Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

Massachusetts Institute of Technology

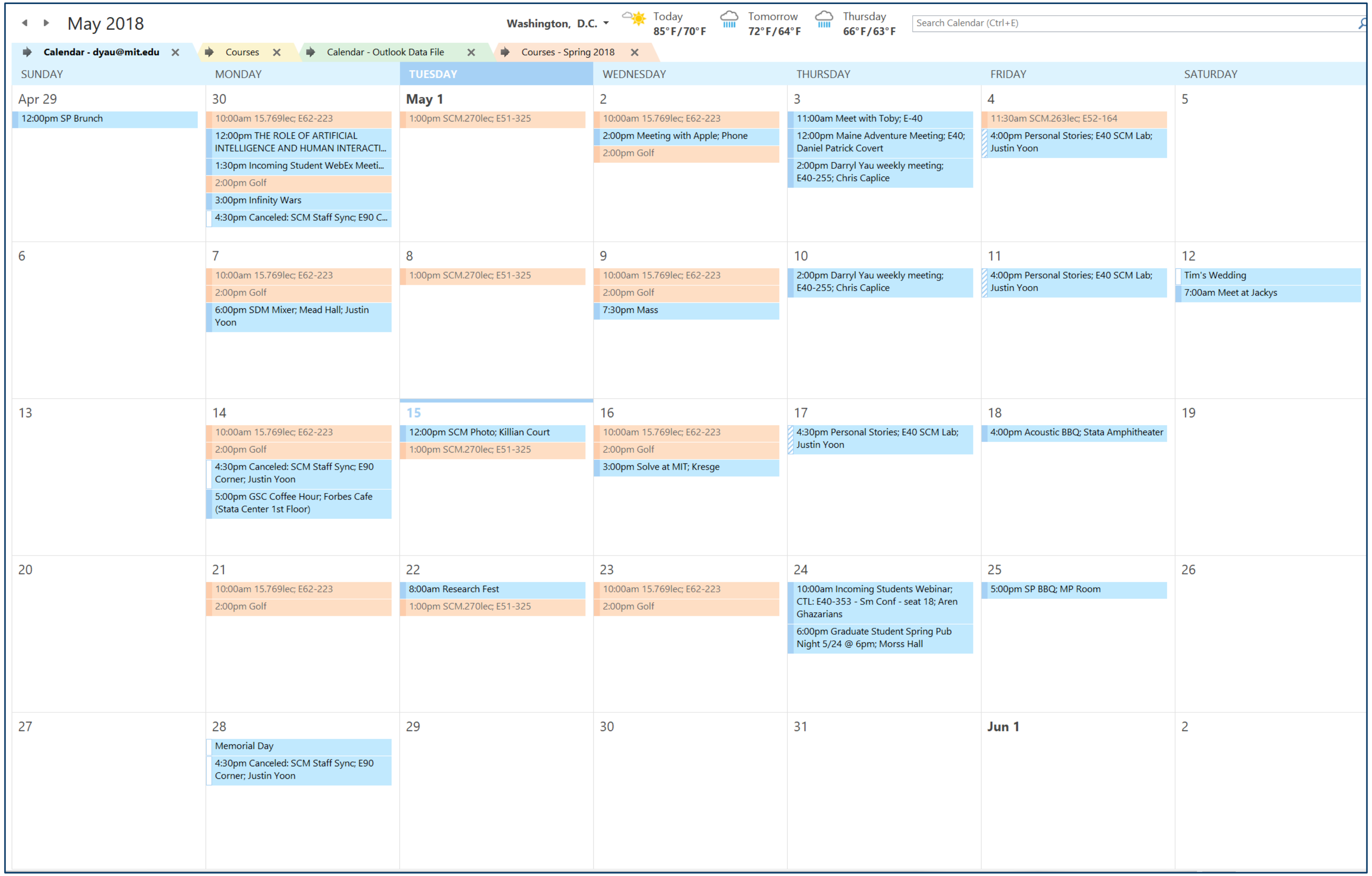
MIT Supply Chain Management

Research Fest 2018

Presented By: Darryl Yau
Advised By: Dr. Christopher Caplice

Darryl Yau

My Typical Schedule



- Always at my meetings
- 100% adherence to schedule
- 100% On Time Delivery (OTD)

Massachusetts Institute of Technology

MIT Supply Chain MANAGEMENT

Supply Chain Example

BUT*…100% adherence to schedule within the supply chain context is almost unheard of*

Period	0	1	2	3	4
Demand	50	100	50	50	100
Production Plan	50	100	50	50	100
Actual Production	40	90	80	20	120
	-10	-20	+10	-20	0

*Supply Chain is very **complex**!*

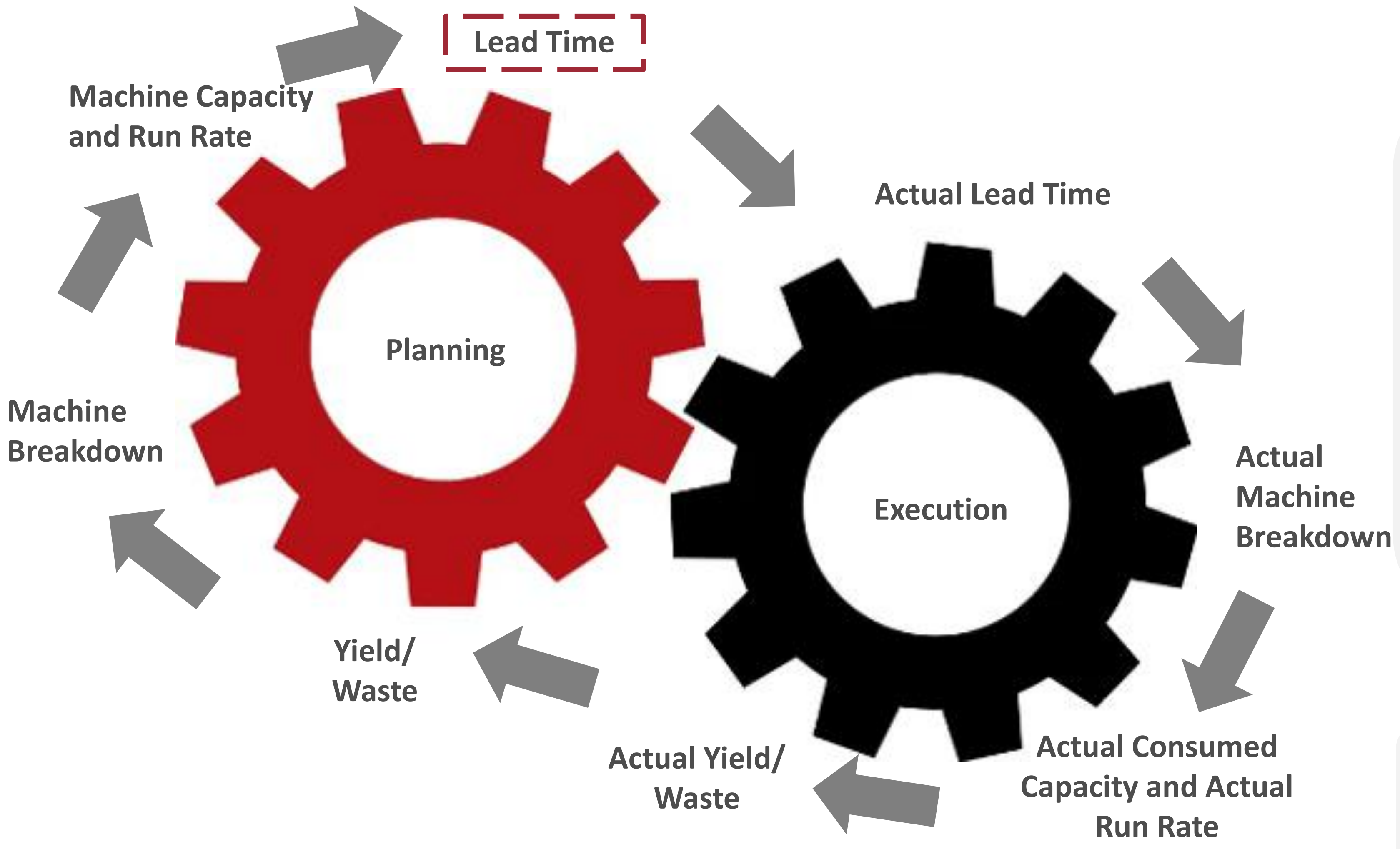

1
2
4
7
8
11
14
15
17
18
23
24

What can we do?

Force operations to conform to the schedule

Create a schedule that is more accurate

Many parameters used during planning process are not given the proper attention it deserves



Consider:

- Values that were not scientifically or accurately set in the first place
- Values that have changed or are changing over time

How do we create a **'self-healing'** supply chain?

Darryl Yau

Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

1

Baseline Current State

To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system are used in predicting lead time and how accurate is the prediction?

2

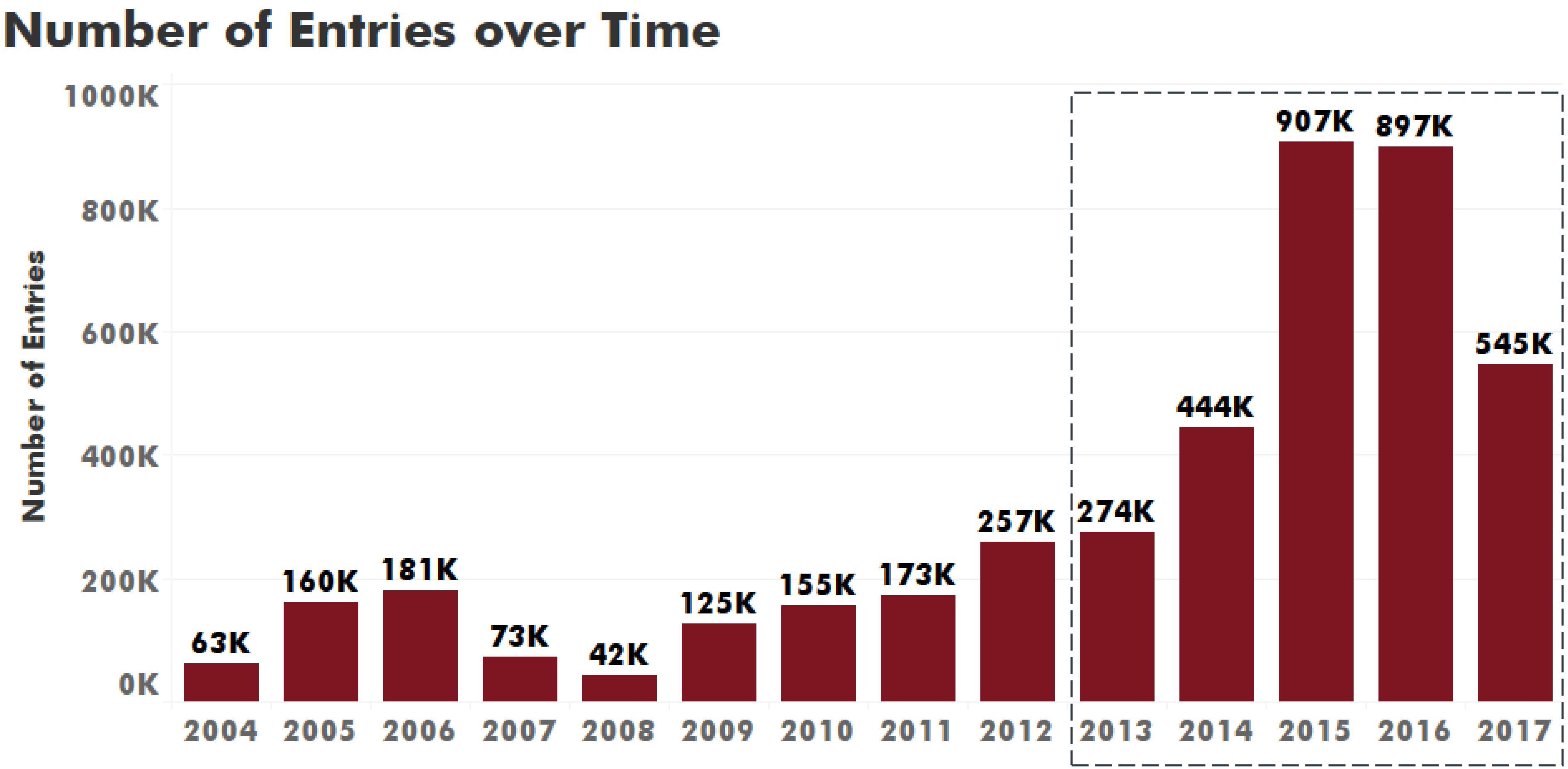
Propose Improved Future State

Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

Today's Agenda

1. Background
2. **Data**
3. Analysis and Results
4. Conclusion

Purchase Order Data (2004–2017)



