Target: Using Analytics to Improve Asset Protection

Saurabh Bodas, Lin Chen, Jake Hill, Shelby Watson
Acknowledgements

- **Ed Tonkon**, Zebra Technologies
- **Jess Pena**, Target
- **Tanner Coghill**, Target
- **Lisa Bruno**, RILA
- **Ellen Jackson**, RILA
- **Dr. Tej Anand**, MSBA Faculty
Master of Science in Business Analytics (MSBA)

- Ranked #3 in Business Analytics worldwide
- 10-month intensive STEM-certified program

**ACADEMIC BACKGROUND**

- 40% ENGINEERING
- 19% BUSINESS
- 15% MATHEMATICS
- 14% ECONOMICS
- 9% COMPUTER SCIENCE
- 3% OTHER

**GENDER**

- 57% MALE
- 43% FEMALE

**AVERAGE TEST SCORES**

- GPA: 3.66
- GMAT: 700
Asset Protection: Countering Theft

$50 Billion
Annual Theft Loss

1.38%
Average Shrink Rate

National Retail Security Survey
Largest Sources of Loss

- Administrative error and Supply Chain Loss
- Internal and External Theft
- Other non-malicious losses
Background and General Observations
Our Main Objectives

1. Track performance of AP teams

2. Optimize resources to prevent the most theft
Understanding the Data

• 1,800+ stores
• 2015-2019
• Broken down into two main segments:
  • Annual Store Data (annual sales, shortage, store attributes, etc.)
  • Weekly Department Data (weekly theft statistics)
• Weekly data is collected as records from individual AP teams
• Annual data is collected from aggregate store records
Exploratory Data Analysis

- Granular Data
- Missing Values
- Addressed through clustering

**Missing Week Distribution Across Stores**

**Missing Week Distribution Across Merch Division**
Addressing Objective 1

Measure an AP Team’s performance against itself

• Trend Extraction

Measure an AP Team’s performance against other similar stores

• Store segmentation
• Assess performance through theft prevention within groups
How to easily interpret a boxplot

Data from above

Side view
Measuring AP Team Performance: Trend Extraction

• Trend is a general direction for the theft time series and could be a good proxy for measuring the performance of Asset Protection team against itself
  • Taking empty package as an example
  • If the trend is always going down with a good amount, performance is improving
  • Otherwise it stays constant or worsens

• Time series is often affected by seasonality and trend need to be extracted first.
Measuring AP Team Performance: Trend Extraction

Original Data

Trend
Measuring AP Team Performance: Trend Extraction

• On average, the value of recorded empty package in 2018 decreased by 15 dollars on a weekly basis.

<table>
<thead>
<tr>
<th>Store</th>
<th>ep_2016</th>
<th>ep_2017</th>
<th>ep_2018</th>
<th>ep_2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>A</td>
<td>3.1079</td>
<td>1.8402</td>
<td>-14.7292</td>
</tr>
</tbody>
</table>

• **Implication:** Give a quantitative measure of reduced dollar amount

• **Elasticity:** This method can check quarter, semi-annual and annual performance of Asset Protection team.

• **Limitation:** It requires high-quality and streamlined data collection for at least 2 years in order to get rid of seasonality effect.
Tracking AP Team Performance

Evaluating AP team performance is tricky:

- Occurrence of crime can be erratic
- Cannot set target theft metrics to be achieved

Best approach: compare each store’s relative performance against all other stores
Tracking AP Team Performance

Target has 1800+ Stores

- Different Geographies
- Store Size
- Store prototype

Riskiness of the Neighborhood

Store segmentation is necessary

How do you quantify risk?
Tracking AP Performance

Why does a particular store prevent more theft than another store?

- More Square Footage
- Riskier Neighborhood
- AP team performs well
Explaining CAP Scores

Each store receives a custom score between 0 - 2000

CRIMECAST Scores: Developed by CAP-Index

Criminology

Data Science

Social disorganization theory

Data Science
Average Prevented Theft (across different CAP scores)

CAP Risk
- VERY LOW
- LOW
- MEDIUM
- HIGH
- VERY HIGH

High Theft

Low Theft

Dec 1, 17  Mar 1, 18  Jun 1, 18  Sep 1, 18  Dec 1, 18  Mar 1, 19  Jun 1, 19  Sep 1, 19  Dec 1, 19
2019 Sales vs. Prevented Theft

High Theft

Low Theft

Low Sales

mean daily sales

High Sales

CAP Risk

VERY LOW
LOW
MEDIUM
HIGH
VERY HIGH
Prevented Theft as % of Sales

Store EK is performing incredibly well!

