Do You Know Who is Your Over-the-Counter Customer
October 16, 2014

Presented by:
Oscar Padilla, VP of Strategy
opadilla@luminarinsights.com
What We’ll Cover Today

i. Big Data and healthcare
ii. Why working with sample data is not good enough
iii. Amazing vast transactional data to derive value
iv. Utilizing retail transaction data to understand your customers
v. Insights that we can derive from transactional data
vi. Outlook and closing remarks
Big Data in Healthcare Stats...

16,000 hospitals worldwide collect data on patients.

An 18% annual compound growth rate is anticipated between 2010 and 2016 for patients that will use remote monitoring devices.

4.9 million patients worldwide will use remote monitoring devices by 2016.

80% of health data is unstructured and stored in hundreds of forms such as lab results, images, and medical transcripts.

Source: Berg Research and IBM
The Big Data revolution is in the early days, and most of the potential for value creations is still unclaimed.

But it has set the industry on a path of rapid change and new discoveries; stakeholders that are committed to innovation will likely be the first to reap the rewards.

Today’s Challenge in Gathering Insights into the Multi-Cultural Market

- Brands are making marketing investment decisions on limited information.
- Targeting assumptions based mostly on survey or sampled methods.
- Limited access to quantitative insights (i.e. “proximity or ethnic encoding methods”).
Sampling is like a Digital Photo…Insights Become Less Precise the Closer You Examine Your Data

Business decisions are inherently weakened if you solely rely on sampling methods.
Focus on Building the Largest Data Set Possible

- Consumer Habits
- Transactional Data (POS, CRM, loyalty e-commerce)
- Digital Media Interactions
- Relevant Analytical Models
- Cultural References

Traditional Sampled Approach

Luminar’s Latino Business Intelligence
So…What’s the Role of Big Data? Understanding the Basics

**Volume**
The sheer amount of data generated that must be ingested and analyzed.

**Velocity**
How fast data is being produced and changed and the speed in which data must be received, understood and processed.

**Variety**
Both structured and unstructured data generated by a wide range of sources.

**Value**
Derived benefits from data to deliver business impact and sustained value.
The “Data Lake” Drives the Analytical Knowledge

A Data Lake is a storage repository that holds a vast amount of raw data

- Data programmers can tap the stream data for real-time analytics
- The data lake accepts inputs from various sources
- Data model emerge with usage over time (machine learning)
- Data scientist use the lake for discovery and ideation
- Take advantage of cluster computing for massive scale & low cost
Ingesting Large Data Set to Derive Value

150 Million Unique Records

17 Million US Adult Latinos
Marketers Are Seeking Precise Answers to Fuel Growth and Increase Efficiencies

- How much have we earned through diverse promotional channels?
- How acculturated is my market?
- What's the Efficiency of our media activities?
- Which marketing drivers have had the greatest effects?
- What's the size of the prize in my trading area?
- Are we optimally allocating our budget across all products?
- How much have we earned through diverse promotional channels?
- What's my market share?
Healthcare Transactional Data Sources

**Activity (claims) and cost data**
- **Owners**: payors, providers
- **Example**: utilization of care, cost estimates

**Pharmaceutical R&D data**
- **Owners**: providers
- **Example**: electronic medical records, medical images

**Clinical data**
- **Owners**: pharmaceuticals, academia
- **Example**: clinical trials, high-throughput screening libraries

**Patient behavior and sentiment data**
- **Owners**: consumers, retailers and stakeholders outside healthcare
- **Example**: patient behaviors and preferences, retail purchase history, user-generated data

Integration of data pools for major opportunity
Patient Behavioral Data & Consumer Data

How we do it: We mash-up retail pharmacy data file against consumer databases and analyze those individuals to produce actionable insights.
A Closer Look at the Process

Luminar 150 million
Transaction datastore
- Zip + household data
- Transactional UPC-level data
- Social/demographics
- Population subsector
- Language of dominance

Retail Pharmacies
Transactional Data
- Item/UPC codes
- Product category
- Competitive landscape

Luminar Analytics Process:
- Data enrichment
- Customer segmentation

Luminar Analytics Process:
- Consumption Analysis
- Segmentation Analysis
- Targeting Solutions

Analysis Report

Luminar BI Portal
Once We Have the Right Data, We Design Solutions

**Consumption Analysis**

What is the Hispanic contribution to client’s business?

**Behavioral Segmentation**

Who are your Hispanic customers?

