Analytics for an Online Retailer: Demand Forecasting and Price Optimization

Kris Johnson – MIT, Operations Research Center

Alex Lee – MIT, Systems Design & Management

Murali Narayanaswamy – Rue La La, VP Pricing & Operations Strategy

Philip Roizin – Rue La La, Chief Financial Officer

David Simchi-Levi – MIT, Operations Research Center

Jonathan Waggoner – Rue La La, Chief Operating Officer

Massachusetts Institute of Technology

Online Retailing: Online Fashion Sample Sales Industry

- Offers extremely limited-time discounts (“flash sales”) on designer apparel & accessories
- Emerged in mid-2000s and has had nearly 50% annual growth in last 5 years
- Key players
  - **Rue La La (US)**
  - Gilt Groupe (US)
  - Markafoni (Turkish)
  - Trendyol (Turkish)

Snapshot of Rue La La's Website

From the Reserve: Watches by Rolex & Cartier ▸

CLOSING IN 2 DAYS, 19:47:42

Judith Ripka Jewelry & Watches ▸

CLOSING IN 2 DAYS, 19:47:42

Check Off His List: Gift Ideas Under $100 ▸

CLOSING IN 2 DAYS, 19:47:42

Saucony Women ▸

CLOSING IN 1 DAY, 19:47:42

Furs by Christian Dior & More: Picks by WGACA ▸

CLOSING IN 1 DAY, 19:47:42

Saucony Men ▸

CLOSING IN 1 DAY, 19:47:42

“Style”

Saucony "Triumph 10" Running Shoe
~~$130.00~~ **$79.90**

Saucony "Progrid Guide 6" Running Shoe
~~$110.00~~ **$65.90**

Saucony "Triumph 10" Running Shoe
~~$130.00~~ **$79.90**

“SKU”

Saucony "Progrid Guide 6" Running Shoe

~~$110.00~~ **$65.90**

Size

5 5.5 6 6.5 7 7.5 8 8.5 9
9.5 10 10.5 11

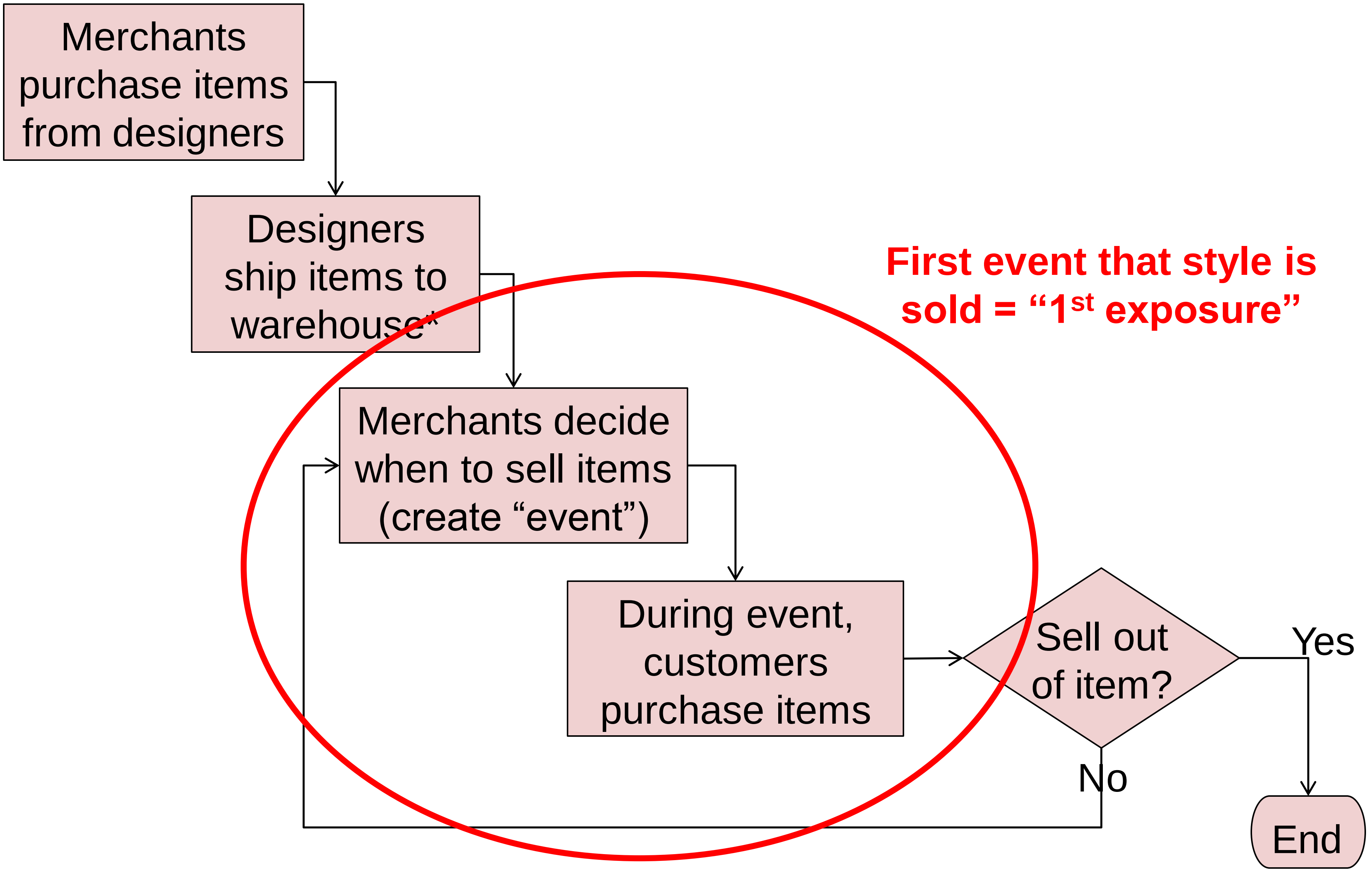
Quantity

1 ▼

ADD TO BAG

Sign up for Quick! Buy It.
Never miss out on something you love.