- Over 4M Line Items (500,000+ Purchase Orders)
- Over 80,000 SKUs

Understanding Different Lead Time Variables Along the Planning Process

Planning Process

Standard Lead–Time Variables

When Order is Placed

When Order is Received

Does not change over time

Lead Time by SKU and Vendor (LT_v)

Lead Time by SKU and Plant (LT_p)

Other Business Constraints/ Decisions

Planned Lead Time

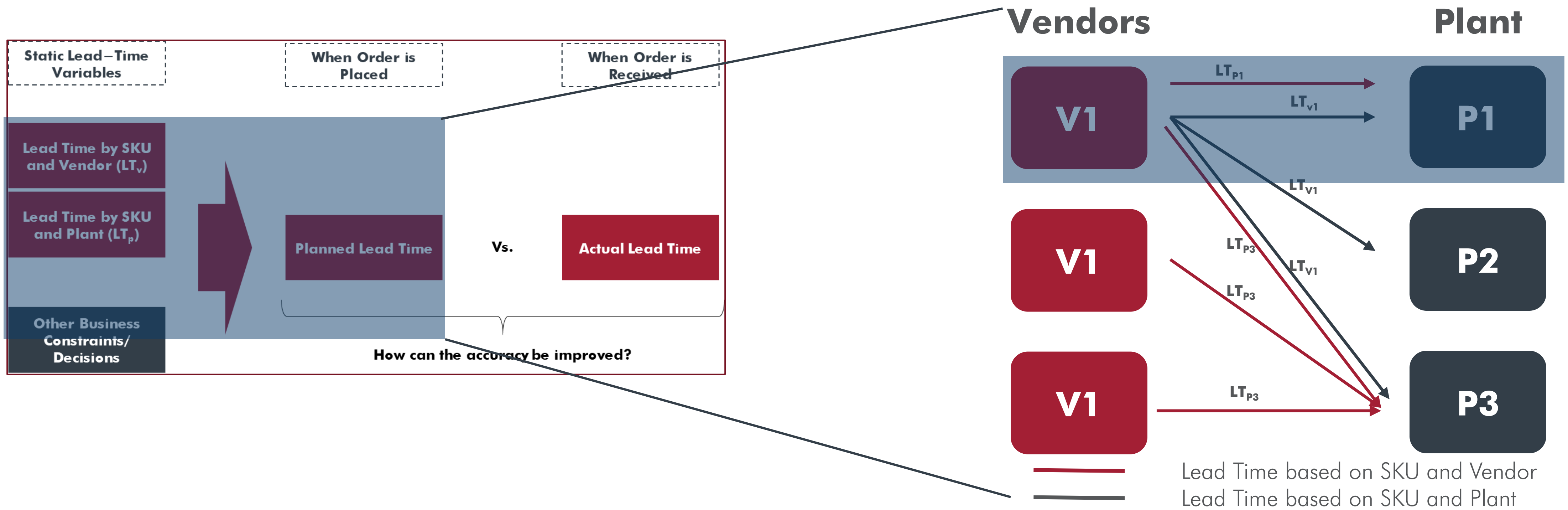
Vs.

Actual Lead Time

How can the accuracy be improved?

Darryl Yau

Conceptualizing How Planned Lead Time is Formulated



$Planned\ Lead\ Time_{SKU-Lane=V1P1} = \beta_0 + \beta_1 LT_{P1} + \beta_2 LT_{V1} + \varepsilon$

Where β_i = coefficients

ε = error or unexplained term

Today's Agenda

1 Background

2 Data

3 **Analysis and Results**

4 Conclusion

Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

1

Baseline Current State

To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system are used in predicting lead time and how accurate is the prediction?

2

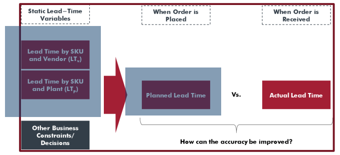
Propose Improved Future State

Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

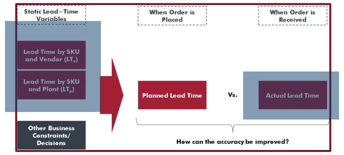
Darryl Yau

Regression Performed Across the Entire Dataset

LT_v and LT_p
vs.
Planned Lead Time



LT_v and LT_p
vs.
Actual Lead Time

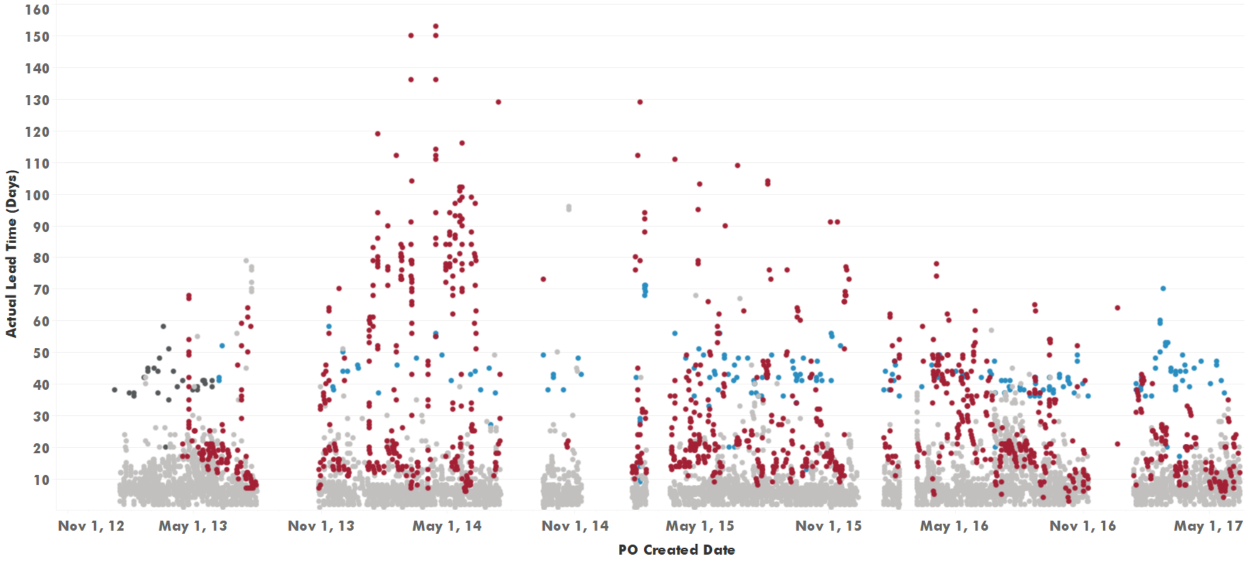


	Dependent Variable	Intercept (β_0)	SKU and Vendor (LT_v) (β_1)	SKU and Plant (LT_p) (β_2)	Adjusted R^2
All	Planned Lead Time	7.086	0.064	0.853	0.253
	Actual Lead Time	7.953	0.224	0.718	0.155
2004-2007	Planned Lead Time	15.827	-0.340	0.051	0.007
	Actual Lead Time	86.269	-2.666	0.049	0.005
2008-2011	Planned Lead Time	20.798	-0.236	0.328	0.032
	Actual Lead Time	73.426	-1.445	0.376	0.014
2012-2015	Planned Lead Time	7.314	0.089	0.909	0.266
	Actual Lead Time	7.042	0.221	0.811	0.204
2016-2017	Planned Lead Time	5.661	0.073	1.122	0.406
	Actual Lead Time	7.537	0.191	0.879	0.249

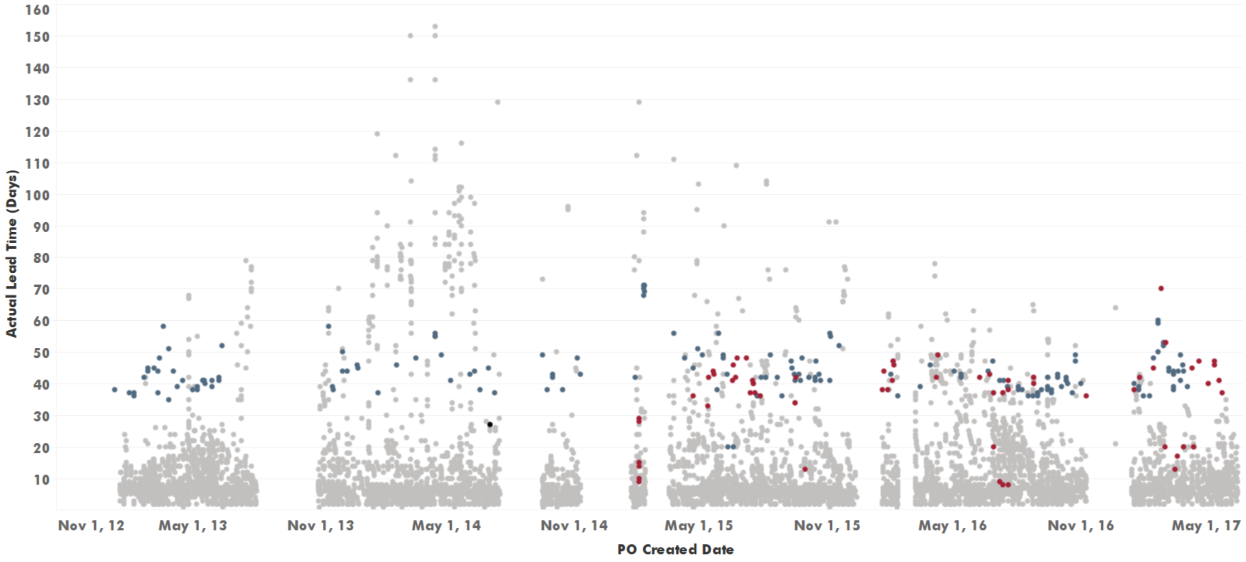
- Poor R^2 values
- Seems to improve over time
- R^2 values for Actual Lead Time consistently worse than R^2 for Planned Lead Time

