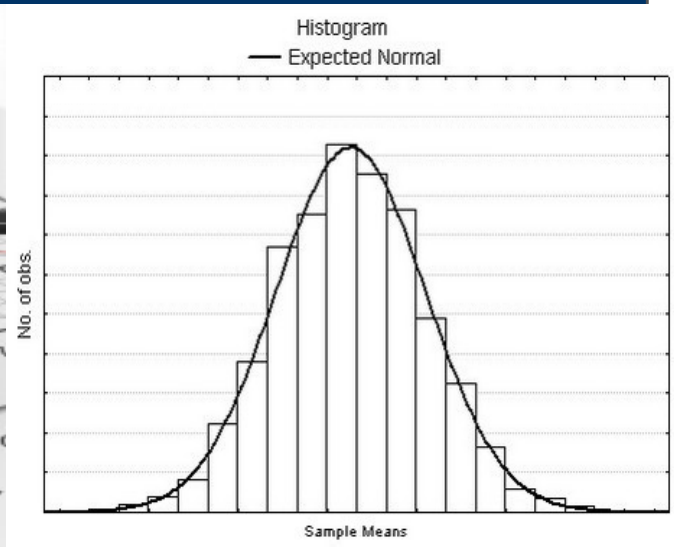
FUZZY ANALYTICS

AN ARTIFICIAL INTELLIGENCE APPROACH TO DETERMINISTIC / STOCHASTIC FEDERATED VIEW OF THE HUGE DATA SYSTEMS

STEP BY STEP APPROACH TOWARDS SUPERIOR LOGICAL DATA QUALITY AND DATA MANAGEMENT



Agenda of the Session

- The Pragmatic Fuzzy Data Analytics Approach
- Why are we here ?
- Where are we now ?
- Fuzzy Analytics Heuristic Model Analysis Approach
- Sample Examples
- Question and Answers

Breaking the Ice : The Pragmatic Fuzzy Big Data Analytics Approach

Key Considerations in Fuzzy Data Analysis

- Components of a Fuzzy Big Data Analysis Plan
- Distinguishing Fuzzy Big Data Types
- Distinguishing Different Types of Analyses

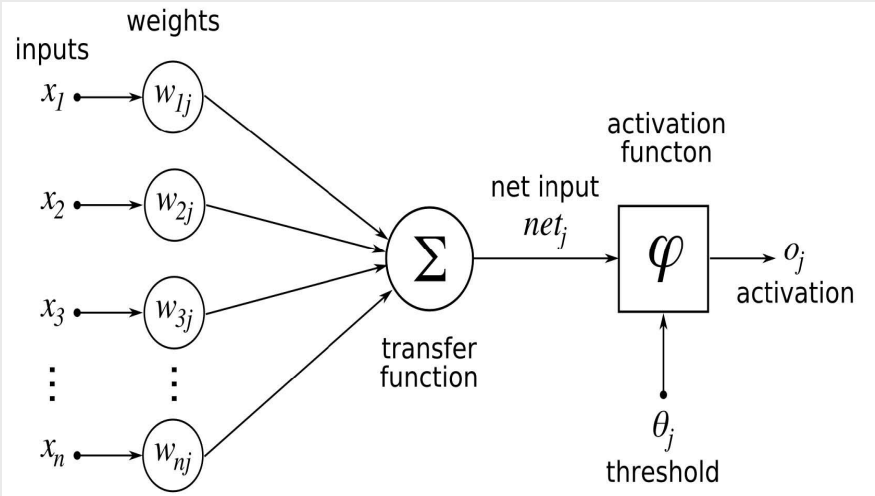
Overview of Different Statistical Tools for Fuzzy Big Data Case Studies:

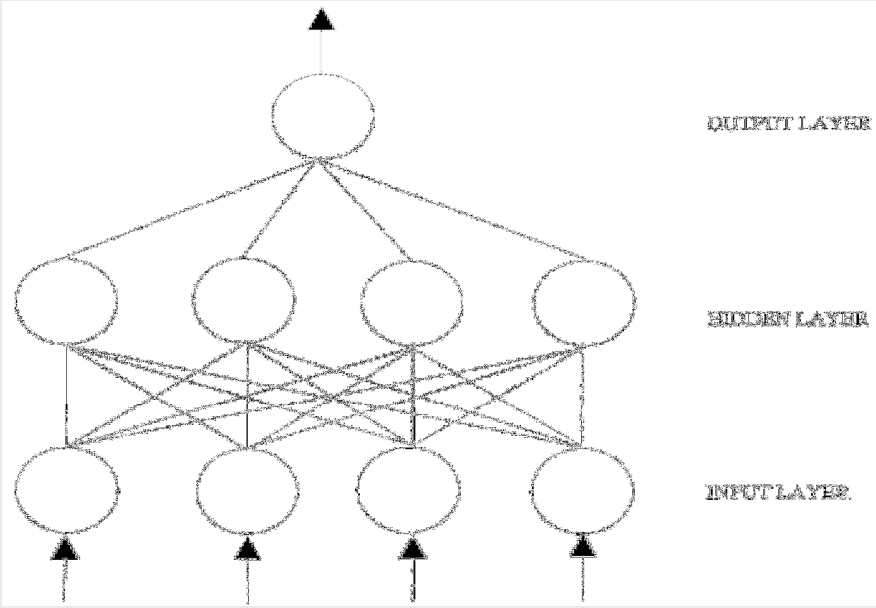
- Stock Data Modeling with Online Behavior Approach
- Sales Statistical Data Model Analysis
- Manufacturing Models Big Data Fuzzy Analytics
- Big Data Factor Analysis a key to Modern AI approach for informed decision making

Where does Fuzzy Analytics fit for Financial Market Researchers ?

- Sum up the experience of seasoned investor
- Indicators for different phases of business life cycle.
  - Recession →consolidation/ fiscal recovery → growth → fiscal decline
- Efficient market hypothesis
- Different methods of forecast eg. GARCH, ARCH, ARIMA, Neural Networks.

Flow Diagram and basic model of Fuzzy Analytics Neural Network





- Data Collection
- Data Preprocessing
- Extract Test Data Set
- Select Network Architecture
- Training
- Forecasting
- Result Analysis

Fuzzy Analytics Applied (for Finance Gurus)

- **Manage Risk** eg. Currency market average daily turnover is $ X trillion as reported in April 2012.
- **Building up portfolio eg.** Hedge funds, mutual funds, fund managers use intelligent system to build up portfolio from different asset classes.
- **Forecast future returns.**
- **Analyze "risk-reward" ratio.**
- **Trend analysis and pattern recognition.**
- **Trading strategies and economic indicators** eg. Projecting Inflation and GDP figures.

What all things Intelligent Systems still can’t capture?

- Market sentiments eg. Natural calamities , Festive Situations etc.
- Emotional attachment to an investment.
- Market reaction to scams and scandals

Why are we here ? - Key Considerations in Fuzzy Analytics of Big Data

- Identify the purpose of Fuzzy Analytics of Big Data
- Understand the sample(s) under Fuzzy Analytics of Big Data
- Understand the instruments being used to collect data for Fuzzy Analytics of Big Data
- Be cognizant of data layouts and formats under Fuzzy Analytics
- Establish a unique identifier if matching or merging is necessary under Fuzzy Analytics
- Plan your work and work your Fuzzy Analytics plan!

Industrial Fuzzy Analytics Modeling Approach

- Big Data Frequency Distributions and Cross -Tabulations :
- Big Data Descriptive Statistics (Means, Std. Deviations, Correlations) :
- Big Data T-tests and Analysis of Variance (ANOVA)
- Big Data Regression Analysis
- Big Data Principal Components/Factor Analysis (Data Reduction)
- Big Data Cluster and Discriminant Analyses (Segmentation)
- Big Data Latent Class Analysis (Classification)
- Big Data Hierarchical Linear Modeling (HLM)
- Big Data Differential Item Functioning (DIF)

Fuzzy Analytics Big Data Correlation Matrix

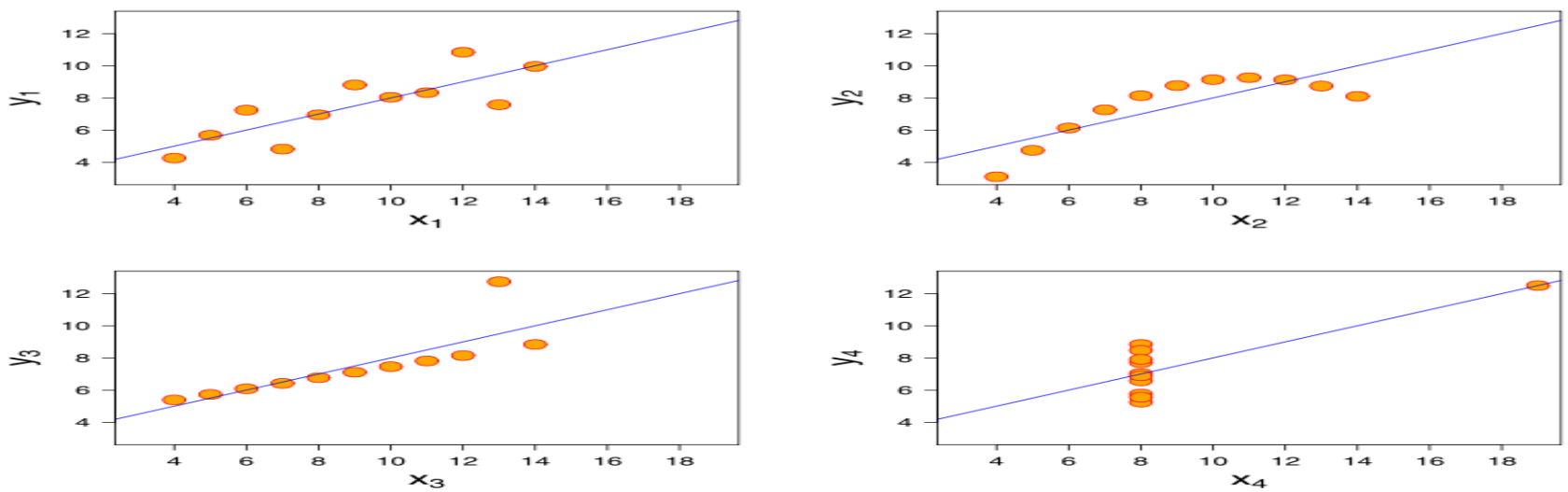
$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}},$$

where $\bar{x}$ and $\bar{y}$ are the sample means of X and Y, and s_x and s_y are the sample standard deviations of X and Y.

This can also be written as:

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}.$$

If x and y are results of measurements that contain measurement error, the realistic limits on the correlation coefficient are not −1 to +1 but a smaller range.[6]



Fuzzy Analytics Big Data Regression Analysis

	Stock	x_i	y_i	$(x_i - \bar{x})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
	1	95	85	17	8	289	64	136
	2	85	95	7	18	49	324	126
	3	80	70	2	-7	4	49	-14
	4	70	65	-8	-12	64	144	96
	5	60	70	-18	-7	324	49	126
Sum		390	385			730	630	470
Mean		78	77					

The regression equation is a linear equation of the form: $\hat{y} = b_0 + b_1 x$. To conduct a regressior analysis, we need to solve for b_0 and b_1. Computations are shown below.

$b_1 = \Sigma [(x_i - \bar{x})(y_i - \bar{y})] / \Sigma [(x_i - \bar{x})^2]$ $b_1 = 470/730 = 0.644$	$b_0 = \bar{y} - b_1 * \bar{x}$ $b_0 = 77 - (0.644)(78) = 26.768$

Fuzzy Analytics Big Data ANOVA (Analysis of Variance)

Measuring Variation Between Groups

The variation between group means is measured with a weighted sum of squared differences between the sample means and the overall mean of all the data. Each squared difference is multiplied by the appropriate group sample size, ni, in this sum. This quantity is called **sum of squares between groups** or **SS Groups**.

$$SSGroups = n_1(x_1 - x)^2 + n_2(x_2 - x)^2 + \ldots + n_k(x_k - x)^2 = \sum_{groups} n_j(\bar{x}_j - \bar{x})^2$$

The numerator of the F-statistic for comparing means is called the **mean square between groups** or **MS Groups**, and it is calculated as -

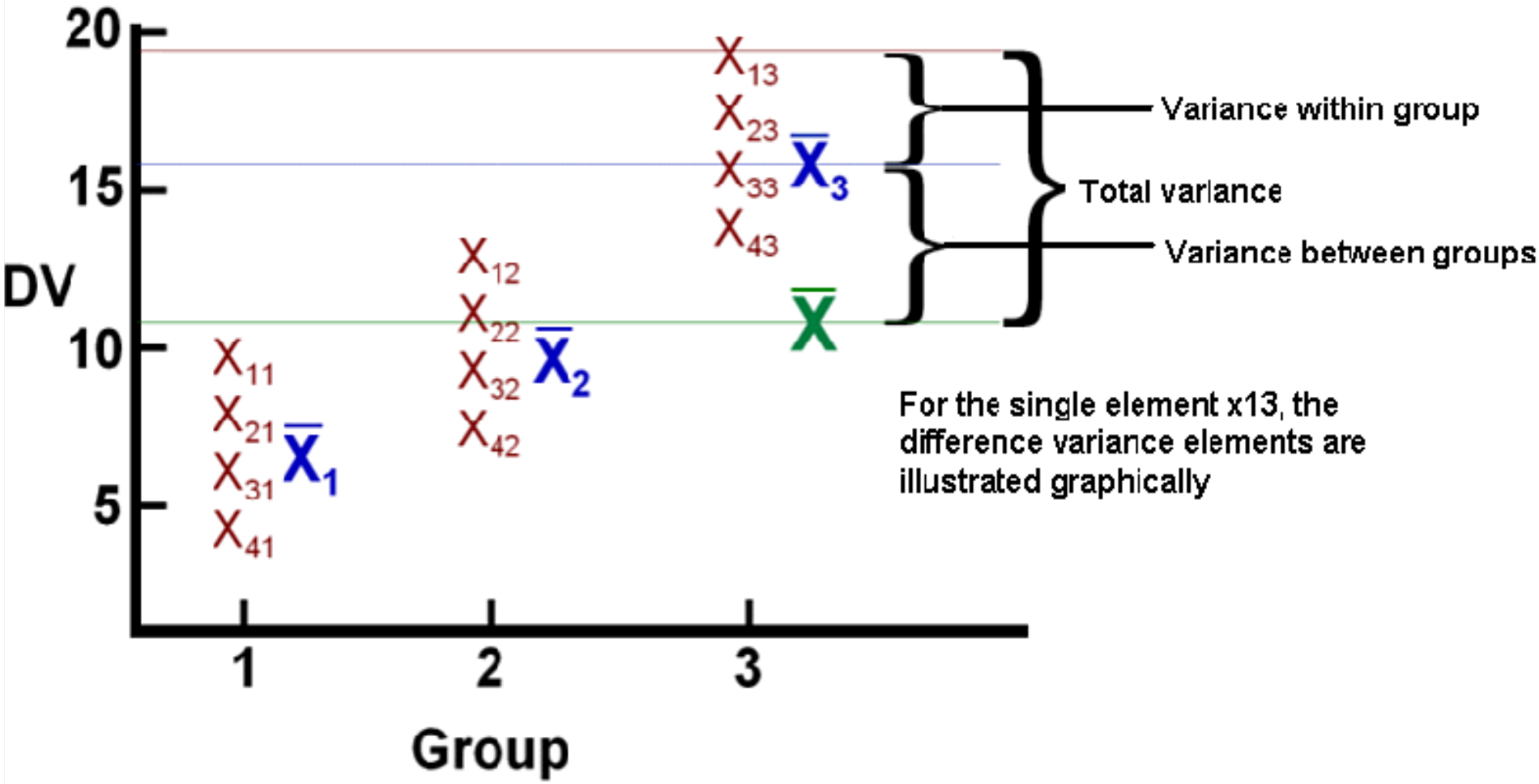
$$MSGroups = \frac{SSGroups}{k-1}$$

Measuring Variation Within Groups

To measure the variation among data points within the groups, find the sum of squared deviations between data values and the sample mean in each group, and then add these quantities. This is called the **sum of squared errors**, **SSE**, or **sum of squares within groups**.