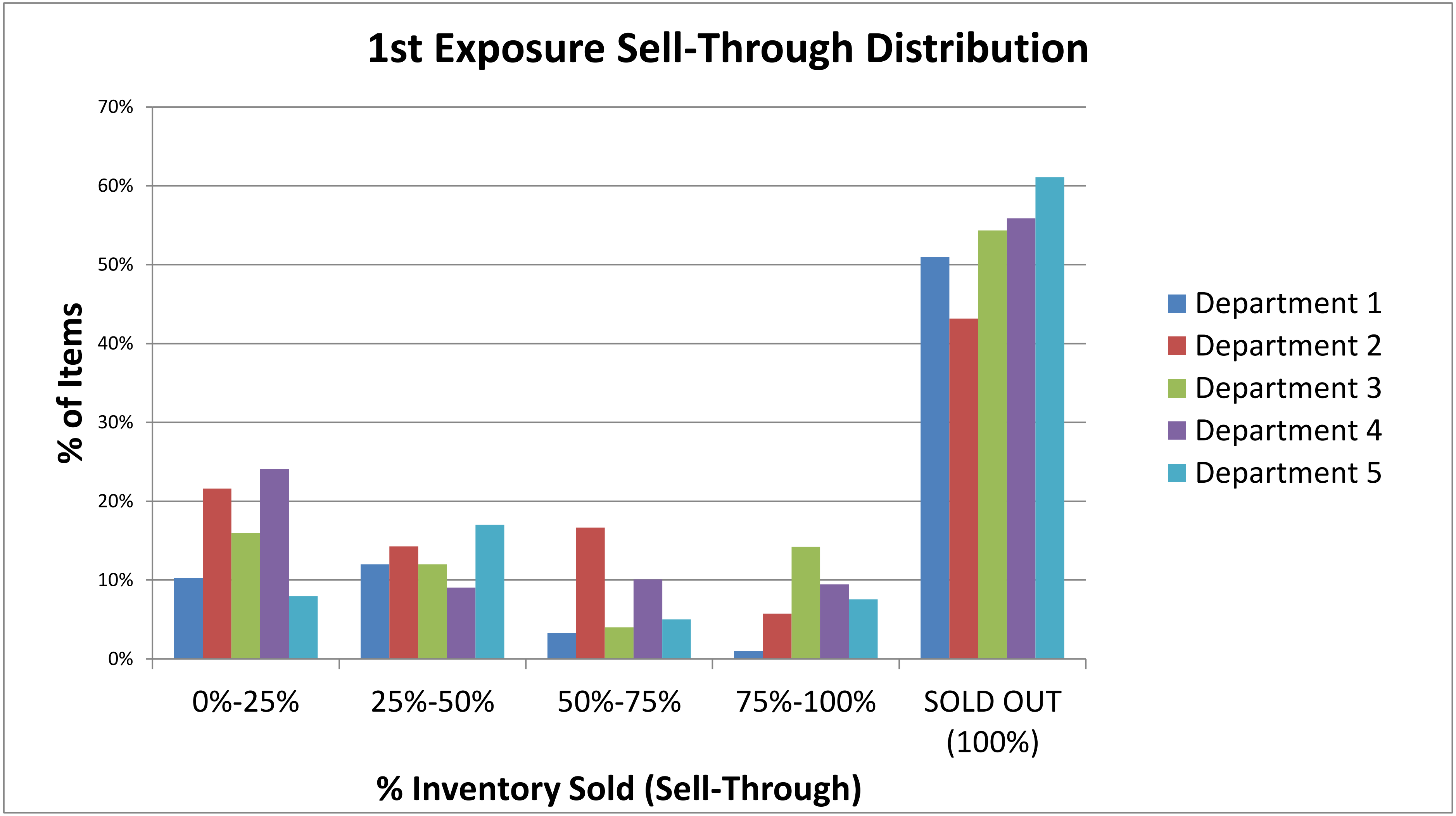
Flash Sales Operations



*Sometimes designer will hold inventory

Video 1 Placeholder:

https://www.youtube.com/watch?v=ahOHAsECeIw&feature=youtu.be



*Data disguised to protect confidentiality

Approach

Goal: Maximize expected revenue from 1st exposure styles

Demand Forecasting

Challenges:

- Predicting demand for items that have never been sold before
- Estimating lost sales

Techniques:

- Clustering
- Machine learning models for regression

Price Optimization

Challenges:

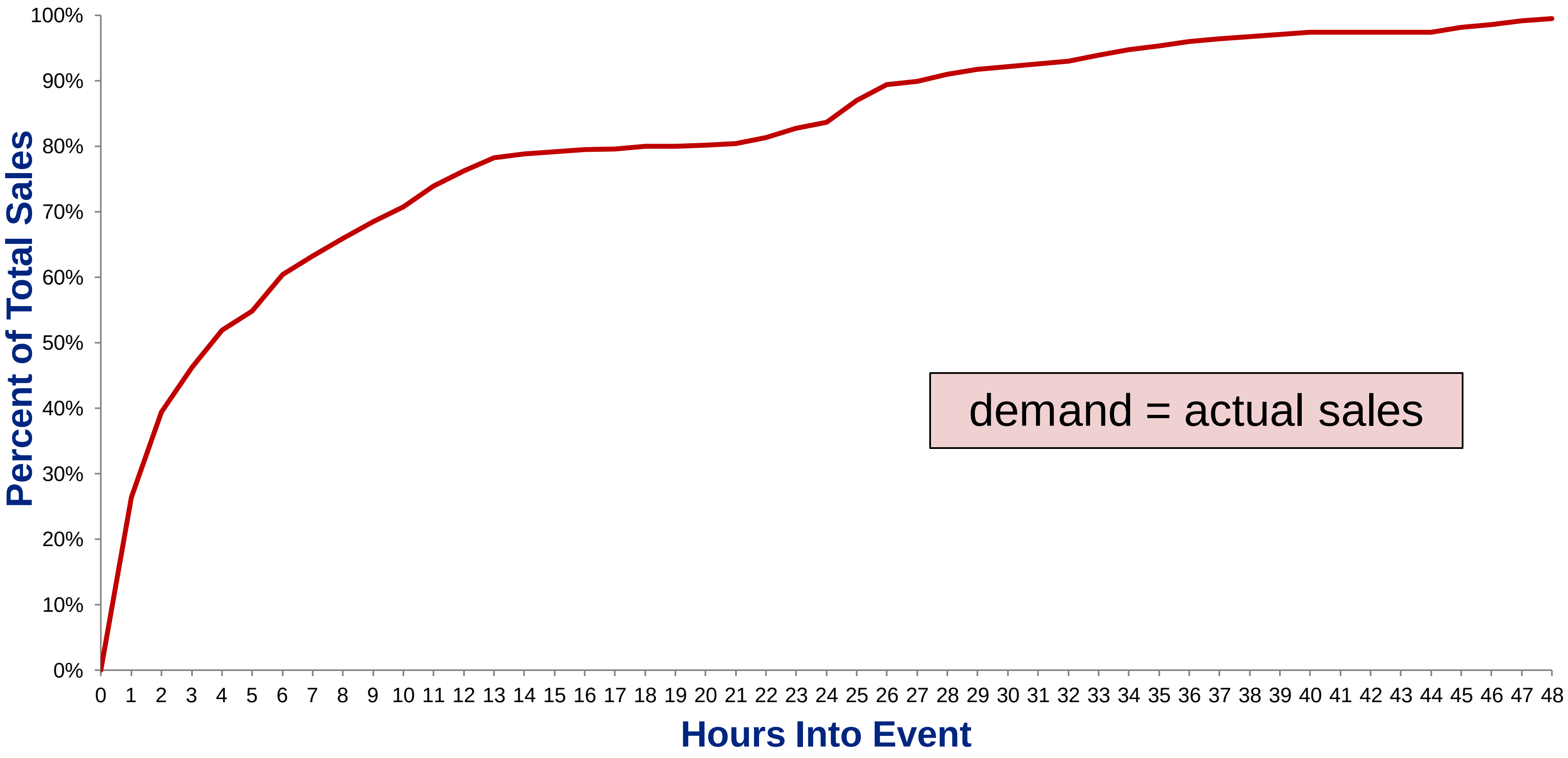
- Structure of demand forecast
- Demand of each style is dependent on price of competing styles → exponential # variables

Techniques:

- Novel reformulation of price optimization problem
- Creation of efficient algorithm to solve daily