High

Low

VERY LOW

LOW

MEDIUM

HIGH

VERY HIGH

Top performing stores (top 1 %ile)
Stores Redistributed: from ‘excellent’ to ‘very poor’

There’s our strong performer again!

How do we move all stores to the ‘excellent’ category?
What Next?

Study best-performing stores

What are they doing differently?

Theft Metrics
- Empty Packaging
- RFID
- Count Updates

AP Staff Data
- Store Manager
- Training Programs
- Team members
Tracking AP team performance

Prevented theft: Store 1
- CAP Risk: HIGH
- store status: good

Prevented theft: Store 2
- CAP Risk: MEDIUM
- store status: poor

The Big Picture: 2019 Sales vs. Prevented Theft (selected stores are highlighted)

The Big Picture: Distribution of Stores (selected stores are highlighted)
Addressing Objective 2

Developing a way to optimize resources for AP Teams

• Data-driven approach

3 Main Steps

• Clustering
• Time Series Forecasting
• Dashboards and Business Optimization
Why Cluster?

Most stores are missing 55+ weekly data points
Worst Case Scenario

Store: BSS
Department: 1

That’s a lot of weeks with zeros!
Best Case Scenario

Store: AHM
Department: 1

Zeros are still causing a lot of variation
Clustering Method Used: Gaussian Mixture Models

K-means (most common)  GMM (most optimal)

It’s just a different shape.
Attributes used for clustering

- **Quarterly theft figures**
  - 13 quarters used

- **Department shortage rates**
  - 26 departments
Clustering stores with similar theft patterns solves the missing data problem
Although clusters 0 and 4 have similar theft figures, their shortage rates differ across departments.
Forecasting Theft: Predicting Future Trends
Optimizing Resource Allocation: Forecasting Theft

![Graphs showing average loss for different clusters from 2016 to 2020](image-url)
Optimizing Resource Allocation: Forecasting Theft

- **Time Series Models:** automated the forecasting process
- **Different Model Families:** ARIMA, TBATS, hybrid, fourier terms, ensemble
- **Benchmark Metrics:** mean, naïve, seasonal naïve

**Purpose:** Update AP hours allocated to each department every week
Good Forecastability

Dept X, 2: Forecasts of $ Loss

$ Median Loss

ARIMA
mean
mean_18
naive
actual_loss

Time

2016 2017 2018 2019 2020
Good Forecastability

Dept X, 2:

Forecasts of $ Loss (zoomed in)
High Variation

Dept Y, 0:

Forecasts of $ Loss

$ Median Loss

2016  2017  2018  2019  2020

Time

ARIMA  mean  mean_18  naive  actual_loss
Noisy/Little Structure

Dept Y, 0: Forecasts of $ Loss (zoomed in)

$ Median Loss

- training data
- test data & forecast
- future forecast

$ Median Loss vs. Time

How do we improve theft forecasts?

1. Prediction intervals for forecasts
2. Data on special events to explain sharp spikes/drops in $ loss
   - weekly promotions
   - anomalous store operations
   - holidays
   - weather forecasts
3. Recalibrate forecasts: COVID-19
Optimizing Resource Allocation: Results and Dashboards
These 5 departments should experience a spike next week

<table>
<thead>
<tr>
<th>Week-on-Week % Change in Forecasted Theft</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEASNL/OUTDR LIVING</td>
</tr>
<tr>
<td>KITCHEN</td>
</tr>
<tr>
<td>TOYS</td>
</tr>
<tr>
<td>SHOES</td>
</tr>
<tr>
<td>PERFORMANCE</td>
</tr>
<tr>
<td>NIT APPAREL</td>
</tr>
<tr>
<td>DOMESTICS</td>
</tr>
<tr>
<td>YOUNG CONTEMPOR.</td>
</tr>
<tr>
<td>KIDS APPAREL</td>
</tr>
<tr>
<td>SWIM</td>
</tr>
</tbody>
</table>

These 5 departments should experience a dip next week
Allocate % of time in labor hours to areas that are predicted to experience that portion of theft

