Quantifying Competitive Retail Spatiotemporal Behavior

Ken Cavagnolo

Disclosure

None of the data, analysis, or results presented today contain proprietary, privileged, or confidential information.

Where proprietary, privileged, or confidential information would be useful to the conversation, generalized information obvious to the public has been substituted.

About Me

"Data Scientist"

Physics

Product

Management

Operations

Logistics

Consulting

Analytics

Applied Data Science Case Study

My humble opinions:

| Data science is science: data→ hypothesis→ testing→ conclusions→ repeat.

| Data science is expansive curiosity and exploration.

| Data science is well-thought out code engineers can productionalize easily.

| Data science chases problems/questions a business doesn't see or want.

| Data science is valuable -- being interesting isn't enough.

Grudge Buying

Insurance | Maintenance | **Essentials** | Services...

Human

Food
Clothing
Water
Shelter

Human

Food
Clothing
Water
Shelter

Economy

Energy
Other
Debatable
Things

Petroleum

| Gasoline and diesel
| Fungible commodities
| 130k retail locations in US
| $1Tn industry all-in
| Low margin products
| Declining need 2%-5% yr^{-1}

Petroleum

Gasoline and diesel
Fungible commodities
130k retail locations in US
$1Tn industry all-in
Low margin products
Declining need 2%-5% yr^{-1}
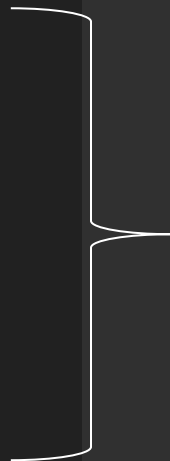
Hyper-competitive market

"We need competitor intelligence."

What is competition?

Competition

| Regulated
| Known players
| Rational players
| Active engagement
| Maximize outcomes
| Have needed resources

Competition

Regulated
Known players
Rational players
Active engagement
Maximize outcomes
Have needed resources
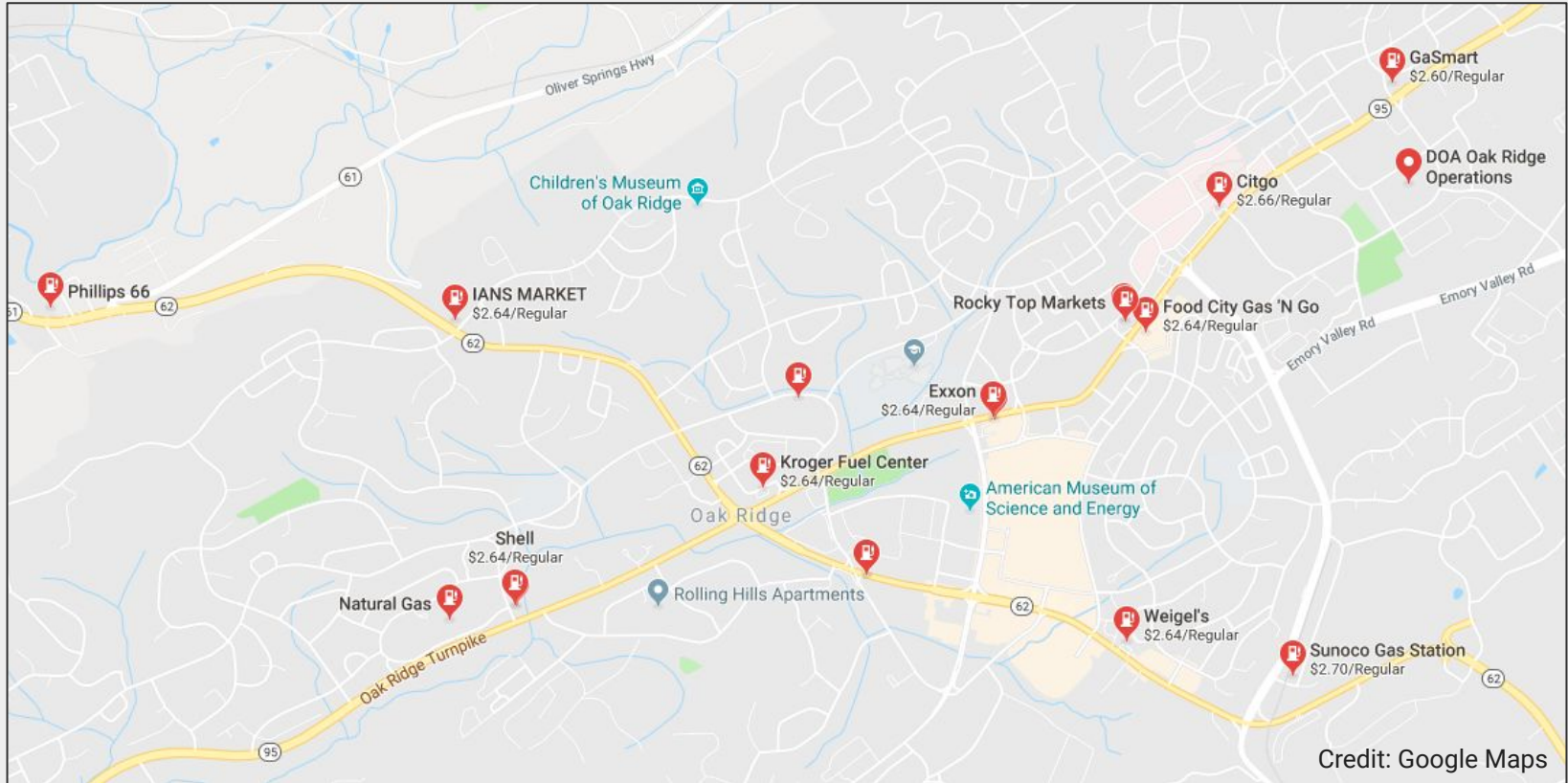
Game
Theory

Key Game Theory Question:

Are players strategically engaged?

Is this…

…competition?

Typical competitive landscape

Credit: Google Maps

Is this competition?