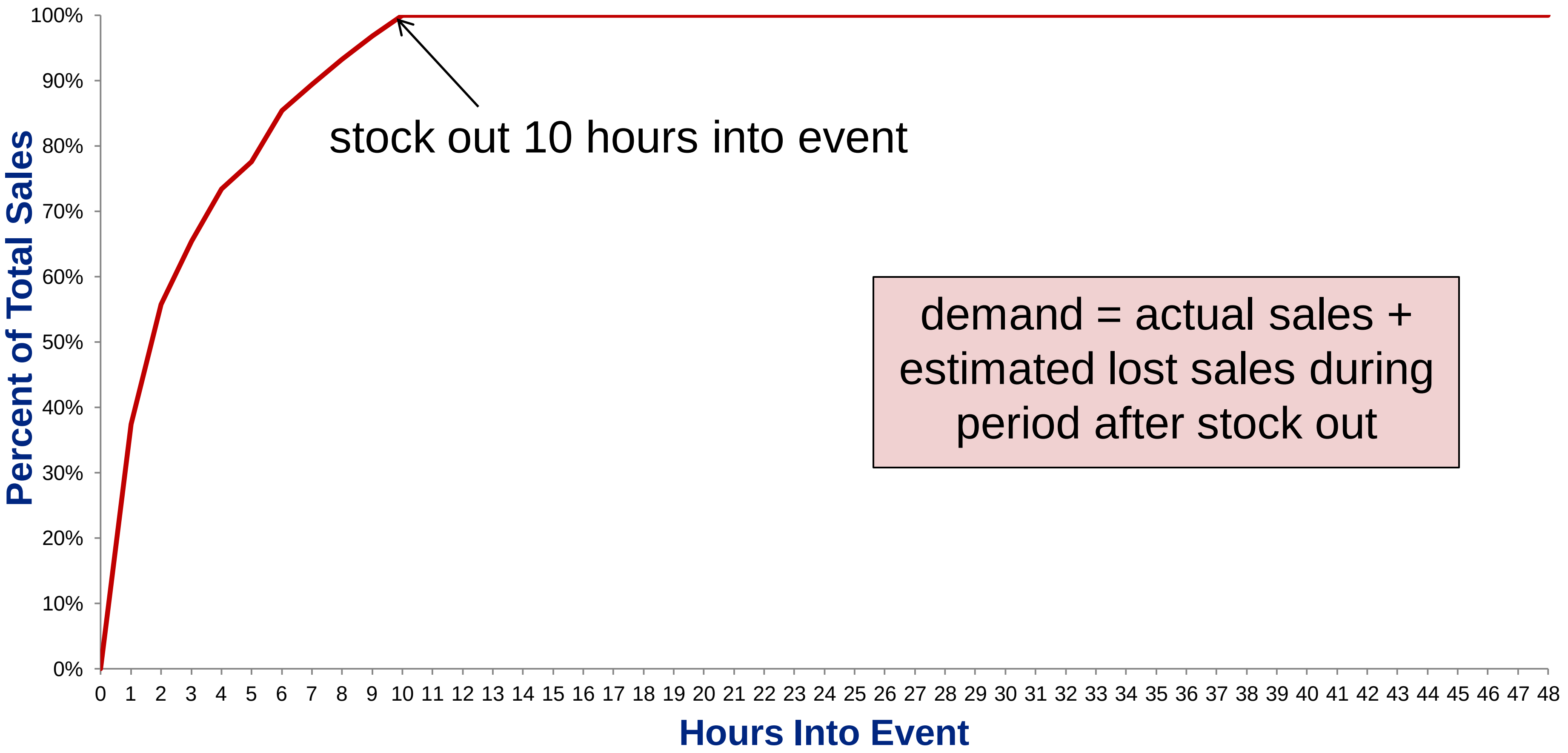
Example Sales Curve for an Item that Doesn’t Sell Out

(sales < inventory)



Example Sales Curve for an Item that Does Sell Out

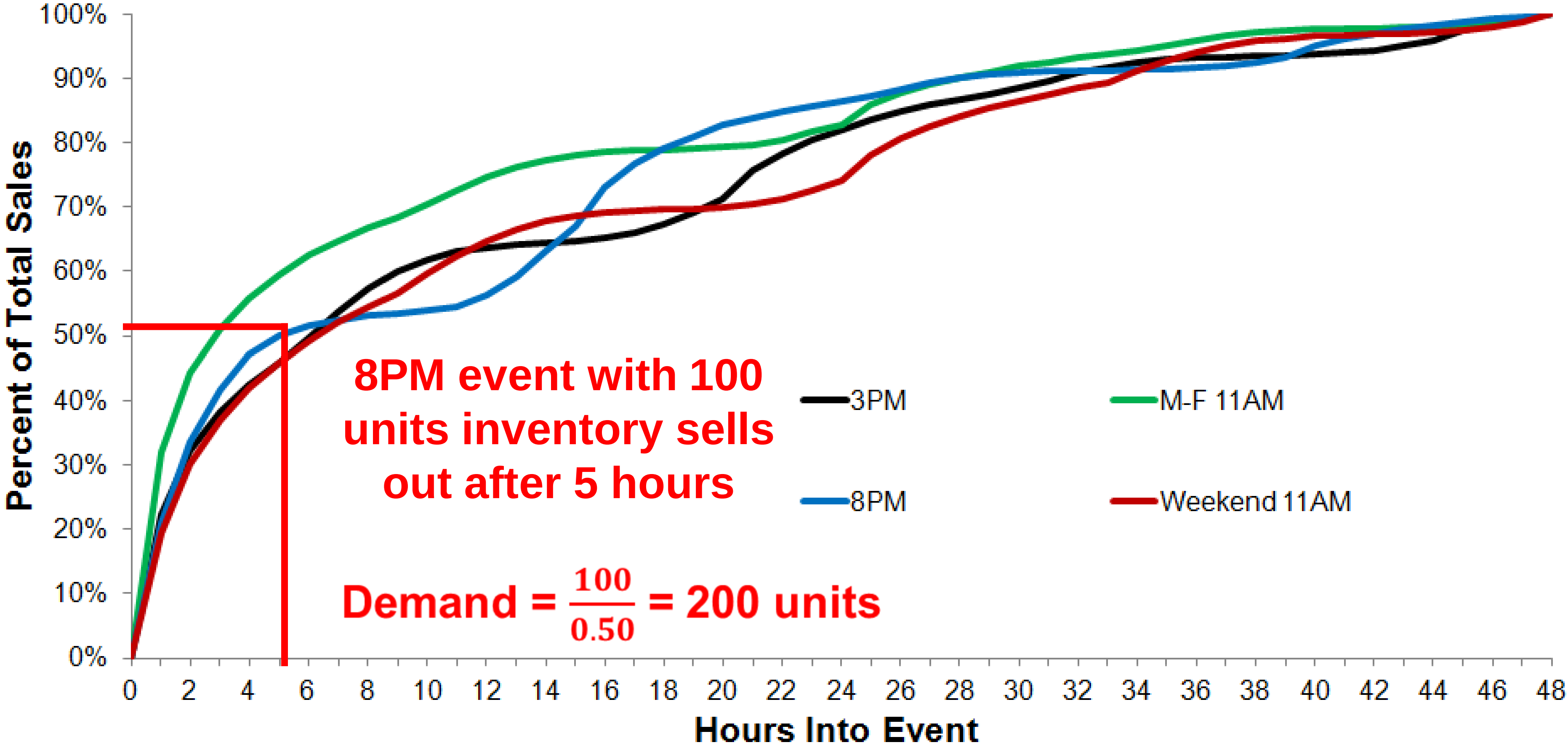
(sales = inventory)



