



service
management
group®

Geospatial Voronoi Analysis

Dan Finkel Lou Bellaire

Service Management Group (SMG)

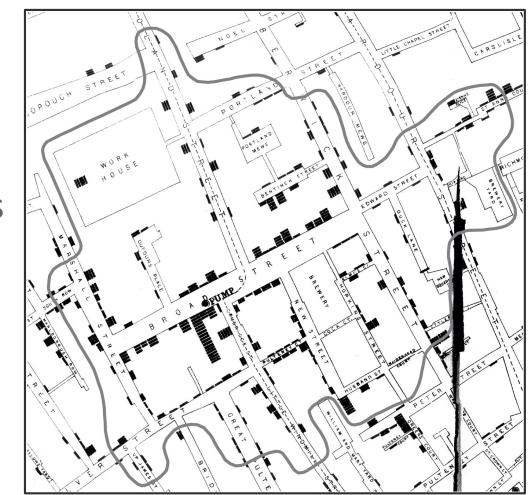
ODSC East

May 5 2017

dfinkel@smg.com
[@danfinkel](https://twitter.com/danfinkel)

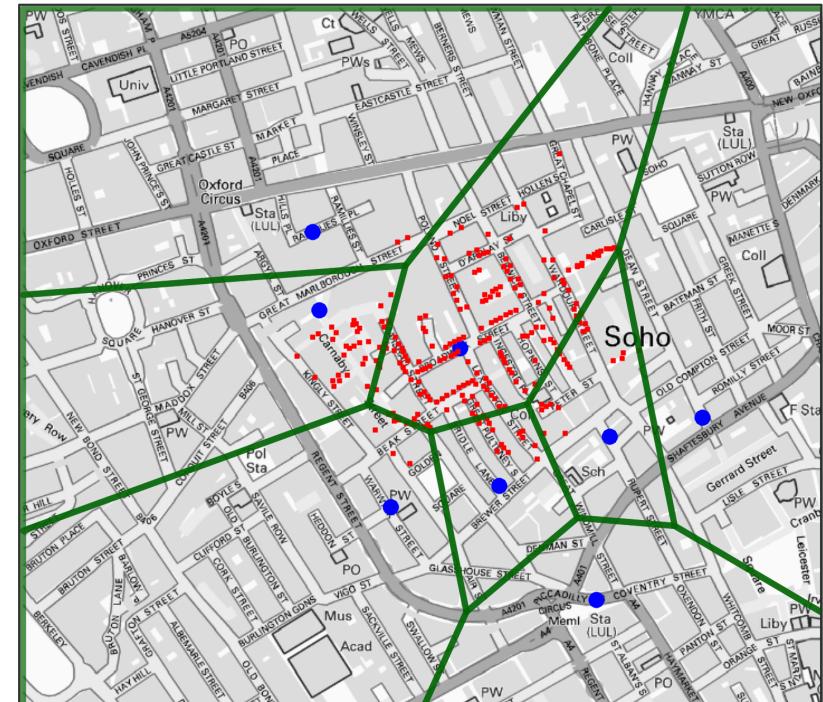
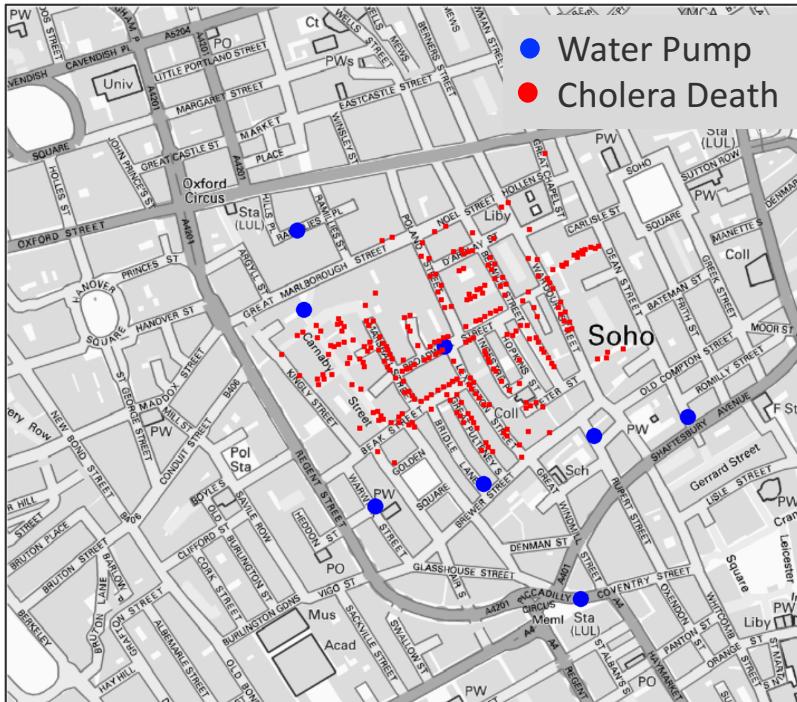
1854 Broad Street Cholera Outbreak

- A major cholera outbreak killed over 600 London Soho residents and caused 75% of residents to flee the area
- A local physician, John Snow, connected the outbreak to filth polluting the water
 - Conclusions were based on graph and pattern analysis
 - Regarded as the founding event of epidemiology
- In the absence of hard biological evidence Snow relied on graph and pattern analysis
- Postscript
 - Snow's evidence caused Broad St well to be disabled
 - After the disease subsided Snow's theory was rejected and air pollution was blamed
 - “To accept Snow's theory of fecal transmission of disease was too unpleasant for most of the public to contemplate”



Original Snow diagrams

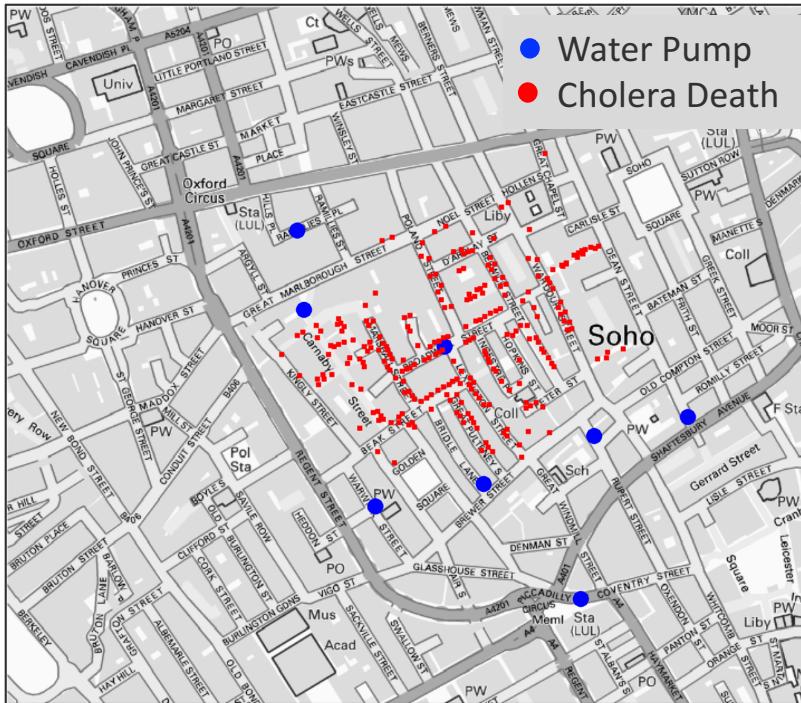
Visualizing the Cholera Outbreak by Water Pump Region



Planar partitions by pump proximity helped Snow argue outbreak was due to polluted water

Data obtained from <http://blog.rtwilson.com/john-snows-famous-cholera-analysis-data-in-modern-gis-formats/>

Visualizing the Cholera Outbreak by Water Pump Region



Data obtained from <http://blog.rtwilson.com/john-snows-famous-cholera-analysis-data-in-modern-gis-formats/>

Voronoi Diagram Overview



Georgy Voronoy
1868 - 1908

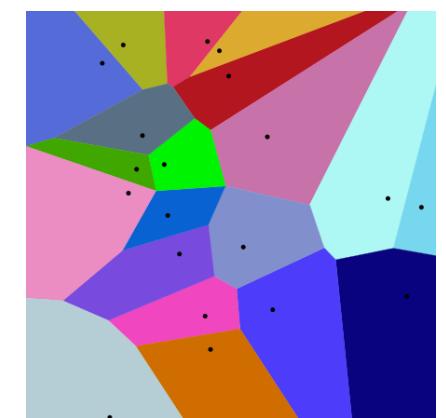
A Voronoi diagram is a planar partition based on distance to a set of points (“seeds”) within the plane

Each seed defines a partition as those points in the plane closer to that seed than any other

$$R_k = \{x \in X \mid d(x, P_k) \leq d(x, P_j) \text{ for all } j \neq k\}$$

Algorithms / implementations

- Typically $O(n \log n)$
- Can scale to multiple dimensions
- Python implementation in `scipy`
 - Utilizes `Qhull` library
 - Computes via its dual (Delaunay triangulation)

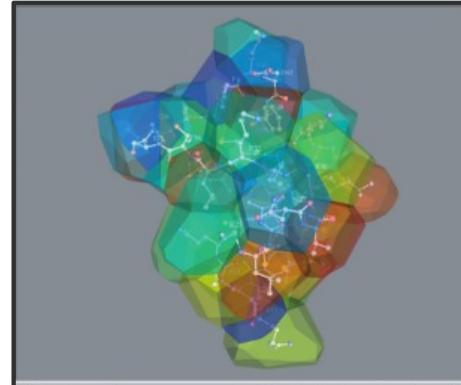


Voronoi Diagram Example (Wikipedia)

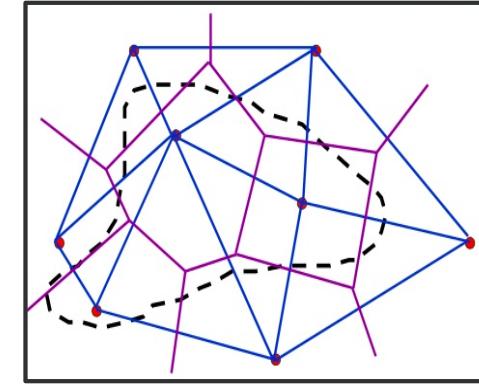
Applications of Voronoi Partitions



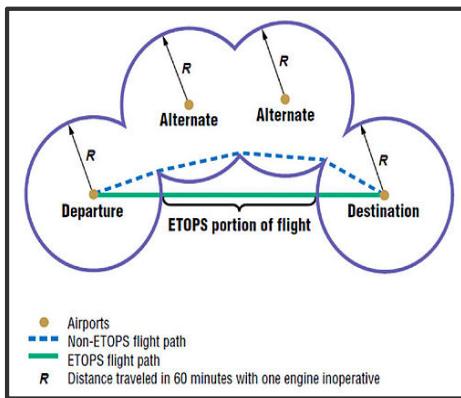
Biology
(Forest Canopies)



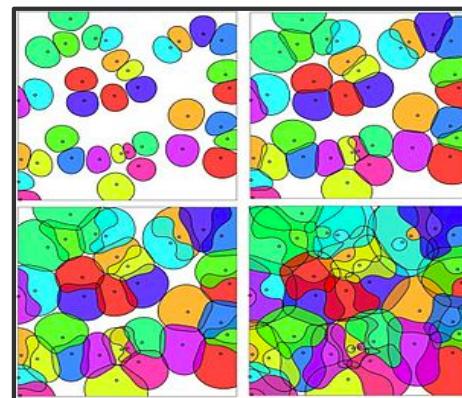
Computational Chemistry



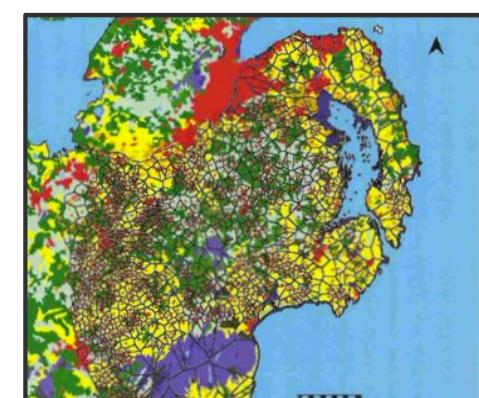
Hydrology
(Rainfall Modeling)