Vendor and Plant appears to be factors contributing to the variability of actual lead time

Colored by Vendor



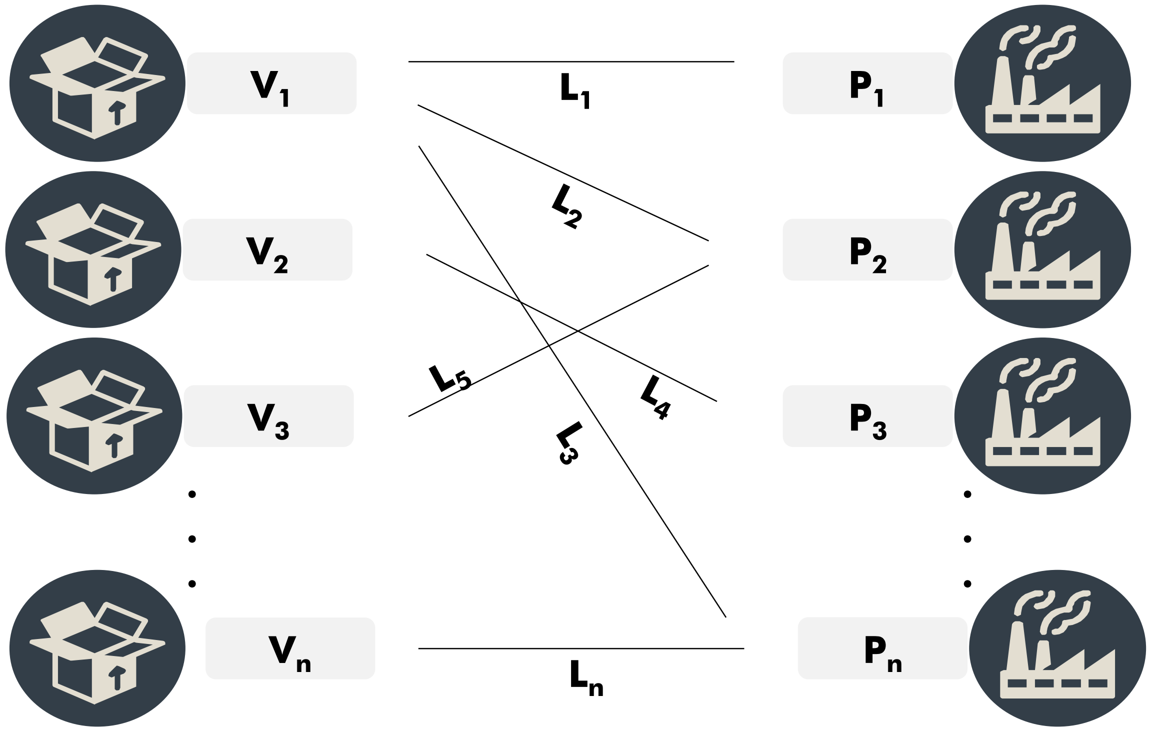
Colored by Plant



Note: Using one SKU as an example

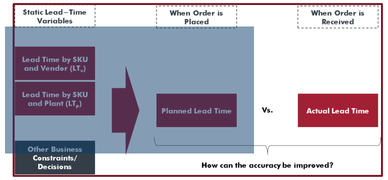
Analyses performed at the SKU–Lane level

Vendors (200+)	Lanes (500+)	Plants (20+)

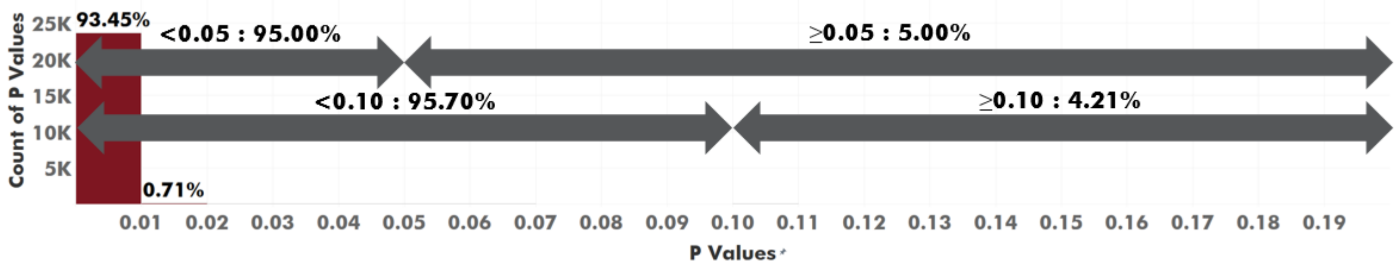


Lane = Unique Vendor and Plant Combination

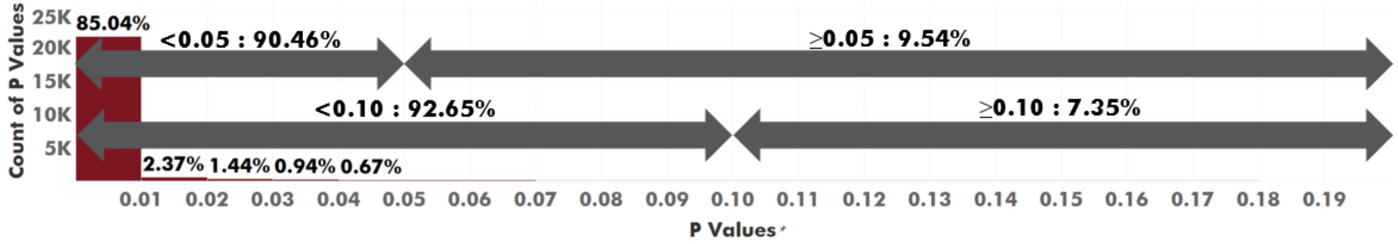
t Test — Null Hypothesis: Are these datasets statistically the same?

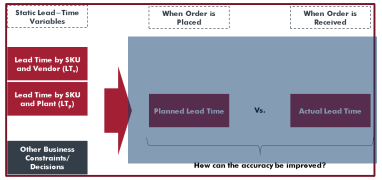


P Values of *T* Tests for LTv vs. Planned Lead TIme

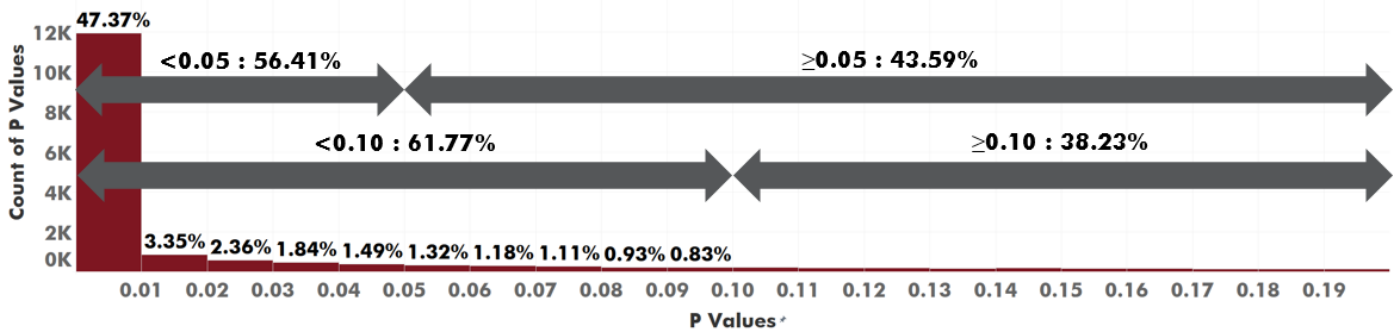


P Values of *T* Tests for LTp vs Planned Lead Time





P Values of *T* Tests for Planned Lead Time vs. Actual Lead Time



- Ran *t* test on over 25,000 SKU-Lanes
  - For LTv and Planned Lead Time
  - For LTp and Planned Lead Time
  - Planned Lead Time and Actual Lead Time
- NOT the same for all tests

Darryl Yau

Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

1

Baseline Current State

To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system are used in predicting lead time and how accurate is the prediction?

2

Propose Improved Future State

The standard lead time variables (LTv and LTp) are **not** good predictors for what is planned

The planned lead times are **not** good predictors for what actually happens

Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

1

Baseline Current State

To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system are used in predicting lead time and how accurate is the prediction?

