## **QN** Immediate Fix Cycle Time Analysis

Raisa Huang

2

Root Cause Analysis of QN Fix Cycle time

Graphical Root Cause Analysis Summary

Compare Fit Model, Partition, Neural Model

Hybrid Text Mining & Data Mining Analysis

Take Away Learnings

#### <sup>3</sup> Histogram – 1<sup>st</sup> Layer of Root Cause Analysis of QN Fix Cycle time

What scenarios impact on QN fix cycle time? The impact is endurable?

SA

- Criteria: within 5 days (In spec, success analysis); over 5 days (out of spec ,failure analysis)
- Use Histogram Conditional Mosaic Plot to conduct both Success Analysis and Failure Analysis



Graph Builder Box Plot – 2<sup>nd</sup> Layer of Root Cause Analysis

- Plot continuous fix cycle time vs. nested structure(categorical country X<sub>3</sub> under containment X<sub>2</sub>)
  - » The cycle time of Replacement is much longer than other containment actions
  - » Containment should be one of important factors to impact on fix cycle time



#### Graph Builder Heatmap – 3<sup>rd</sup> Layer of Root Cause Analysis

- Add **back categorical** "defect type (X<sub>1</sub>)" on Y axis, color for fix cycle time
- Use 8 x 9 layout (balanced) to quickly catch out the max / min cycle time scenarios
  - Replacing TW Damage parts is the worst case for cycle time
  - **Replacing USA Dimension** issue parts is the 2<sup>nd</sup> worst scenario.



#### Pareto Chart – 4<sup>th</sup> Layer of Root Cause Analysis

- Add additional factor "**workstation** (X<sub>4</sub>)" in Pareto Chart to visualize frequency event
  - Replacing TW damage & CVD-SF
  - Replacing USA Dimension & CVD-Lid

» Per previous inference : Except for Dimension and Damage, other defects are easy to quickly fix





#### Tabulate – 5<sup>th</sup> Layer of Root Cause Analysis

- Show average and count on Tabulate table to do further comparison
- FA: CVD-SF Replacement TW damage issue

7

SA: CVD-MT MFG rework USA Workmanship issue

				immedia	nte f	ix cycle	time							
Work Station														
			CVD	-LID			CVD	)-SF						
		D	efec	t type			Defec	t type						
		Dimens	ion	Damag	je	Dimen	sion	Damag	je		immedia	ite fi	c cycle ti	me
Containment	Country	Mean	Ν	Mean	Ν	Mear	N	Mean	Ν		W	ork St	tation	
Replacement	TW	15	7	4	1	8	3	34	6		CVD-MT			
	USA	16	6	7	6		0	2	1		D	ofert	tune	
	All	15	13	7	7	8	3	29	7			1.1	type	
		1 1						1			Workman	iship	Functio	nal
						Co	Intai	nment	C	ountry	Mean	N	Mean	N
						M	FG re	work	U	5A	2	20	3	16
									T	V	2	8	1	2
		• • •		•					A	l .	2	28	3	18
								• • •				•		

#### Root Cause Analysis Summary

- Use different Graphical JMP Platforms in Engineering and Logical Sequence to conduct deeper Root Cause Analysis
  - 1<sup>st</sup> Layer Histogram: set Conditional Mosaic to investigate both SA and FA
  - 2<sup>nd</sup> Payer Box plot: know how to investigate the process special variations (skewness, outliers)
  - 3<sup>rd</sup> Layer Heatmap plot: narrow down the SA/FA root cause analysis scope to Defect Type X Country Square
  - 4<sup>th</sup> Layer Pareto Chart: conduct 2-dimensional Pareto Chart from previous Heatmap results
  - 5<sup>th</sup> Layer Tabulate: visualize the Pivot Table on integrating the previous layers of Root Cause Analysis
- Identify the Potential inputs (X<sub>s</sub>) to Predict the QN fix Cycle Time
  - 1<sup>st</sup> Layer Histogram: Defect type (X<sub>1</sub>)
  - 2<sup>nd</sup> Layer Box plot: Containment (X<sub>2</sub>), Country (X<sub>3</sub>)
  - **3<sup>rd</sup> Layer Heatmap**: Defect type (X<sub>1</sub>), Containment (X<sub>2</sub>), and Country (X<sub>3</sub>)
  - 4<sup>th</sup> Layer Pareto Chart: Defect type (X<sub>1</sub>), Containment (X<sub>2</sub>), and Country (X<sub>3</sub>), Workstation (X<sub>4</sub>)
  - 5<sup>th</sup> Layer Tabulate: Narrow Down to Damage (Defect type X<sub>1</sub>), Replacement (Containment X<sub>2</sub>), TW (Country, X<sub>3</sub>), CVD-SF (Workstation X<sub>3</sub>)
- Next Step: Build a model to predict the QN fix Cycle Time (Validation of Root Causes)

