MIT Center for Transportation & Logistics

Enhancing S&OP Performance with Analytics

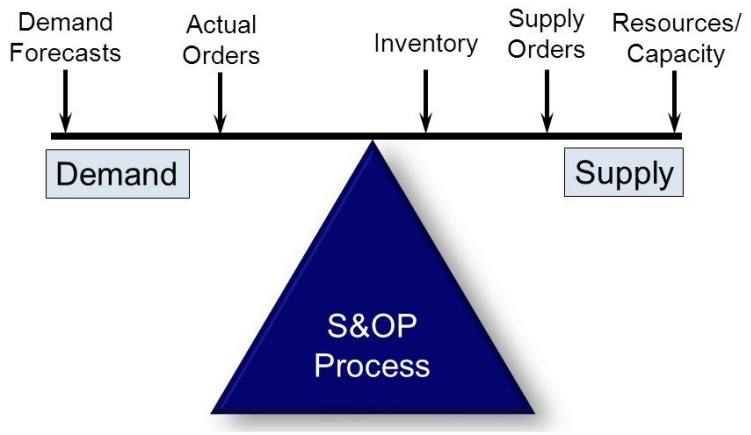
Presenters: *Deepti Kidambi & Minhaaj Khan*
Advisor: *Dr. Tugba Efendigil*

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
- Profit Margin by 3.8%, $1.8MM
- $17MM for the year