Aviation
(Flight Diversion)



Wireless Network Capacity



Archeology

Constructing Voronoi Polygons in Python

Simple Functionality

1. Import the library

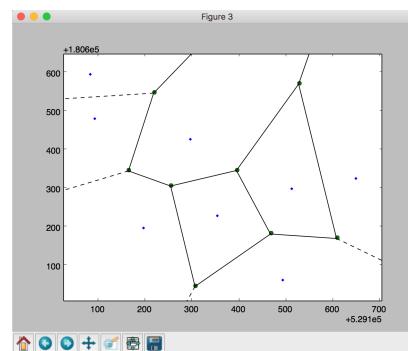
```
> from scipy.spatial import Voronoi, voronoi_plot_2d
```

2. Instantiate

```
> vor = Voronoi(pumps_xy)
```

3. Plot / Visualize

```
> voronoi_plot_2d(v)
```



Constructing Voronoi Polygons in Python

Simple Functionality

1. Import the library

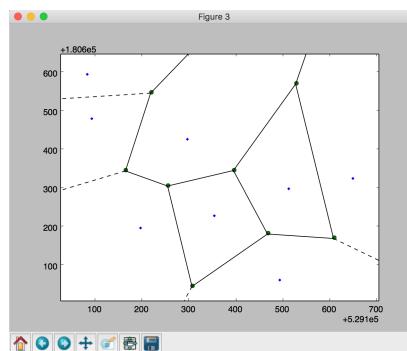
```
> from scipy.spatial import Voronoi, voronoi_plot_2d
```

2. Instantiate

```
> vor = Voronoi(pumps_xy)
```

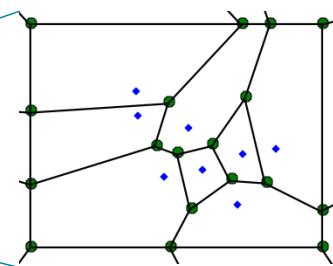
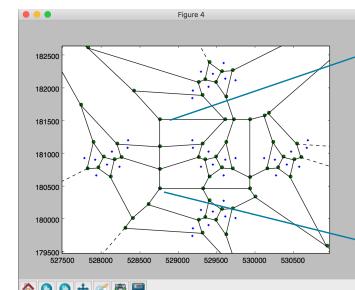
3. Plot / Visualize

```
> voronoi_plot_2d(v)
```



Additional Modifications

Mirror pumps and build Voronoi off expanded set
(to create hard boundaries)



Convert Voronoi regions into polygons

```
def build_vor_polys(vor, soho_poly, pumps_xy):
    polys = []

    # Build all lines in the voronoi region set
    lines =
        [shapely.geometry.LineString(vor.vertices[line])
         for line in vor.ridge_vertices
         if -1 not in line]

    # polygonize the lines and check if
    # they are original or copies
    for poly in shapely.ops.polygonize(lines):
        if is_in(poly, pumps_xy):
            poly = poly.intersection(soho_poly)
            x, y = poly.exterior.xy
            polys.append(poly)

    return polys
```

Shapely Geospatial functions

- Compute area
- Intersect, overlap with other polygons
- Join / union with other objects

All code available on <http://github.com/danfinkel/python>

Outline

- Introduction
- • SurveyMini Overview
- Geospatial Research Examples
 - MLB Stadium Visits
 - Fast Food Store Locations
- Summary



5-star human experiences matter

We believe the more you know about your customers and employees, the more 5-star human experiences you'll create in the future. The more we know about your business, the better and more actionable our insights.

Service Management Group

Market Research and Consulting Services



service
management
group®

The Service Profit Chain



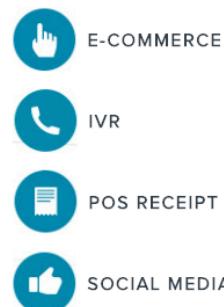
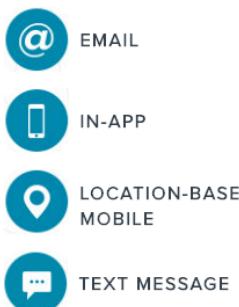
Highly Engaged
Employees

Highly Satisfied
Customers

- High frequent visitors
- Willing to spend more
- Recommend you more often

SMG was founded in 1991 to help clients successfully execute the Service Profit Chain and derive key performance measures

SMG Measurement Streams



Through innovative technology SMG models and measures the service profit chain and develops novel insights that drive client business



We're constantly innovating to help our clients win consumer loyalty and increase profitable sales

MEASUREMENT

TECHNOLOGY

INSIGHTS

- Research on research
- Thought leadership

- Continuous innovation
- Everything mobile

- Change agents
- Business impact



SurveyMini By the Numbers

1 million+

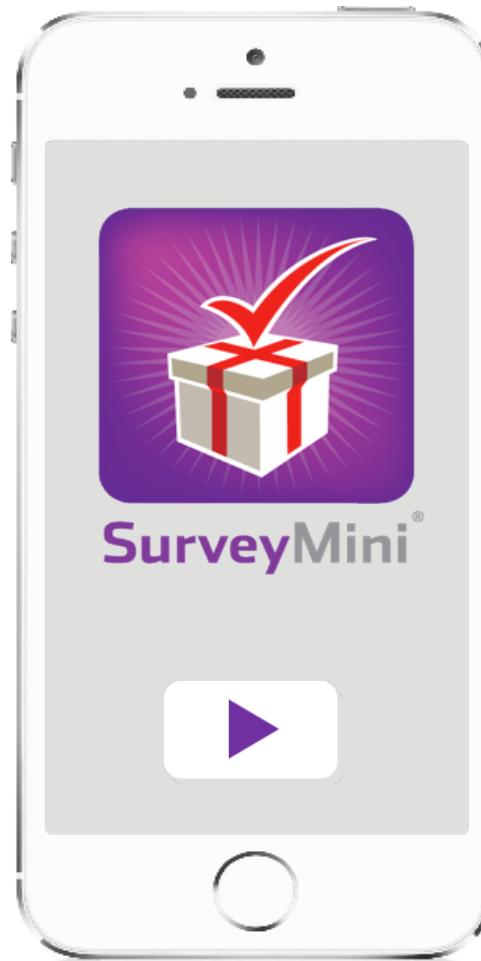
App Downloads

140,000+

Daily Active Users

600,000+

Locations Mapped



20 million+

Location Records Daily

2.5 million+

Daily Detected Visits

60,000+

Daily Triggered Surveys

SMG Data Scientists Harness SurveyMini Data to Measure and Assess the Service Profit Chain for Over 100 Major Restaurant and Retail Brands

Outline

- Introduction
- SurveyMini Overview
- Geospatial Research Examples
 - • MLB Stadium Visits
 - Fast Food Store Locations
- Summary

MLB Stadium Market Research

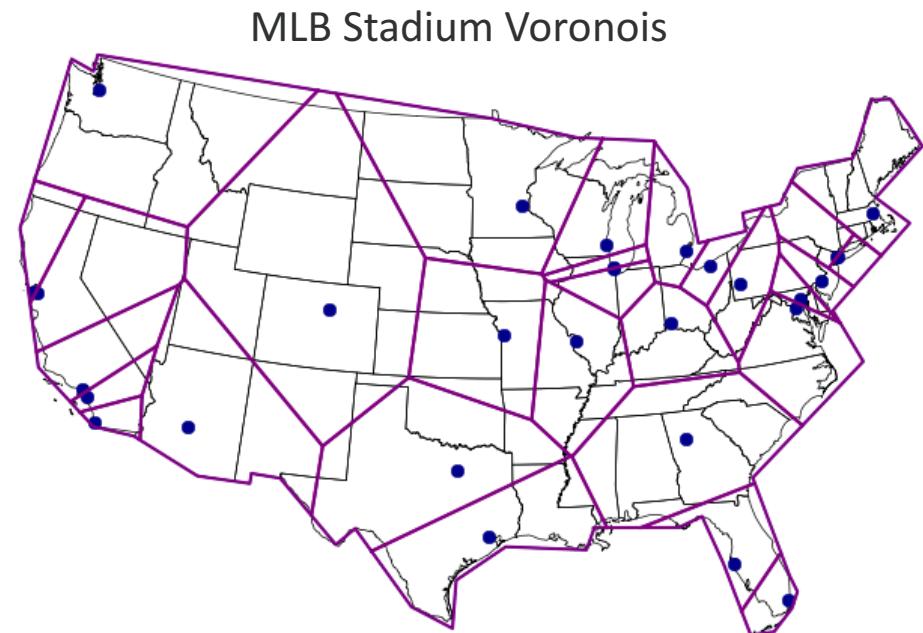
Voronoi Analysis



- Currently there are 30 Major League Baseball (MLB) teams
 - 29 US-based teams
 - Each team plays 162 games from April – September
 - Playoffs / World Series take place every October
- Many interesting geospatial analysis questions to be explored with an app-based panel
 - How many “home” fans do teams draw?
 - How far do fans travel to attend games?

MLB Stadium Market Research

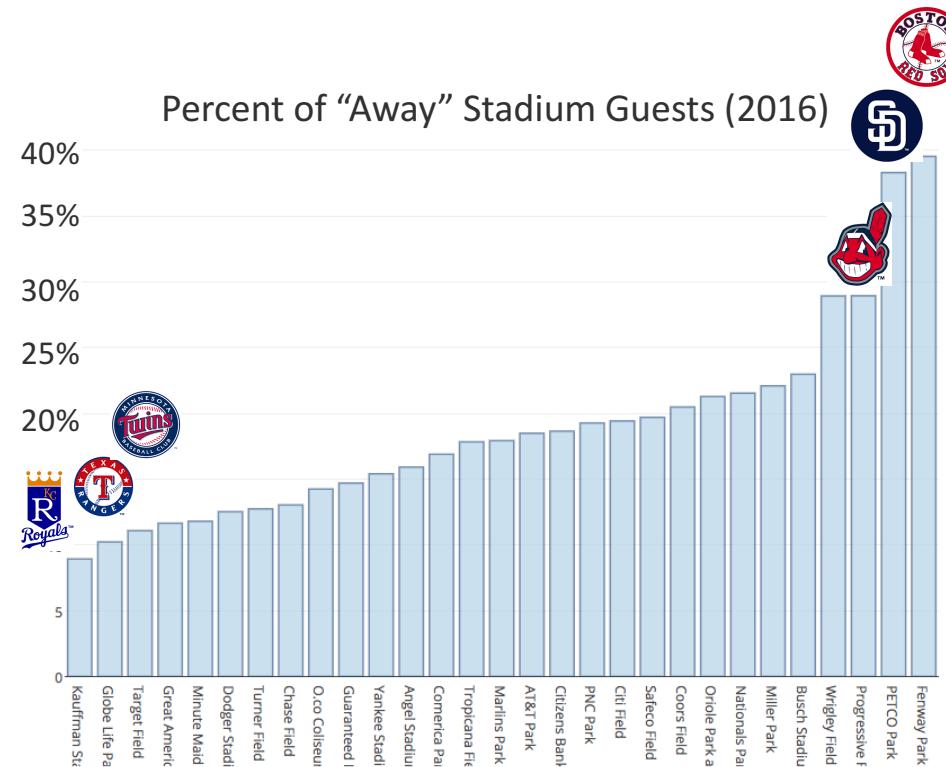
Voronoi Analysis



- Currently there are 30 Major League Baseball (MLB) teams
 - 29 US-based teams
 - Each team plays 162 games from April – September
 - Playoffs / World Series take place every October
- Many interesting geospatial analysis questions to be explored with an app-based panel
 - How many “home” fans do teams draw?
 - How far do fans travel to attend games?