The week-on-week % change from the previous slide is reflected here

Store: AHM
Total Staff Hours

Allocation of Staff Hours

<table>
<thead>
<tr>
<th>Category</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOME ELECTRONICS</td>
<td>xxx</td>
</tr>
<tr>
<td>STORAGE/UTILITY</td>
<td>xxx</td>
</tr>
<tr>
<td>BEAUTY</td>
<td>xxx</td>
</tr>
<tr>
<td>BABY</td>
<td>xxx</td>
</tr>
<tr>
<td>MENS</td>
<td>xxx</td>
</tr>
<tr>
<td>SPORTING GOODS</td>
<td>xxx</td>
</tr>
<tr>
<td>MOBILE/SERVICES</td>
<td>xxx</td>
</tr>
<tr>
<td>READY-TO-WEAR</td>
<td>xxx</td>
</tr>
<tr>
<td>TOYS</td>
<td>xxx</td>
</tr>
<tr>
<td>KITCHEN</td>
<td>xxx</td>
</tr>
<tr>
<td>JEWELRY/ACCESS</td>
<td>xxx</td>
</tr>
<tr>
<td>YOUNG CONTEMPORARY</td>
<td>xxx</td>
</tr>
<tr>
<td>DOMESTICS</td>
<td>xxx</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>xxx</td>
</tr>
<tr>
<td>ENTERTAINMENT</td>
<td>xxx</td>
</tr>
<tr>
<td>INT/HOS/SLEEP</td>
<td>xxx</td>
</tr>
<tr>
<td>OTC</td>
<td>xxx</td>
</tr>
<tr>
<td>NIT APPAREL</td>
<td>xxx</td>
</tr>
<tr>
<td>KIDS APPAREL</td>
<td>xxx</td>
</tr>
<tr>
<td>SHOES</td>
<td>xxx</td>
</tr>
<tr>
<td>DECORATIVE HOME</td>
<td>xxx</td>
</tr>
<tr>
<td>HSHLD/PAPR</td>
<td>xxx</td>
</tr>
<tr>
<td>SWIM</td>
<td>xxx</td>
</tr>
<tr>
<td>STATIONERY</td>
<td>xxx</td>
</tr>
<tr>
<td>PETS</td>
<td>xxx</td>
</tr>
<tr>
<td>SEASNL/OUTDR LIVING</td>
<td>xxx</td>
</tr>
</tbody>
</table>
This cluster’s forecast has a pretty good fit, only slightly under-estimating the actual theft.
A positive long term trend may suggest [...]
<table>
<thead>
<tr>
<th>Department</th>
<th>Week-on-Week % Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEASNL/OUTDR LIVING</td>
<td>12.9%</td>
</tr>
<tr>
<td>KITCHEN</td>
<td>5.6%</td>
</tr>
<tr>
<td>TOYS</td>
<td>3.9%</td>
</tr>
<tr>
<td>SHOES</td>
<td>2.7%</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>2.1%</td>
</tr>
<tr>
<td>NIT APPAREL</td>
<td>-0.2%</td>
</tr>
<tr>
<td>DOMESTICS</td>
<td>-1.1%</td>
</tr>
<tr>
<td>YOUNG CONTEMPOR..</td>
<td>-1.3%</td>
</tr>
<tr>
<td>KIDS APPAREL</td>
<td>-4.1%</td>
</tr>
<tr>
<td>SWIM</td>
<td>-26.4%</td>
</tr>
</tbody>
</table>

**These 5 departments should experience a spike next week**

**These 5 departments should experience a dip next week**
Conclusions and Implementation

Measuring AP Team Performance
- Trend Extraction
- CAP Score Segmentation

AP Team Resource Optimization
- Clustering
- Time Series Forecasting
- Resource Allocation Dashboard

Implementation
- Corporate level
- Trickle-down to store level
Contact Information

Saurabh Bodas
- Saurabh.bodas@utexas.edu
- linkedin.com/in/saurabh-bodas

Lin Chen
- jacob.hill@utexas.edu
- linkedin.com/in/jake-hill/

Jake Hill
- cllin.chen@utexas.edu
- linkedin.com/in/linchenkaren/

Shelby Watson
- shelby.Watson@utexas.edu
- linkedin.com/in/shelbyewatson/