Estimating Lost Sales

- Use data from items that did not stock out to predict lost sales of items that did stock out
- For each event length…
  - Aggregate hourly sales given set of characteristics, i.e. event start time of day
  - Create sales curve for each set of characteristics
    - Results in hundreds of sales curves
    - Use clustering to help further aggregate

Example Clustering Results: Demand Curves for 2-Day Events



Forecasting Model: Explanatory Variables Included

Products

- Department
- Class
- Color Popularity
- Size Popularity
- Brand Type A/B
- Brand Popularity

Combination

- Price
- % Discount = (1 – Price / MSRP)
- # Concurrent Events in Department
- # Styles Sold in Same Subclass and Event (i.e. # Competing Styles)
- Relative Price of Competing Styles
- # Branded Events in Previous 12 Months

Events

- Year
- Month
- Week Day / Time
- Event Type
- Event Length

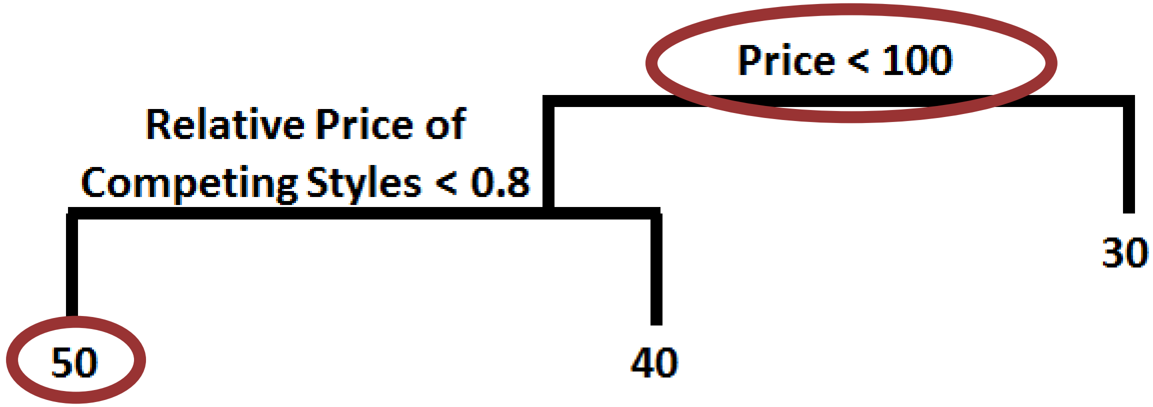
Each input is calculated for a unique {style, event} pair.

Forecasting Model Approach

- Separate data by department; for each department…
  - Randomly divide into training & testing data sets
  - Apply several machine learning techniques to training data
    - Linear regression
    - Power regression
    - Semi-logarithmic regression
    - Regression trees
  - Use cross-validation to choose best model

Regression Tree – Illustration

If condition is true, move left; otherwise, move right



Demand prediction

Approach

Goal: Maximize expected revenue from 1st exposure styles

Demand Forecasting

Challenges:
- Predicting demand for items that have never been sold before
- Estimating lost sales

Techniques:
- Clustering
- Machine learning models for regression

Price Optimization

Challenges:
- Structure of demand forecast
- Demand of each style is dependent on price of competing styles → exponential # variables

Techniques:
- Novel reformulation of price optimization problem
- Creation of efficient algorithm to solve daily

Complexity

- Three of the features used to predict demand are associated with pricing
  - Price
  - % Discount = $\dfrac{1 - \text{Price}}{\text{MSRP}}$
  - Relative Price of Competing Styles =

$$\frac{\text{Price}}{\text{Avg. Price of Competing Styles}}$$

- **Pricing must be optimized concurrently for all competing styles**

Key Observation

- Demand depends only on *average* price of competing styles
- Let N = # competing styles (to be priced concurrently), and let k = the sum of prices of all styles
  - Average price = $\frac{k}{N}$
  - Relative price of competing styles = $\frac{price}{k/N}$
- Finite set of possible prices
  - Prices must end in \$4.90 or \$9.90
  - Consists of lower bound, upper bound, and every increment of \$5.00 between the bounds
  - Ex: {\$24.90, \$29.90, \$34.90, \$39.90}