MLB Stadium Guests

Locals vs Out-of-Towners



Binning patrons by Voronoi region illustrates different regional visit patterns across MLB teams

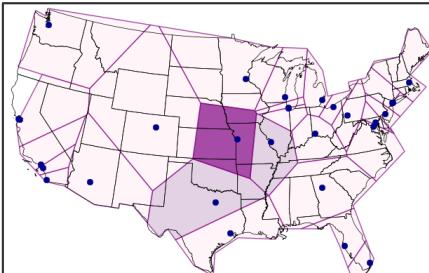
MLB Stadium Guests

Locals vs Out-of-Towners

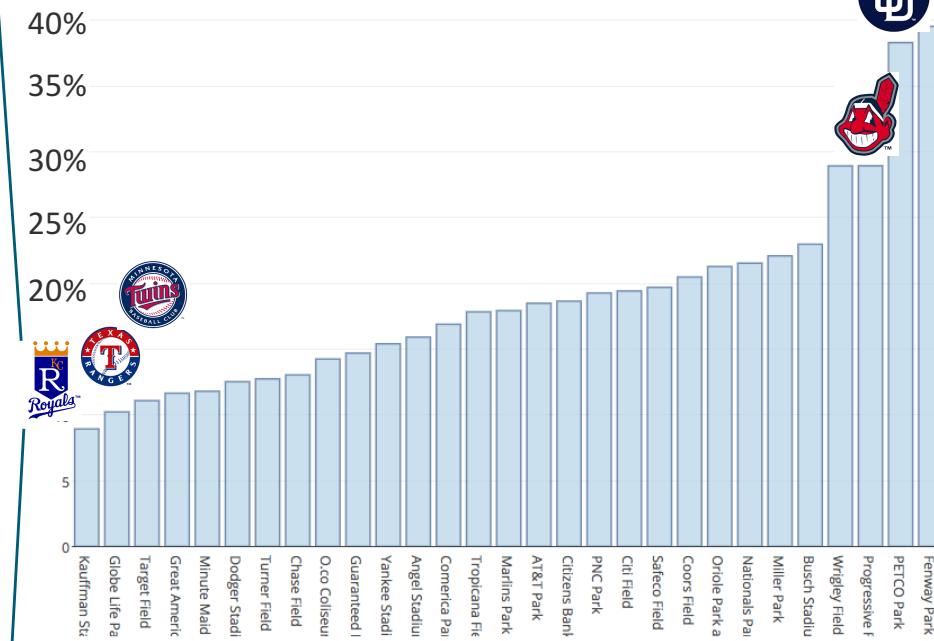
KC Royals Marketing



Where Do Royals Game Patrons Come From?



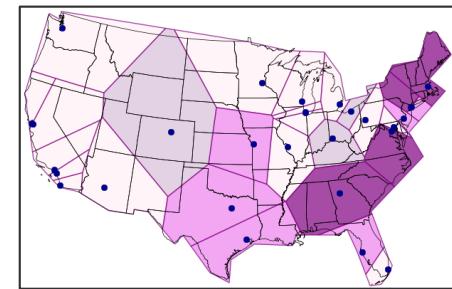
Percent of "Away" Stadium Guests (2016)



Binning patrons by Voronoi region illustrates different regional visit patterns across MLB teams



Where Do Red Sox Game Patrons Come From?



Boston Red Sox Marketing

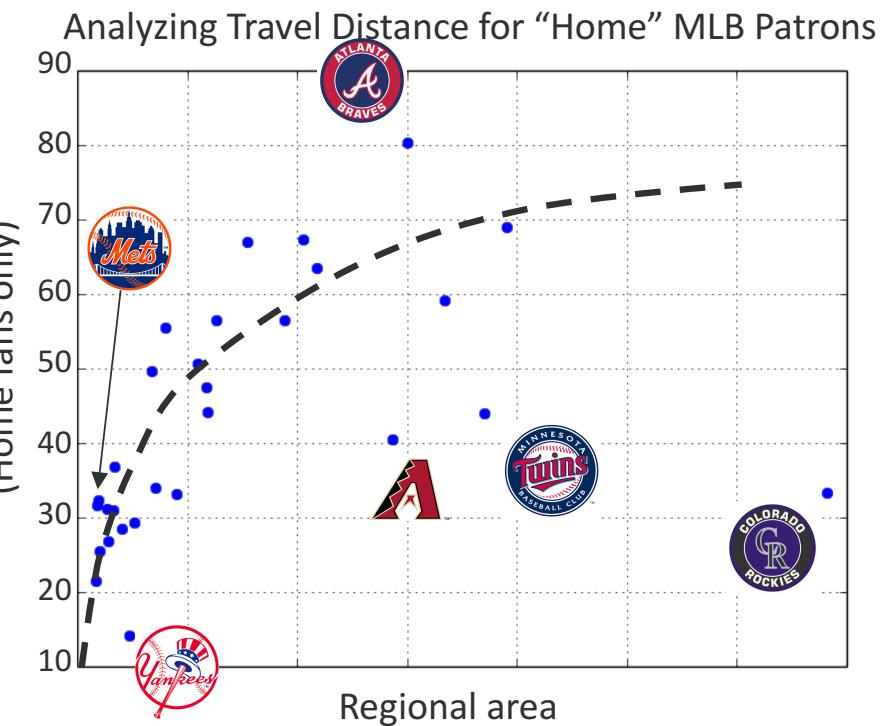


MLB Stadium Guests

Measuring Travel Distances



- Patron's travel distance to stadium trends with Voronoi area
- Additional considerations for outliers
 - Population density
 - Marketing strategies
 - Competing minor league teams

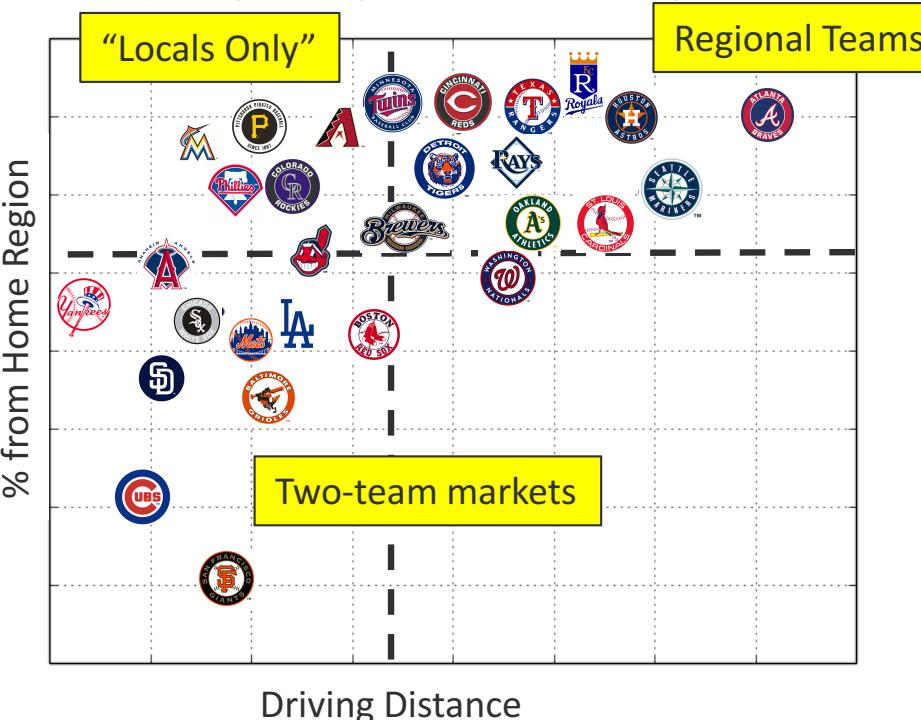


MLB Stadium Guests

Measuring Travel Distances



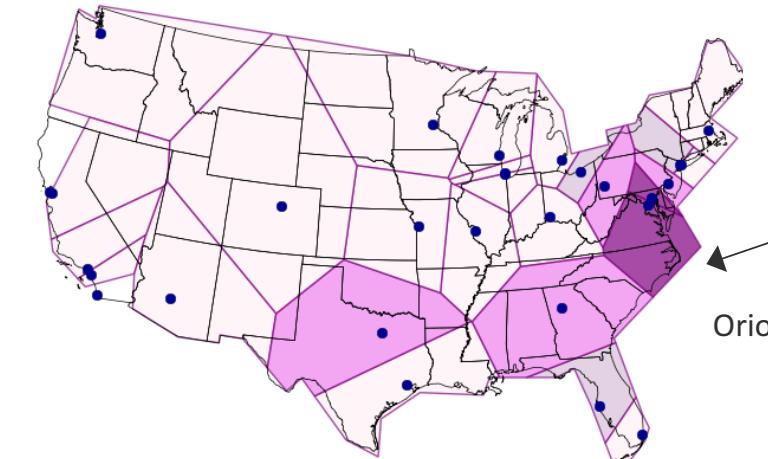
Comparing Driving Distance / Drawing Home Fans



- Patron's travel distance to stadium trends with Voronoi area
- Additional considerations for outliers
 - Population density
 - Marketing strategies
 - Competing minor league teams

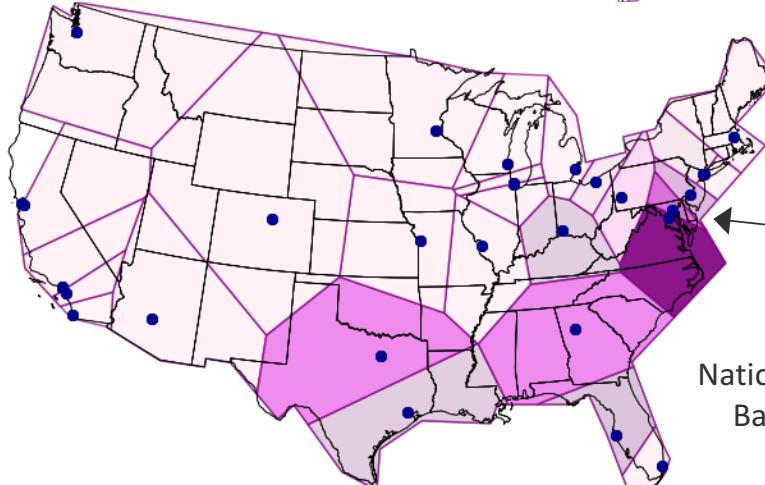
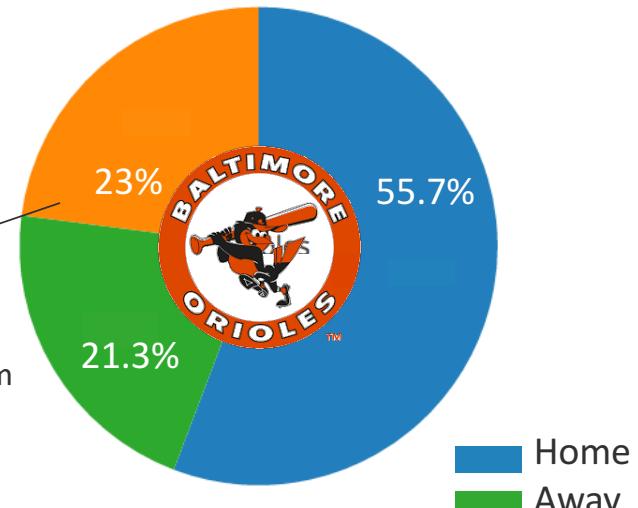
MLB Stadium Guests

Identifying Asymmetric Patterns

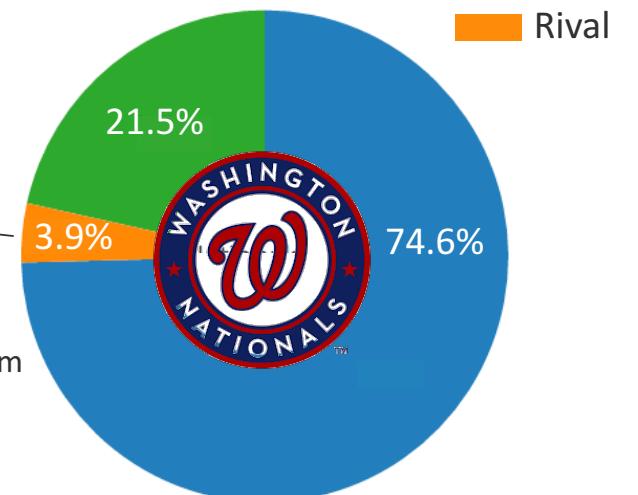


Orioles guests from DC region

Home Region Breakout for Team Patrons



Nationals guests from Baltimore region



Baltimore Orioles and Washington Nationals draw differently from each other's home regions

Outline

- Introduction
- SurveyMini Overview
- Geospatial Research Examples
 - MLB Stadium Visits
 - Fast Food Store Locations
- Summary



Taco Bell USA

Overview

- Owned and operated by Yum! Brands
- Founded in 1962 by Glen Bell
- Over 6,000 continental US locations
- Taco Bell sells over 5 million tacos per day
- First tacos – 19 cents (!)