Key Research Questions

1

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

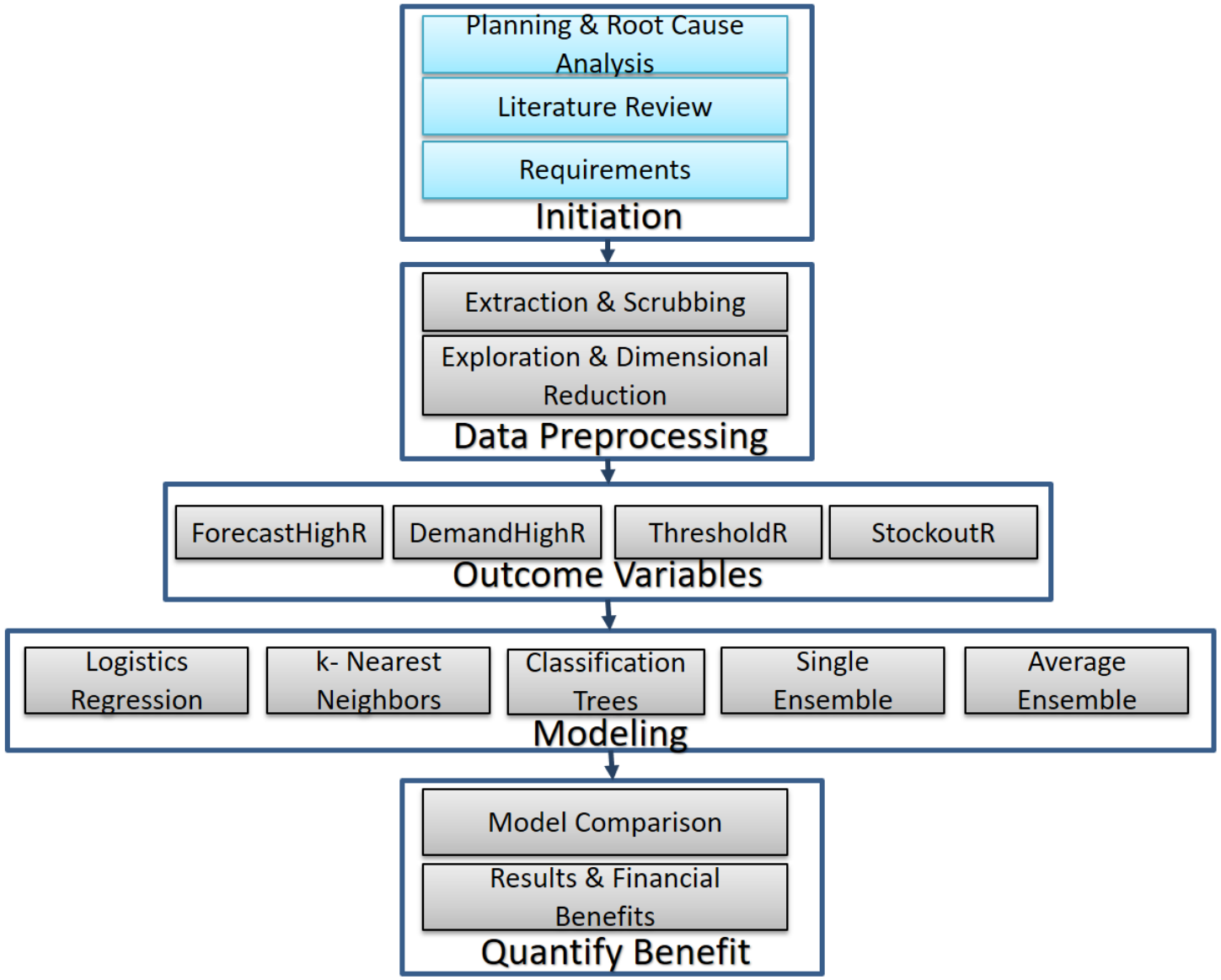
2

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

3

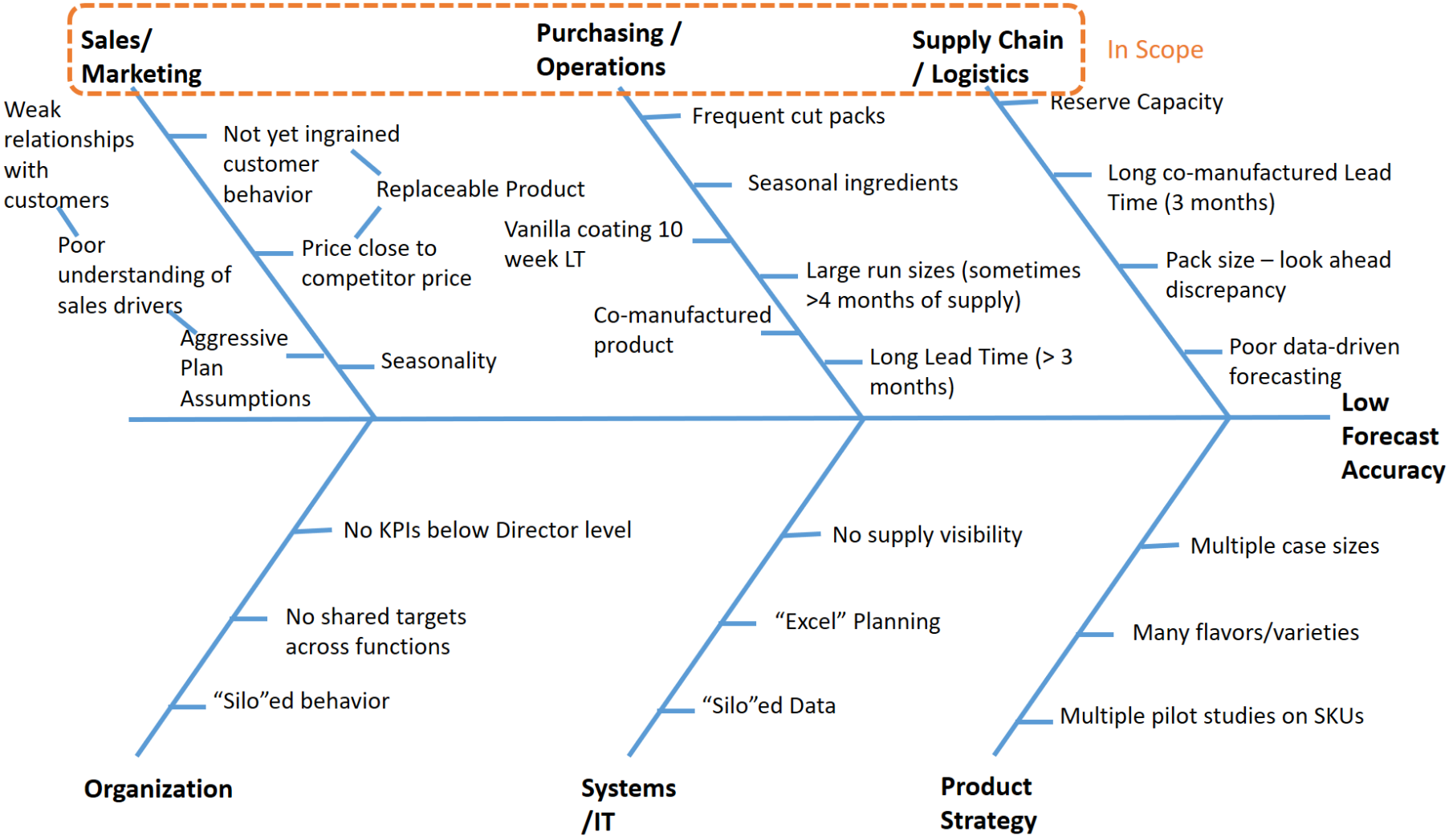
What factors are important to the success of other CPG companies that want to pursue a similar ***risk assessment methodology*** in their S&OP plan?