## Model Selection and Comparison

The fit model challenge:

- Skewed distribution: log transformation -> no help
- All input variables are categorical type (filter out 60% of workstation category, R–square increase by 6%)
- Dependency among categorical variables (low risk)





- Partition tree model:
  - » Distribution free model
  - » Split base on data available
  - » Little overfit concern
  - » Recursive split
  - » Random Forest Predictor Screening

⊿ 💌 Predictor Screening								
	in	Copy Selected						
Predictor	Contribution	Portion		Rank ^				
Defect type	1864.08	0.3384		1				
Work Station	1780.23	0.3232		2				
Country	1038.42	0.1885		3				
Containment	702.34	0.1275		4				
fault by (ref)	123.17	0.0224		5				

- Neural Network model:
  - » Strong transformation model
  - » Two steps (training & validation) model
  - » Significant overfit concern



Y	Log(immedicycle time 2)
	optional
Weight	optional numeric
Freq	optional numeric
By	optional

Summary of Fit

Root Mean Square Error

Observations (or Sum Wats)

Mean of Response

0.303036 0.247055

0.933638

1.268714

270

RSquare

RSquare Adj

#### -Construct Model Effects

Add	Defect type
Cross	Containment
Nest	Work Station
Macros 🕶	fault by (ref)

#### Use **Log transformation** of the cycle time variable to transform the skewed cycle time distribution

•	Remembered Settings									
					Work		immediate fix	immediate fix cycle	immediate fix cycle	
	Setting	Defect type	Containment	Country	Station	fault by (ref)	cycle time 2	time 2 Lower CI	time 2 Upper Cl	Desirability
)	_Optimal_	Wrong part	Supplier rework	KR	CCT-Staging	g Supplier fault	0.5518602	0.2182356	1.3955084	0.895239
	Setting	Defect type	Containment	Country	Station	fault by (ref)	cycle time 2	time 2 Lower Cl	time 2 Upper Cl	Desirability
)	Optimal	Dimension	Replacement	JP	CVD-MT	MFG fault	30.040394	9.7730486	92.338152	0.695992

#### (2) Partition Tree Model – Model Improvement & Comparison

#### Baseline Model

11

- Model is not adequate (R square 37.9%)
  - Originally 4 input factors

#### Model Augmentation

- Discuss with SME and select a new output variable (Y)
- Add 5<sup>th</sup> input variable (Workstation ,X)
- R square has been improved around 20%

#### Model Simplification

- Utilize Pareto principle and data filter to screen any minor data category
- Total sample size decrease to 270 from 426
- R square improve around 6%

Select Columns	Cast Selected Columns into Roles	Action —
28 Columns	Y. Response A QN age (day)	OK
II. fault by	optional	Cancel
L Defect type		Cancer
QN age (day)		
L Country	X Faster II Defect type	Remove
🔥 UD code	A, Tactor	
🔥 Work Center		Recall
IL WK	ob code	Help
IL ON type	Country	
17.21902	loptional	

			Number	
RSquare	RASE	N	of Splits	AICc
0.379	27.269101	426	37	4111.52





			Number	
RSquare	RASE	N	of Splits	AICc
0.623	4.5468444	270	35	1670.14

## Model Augmentation (R-square improved by 20%)

## • 0% R-square Improvement

**Column Contributions** 

Term

Country

UD code

Defect type

fault by (ref)