Key Idea for Algorithm

- Formulate integer optimization problem for each value of k, (IP_k)

Maximize Revenue

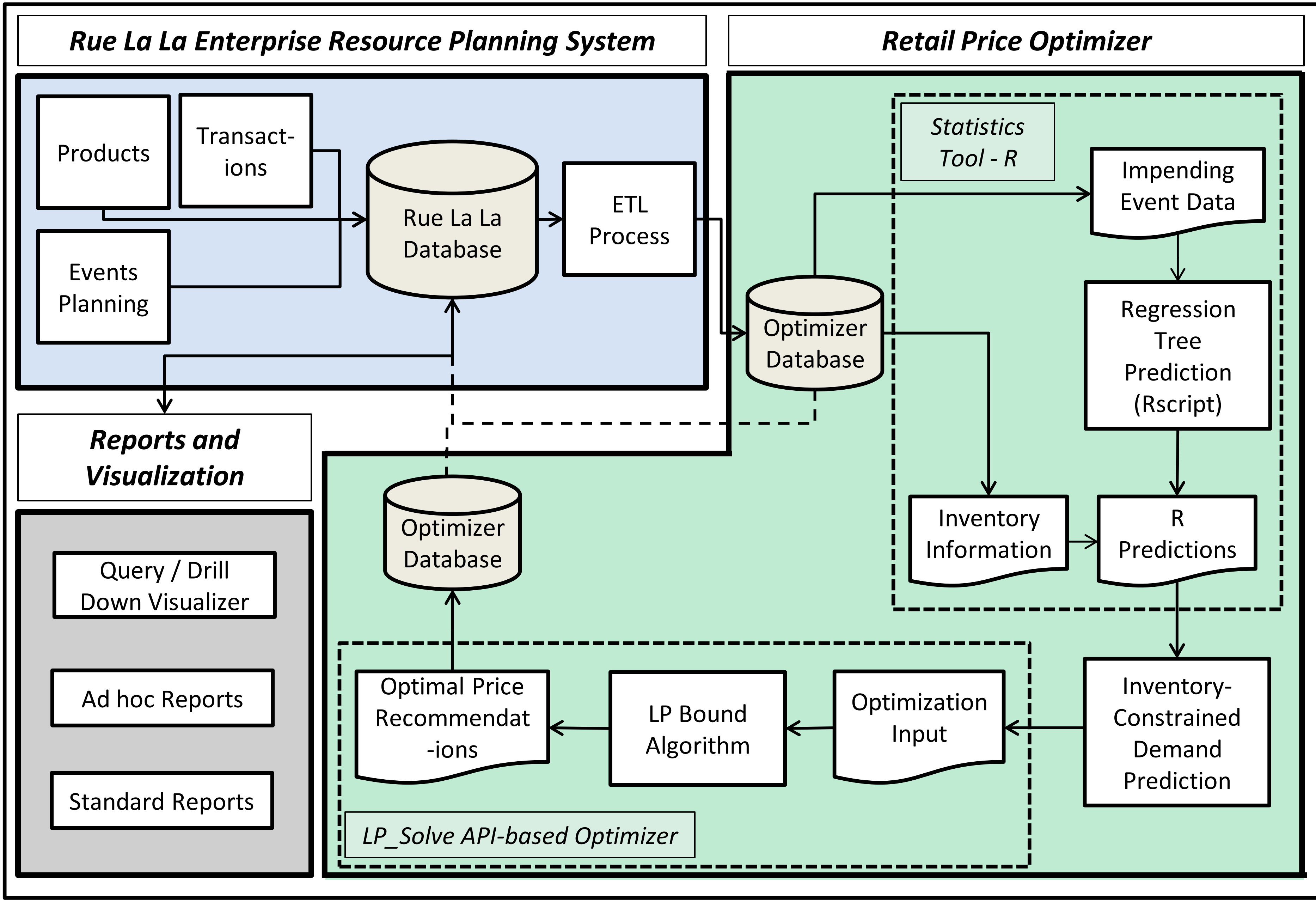
s.t. *1) Each style must be assigned exactly one price*

2) Sum of prices of all styles must = k

- Can show that optimal objective of (IP_k) and its linear relaxation only differ by the revenue associated with a single style!
  - Independent of problem size
- Use this to develop efficient algorithm to solve on daily basis