Methodology - Initiation



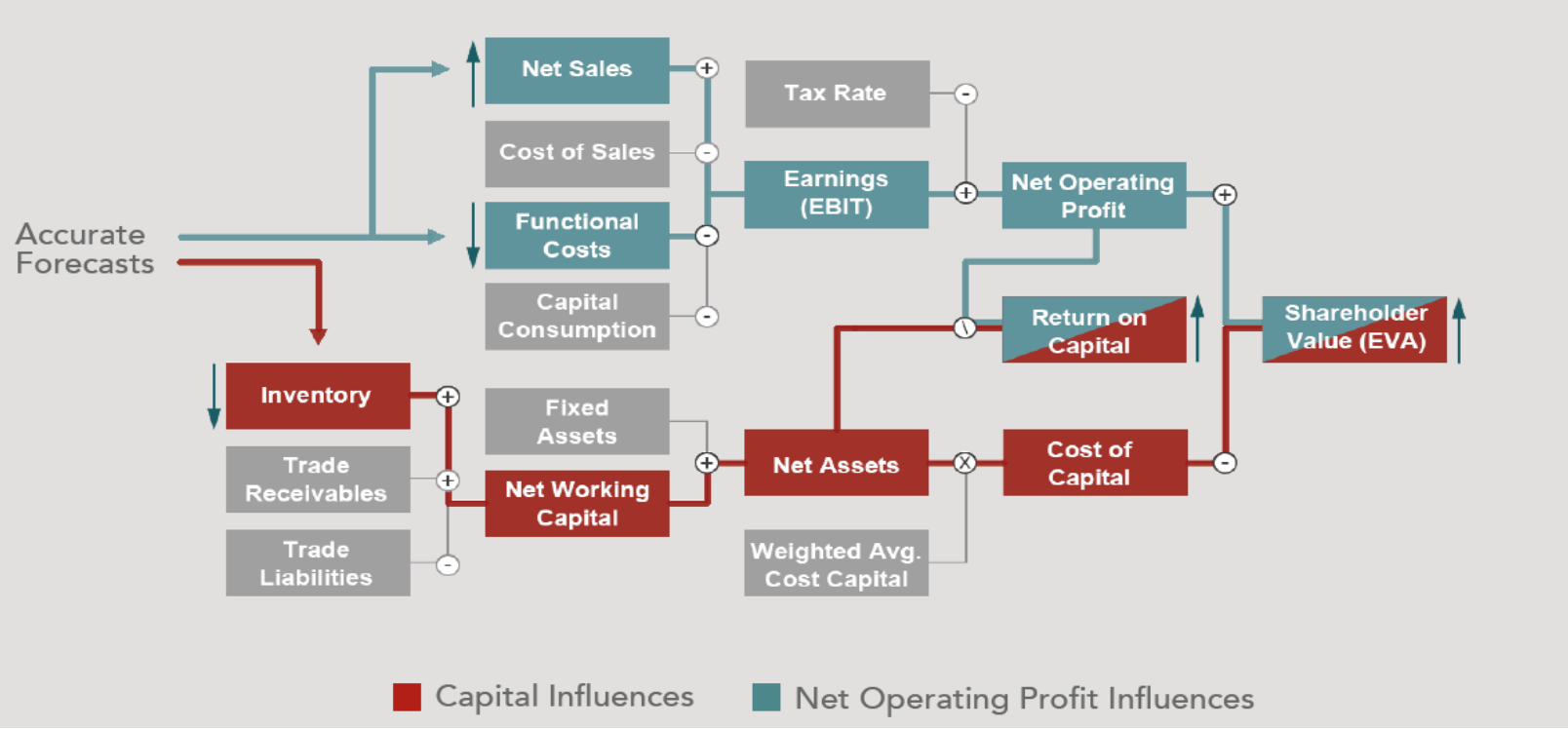
Methodology - Initiation

- Root Cause Analysis



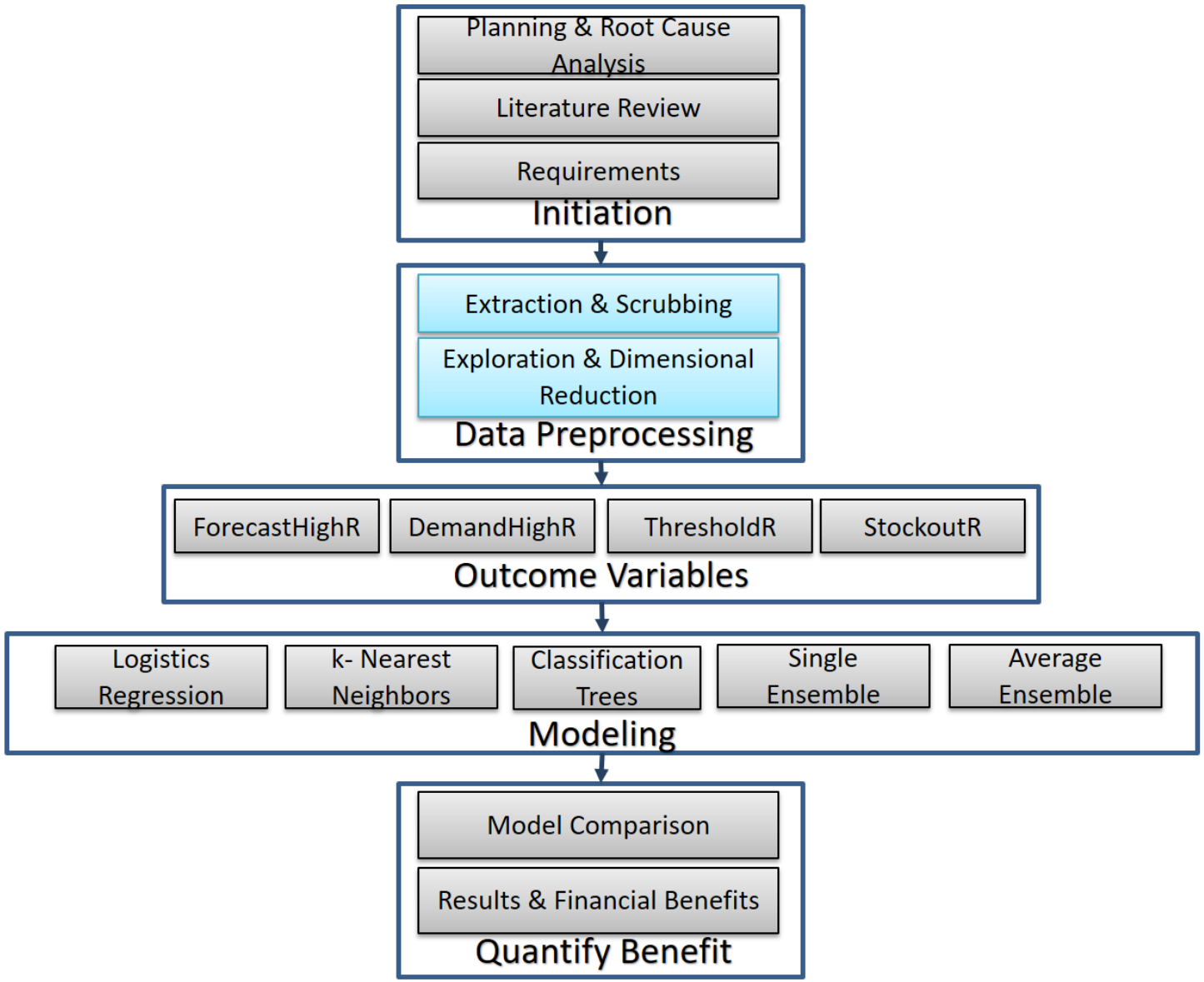
Methodology - Initiation

- 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 https://hbr.org