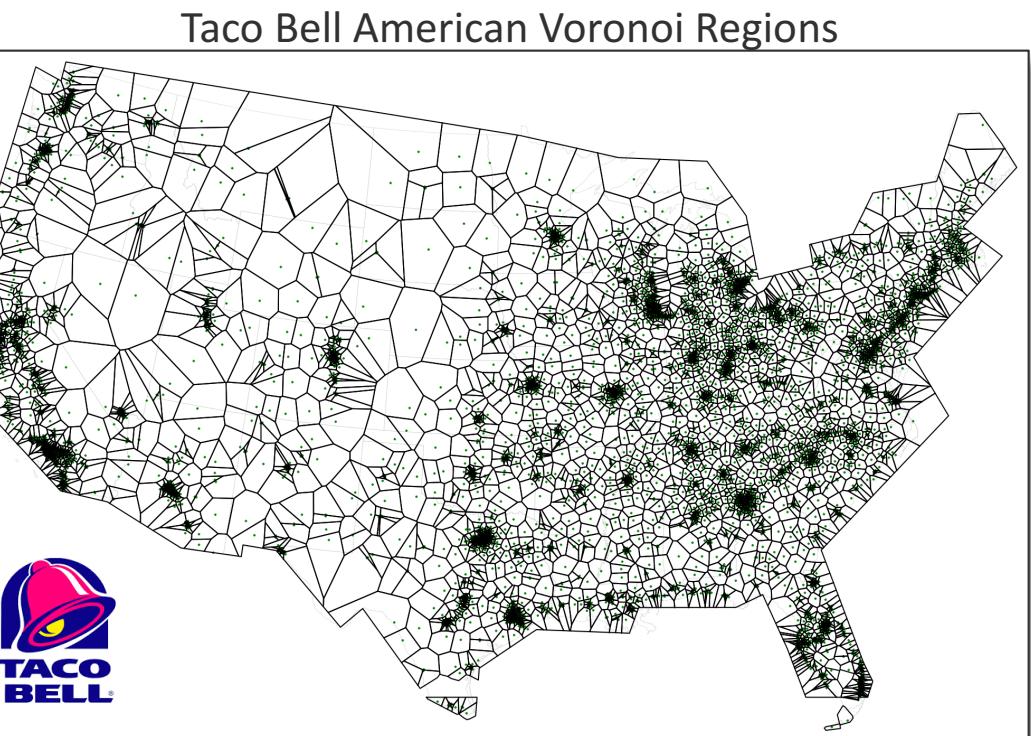
1960's



1990's



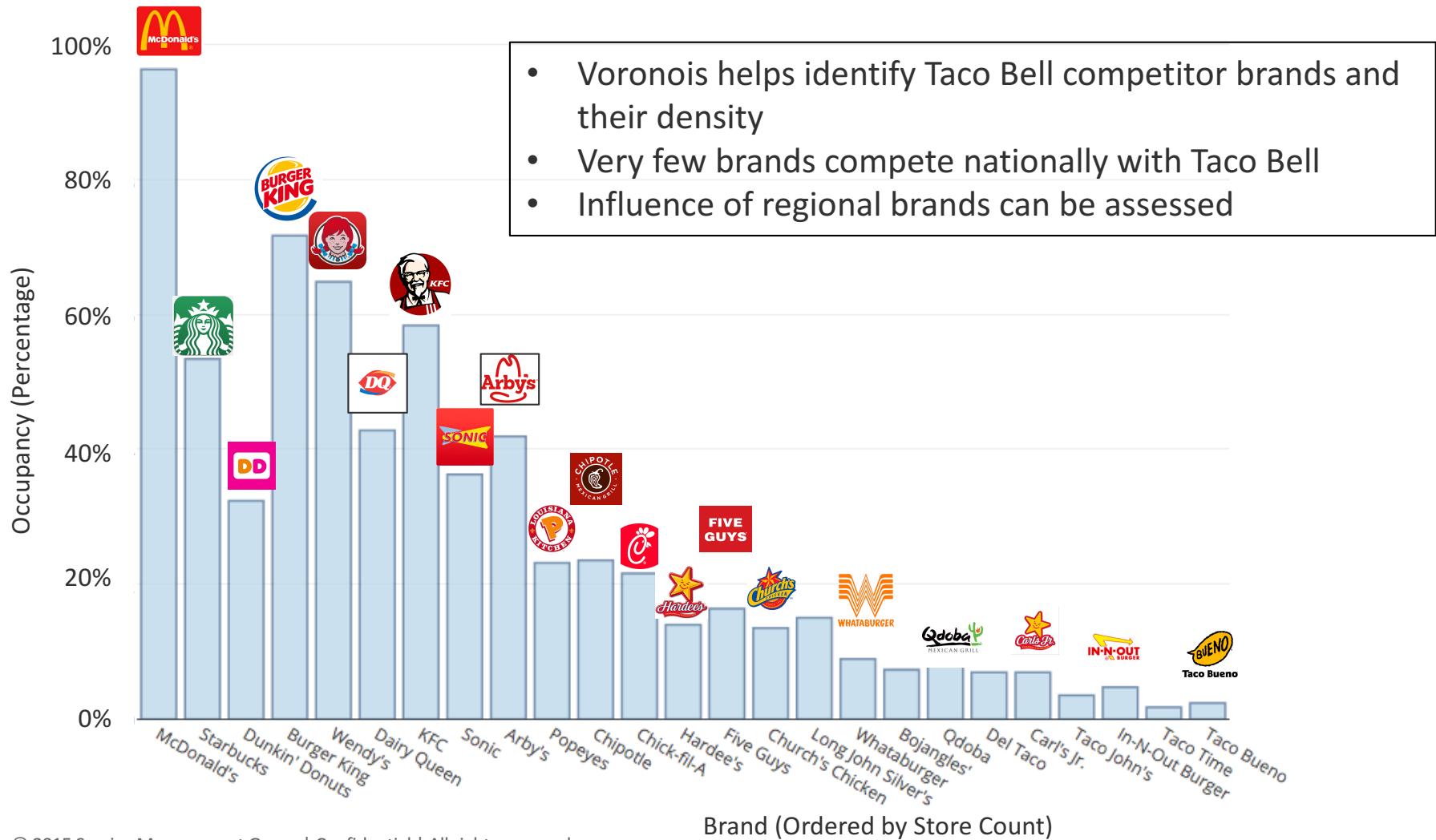
2010's



Taco Bell USA

Brand Competitor Density

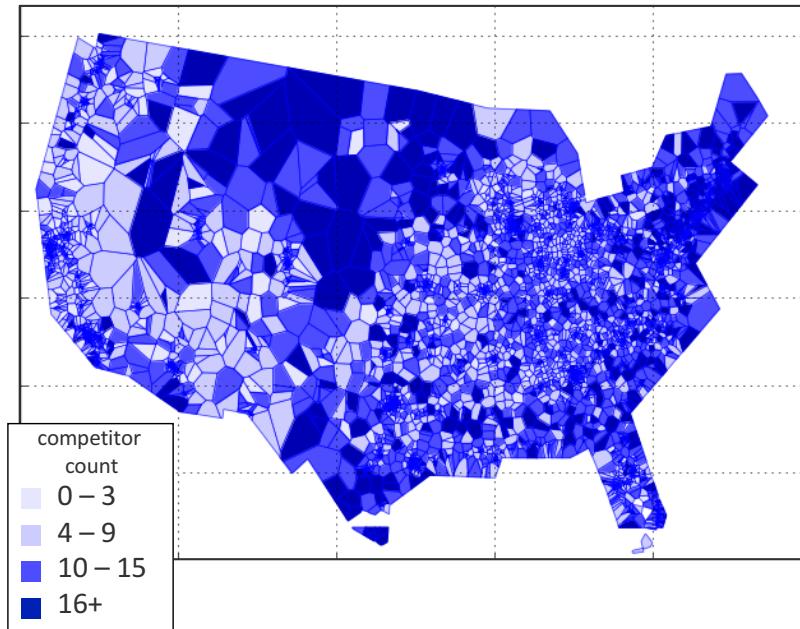
Occupation Percentage of Taco Bell Voronoi Regions