Prior selection for driving distance

Find Patterns

| Assume if there is competition, it's primarily on price
| Price competition suggests structured price changes
| Structured price changes should be discoverable
| Compare the time series:

- Fourier Transforms (FFT)
- Piecewise Aggregate Approximation (PAA)
- Symbolic Aggregate Approximation (SAX)
- Dynamic Time Warping (DTW)
- Many others

Find Patterns

| Assume if there is competition, it's primarily on price
| Price competition suggests structured price changes
| Structured price changes should be discoverable
| Compare the time series:

- Fourier Transforms (FFT)
- Piecewise Aggregate Approximation (PAA)
- Symbolic Aggregate Approximation (SAX)
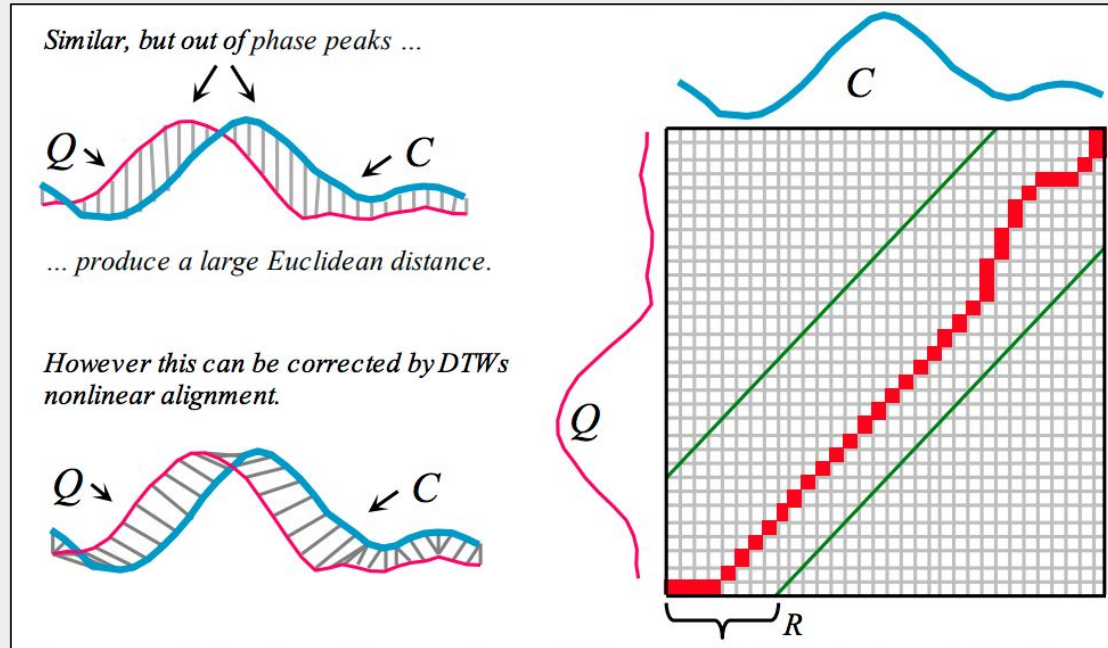- **Dynamic Time Warping (DTW)**
- Many others

Why DTW?

Pros

| Captures lagged behavior

| Insensitive to magnitudes

| Subsequence matching

| Elegant implementations

Cons

| Not a metric!

| Dissimilarity is relative and arbitrary
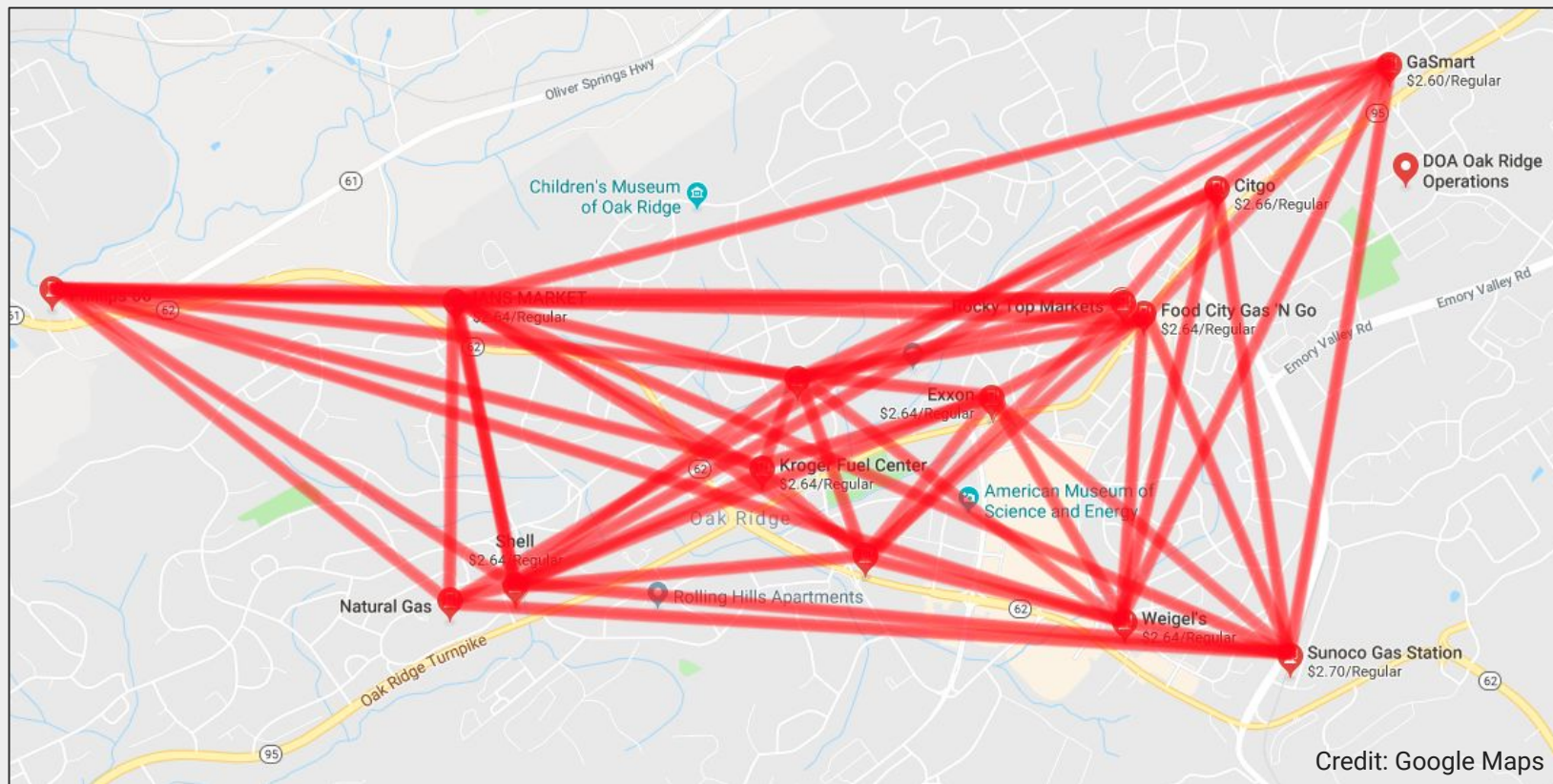
| Noise will hurt measure