$$SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2 = \sum_{allgroups} (n_j - 1)s_j^2$$

ANOVA Analysis : Basis for Factor analysis



ANOVA Analysis : Basis for Factor analysis

$$F_{\alpha, n_1 - 1, n_2 - 1}$$

$$F_{\alpha, n_1 - 1, n_2 - 1}$$

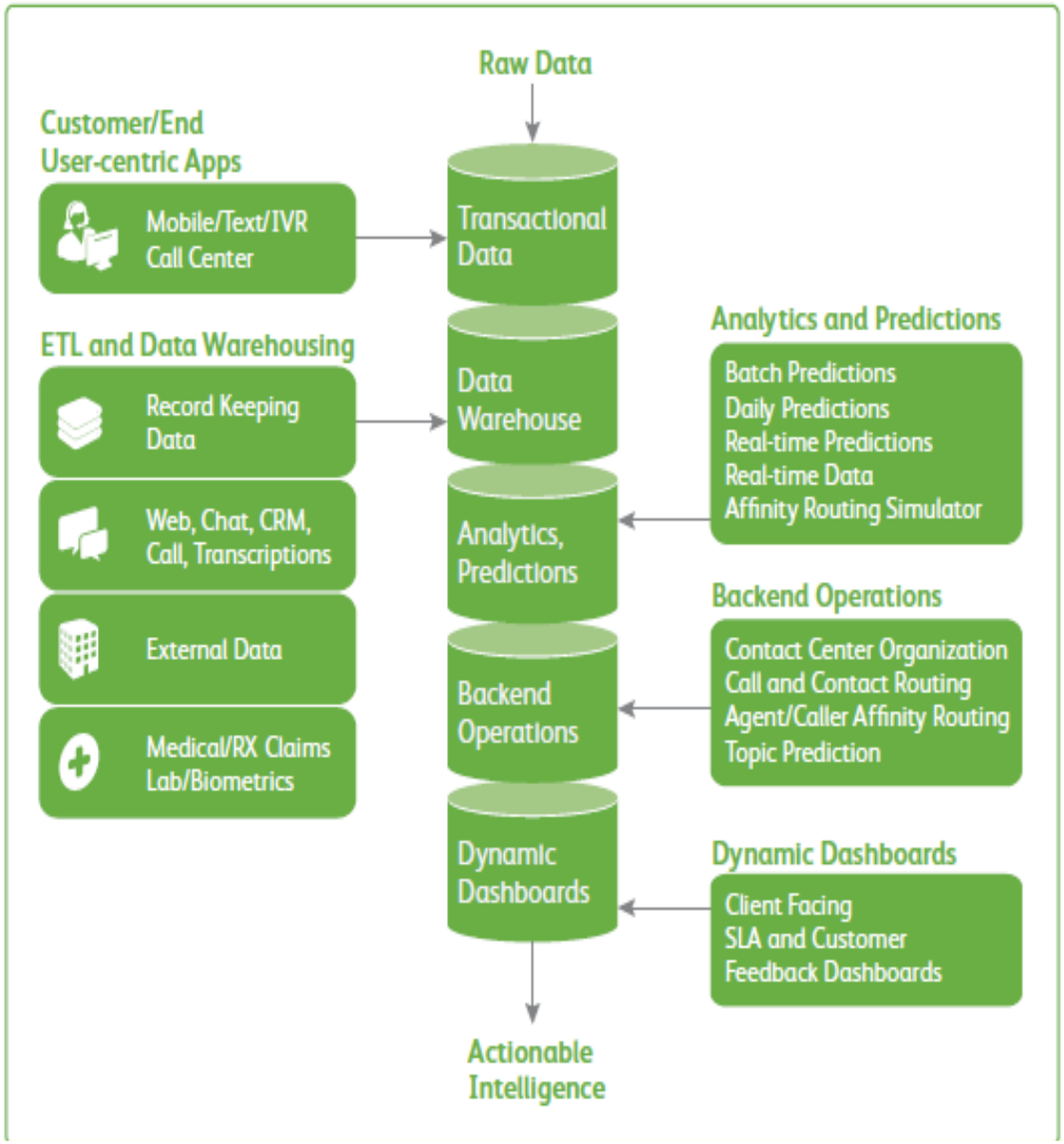
Critical Values for F Distributions

v2	α	v1: 1	2	3	4	5	6	7	8	9	10	15	20
1	0.1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	61.22	61.74
1	0.05	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	245.95	248.02
1	0.01	4052.2	4999.3	5403.5	5624.3	5764.0	5859.0	5928.3	5981.0	6022.4	6055.9	6157.0	6208.7
1	0.001	405311.6	499725.3	540256.5	562667.8	576496.1	586032.9	593185.4	597953.8	602245.3	605583.2	616073.6	620842.0
2	0.1	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44
2	0.05	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.43	19.45
2	0.01	98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.43	99.45
2	0.001	998.38	998.84	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31	999.31
3	0.1	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18
3	0.05	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66
3	0.01	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	26.87	26.69
3	0.001	167.06	148.49	141.10	137.08	134.58	132.83	131.61	130.62	129.86	129.22	127.36	126.43
4	0.1	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84
4	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80
4	0.01	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.20	14.02
4	0.001	74.13	61.25	56.17	53.43	51.72	50.52	49.65	49.00	48.47	48.05	46.76	46.10

More Advanced the Analysis, Greater the Amount of Preparation

- Most analyses can be executed straight from a working data file
- Some analyses may require transformations of the raw data, subsets, or specific input data to comply with statistical software

Big Data Analytics : A Practical Research Example : *Case Study 1*



Big Data Analytics in action:

- Identifies future at-risk borrowers so that proactive measures can be taken to prevent delinquency and default
- Supports investigations into government regulatory filings fraud
- Defines custom collection strategies based on risk score
- Prioritizes borrowers in the inbound ACD queue based on risk score
- Engages high-risk borrowers in multi-channel strategy to prevent delinquency
- Call Center Capacity Analytics

Big Data Analytics : A Practical Research Example : *Case Study 1*

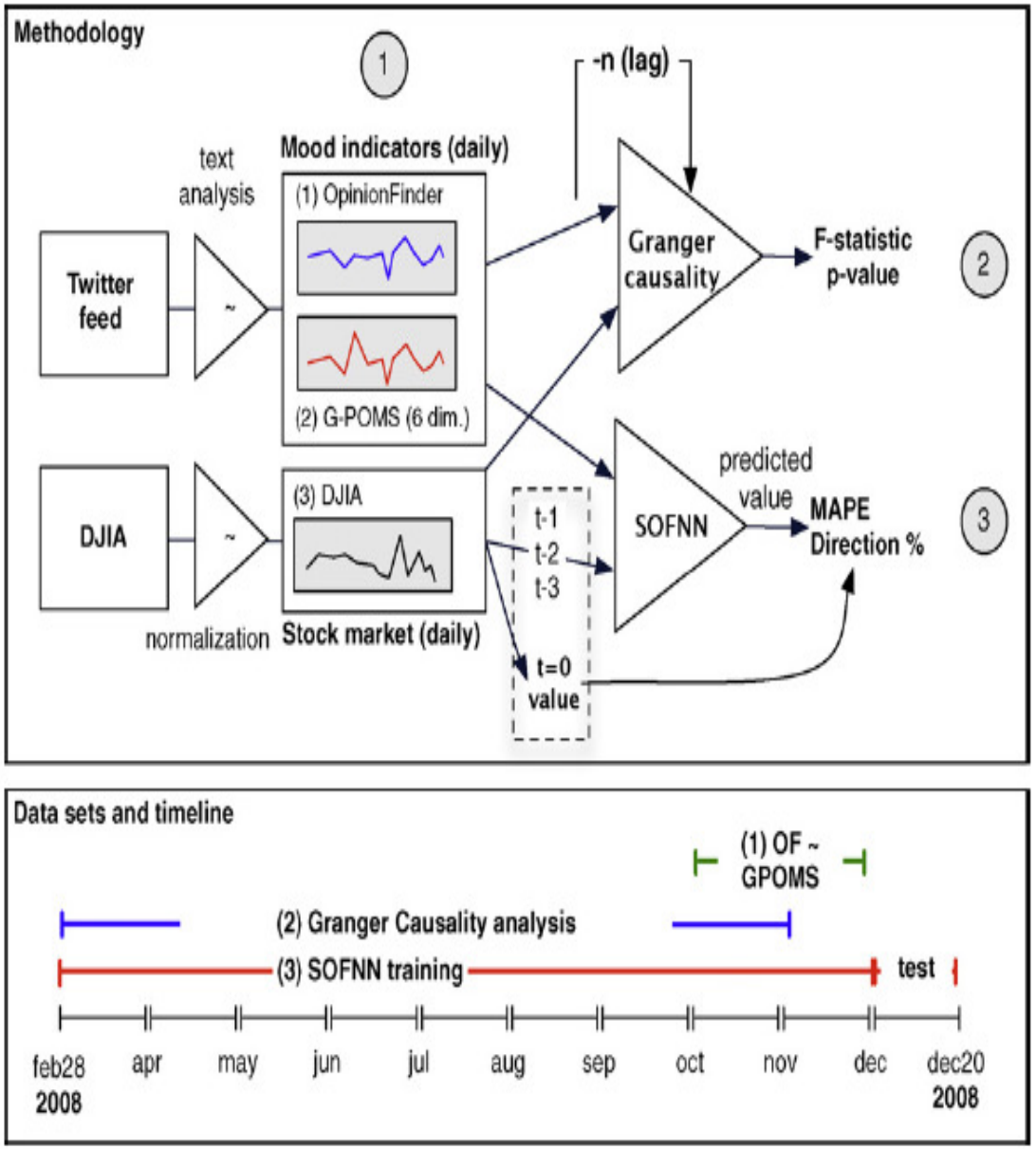
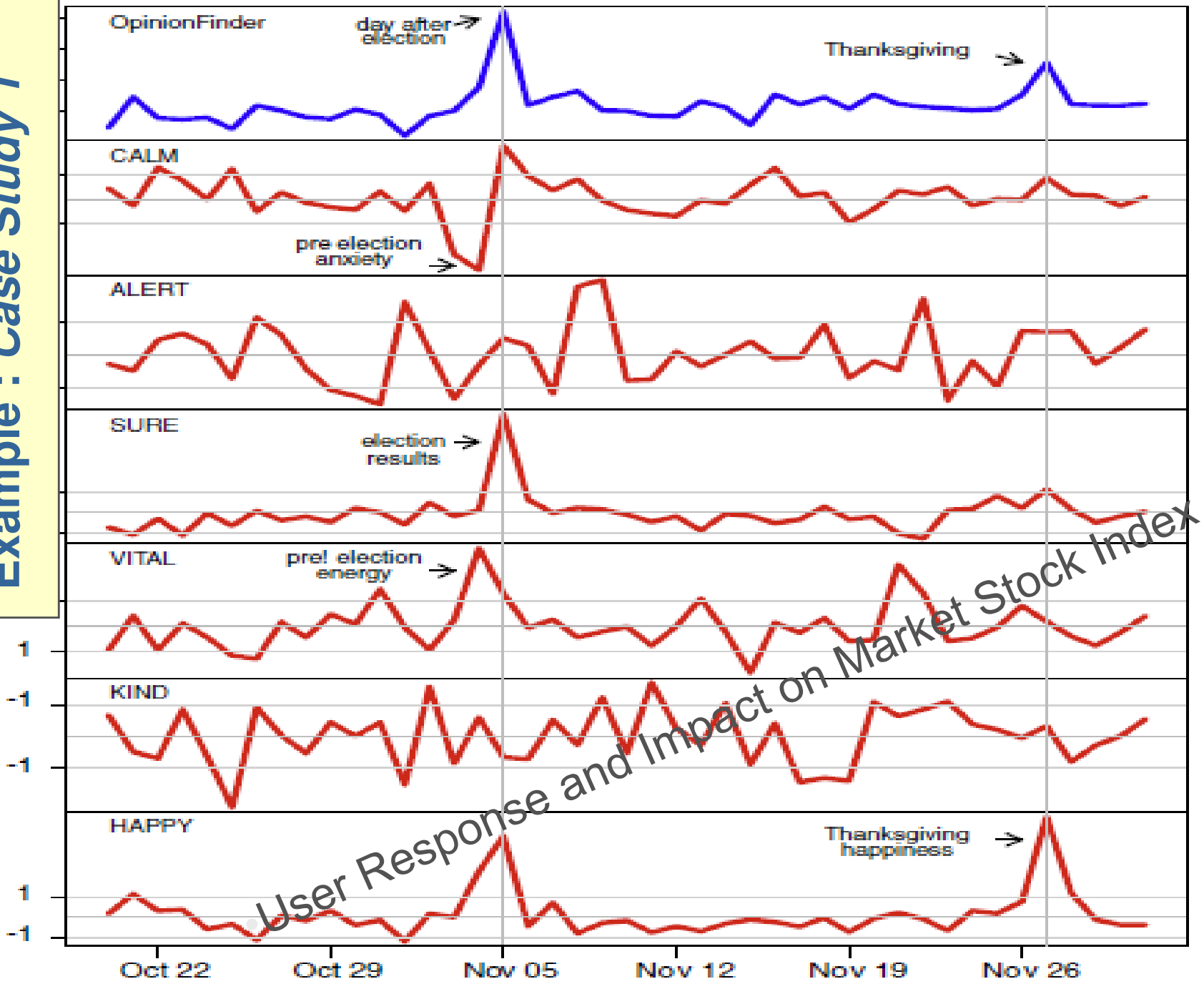
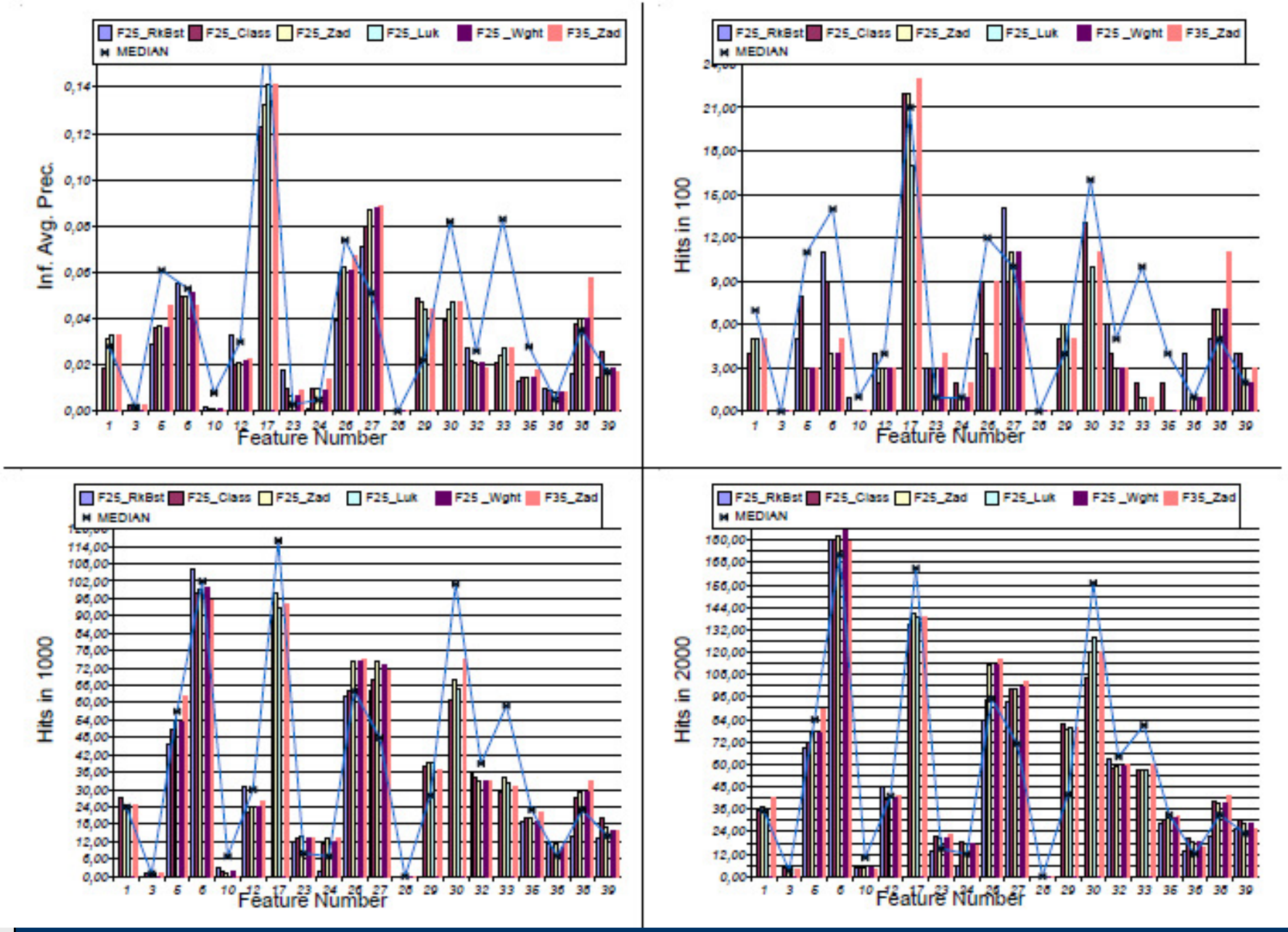


