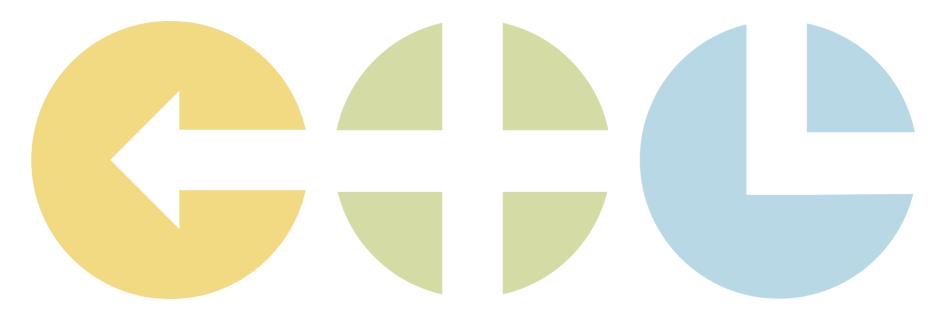
# Enhancing S&OP Performance with Analytics



Presenters: *Deepti Kidambi & Minhaaj Khan* Advisor: *Dr. Tugba Efendigil* 

#### Agenda

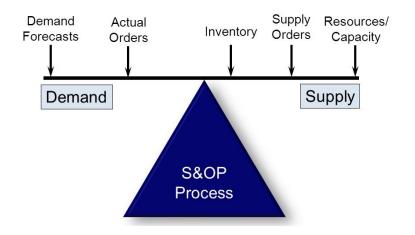
- Introduction
- Key Research Questions
- Methodology
- Results
- Conclusion



Introduction

#### S&OP process and risks to SC

#### Nuances of the CPG Industry

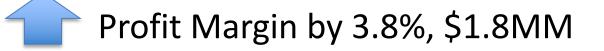






#### Introduction

• Sponsor and scope









### **Key Research Questions**

Can predictive analytics models effectively *predict risk patterns* in the S&OP plan?



How much can these models improve *consensus forecast accuracy* and what is the *financial impact* of this improvement?



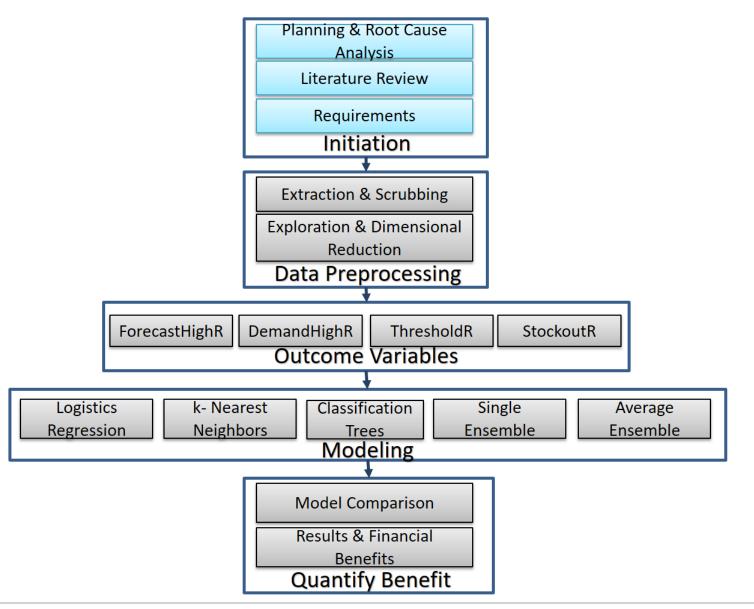
What factors are important to the success of other CPG companies that want to pursue a similar *risk assessment methodology* in their S&OP plan?