Find the Warping Path

Searching and Mining Trillions of Time Series Subsequences under Dynamic Time Warping
Rakthanmanon, Campana, Mueen, Batista, Westover, Zhu, Zakaria, Keogh

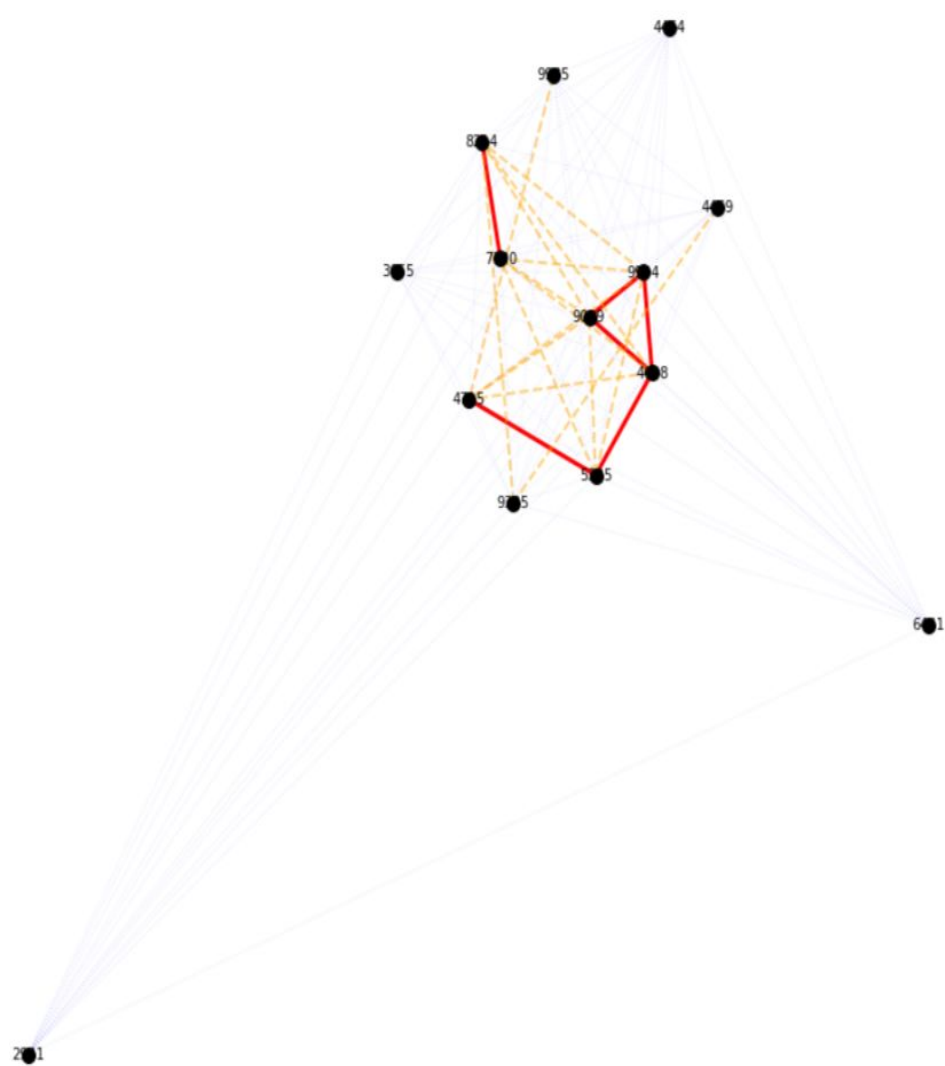# Typical Comparisons

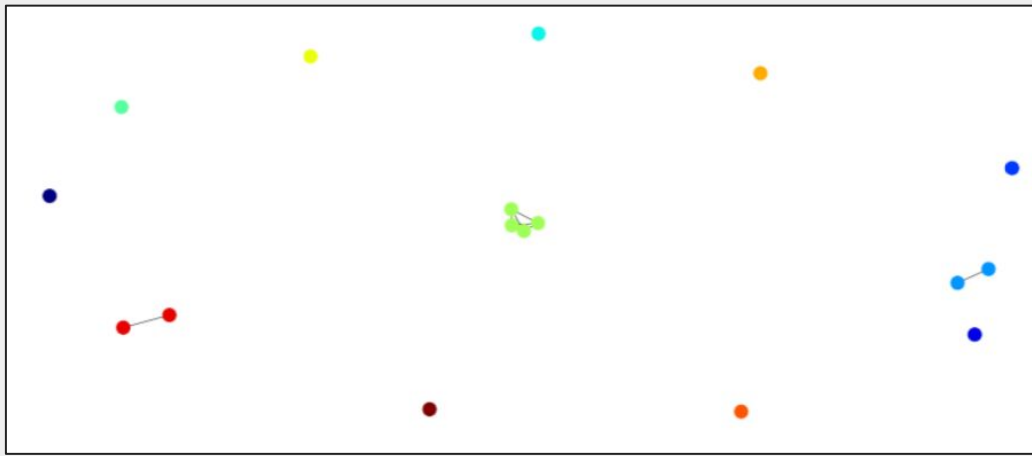# A Graph Problem

Credit: Google Maps

Pricing Graph

$1/(dtw)^2$ edge weights

A pattern is emerging

Is there logical structure?
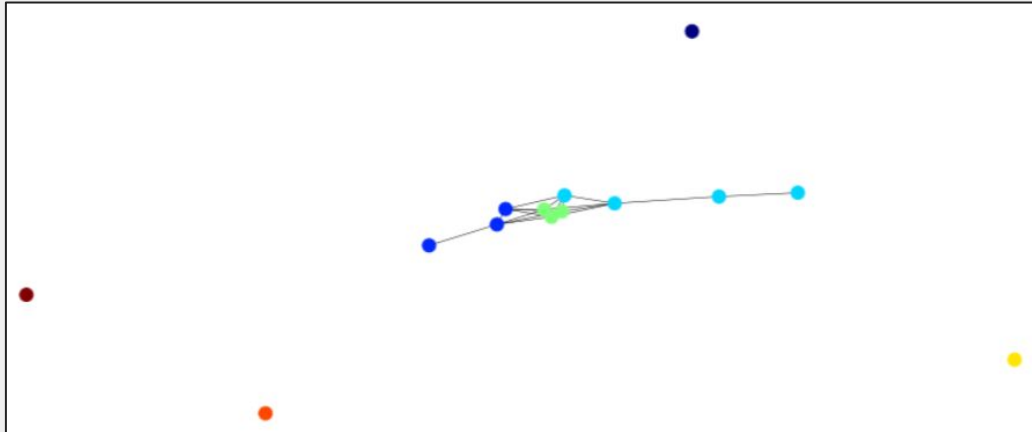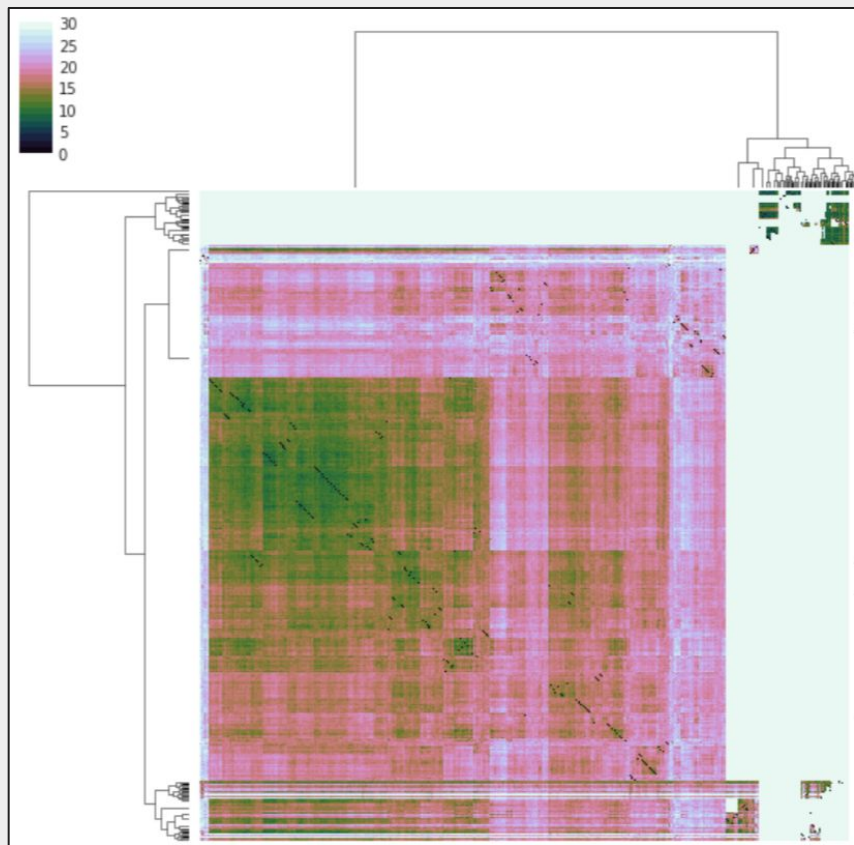
Find communities

Communities

Fuel sellers clump together

Orbiting entities easily explained

Tight knit entities are mimics

Clear 'leaders' and 'followers'

Full Market View

All-v-all DTW comparison

DTW > 15 is dissimilar for our purposes

Market appears oddly uniform

25%

of the entire market is similar in pricing behavior.

No collusion.

"We need competitor intelligence."

Are we competing?

Supergames Literature

| "...tacit collusion" due to lack of pricing sophistication [1]

| "...price leadership... signals volatile costs" to market [2]

| Equilibrium is a desirable state, arguably most desirable [3]

| Communities dictate the games played and by whom

| Easy to take prices **UP**, hurts to take them **DOWN**

1: Borenstein & Shepard 1996, Dynamic Pricing in Retail Gasoline Markets
2: Russell et al. 2010, Pricing in Retail Gasoline Markets