Fig. 1. Diagram outlining 3 phases of methodology and corresponding data sets: (1) creation and validation of OpinionFinder and GPOMS public mood time series from October 2008 to December 2008 (Presidential Election and Thanksgiving), (2) use of Granger causality analysis to determine correlation between DJIA, OpinionFinder and GPOMS public mood from August 2008 to December 2008, and (3) training of a Self-Organizing Fuzzy Neural Network to predict DJIA values on the basis of various combinations of past DJIA values and OF and GPOMS public mood data from March 2008 to December 2008.

Big Data Analytics : A Practical Research Example : *Case Study 1*



User Response and Impact on Market Stock Index


F25_RkBst
F25_Class
F25_Zad
F25_Luk
F25 _Wght
F35_Zad
MEDIAN
Inf. Avg. Prec.
Hits in 100
Hits in 1000
Hits in 2000
Feature Number

Fuzzy Analytics Model : Retail Sales Case Study 2

Let us assume the following variables :-

C = Cash Collection per Counter at one business unit

TC = Total Cash Collection at one business unit

t = Unit Transaction Amount per POS Counter at one business unit

FSH = FASHION Division

S (Total_UAE) = Total Sales in UAE

S (Total_Kuwait) = Total Sales in Kuwait

S (Total_South_Africa) = Total Sales in South Africa

S (Total Group) = Total Sales in ABCD LLC Group

Suppose there are 5 Business Locations of Business Units in UAE namely A , B , C , D and E.

For Location A has 10 POS counters namely Counter 1 , Counter 2 , Counter 3 , ..., Counter 10.

Total Cash Collected at midnight at Outlet : $TC(A) = \sum_{i=1}^{n=10} C(i)$

Fuzzy Analytics Model : Retail Sales Case Study 2

Let us assume the following variables :-

C = Cash Collection per Counter at one business unit

TC = Total Cash Collection at one business unit

t = Unit Transaction Amount per POS Counter at one business unit

FSH = FASHION Division

S (Total_UAE) = Total Sales in UAE

S (Total_Kuwait) = Total Sales in Kuwait

S (Total_South_Africa) = Total Sales in South Africa

S (Total Group) = Total Sales in ABCD LLC Group

Suppose there are 5 Business Locations of Business Units in UAE namely A , B , C , D and E.

For Location A has 10 POS counters namely Counter 1 , Counter 2 , Counter 3 , ..., Counter 10.

Total Cash Collected at midnight at Outlet : $TC(A) = \sum_{i=1}^{n=10} C(i)$

Fuzzy Analytics Model : Practical Example Contd…

Total Cash per Business Unit

= Summation of all unit Transactions over all POS counters at midnight at Outlet : TC(A)

$$= \text{Total Cash Collected at midnight at Outlet : } TC(A) = \sum_{i=1}^{n=10} C(i)$$

$$= \sum_{i=1}^{n=10} \sum_{j=1}^{j=N} t(j) \text{ in UAE Dirhams}$$

Similarly for other locations :

$$TC(B) = \sum_{i=1}^{n=10} \sum_{j=1}^{j=} t(j) \text{ in UAE Dirhams}$$

$$TC(C) = \sum_{i=1}^{n=10} \sum_{j=1}^{j=} t(j) \text{ in UAE Dirhams}$$

Fuzzy Analytics Model : Practical Example Contd…

$$TC(D) = \sum_{i=1} \sum_{j=1} t(j) \text{ in UAE Dirhams}$$

n= 10 j =

$$TC(E) = \sum_{i=1} \sum_{j=1} t(j) \text{ in UAE Dirhams}$$

S (Total_UAE) = Total Sales in UAE =

k = 5 n= 10 j =

$$\sum_{k=1} \sum_{i=1} \sum_{j=1} t(j) \text{ in UAE Dirhams for 5 Business Units}$$

k = X n= 10 j =

$$\sum_{k=1} \sum_{i=1} \sum_{j=1} t(j) \text{ in UAE Dirhams for X no of Business Units}$$

Fuzzy Analytics Model : Practical Example Contd…

Total Sales in ABCD Group

= S (Total_UAE) + S (Total_South_Africa) + S (Total_Kuwait)

=

$$\sum_{k=1}^{k=X} \sum_{i=1}^{n=10} \sum_{j=1}^{j=} t(j)\ dl$$

in UAE Dirhams for Y no of Business Units in three countries

Now t(j) = f (X,Y) = A(X,Y) , i.e. every transaction unit is an Account function of Fashion Products and Food items defined in different classes called Transaction types like Credit Card - Visa , Void , Foreign Currency types.

Statistical Modeling and Computational Techniques of Data Analysis : Practical Example Contd ….

Thus , Total Sales in ABCD LLC Group

= S (Total_UAE) + S (Total_South_Africa) + S (Total_Kuwait)

=

$$\sum_{k=1}^{k=X} \sum_{i=1}^{n=10} \sum_{j=1}^{j=} t(j)\ dl$$

=

$$\sum_{k=1}^{k=X} \sum_{i=1}^{n=10} \sum_{j=1}^{j=} A(x,y)\ dl$$

Total Sales Revenue (ABCD Group) = P

=

$$\sum_{k=1}^{k=X} \sum_{i=1}^{n=10} \sum_{j=1}^{j=} A(x,y)\ dl$$

Total Expenses = Summation of all expenses from GL Expenses = Q (say)

Gross Margin = P - Q

Recommendations from the Fuzzy Analytics Model

- End to end Account Creation the way ERP is done to reflect homogeneous Solution Architecture in terms of Business Intelligence and Reporting.
- Report Formats to be mapped as per the Manual Reporting executed at the branches as provided by Tiger Inc Finance Department.
- Regarding the Reporting Architecture it is indispensable to present the same in a Web-Intelligence mechanism
- The Finance team must have complete transparency in getting all required data with no latency gap.

Input Financial Statements for Fuzzy Analytics

	1993	1994	1995	AVG		
ASSETS						
Cash	6.2	9.4	13.1	9.57		
Accounts receivable	11.3	11.8	12.5	11.87	96.17	12.34
Inventory	6.2	6.6	7	6.60		
Prepaid expenses	0.3	0.6	2.2	1.03		
Other current	0	0.9	0.9	0.60		
Total current	24.0	29.3	35.7	29.67		
Fixed assets	35.3	36.1	36.4	35.93		
TOTAL	59.3	65.4	72.1	65.60		
LIABILITIES AND SHAREHOLDERS' EQUITY						
Accounts payable	7.5	7.9	8.3	7.90		8.21
Other current	0.7	1.3	2.2	1.40		
Total current	8.2	9.2	10.5	9.30		
Long term debt	16.8	20.4	24.3	20.50		
Shareholders' equity	34.3	35.8	37.3	35.80		
TOTAL	59.3	65.4	72.1	65.60		

	1993	1994	1995	AVG
Revenues	91.2	95.8	101.5	96.17
Cost of goods sold	27.4	28.7	30.5	28.87
Gross income	63.8	67.1	71	67.30
Operational expenses	50.1	52.7	55.8	52.87
Operating Income	13.7	14.4	15.2	14.43
Interest	0.9	1.0	1.6	1.17
Income before tax	12.8	13.4	13.6	13.27
Income tax	4.2	5.2	5.2	4.87
Net income	8.6	8.2	8.4	8.40

Working 1: Calculation of Profit from Additional Sales

		1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Equivalent Additional Units		1,650	3,000	3,975	3,975	3,975	3,975	3,975	3,975	3,975	3,975
Price	1.70	1.75	1.80	1.86	1.91	1.97	2.03	2.09	2.15	2.22	2.28
Additional Revenue		2,889	5,411	7,384	7,606	7,834	8,069	8,311	8,560	8,817	9,082
Avg Profit Margin (1993-95)	15.0%										
Profit on Added Sales		**434**	**812**	**1,108**	**1,142**	**1,176**	**1,211**	**1,247**	**1,285**	**1,323**	**1,363**

Working 2: Calculation of Additional Working Capital Needs

		1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Avg Inventory Days (1994-96)	24.7										
Inventory Requirements		182	341	466	480	494	509	524	540	556	573
Avg Receivables/Sales (1993-95)	12.34%										
Receivables Increase		357	668	911	939	967	996	1,026	1,056	1,088	1,121
Avg Payables/Sales (1993-95)	8.22%										
Payables Increase		(237)	(445)	(607)	(625)	(644)	(663)	(683)	(703)	(724)	(746)
Working Capital Needs		301	564	770	793	817	842	867	893	920	947
Change in Working Capital		**301**	**263**	**206**	**23**	**24**	**25**	**25**	**26**	**27**	**28**

Working 3: Calculation of Company WACC

Return on Equity

Risk-free Rate	8.06%	(10-year Bond)
Beta	0.85	
Market Risk Premium	6%	
Return on Equity	**13.16%**	
WACC	**10.5%**	

Return on Debt

Risk-free Rate	8.06%	(10-year Bond)
Default Spread	2%	
Return on Debt	**10.06%**	
Weight of Debt	0.4	