2

3

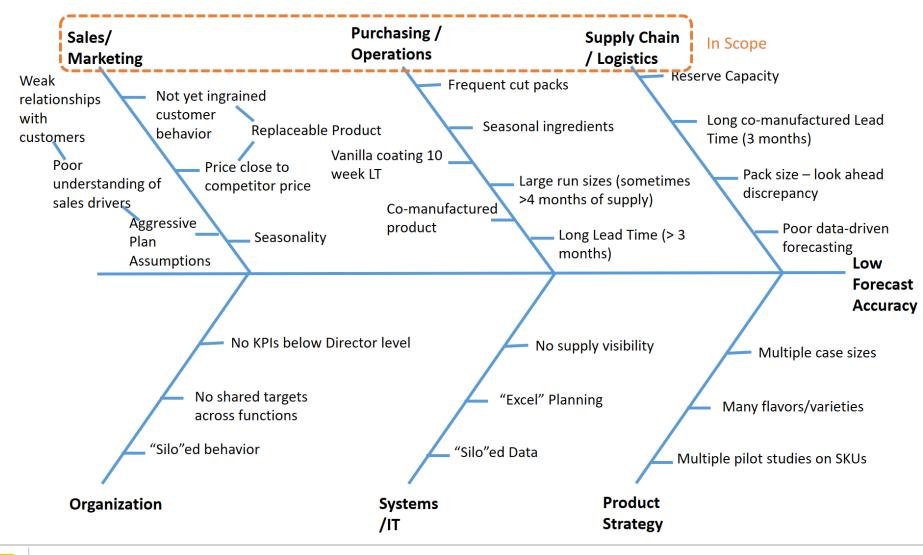
#### Methodology - Initiation



MIT Center for Transportation & Logistics

### Methodology - Initiation

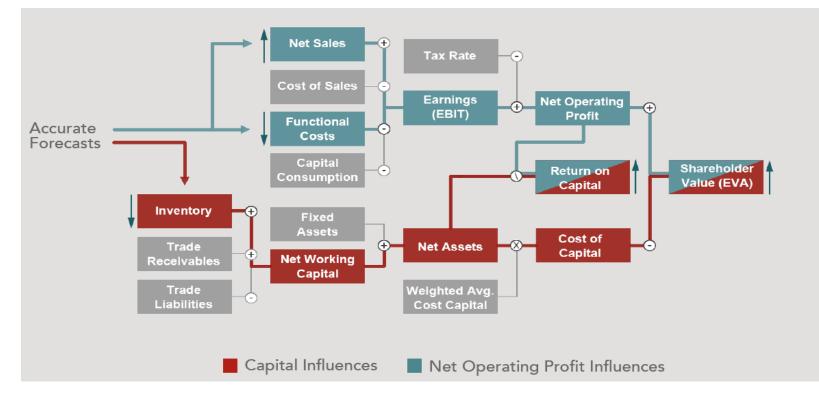
Root Cause Analysis



MIT Center for Transportation & Logistics

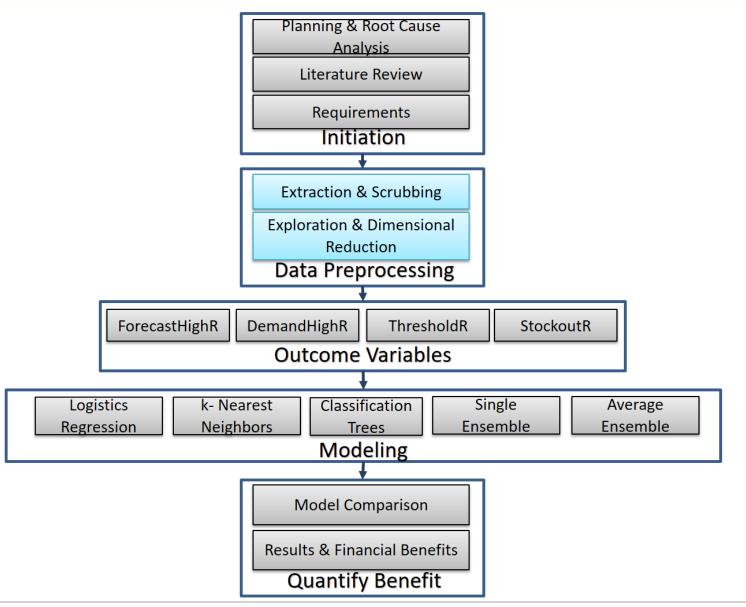
# Methodology - Initiation

#### Literature Review



- E2OPEN (2016) Forecast Accuracy: Why It Matters and How To Improve It. Retrieved from https://www.e2open.com/resources/forecast-accuracy-why-it-matters-and-how-to-improve-it
- Chambers, J., Mullick, S., & Smith, D. (1971 Jul.) How to Choose the Right Forecasting Technique. Harvard Business Review. Retrieved from https://hbr.org
- Davenport, T. (2006) Competing on Analytics. Harvard Business Review. Retrieved from <a href="https://hbr.org">https://hbr.org</a>
- Hinkel, J., Merkel, O., & Kwasniok, T. (2016, Apr. 13) Good Sales and Operations Planning Is No Longer Good Enough. Retrieved from http://www.bain.com
- Myerholtz, B., & Caffrey, H. (2014, Nov. 4) Demand Forecasting: The Key to Better Supply-Chain Performance. Retrieved from https://www.bcg.com

#### Methodology - Data Preprocessing





### Methodology – Data Preprocessing

Forecast

Min8Wks