3: Nash 1950, Equilibrium points in n-person games

Game Types

| The game of **UP** prices go... up

| The game of **DOWN** prices go... down

| Subject matter experts selected one week as maximum game length

| A game begins when local competitive market departs from equilibrium

| DTW subsequence matches highlight 'motifs' or flavors of games:

  ○ Expired -- game time ends without returning to equilibrium
  ○ All-in -- all non-initiator competitors return to a state of equilibrium
  ○ Abandoned -- the game initiator returns the system to equilibrium

"Would you like to play a game?"

UP

DOWN
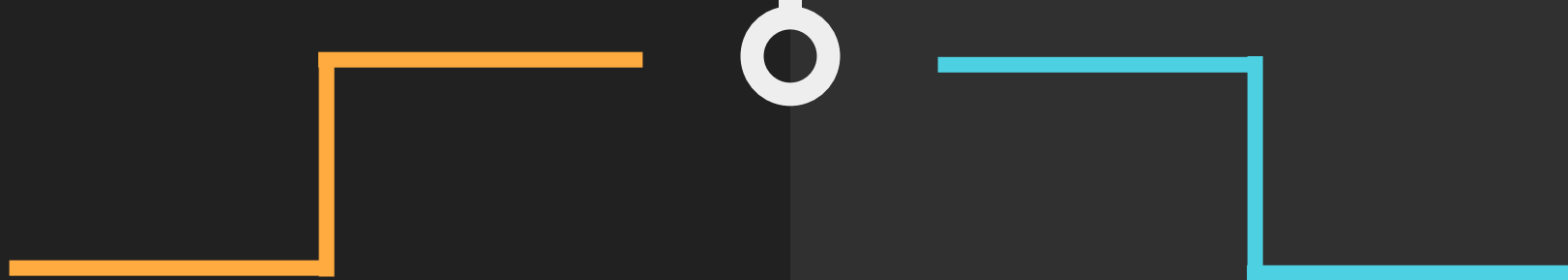
START

Goals:
- Marketshare
- Gross Profit

All-in

Abandoned

Expired

Typical Game Playing

2017: 28k Games

```
Game: down          Outcome: abandon          906     3.3% (loss)
Game: down          Outcome: all-in          7150    25.8% (loss)
Game: down          Outcome: expired         5253    18.9% (win)

Game: up            Outcome: abandon         1837     6.6% (loss)
Game: up            Outcome: all-in          7397    26.7% (win)
Game: up            Outcome: expired         5204    18.8% (loss)
```

Game Times

Working hours mostly

Rarely weekends

8 hour shifts

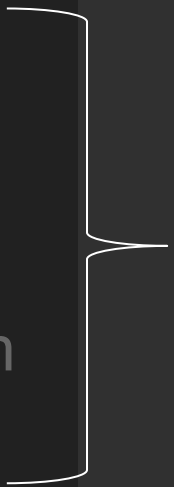This looks like people

Does any of this matter?

