The Path to Actionable Insights

Cheryl Pammer,   Minitab LLC

Minitab

Agenda

- The Digital Transformation and YOU
- Machine Learning and the Digital Transformation
- Success Story
- Where Are We on this Journey?

Digital Transformation

- A new level of organizing and controlling the entire value chain with the life cycle of products and services.
- The availability of all relevant information in real time.

Configuring the Digital Transformation

- Where will sensors be placed?
- What data will be collected and at what frequency?
- What analytic methods will be used?

Identifying and predicting defects is great, but not if you don't know what is driving them and how to reduce or eliminate them.

Introduction to Machine Learning

MACHINE LEARNING = **MODERN STATISTICS** + **COMPUTER SCIENCE**

Classic statistical techniques, such as linear regression, can struggle with these issues. Machine learning methods were created to solve these problems and they are often able to do this **automatically**.

Machine Learning is Not New

Logistic Regression
1958

Support Vector Machine
1963

K-nearest neighbors
1967

Artificial Neural Networks
1975

Decision Tree
1986

Random Forest
1995

What Has Changed?

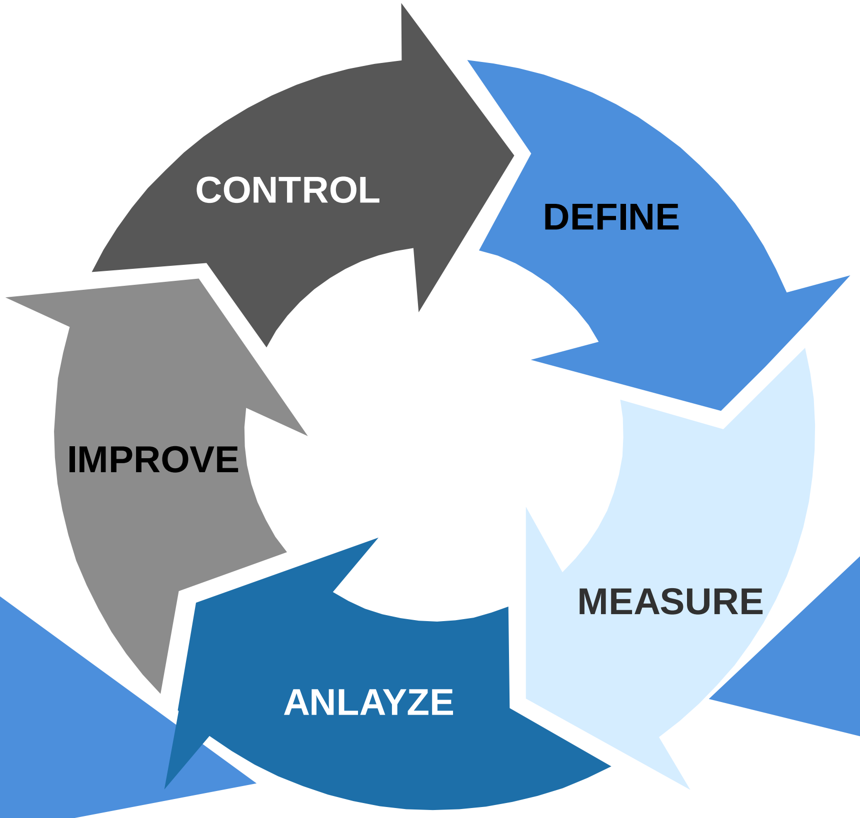
The volume of data available now due to digitalization.

- Compute power and storage has become a commodity with faster access.
- More data is available.
- This will continue to grow as more data is stored in digital form.

Machine Learning Enhances Lean & CI Programs

ANALYZE

ML Value-Add
What are the root causes of the problem?
Explore unknown factors impacting performance, broaden to thousands of variables in real-time.



MEASURE

ML Value-Add
How big is the problem?
Add new data layers and methods to better define the problem you are looking to understand.

The Analytics Roadmap

DESCRIPTIVE
Report

DIAGNOSTIC
Discover & Explore

PREDICTIVE
Forecast

PRESCRIPTIVE
Anticipate

Learning from the past.
Discovering the present.
Predicting the future.
Prescribing action.