**Targeting Solution**

Where is a client most likely to win with Hispanic customers?
**Consumption Analysis: OTC of Allergy Medication**

**Objective:** Gain a better understanding of the consumption of OTC allergy medication among Asian-Americans in the New York DMA

**Methodology:** Produced volumetric and shopping behavior analysis using Luminar transactional data with an ethnic data overlay

- We identified Asians-Americans in Luminar transactional datastore
- Estimated the total 2013 consumption volume and dollars for three major allergy OTC medications
- Estimated the average frequency of purchase and latency times between purchases
- Calculated the average spend per buying occasion
- Ran growth scenarios to identify sources of growth
## Total Market Overview of Allergy OTC Medication

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total market sales 2013</td>
<td>$24 Million</td>
</tr>
<tr>
<td>Total customers</td>
<td>0.9 Million</td>
</tr>
<tr>
<td>Penetration (top 3 brands)</td>
<td>6.3%</td>
</tr>
<tr>
<td>Average spend per customers</td>
<td>$27.2</td>
</tr>
</tbody>
</table>

Only 6% of the NY DMA adult population consumes major Allergy Brands (Claritin, Allegra, Zyrtec)

*Source: 2014 Luminar analysis. Allergy OTC Asian consumption*
Asians Underperform in the NY DMA on the Consumption of OTC Allergy Medication

Nevertheless, there is an opportunity for market growth among Asians...

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian market sales 2013</td>
<td>$1.1 Million</td>
</tr>
<tr>
<td></td>
<td>(4.8%)</td>
</tr>
<tr>
<td>Asian customers</td>
<td>48,000</td>
</tr>
<tr>
<td></td>
<td>(5.4%)</td>
</tr>
<tr>
<td>Asian population in NY DMA</td>
<td>7%</td>
</tr>
<tr>
<td>Penetration (3 top brands)</td>
<td>4.8%</td>
</tr>
<tr>
<td>Average spend per Asian customers</td>
<td>$23.8</td>
</tr>
</tbody>
</table>

Source: 2014 Luminar analysis. Allergy OTC Asian consumption
Only Caucasians Over Perform in the NY DMA In the Consumption of OTC Allergy Medication

**Revenue Mix**

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>Asian</th>
<th>Hispanic</th>
<th>Caucasian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>3.5%</td>
<td>4.8%</td>
<td>7.4%</td>
<td>84.3%</td>
</tr>
</tbody>
</table>

**Customer Mix**

<table>
<thead>
<tr>
<th></th>
<th>AA</th>
<th>Asian</th>
<th>Hispanic</th>
<th>Caucasian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>4.8%</td>
<td>5.4%</td>
<td>7.8%</td>
<td>82.0%</td>
</tr>
</tbody>
</table>

Customer to Revenue Index

|        | 74  | 88   | 95     | 103      |

Source: 2014 Luminar analysis. Allergy OTC Asian consumption
Claritin leads the market of the top 3 brands with 39%.
Nevertheless Allegra is the #1 brand among African Americans.
Each brand has its best performance in different ethnic markets.

<table>
<thead>
<tr>
<th>BRAND</th>
<th>AFRICAN AMERICAN</th>
<th>ASIAN</th>
<th>HISPANIC</th>
<th>CAUCASIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allegra</td>
<td>115</td>
<td>106</td>
<td>106</td>
<td>99</td>
</tr>
<tr>
<td>Claritin</td>
<td>94</td>
<td>103</td>
<td>106</td>
<td>100</td>
</tr>
<tr>
<td>Zyrtec</td>
<td>91</td>
<td>88</td>
<td>83</td>
<td>103</td>
</tr>
</tbody>
</table>

Market Share by Ethnicity

AFRICAN AMERICAN: 39%, 37%
ASIAN: 36%, 41%
HISPANIC: 36%, 42%
CAUCASIAN: 33%, 39%, 28%
Behavioral Segmentation: Cluster Analysis

- We disaggregated Hispanics in Luminar’s datastore within given zip code trade areas
- We analyzed consumption in the major product categories, including:

<table>
<thead>
<tr>
<th>BEAUTY-ACCESSORIES</th>
<th>PERSONAL CARE/INTIMATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEAUTY-COSMETICS/BEAUTY AIDS</td>
<td>HEALTH-NUTRACEUTICALS &amp; VITAMINS</td>
</tr>
<tr>
<td>HEALTH-MEDICAL SUPPLIES</td>
<td>HEALTH-FEMALE WELLNESS</td>
</tr>
<tr>
<td>PERSONAL CARE/WOMEN</td>
<td>HEALTH-NEW AGE</td>
</tr>
<tr>
<td>PERSONAL CARE/MEN</td>
<td>BEAUTY-PHYSICAL ENHANCEMENT</td>
</tr>
</tbody>
</table>