RSquare

0.368

Number

of Splits

12

Column Contributions						🖌 QN age (day)		
Term	Number of Splits		ss				Port	ion
Defect type	17	114415.	148				0.5	911
Country	14	59520.9	863				0.3	075
UD code	5	19591.3	368				0.1	012
fault by (ref)	1	50.4751	834				0.0	003
	1				Number			
RSquare	e	RASE		Ν	of Splits	A	ICc	
0.37	27.2	69101	4	426	37	411	1.52	

SS

15 2430.52733

12 2407.45243

8 1829.65564

5 168.540489

RASE

5.248367

🚄 immediat... cycle time

Number

N of Splits

426

Portion

0.3555

0.3522

0.2676

0.0247

AICc

40 2714.91

- 16% R-square Improvement
- Add X factor: MFG Workstation (the NO.2 ranking, around 28%)
- UD code less critical after adding workstation

Y, Response a immediat... cycle time Column Contributions

17 4546.69243

15 2794.18279

14 1584.96213

8 791.891966

2 38.4684764

RASE

4.548719

SS

Number

of Splits

Term

Country

UD code

Defect type

Work Station

fault by (ref)

RSquare

0.525

•	Another	<b>4%</b>	R-Sq	uare
	improve			

Change X factor: Containment from UD code

Builds a decision tree to predict a response.	
Select Columns	Cast Selected Columns into Roles Action
36 Columns	Y. Response 🖌 immediat cycle time 🛛 OK
Enter column name P 🔻	optional
fault by (ref) Defect type Containment Country immediate fix cycle time BU Work Station	X, Factor II, fault by (ref) L Defect type Containment L Containment Help L Work Station

#### **Column Contributions**

Portion

0.4660

0.2864

0.1625

0.0812

0.0039

AICc

56 2634.23

Term	Number of Splits	SS		Portion
Defect type	16	4612.40123		0.4390
Work Station	17	4260.34495		0.4055
Containment	6	894.132282		0.0851
Country	10	603.876596		0.0575
fault by (ref)	3	134.900637		0.0128
			Number	

			Number	
RSquare	RASE	N	of Splits	AICc
0.566	4.3510406	426	52	2585.74

Top two input (X) factors in rankings: Defect type & Country → Defect type & Workstation

426

Number

N of Splits

## Model Simplification (R-square improved by 6%)

0.4055

0.0851

0.0575

0.0128

⊿ 💌

- Previous model augment includes all categories & data
  - Plus: considering all scenarios

13

Work Station

Containment

Country fault by (ref)

Drawback: too many categories might dilute prediction power



17 4260.34495

6 894.132282

10 603.876596

3 134.900637

- Simplify dataset by filtering out minor categories with fewer counts to improve prediction power
  - **Remove 60% categories of Workstation**

	1						
Data Filter	Ide				Work	immedia	ate fix
Clear Favorites 🕶	fault by (ref)	Defect type	Containment	Country	Station	cycle t	ime
Select of Show of Include	Supplier fault	Functional	Replacement	USA	CVD-MT		47
270 matching rows	AMAT fault	Damage	MFG rework	TW	CVD-SF	[	
	MFG fault	Workmanship	Supplier rework	KR	CVD-LID		
fault by (ref) (3)		Dimension		DE	CCT-Staging		
AMAT fault 41		Cosmetic		JP	PVD-M	h .	
MFG fault 11		Labeling		<b>F</b>	CVD-PK		
Supplier fault 374		Missing part					
Defect type (11) ×		Wrong part					0
Damage 67			-				-
Dimension 63							
Functional 60			Numb	er			
Cosmetic 54	PCourse	DACE	M of Sol	ite A	IC c		
Workmanship 50	nsquare	INASE.			icc		
Missing part 40	0.623	4.5468444	270	35 1670	.14		
Wrong part 36							
Labeling 35	Column Co	ntribution	s				
DOA 7 ~		Number					
Containment (4)	T	Number				Denting	
Poplacement 224	Term	or Splits				Portion	
MEG rework 120	Defect type	9 45	12.26539			0.4900	
Supplier rework 49	Work Station	14 28	303.04562			0.3044	
use as is 14	Containment	3 11	50.87421			0.1250	
	fault by (ref)	3 38	34.822466			0.0418	
	Country	6 34	8.534565			0.0389	
	country	0 0.			1	0.0000	