- 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

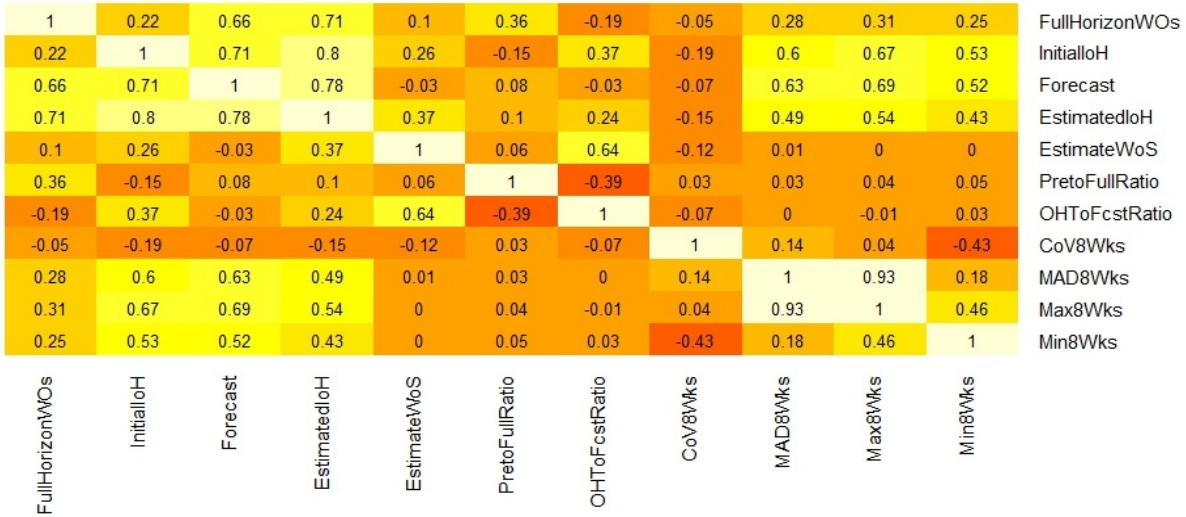


Methodology – Data Preprocessing

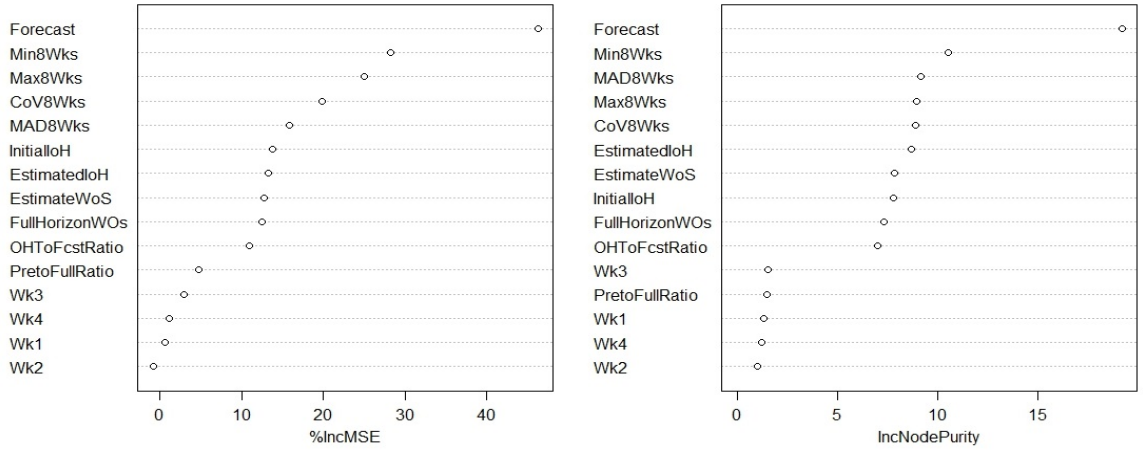
- S&OP Excel files from Sep 2016-Nov 2017
- 2,477 records for a protein bar brand

