Demand Forecasting to Increase Profits on Perishable Items

Ankur Pandey
Arun Chaubey
Sanchit Garg
Shahid Siddiqui
Sharath Srinivas

Forecasting Analytics, ISB
GOAL: Maximize profits when selling perishable items such as fruits and vegetables
Data:
• Hypermart customer Transaction data from 8/1/2011 to 8/31/2012
• Each transaction includes the customer ID, SKU and purchase quantity
• 5 SKUs were explored: Banana, Apple, Onion, Tomato, Papaya

Caveats:
• Transaction data includes only loyalty card purchases
• Data does not include promotions
• Data includes only customer demand and not indicate inventory levels, procurement etc.
• Infrequent visits by customers to the store
Hypothesis: There is weekly seasonality
Higher demand on:
- Sunday
- Saturday
- Wednesday
**Goal**

**Data**

Explore & Visualize

Pre-process

Forecasting

Evaluate

---

**Aggregate daily customer transaction volumes**

---

**Remove Outliers**

---

**Partition – Training and Validation**

---
Naïve Model: Forecast for Aug 2012

- Naïve model can only be used as a benchmark
- Accuracy of the model is very low

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>22.632</td>
</tr>
<tr>
<td>Average Error</td>
<td>1.004903</td>
</tr>
<tr>
<td>MAPE</td>
<td>50.15%</td>
</tr>
<tr>
<td>RMSE</td>
<td>791.4196</td>
</tr>
</tbody>
</table>
**Linear Regression Model (With no Dummy Variables): Forecast for Aug 2012**

- **MAE**: 21.26407
- **Average Error**: 16.19199
- **MAPE**: 35.33%
- **RMSE**: 896.3684

**ACF Plot for Residual**

- ACF Plot shows that there is seasonality left in the residuals
**Goal**

**Data**

**Explore & Visualize**

**Pre-process**

**Forecasting**

**Evaluate**

---

**Linear Regression Model (With Dummy Variables): Forecast for Aug 2012**

- **ACF Plot** shows that there is seasonality left in the residuals.

- **MAE**: 19.8724
- **Average Error**: 17.33157
- **MAPE**: 36.46%
- **RMSE**: 552.8134

---

**ACF Plot** shows that there is seasonality left in the residuals.
Holt Winters: Forecast for Aug 2012

Time Plot of Actual Vs Forecast (Validation Data)

- **ACF Plot for Error**
  - ACF
  - Lags: 0, 1, 2, 3, 4, 5, 6, 7
  - ACF values: 0, 0.5, 1

- **Evaluation Metrics**
  - **MAE**: 15.04821
  - **Average Error**: 2.589459
  - **MAPE**: 34.84%
  - **RMSE**: 350.3683

- **Seasonality handled by Holt Winters method**
- Sales of individual SKUs categorized on Day of the week is noisy
- Instead we forecast amount of Footfall in the store
- Use footfall as a proxy to forecast the SKU quantity demand
Step 1: Forecast Footfall
Step 2: Forecast Sales of SKU 1000

y = 0.3095x + 2.741
R² = 0.5501

Series1

-- Linear (Series1)
2 Staged Model: Forecast for Aug 2012

- **MAE**: 13.27165
- **Average Error**: -2.95764
- **MAPE**: 33.33%
- **RMSE**: 286.4652
**Two step process:**

1. **Determine a cost-metric** (e.g. profit: See Appendix for profit calculations)

2. **Evaluate the effect of different forecasting methods on the metric**

---

**Achieved 17% improvement** in profitability by **leveraging** advanced forecasting techniques at SKU level (compared to baseline-naïve forecasting)
Recommendations

• By forecasting demand at the SKU level, the store can increase profitability by:
  – Reducing wastage
  – Reducing lost sales

• Two-staged model offers best performance in terms of profitability improvement

• **Key Learning:** Handling noisy data, “Torture the data, and it will confess to anything!”
A Model of System Mechanics

- Suppose $Q = 100$
  - suppose demand turns out to be $D = 110$
  - profit = $70,000(100) - $55,000(100) = $1,500,000$

- Now suppose $Q = 100$ and demand turns out to be $D = 90$
  - profit = $70,000(90) + $20,000(10) - $55,000(100) = $1,000,000$

- The General Formula:

<table>
<thead>
<tr>
<th>If $D \geq Q$</th>
<th>Sales</th>
<th>Salvage</th>
<th>COGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P \times Q$</td>
<td>0</td>
<td>$C \times Q$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>If $D &lt; Q$</th>
<th>Sales</th>
<th>Salvage</th>
<th>COGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P \times D$</td>
<td>$S \times (Q - D)$</td>
<td>$C \times Q$</td>
<td></td>
</tr>
</tbody>
</table>

- As a single formula...
  - profit($Q, D$) = $IF(D < Q, (P \times D) + S \times (Q - D) - (C \times Q), (P \times Q) - (C \times Q))$