"Yes!"   Patterns mean we can model and make better decisions.

"Maybe."   When our strategy is competitor dependent.

"No."   If demand isn't impacted, we don't care.

Elasticity

| Stationarity
| Periodogram
| Deseasonalize
| Magnitude binning
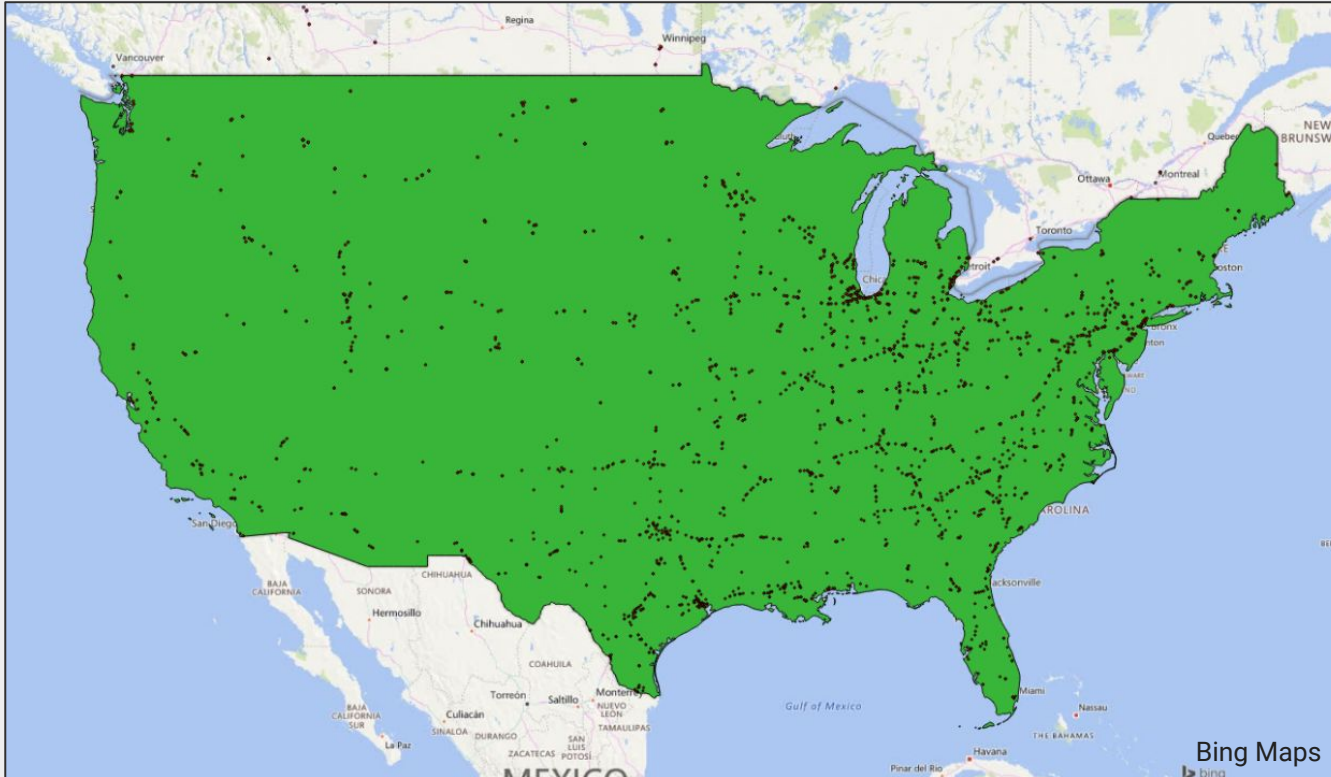| Select by non-equilibrium
| Regress

Demand is impacted

"We need competitor intelligence."

Who is a competitor?

First Law of Geography

"Everything is related to everything else, but near things are more related than distant things."
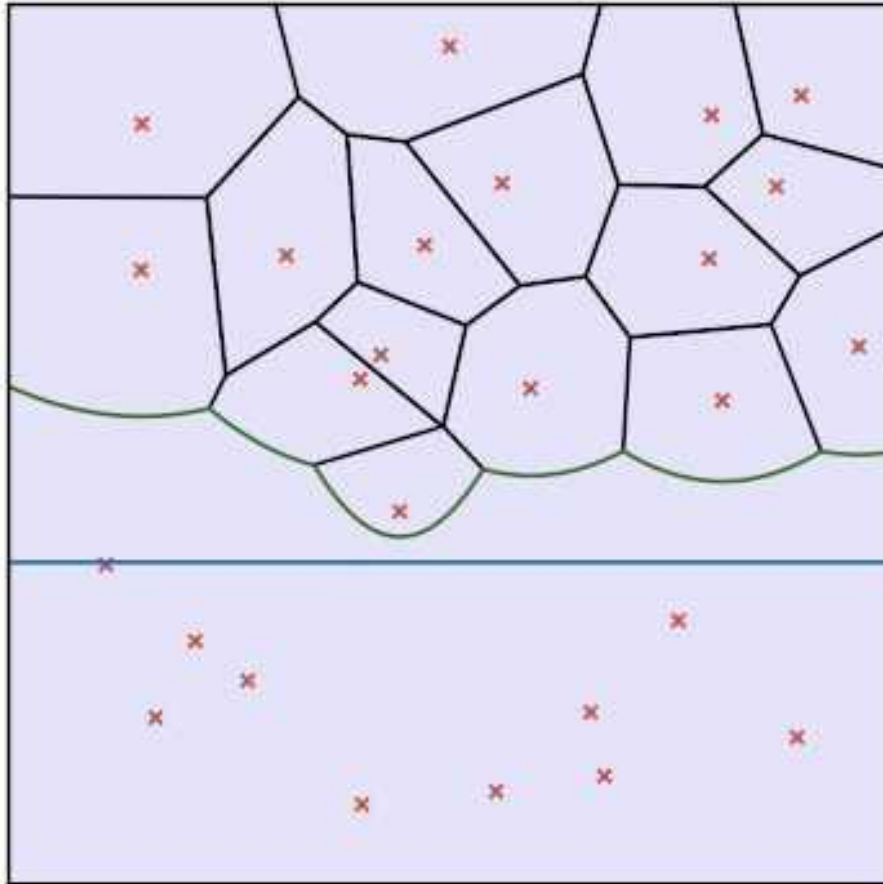