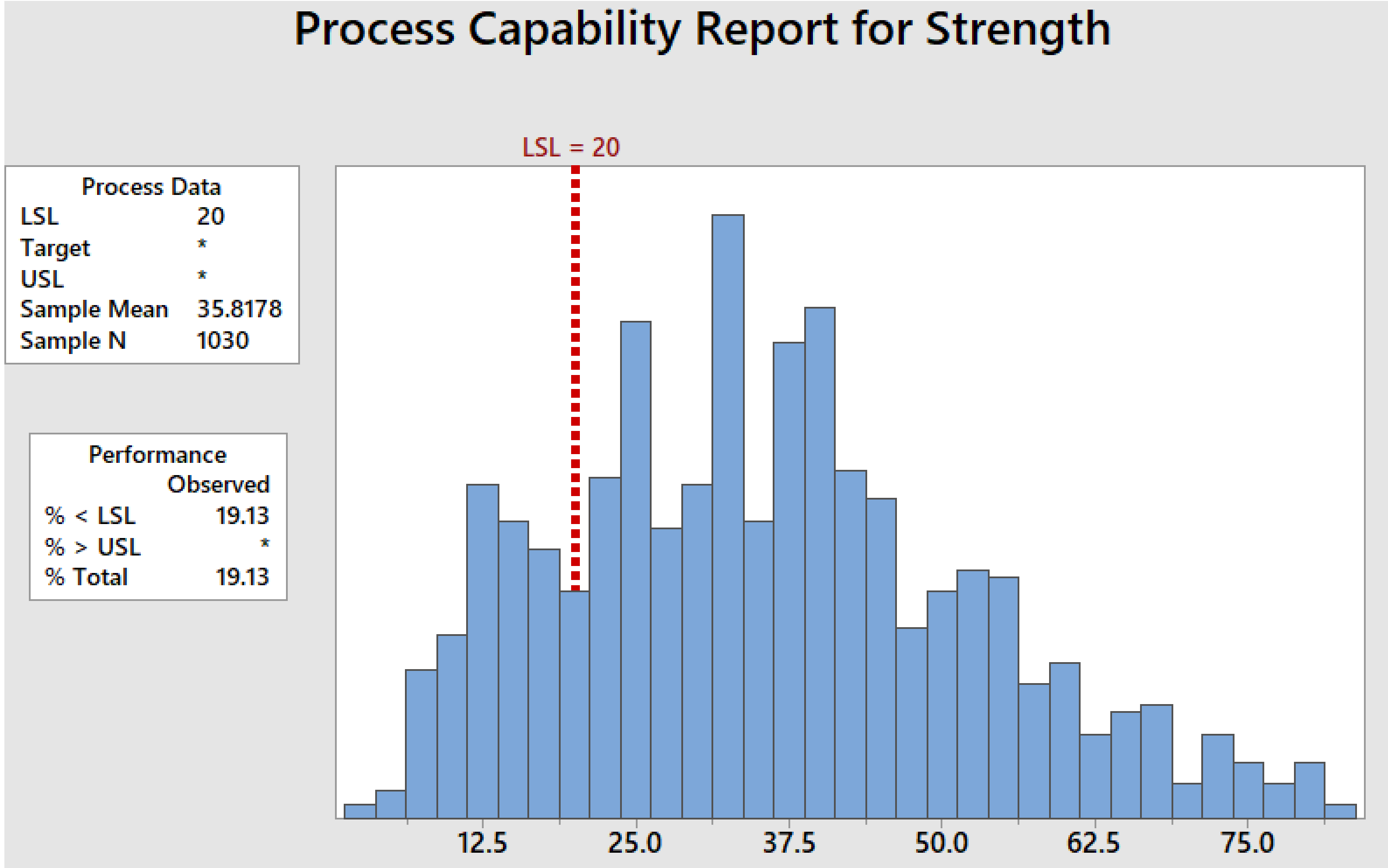
Example: Concrete Strength

How can we improve concrete strength?

- Investigate process data to determine the relative importance and nature of the relationship between potential predictors and concrete strength.
- Design an experiment around potentially optimal factors and settings found in the first stage.
- Use the model obtained from the DOE to perform simulation to further explore the complex behavior of the system.

Original data source: I-Cheng Yeh, "Modeling of strength of high performance concrete using artificial neural networks," Cement and Concrete Research, Vol. 28, No. 12, pp. 1797-1808 (1998).

Descriptive: Histogram and Descriptive Statistics



- Defect rate is 19.13%.
- Process is not capable of meeting lower specification of 20 MPa.

Diagnostic: Regression Model

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	44	232787	5290.6	95.82	0.000
Cement	1	693	693.0	12.55	0.000
Blast_Furnace_Slag	1	356	355.9	6.45	0.011
Fly_Ash	1	144	143.8	2.61	0.107
Water	1	1562	1561.6	28.28	0.000
Superplasticizer	1	928	928.4	16.82	0.000
Coarse_Aggregate	1	512	512.2	9.28	0.002
Fine_Aggregate	1	785	785.4	14.22	0.000
Age	1	2	1.8	0.03	0.856
Cement*Cement	1	439	439.4	7.96	0.005
Blast_Furnace_Slag*Blast_Furnace_Slag	1	118	117.9	2.14	0.144
Fly_Ash*Fly_Ash	1	16	16.1	0.29	0.590
Water*Water	1	1414	1414.5	25.62	0.000
Superplasticizer*Superplasticizer	1	467	467.1	8.46	0.004
Coarse_Aggregate*Coarse_Aggregate	1	243	243.4	4.41	0.036
Fine_Aggregate*Fine_Aggregate	1	855	854.5	15.48	0.000
Age*Age	1	12186	12185.8	220.70	0.000
Cement*Blast_Furnace_Slag	1	163	163.2	2.96	0.086
Cement*Fly_Ash	1	16	16.3	0.30	0.587
Cement*Water	1	1867	1866.7	33.81	0.000
Cement*Superplasticizer	1	1409	1409.4	25.53	0.000
Cement*Coarse_Aggregate	1	284	283.8	5.14	0.024
Cement*Fine_Aggregate	1	504	504.1	9.13	0.003
Cement*Age	1	82	82.4	1.49	0.222
Blast_Furnace_Slag*Fly_Ash	1	3	3.2	0.06	0.809
Blast_Furnace_Slag*Water	1	804	804.3	14.57	0.000
Blast_Furnace_Slag*Superplasticizer	1	883	883.0	15.99	0.000
Blast_Furnace_Slag*Coarse_Aggregate	1	210	210.4	3.81	0.051
Blast_Furnace_Slag*Fine_Aggregate	1	215	214.7	3.89	0.049
Blast_Furnace_Slag*Age	1	385	385.3	6.98	0.008
Fly_Ash*Water	1	756	756.4	13.70	0.000
Fly_Ash*Superplasticizer	1	1465	1464.9	26.53	0.000
Fly_Ash*Coarse_Aggregate	1	41	40.5	0.73	0.392
Fly_Ash*Fine_Aggregate	1	66	66.0	1.19	0.275
Fly_Ash*Age	1	354	353.9	6.41	0.012
Water*Superplasticizer	1	513	513.0	9.29	0.002
Water*Coarse_Aggregate	1	1364	1363.6	24.70	0.000
Water*Fine_Aggregate	1	1399	1398.8	25.33	0.000
Water*Age	1	1	0.6	0.01	0.916
Superplasticizer*Coarse_Aggregate	1	700	700.3	12.68	0.000
Superplasticizer*Fine_Aggregate	1	1010	1010.0	18.29	0.000
Superplasticizer*Age	1	145	145.3	2.63	0.105
Coarse_Aggregate*Fine_Aggregate	1	433	432.7	7.84	0.005
Coarse_Aggregate*Age	1	9	9.5	0.17	0.678
Fine_Aggregate*Age	1	96	96.3	1.74	0.187
Error	985	54386	55.2		
Lack-of-Fit	951	53532	56.3	2.24	0.002
Pure Error	34	854	25.1		
Total	1029	287173			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)	10-fold S	10-fold R-sq
7.43062	81.06%	80.22%	78.82%	7.73574	78.54%

