Journey from Descriptive Analytics to Predictive Analytics

Balaji Sundaram, BI Analyst, Benjamin Moore & Co
Sree Rajitha Indraganti, Lead BW Analyst, Benjamin Moore & Co
Session ID # 83598
About the Speakers

Speaker Name
• Balaji Sundaram - Business Intelligence Analyst
• 5 years of experience in BI
• Certified Data Scientist

Speaker Name
• Sree Rajitha Indraganti – Business Warehouse Lead
• 10 years of experience in SAP BW
Agenda

• Introduction
• Reporting System Architecture
• Descriptive to Predictive Analytics
• Benefits Realized
• Roadmap
Benjamin Moore & Co. Company Overview

• Benjamin Moore & Co., a Berkshire Hathaway company, was founded in 1883.
• One of North America's leading manufacturers of premium quality residential, commercial and industrial maintenance coatings, Benjamin Moore & Co. maintains a relentless commitment to innovation and sustainable manufacturing practices.
• The Benjamin Moore premium portfolio spans the brand’s flagship paint lines including Aura, Regal Select, Natura and ben. The Benjamin Moore & Co. family of brands includes specialty and architectural paints from Coronado, Corotech, Lenmar and Insl-x.
• Benjamin Moore & Co. coatings are available primarily from its more than 5,000 locally owned and operated paint and decorating retailers.
Benjamin Moore & Co. Company Overview

“Benjamin Moore’s primary goal: turn out the best paint in the world and have the best retailer organization in the world”

- Warren Buffett, Berkshire Hathaway

Benjamin Moore has ranked “Highest in Customer Satisfaction with Interior Paints Four Years in a Row”. 
Wonder What Your Customer really wants

Give them before they could ask
Modern BI and Analytics Platforms

Today 3 to 5 Years

Pervasive Autodescriptive Diagnostic Predictive, Prescriptive

IT-Led Descriptive

Business-Led Descriptive/Diagnostic

Semantic Layer-Based Platforms

Visual-Based Data Discovery Platforms

Pervasiveness of ML-Enabled Advanced Insight on All Data

Time to Advanced Insight

Source: Gartner
From Traditional Reporting Process

QUERY ➔ DATABASE (BW/HANA) ➔ RESULT
Reporting Process Towards…

- Past Data
- New Data
- Training
- Model
- Result
- Action
Old Vs New Architecture

Old

Informatica

- Oracle
- CRM
- Ecommerce
- DB2
- Flat Files

Stage

- Hist
- EDW Oracle
- Non-EDW Oracle Loads

New

SAP BW/HANA

- Stage
- CRM
- Ecommerce
- DB2
- Flat Files

SAP Data Services / Workbench

- Stage
- SAP HANA
- Non-EDW Oracle Loads
- Sybase
Reporting System Architecture – Pre SAP Initiative

- Reporting Layer
  - SAP BusinessObjects BI Enterprise 4.2 SP03
    - Web Intelligence
  - Oracle
    - Data Marts
    - Tables and Views
    - Staging
  - Flat Files
  - Legacy System
  - 3rd Party Systems

- Data Warehouse Layer

- Source System Layer

---

Benjamin Moore & Co.
SINCE 1883
Project Delivery – Tracking / Keeping Pace

• Project life cycle
  – Phase 1 (SAP Data)
    • Began July 2015
    • Sprint 1 go live Jan 2016
    • Sprint 2 Phase 1 go live Jan 2017
    • Sprint 2 Phase 2 go live Dec 2017
  – Phase 2 (Non SAP data)
    • Began October 2017
    • Sprint 1 go live April 2018
    • Sprints 2 and 3 go live May 2018
    • Sprint 4 to be completed June 2018

• Delivered using Agile methodology for SAP BW and Scrum approach for HANA

• Each sprint had a series of associated RICEFs
  – Each RICEF had a series of associated tasks/ effort
  – Delivered End to End to a final Business Objects report

• Daily standups with visual progress tracking
Benjamin Moore & Co. embarked on a business transformation effort – **Prism** – with a multiphase implementation of SAP ECC, beginning in 2014, with these primary objectives:

- Advance and standardize business processes to support future growth
- Enable enhanced analytics and data-driven decision-making