2

Propose Improved Future State

Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

Time Series Analysis — Forecasting Methods

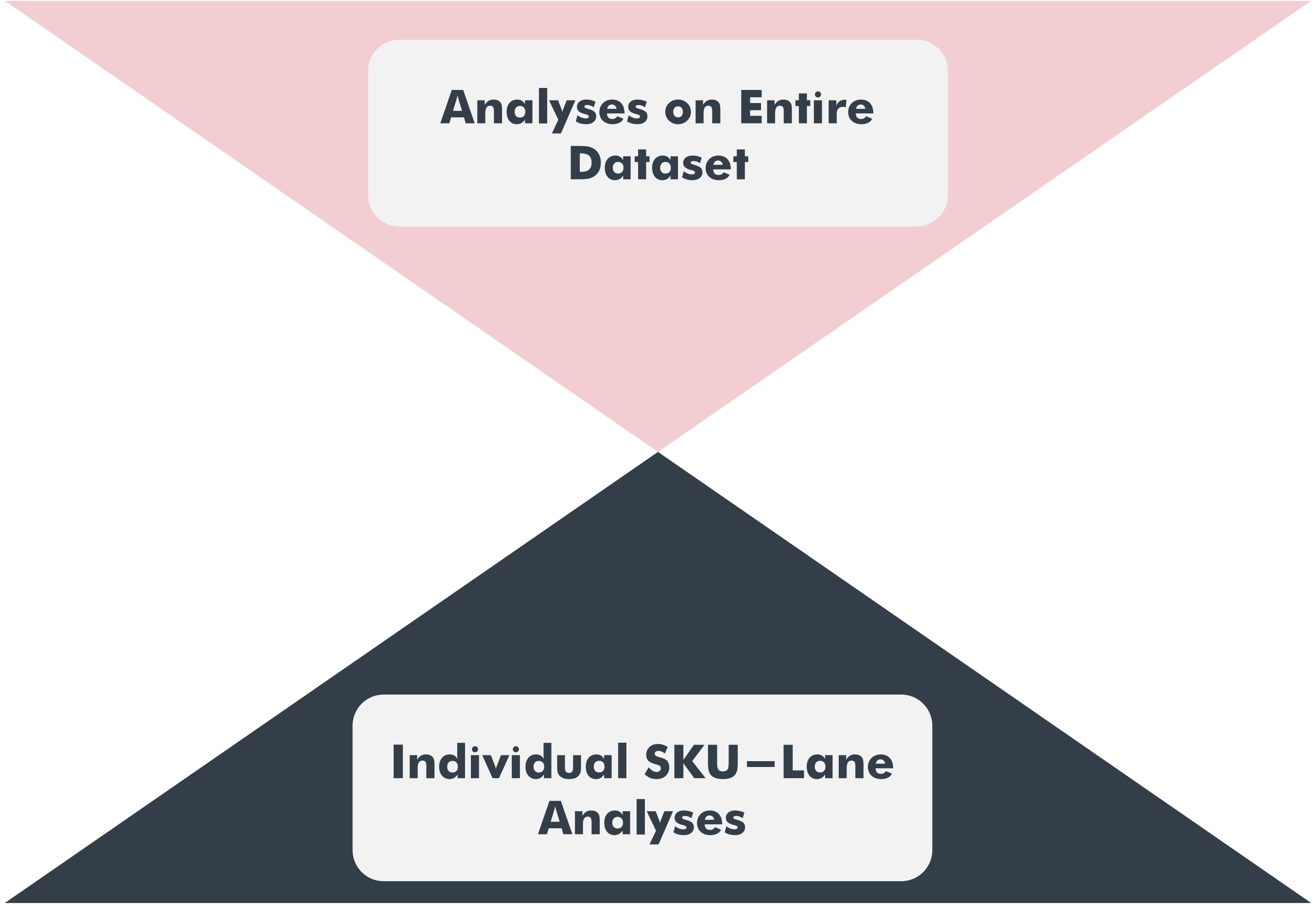
Method	Description	Equation
Method 1 – Naïve	Only the last data point is considered.	$\hat{x}_{t,t+1} = x_t$
Method 2 – Simple Mean	All the data points are considered. Any trend in the underlying data will lead to severe lagging.	$\hat{x}_{t,t+1} = \frac{\sum_i^t x_i}{t}$
Method 3 – Moving Average	Only the last n data points are considered.	$\hat{x}_{t,t+1} = \frac{\sum_{i=t+1-n}^t x_i}{n}$
Method 4 – Single Exponential Smoothing	This model is used to capture level of the time series. However, data is treated differently depending on its age.	$\hat{x}_{t,t+1} = \alpha x_t + (1-\alpha)\hat{x}_{t-1,t}$
Method 5 – Holt's Method (level and trend)	This model is used to forecast time series with a linear trend. A form of exponential smoothing, a higher weight is given to data that is more recent.	$\hat{x}_{t,t+\tau} = \hat{a}_t + \tau\hat{b}_t$ $\hat{a}_t = \alpha x_t + (1-\alpha)(\hat{a}_{t-1}+\hat{b}_{t-1})$ $\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1}$
Method 6 – Holt-Winter's Method (level, trend, and seasonality)	Model is used to forecast time series with both a linear trend and seasonality. A form of exponential smoothing, a higher weight is given to data that is more recent.	$\hat{x}_{t,t+\tau} = \hat{a}_t + \tau\hat{b}_t + \hat{F}_{t+\tau-P}$ $\hat{a}_t = \alpha(\frac{x_t}{\hat{F}_{t-P}}) + (1-\alpha)(\hat{a}_{t-1}+\hat{b}_{t-1})$ $\hat{b}_t = \beta(\hat{a}_t - \hat{a}_{t-1}) + (1-\beta)\hat{b}_{t-1}$ $\hat{F}_t = \gamma(\frac{x_t}{\hat{a}_t}) + (1-\gamma)\hat{F}_{t-P}$

vs.

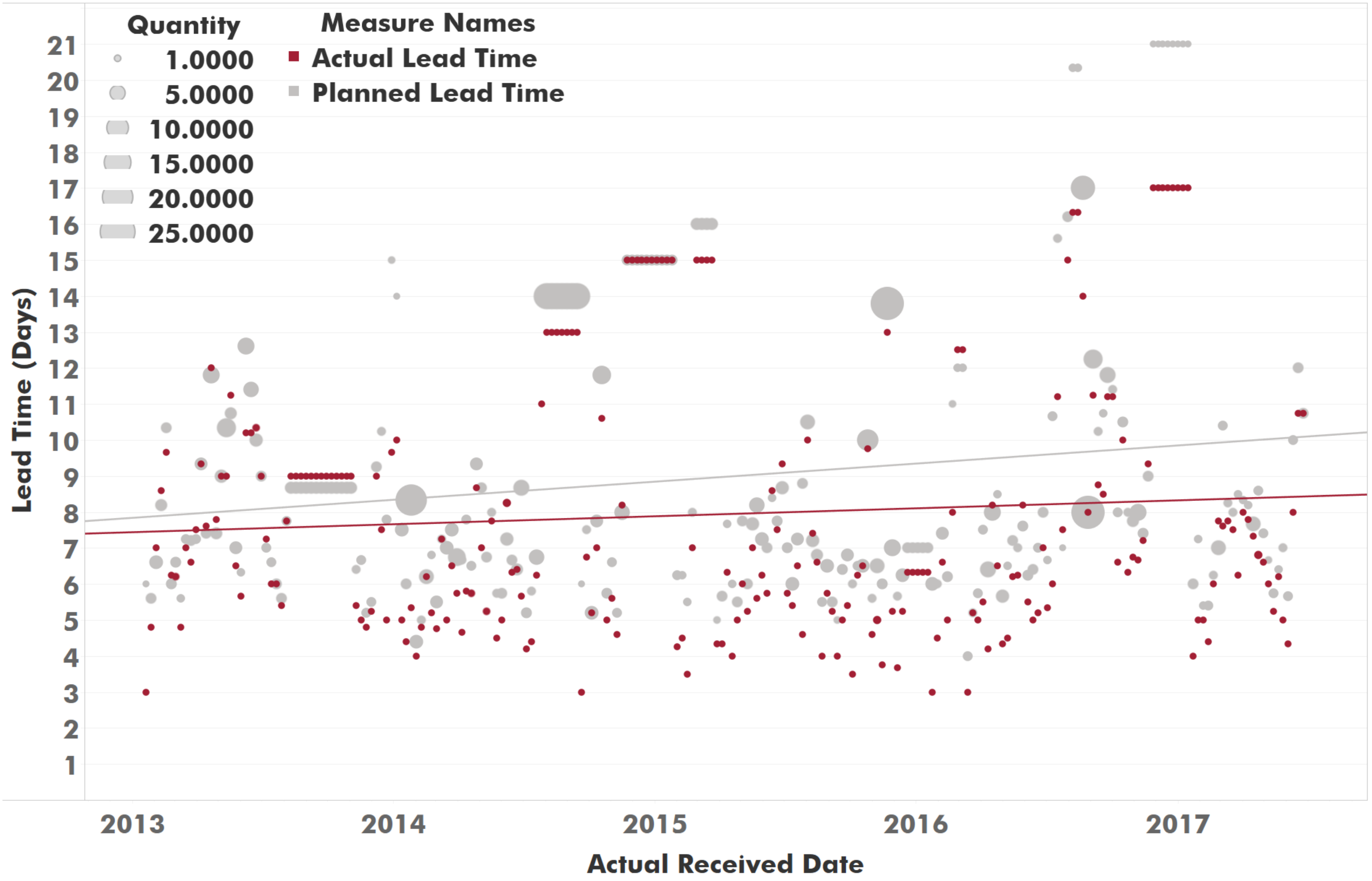
Baseline
Baseline 1 – Planned Lead Time
Baseline 2 – Regression Analyses