Evaluation of Expansion Project at Tiger Inc

All amounts are in Thousands of Dollars

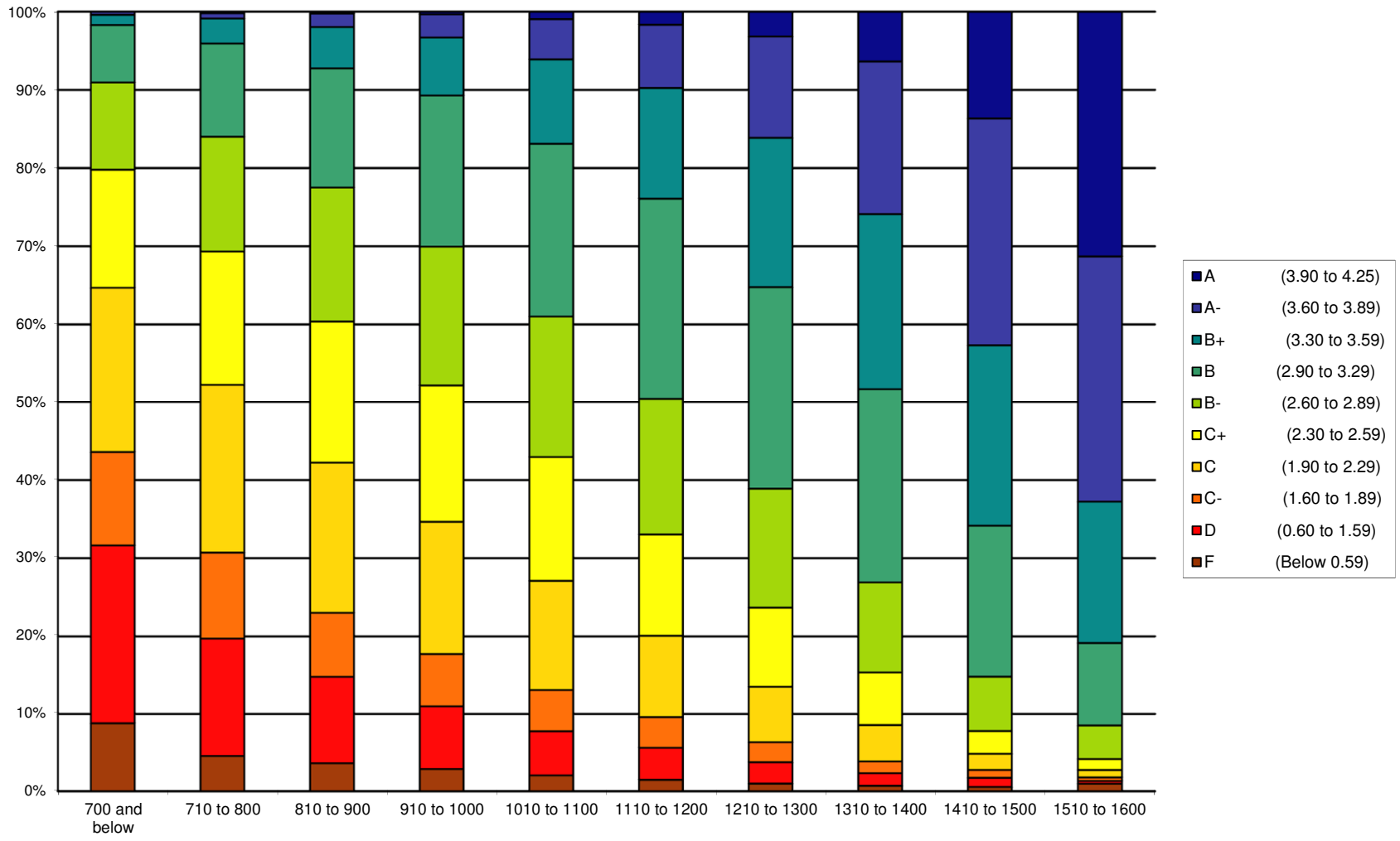
	1995*	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Ignored
Project consultancy												40,000
Salaries and administrative expenses												223,000
Building Expansion	(1,300)										1,300	
Spiral Freezer	(1,600)											
Processing Line	(1,300)											
Warehouse	(600)										600	
Contingency needs	(400)										400	
Land Cost	(250)										250	
CF from Added Sales (W-1)		282	528	720	742	764	787	811	835	860	886	
Working Capital Inv (W-2)		(301)	(263)	(206)	(23)	(24)	(25)	(25)	(26)	(27)	920	
Depreciation Tax Shield		102	102	102	102	102	102	102	102	102	102	
Capital Expenditures		(290)	(290)	(290)	(290)	(290)	(290)	(290)	(290)	(290)	(290)	
Tax Shield on CCA		152	269	208	161	126	100	80	65	53	606	
Cost Savings		81	129	138	138	138	138	138	138	138	138	
Other Savings		90	92	95	98	101	104	107	110	114	117	
Net Cash Flow	(5,450)	115	566	767	928	917	916	922	933	950	5,028	
Hurdle Rate	18%											
WACC (W-3)	10.5%											
NPV at Hurdle Rate	(1,548)											
NPV at WACC	481											
IRR	11.9%											

Assumptions:

Tax Rate	35%
Inflation Rate	3%

* For simplicity it is assumed that the expansion will be completed by year end 1995 and regular sales will incur in 1996
* if the inflation is increased from 3 to 5, then price can be moved-up, so the NPV & IRR will increase
* or considering the business potential, executive committie should approve lower (11%) hurdle rate

Federated Statistical View of Sales Tiger Inc Data Analysis



Fuzzy Analytics Project Phases

- Big Data Source Analysis Phase
- Statement of research questions
- Methods used to answer research questions
- File restructuring procedures (syntax creation, adding new variables as needed)
- Algorithms for scoring, equating, etc.
- Data cleaning procedures (e.g. removing outliers)
- Quality control procedures at every step in the project

Sample Project Gantt Chart : A practical suggestion to Business

Data Gathering Phase (3 Months)

Sessions on Data Analysis (3 Months) : Heavy duty statistical data analysis support

Iterative Data Analysis Checks till Finalization / Data Quality (4-5 Months)

Final Report Preparation : 2 Months

Fuzzy Analytics : Manufacturing Segment : Case Study 3

- What is the role of Fuzzy Analytics in data mining?
- What is the nature of Fuzzy Analytics' contribution towards Business?
- What is the role of an Fuzzy Analytics in manufacturing?

Fuzzy Analytics in Manufacturing : Merits Unleashed

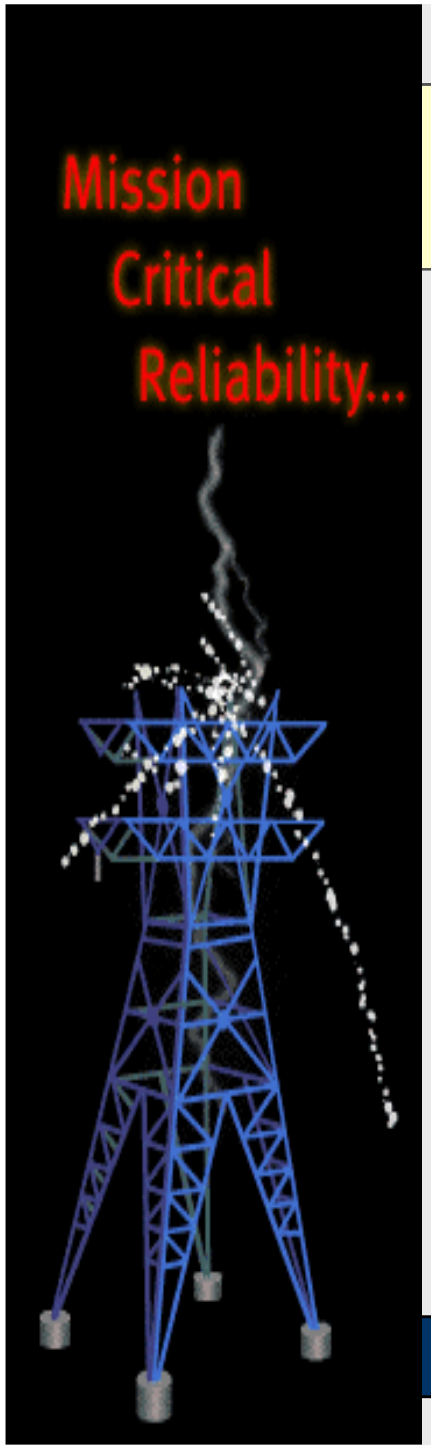
- Production Scheduling
- Advanced Planning and Scheduling
- Production Reporting
- Inventory Management
- Production Accounting
- Production Capacity Planning
- Materials Requirements Planning
- Production Process Control

Fuzzy Analytics : From Mfg Plant Critical Infrastructure Big Data Perspective

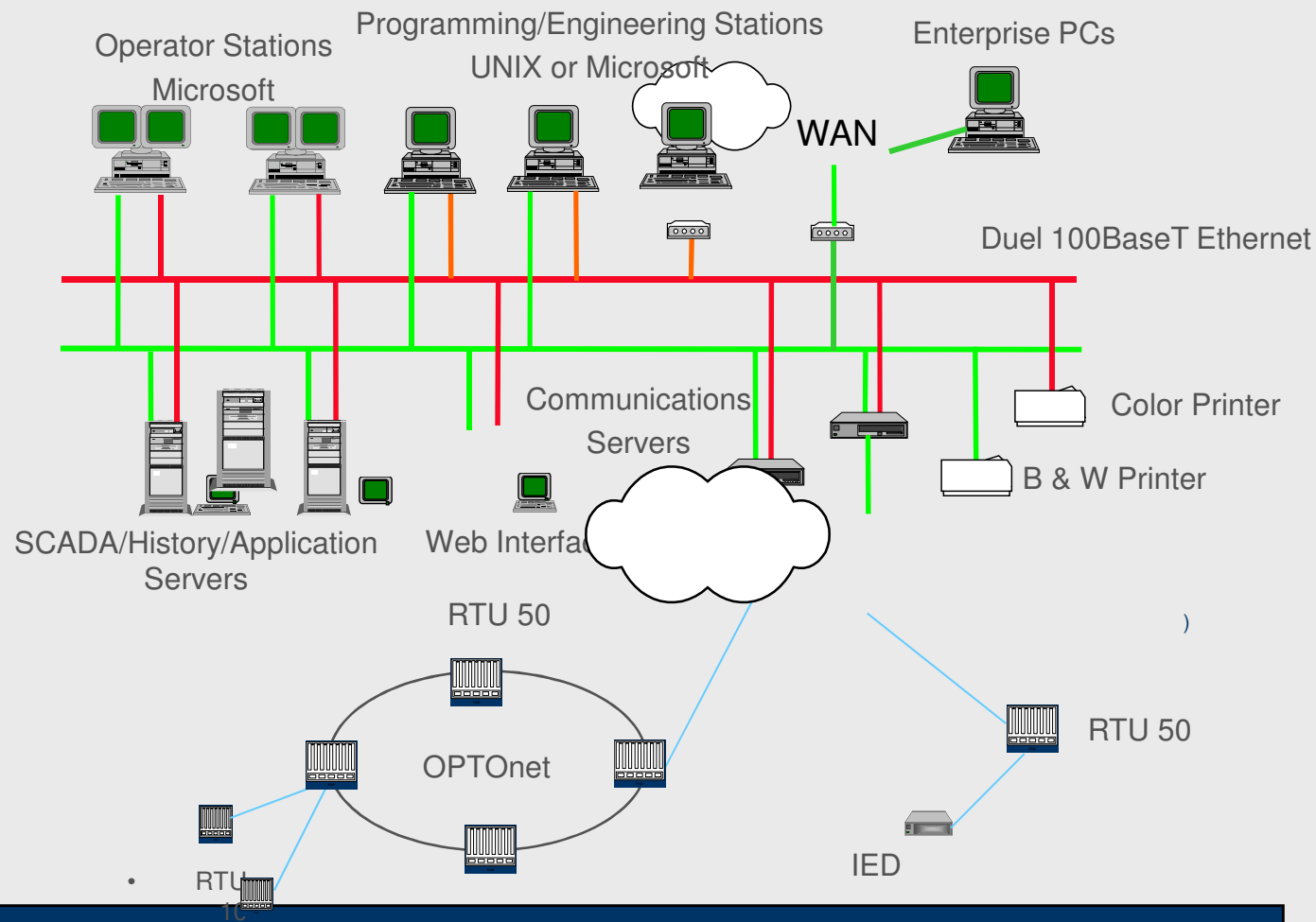
- **<u>Traditional SCADA</u> for pipelines, power distribution or anything that spans great distances**
  - Typical
  - Stand-Alone
  - Networked
- **<u>Local control SCADA</u> for plant or unit operations or any architecture that is within a defined boundary**
  - DCS
  - PC based
  - HMI / PLC

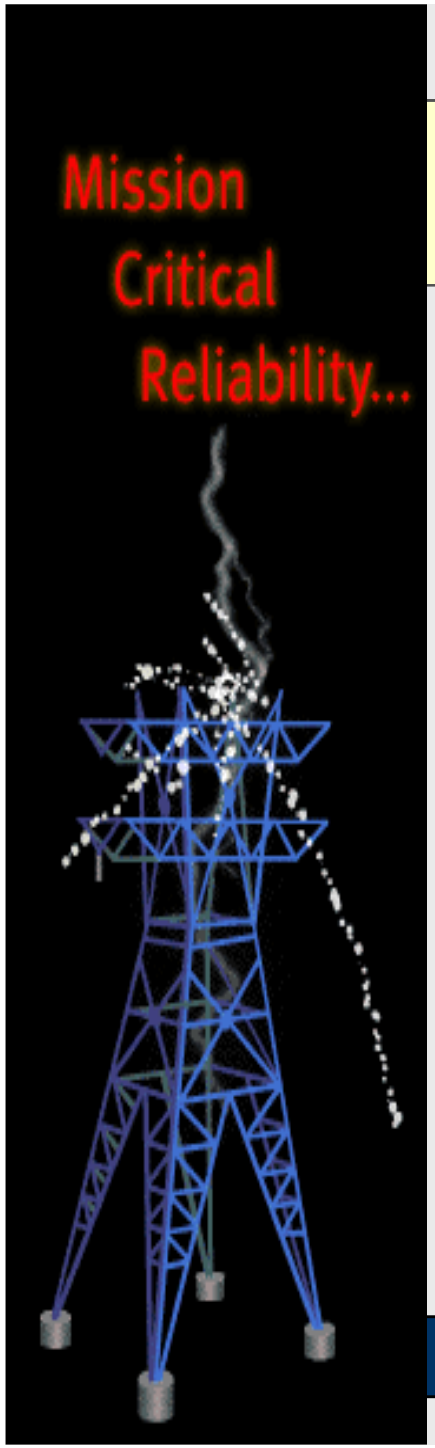
Fuzzy Analytics …

The dashboard solutions, from the production line to the bottom line.