Total amount (N) decreases to 270 from 426 (156

#### Partition Tree Model Optimization – Min & Max QN Cycle Time

- The major contributor are Defect type & Workstation ~ 80% (Pareto Concept)
- According to prediction profiler of the method,
  - The best scenario (min cycle time) :Labeling, PVD-M
  - The worst scenario (max cycle time) :Damage, CVD-MT

Column Contributions										
Term	Number of Splits	SS		Portion						
Defect type	9	4512.26539		0.4900						
Work Station	14	2803.04562		0.3044						
Containment	3	1150.87421		0.1250						
fault by (ref)	3	384.822466		0.0418						
Country	6	358.534565		0.0389						



Doesn't country impact QN fix cycle time? Is it right?

#### 15 Model Limitations: Recursive Partitions Column Contributions Number Recursive partitions (sequential dependency risk) of Splits SS Portion Term 9 4512.26539 0.4900 Defect type Factor "country" is split 6 times, and only 1 time 0.3044 Work Station 14 2803.04562 Containment 3 1150.8742 0.1250 happened in the higher cycle time cluster. fault by (ref) 3 384.822466 0.0418 6 358.534565 0.0389 Country All Rows Such recursive dependency limitation may impact 270 LogWorth Difference Count 5.37063 6.2888889 8.0095095 Mean the predictive model Defect type(Workmanship, Defect type(Dimension, Damage) Std Dev 7.4153129 Cosmetic, Functional, Wrong part, 88 LogWorth Difference Count Labeling, Missing part) 9.9090909 6.82274 Mean 2.171908 182 LogWorth Difference Count Std Dev 10.50958 Mean 4.5384615 4.1898365 5.13333 Work Station(CCT-Staging, CVD-Std Dev 4.3792774 LID, CVD-MT, CVD-SF, PVD-M) Defect type(Dimension) Count 170 LogWorth Difference Count 20 LogWorth Difference Mean 4.2 5.4641383 2.86316 Mean 12.8 1.0241808 5.36264 Containment(Supplier rework, Std Dev 6.8333547 Std Dev 4.1031089 MFG rework) 95 LogWorth Difference Count Containment(Replacement) 6 Mean 2.9368421 1.5480432 1.48192 Country(KR, JP, TW) Country(USA, DE) Count 75 LoaWorth Difference Count 13 LogWorth Difference Count Country(JP, TW) Std Dev 3.1515296 Mean 5.8 2.3076389 3 Mean 10.923077 1.011907 7.35 Mean 16.285714 33 LogWorth Difference Count Std Dev 4.6031716 Std Dev 7.7616744 Std Dev 2.3603874 Mean 1.969697 2.2843331 1.44615 Candidates Std Dev 1.5100672 Work Station(CCT-Staging, CVD-Work Station(PVD-M, CVD-MT) LID, CVD-SF) Count 45 LogWorth Difference 30 LogWorth Difference 7 2.2116705 5.41353 Mean 4 1.2289428 2.91667 Std Dev 4.9313469 Std Day 3 4038518 Country(DE, USA, KR) Defect type(Labeling, Functional, Defect type(Missing) Defect type(Workmanship, Missing Cosmetic, Wrong part) Workmanship) Count 62 LogWorth Difference 38 LogWorth Difference part, Cosmetic, Labeling) Count Mean 3.4516129 2.1365378 2.53066 24 LogWorth Difference Mean 6.1578947 1.3650477 3.96364 Mean 11.571429 Mean 4.5833333 1.3820103 2.92308 Std Dev 3.651532 Std Dev 4.1167182 Std Dev 6.7046537 Std Dev 3.549852 Candidates 3 Country(TW, USA) Country(KR. DE) Work Station(CVD-SF, CCT) Country(JP, TW, DE, KR) Country(USA) Count 33 LogWorth Difference Count Staging, CVD-LID, PVD-M) 11 LogWorth Difference Count 13 LogWorth Difference Mean 5.6363636 0.3384261 9.6 Mean Mean 5.9230769 1.1149059 Count 23 LogWorth Difference Mean 3 0.6346458 1.46667 Std Dev 3.7149881 Std Dev 5.4129474 Std Dev 1.9493589 Std Dev 4.0918932 Mean 5.0434783 1.1627271 3.2197 Candidates Std Dev 4.2584489 Work Station(CVD-MT) Work Station(PVD-M Work Station(CVD-SE\_CVD-LID) Work Station(CCT-Staging) Work Station(CCT-Staging) Work Station(CVD-SF, CVD-LID) 25 LogWorth Difference Count Count Count Count Count Mean 5.36 0.1461622 0.62281 Mean 6.5 Country(DE, KR) Country(USA) Mean 3.6666667 Mean 3.4 7.5 Std Dev 3.5223098 Std Dev 4.4077853 dean. Std Dev 1.0954451 Std Dev 2.3380904 Std Dev 3.3615473 Std Dev 3.8544964 Count 11 LogWorth Difference Count 12 LoaWorth 5 Candidates Mean 3.3636364 1.1063871 4.46667 Mean 6.5833333 1.271219 Candidates Candidates Candidates Candidates Country(USA) Country(TW) Std Dev 4.2254639 Std Dev 3.8247598 Count 19 Count Mean 5.2105263 Mean 5.8333333 Std Dev 3 4893321 Std Dev 3.920034 Candidates Candidates