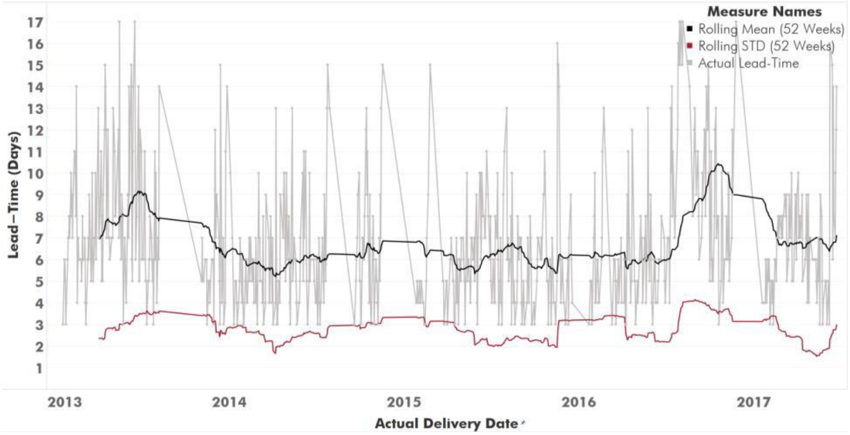
Darryl Yau

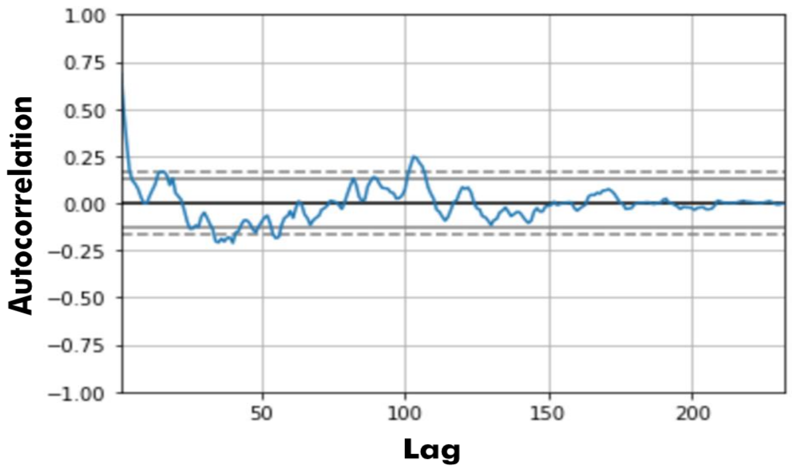
Bottom–up and Top–down Analyses



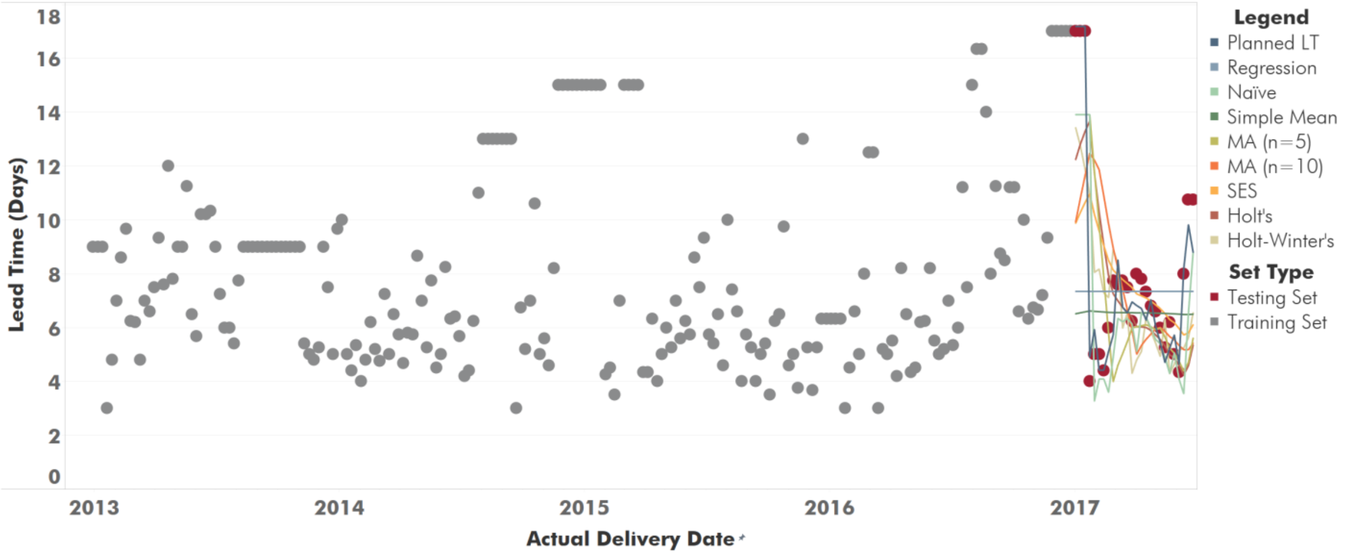
Analyzing one SKU–Lane

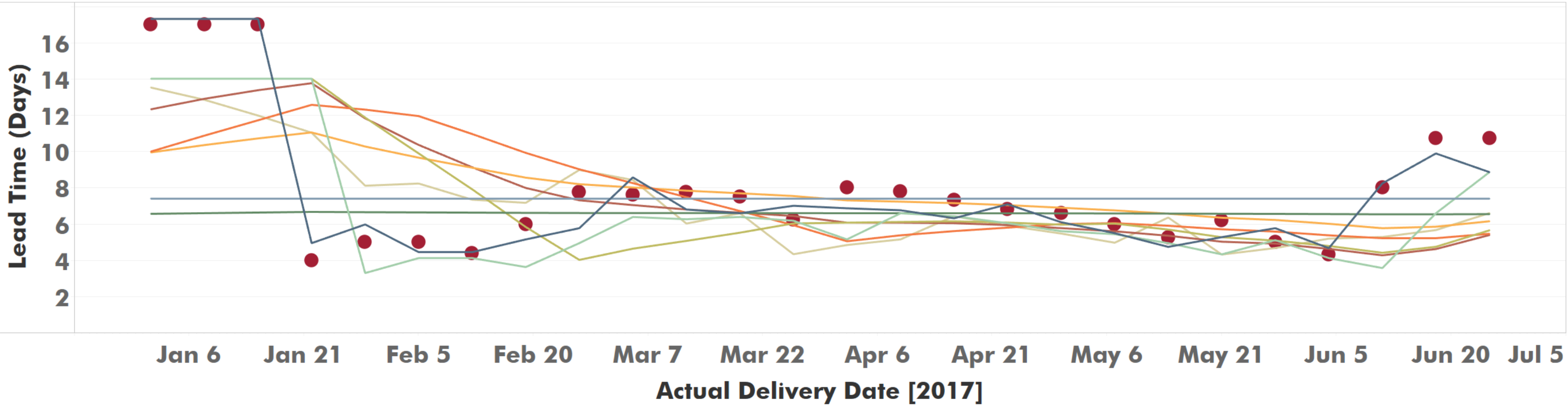






Forecasting on One SKU–Lane





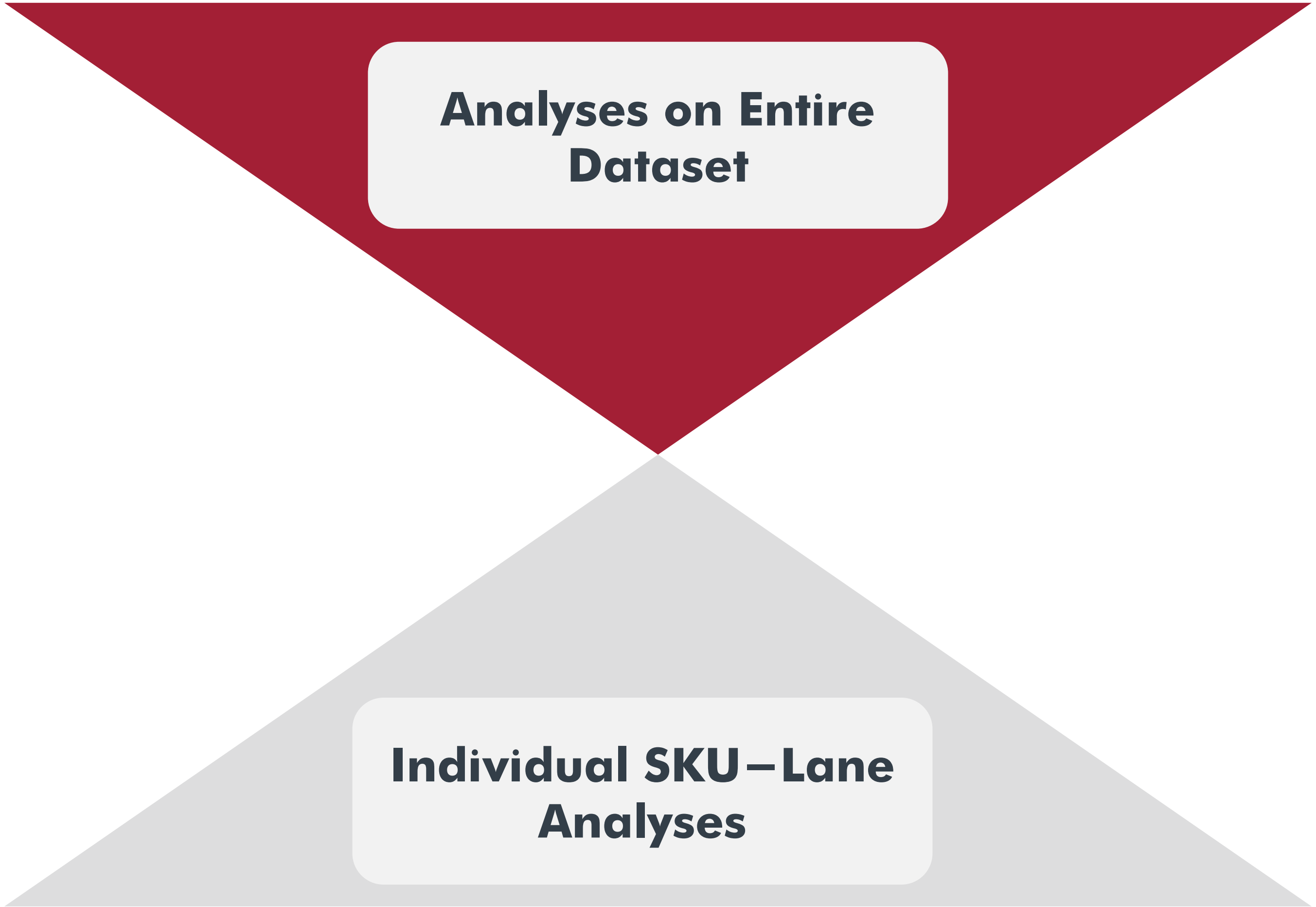
Massachusetts Institute of Technology

MIT Supply Chain MANAGEMENT

Forecasting on One SKU–Lane

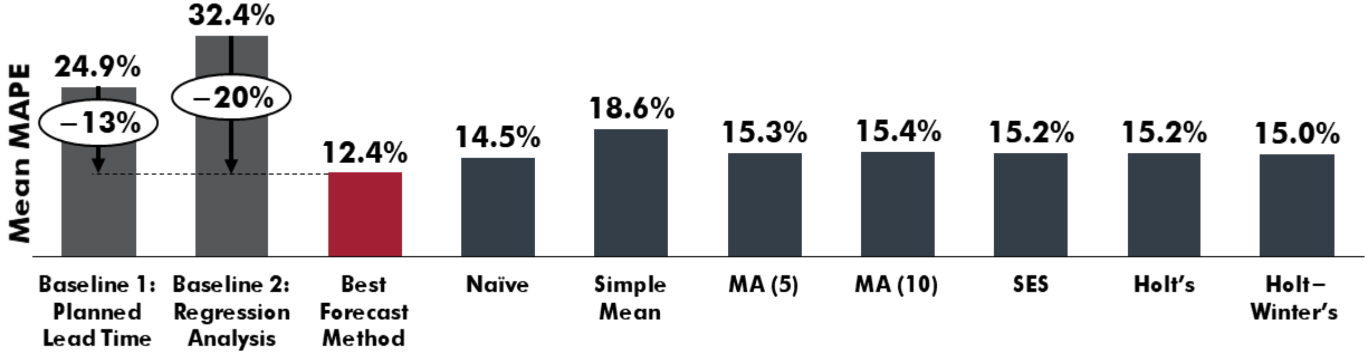
	Mean (Days)	Coefficient of Variation	Standard Deviation (Days)	Mean Deviation (MD) (Days)	Mean Average Deviation (MAD) (Days)	Mean Absolute Percent Error (MAPE)	Root Mean Squared Error (RMSE) (Days)	Mean Percent Error (MPE)
Baseline 1: Planned Lead Time	9.23	0.49	4.56	-1.34	1.40	18.45%	1.85	-17.71%
Baseline 2: Regression Analysis	8.98	0.00	0.00	-1.09	3.21	46.42%	3.84	-32.99%
Naïve Approach	8.13	0.50	4.05	-0.24	1.29	24.12%	**2.81**	-10.29%
Simple Mean Approach	8.00	0.01	0.04	-0.11	2.63	34.88%	3.70	-18.58%
Moving Average (n=5)	8.81	0.45	4.00	-0.93	2.45	44.73%	3.96	-27.59%
Moving Average (n=10)	9.34	0.34	3.16	-1.45	3.40	56.61%	4.62	-40.09%
Simple Exponential Smoothing ($\alpha = 0.1$)	9.53	0.20	1.93	-1.64	3.28	52.86%	3.93	-40.89%
Holt's Method ($\alpha = 0.2$ $\beta=0.05$)	9.13	0.39	3.55	-1.24	2.59	45.94%	4.08	-33.51%
Holt-Winter's Method ($\alpha = 0.2$ $\beta=0.05$ $\gamma=0.1$)	8.66	0.36	3.12	-0.77	2.21	37.33%	3.00	-22.62%

Bottom–up and Top–down Analyses

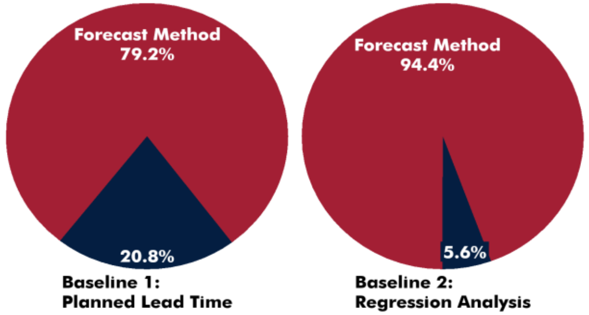


Analysis of Entire Dataset

Mean MAPE of Forecast Methods vs. the Baselines



Which had the Lower RMSE Result? (Forecast Method vs Baselines)