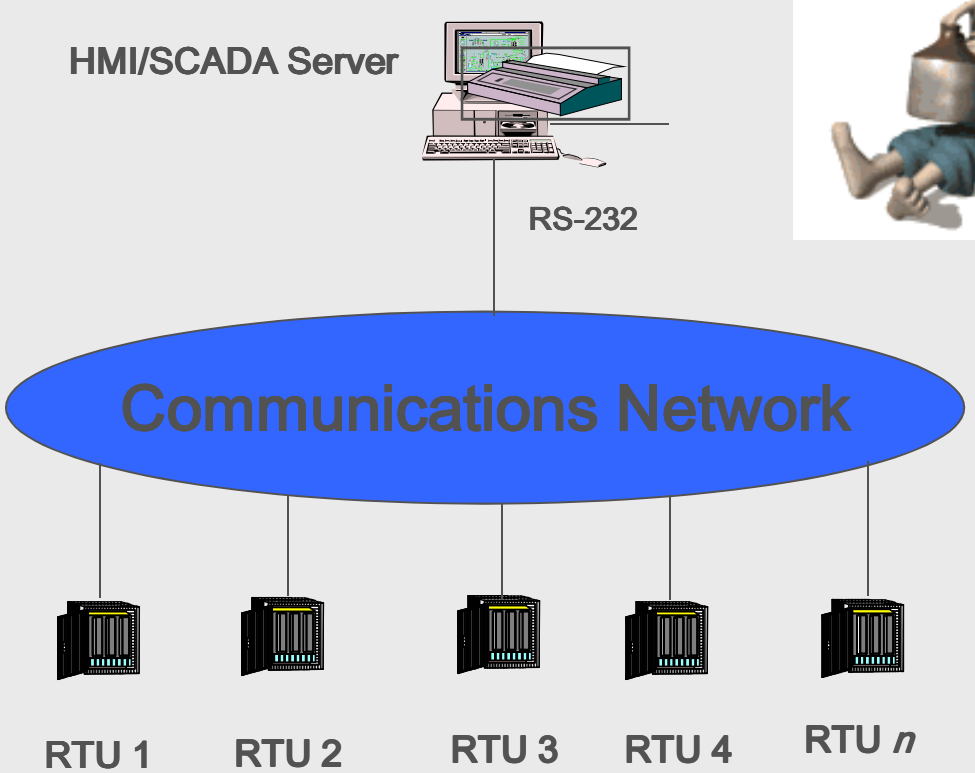


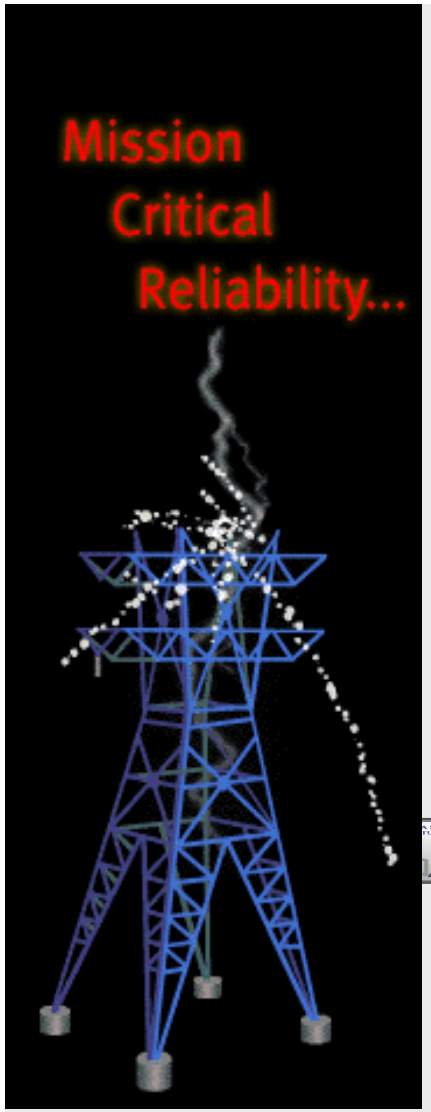
Typical SCADA Platform



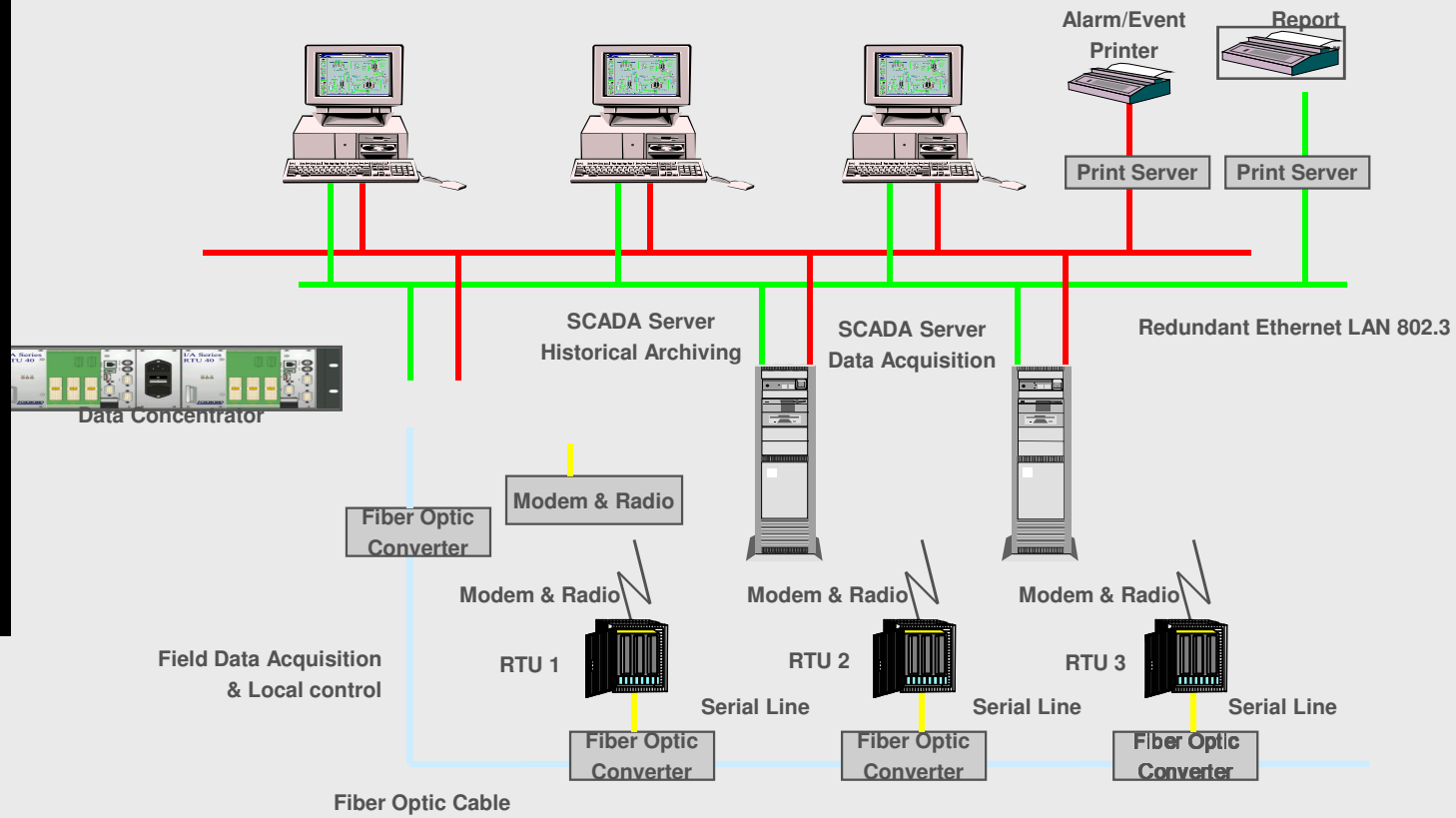


Stand Alone SCADA

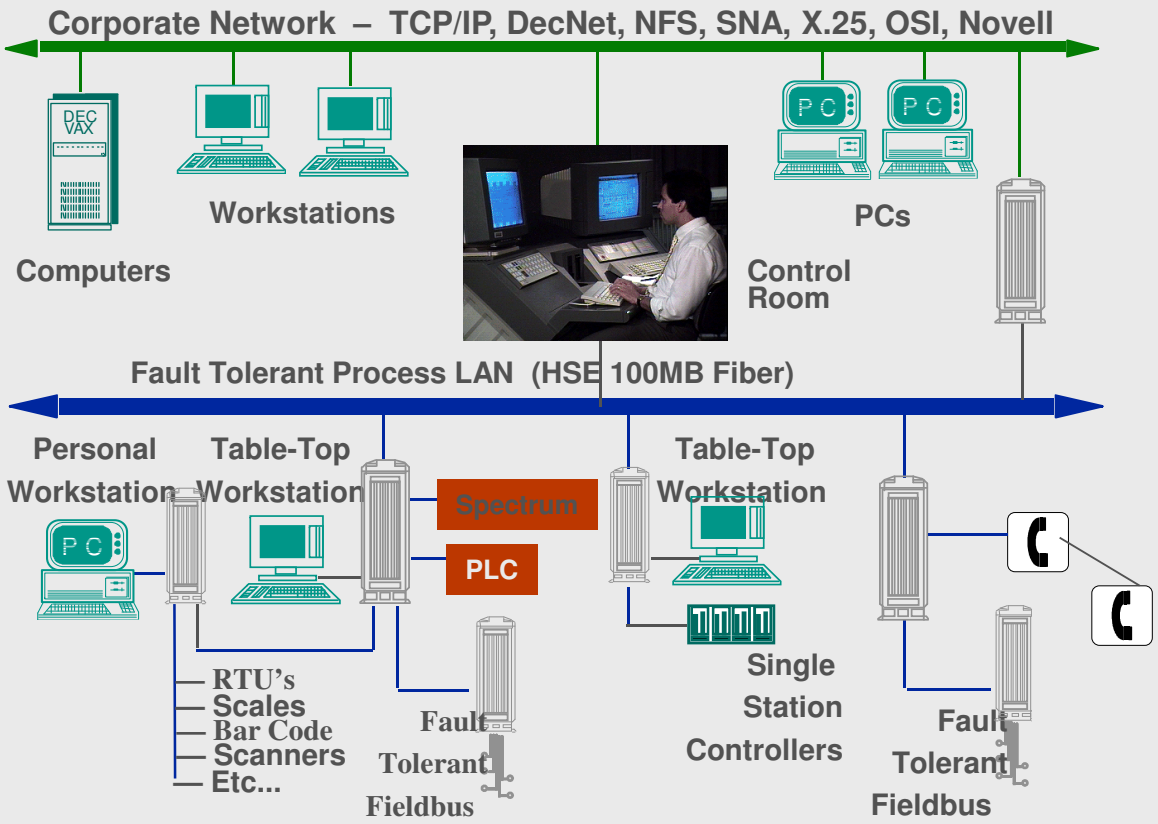




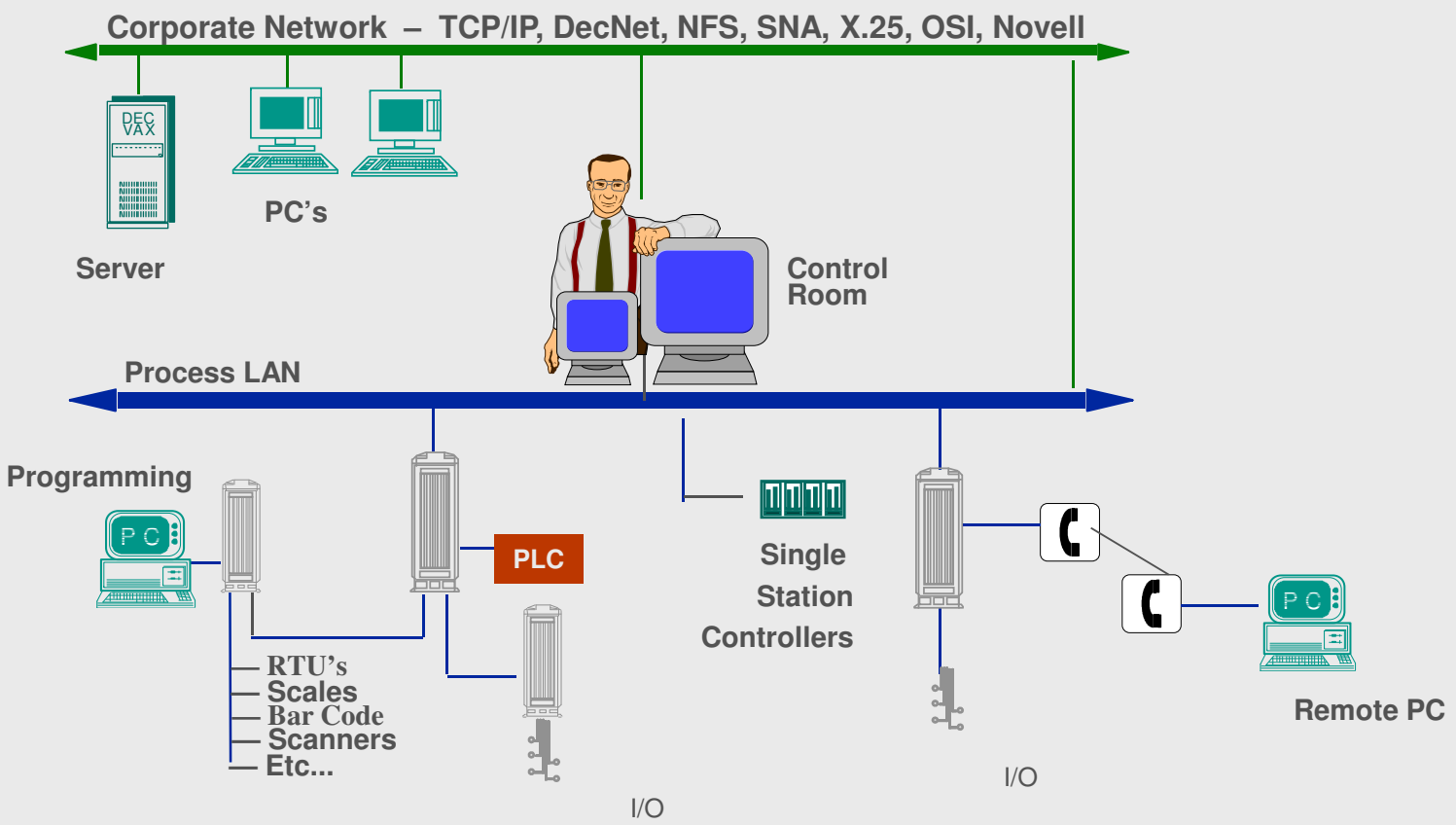
Fuzzy Networked SCADA Data

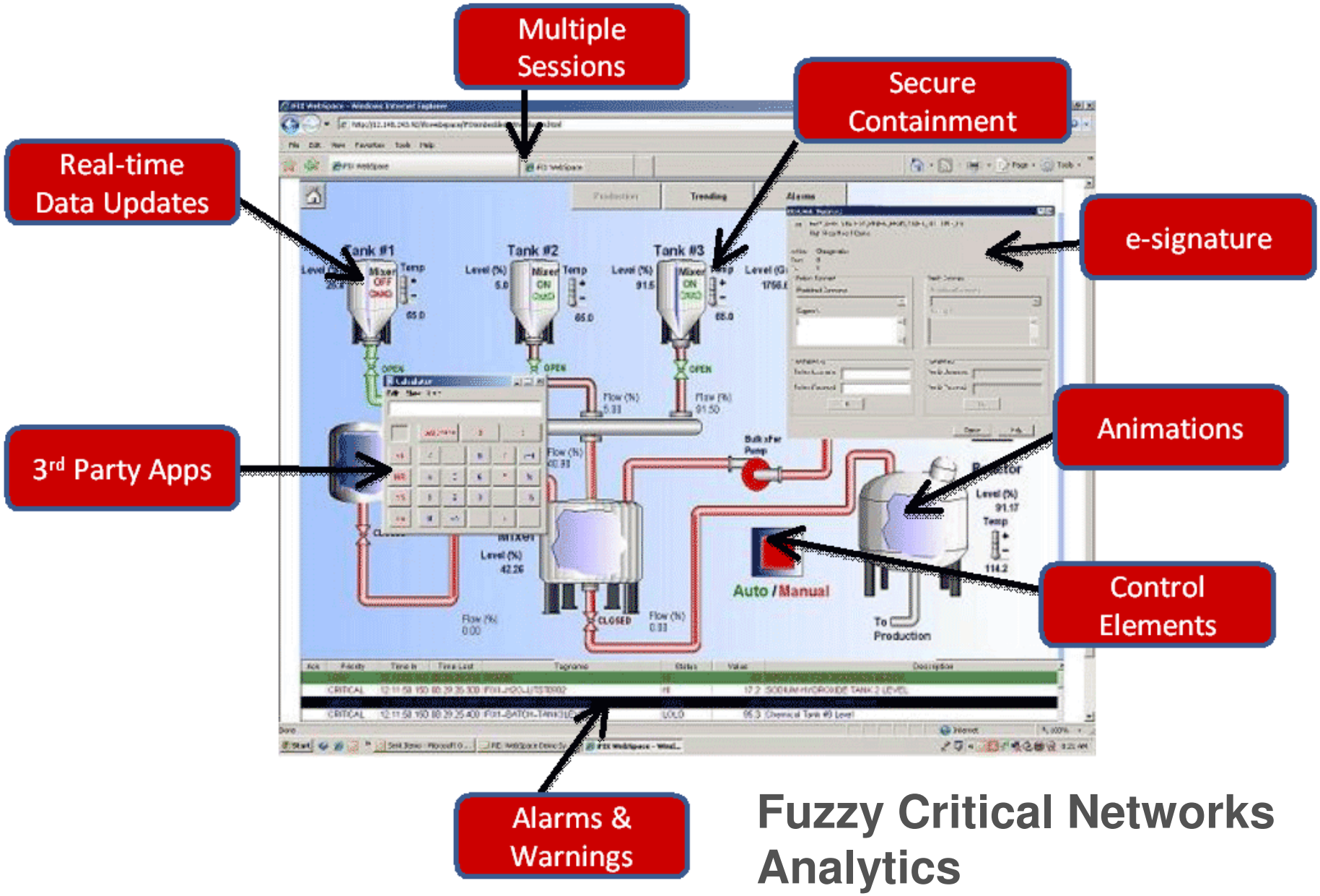


Production Plant Control



Networked HMI/PLC

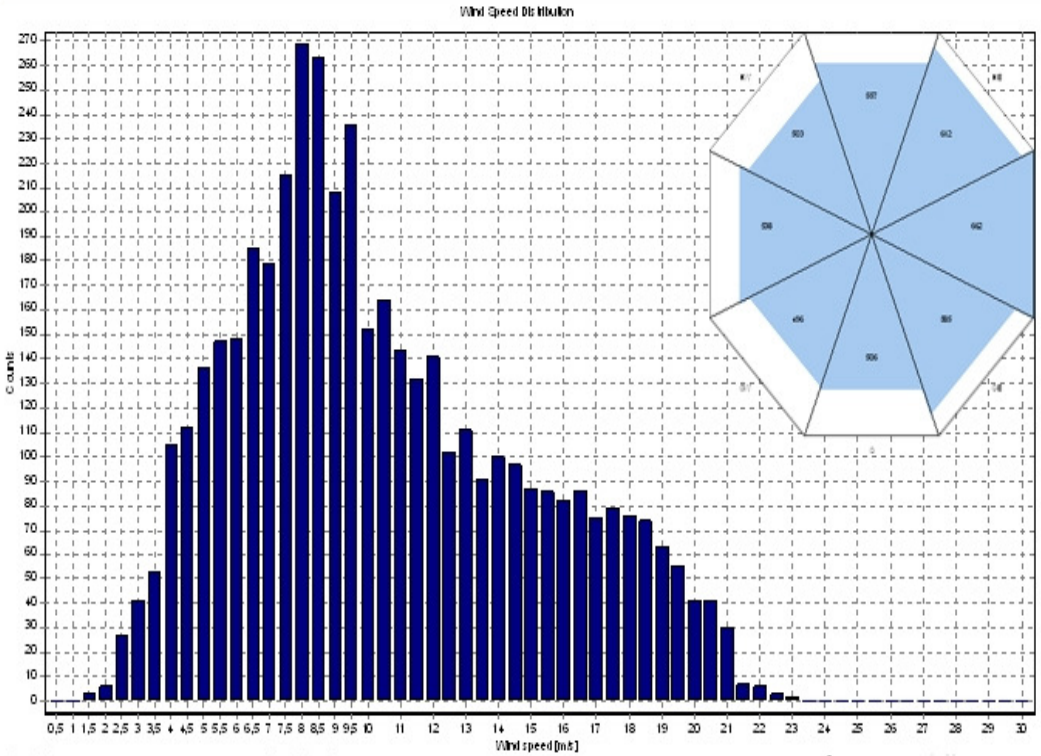




Fuzzy Critical Networks Analytics

RCP ver. 1.0.1.110

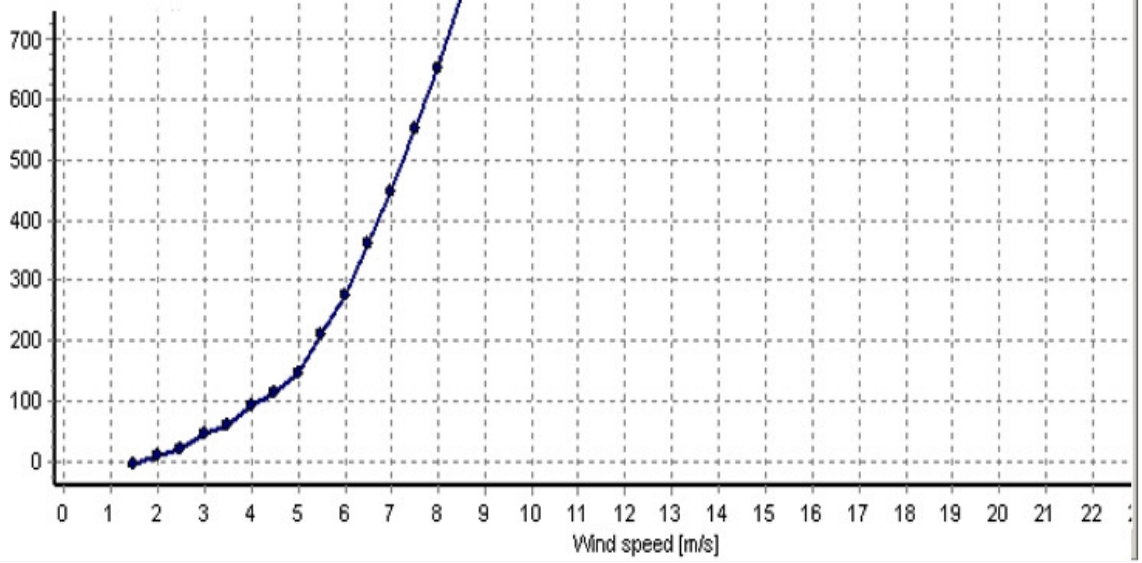
File Window Help



: Power Curve Report
: DEIF WIND PARK\Line A\Mao 013

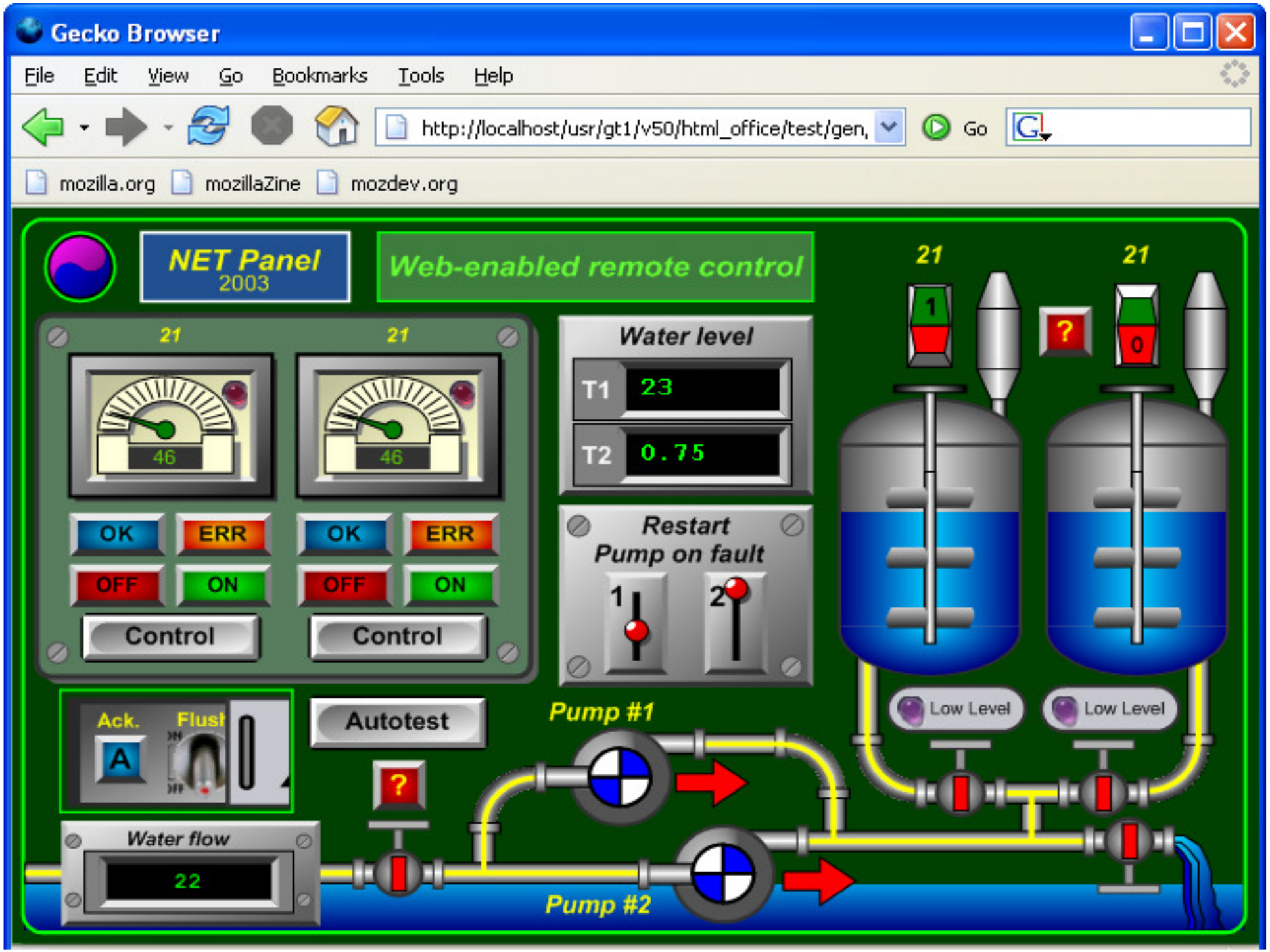
: 08-02-2010
: 12-03-2010 23:59:59

: 792,00 h
: 93,64 %

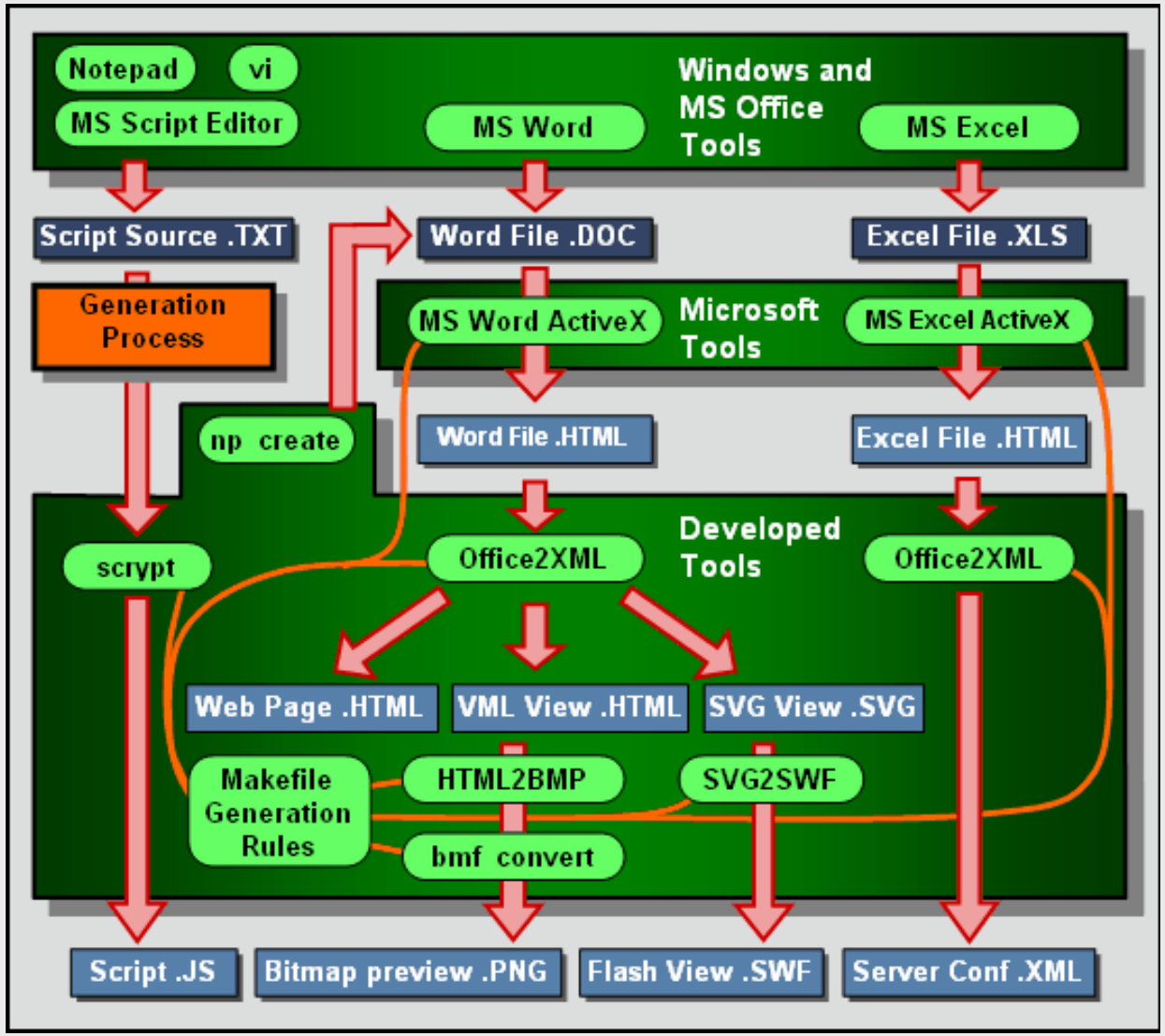


Report Type	: Stop Report
Turbine ID	: DEIF WIND PARK\Line A\Mao 013
From	: 08-02-2010
To	: 12-03-2010 23:59:59
Report TimeSpan	: 792,00 h
Report Certainty Coefficient	: 75,51 %

Stop Condition no.	Stop condition text	Number of occurrences	Down time [h]
GRID FAILURE			
-	-	-	-
WEATHER CONDITIONS STOP			
16002	20 s Average Windspeed Over Limit	326	118,30