## Neural Network (Artificial Intelligence)

- Observe severe overfit concern between training and validation R square
  - Overfit: > 20% R-Sq between Training Set (building model), and Validation Set (fitting model)
  - Too aggressive Black-Box transformation to build a model with training set (too good to be true)
- The major contributor : Workstation

Training		⊿ Validation							
immediate f	fix cycle time	⊿ immediate f	fix cycle time						
Measures	Value	Measures	Value						
RSquare	0.659755	RSquare	0.4449701						
RASE	4.5389141	RASE	4.7913103						
Mean Abs Dev	3.0979374	Mean Abs Dev	3.5440675						
-LogLikelihood	527.69274	-LogLikelihood	268.71682						
SSE	3708.3134	SSE	2066.0989						
Sum Freq	180	Sum Freq	90						

#### Variable Importance: Independent Uniform Inputs

Summary Report												
Column	Main Effect	Total Effect	.2	.4	.6	.8						
Work Station	0.081	0.566										
fault by (ref)	0.05	0.544										
Country	0.043	0.352										
Defect type	0.108	0.336										
Containment	0.104	0.289										



#### Model Comparisons and Selection



- Root Cause Analysis: Damage issue (defect type), Replacement (containment), TW (country), CVD-SF (workstation) is the worst scenario with longer QN fix cycle time
  - Neural Model has the identical scenario as the graphical root cause analysis
  - Only concern on the Overfit risk
- The 3 models have very close prediction on the worst cycle time within 1.2 Days