Max8Wks

CoV8Wks

MAD8Wks

InitialloH

Wk3

Wk4

Wk1

Wk2

- S&OP Excel files from Sep 2016-Nov  $\bullet$ 2017
- 2,477 records for a protein bar brand

Predictors					
FullHorizonWOs	CoV8				
PretoFullRatio	MAD8Wks				
InitialIoH	Min8Wks				
OHToFcstRatio	Max8Wks				
Forecast	Wk1-Wk4				
EstimatedIoH	Promo (unavailable)				

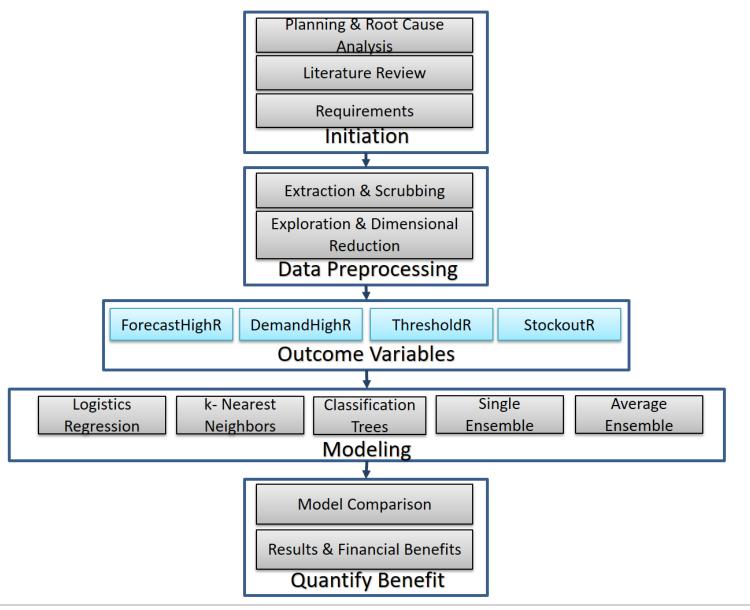
#### Variable Correlation Heat Map

1	0.22	0.66	0.71	0.1	0.36	-0.19	-0.05	0.28	0.31	0.25	FullHorizonWOs
0.22	1	0.71	0.8	0.26	-0.15	0.37	-0.19	0.6	0.67	0.53	InitialloH
0.66	0.71	1	0.78	-0.03	0.08	-0.03	-0.07	0.63	0.69	0.52	Forecast
0.71	0.8	0.78	1	0.37	0.1	0.24	-0.15	0.49	0.54	0.43	EstimatedIoH
0.1	0.26	-0.03	0.37	1	0.06	0.64	-0.12	0.01	0	0	EstimateWoS
0.36	-0.15	0.08	0.1	0.06	1	-0.39	0.03	0.03	0.04	0.05	PretoFullRatio
-0.19	0.37	-0.03	0.24	0.64	-0.39	1	-0.07	0	-0.01	0.03	OHToFcstRatio
-0.05	-0.19	-0.07	-0.15	-0.12	0.03	-0.07	1	0.14	0.04	-0.43	CoV8Wks
0.28	0.6	0.63	0.49	0.01	0.03	0	0.14	1	0.93	0.18	MAD8Wks
0.31	0.67	0.69	0.54	0	0.04	-0.01	0.04	0.93	1	0.46	Max8Wks
0.25	0.53	0.52	0.43	0	0.05	0.03	-0.43	0.18	0.46	1	Min8Wks
FullHorizonWOs	InitialloH	Forecast	EstimatedIoH	EstimateWoS	PretoFullRatio	OHT oF cstRatio	CoVBWks	MAD8Wks	M ax8Wks	Min8Wks	