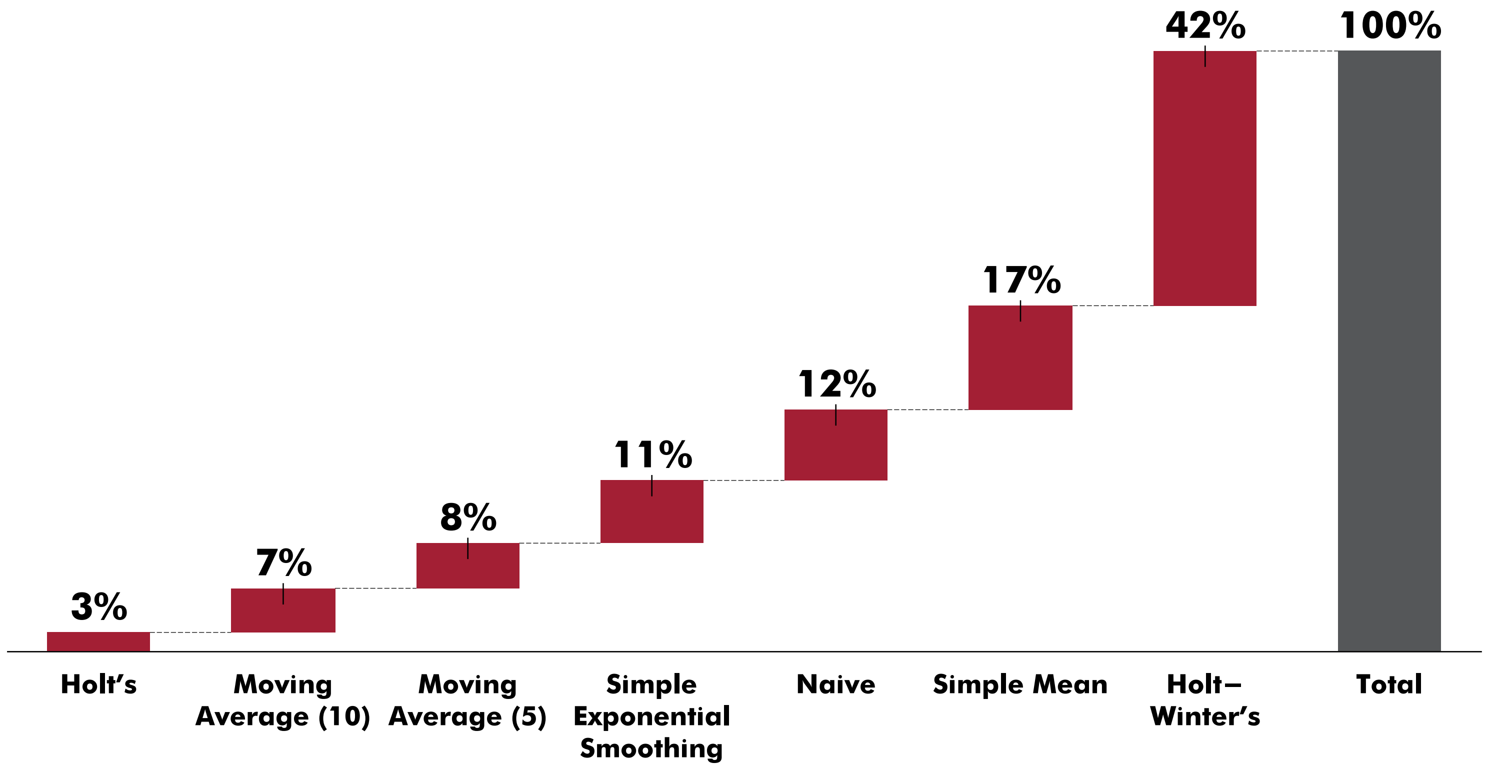


- Ran for over 2,500 SKU-Lane Combinations
- Best Forecast Method had a lower average MAPE than both baselines
- Using a single method had a lower average MAPE than both baselines

- Best Forecast Method, on average, performed better than both baselines

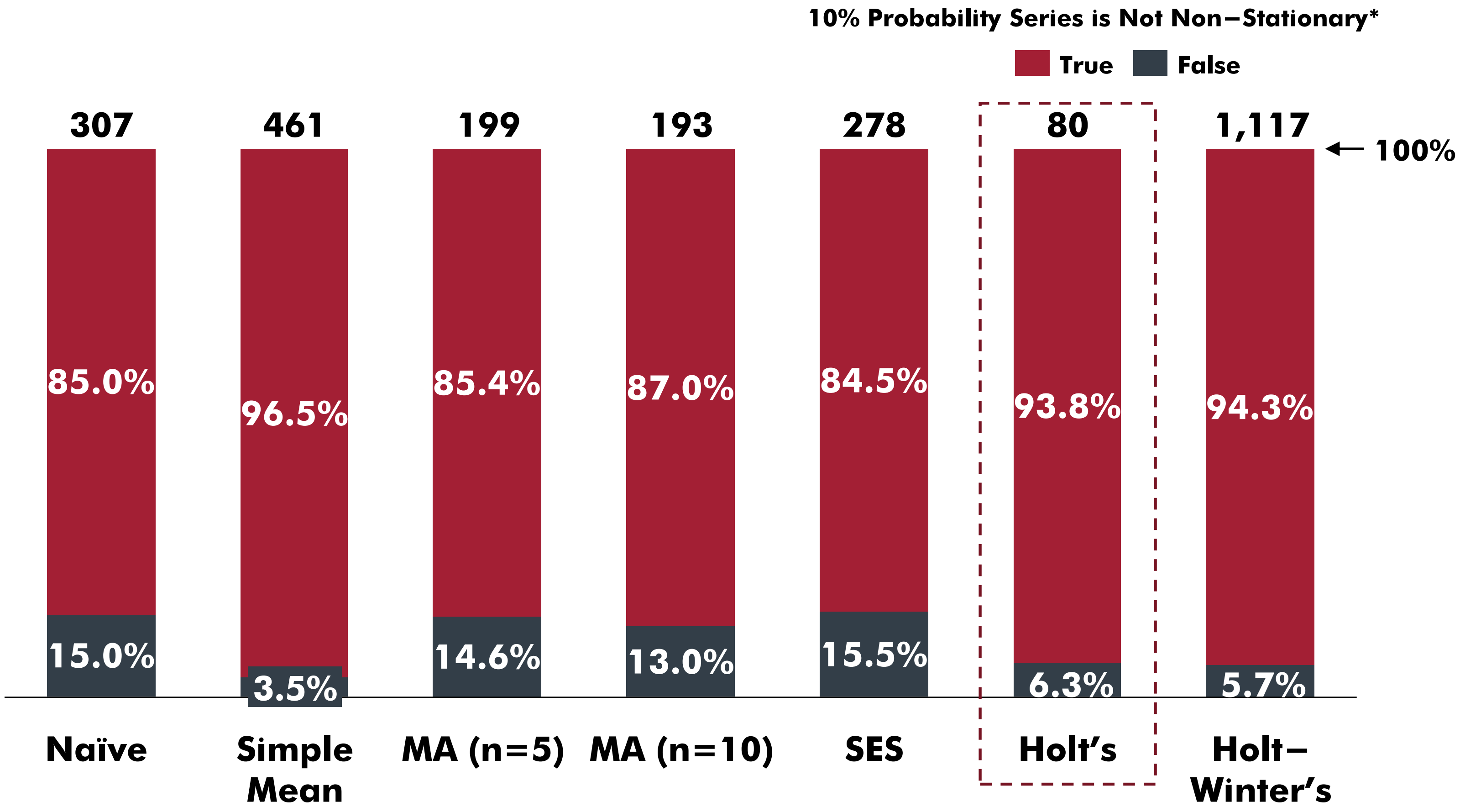
Which Forecast Method?

Forecast Method with the Lowest RMSE Value (gaps not filled)



- Holt-Winter's Method regularly performed better than other methods
- Holt's Method regularly performed worse than other methods

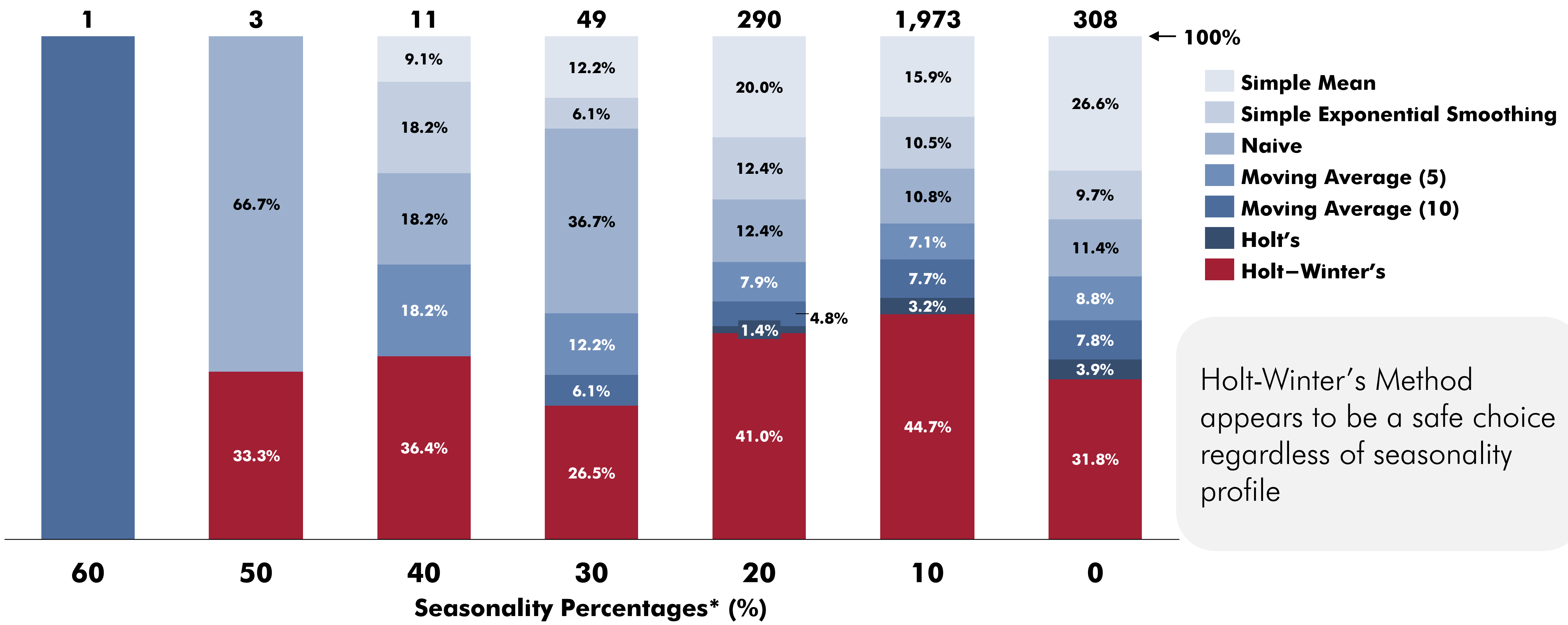
Trend did not appear to be a big factor in this dataset



*Based on Dickey–Fuller Test for Stationarity

Of the **five** SKU-Lanes that may have a trend in the data, only 1 appeared to have a significant trend.