Challenges with Regression Model

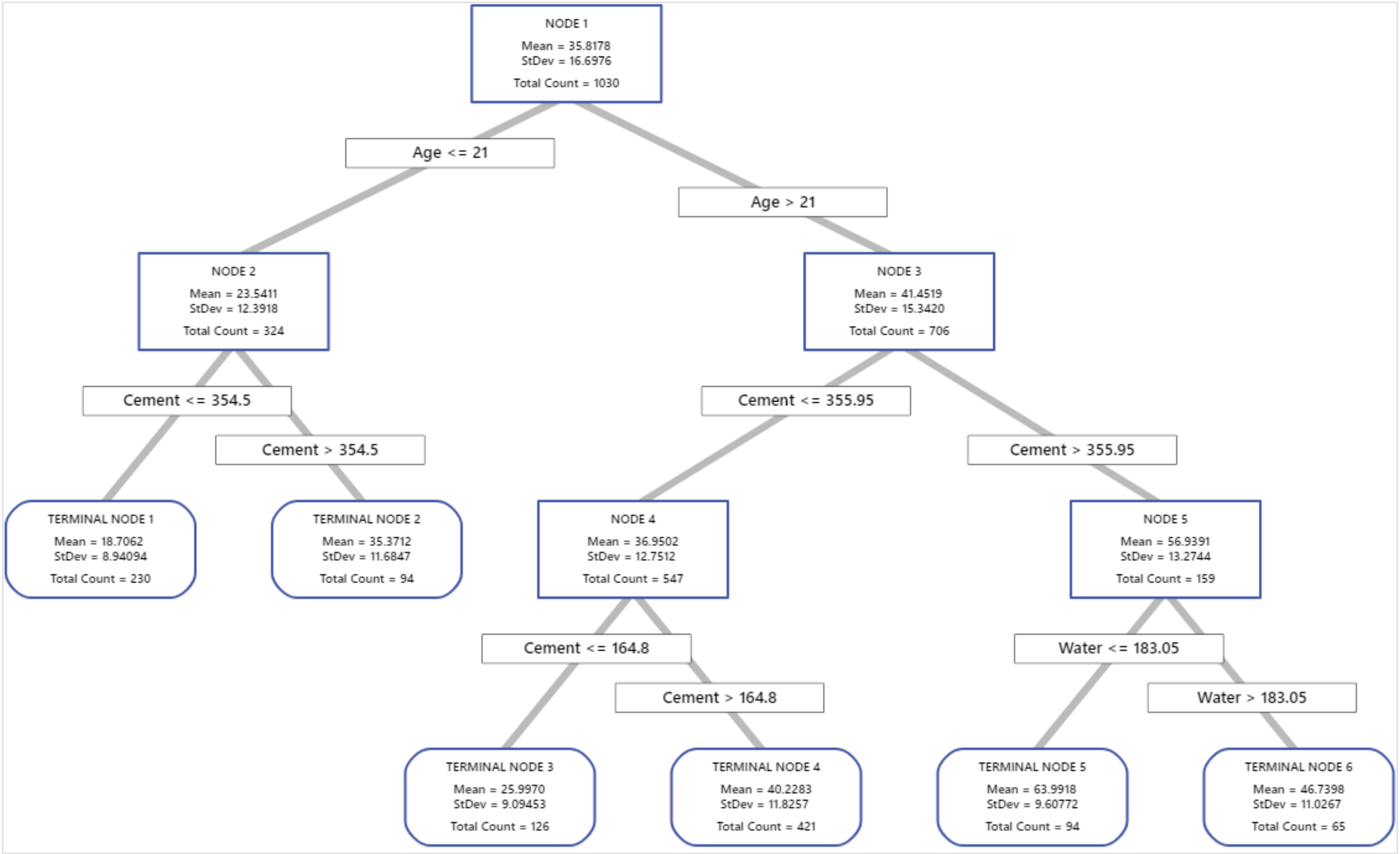
- Many significant p-values.
- Determination of predictors to include in model is challenging.
- Relationships appear nonlinear.
- Complex interactions exist.
- Local effects may exist.
- Predictive accuracy may be compromised (R-sq(pred) = 78.54%).

Diagnostic: CART Model

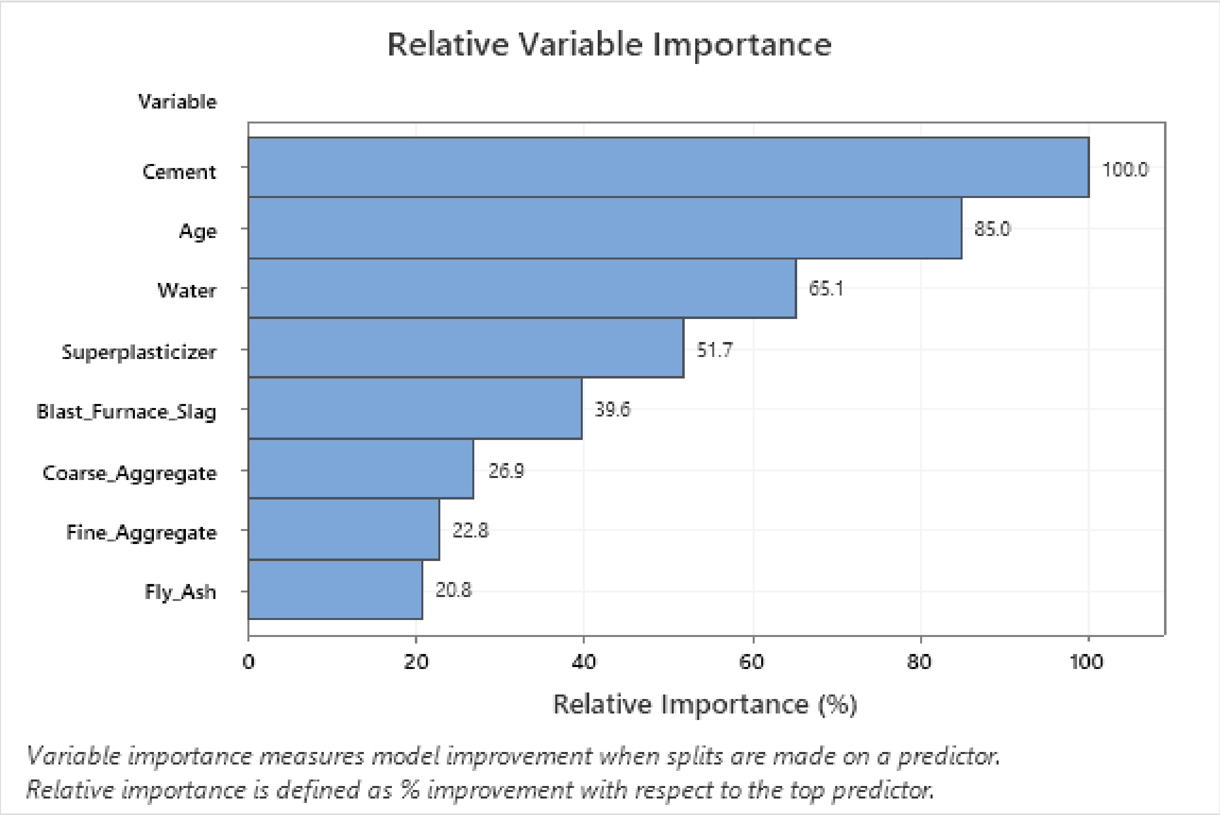
Benefits of CART Model

- Able to determine key drivers
- Handles nonlinear relationships and complex interactions.
- Better predictive accuracy.
- Can visualize important relationships.

