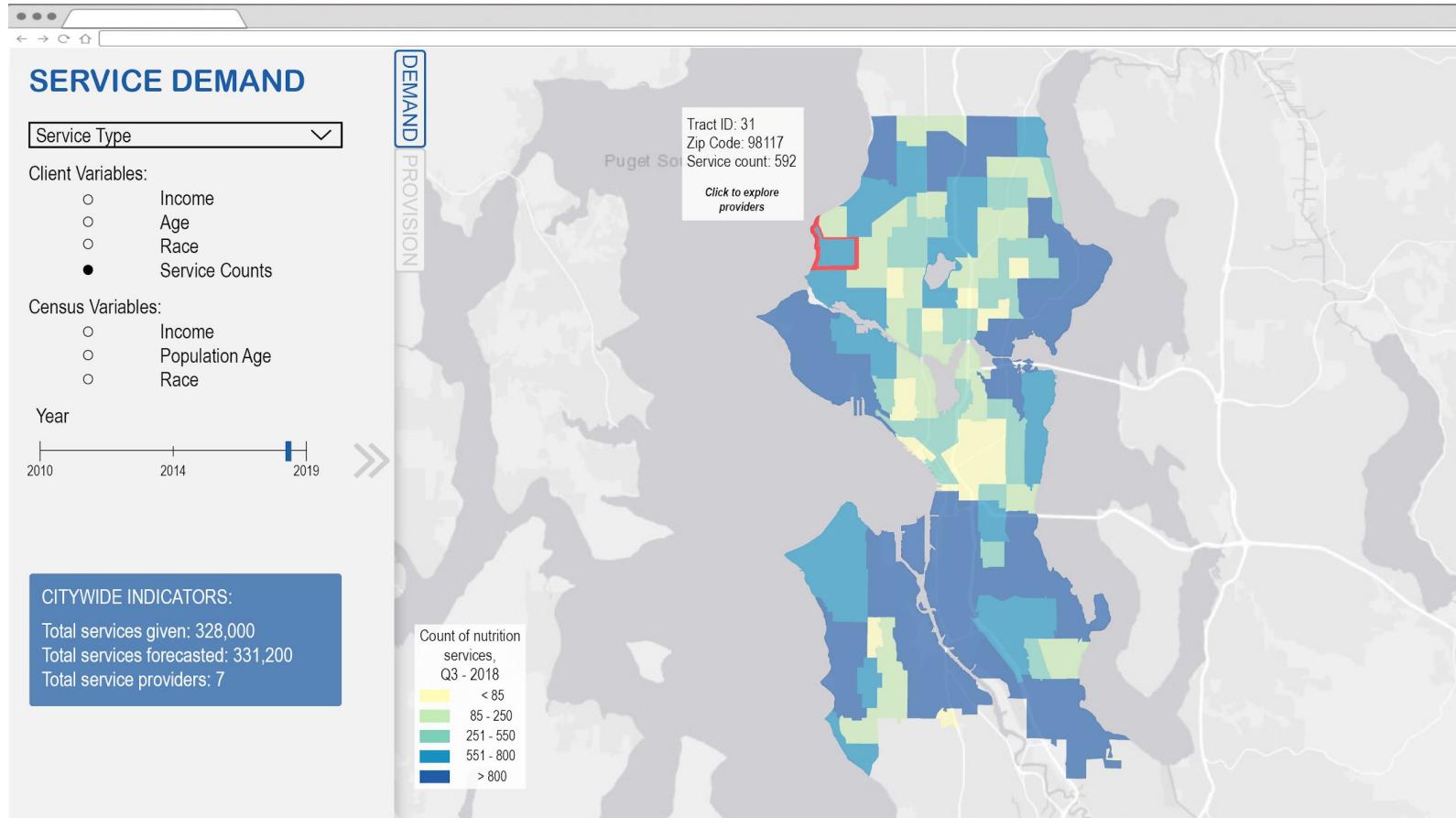
Mean Absolute Error (MAE) for 2016 test set

Vertical line represents mean observed counts



Figure 15

# Dashboarding and demand/supply alignment



# Dashboarding and demand/supply alignment

## SERVICE DEMAND

Service Type

Client Variables:

- Income
- Age
- Race
- Service Counts

Census Variables:

- Income
- Population Age
- Race

Year

2010      2014      2019

**CITYWIDE INDICATORS:**

Total services given: 328,000  
Total services forecasted: 331,200  
Total service providers: 27

**DEMAND**

Puget Sound map showing service distribution by tract. A callout for Tract ID: 31, Zip Code: 98117, Service count: 592. Click to explore providers.

**Observed vs. forecasted services**

The chart displays two lines: a solid green line for Tract 31 and a dashed blue line for Citywide. Both lines show a general upward trend over time, with some fluctuations. The legend indicates that dotted lines represent forecasts.

Period	Tract 31 (Observed)	Citywide (Observed)	Tract 31 (Forecast)	Citywide (Forecast)
Q1, 2017	~160,000	~320,000	~160,000	~320,000
Q3, 2017	~180,000	~330,000	~180,000	~330,000
Q1, 2018	~180,000	~330,000	~180,000	~330,000
Q3, 2018	~160,000	~340,000	~160,000	~340,000

**Client vs. census composition**

This bar chart compares the percentage of clients (yellow bars) and the percentage of the census population (blue bars) across four demographic groups. The y-axis represents the percentage from 0% to 100%.