#### Forecast Min8Wks MAD8Wks Max8Wks CoV8Wks EstimatedIoF EstimatedIoH EstimateWoS EstimateWoS InitialloH FullHorizonWOs FullHorizonWOs OHToFcstRatic OHToFcstRatio PretoFullRatio Wk3 PretoFullRatio Wk1 Wk4 Wk2 10 40 15 0 20 30 0 5 10 %IncMSE IncNodePurity

#### Variable Importance Plot

#### Methodology – Outcome Variables



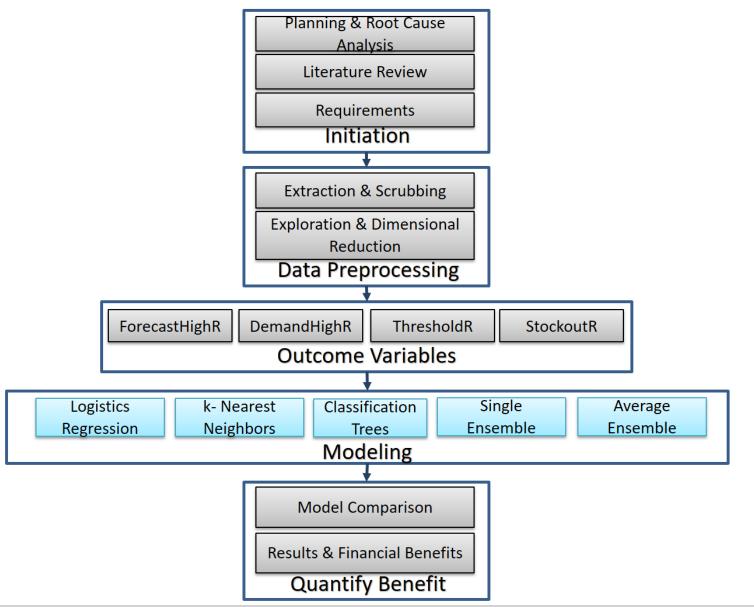
MIT Center for Transportation & Logistics

### Methodology – Outcome Variables

Outcome Variables	Definition
ForecastHighR	<ul> <li>(Forecast-ActualDemand)/Forecast &gt; 0.5</li> <li>Forecast&gt;100</li> </ul>
DemandHighR	<ul> <li>(ActualDemand-Forecast)/Forecast &gt; 0.5</li> <li>Forecast&gt;100</li> </ul>
ThresholdR	<ul> <li>WoS &lt; 4-week threshold</li> </ul>
StockoutR	<ul> <li>Weekly demand &gt; Weekly supply across the entire network</li> </ul>



