From Cluster To Counter

A Segmented Approach To Planogram Design

CCSWB – ESTEFANÍA ELIZONDO; XIMENA ZUNIGA, JESSICA JOHNSON
MIT ADVISOR – Prof. STEPHEN GRAVES

Andrea Valacchi

Tommaso Serafin

Problem Statement

Small Regional Key Accounts

Undifferentiated Standard Planogram

Unviable to be store-specific

Art & Science

Optimal Planogram

- Tailored to common store specificities
- Streamlined creation process
- Economically viable level of granularity
- Compatibility with current planograms

Data

Historic Sales data

Product Attributes

Customers

Profitability

Projected Census Data

Income

Education

Age

Pilot set of **157** Regional Key Accounts

Methodology

Order Of Entry Per Store

1. **Historical Sales Data**
   EQ
   COGS
   Store
   Brand
   Format
   Size
2. **Velocity Per SKU**
   Derived metric
   $\frac{\text{Sales}}{\text{Fixed time frame}}$
3. **Order Of Entry**
   *"**Prioritized sequence** in which **products should be allocated** shelf space within a planogram"*

New Metric: Profitability ***($/Unit)*** X Velocity ***(Units/Week)*** = **Profit per week** ***$/Week***

Currently used in creating Planograms

*Planograms are now designed with **profitability** in mind, **not only velocity***

Order Of Entry Per Segment

OOE per Store + **Cluster Assignment** → **OOE per Segment**

Wisdom Of The Crowds

Tradeoff: **Optimal** vs **Robust**

Profit-Based Ranking	VS	**Ranked Majority Voting**
✓ *Maximizes returns*		✓ *Robust to Outliers*
✗ *Is highly sensible to outliers*		✗ *Profit is not maximized*

No clear winner, so both methods are retained in the final pipeline

Segmentation

Data Preprocessing

Top-Down Processing

Census Data

Selected 18 features out of the 565 available

Shopper Demographics

Bottom-Up Processing

Historical Sales Data

Engineered composite features to capture trends

Relationship with CCSWB

Model Selection

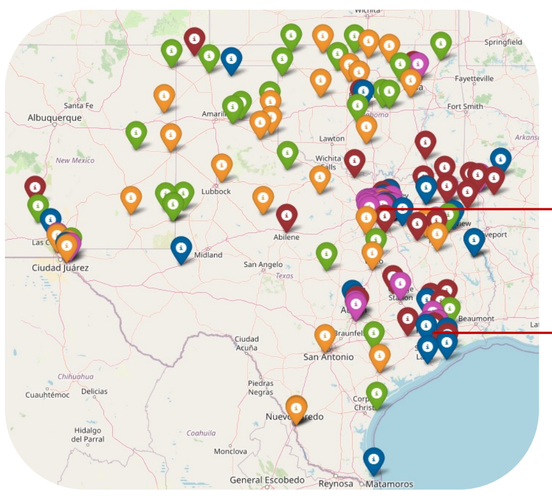
*Clusters were interpreted by fitting **decision trees** on the features to recover cluster assignments*

K-Means Clustering	VS	**Gaussian Mixture Models**
✓ *Simplicity and Speed*		✓ *Flexible cluster shape*
✗ *Assumes spherical clusters*		✓ *Probabilistic Assignment*
✗ *Binary Assignment*		✗ *Computationally Intensive*

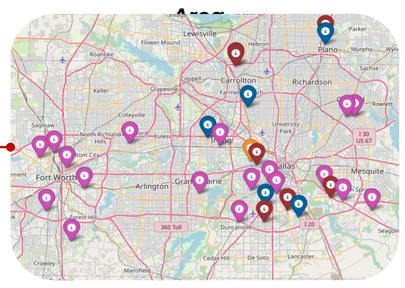
Produces clusters better aligned with business knowledge

Segmentation

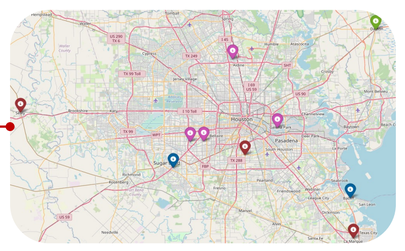
 # Clusters = 5



*Focus On: **Dallas Fort-Worth Area***



*Focus On: **Houston Area***



Impact

Direct Impact

+ $1.4 Million

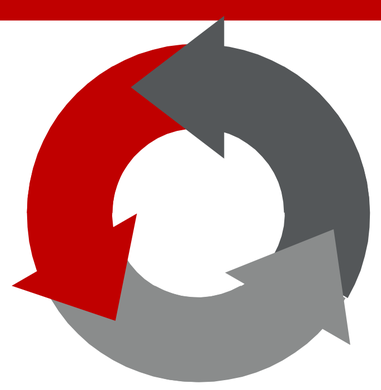
+ 15%

In Gross Profit Across the 157 Stores

Indirect Impact

Modular Pipeline

- ✓ Modular design to ensure integration with current and future systems
- ✓ New Inputs can be provided without hindering scalability



Project Spillover

- ✓ Sociodemographic features now available at the census Tract level
- ✓ Developed best practices for Geospatial data manipulation

Knowledge Transfer Sessions

- ✓ Avoid the loss of acquired informal knowledge

Standard Planogram