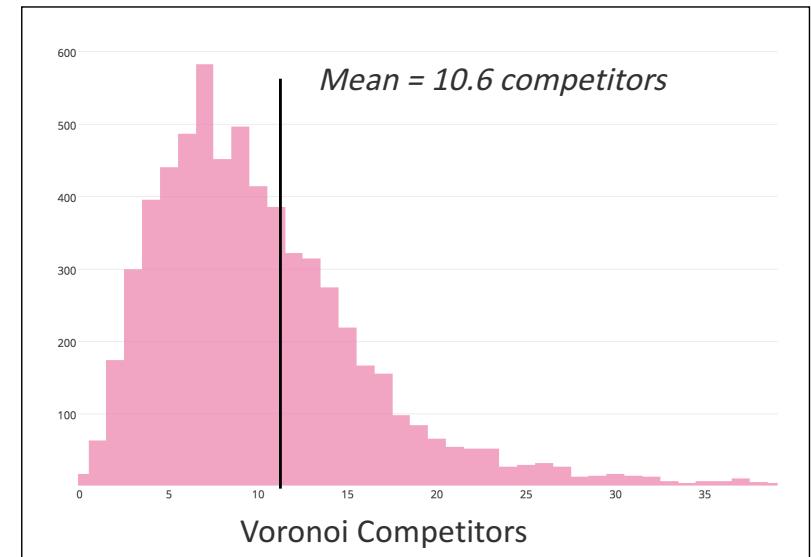
Taco Bell USA

Regional Analysis Based on Competitor Density

Taco Bell Competitor Count Density



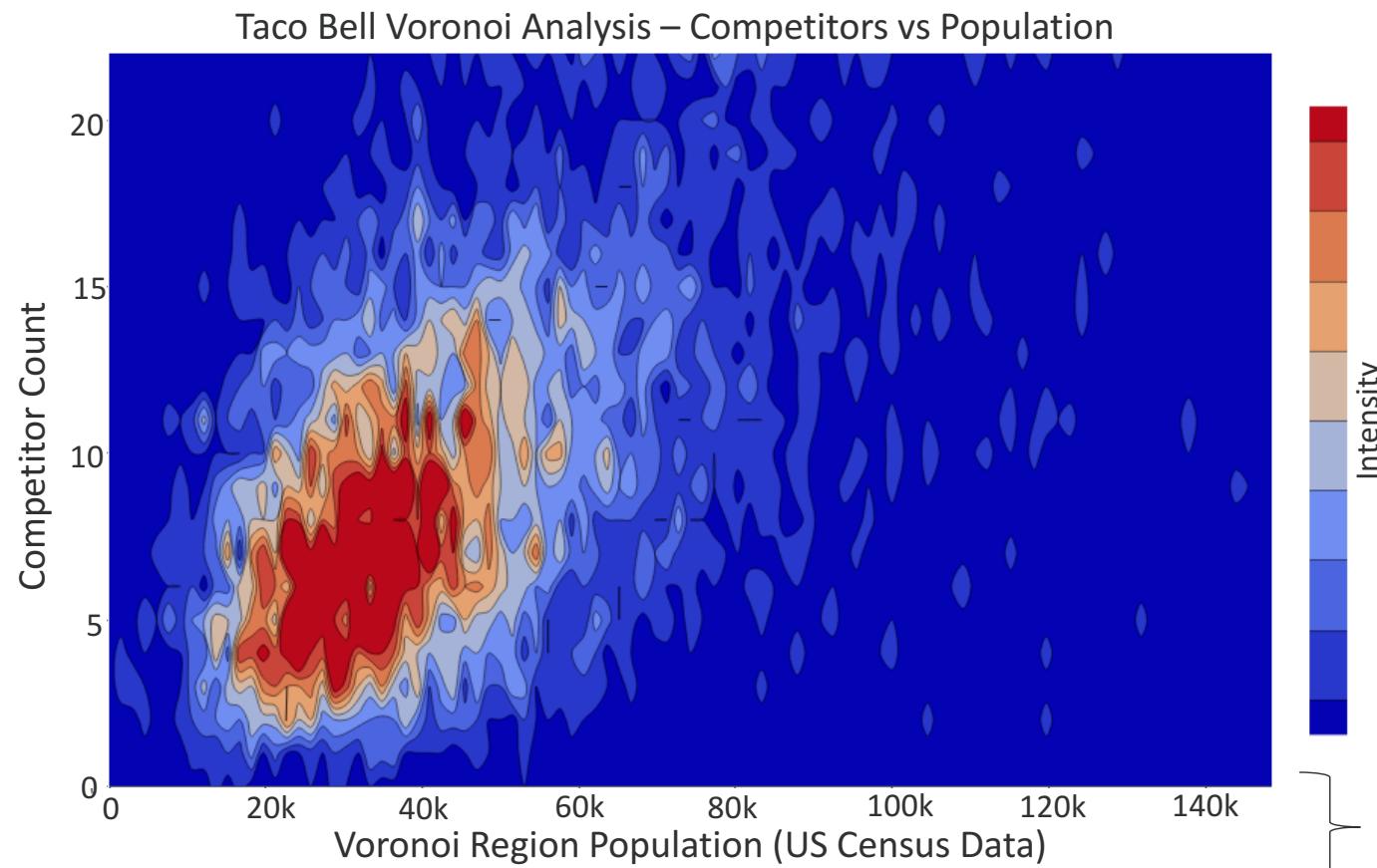
Histogram of Competitors in Taco Bell Voronoi



While competitor density is an important factor influencing a client's regional growth strategy, there are others we can consider

Taco Bell USA

Competitors vs Population



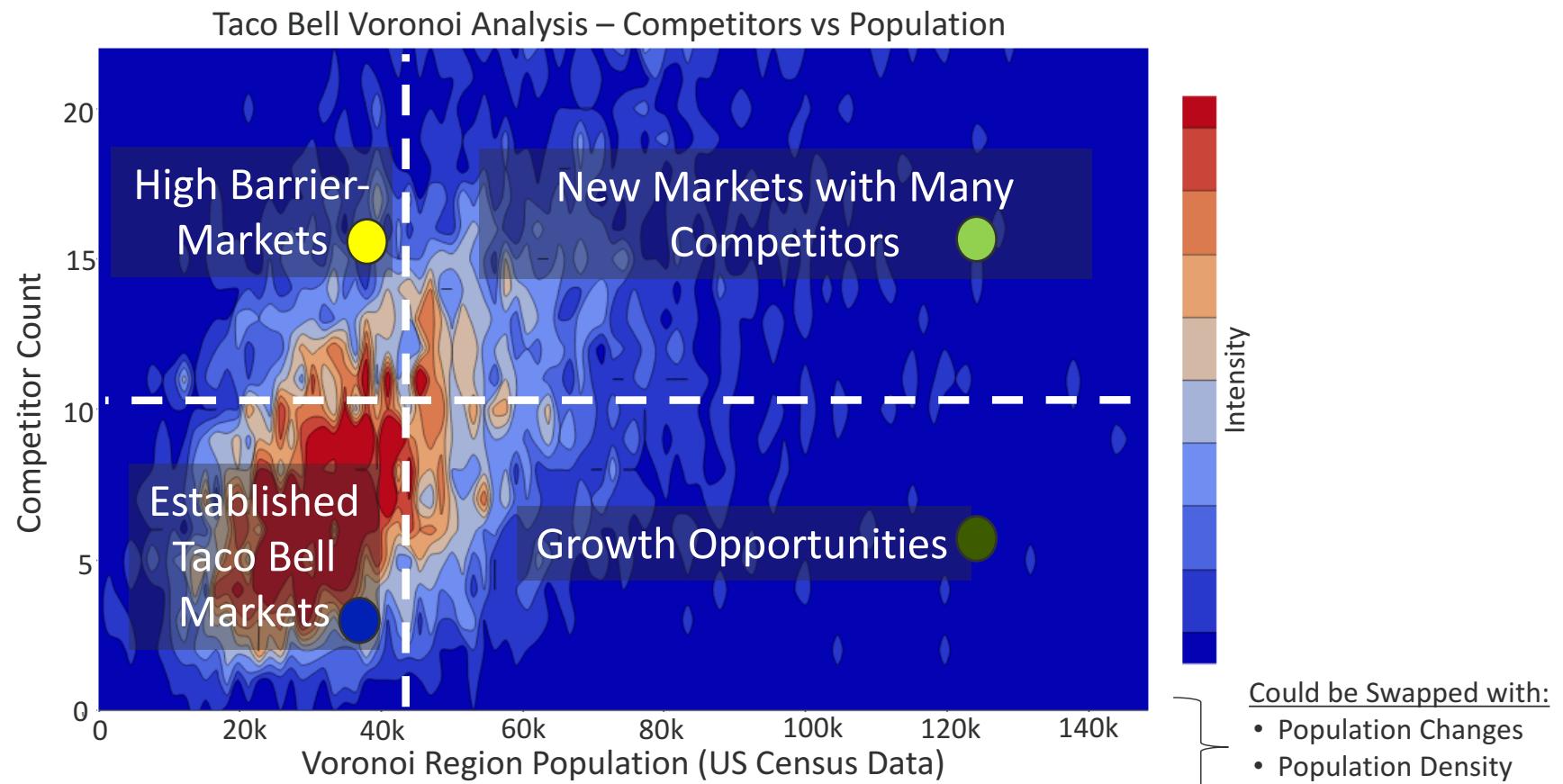
Could be Swapped with:

- Population Changes
- Population Density

- US census data used to estimate population of Taco Bell Voronoi regions
- Data can be used to identify candidate regions/locations for growth

Taco Bell USA

Competitors vs Population

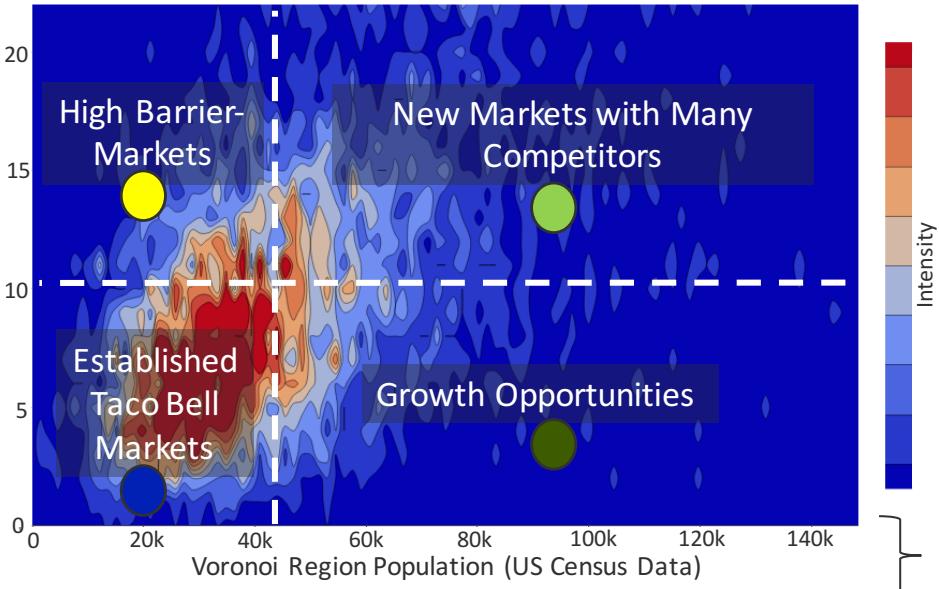


- US census data used to estimate population of Taco Bell Voronoi regions
- Data can be used to identify candidate regions/locations for growth

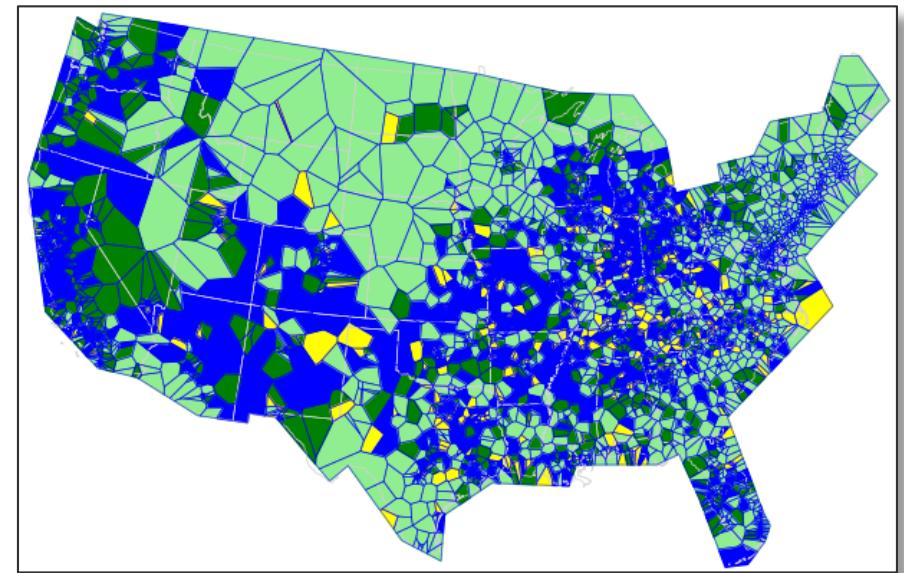
Taco Bell USA

Competitors vs Population

Taco Bell Voronoi Analysis – Competitors vs Pop



Mapped Competitor / Population Score

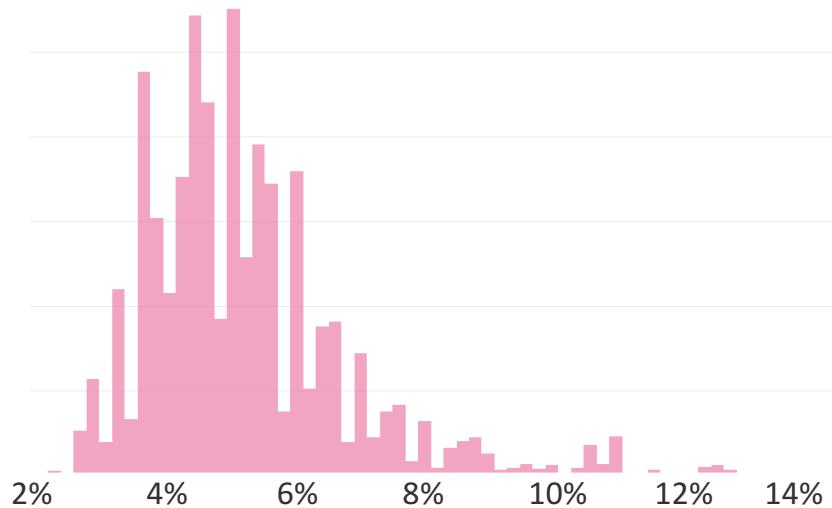


- US census data used to estimate population of Taco Bell Voronoi regions
- Data can be used to identify candidate regions/locations for growth