Alternative Tree Diagram

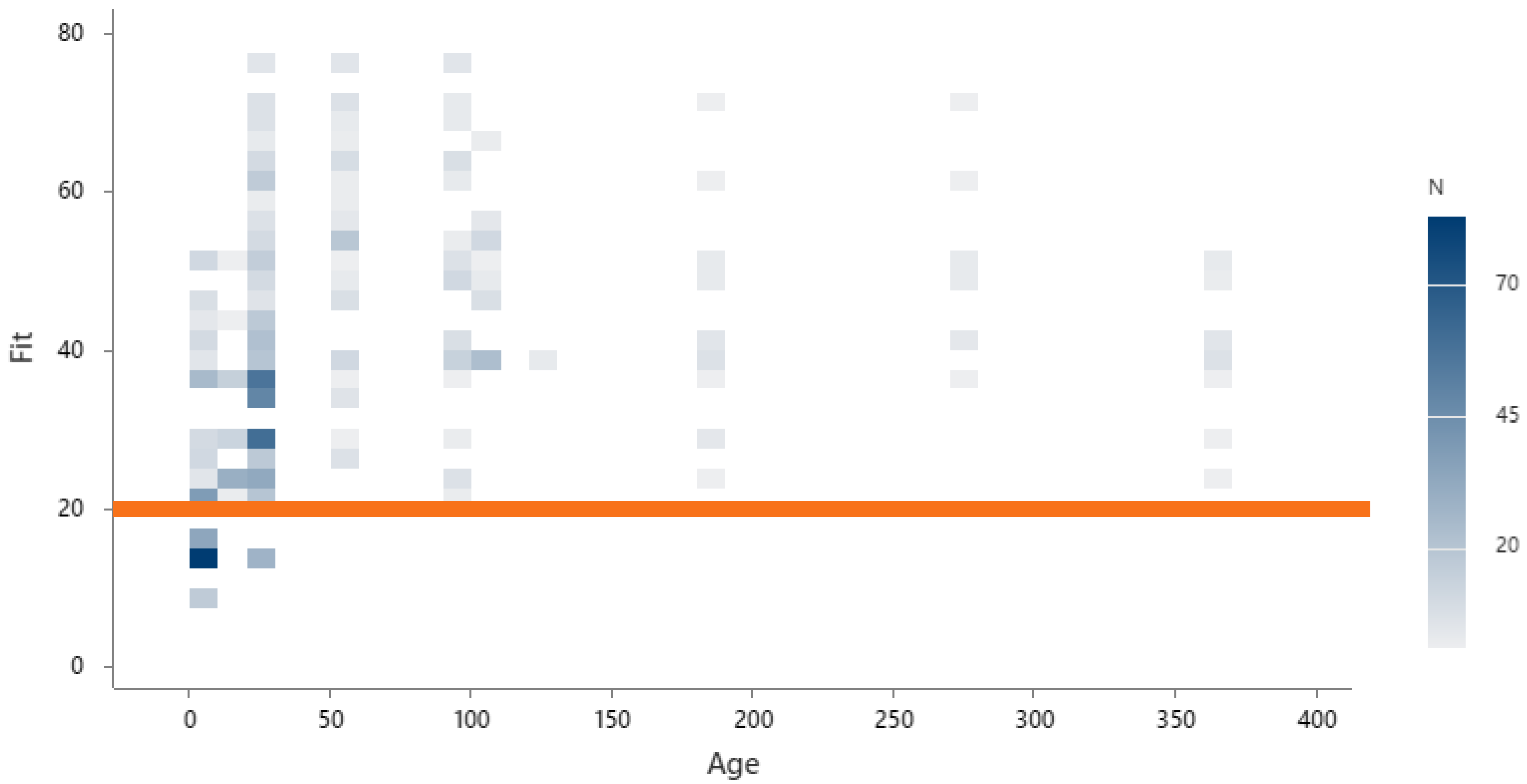


Diagnostic: CART Variable Importance



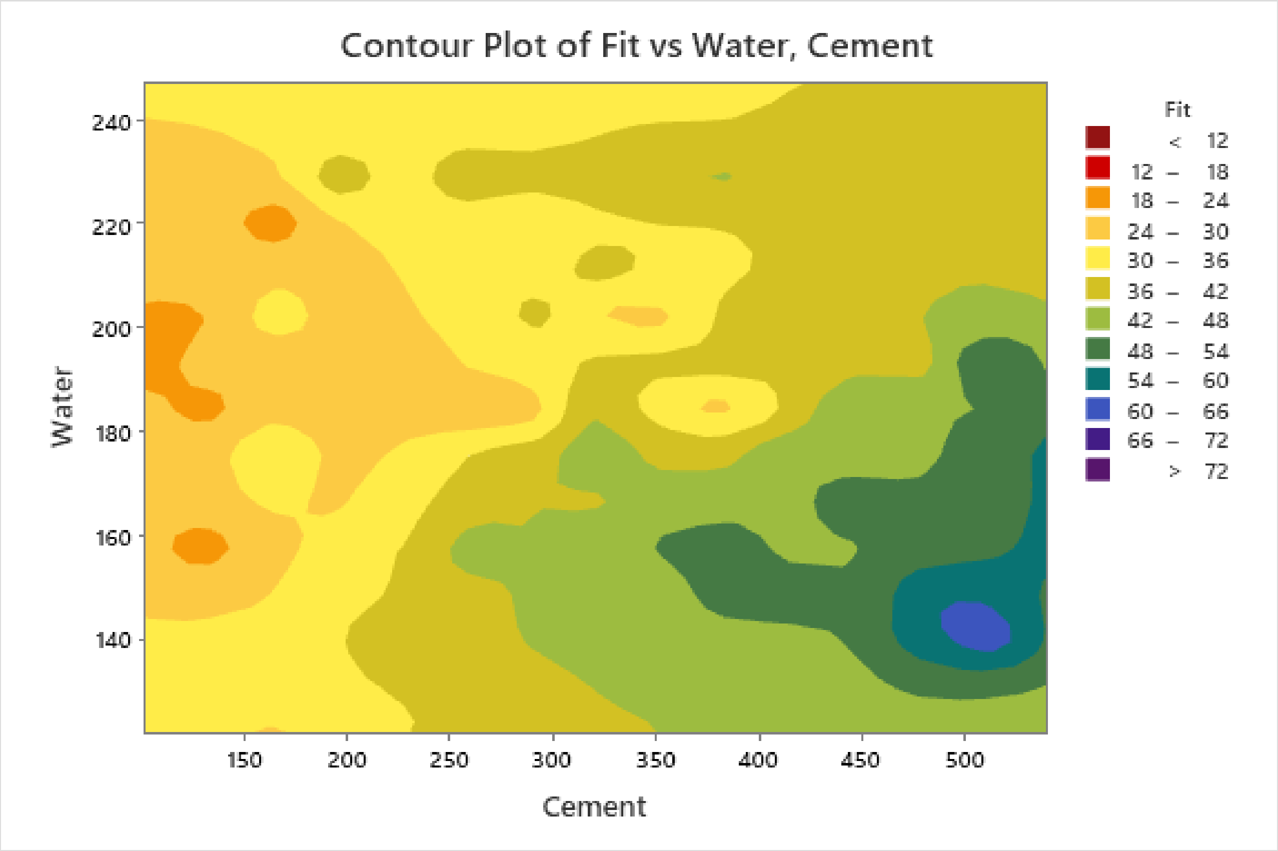
- Age and Cement are most important drivers of compressive strength.
- Importance scores are relative with the most important variable having an importance score of 100. For example, Cement is roughly twice as important as Superplasticizer.