16001	3 s Windspeed Over Limit	11	21,61
20001	Windspeed Too Low For Turbine Operation	4	1,34
OPERATOR STOP			
15003	Pitch Battery Test Performed	4	0,08
18001	SCADA System Stop	2	0,07
TURBINE FAILURE			
15002	Extreme Sudden Change in Wind Direction	37	1,25
CONTROLLER MALFUNCTION			
-	-	-	-
	In total:	384	142,65


Gecko Browser
File Edit View Go Bookmarks Tools Help
http://localhost/usr/gt1/v50/html_office/test/gen,
Go
mozilla.org
mozillaZine
mozdev.org
NET Panel
2003
Web-enabled remote control
21
21
46
46
OK
ERR
OFF
ON
Control
Water level
T1
23
T2
0.75
Restart
Pump on fault
1
2
Ack.
Flush
A
Autotest
Pump #1
Pump #2
Water flow
22
Low Level
Low Level

PLC Fuzzy Data Analytics Architecture



Critical Fuzzy Factor Analytics : ANOVA

NWA Quality Analyst Graphics - [StblANCOVAReport.JJ.nwg]

File Edit View Settings Window Help

CS-303 Stability Studies
RESULTS predicted by Time in Months
ANCOVA - Analysis of Covariance

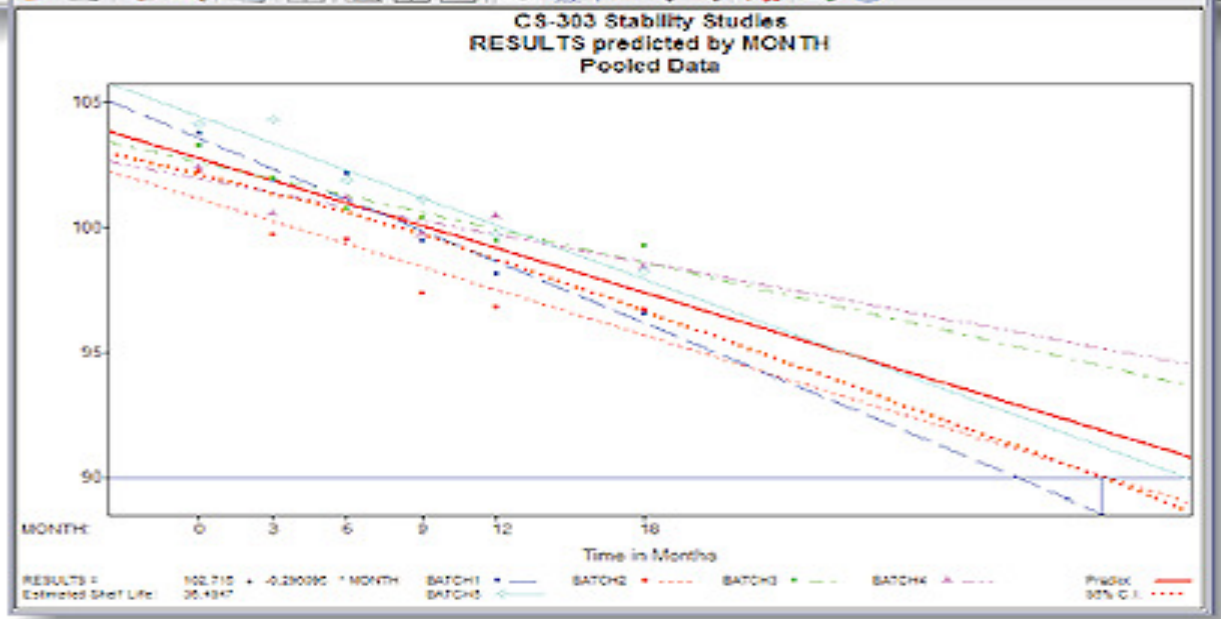
Batch	df	SSx	xy	SSy	Slope (b coeff)	Intercept (a coeff)	Sb	Deviation from Regression df	SS	MS
