With its wide range of customers through different trade channels, geographica

tions, and product portfolios, CCSWB aims to create tailored service experiences for
its customers by understanding their characteristics. Therefore, sales potential is introduced to measure the
unrecognized room-to-grow at a customer.

How much more can we sell?

It is a win-win solution for both CCSWB and its customers because it drives extra sales for both sides.

Problem

Objective

• Identify internal & external factors that can affect sales potential
• Calculate a monthly, customer-level true sales potential by product category
• Provide guidance to optimize sales and resource allocation

Methodology

1. Segmentation: The data are first filtered to 26 trade channels that CCSWB already defined.
2. Clustering: For each trade channel, we use k-means clustering to group similar customers together based on features.
3. K-Nearest Neighbors: We find the closest "neighbors" of a target customer based on the difference of features within its assigned cluster.
4. Potential Calculation: The potential is the gap of forecast from the target customer to the weighted average of its "neighbors".

Prediction

To accurately estimate potential, we need an accurate forecast of sales volume. During data exploration, we observe very distinct purchasing habits across channels and beverage categories. Therefore, our ensemble model tackles this problem by taking consideration of sales propensity.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code Outline</td>
<td>1st On-site Visit</td>
<td>2nd On-site Visit</td>
<td>Hand-over</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st Revision Accuracy</td>
<td>2nd Revision Methodology</td>
<td>3rd Revision Validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Diag 1: Impact of Segmentation

Diag 2: Impact of Clustering

Diag 3: Impact of K-Nearest Neighbors

Diag 4: Impact of Potential Calculation

Validation

1. Back-testing: palo pinto (Jan 2020)
   - Sales opportunity captured: +61% (6,350 vs. 10,150)
   - 93% supported plan

2. On-Shelf-Availability (OSA) vs. Potential
   OSA = % of items that are in-stock on the shelf
   - Low OSA = high potential
   - High OSA = low potential

   Assumption
   - Low OSA = high potential
   - High OSA = low potential
   - (Demand not met, we can sell much more!)

   Results
   - Correlation can be proved in several channels.
   - Limitations: many other factors also affect OSA.

Result

Our results are very promising. We compared the OS2 across all three models. It turned out that ensemble model outperforms the other two (or equivalent) in all trade channels.

Impacted Customers: 13,700

Correctly Identify: 99%

Out-of-Sample R-squared: +37.9% than baseline

Out-of-Sample Area Under Curve: +29.2% than baseline

Future Steps

- Flexibility for add-on constraints
- Optimizes resource, considers profitability
- Pilot launch / Proof of Value
- Automation / Run-time Efficiency
- Internal CIMA System integration