Holt–Winter's Method appears to perform well regardless of the level of seasonality in the data



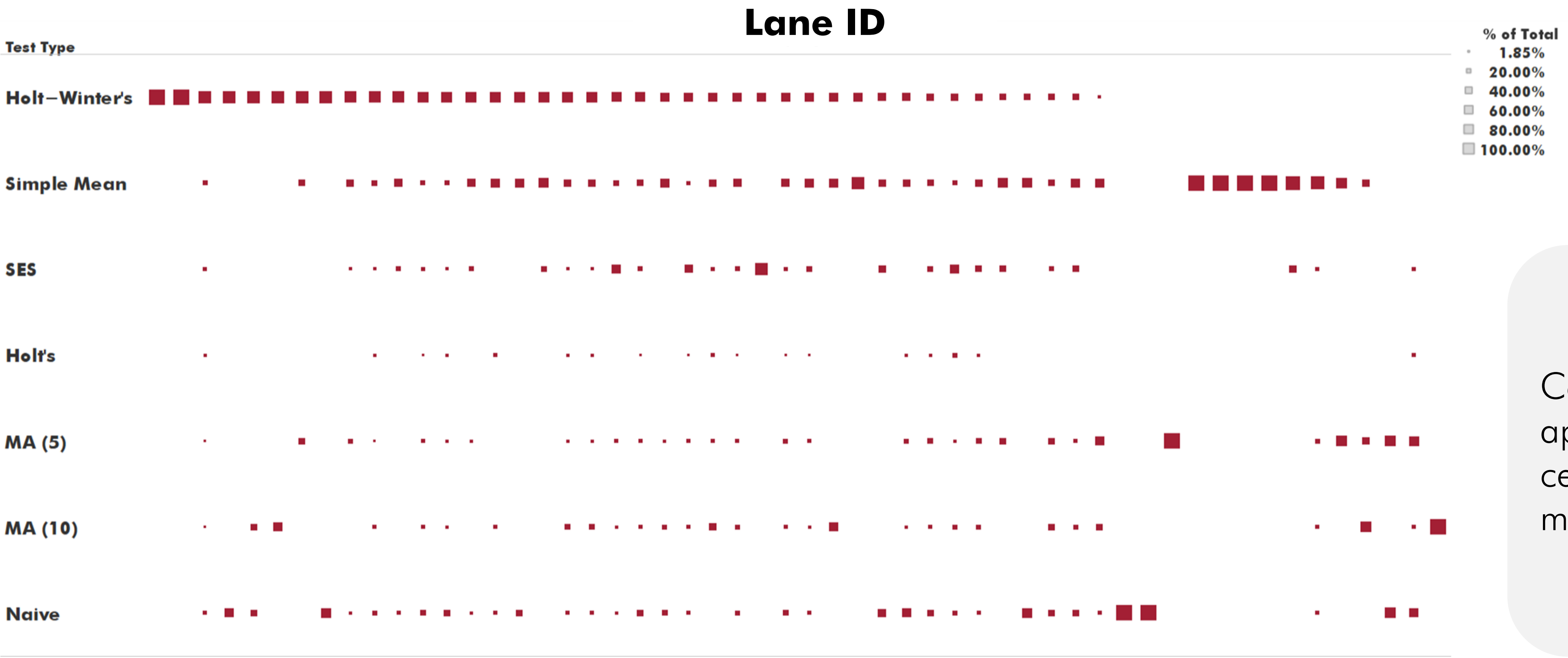
Holt-Winter's Method appears to be a safe choice regardless of seasonality profile

Massachusetts Institute of Technology

MIT Supply Chain Management

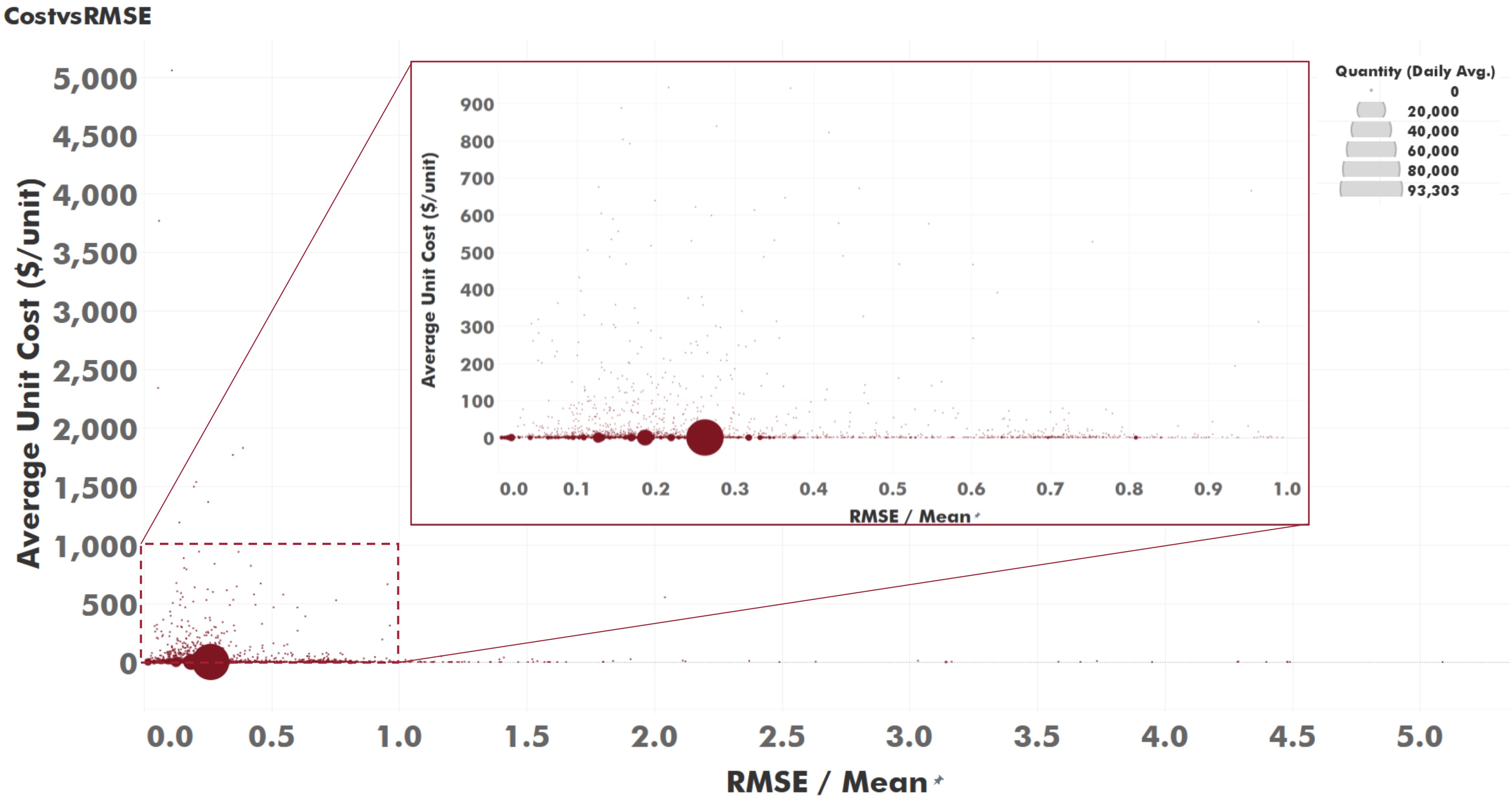
*Based on Time Series Decomposition

Signs of 'Lane Profile'



Certain Lanes appear to favor certain forecasting methods

Cost of item appears to be a factor in how well they currently plan



- Higher cost items have lower RMSE.
- Lower cost items have higher RMSE

Financial Implications

$$E[safety\ stock\ cost] = chk\sigma_{DL}$$

$$\sigma_{DL} = \sqrt{\mu_L \sigma_D^2 + \mu_D^2 \sigma_L^2}$$

Notations:

c: *unit cost* ($/unit)

h: *holding rate* ($/$ value/time). *For this analysis, h is assumed to be* 20%

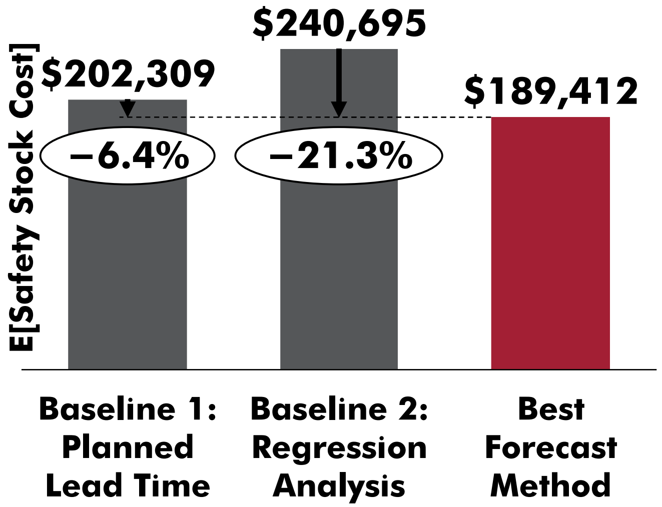
k: *safety factor*. For this analysis, service level is assumed to be 95%, thus k = 1.645

σ_{DL}: standard deviation of demand over lead time

σ_i: *standard deviation of demand* (D) *or lead time* (L)

μ_i: *mean of demand* (D) *or lead time* (L)

Estimated Safety Stock Costs



Potential cost savings by reduction in their safety stock

Bringing it together....

Planning Process

Standard Lead–Time Variables

When Order is Placed

When Order is Received

Predictive Lead Time Variable

Other Business Constraints/ Decisions

Planned Lead Time (based on Predictive LT)

Vs.

Actual Lead Time

Accuracy Improved

Improving Supply Chain Planning with Advanced Analytics

Analyzing Lead Time as a Case Study

1

Baseline Current State

Using historical data to predict lead times can reduce the error between plan and actual

Reduces Safety Stock costs and manual labor costs

To what extent are the lead time variables found in the Enterprise Resource Planning (ERP) system are used in predicting lead time and how accurate is the prediction?

2

Propose Improved Future State

Can predictive analytics on historical lead time data be used to improve the forecast accuracy and what are the benefits in doing so?

Today's Agenda

1. Background
2. Data
3. Analysis and Results
4. **Conclusion**

Conclusions and Considerations

Using Predictive Lead Time can reduce the error between plan and actual

- **Safety Stock Cost** ↓
- **Manual Labor in Planning and Re-Planning** ↓
- **Manual Labor in PO Management** ↓

Consider:

- **Assigning Forecast Method by Lane**
- **Implementing on High Volume, Low Cost Items**
- **Categorizing SKU-Lanes by Trend, Seasonality**

Questions?

