Driving Season Ticket Sales

Cleveland Indians Customer Prospect Analysis

Wharton Customer Analytics Initiative
Successful Applications of Customer Analytics
5/1/14
“What gets measured, gets managed”
- Peter Drucker

To assist the organization achieve its Guiding Commitments, by **creating and implementing data-supported decision making processes, tools and systems**, maximizing asset value and enhancing strategic relationships.
Re-evaluating our current model for finding new Season Ticket holders

As we’ve created a new central analytics capability, we’ve utilized that resource to determine a new path for finding our best Season Ticket prospects:

To assist the organization achieve its Guiding Commitments, by creating and implementing data-supported decision making processes, tools and systems, maximizing asset value and enhancing strategic relationships.

A New Model for Season Ticket Prospecting
Different ‘Go to Market’ approach for different ticketing products

We have four distinct ticketing products, which require a different go to market platform, resources, and capability:

<table>
<thead>
<tr>
<th>Product</th>
<th>Marketing / Sales Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mass Marketing</td>
</tr>
<tr>
<td>Single Game Tickets</td>
<td>✔</td>
</tr>
<tr>
<td>Group Tickets</td>
<td>✗</td>
</tr>
<tr>
<td>Season Tickets</td>
<td>✗</td>
</tr>
<tr>
<td>Premium Tickets</td>
<td>✗</td>
</tr>
</tbody>
</table>
Our Seasons business has experienced a fundamental shift in the past decade.

Due to unique factors in CLE and variable team performance, we have a very new reality in our base of selling season tickets:
The 1990’s: A Perfect Storm

The 1990’s brought about a confluence of unrepeatable events that led to unprecedented success for the Indians:

- Nucleus of Young Players
- New Ballpark
- Revenues to support Top 5 payroll
- Post-Strike MLB
- Browns Leave CLE
- Pre-Lebron Cavs Era
- Robust 1990’s CLE Economy
- No Indians Post-Season Since 1954

455 Consecutive Sell-Outs
+ 6 Post-Season Appearances in 7 years
+ 2 World Series Appearances
+ Top 5 MLB Payroll (as high as 3rd)
+ New Standard for Fans to Measure Team Success
Changing Cleveland Market Dynamics Adds to Indians Difficulties

• Significant Economic Downturn
  • Acquisition or relocation of at least four Fortune 500 companies since 2000
  • Recent global economic crisis further impacted CLE auto, manufacturing and financial corporate base leading to loss of sponsorship and unemployment among fan base
  • Cleveland is now the 2nd poorest big city in the U.S.
  • Cleveland poverty rate has grown twice as fast as U.S. rate since 2007
  • Cleveland ranks 3rd in foreclosures according to recent GAO study
• Cleveland had 2nd largest population decline since 2000 (New Orleans #1)
  • CLE now the 43rd largest city and the 29th largest Arbitron market in the U.S.
  • CLE was the 6th largest U.S. city in 1940 and in the top 10 as recently as 1970
• Return of NFL Football to Cleveland in 1999
The unique challenge of prospecting for Season Ticket Holders

Our season ticket holder prospects represent less than 1% of CLE DMA, presenting unique challenges of efficiently finding the best people to target and sell to:

- **Reality**: Season ticket holders are lifeblood of baseball
  - Majority of revenue, <1% of local CLE DMA
- **Challenge**: Create a new model that significantly increases predictability of a prime selling prospect to spend more time SELLING
Changing the resource allocation model to focus on selling

We currently spend the majority of our internal selling resources trying to find the right people and less time on effectively selling our proposition. We want to flip that model and change the dynamic.
**Key questions we considered**

- **Questions:**
  - WHAT attributes make our season ticket holders *unique*?
  - HOW do we more *efficiently* find them?
  - HOW can we *predict* the level of spend and type of product that they’d be open to *before* we have our first contact?

- **Approach:**
  - Determine a new modeling approach for using better, different past information to more precisely predict who we should target, how much they are likely to spend, and what product they’re likely to purchase.
Prospecting Background
What are Prospector & Renewal Models?

**Prospector model**: new season ticket sales

**Renewal model**: retaining current season ticket holders

<table>
<thead>
<tr>
<th>Unique ID</th>
<th>Predictor Variables</th>
<th>Response Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Account ID</td>
<td>Spend last Year</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$100</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>$3000</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>$0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Response Variables

Priority – Purchase

Capacity – Buyer Spend

Revenue – Lead Spend

= Priority * Capacity
3rd Party Score Results

Cleveland Indians
Motivating Questions
Lead Scoring

• With finite time and resources, who should our salespeople contact?