Bing Maps

Dividing up a space

| Desire a unique representation of the space

| Want to capture 'nearness' and 'farness'

| Prefer to avoid bias of large areas for rural locales

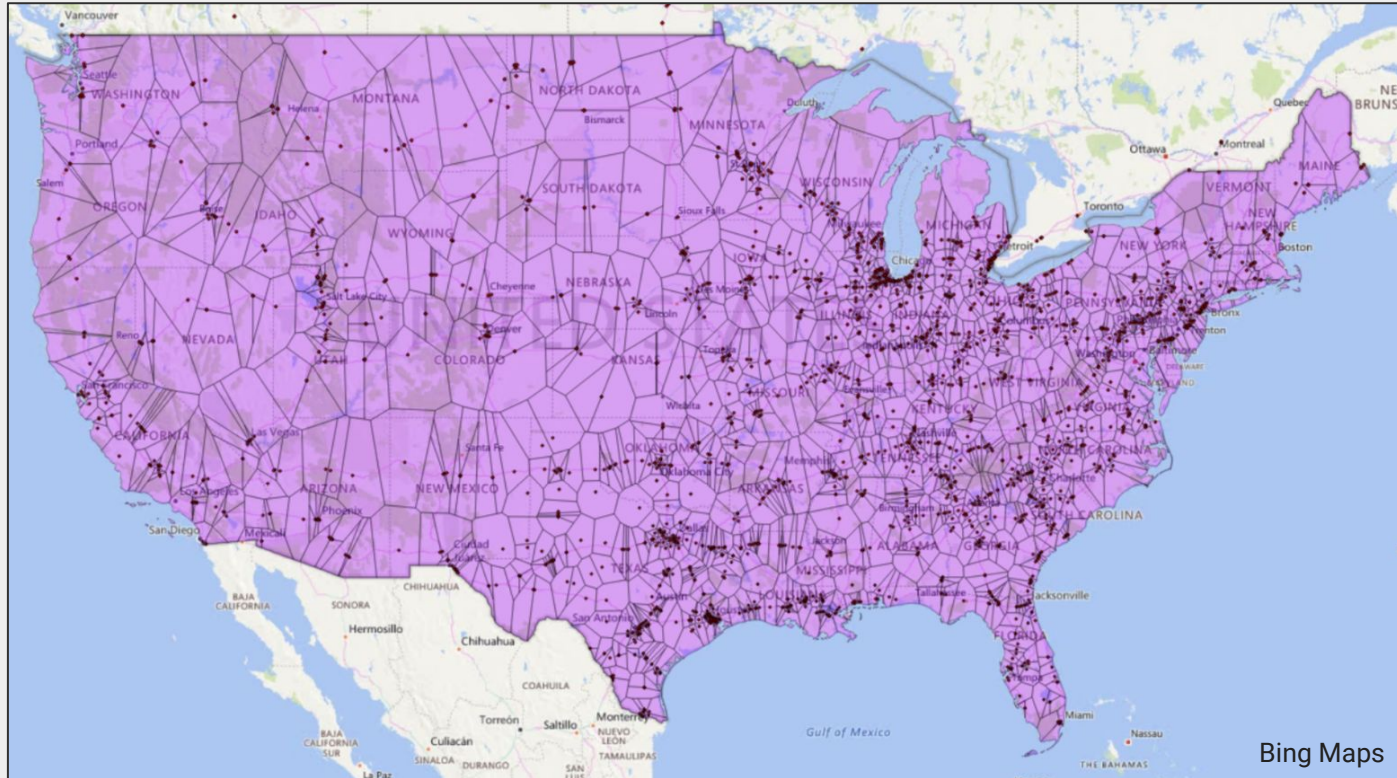| Regions can be culled for road network access
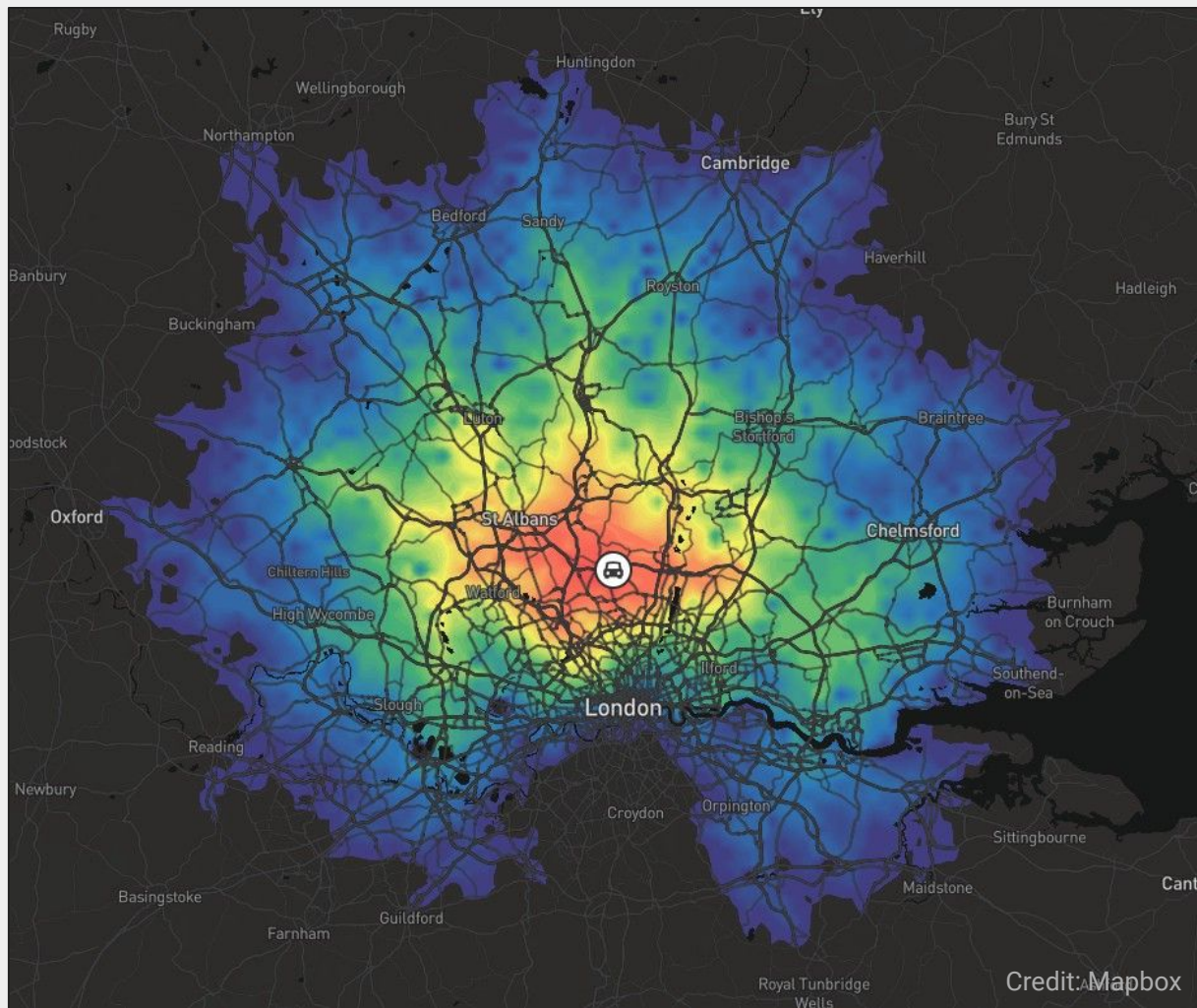
Voronoi Tessellation