IMPLEMENTATION & IMPACT

Pricing Decision Support Tool



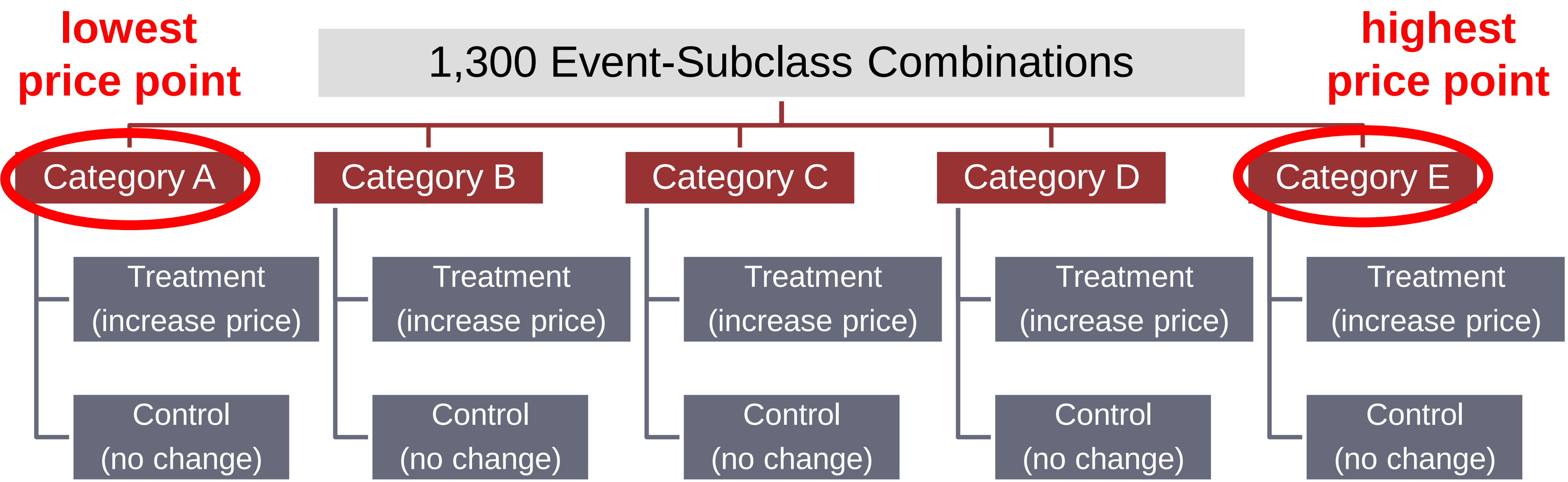
Video 2 Placeholder:

https://www.youtube.com/watch?v=lc4wV6O_YDA&feature=youtu.be

Live Tests

- Motivated by historical analysis
  - Suggests model recommended price increases will increase revenue by ~10% with little to no impact on demand
- Set lower bound on price = merchant suggested price
  - Model only recommends price increases (or no change)
- Identified ~1,300 event-subclass combinations where tool recommended price increases for at least one style

Live Tests


lowest
price point
highest
price point
1,300 Event-Subclass Combinations
Category A
Category B
Category C
Category D
Category E
Treatment
(increase price)
Control
(no change)

Mann-Whitney / Wilcoxon Rank Sum Test

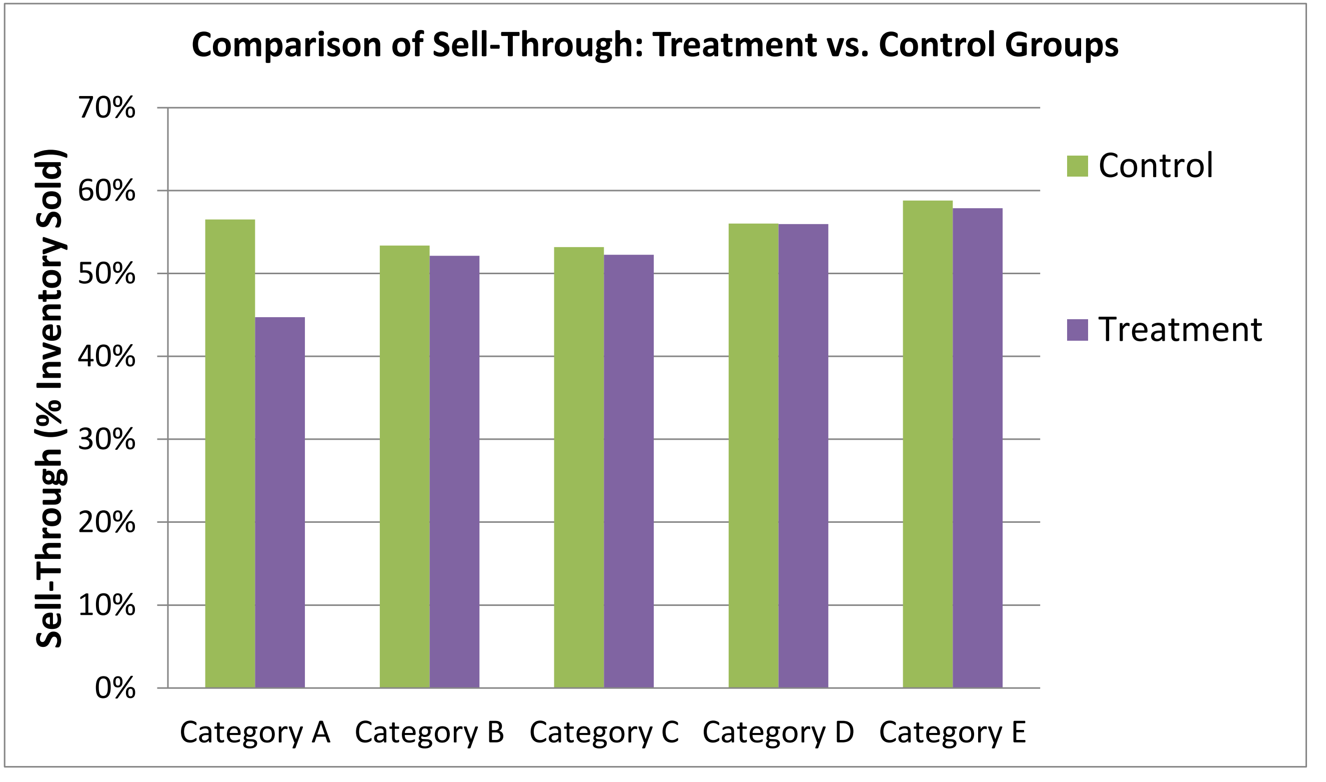
- Hypothesis test that assumes no particular distributional form on treatment or control groups
  - H_0: raising prices has no effect on sell-through
  - H_A: raising prices decreases sell-through
- Idea of test
  - Combine sell-through data of treatment and control groups
  - Order data and assign rank to each observation
  - Sum ranks of all treatment group observations
  - If sum is too low, reject H_0

