Big Data for Transportation Models

Neal Bowman
Sales Engineer
neal.bowman@streetlightdata.com

STREETLIGHTDATA

Agenda for Today's Presentation

I. **What is Big Data?**

II. **Why Am I Talking About It?**

III. **Different Data for Different Models**

IV. **Real-World Examples**

V. **Q&A**

STREETL GHTDATA

What is Big Data?

Big Data is A Meaningless Buzzword...

...But Geospatial Records Created by Mobile Devices Can Enhance and Improve Data-Driven Modeling

LBS Data Location
Circle radii vary: they accurately reflect the spatial precision of each unique data point

Navigational GPS Location
Circle enlarged for visibility

Note: This image shows a filtered subset of data to improve visibility
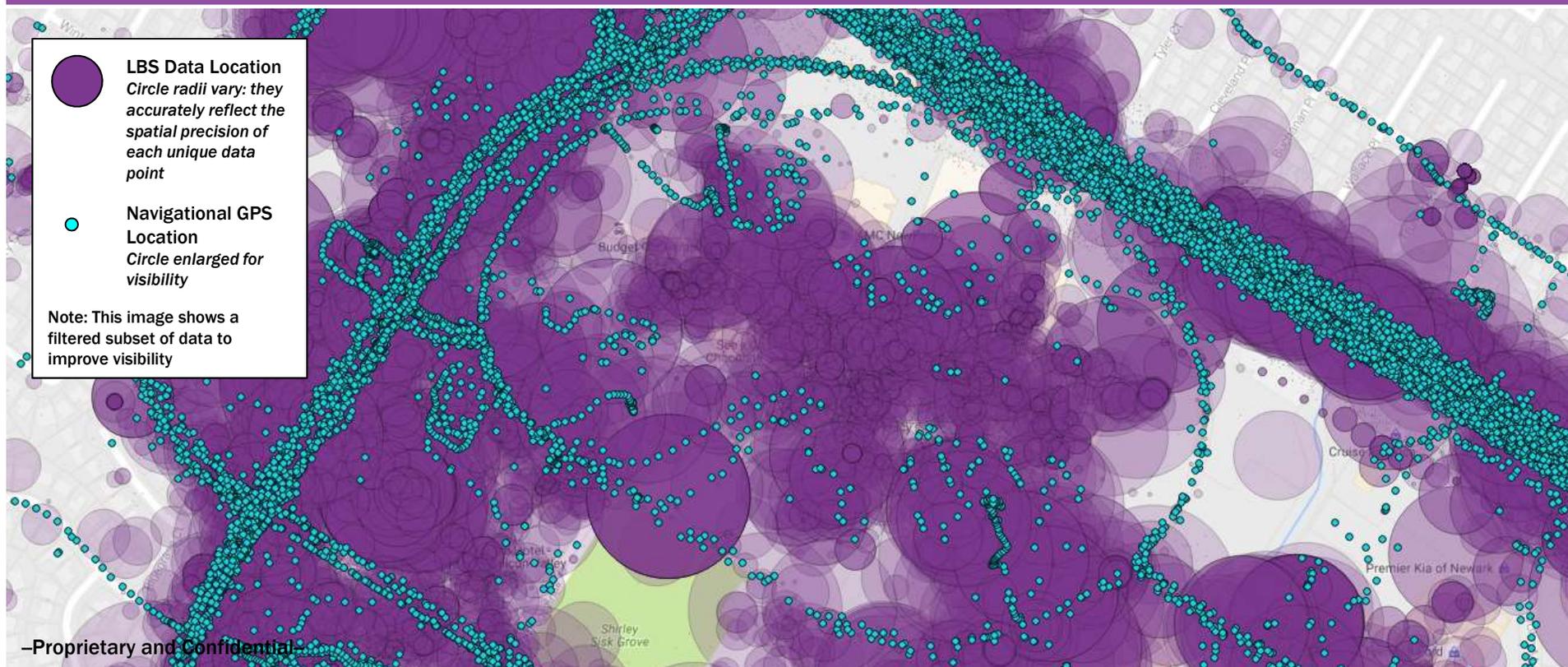
Today, We'll Discuss Three Types of Big Data