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

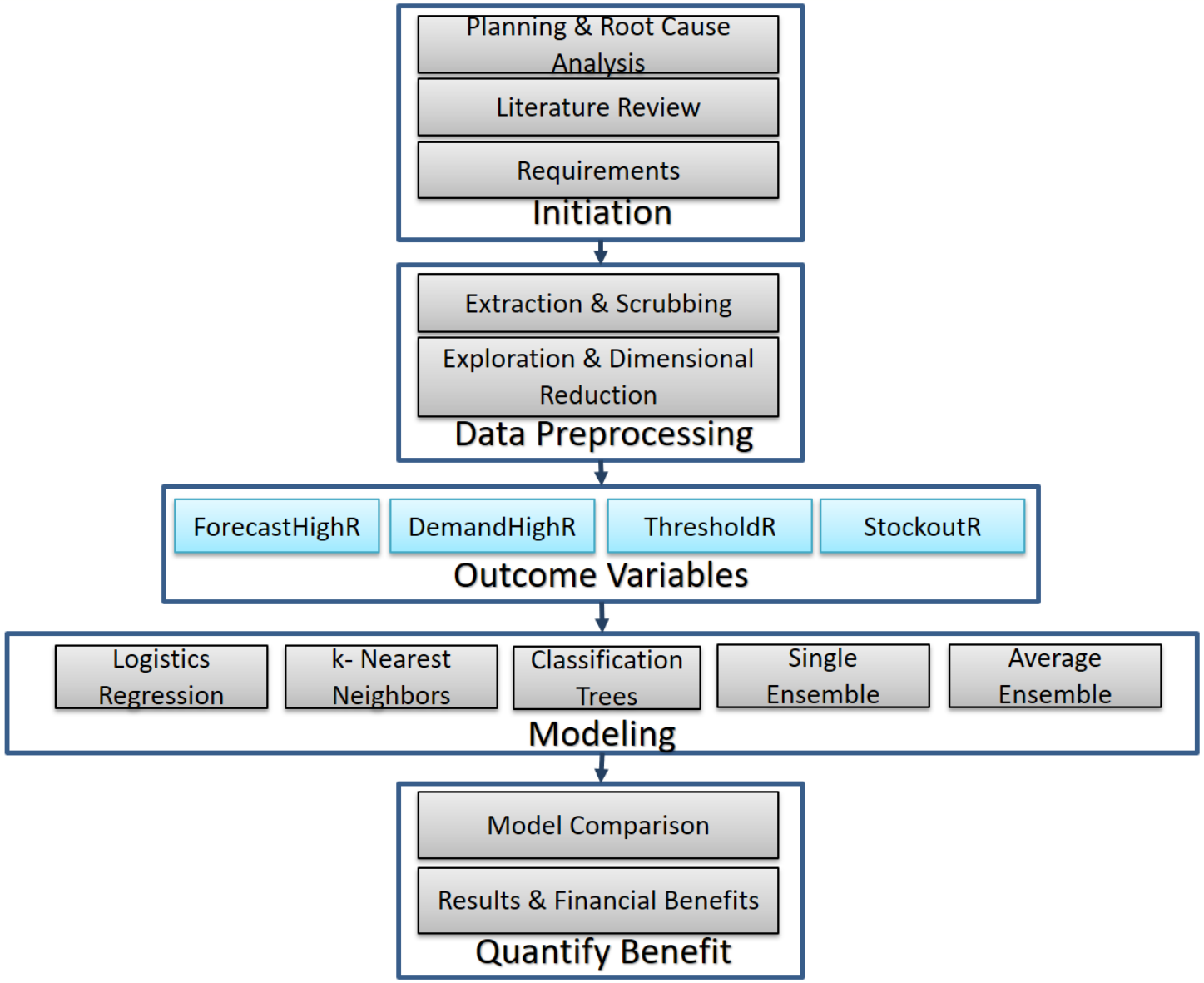
Variable Correlation Heat Map



Variable Importance Plot



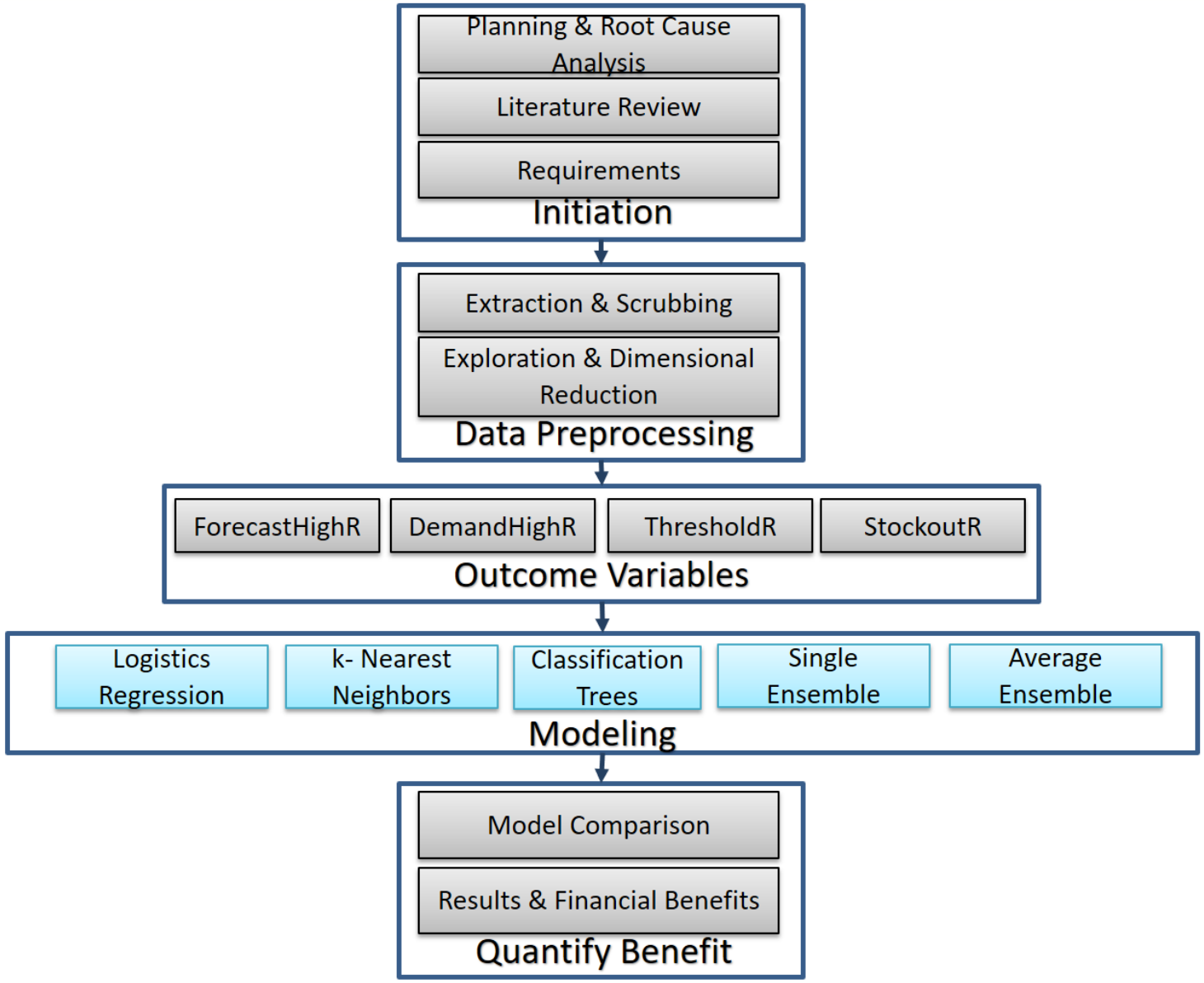
Methodology – Outcome Variables



Methodology – Outcome Variables

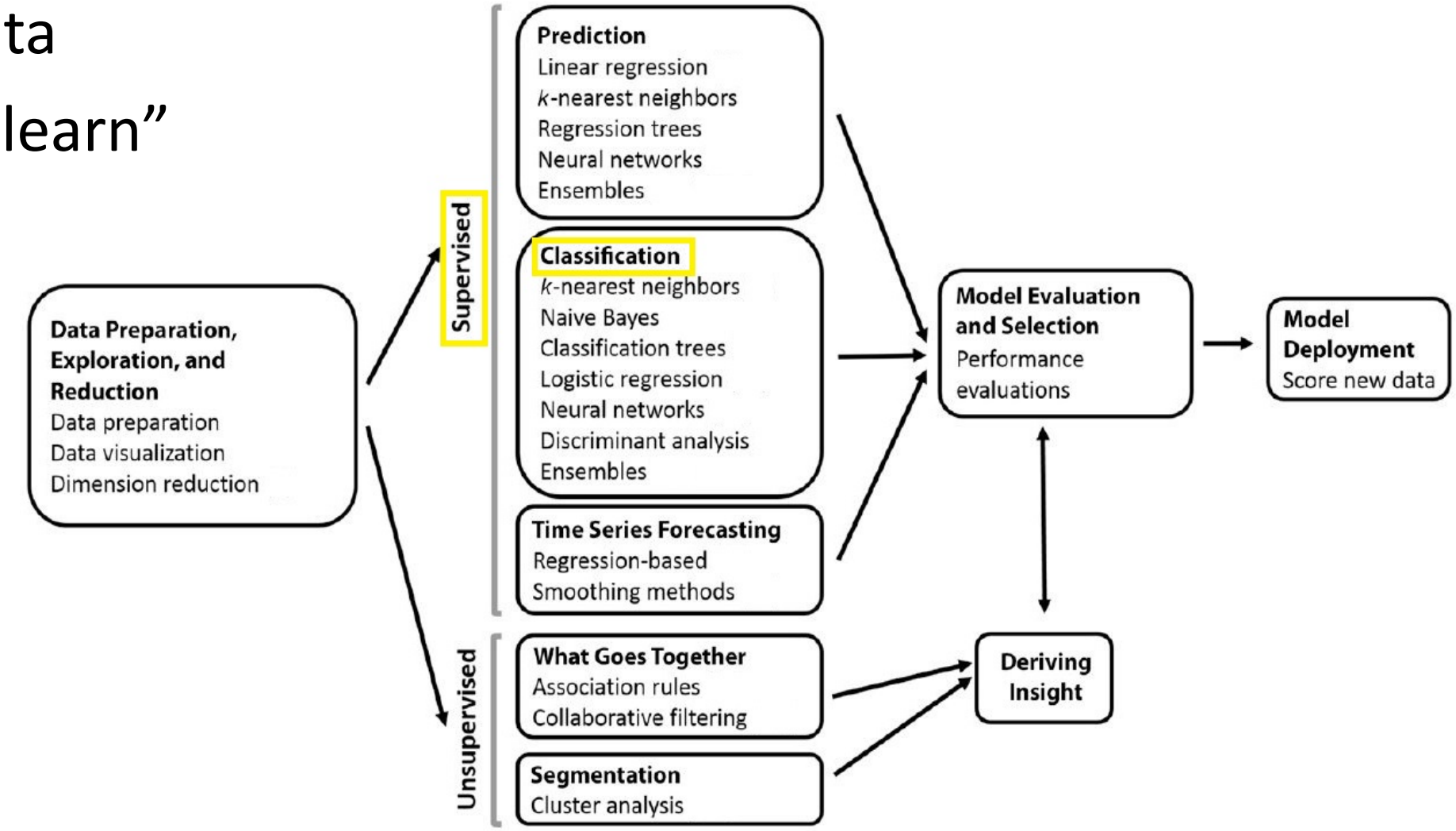
Outcome Variables	Definition
ForecastHighR	• *(Forecast-ActualDemand)/Forecast* > 0.5 • Forecast>100
DemandHighR	• *(ActualDemand-Forecast)/Forecast* > 0.5 • Forecast>100
ThresholdR	• WoS < 4-week threshold
StockoutR	• *Weekly demand > Weekly supply* across the entire network