**Source System Layer**
- SAP ECC 6.0 SP07 on HANA
  - Standard and Generic Data Sources
  - FM Extractors
  - ABAP Dictionary Tables
- SAP BW 7.4 SP11 on HANA
  - BEx Queries
  - Composite Providers
  - Open ODS Views
  - Advanced DSOs
  - InfoObjects

**Data Warehouse Layer**
- HANA Enterprise Cloud

**Reporting Layer**
- SAP Business Objects BI Enterprise 4.2 SP03
  - Web Intelligence
- Dashboards and Visualizations
  - QlikSense

**Additional Systems**
- Oracle
  - Data Marts
  - Tables and Views
  - Staging
- Flat Files
- Non SAP
- 3rd Party Systems
- On Premise

**Reporting Systems**
- SAP Business Objects BI Enterprise 4.2 SP03
- QlikSense

Reporting Flow

• BW BEx Queries used as foundation for WebI reports, via BICS connection, eliminating the need for Universes

• BOBJ Web Intelligence as the reporting front end interface for all pre-built and ad-hoc user reporting

• WebI integrated with third-party tool to provide flexible and dynamic broadcasting of reports

• Ad-hoc: Users can copy and modify reports in personal folders or create new reports
## DSR109 BMNA Daily Sales (Year-To-Date)

### US & Canada Daily Sales Report as of Tuesday, 7/18/18 • Year-to-Date (TY vs LY)

<table>
<thead>
<tr>
<th>Category (US Calendar)</th>
<th>TY</th>
<th>LY</th>
</tr>
</thead>
<tbody>
<tr>
<td>West &amp; Central CA</td>
<td>1.3% YTD</td>
<td>1.3% YTD</td>
</tr>
<tr>
<td>Eastern CA</td>
<td>2.1% YTD</td>
<td>2.1% YTD</td>
</tr>
<tr>
<td>Great Lakes</td>
<td>2.7% YTD</td>
<td>2.7% YTD</td>
</tr>
<tr>
<td>Rocky Mtns/TX</td>
<td>4.2% YTD</td>
<td>4.2% YTD</td>
</tr>
<tr>
<td>Specialty Coatings</td>
<td>-5.4% YTD</td>
<td>-5.4% YTD</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.7% YTD</td>
<td>0.7% YTD</td>
</tr>
<tr>
<td>New York Metro</td>
<td>-2.2% YTD</td>
<td>-2.2% YTD</td>
</tr>
<tr>
<td>Mid Atlantic</td>
<td>3.8% YTD</td>
<td>3.8% YTD</td>
</tr>
<tr>
<td>South East</td>
<td>6.7% YTD</td>
<td>6.7% YTD</td>
</tr>
<tr>
<td>International</td>
<td>24.4% YTD</td>
<td>24.4% YTD</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>TY Gallons</th>
<th>Chg %</th>
<th>Chg %</th>
<th>TY Dollars</th>
<th>Chg %</th>
<th>Chg %</th>
<th>TY PPG</th>
<th>Chg %</th>
<th>Chg %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>25,482,028</td>
<td>1.3%</td>
<td></td>
<td>599,424</td>
<td>6.0%</td>
<td></td>
<td>2,319</td>
<td>4.3%</td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>21,254,903</td>
<td>1.1%</td>
<td></td>
<td>567,039</td>
<td>5.3%</td>
<td></td>
<td>22,12</td>
<td>4.3%</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>3,495,670</td>
<td>0.5%</td>
<td></td>
<td>91,201</td>
<td>7.6%</td>
<td></td>
<td>20,90</td>
<td>0.1%</td>
<td></td>
</tr>
<tr>
<td>Int'l</td>
<td>670,754</td>
<td>26.6%</td>
<td></td>
<td>17,193</td>
<td>27.0%</td>
<td></td>
<td>20,69</td>
<td>4.9%</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>27,047</td>
<td>2.1%</td>
<td></td>
<td>661</td>
<td></td>
<td></td>
<td>14,97</td>
<td>(0.4%)</td>
<td></td>
</tr>
</tbody>
</table>

- TY Gallons: Total Gallons sold during the year-to-date.
- Chg %: Change in Gallons from LY.
- TY Dollars: Total Dollars sold during the year-to-date.
- Chg %: Change in Dollars from LY.
- TY PPG: Total PPG (Price Per Gallon) during the year-to-date.
- Chg %: Change in PPG from LY.
Challenges

• Change Management
Challenges

• Data Literacy

Business

• Knows what the data means
• Knows how to interpret the data to make decisions
• DOESN'T know how to think about data technically (organize, classify, etc.)