Diagnostic: Binned Scatterplot of CART Model Fits



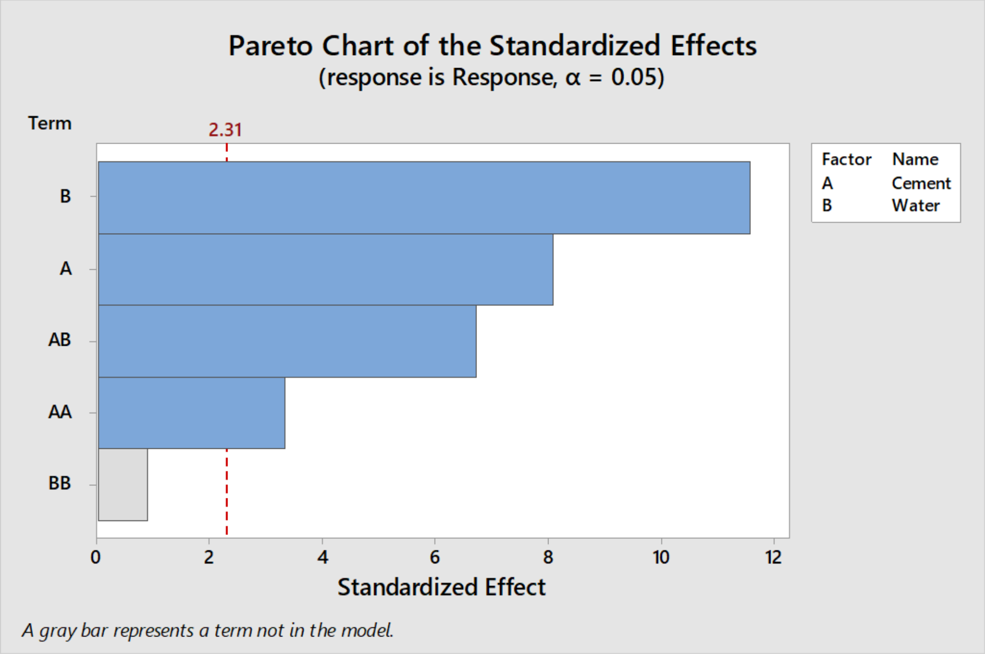
- Age is an important driver of compressive strength, but it is not controllable in practice.
- The effect of age is nonlinear.

Diagnostic: Contour Plot of CART Model Fits



- Water and Cement appear to interact.
- Low values of Water and High values of Cement appear to maximize strength.
- Use DOE to fine tune optimal levels of water and cement and validate results.

Diagnostic: Analyze Response Surface Design



Model Summary

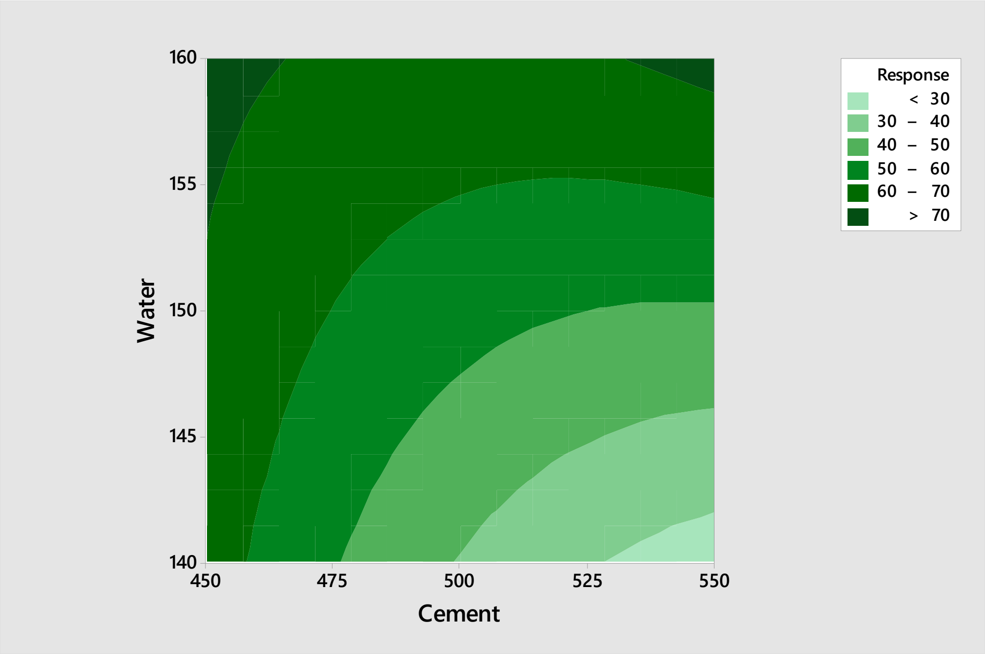
S	R-sq	R-sq(adj)	R-sq(pred)
2.97102	96.96%	95.43%	90.13%

Regression Equation in Uncoded Units

Response = 1984 - 5.375 Cement - 8.58 Water + 0.002186 Cement*Cement + 0.01996 Cement*Water

- DOE confirmed the relationship between Strength and Concrete and Water.
- Predictive accuracy of model is adequate (R-sq(pred) = 90.13%.)