• What change in revenue can be expected from additional data and different modeling techniques?
Fan Understanding

- What customer behavioral and/or demographic characteristics are most relevant for selling new season tickets?
Data Management

• What is the level of some of our current data quality?

• What is the value of...
  • CRM data?
  • Email data?
  • Demographic / Acxiom data?
  • Purchase data from other venues (LiveAnalytics)?
Sales Management

• Given different staffing levels, how many leads is it worthwhile to call?

• Should our focus be on lapsed season ticket holders or potential newer customers?

*If more granular CRM data...*

• Effectiveness of sales departments
• Effectiveness of contact tactics
• Actual vs. expected salesperson performance
Data
### Historical Data Availability at Customer Level

#### Table

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>STH Sales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRM Activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ticket Forwards</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Shows</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singles Online</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singles Walkup</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd Party Scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Email Activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StubHub Resellers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>StubHub Buyers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LiveAnalytics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = will be available

---

**Available**

**Poor or Incomplete**

**Not Available**

Cleveland Indians
Prospect Population

CRM Contacts
185,649 Total Accounts
137,873
60.5%

3rd Party Scores
88,898 Total Accounts
42,069
18.5%

New STH Buyers
1,316 Total Accounts
958
0.1%

For example, 25% of CRM contacts (46,818 / 185,649) went to scored leads.

For example, 159 accounts (12% of all new STHs) bought season tickets without CRM contact.
Data Sets for Models

Train Data 2012-13:

Prospects: 23,737
Close rate: 2.7% (631 buyers)
Avg. lead spend: $67
Avg. buyer spend: $2,537

Test Data 2012-13:

Prospects: 24,198
Close rate: 2.8% (685 buyers)
Avg. lead spend: $70
Avg. buyer spend: $2,465

3rd Party Data 2012-13:

Prospects: 23,616
Close rate: 0.4% (103 buyers)
Avg. lead spend: $8.68
Avg. buyer spend: $1,991
Email Activity

- Open
- Click
- Mark as spam
- Unsubscribe
- Bounce

Total emails sent by month:

Distribution of # of emails received by individual in 2013:

Cleveland Indians
## Ticketmaster’s LiveAnalytics Data

### LiveAnalytics Demographics (LAD) (Via Acxiom)

- Address
- Market demographics
- Gender
- Age
- Household
- Home ownership
- Residence
- Occupation
- Income / financial
- Education
- Mail buyer / responder / donor
- Interest in music, sports, culture
- Vehicle ownership
- Personicx segment

48 variables

### LiveAnalytics Events (LAE)

- Tickets, spend, price, events
- Distance traveled
- Sales channel
- Delivery method
- Parking / group buyer
- Purchase period
- Concert, arts, family, sports events
- Recency, frequency, monetary scores
- Income / financial

(180 variables)
Email Domain

- .edu:
  - # of Leads: $17
  - Avg. New STH Revenue: $17

- .net:
  - # of Leads: $51
  - Avg. New STH Revenue: $51

- .org:
  - # of Leads: $25
  - Avg. New STH Revenue: $25

- [other].co.:
  - # of Leads: $123
  - Avg. New STH Revenue: $123

- aol:
  - # of Leads: $50
  - Avg. New STH Revenue: $50

- gmail:
  - # of Leads: $90
  - Avg. New STH Revenue: $90

- other:
  - # of Leads: $19
  - Avg. New STH Revenue: $19

- yahoo:
  - # of Leads: $43
  - Avg. New STH Revenue: $43
## Reference Predictions