Mann-Whitney / Wilcoxon Rank Sum Test

1,300 Event-Subclass Combinations

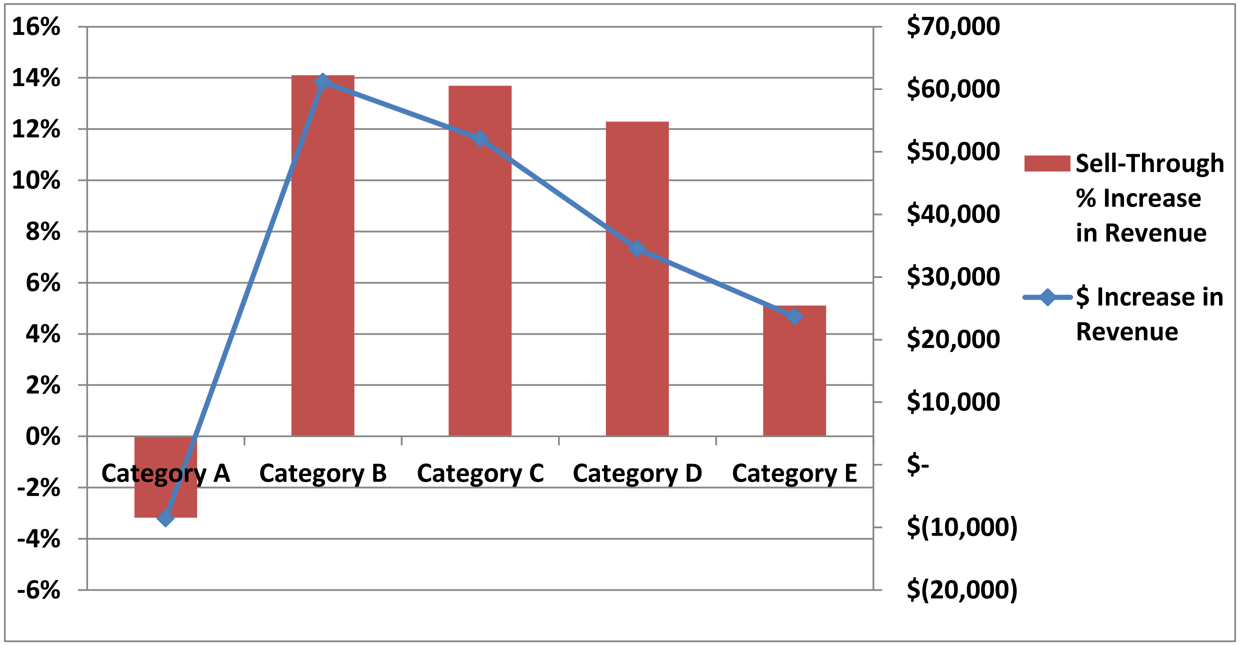
- Category A
  - Treatment (increase price)
  - Control (no change)
  - **Rejects H_0, α = 1%**
- Category B
  - Treatment (increase price)
  - Control (no change)
  - **Does not reject H_0, α = 10%**
- Category C
  - Treatment (increase price)
  - Control (no change)
  - **Does not reject H_0, α = 20%**
- Category D
  - Treatment (increase price)
  - Control (no change)
  - **Does not reject H_0, α = 20%**
- Category E
  - Treatment (increase price)
  - Control (no change)
  - **Does not reject H_0, α = 20%**

Visual Comparison



Revenue Impact

- Treatment group's increase in revenue, assuming demand is impacted by price increases as shown on previous slide



Video 3 Placeholder:

https://www.youtube.com/watch?v=AzJhAxkpkEU&feature=youtu.be

Conclusion

- Created and implemented pricing decision support tool that recommends prices for 1st exposure styles
  - Used clustering to estimate lost sales
  - Built regression trees to predict demand
  - Developed efficient algorithm to solve multi-product price optimization problem
- Implementation of these analytics techniques shows expected increase in revenue of ~10% with little impact on demand

Our Team

Murali Narayanaswamy – VP Pricing & Strategy
Philip Roizin – Chief Financial Officer
Jonathan Waggoner – Chief Operating Officer

Kris Johnson – Operations Research Center
Alex Lee – Systems Design & Management
David Simchi-Levi – Operations Research Center

Deb Mohanty
Hemant Pariawala

Marjan Baghaie, Andy Fano
Paul Mahler, Matt O'Kane