Predictive: Response Surface Design Contour Plot

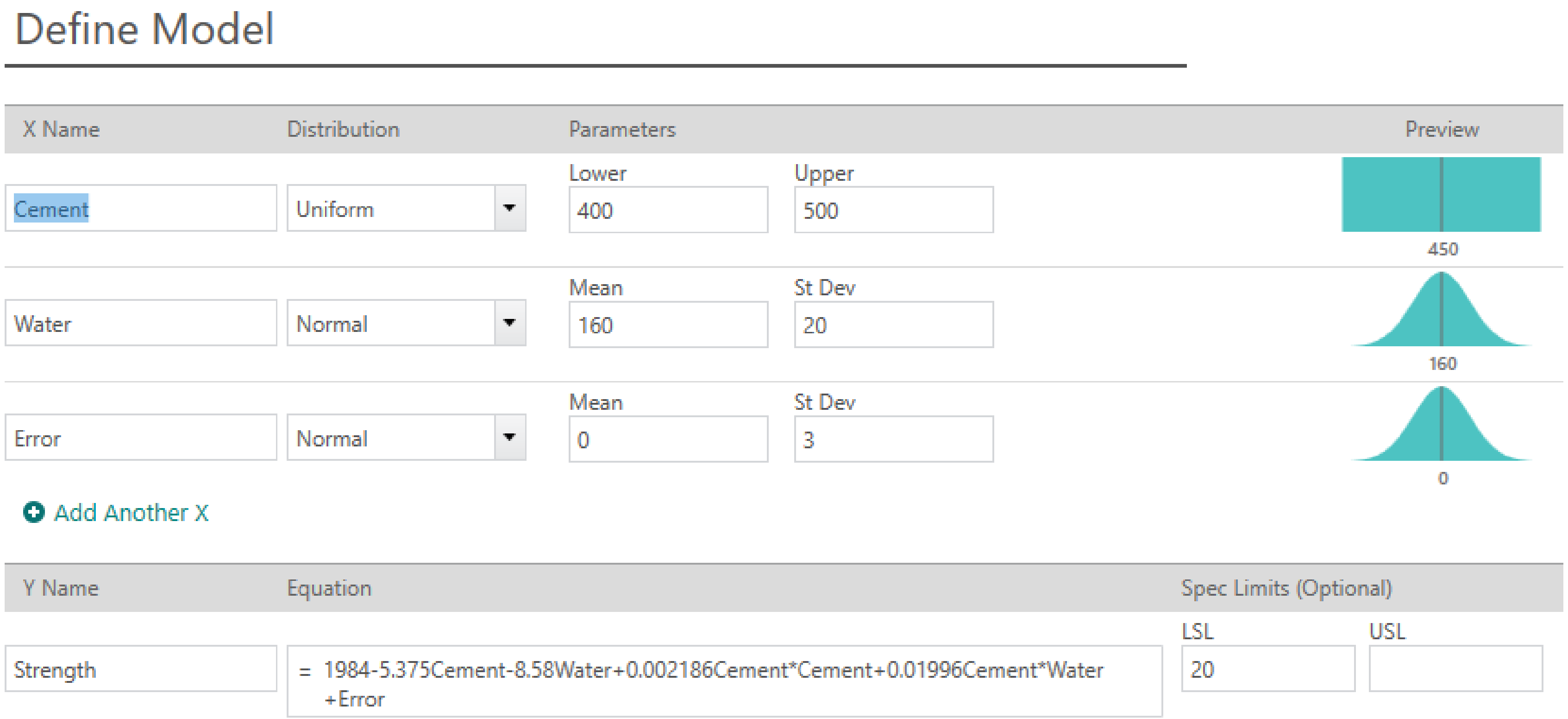


Prediction

Fit	SE Fit	95% Lower Confidence Bound	95% Lower Prediction Bound
72.8312	2.57298	68.0467	65.5227

- The optimal settings for concrete strength are at Water = 160 and Cement at either 450 or 550.
- When you **fix** Water at 160 and Cement at 450, the predicted strength is 72.8.
- You can be 95% confident that fixing Water at 160 Cement at 450 will result in an individual concrete sample that exceeds 65.52 in strength.

Prescriptive: Monte Carlo Simulation



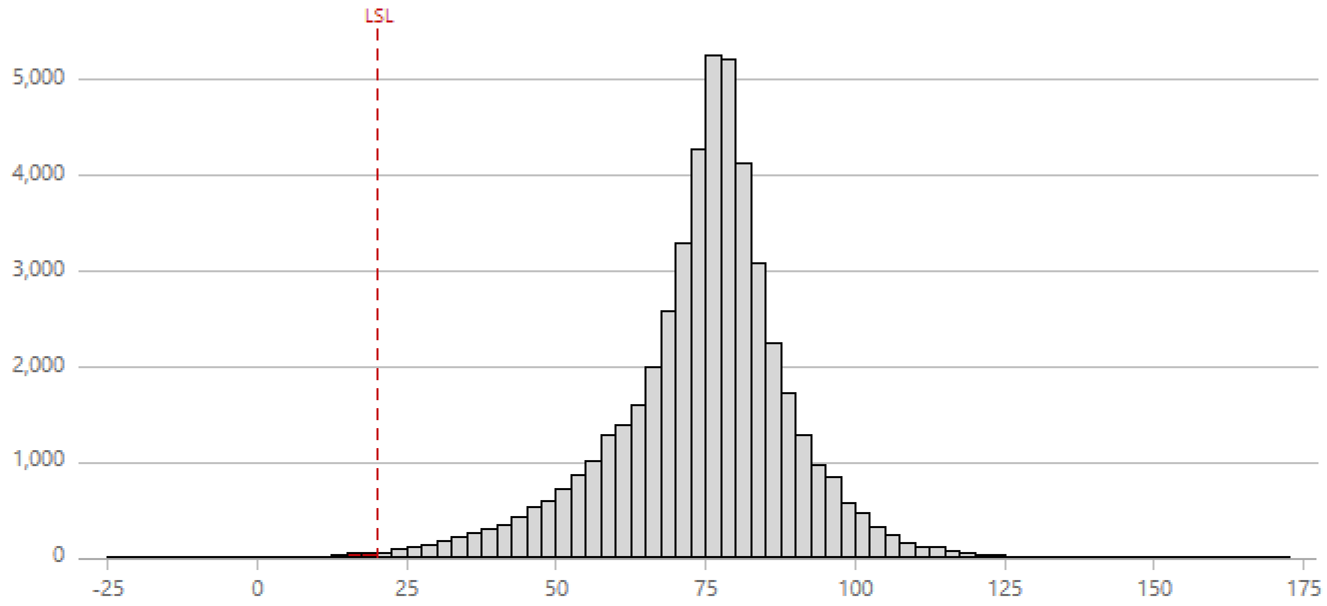
Use updated information to simulate process improvement:

- Use model equation, optimal settings, and Error are from DOE.
- Current process data provides distribution information for Cement and Water.

Prescriptive: Monte Carlo Simulation

Simulation Results

Strength



The simulation indicates that you can expect 0.54% of the *Strength* values to fall outside of the specification limits. This corresponds to a Cpk of 0.8256. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

0.8256

% Out of Spec

0.54%

Summary Statistics

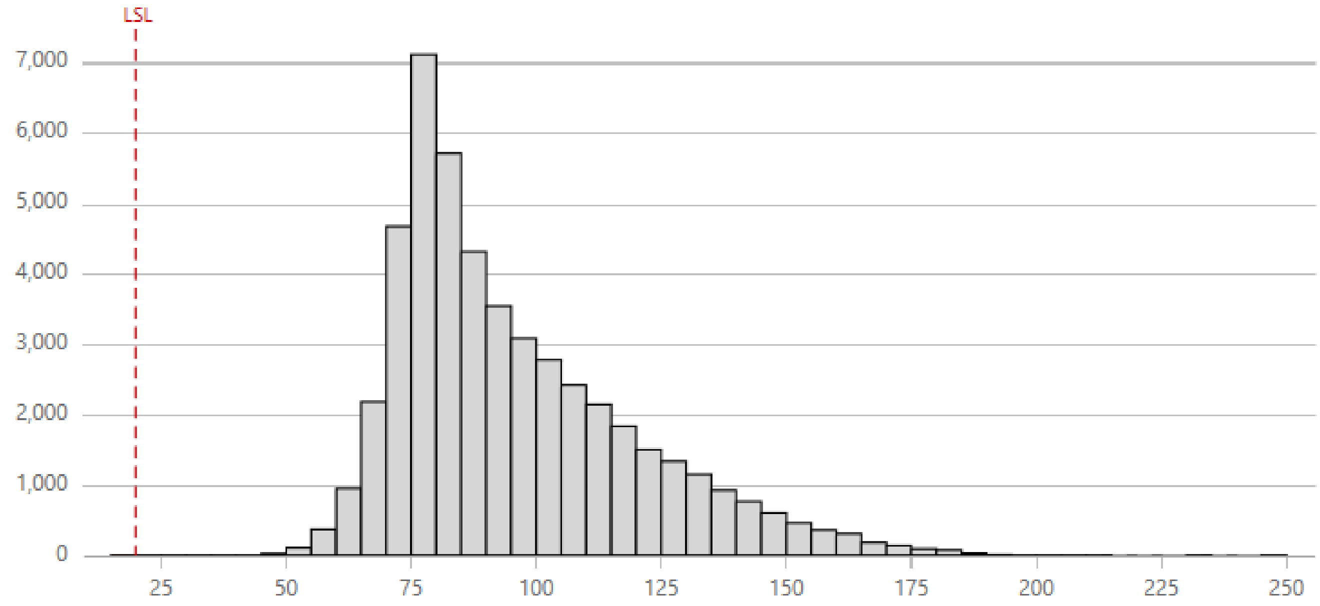
N	50,000
Mean	74.1529
Standard Deviation	15.751

- Defect rate has improved, but still too high.
- Use Parameter Optimization to further optimize settings.