		Summary of Fit					Remembered Settings										
Fit Mode		RSquare RSquare Adj	Guara Error	0.30303	6 5		0	Setting _Optimal_	Defect type Wrong part	Containmer Supplier rew	nt Country	Work Station CCT-Staging	<b>fault by (ref)</b> Supplier fault	immediate fix cycle time 2 0.5518602	immediate fix cy time 2 Lowe 0.2182	r Cl immediate fix cycle r Cl time 2 Upper Cl 356 1.3955084	Desirability 0.895239
		Mean of Res Observation	sponse s (or Sum Wgts)	1.26871	4 10		0	Setting Optimal	Defect type Dimension	Containm Replaceme	ent Country ent JP	Station CVD-MT	<b>fault by (ref)</b> MFG fault	<b>cycle time 2</b> 30.040394	time 2 Lower 9.77304	CI time 2 Upper CI 86 92.338152	Desirability 0.695992
								Setting	Defect type	Work Station	Containment	Country	fault by (ref)	mmediate fix	Desirability		
Partition		RSquare	RASE	N	of Spli	er ts AICo	•	_Optimal_	Labeling	PVD-M	MFG rework	JP	Supplier fault	0.8	0.857144		
		0.623	4.5468444	270	3	35 1670.14	<u>+</u>	Setting	Defect type	Station	Containment	Country	fault by (ref)	cycle time	Desirability		
	Train	ing		⊿ Valida	ation			Rememb	pered Settin	as	Replacement	KI	WI C IBUIL	20.0	0.201314		
	⊿imr	<mark>nediate f</mark> iz	x cycle time	⊿ imn	nediate f	ix cycle tim	e					Work		immediate fix			
	Mea	isures	Value	Meas	ures	Value		Setting	Defect type	Containme	ent Country	Station	fault by (ref)	cycle time	Desirability		
Neural	RASI	Lare E 4	4.5389141	RASE	are	4.7913103		_opumai_	workmanship	Supplier rev	WORK KR	Work	g supplier lault	immediate fix	0.999797		
	Mea	n Abs Dev	3.0979374	Mean	Abs Dev	3.5440675		Setting	Defect type	Containm	nent Country	/ Station	fault by (ref)	cycle time	Desirability		
	-Log SSE	Likelihood	527.69274 3708.3134	-Logi SSE	.ikelihood	268.71682 2066.0989	0	_Optimal_	Damage	Replacem	ient TW	CVD-SF	Supplier fault	29.019381	0.666909		
	Sum	Freq	180	Sum I	Freq	90				• • •					•		

## Text Mining and Data Mining Hybrid

- Search keywords from QN Database (Categorical and Text Variables)
- Convert the Keywords information to Binary Indicators
- Conduct the further Data Mining- Root Cause Analysis On F10246 Case

Text Explorer for ON long text		Sum(replac· Ind	icator) & Sum(rework· I	Indicator) vs. dimens· Indicator	
		0	<mark>F10246 i</mark>	indicator 1	immediate fix cycle time
Number of Terms of CasesNumber Tokens Tokens per CaseNumber of Non- Empty CasesPortion of Non- Empty Cases737108488045.18521081.0000CeplacsupplierSupplierissureworklabelmfgprovidpart	replac-Indicator	dimens- Indicator: 0 Mean(immediate fix cycle time): 4.33 Sum(replac- Indicator): 43	dimens- Indicator: 1 Mean(immediate fix cycle time): 11.39 Sum(replac- Indicator): 13	dimens- Indicator: 0 Mean(immediate fix cycle time): 22.22 Sum(replac- Indicator): 8 Gum(replac- Indicator): 8	225 - 18.1 13.8 immediate fix cycle time 9 9 5.0
damag cabl festo screw sf hole inform qti mic- miss tc wrong materi hc plan qn susceptor 11 attach- incorrect instal power function one open box pcs refer scratch 4 ac check draw help make pin workmanship actual correct lid m16 process receiv recoveri y 0246 2 amat locat posit see side spare surfac version 0191 3 23 connect find heater result swap termin turn	k- Indicator	dimens- Indicator: 0 Mean(immediate fix cycle time): 4.33 Sum(rework- Indicator): 22	dimens Indicator: 1 Mean(imm fix cycle time): 11.39 Sum(rework Indicator): 4	dimens- Indicator: 0 Mean(immediate fix cycle time): 22.22 Sum(rework- Indicator): 10	
Over 5 days fix cycle time have strong relationship with "Replace, rework, dimension, F10246"	rewor				

## Take Away Learnings

- JMP Graphical Platforms are powerful to conduct deeper root cause analysis through Engineering, Logical, Data-Driven process
- Compare and Select more appropriate JMP Model from Classical Fit Model to modern Partitions and Neural Network by knowing the model limitations and risks connecting to previous Graphical Root Cause Analysis
- Conduct the Hybrid Text Mining and Data Mining Root Cause Analysis on the Complicated QN Database

• THANK GCI MBB Charles Chen as my Project Mentor

# Thank You