Cellular Tower

GPS
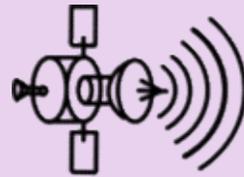
Location-Based Services

The Impact of Spatial Precision is Significant

5 Meters

20 Meters

50 Meters

300 Meters
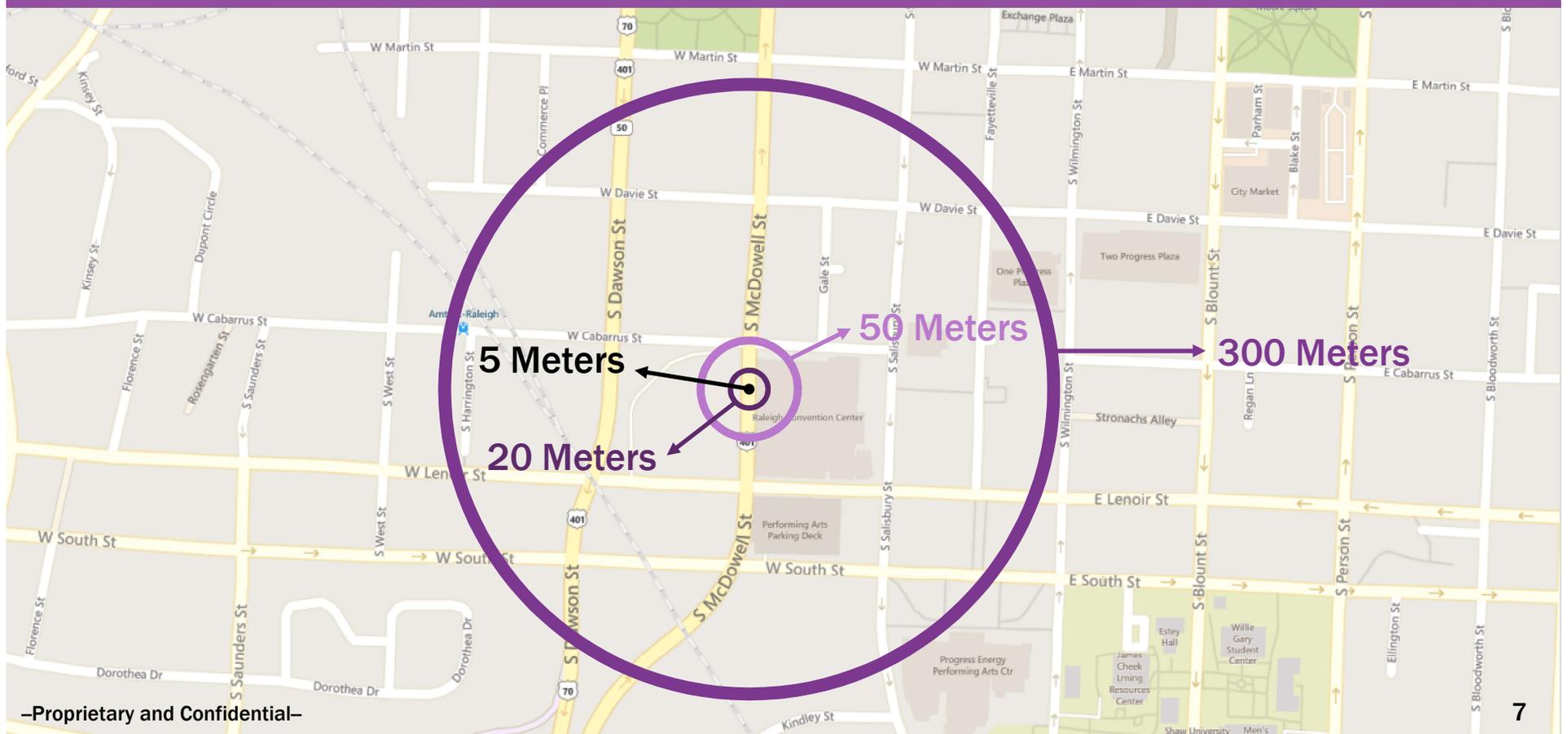
This Is What You Get With 1000m Spatial Precision
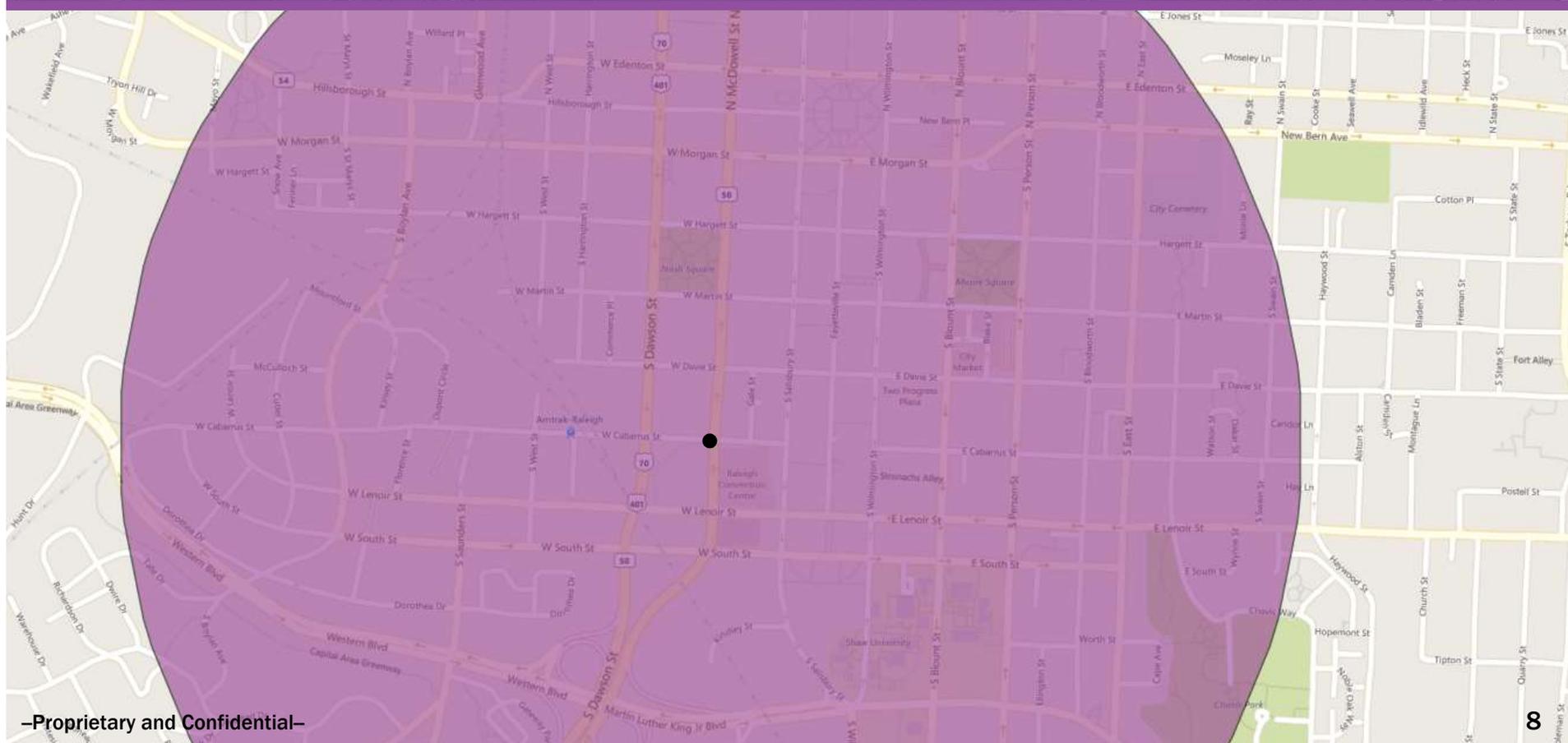
Different Types of "Big" Geospatial Data Offer Different Benefits for Modelers