Taco Bell USA

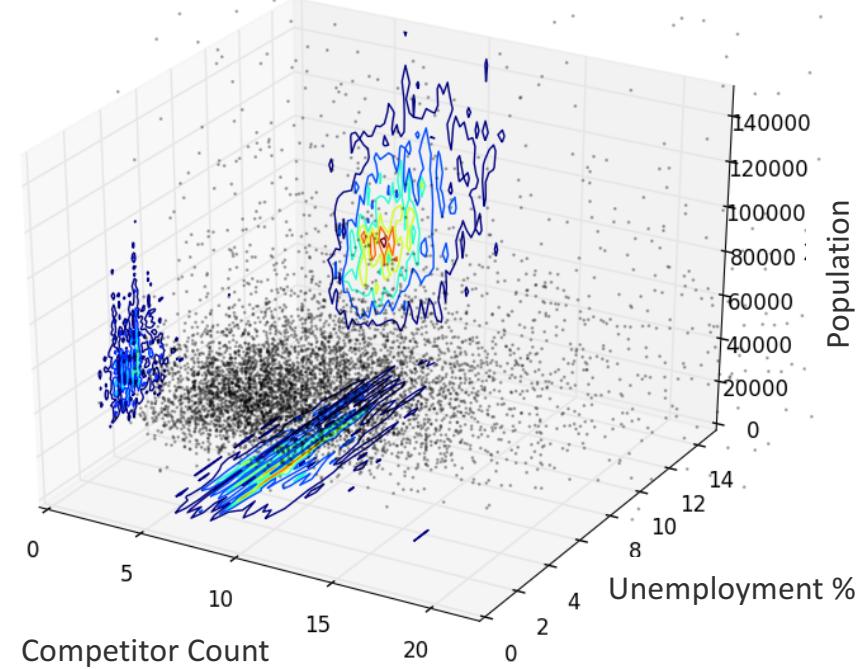
Incorporating Unemployment Data

Unemployment Rate in Taco Bell Counties (Jan 2017)



- Unemployment data published monthly by API by Bureau of Labor Statistics
- Low unemployment
 - Many potential customers
 - Challenging hiring environment
- High unemployment
 - Less disposable income
 - Hiring opportunities

Taco Bell Store Analysis
Competitors x Population x Unemployment



Combining unemployment levels with population data and competitor can lead to new business opportunities

Outline

- Introduction
- SurveyMini Overview
- Geospatial Research Examples
 - MLB Stadium Visits
 - Fast Food Store Locations
- • Summary

Summary

The article discusses the state of location intelligence analytics and BI in 2017. It highlights research & development (R&D) and executive management as the most important functions. Energy and transportation industries place the most significant emphasis on location intelligence analytics and BI. 63% of respondents perceive location intelligence as being critically (20%) and very important (43%) to their ongoing business operations. Executive management is most interested in cloud-based deployments of location intelligence analytics, and BI gives cloud its highest criticality ranking.

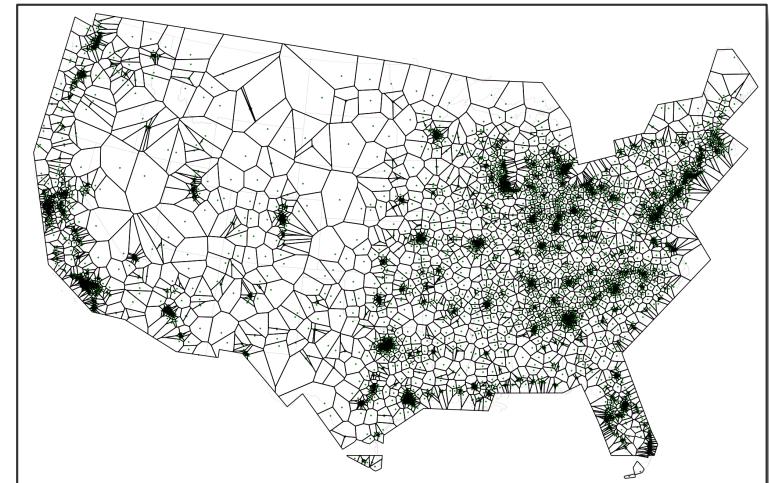
Key points from the article:

- Research & Development (R&D) and executive management are two functions that place the greatest importance on location intelligence analytics and business intelligence (BI).
- Energy and transportation industries place the most significant emphasis on location intelligence analytics and BI today.
- 63% of respondents perceive location intelligence as being critically (20%) and very important (43%) to their ongoing business operations.
- Executive management is most interested in cloud-based deployments of location intelligence analytics, and BI gives cloud its highest criticality ranking.

These and many other insights are from an excellent research study recently published by Dresner Advisory Services titled *2017 Location Intelligence Market Study Report* (92 pp., PDF, client access). Dresner Advisory Services defines location intelligence as a form of business intelligence.

- Location intelligence is a rapidly growing data science application
 - Traditional physical science analytic techniques are being applied to soft-science problems
- Service providers want to understand behavioral attributes of their customers!

- Voronoi partitioning creates “natural” geospatial boundaries for customer regions
- Geospatial technologies (databases, file formats, open-source data science tools) are a “perfect storm” of opportunity for data science!
- Customer journeys will continue to evolve to combine offline and online behavior





Any Questions?

Thank you!

Dan Finkel

@danfinkel

Service Management Group

dfinkel@smg.com

<http://github.com/danfinkel/python/>



Back-ups

Word clouds of open ends from locals vs out-of-towners

KC Grocery overview

List of brands

Voronoi of region

1st show that home customer % is a function of competition graph
(turn the scatter plot into a bar plot histogram)

Next Show survey data similarities / differences

Highlight Price Chopper

Lastly show convenience stats

- Which brands get most convenience selections
- Does convenience choice bias towards home region shoppers?
 - Conclusion → Convenience == Grocery shopping close to home

DSG / Sports Authority MA example

Show which regions have golf courses

Light up regions with ice rinks

Maybe another example

Make national voronoi for AMC and Regal (any of our movie theatres)

Measure which restaurant brands have the best coverage

Chili's, Applebee's, TGIF, California Pizza Kitchen



Summary

Thanks!

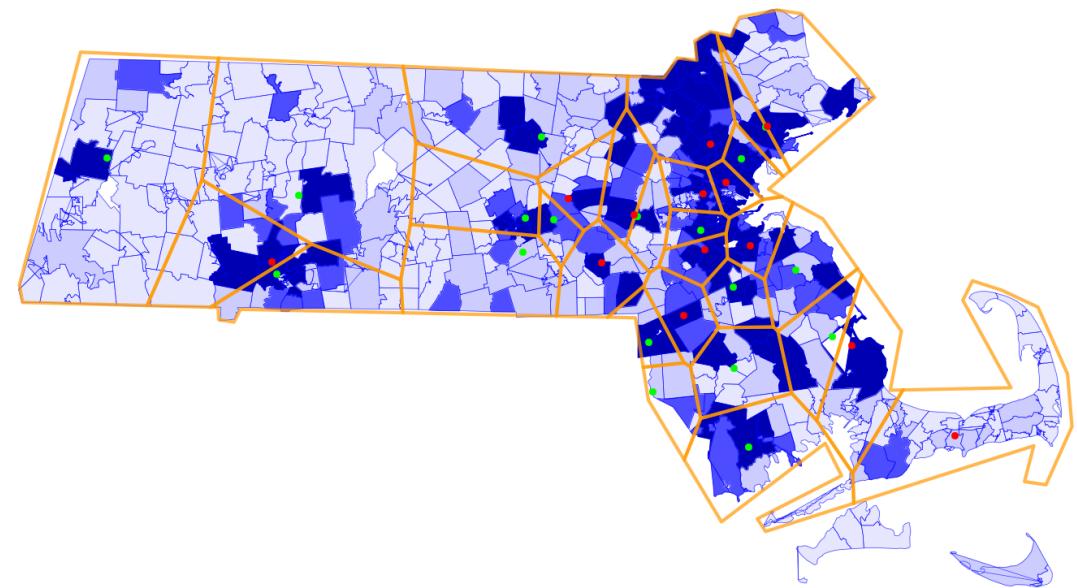
Dan Finkel

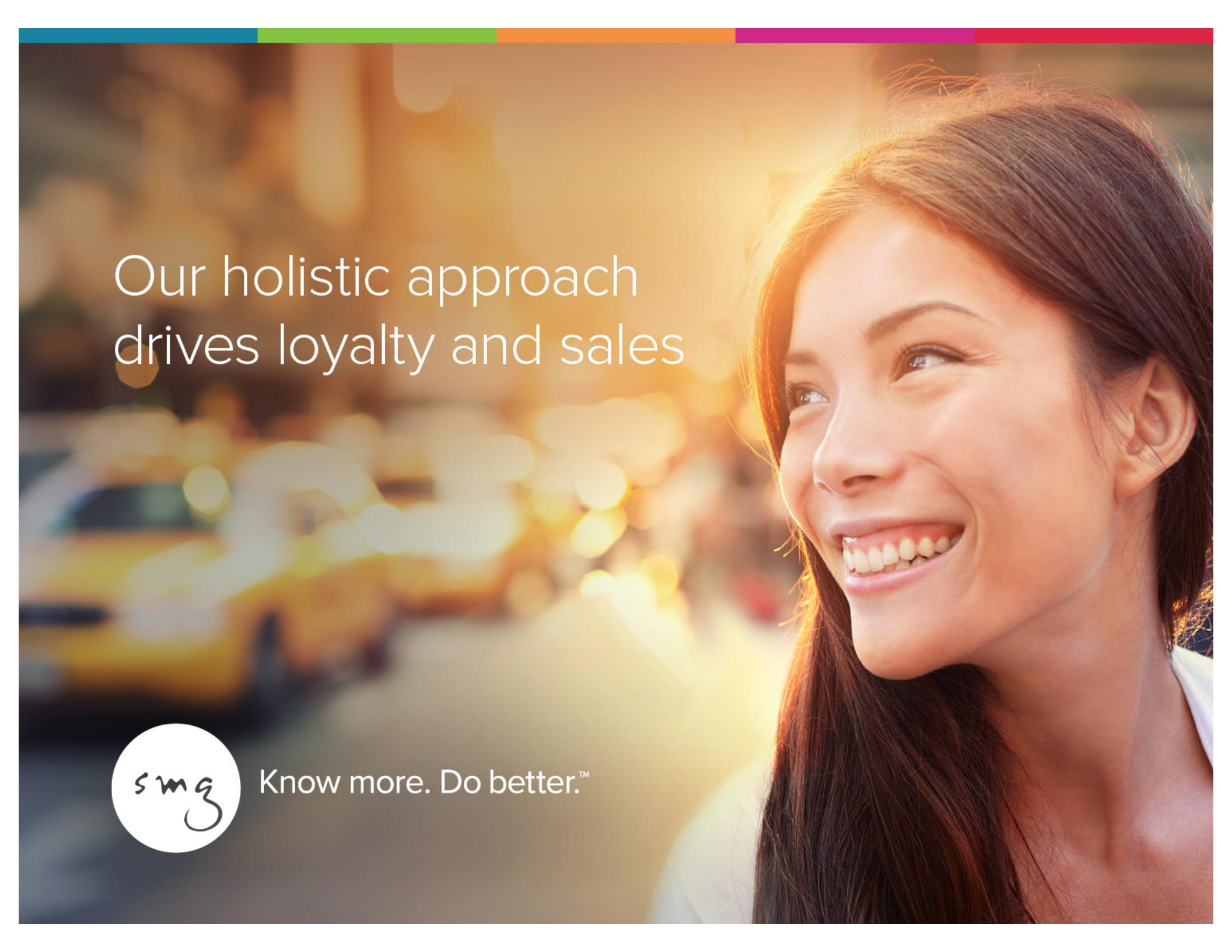
dfinkel@smg.com

@danfinkel

<http://github.com/danfinkel/python>

• asdf





Our holistic approach
drives loyalty and sales



Know more. Do better.[™]



Office DEPOT OfficeMax

ZALES
THE DIAMOND STORE®

Marshalls

jenny CRAIG



HELZBERG
DIAMONDS.