Simple

Interpretable

Fast

Tessellated USA
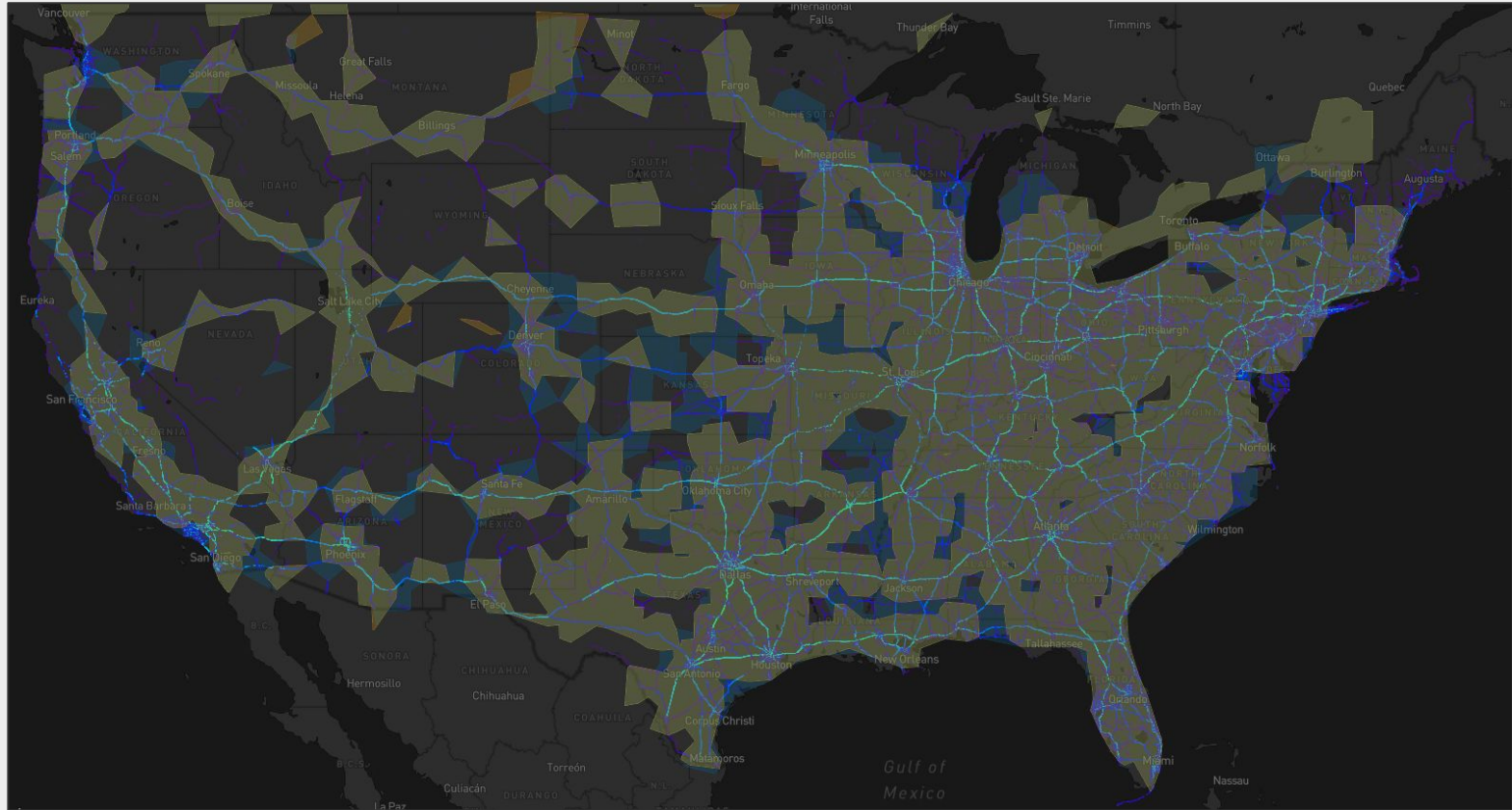
Need to correct for road network drive times