Backup Slides

Today's Agenda

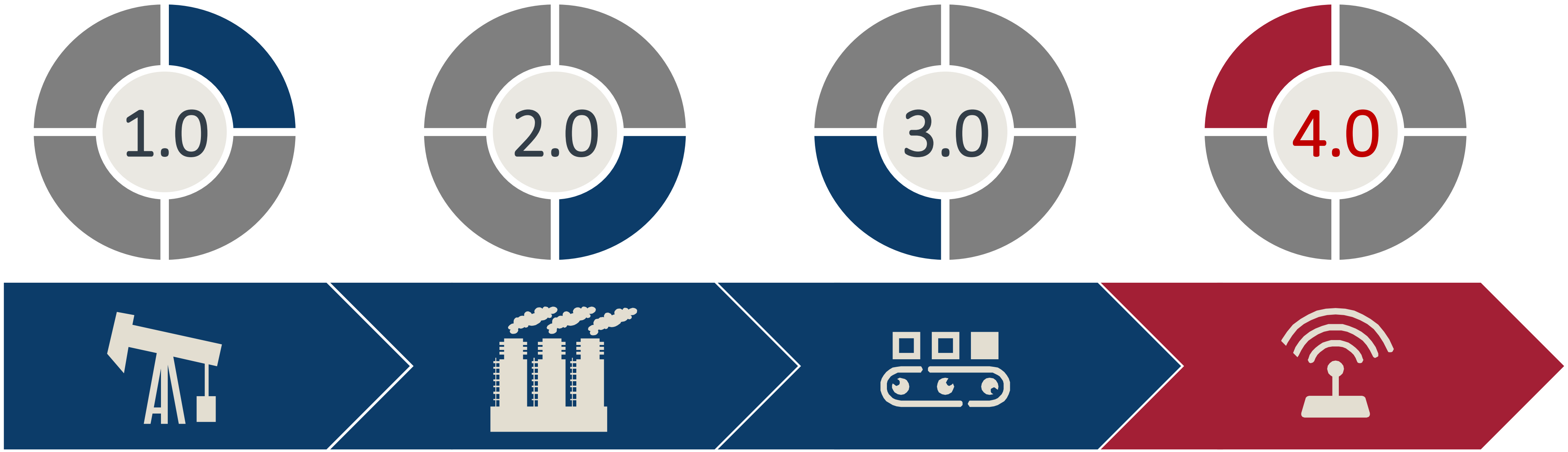
1 Background

2 Data

3 Analysis and Results

4 Conclusion

Industry 4.0



1784 – Mechanization, water power, steam power

1870 – Mass production, assembly line, electricity

1969 – Computer and automation

Present – Cyber physical systems / digital transformation

Important for 2 reasons:

1. Access to more data for analyses
2. Evolution of a "digital supply chain's" role in planning

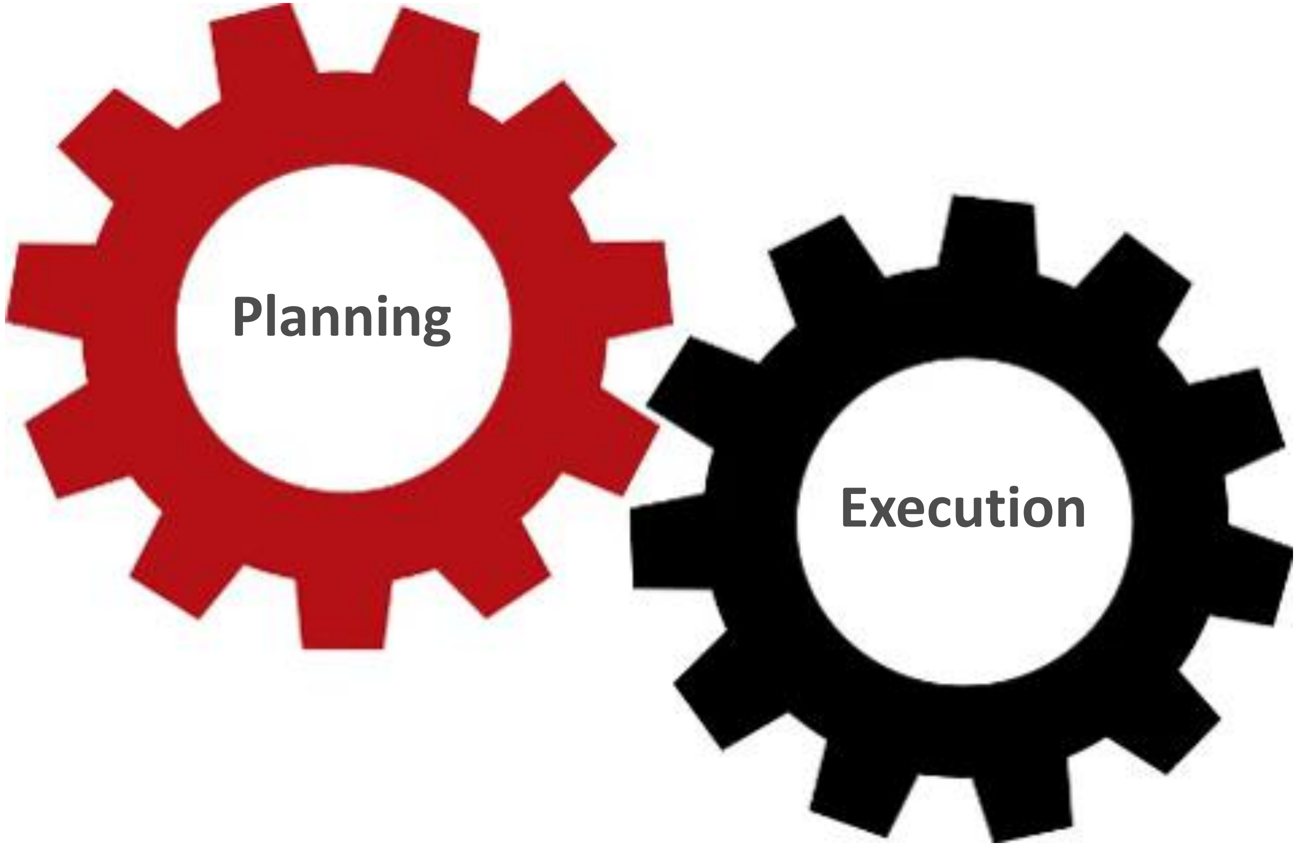
Digital Supply Chain

'Digital Copy' implies a level of detail in their similarity

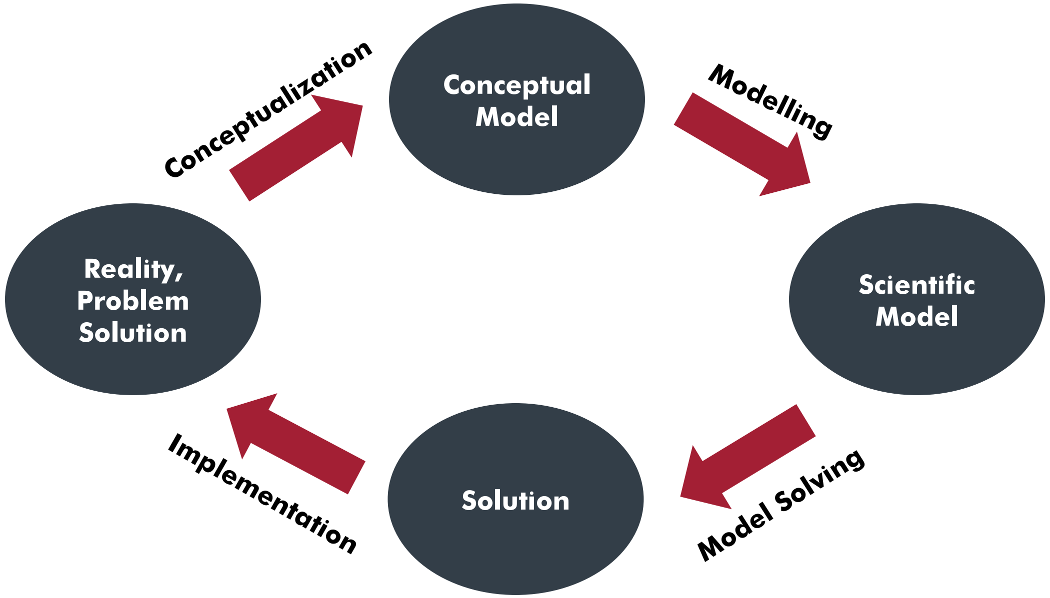
≈

Supply Chain Planning Systems (e.g., ERP, APS) are becoming increasingly more complex in order to more accurately model the complexities of the physical supply chain

… because a plan that does not reflect reality will much manual intervention during execution



How abstract should we conceptualize the problem?



More abstract?

- Might solve the wrong problem
- Require more manual labor to supplement the decision making process

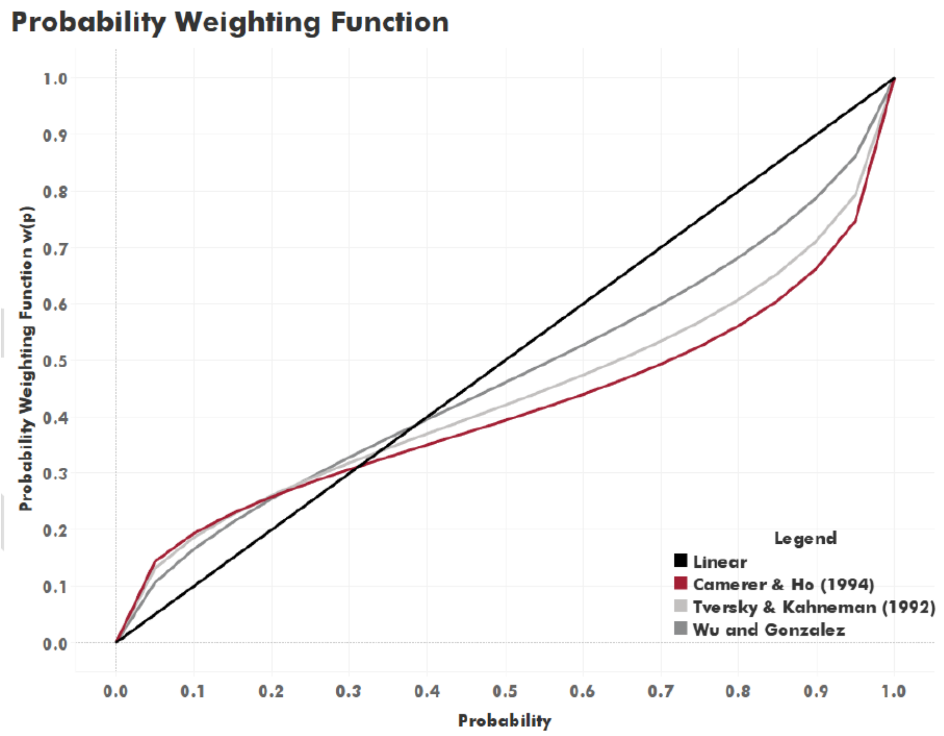
Errors of the Third Kind

Ian Mitroff

- First and Second Kind were about Accuracy – False Positive and False Negatives
- Third Kind (Mitroff, 1974) – Solving the wrong problem by choosing the wrong problem representation
  - Could be more problematic than first and second kind errors

Humans Making Decisions?

- Kahneman & Tversky, 1979
  - Prospect Theory – People make decisions based on potential value rather than the outcome
- Wu and Gonzalez, 1999
  - Further studies on Prospect Theory. Analyzed different probability weighting functions
- Schweitzer & Cachon, 2000
  - Managers consistently deviated from the optimal order point for newsvendor problem, even with feedback and additional training



To Summarize...

We need a planning process that is:

Data Driven

More Complex, More Like the Physical

Less Human Intervention