#### Methodology – Modeling



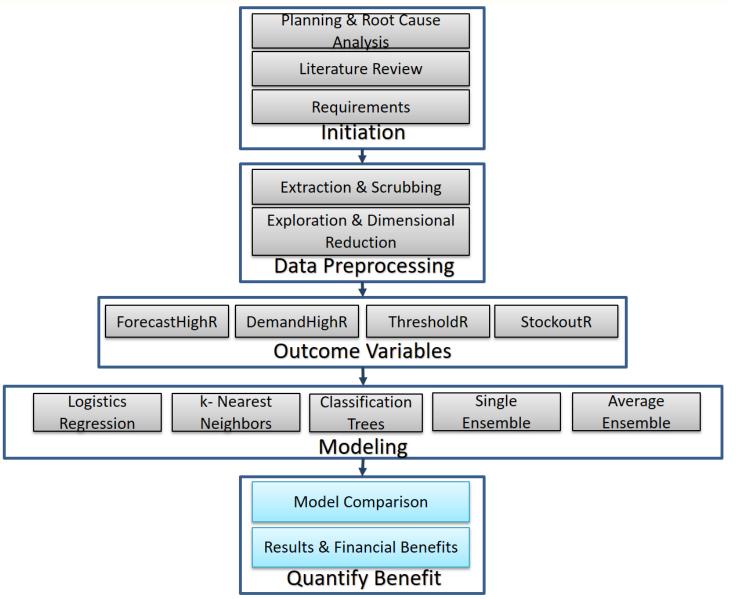


# Methodology – Modeling

• Historical Data Prediction Linear regression k-nearest neighbors • Algorithms "learn" **Regression trees** Neural networks Ensembles Supervised • Predictions Classification k-nearest neighbors **Model Evaluation** Naive Baves Model and Selection Data Preparation, Classification trees Deployment Performance Exploration, and Logistic regression Score new data evaluations Reduction Neural networks Data preparation Discriminant analysis Data visualization Ensembles **Dimension reduction Time Series Forecasting Regression-based** Smoothing methods What Goes Together Unsupervised Deriving Association rules Insight Collaborative filtering Segmentation **Cluster analysis** 



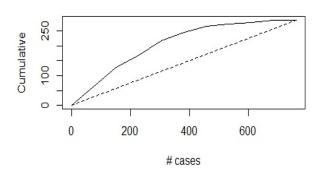
#### Methodology – Quantify Benefit

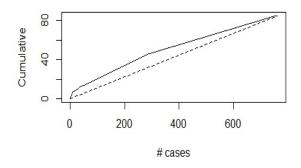


MIT Center for Transportation & Logistics

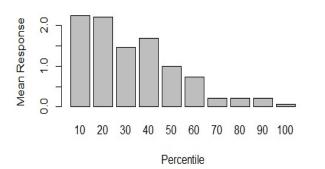
### Methodology – Model Analysis

#### Lift and Decile-wise Lift Chart

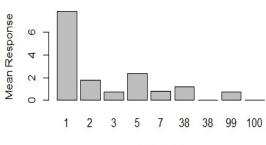




Decile-wise lift chart







Percentile

#### **Confusion Matrix**

(1) Logistic Regression Cutoff=.45	(2) k-nearest neighbors k=7	(3) Classification Trees Cutoff=.3
Reference	Reference	Reference
Prediction 0 1	Prediction 0 1	Prediction 0 1
0 393 71	0 395 97	0 351 72
1 82 216	1 80 190	1 124 215
<b>Accuracy</b> : 0.7992	<b>Accuracy</b> : 0.6620	<b>Accuracy</b> : 0.7428
<b>P-Value</b> : <2e-16	<b>P-Value :</b> <2e-16	<b>P-Value</b> : <1.819e-12
Sensitivity : .753	Sensitivity : .662	Sensitivity : .749
(4) Single Ensemble Cutoff=.35	(5) Average Ensemble (Model	1, 2 & 3) Cutoff=0.4
(4) Single Ensemble Cutoff=.35 Reference	(5) Average Ensemble (Model Reference	1, 2 & 3) Cutoff=0.4
		1, 2 & 3) Cutoff=0.4
Reference	Reference	1, 2 & 3) Cutoff=0.4
Reference Prediction 0 1	Reference Prediction 0 1	1, 2 & 3) Cutoff=0.4
Reference Prediction 0 1 0 381 70	Reference Prediction 0 1 0 373 59	1, 2 & 3) Cutoff=0.4
Reference Prediction 0 1 0 381 70 1 94 217	Reference Prediction 0 1 0 373 59 1 102 228	1, 2 & 3) Cutoff=0.4

# Methodology – Model Performance Comparison

#### **Forecast Accuracy**

ForecastHighR	DemandHighR	ThresholdR	StockoutR
79.92%	82.53%	88.71%	72.51%
66.20%	85.25%	89.37%	76.82%
74.28%	81.41%	87.93%	74.66%
78.48%			
78.87%			
	79.92% 66.20% 74.28% 78.48%	79.92%       82.53%         66.20%       85.25%         74.28%       81.41%         78.48%       1	79.92%       82.53%       88.71%         66.20%       85.25%       89.37%         74.28%       81.41%       87.93%         78.48%