BATCH1	5	210	-86	37.1333	-0.409524	103.543	0.0477379	4	1.91428	0.47857
BATCH2	5	210	-63.5	23.1883	-0.302381	101.136	0.0588954	4	3.98713	0.996782
BATCH3	5	210	-46.6	11.7483	-0.221905	102.559	0.040935	4	1.40756	0.351891
BATCH4	5	210	-38.8	9.13333	-0.184762	101.911	0.0483509	4	1.96458	0.491144
BATCH5	5	210	-76	28.9734	-0.361905	104.429	0.0418127	4	1.46858	0.367144
MONTH	5	210	210	210	1	0	0	4	0	0
								24	10.7421	0.447588
Pooled within Groups	30	1260	-100.9	320.177	-0.0800794	8.64063	0.0924189	29	312.097	10.762
					Difference Between Slopes			5	301.355	60.2709
Between Groups	5	0	0.000083519	42665.2						
Within + Between	35	1260	-100.9	42985.3	-0.0800793	85.5962	1.0016	34	42977.3	1264.04
					Difference Between Pooled and Total			5	42665.2	8533.03

F Crit F prb = .26

Comparison of Slopes: F(5, 2
Comparison of Intercepts: F(5, 2

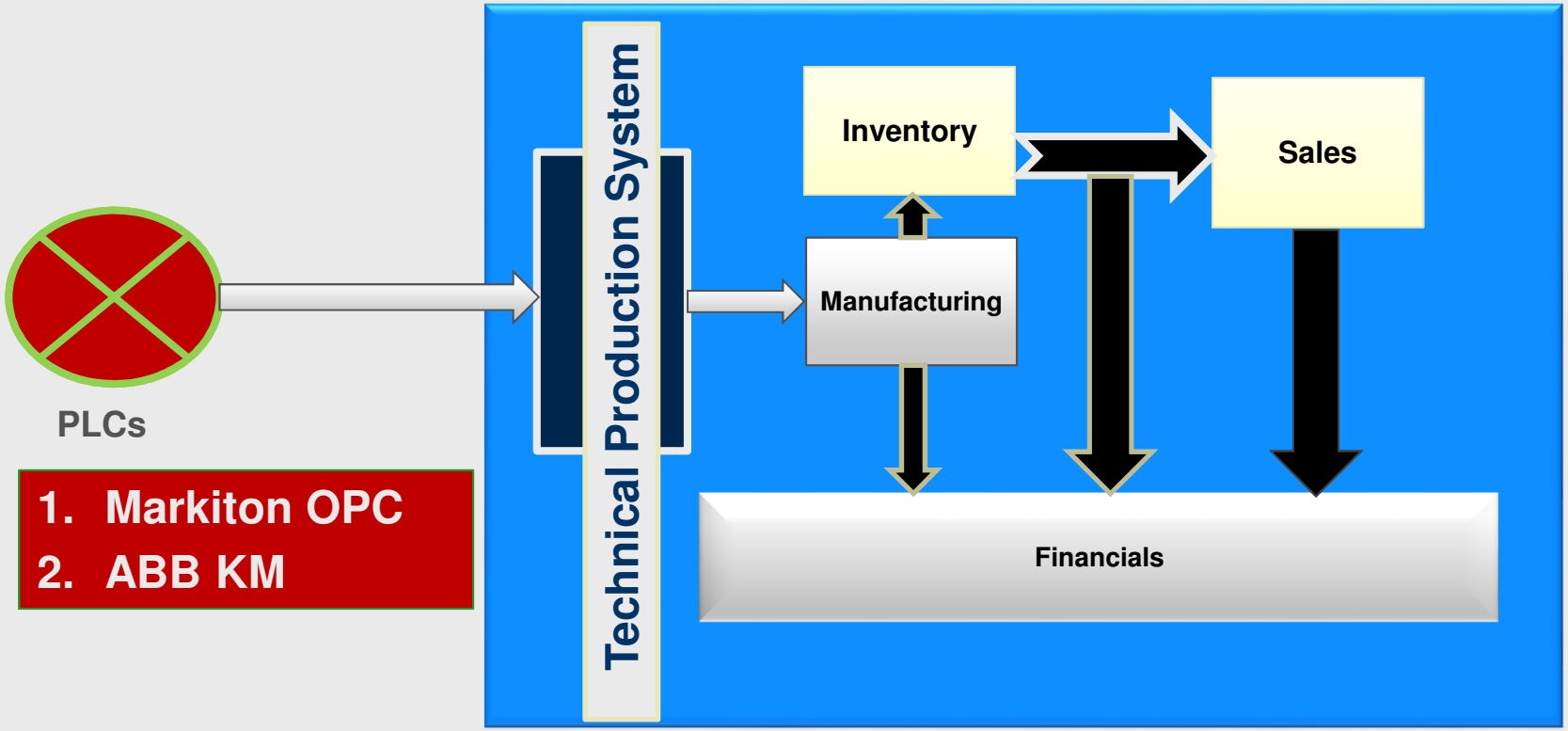
NWA Quality Analyst Graphics - [PooledData.nwg]

File Edit View Settings Window Help



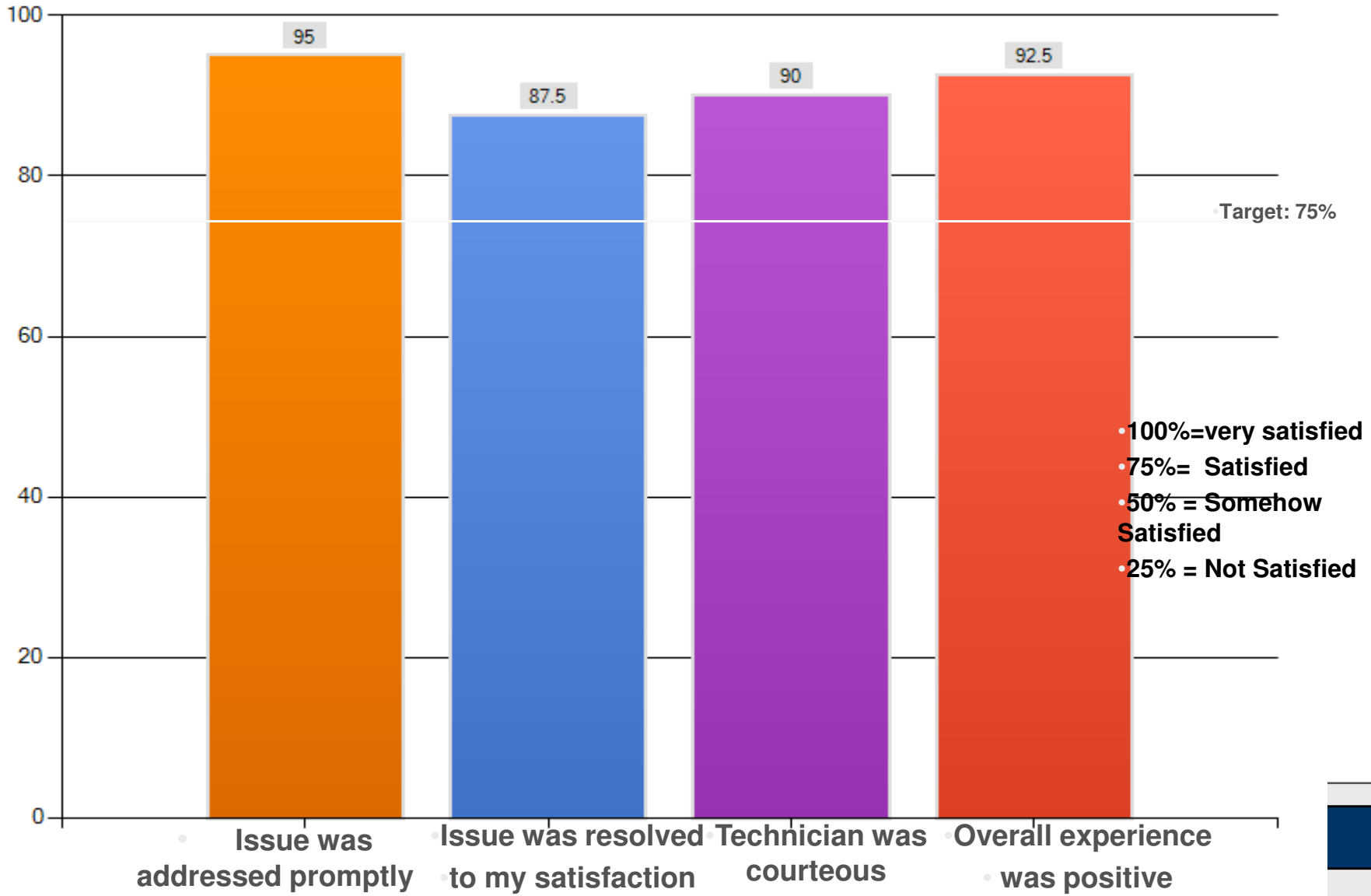
ERP Production Application Solution Architecture