Methodology – Modeling

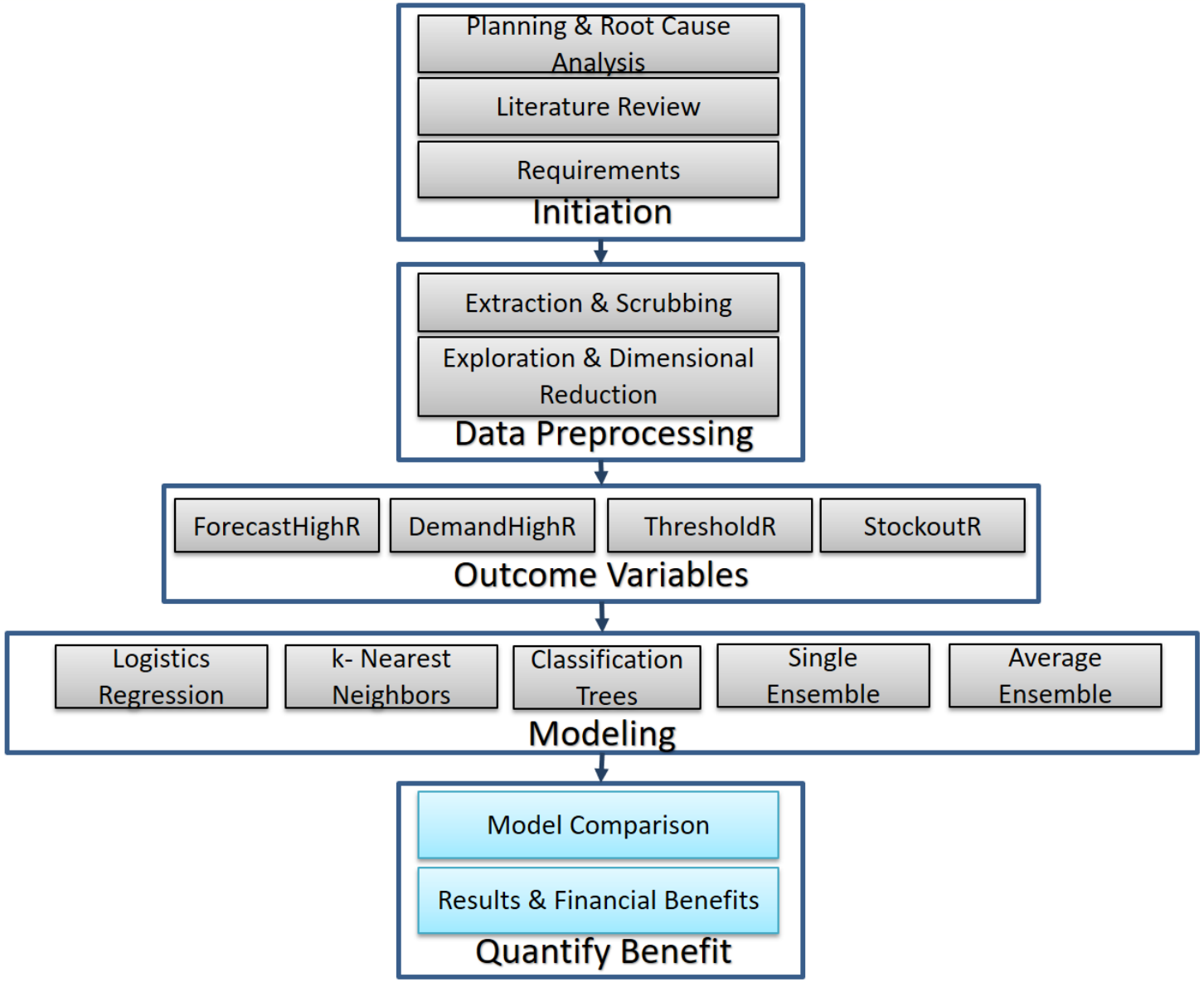


Methodology – Modeling

- Historical Data
- Algorithms “learn”
- Predictions

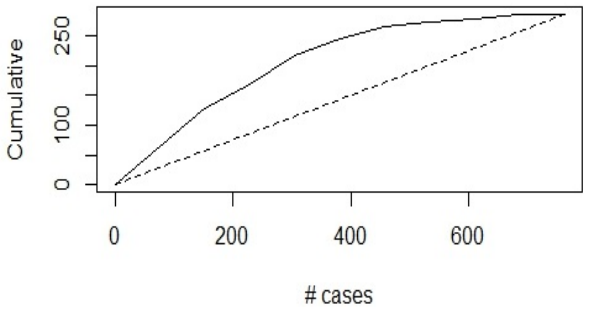


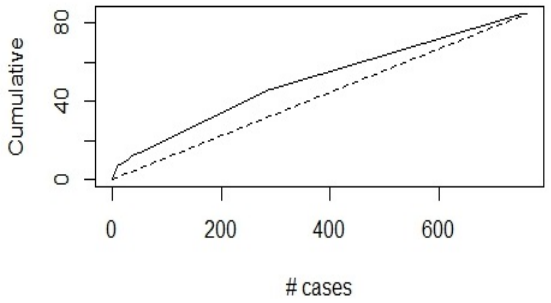
Methodology – Quantify Benefit

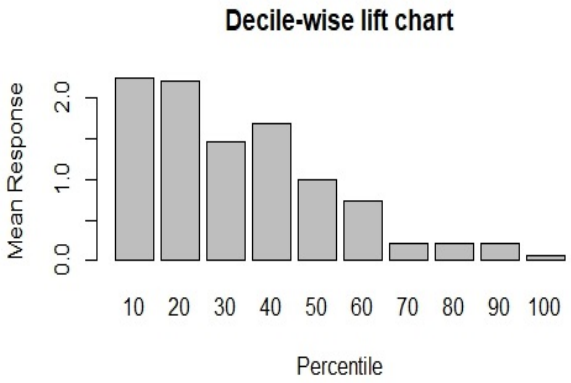


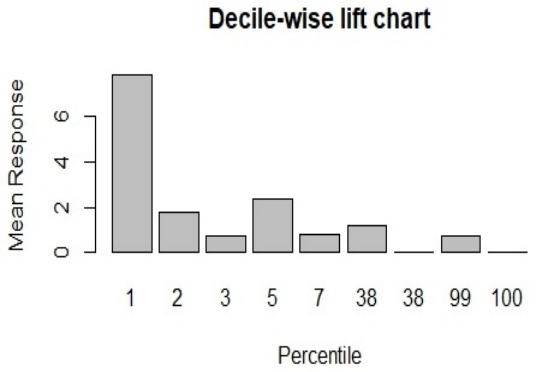
Methodology – Model Analysis

Lift and Decile-wise Lift Chart









Confusion Matrix

(1) Logistic Regression Cutoff=.45

```
          Reference
Prediction   0   1
         0 393  71
         1  82 216

Accuracy : 0.7992
P-Value : <2e-16
Sensitivity : .753
```

(2) k-nearest neighbors k=7

```
          Reference
Prediction   0   1
         0 395  97
         1  80 190

Accuracy : 0.6620
P-Value : <2e-16
Sensitivity : .662
```

(3) Classification Trees Cutoff=.35

```
          Reference
Prediction   0   1
         0 351  72
         1 124 215

Accuracy : 0.7428
P-Value : <1.819e-12
Sensitivity : .749
```

(4) Single Ensemble Cutoff=.35

```
          Reference
Prediction   0   1
         0 381  70
         1  94 217

Accuracy : 0.7848
P-Value : <2e-16
Sensitivity : .756
```

(5) Average Ensemble (Model 1, 2 & 3) Cutoff=0.4

```
          Reference
Prediction   0   1
         0 373  59
         1 102 228