|	Cellular	Navigation-Based GPS	Location-Based Services	Key Takeaway
Spatial Precision	200-1000 meters	5 meters	5 meters – 50 meters	Cellular data doesn't show exact routes, origins, or destinations because it is spatially imprecise.
Frequency of Data Pings	Irregularly; every 15 min – hours	Regularly; every 1 sec – 1 min	Variable; sometimes triggered by location change	Infrequent pings mean cellular data is messy; a LOT additional processing is required to clean incomplete trips.
Accuracy of Output Metrics	Medium-Low for small geographies	High, assuming good normalization and contextualization algorithms	Medium-high, assuming good normalization & contextualization algorithms.	Lack of spatial precision and infrequent pinging of cellular data leads to inaccurate Metrics.
Type of Trip	Blends personal and commercial trips	Differentiates personal and commercial trips	Blends personal and commercial trips (for now)	Only GPS data segments commercial traffic from personal traffic. Commercial fleet management systems are tagged.
Sample Size	Varies by telco: ~10% of population for small telcos and ~25-30% for large telcos	Varies by region; ~1-4% personal trips; ~10-12% of commercial trips (for INRIX)	Medium – 30M+ US devices in our database (>10% of US adult population)	Cellular provides largest sample, but GPS and LBS sample sizes are statistically valid for any transportation project.