Bing Maps

Isochrones

Realistic drive times

Encode road network

Reasonable constraint

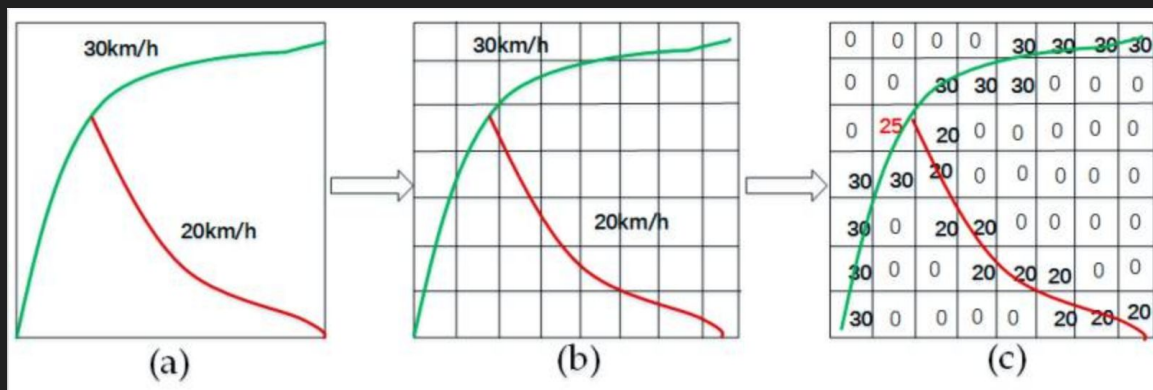Credit: Mapbox

IsoVorochrones (IVC's)

Apply Game Logic

"We need competitor intelligence."

How to model competitor behavior?

Be lucky

| Parallel work on traffic and demand patterns
| Analysis focused on macrocellular, FAF4, HPMS, and CalTrans data sets
| Analog problem in traffic and weather pattern prediction [1, 2, 3, 4]:

Yu et al. 2017

1: Ma et al. 2015, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data.
2: Shi et al. 2015, Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting
3: Yu et al. 2017, Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks.
4: Wang et al. 2017, Spatiotemporal Modeling and Prediction in Cellular Networks: A Big Data Enabled Deep Learning Approach.

The Model

Forecast

Encode

Image Sequences

ConvLSTM

ConvLSTM

ConvLSTM

ConvLSTM

Image Stack
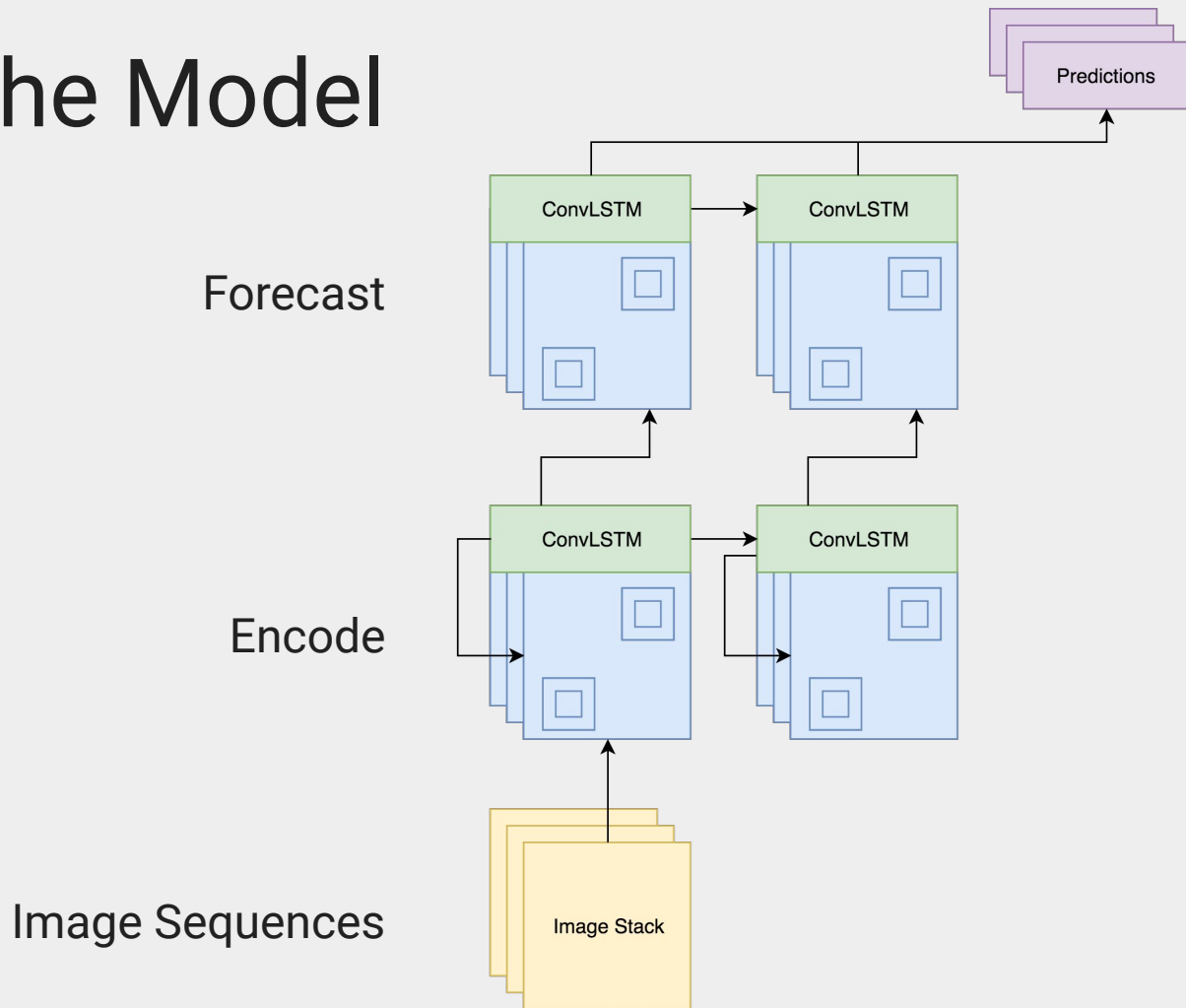
Predictions

Predictions

| Outputs are 2D images in set time steps

| If value in time step, that's when a move occurs

| Values represent who and how much

Future Work

Scenario Engine:

1. Take current market state and fake a range of changes

2. Make predictions for each change going out 48 hours

3. Combine predictions into scenarios and gauge value of each

4. Select a set of strategic moves based on desired outcomes

End

Thank you.