Prescriptive: Monte Carlo Simulation

Parameter Optimization Results

Strength



Name	New Settings
Cement	(350.055; 450.055)
Water	(150.142; 20)
Error	(0; 3)

Process Performance (Cpk)

1.68

% Out of Spec

0.00%

Summary Statistics	
N	50,000
Mean	96.118
Standard Deviation	24.8265

To improve performance, maintain cement amount between 350 and 450 and water at 150.

Where Are We on the Analytics Journey?

How many years do you think it will take your overall industry to realize the potential of the Digital Transformation?

From Minitab LLC Research

- 493 Director and above level executives responded to the survey.
- *Region:* 45% of the respondents were from US
- *Industry*: Manufacturing and Healthcare made for 21% of the respondents

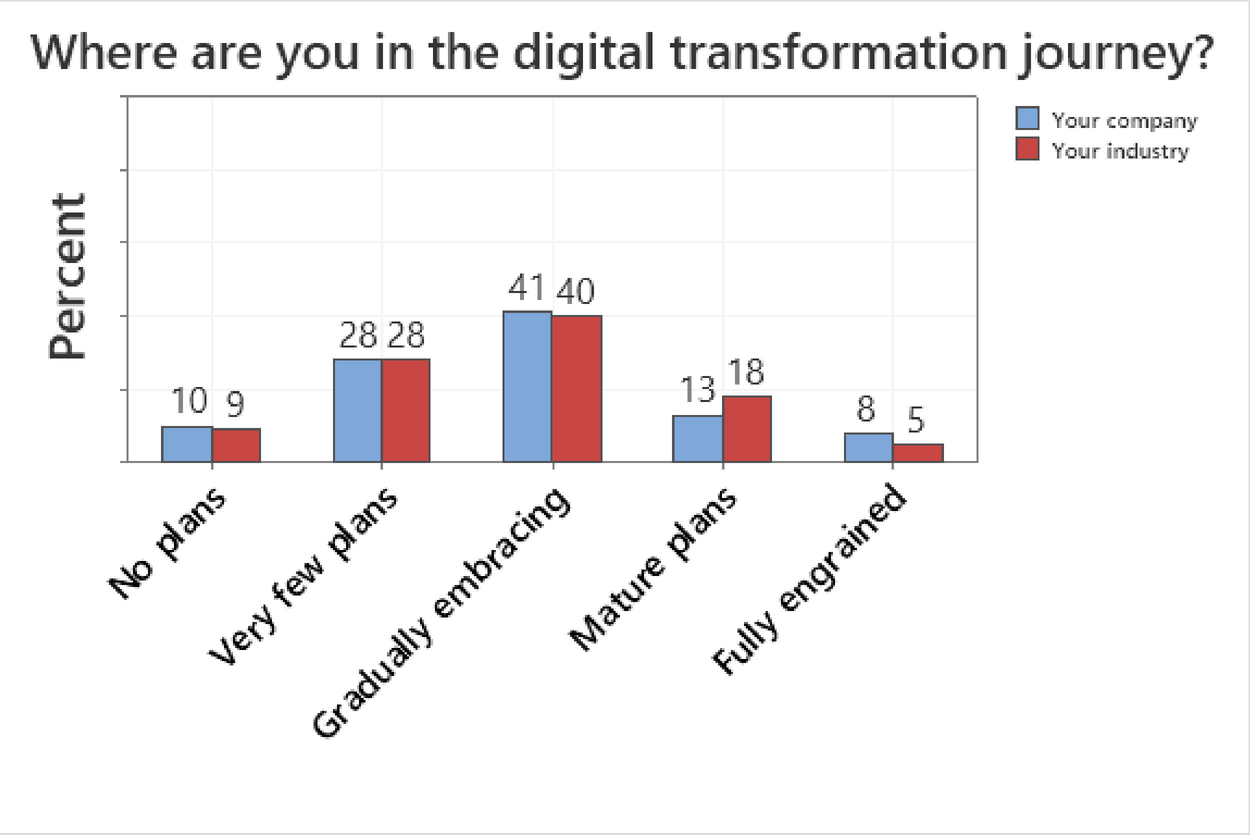
Evolution of the Digital Transformation

“I think we’re still on the precipice of technological transformation and I think there will be additional tools within 5-10 years that will be astounding and compelling.”

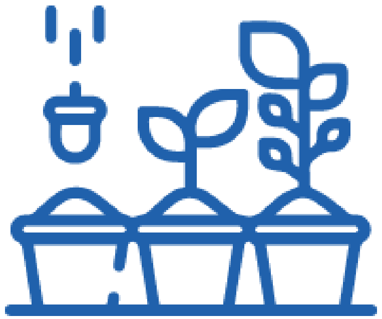
– CFO, Manufacturing

8

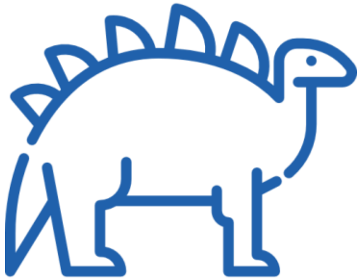
The average number of years Decision Makers think it will take for their overall industry to realize the potential of Digital Industrial Transformation and the IIoT.