But Why Should You Listen to Me?

STREETLIGHTDATA

Who Is StreetLight Data?

StreetLight Data is a mobility analytics provider that makes it **easy, affordable and intuitive for the transportation industry to use Big Data.**
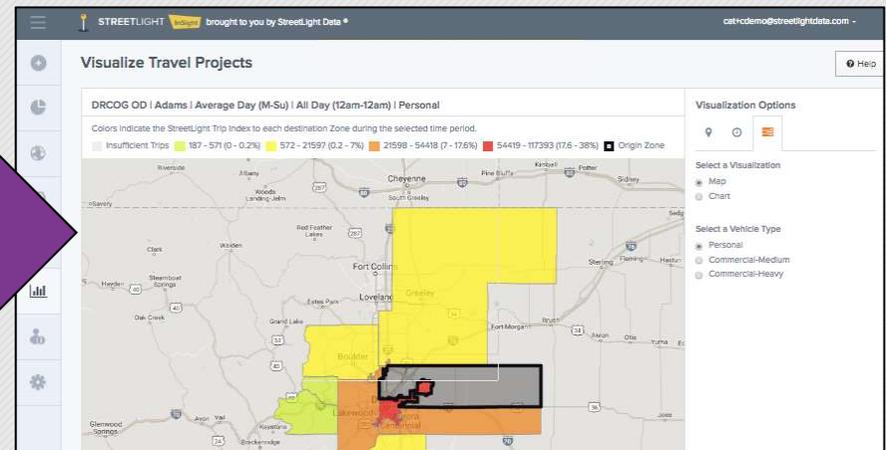
With our easy-to-use web app, you can accurately measure transportation behavior for almost any project in just a few minutes.

We are the First Company to Provide Transportation Metrics Derived from Big Data via a Web App

Input:	Processing:	Web App Output:
Big Data	**RouteScience® Technology**	***StreetLight InSight* Metrics**

Anonymous and accurate Location Data

Road network, land use, parcel, census and more Contextual Data

CLEAN
TRIPIFY
CONTEXTUALIZE
AGGREGATE

- **Core Metrics:**
  Origin-Destination, Select Link, Zone Activity

- **Premium Add-On Metrics:**
  Trip Attributes, Traveler Attributes, Commercial Tours

- **Customization:**
  Day Parts, Day Types, Data Period, Trip Type

- **Speed:**
  Most Metrics are Ready in Minutes

STREETLIGHTDATA

Drilling Down on GPS Data from Our Partner INRIX

| Analyze	To Understand
Precise Origins & Demographics	• Trip generation • Transportation equity • Transit and Facility Accessibility
Precise Destinations & Day/Time	• Commuting & work locations • Transportation centers • "Last mile" and parking • TDM Solutions
Actual Routes and Trip Duration	• Actual travel times ,VMT • Variations, reliability, circuitry • Route choice, diversion, arterials
Vehicle Type and Class	• Personal vs. commercial trips • Criteria emissions • Commercial weight classes
Custom Date Ranges	• "Before and after" studies • Seasonal variance • Trends over time