IT

• Knows how to build and manage data systems
• Knows how to build and update reports
• DOESN'T know what the data means or what decisions should be made from it
• More Training Sessions
• Increased number of Powers users
• Increased number of Business users
• Faster reporting than before
• Additional IT resources
– Self Service Analytics was getting matured
– Descriptive Analytics and Ad-Hoc reporting increased
Challenges

• Heavy Usage of Excel
Challenges

• Heavy Usage of Excel

• Inconsistency in Metrics

Anyone and Everyone can manipulate the data
Different People were getting different results

Can we trust our data?
Are our conclusions accurate?

UN-FUN FACT:
ONLY 51% of business leaders surveyed reported having the necessary information to feel confident in making decisions.

Challenges

• Heavy Usage of Excel
• Inconsistency in Metrics
• Data Trust

UN-FUN FACT:
ABOUT 90% of spreadsheets contain errors.

Modern BI and Analytics Platforms

- **Today**: Visual-Based Data Discovery Platforms
- **3 to 5 Years**: Pervasive Autodescriptive Diagnostic Predictive, Prescriptive

**Pervasiveness of ML-Enabled Advanced Insight on All Data**
- **Low**
  - IT-Led Descriptive
  - Business-Led Descriptive/Diagnostic
  - Pervasive Autodescriptive Diagnostic Predictive, Prescriptive

**Time to Advanced Insight**
- **Months**: Semantic Layer-Based Platforms
- **Days/Hours**: Visual-Based Data Discovery Platforms
- **Instant/In-Line**
Predictive Analytics

• First need from Business
  o Forecast the need for Raw Materials

• Business Involved
  o Procurement and Supply Chain

• Partially Satisfy the requirement using - Excel and SAP BO
Challenges

- Not dynamic in generating results
- Not powerful enough to handle large data sets
- Capable of performing only algorithms
Business users realized the need and advantages of predictive Analytics

Preliminary Evaluation and POC’s
  - SAP Predictive Analytics
  - Text Analytics via HANA Libraries
  - Text Analytics via HDInsight and Qliksense
Reporting System Architecture – Phase 2

Reporting Layer
- SAP Business Objects BI Enterprise 4.2 SP03
  - Web Intelligence

Data Warehouse Layer
- SAP BW 7.4 SP11 on HANA
  - BEx Queries
  - Composite Providers
  - Open ODS Views
  - Advanced DSOs
  - InfoObjects
- SAP ECC 6.0 SP07 on HANA
  - Standard and Generic Data Sources
  - FM Extractors
  - CDS Views
  - ABAP Dictionary Tables

Source System Layer
- SAP HANA Enterprise
  - Calculation Views
  - DLM
  - HOT DATA
- IQ
  - Cold Data

Dashboards and Visualizations
- QlikSense

Flat Files
- Non SAP

3rd Party Systems
- HANA Enterprise Cloud
Benefits Realized

• High Performance: power of HANA for better performance
• Integration: SAP and Non-SAP on a single HANA platform for better data integration, support and maintenance
• Predictive and agile analytics with historical sales data
• Improved use of previously unused data
• Multi temperature data management using DLM
• Known Customer Needs
  o Publish forecast results in DSR App
  o Publish retailer churn analysis SMD App

• Proof of Concept
  o Real Time Predictive Analytics
  o Integrated R with Qliksense
  o Integrated Rapidminer with Qliksense
  o Integrated HANA with R & Rapidminer

• Business Unit
  o Presented Live Demo to FP&A and Pricing Team
Data Scientist

Database (HANA, Azure, Big Data)

Applications

Advanced Analytics

QlikSense

Social Media Analysis
Retailer/Contractor Churn Analysis
Forecast Analysis

R

Trained Models
New Customized Models (through APIs)