Key Research Findings

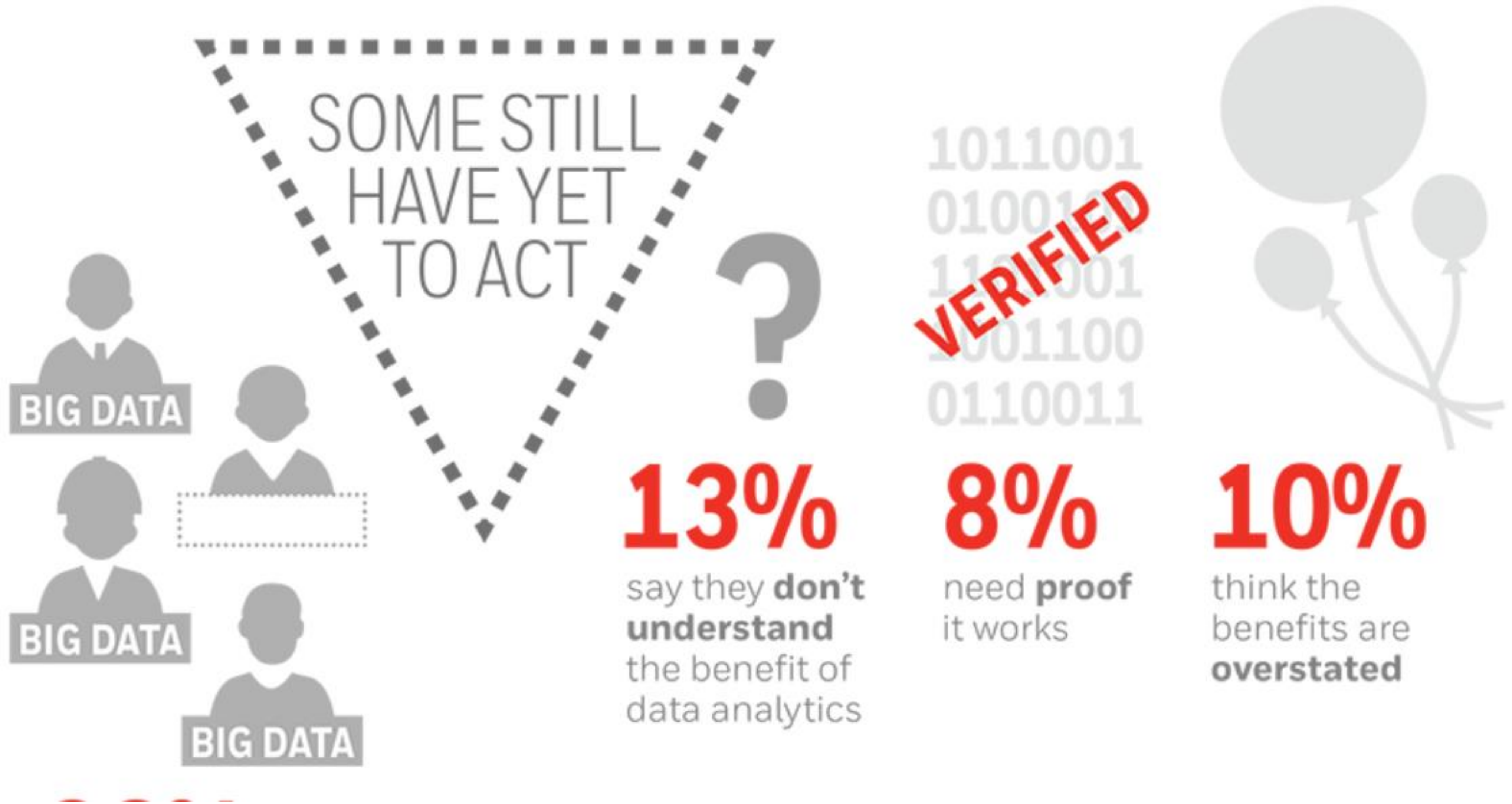


- **Early Days for Digital** – Digital transformation seen as "early days;" transformation gradual and incremental.



- **Obsolete Too Soon** – Leaders concerned that constantly evolving technology and its application may make today's solution obsolete before they see ROI.

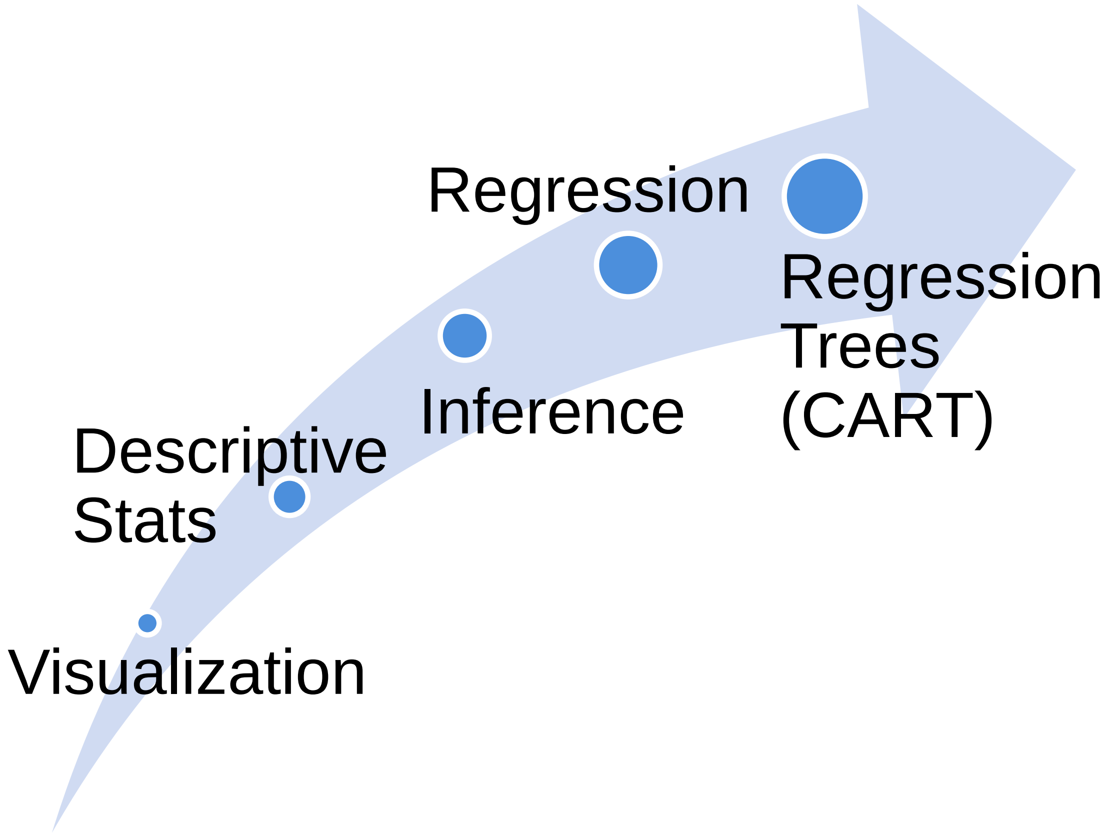
Key Research Findings



Begin With Creating a Culture of Analytics

Transformational organizations develop a **culture of analytics** throughout their organization.

- Analytics should not reside solely with Data Scientists.
- Goal should be to move everyone along analytics curve.
- Quality professionals are well-suited for leading the transformation.



Take-aways

- This truly is a journey; no one is there yet.
- Leverage your Quality groups and resources, in whom you have already invested heavily.
- Technology is important, but the teams you assemble to process, digest and ultimately act on information is what will set you apart from your competitors.
- Your process improvement experts, armed with easy-to-use analytics tools, will propel your initiative forward.

Cheryl Pammer
cpammer@minitab.com