STREETLIGHTDATA
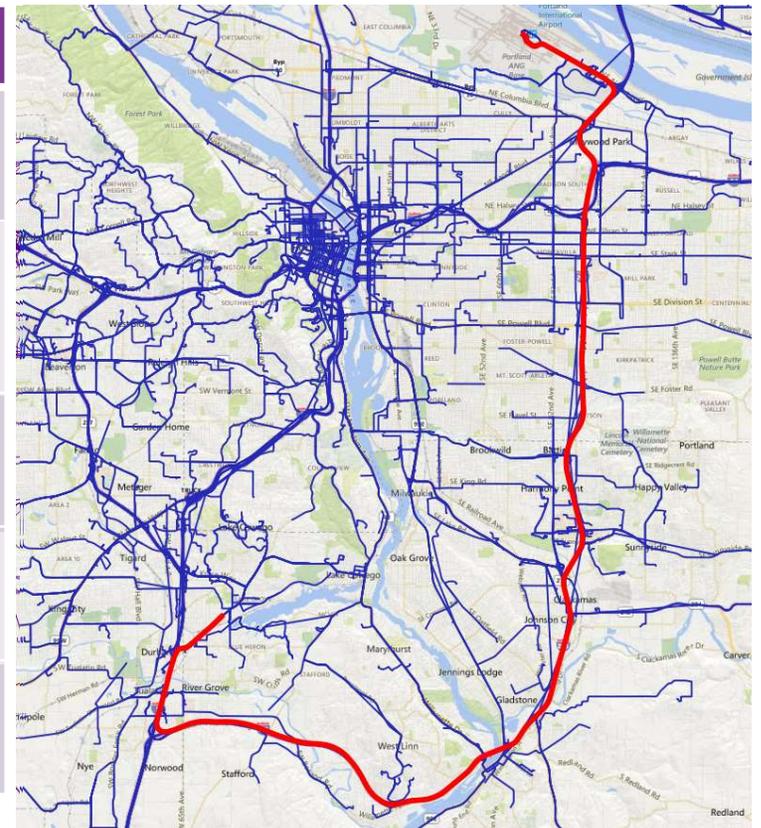
Drilling Down on LBS Data from Our Partner Cuebiq

Our New Data Partner: Cuebiq

- Location Based-Services Data (SDK provider for mobile apps)
- Archival and anonymous
- Spatial precision 50% better than 25m, 75% better than 50m
- Typically, devices collect 30 distinct activity locations per day
- 30M device sample size → 10% of adult US population
- "Privacy first" policy

New and Enhanced *StreetLight InSight* Metrics

- True Trip Purpose (Work, Home, Other)
- Additional Trip Purposes
- Activity-Based Modeling, including personal travel days
- Improved Normalization
- Enhanced Demographics
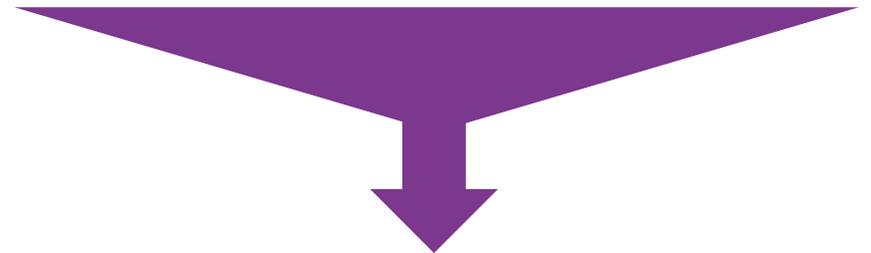
C U E B I Q

STREETLIGHTDATA

We Turn Their GPS and LBS Data into Metrics You Can Actually Use – and Plug into Modeling Software

StreetLight InSight Metrics

Core Metrics:
.shp, .csv, visualizations

Origin/Destination Matrix:
Relative volume and average travel time of trips between Origin and Destination Zones

Origin/Destination + Middle Filter (*Select Link*)
Relative volume and average travel time of trips that pass through Middle Filter Zones, or links, when traveling between Origin and Destination Zones

Zone Activity Analysis
Relative volume, average travel time, and average length of trips that originate in, have destinations in, or pass through each Zone analyzed

Premium Bundles:
.shp, .csv

Premium A: Trip Attributes
Trip time, trip length, trip speed, and trip circuitry

Premium B: Traveler Attributes
Simple trip purpose, household income, race, education level of head of household, and family status

Premium C: Commercial Tours (*2 Month Lead Time; Not Yet in App*)
The relative volume of tours (within 24 hours) and dwell times at each intermediary stop for tours that begin or end in specified Zones