- Developed behavioral segmentation to identify groups of Latinos with similar consumption patterns
- Developed different personas based on ACTUAL transactional behaviors
**Cluster 1**: Young and active, Spanish-Dominant women with medium household income and high consumption in personal care, beauty, cosmetics, beauty aids, and health and nutritional products.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency</th>
<th>Cluster 1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>65%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>35%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Age Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-35</td>
<td>49%</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td>36-55</td>
<td>32%</td>
<td>39%</td>
<td></td>
</tr>
<tr>
<td>56-More</td>
<td>19%</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>Household Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (&lt;$30k)</td>
<td>25%</td>
<td>26%</td>
<td></td>
</tr>
<tr>
<td>Medium ($30k-$50k)</td>
<td>38%</td>
<td>24%</td>
<td></td>
</tr>
<tr>
<td>High (&gt;=$50k)</td>
<td>37%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>61%</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>39%</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Presence of children in HH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>39%</td>
<td>41%</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>61%</td>
<td>59%</td>
<td></td>
</tr>
<tr>
<td>Language of Preference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>21%</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>79%</td>
<td>73%</td>
<td></td>
</tr>
</tbody>
</table>
**Cluster 2**: Married mature men, conscientious and generous, half without children in HH. Average HH income, heavy altruistic behavior and high consumption in health – medical supplies and personal care

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cluster 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># Individuals</td>
<td>731.133</td>
<td>3.141.241</td>
</tr>
<tr>
<td>Size</td>
<td>23%</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency</th>
<th>Cluster 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female</td>
<td>39%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>61%</td>
<td>50%</td>
</tr>
<tr>
<td>Age Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18-35</td>
<td>36%</td>
<td>42%</td>
</tr>
<tr>
<td></td>
<td>36-55</td>
<td>35%</td>
<td>39%</td>
</tr>
<tr>
<td></td>
<td>56-More</td>
<td>29%</td>
<td>19%</td>
</tr>
<tr>
<td>Household Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (&lt;$30k)</td>
<td>20%</td>
<td>26%</td>
<td></td>
</tr>
<tr>
<td>Medium ($30k-$50k)</td>
<td>30%</td>
<td>24%</td>
<td></td>
</tr>
<tr>
<td>High (&gt;=$50k)</td>
<td>50%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>Marital Status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>74%</td>
<td>67%</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>26%</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Presence of children in HH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>51%</td>
<td>41%</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>49%</td>
<td>59%</td>
<td></td>
</tr>
<tr>
<td>Language of Preference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>25%</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>75%</td>
<td>73%</td>
<td></td>
</tr>
</tbody>
</table>
Targeting Solutions: Identifying Lookalikes and “Act alikes”

- Leverage large transactional data sets to first identify behavioral characteristics
- Develop propensity models of your best customers
- Apply these models to develop
  - Look Alikes: in path-to-purchase
  - Act alikes: have ideal characteristics but are not in path to purchase
- Close the loop by finding these targets offline and online

Diagram:
- High Value Prospects
- In market consumers (lookalikes)
- High value prospects audience (act alikes)

Total Market
Identifying High-Value Prospect

Known Customers

Non-Customers

Propensity Model

Scoring Process

High Value Prospects After Scoring

High

Med

Low
Multi-channel Targeting Solution

**Your First Party Data**
- Ingest Client CRM Data

**Luminar Offline Data Environment**
- 140 Million US adults with 1200+ data attributes

**Custom Modeling**
- Formulate appropriate models based on your needs

**Identify High-value Prospects**
- Score Luminar Data store
- Customized Audience Reach

**Targeting Channels Based on Behavioral Models**

**High Value Prospects After Scoring**

Client CRM Data Set
The Healthcare Ecosystem & Future Opportunities

Source: Endeavour Partners 2014
Three Key Closing Remarks…

- Marketing to the multicultural sector is not just about language, acculturation or relevancy; it’s about having precise **measurability that can prove efficiencies and ROI**

- Clients are sitting on a **wealth of valuable data assets** – it’s important to identify those data assets that can be used to support marketing and growth strategies

- It not just about focusing what we know about customers, the true opportunities come from seeing something you never seen before…**understanding the unknowns**
Two Great TED Talk Speakers on Data & Healthcare

**John Wilbanks**
Data Commons Advocate

🔗 Website: WeConsent.us  📝 Blog: Del-fi

**Joel Selanikio**
Health and technology activist

🔗 Company: DataDyne  📥 LinkedIn: Joel Selanikio  🌐 Twitter: @iselanikio
Thank you

Oscar E. Padilla
Vice President of Strategy
opadilla@luminarinsights.com