Be Happy
HomeGoods



REGAL
CINEMAS



Michaels
Where Creativity Happens®



T.J.maxx®

aaronbrothers
ART & FRAMING

Jack Wills



Chick-fil-A

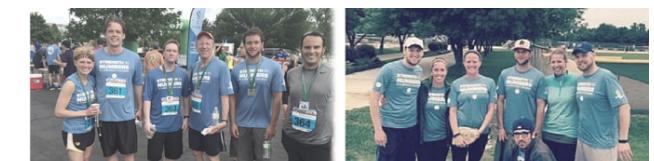
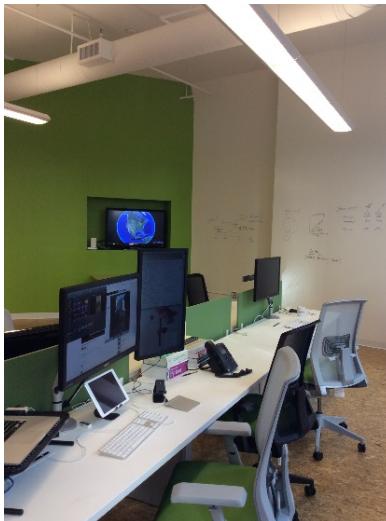


The Cheesecake Factory



OUTBACK STEAKHOUSE®

Culture is important to us



DIVISION E CHAMPIONS

#StrengthInNumbers



Technology as a differentiator

Mobile technology is changing the world

Military Intelligence



"Network"

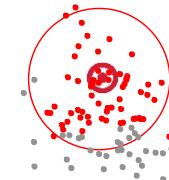
Advanced Mobile Research

Brand Geek Insights



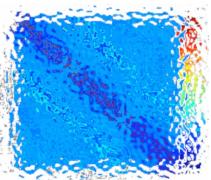
Smart collection Interactive discovery

PinPoint Technology

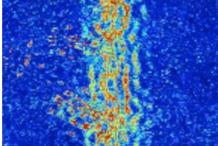


Accurate locationing Advanced processing

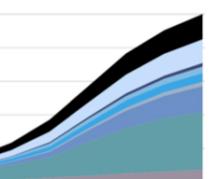
Future Health



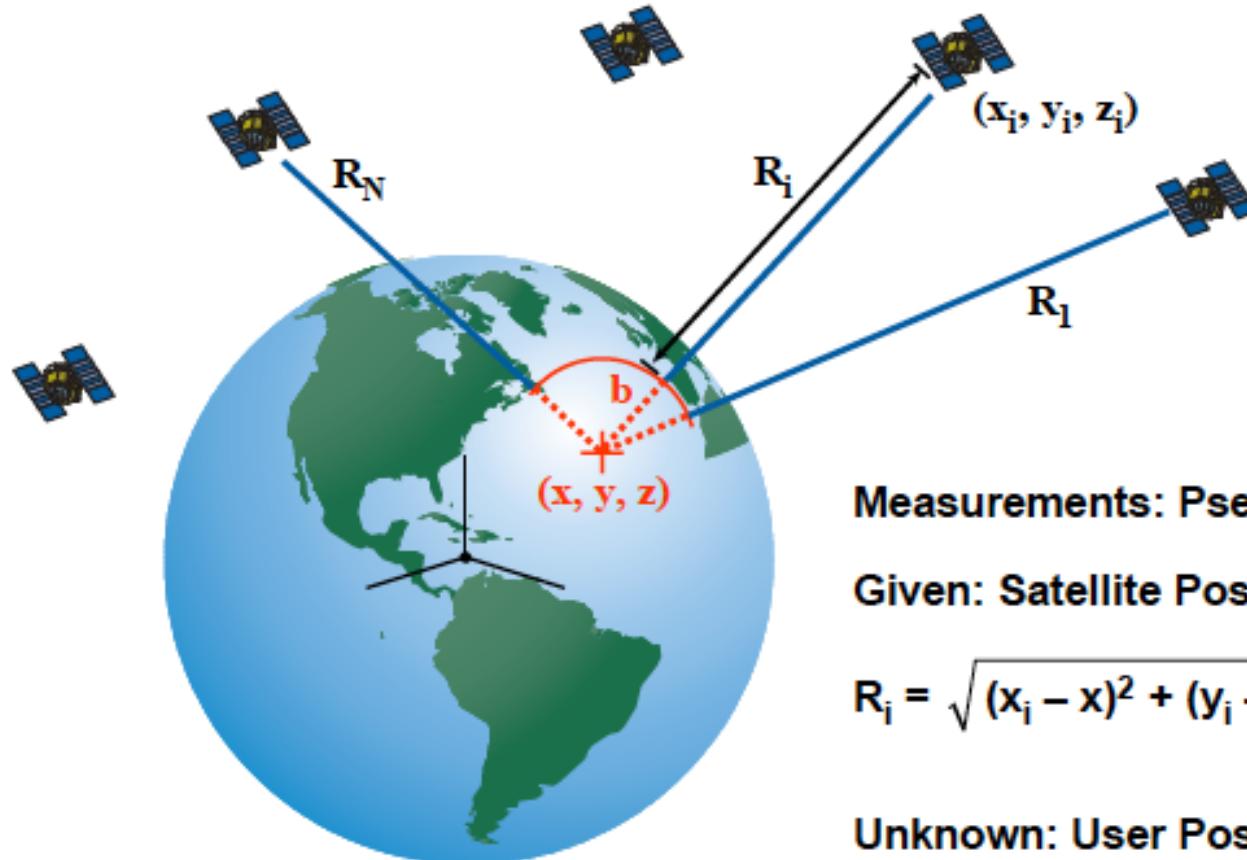
National Defense



Financial Services



Fundamentals of modern positioning, navigation, and timing



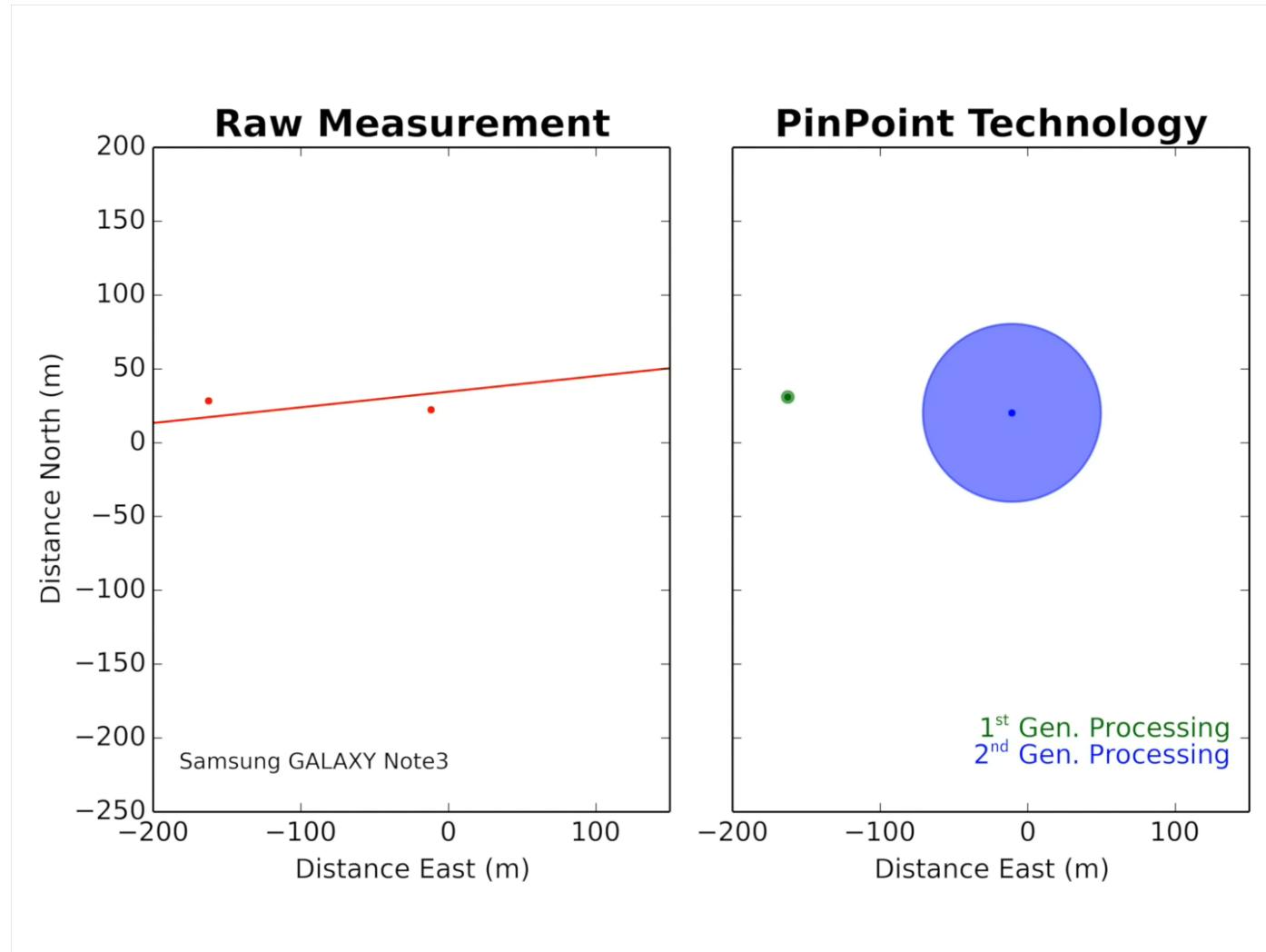
Measurements: Pseudoranges $\{R_i\}$

Given: Satellite Positions $\{(x_i, y_i, z_i)\}$

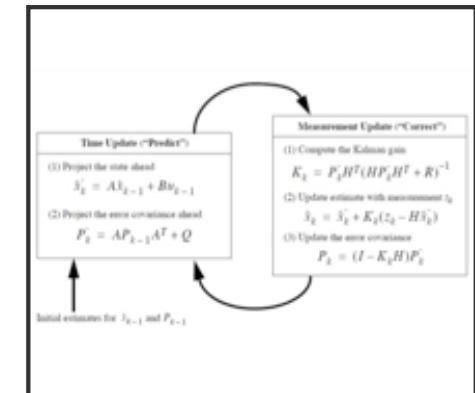
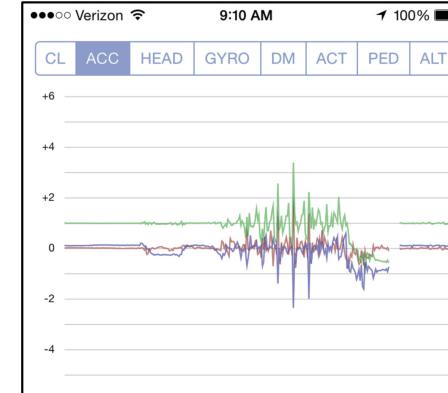
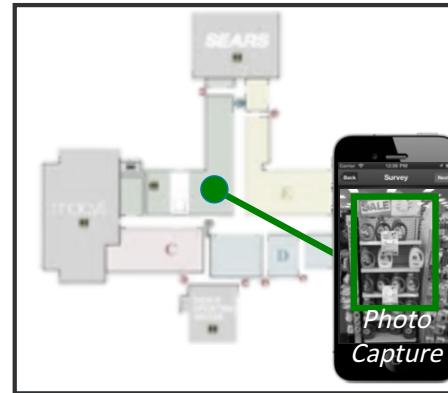
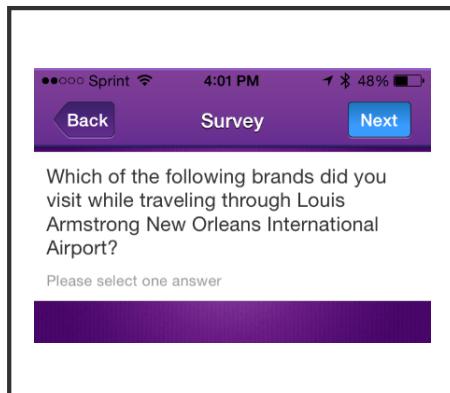
$$R_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - b, \quad i = 1, 2, \dots, N$$