Demographic Group	Clients (%)	Census (%)
below poverty line	~80%	~55%
> 65 yrs	~80%	~45%
white	~20%	~70%
black	~10%	~5%

**Auditing metrics**

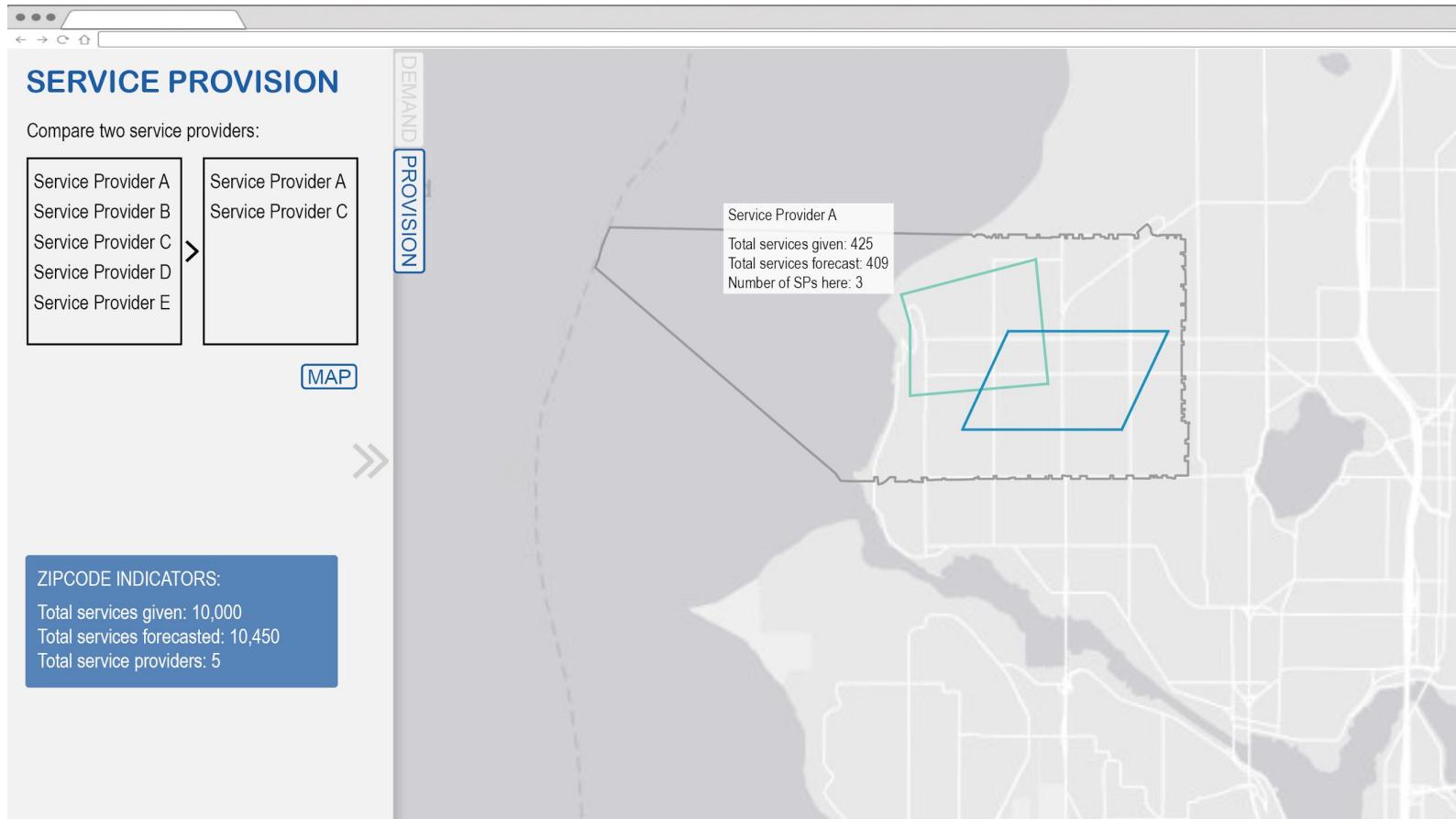
Placeholder for auditing metrics.

**Activities of daily living**

This bar chart shows the percentage of clients (y-axis, 0% to 100%) who engage in five daily activities. The x-axis categories are eating, driving, cooking, bathing, and walking.

Activity	Percent of clients
eating	~15%
driving	~35%
cooking	~35%
bathing	~35%
walking	~45%

# Dashboarding and demand/supply alignment



# Dashboarding and demand/supply alignment

## SERVICE PROVISION

Compare two service providers:

- Service Provider A
- Service Provider B
- Service Provider C
- Service Provider D
- Service Provider E

>

- Service Provider A
- Service Provider C

**MAP**

**ZIPCODE INDICATORS:**

Total services given: 10,000  
Total services forecasted: 10,450  
Total service providers: 5

## DEMAND PROVISION

Observed vs. forecasted services

Number of services

\*dotted lines represent forecast

Period	Provider A (Obs)	Provider A (Forecast)	Provider B (Obs)	Provider B (Forecast)	Zip code (Obs)	Zip code (Forecast)
Q1, 2017	5,250	5,300	4,800	4,900	4,800	4,900
Q3, 2017	5,500	5,600	5,000	5,100	5,000	5,100
Q1, 2018	5,800	5,900	5,300	5,400	5,300	5,400
Q3, 2018	6,000	6,100	5,600	5,700	5,600	5,700

Client vs. census composition

Percent

clients

census

Category	Clients (%)	Census (%)
below poverty line	~70	~55
> 65 yrs	~70	~45
white	~30	~65
black	~10	~5

Auditing metrics

Activities of daily living

Percent of clients

Activity	Percent of Clients (%)
eating	~20
driving	~35
cooking	~35
bathing	~35
walking	~40

# Comprehensive Planning & Forecasting

## Housing

What trends today could we never have forecasted in 2001?

# Comprehensive Planning & Forecasting

What **Housing** trends do you think are going to be important in 2041?