End User

Developer

Data Scientist


Forecast Analysis

Applications

Advanced Analytics

QlikSense

Social Media Analysis
Retailer/Contractor Churn Analysis
Forecast Analysis

R

Trained Models
New Customized Models (through APIs)
Modern BI and Analytics Platforms

Today

3 to 5 Years

IT-Led
Descriptive

Business-Led
Descriptive/Diagnostic

Pervasive
Autodescriptive
Diagnostic
Predictive, Prescriptive

Semantic Layer-Based Platforms

Visual-Based Data Discovery Platforms

Pervasiveness of ML-Enabled Advanced Insight on All Data

Time to Advanced Insight

Months

Days/Hours

Instant/In-Line

IT-Led
Descriptive
Business Requirement for AI

- Retailer Churn
- Price Elasticity
- Forecast Gallons
Let's look at some numbers.

8 Stores closed after renovating – We could have saved $175k
Let's look at some Numbers

8 Stores closed after renovating – We could have saved $175k
1.25% Gallons were lost due to the churned customers
Let's look at some numbers.

8 Stores closed after renovating – We could have saved $175k

1.25% Gallons were lost due to the churned customers

$3M would have been the profit margin from all discontinued stores
First Predictive Model

• Develop the model within Enterprise
• Predictive Model – R/ Python
• Semi-supervised Learning
  o Clustering the Retail Outlets (Unsupervised Learning)
  o Random Forest for building the tree (Supervised Learning)
• AI Platform – SAP Leonardo
• Visualization – Qliksense
First Predictive Model

Churn Model

- Promotions
- CRM
- Demographics/External
- Accounts Receivable
- Sales

Tidbits
- U.S. companies lose $136.8 billion per year due to avoidable consumer switching. - CallMiner

Tidbits
- Churn can increase by up to 15% if businesses fail to respond to customers over social media. - Gartner

First Predictive Model
Confusion Matrix – Preliminary Results based on data through 2017

<table>
<thead>
<tr>
<th></th>
<th>Actual Discontinue</th>
<th>Actual Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Discontinue</td>
<td>34</td>
<td>52</td>
</tr>
<tr>
<td>Predicted Active</td>
<td>7</td>
<td>1742</td>
</tr>
<tr>
<td>Total</td>
<td>41</td>
<td>1794</td>
</tr>
</tbody>
</table>

Label
No – Discontinue
Yes - Active
The work in progress model predicted 74% of Churned “Paint and Decorating” Retailers out of all the churned “Paint and Decorating” Retailers in 2018.
The *work in progress* model predicted 74% of Churned “Paint and Decorating” Retailers out of all the churned “Paint and Decorating” Retailers in 2018.

<table>
<thead>
<tr>
<th></th>
<th>Actual Discontinue</th>
<th>Actual Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted Discontinue</td>
<td>74%</td>
<td>0</td>
</tr>
<tr>
<td>Predicted Active</td>
<td>26%</td>
<td>0</td>
</tr>
</tbody>
</table>

$3.75 \text{ M}$
Where are we heading to...
Our Vision

- **2015**
- **2016**
- **2017**
- **2018**
- **2019**
- **2020**

**Modern BI and Analytics Platforms**

Today

3 to 5 Years

- **Pervasive Autodescriptive**
- **Diagnostic**
- **Predictive, Prescriptive**

**IT-Led Descriptive**

- **Months**

**Business-Led Descriptive/Diagnostic**

- **Days/Hours**

**Pervasiveness of ML-Enabled Advanced Insight on All Data**

- **Source: Gartner**
Q&A

For questions after this session, contact us at

Balaji Sundaram
Email: Balaji.Sundaram@benjaminmoore.com
LinkedIn: https://www.linkedin.com/in/balaji-sundaram-994b4665/

Sree Rajitha Indraganti
Email: Sree.rajitha@benjaminmoore.com
LinkedIn: https://www.linkedin.com/in/sree-rajitha-indraganti-7430158a/
Take the Session Survey.

We want to hear from you! Be sure to complete the session evaluation on the SAPPHIRE NOW and ASUG Annual Conference mobile app.
Presentation Materials

Access the slides from 2019 ASUG Annual Conference here:

http://info.asug.com/2019-ac-slides
Let’s Be Social.

Stay connected. Share your SAP experiences anytime, anywhere.
Join the ASUG conversation on social media: @ASUG365 #ASUG