Project Options
Available for All Metrics

- **Custom Day Types**
  Define Weekends and Weekdays

- **Custom Day Parts**
  Six day parts; customize with bins as small as one hour

- **Data Period**
  Choose the months to analyze from January 2014 through 1-2 months ago

- **Trip Type**
  Commercial vs. Personal

- **Truck Type**
  Heavy Duty vs. Medium Duty

STREETLIGHTDATA

Different Models, Different Data

Micro Simulation

GPS

No One Type of Data is Ideal for Every Model: Choose The Right Combination for Your Application

Meso Simulation

GPS

LBS

No One Type of Data is Ideal for Every Model: Choose The Right Combination for Your Application

Macro Simulation

GPS

LBS

Cell.

Real-World Examples

STREETLIGHTDATA

Calibrating a Travel Demand Model in North Carolina

Transport. Demand Modeling
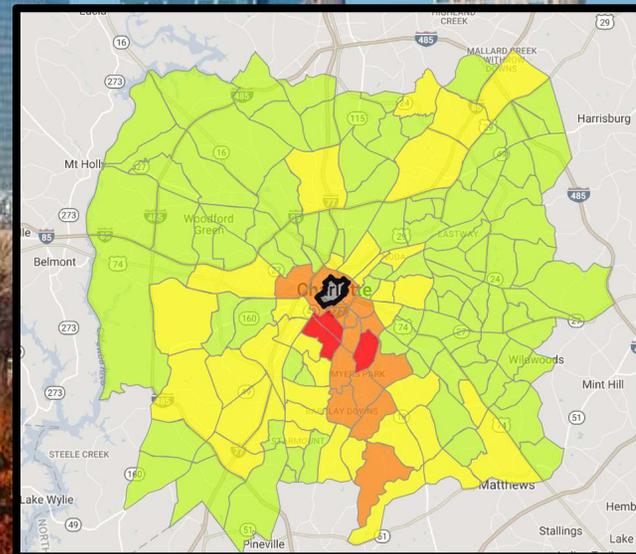
Origin-Destination for North Carolina MPO

Need: Accurate O/D for calibration or transportation demand model without expensive/time consuming survey for personal and medium/heavy duty commercial trips.

Zones: 155 census tracts

Metric: O/D with demographics

Months: Spring 2016 (3 months)

Time to Run: 9 minutes

A Corridor Impact Study for a New Route

Corridor + Facility Improvement

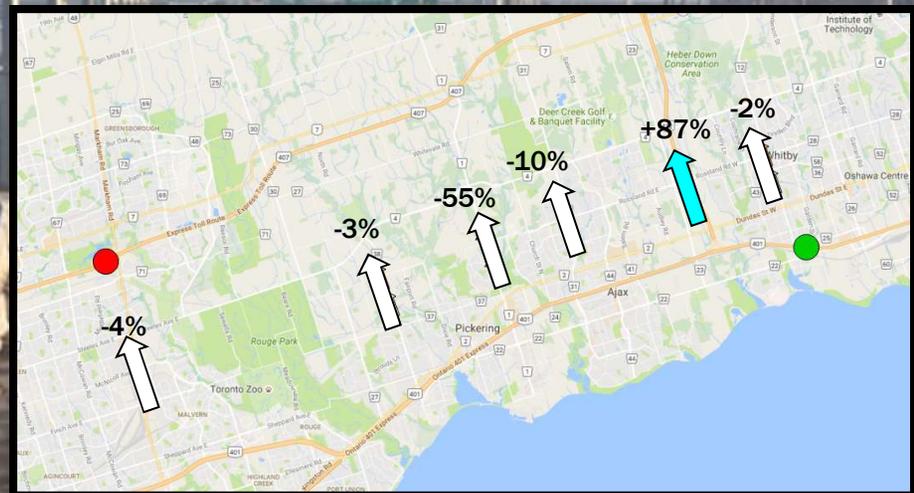## Corridor Impact in Canada

Need: Route choice between 2 highways to measure the impact of offering a new option.

Zones: 6 routes, 3 origin zones, 6 destinations zones.

Metric: O/D with Middle Filter

Months: May, June 2016

Time to Run: 2 min for each month.

Questions?