p=value							
Models	ForecastHighR	DemandHighR	ThresholdR	StockoutR			
(1)Logistic Regression	<2e-16	0.017	0.574	0.584			
(2)k-nearest neighbors	<2e-16	5.70E-06	0.348	0.007			
(3)Classification Tree	1.819E-12	0.109	0.807	0.132			
(4)Single Ensemble	<2e-16						
(5)Average Ensemble (Models 1, 2 & 3)	<2e-16						

n\_\/alua

#### Sensitivity

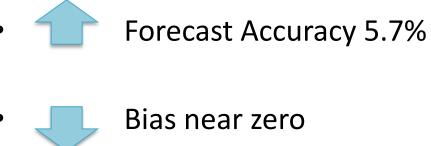
Models	ForecastHighR	DemandHighR	ThresholdR	StockoutR	
(1)Logistic Regression	75.30%	28.50%	8.20%	6.90%	
(2)k-nearest neighbors	66.20%	63.00%	15.30%	34.70%	
(3)Classification Tree	74.90%	35.00%	10.60%	29.70%	
(4)Single Ensemble	75.60%				
(5)Average Ensemble (Models 1, 2 & 3)	79.40%				

### Methodology – Quantify Business Output

- Model Tested on S&OP Plans from Feb-Apr 2018
- Output from model (right) used for ForecastHighR risk mitigation

ItemID	Week	ModelProbability	ModelPrediction Modify	ForecastOrig	ForecastDMModel
4969	2/4/2018	65%	1 Y	211	141
4969	2/11/2018	71%	1 Y	211	141
4969	2/18/2018	77%	1 Y	211	141
4969	2/25/2018	77%	1 Y	191	127
4969	3/4/2018	81%	1 Y	160	107
4969	3/11/2018	77%	1 Y	160	107
4969	3/18/2018	72%	1 Y	160	107
4969	3/25/2018	86%	1 Y	353	235
4969	4/1/2018	74%	1 Y	257	171
4969	4/8/2018	68%	1 Y	257	171
4969	4/15/2018	73%	1 Y	257	171
4969	4/22/2018	69%	1 Y	120	80
4969	4/29/2018	70%	1 Y	138	92
4970	2/4/2018	32%	0	149	149
4970	2/11/2018	37%	0	149	149
4970	2/18/2018	32%	0	149	149
4970	2/25/2018	18%	0	148	148
4970	3/4/2018	33%	0	154	154
4970	3/11/2018	23%	0	149	149
4970	3/18/2018	23%	0	160	160
4970	3/25/2018	23%	0	153	153
4970	4/1/2018	17%	0	122	122
4970	4/8/2018	16%	0	122	122
4970	4/15/2018	17%	0	122	122
4970	4/22/2018	18%	0	119	119
4974	2/4/2018	54%	1	149	149
4974	2/11/2018	35%	0	145	145
4974	2/18/2018	31%	0	147	147
4974	2/25/2018	55%	1 Y	150	100
4974	3/4/2018	59%	1 Y	156	104
4974	3/11/2018	63%	1 Y	156	104
4974	3/18/2018	64%	1 Y	158	105
4974	3/25/2018	64%	1 Y	157	105
4974	4/1/2018	53%	1 Y	124	83
4974	4/8/2018	63%	1 Y	124	83

# Methodology – Quantify Benefits



Gross Profit \$1.8MM (protein bar brand)

Accuracy	Feb	Mar	Apr	Total
Baseline	50.4%	55.3%	44.8%	50.0%
Predictive Model	54.1%	57.2%	55.8%	55.7%
Improvement	3.6%	1.9%	10.9%	5.7%

Improvement in Model Accuracy & Bias

Bias	Feb	Mar	Apr	Total
Baseline	-1.0%	2.4%	8.5%	3.7%
Predictive Model	-3.9%	-1.0%	3.4%	-0.3%

#### Increase of \$17MM in annual gross profit



# Conclusion

**Supervised classification models effectively predict risks** in the S&OP plan, even without big data.



Three steps to gain large increase in profit and competitive advantage:

- Capture planning data
- Leverage predictive analytics
- Buy in from key stakeholders

2

3









# Questions?

Presenters: *Deepti Kidambi & Minhaaj Khan* Advisor: *Dr. Tugba Efendigil*