Accuracy : 0.7887
P-Value : < 2.2e-16
Sensitivity : .794
```

Methodology – Model Performance Comparison

Forecast Accuracy

Models	ForecastHighR	DemandHighR	ThresholdR	StockoutR
(1)Logistic Regression	79.92%	82.53%	88.71%	72.51%
(2)k-nearest neighbors	66.20%	85.25%	89.37%	76.82%
(3)Classification Tree	74.28%	81.41%	87.93%	74.66%
(4)Single Ensemble	78.48%			
(5)Average Ensemble (Models 1, 2 & 3)	**78.87%**			

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**			

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

- ↑ Forecast Accuracy 5.7%
- ↓ Bias near zero
- ↑ Gross Profit $1.8MM (protein bar brand)

Improvement in Model Accuracy & Bias

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%**

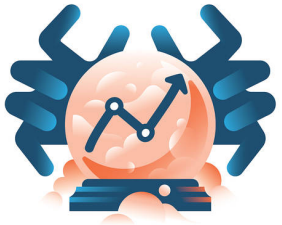
Bias	Feb	Mar	Apr	Total
Baseline	-1.0%	2.4%	8.5%	3.7%
Predictive Model	-3.9%	-1.0%	3.4%	**-0.3%**

Increase of $17MM in annual gross profit

Conclusion

1

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



2

Potential to deliver ***substantial improvement*** in ***forecast accuracy*** and ***gross profit***.

3

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

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

Questions?

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