Unknown: User Position (x, y, z)
Receiver Clock Bias b

PinPoint Technology converts raw measurements to accurate position estimates

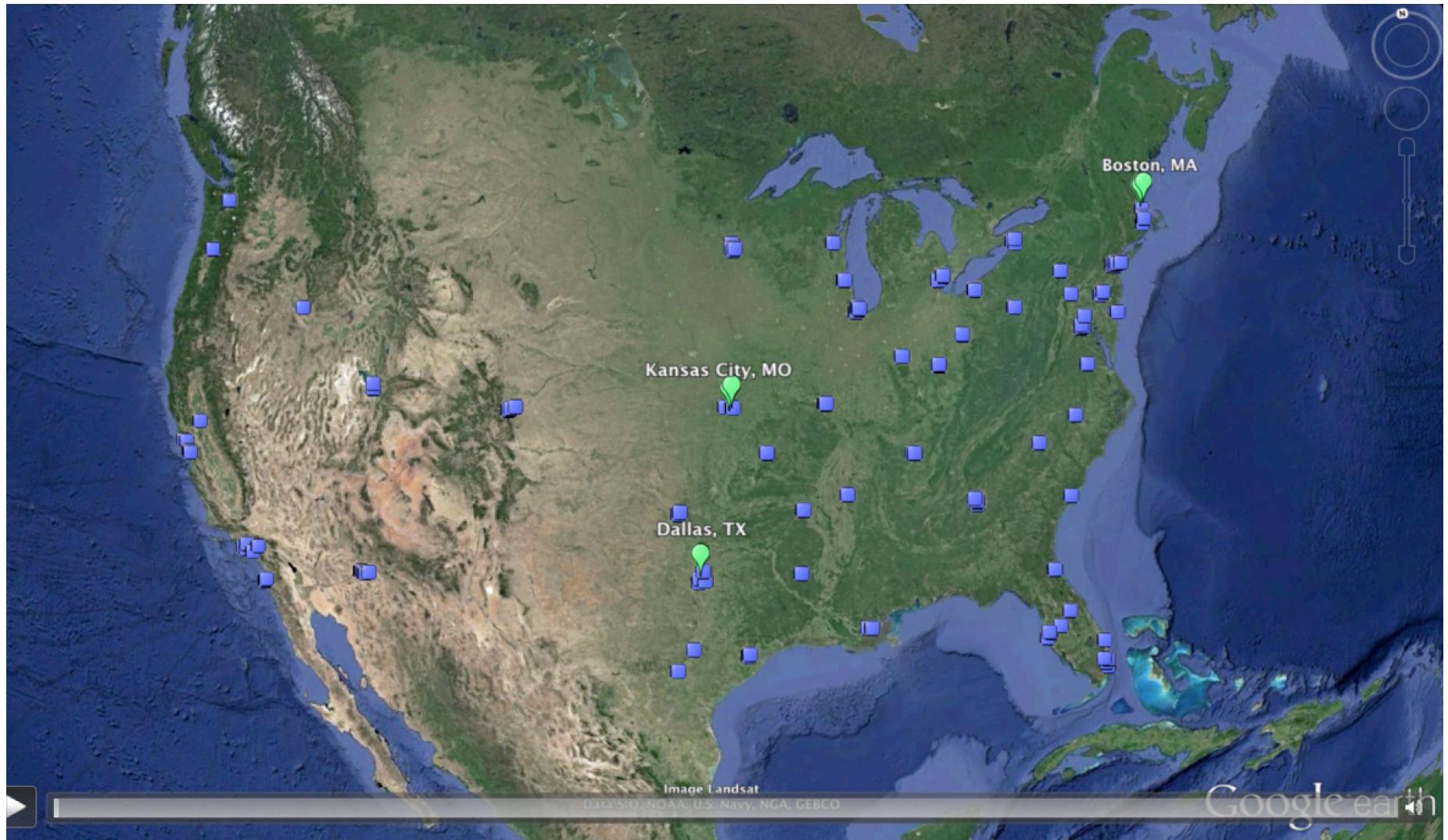


A suite of four technology components are in development

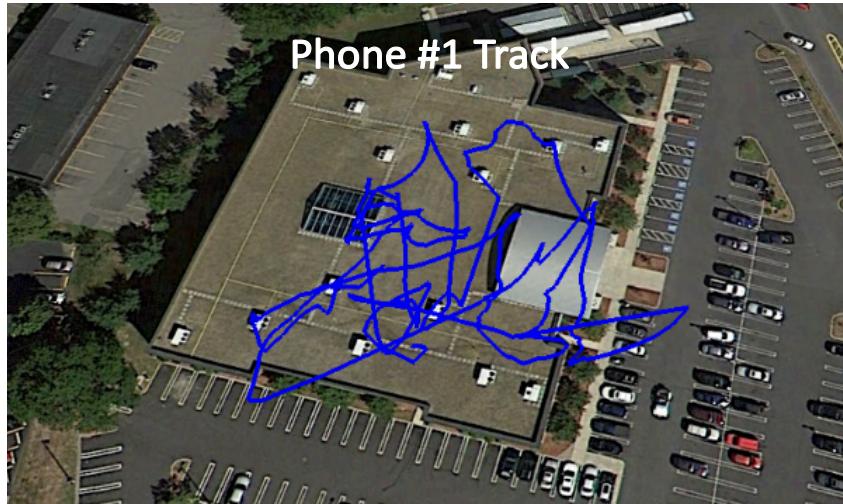


- Query nearby places
- Digital library
- Photo refinement
- Core motion
- Auxiliary sensors
- Enhanced algorithms

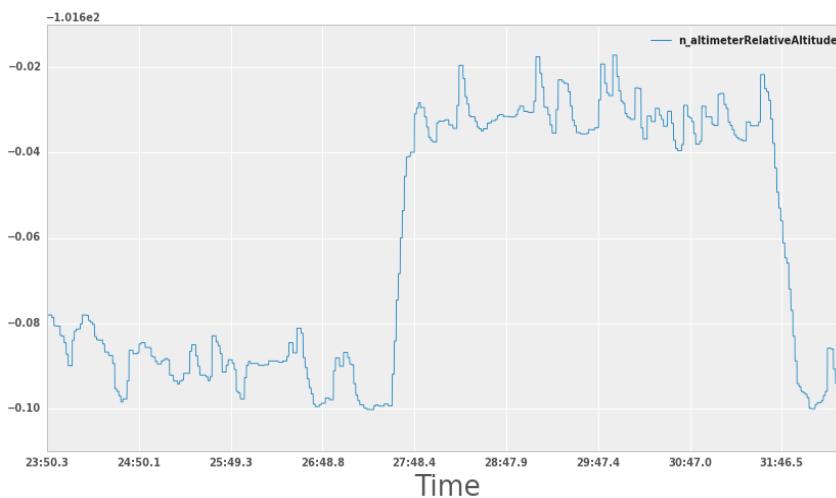
We have R&D test activities in progress



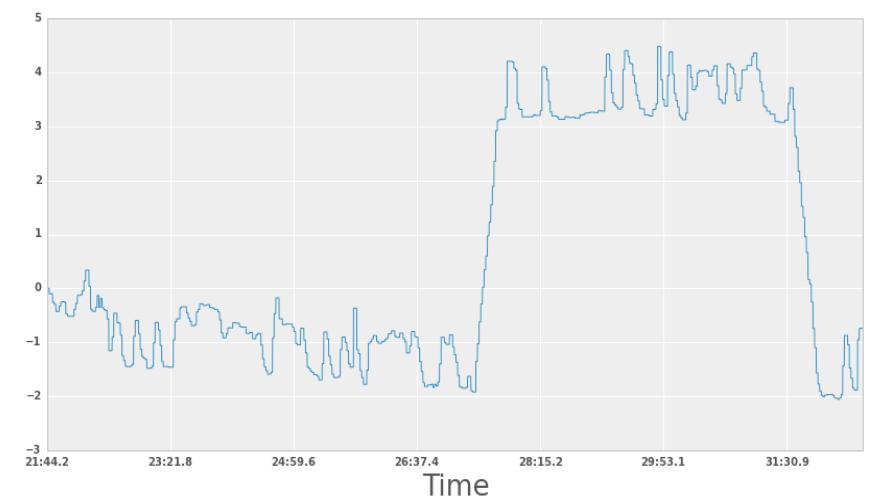
Auxiliary sensor data evaluation results



Altimeter: Relative altitude



Altimeter: Relative altitude



Shopper's journey





Together we can
Know more. Do better.

Let's get started.



service
management
group®

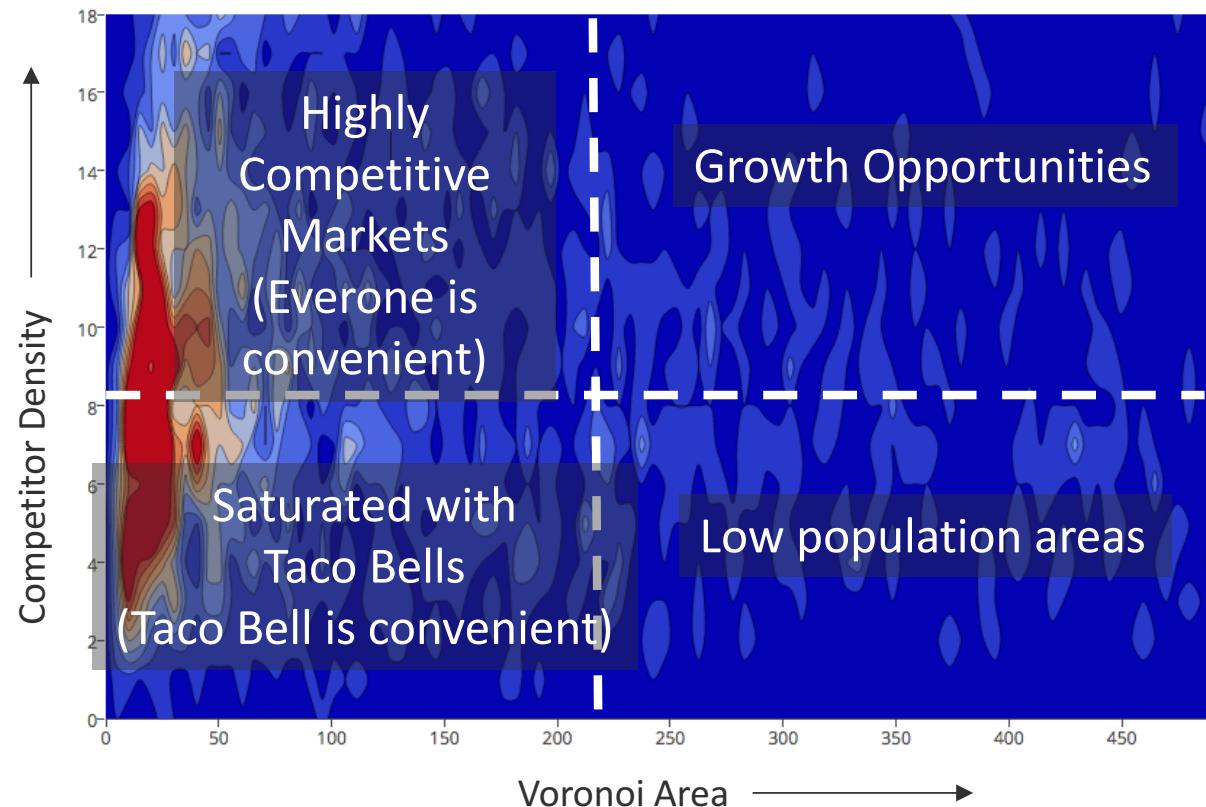


Our model

SMG was founded on the principles of *The Service Profit Chain*, which demonstrates highly engaged employees create highly satisfied customers. Highly satisfied customers are more loyal, which means they're more frequent, they spend more per visit, and they recommend you more often.

Taco Bell USA

Labeling Stores



asdf