<table>
<thead>
<tr>
<th>Data Population</th>
<th>Fixed Average</th>
<th>Lapsed Years</th>
<th>Cluster &amp; Logistic</th>
<th>Two-stage</th>
<th>3rd party</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical buyers and leads</td>
<td><strong>Historical buyers and leads</strong></td>
<td><strong>None</strong></td>
<td><strong>Lapsed years</strong></td>
<td><strong>Base + Acxiom + LA Events</strong></td>
<td><strong>Ticketing + Acxiom</strong></td>
</tr>
<tr>
<td><strong>Predictor Variables</strong></td>
<td><strong>Revenue</strong></td>
<td><strong>Revenue</strong></td>
<td><strong>Priority</strong></td>
<td><strong>Priority, Capacity, and Revenue</strong></td>
<td><strong>Priority and Capacity (binned)</strong></td>
</tr>
<tr>
<td><strong>Response Variable(s)</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>K-means clustering and logistic regression</strong></td>
<td><strong>Decision trees and rule-based models</strong></td>
<td><strong>Random forests and SVMs</strong></td>
</tr>
<tr>
<td><strong>Methodology</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>Averaging</strong></td>
<td><strong>Averaging</strong></td>
</tr>
</tbody>
</table>
Two-Stage Model Construction

Stages / Outcomes:

1. **Priority** – Likelihood to purchase season tickets – Decision tree
2. **Capacity** – Estimated season ticket spend if purchase a ticket – Rule-based model

**Expected Spend** = Priority * Capacity

Three models for three sets of predictor variables:

<table>
<thead>
<tr>
<th>Base</th>
<th>LAD</th>
<th>LAE</th>
<th>Model Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td></td>
<td></td>
<td>Base</td>
</tr>
<tr>
<td>☐</td>
<td>☐</td>
<td></td>
<td>Base + LAD</td>
</tr>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>Base + LAD + LAE</td>
</tr>
</tbody>
</table>
Two-Stage Base Model: Example Priority Tree
Two-Stage Base Model: Example Capacity Rule-Based

Model:

Rule 1: [329 cases, mean 1638.196, range 16 to 10000, est err 1149.617]
   if
       CI_EmailDomain in {aol, .edu, gmail, other, yahoo}
   then
       outcome = 862 + 868 STH_AvgFSE - 592 EML_TixImpOpnRt_LY
               - 59 STH_TotalYears

Rule 2: [250 cases, mean 3179.533, range 17 to 10000, est err 2279.088]
   if
       STH_AvgFSE <= 0.74074
       CI_EmailDomain in {.net, .org, [other].com}
   then
       outcome = 2465.344 - 2005 STH_AvgFSE - 1468 EML_OthImpOpnRt_LY
               - 11 EML_TixReceived_LY + 711 EML_OthOpenPer_LY
               - 367 EML_TixOpenPer_LY - 99 STH_TotalYears

Rule 3: [52 cases, mean 5133.067, range 360 to 10000, est err 2504.934]
   if
       STH_AvgFSE > 0.74074
       CI_EmailDomain in {.net, .org, [other].com}
   then
       outcome = 1985.245 + 584 STH_AvgFSE + 45 STH_AvgPrice
               - 745 EML_TixOpenPer_2Y - 1116 EML_OthImpOpnRt_LY
               + 540 EML_OthOpenPer_LY - 279 EML_TixOpenPer_LY
               - 76 STH_TotalYears
Two-Stage Base + LAD + LAE Model: Variable Importance
Business Impact
Model Performance: Absolute – Test Data

Mean Absolute Error

Mean Absolute Percent Error

Correlation
Model Performance: Rank – Test Data
## Model Performance: Revenue Outcomes

<table>
<thead>
<tr>
<th>Selecting leads by...</th>
<th>Led to an average...</th>
<th>Compared to...</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd party score</td>
<td>86% revenue increase</td>
<td>Random selection</td>
</tr>
<tr>
<td>Lapsed STH years</td>
<td>126% revenue increase</td>
<td>3rd party score</td>
</tr>
<tr>
<td>Model w/ Acxiom</td>
<td>8% revenue increase</td>
<td>Only internal data</td>
</tr>
<tr>
<td>demographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model w/ LiveAnalytics events</td>
<td>4% revenue increase</td>
<td>Internal + Acxiom data</td>
</tr>
<tr>
<td>Optimal internal model</td>
<td>51% revenue increase</td>
<td>Lapsed STH years</td>
</tr>
</tbody>
</table>
Prospector Implementation

Predicted Lead Revenue (log scale)

# 1 ranked lead  # ~730K ranked lead

Score
5 Star
4 Star
3 Star
2 Star
1 Star

Cleveland Indians
Recap

- **Lead Scoring** – 51% revenue increase
- **Fan Understanding** – reference documentation
- **Data Management** – 4% revenue for additional event data
- **Sales Management**
  - **Department Structure** – focus on very top customers
  - **CRM Efforts** – evaluation framework