: An outcome of Fuzzy Analytics of Critical Networks



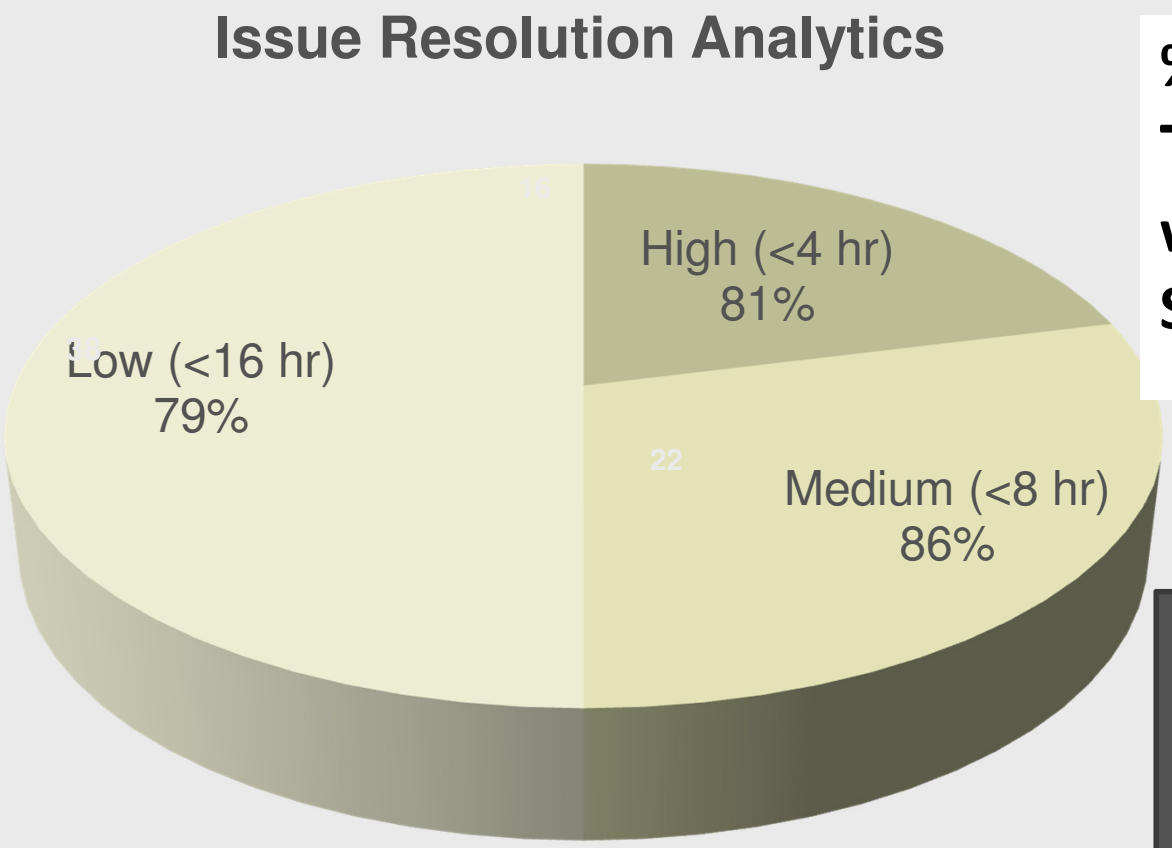
1. Markiton OPC
2. ABB KM

Critical Infrastructure Issues Resolution Survey



- 100%=very satisfied
- 75%= Satisfied
- 50% = Somehow Satisfied
- 25% = Not Satisfied

Critical Infrastructure Issues Resolution Time

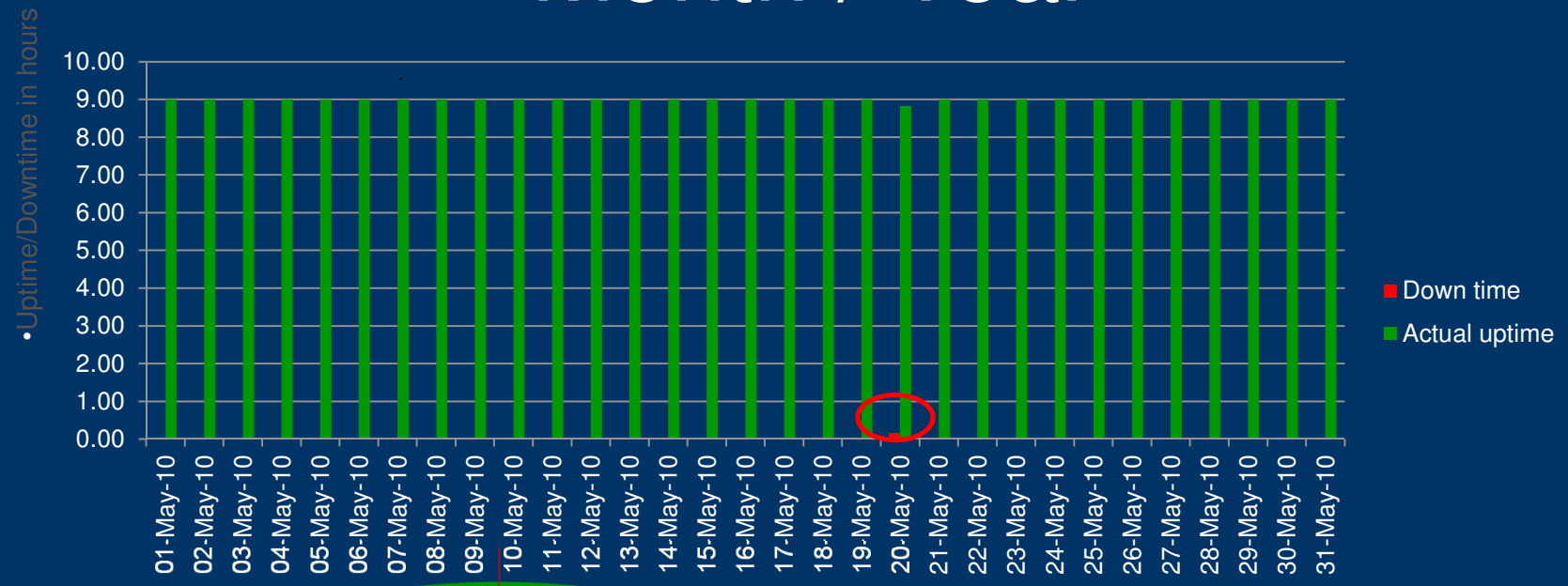


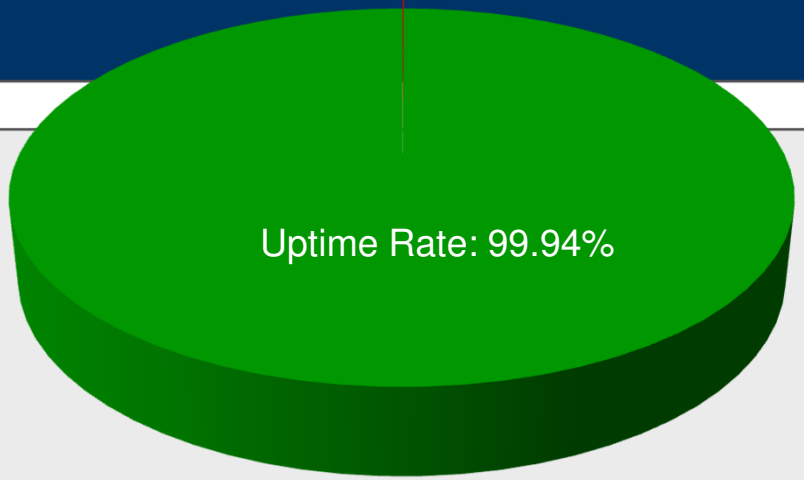
% Tickets within SLA

Average : 82%

Target: >75%

Systems Uptime/Downtime* Month / Year





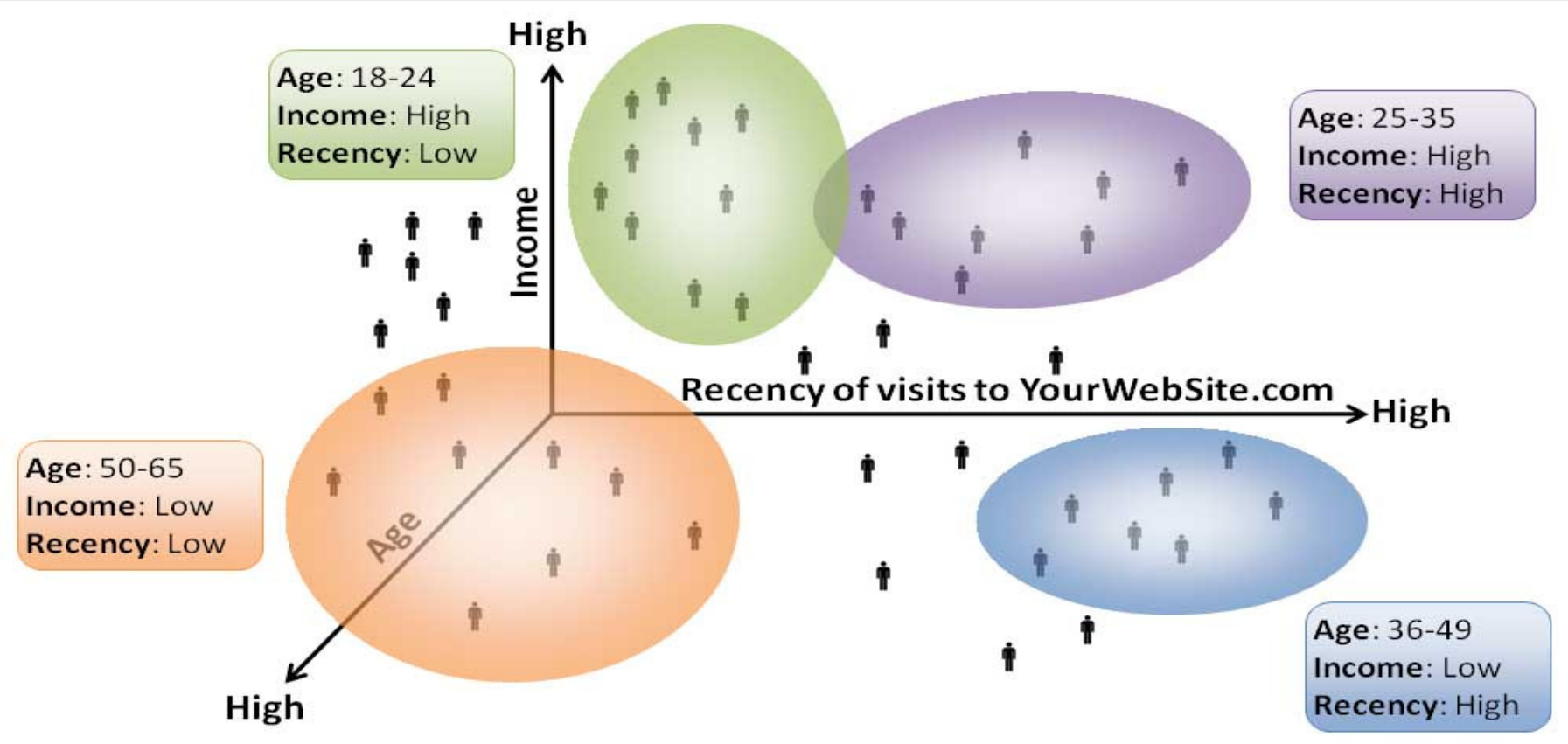
- Uptime – 1850/1860min (99.94%)
- Downtime – 10/1860min (0.06%)
- Target – 99.9%

* 8 AM to 9 PM

Fuzzy SCADA Analytics : Beyond the Big Data Analytics

For a large scale manufacturer , informed decision making is leveraged in the following areas :-

- Distributed Control into the field
- Distributed Computing
- Distributed I/O into the field
- Services, Applications, Solutions
- Industry Standards
- Configurator tools
- Asset Management
- Performance guarantees ($)
- Life Cycle Costs (scalability / migration / upgrades)
- Digital bus selection



FUZZY ANALYTICS : BIG DATA FACTOR ANALYSIS

Fuzzy Analytics : Big Data Factor Analysis - as an exploratory and confirmatory multivariate technique

- **Big Data Factor analysis is commonly used in:**
  - Data reduction
  - Scale development
  - The evaluation of the psychometric quality of a measure, and
  - The assessment of the dimensionality of a set of variables.

Fuzzy Analytics Big Data factor analysis as an exploratory and confirmatory multivariate

- **Regardless of purpose, factor analysis is used in:**
  - the determination of a small number of factors based on a particular number of inter-related quantitative variables.

- **Unlike variables directly measured such as speed, height, weight, etc., some variables such as egoism, creativity, happiness, religiosity, comfort are not a single measurable entity.**

- **They are constructs that are derived from the measurement of other, directly observable variables**

Fuzzy ANALYTICS : BiG DATA Factor Analysis

- Generally, the number of factors is much smaller than the number of measures.
- Therefore, the expectation is that a factor represents a set of measures.
- From this angle, factor analysis is viewed as a data-reduction technique as it reduces a large number of overlapping variables to a smaller set of factors that reflect construct(s) or different dimensions of contruct(s).

Fuzzy Analytics : Big Data Factor Analysis

- The assumption of factor analysis is that underlying dimensions (factors) can be used to explain complex phenomena.
- Observed correlations between variables result from their sharing of factors.
- Example: Correlations between a person's test scores might be linked to shared factors such as general intelligence, critical thinking and reasoning skills, reading comprehension etc.

Ingredients of a Good Factor Analysis Solution

- A major goal of factor analysis is to represent relationships among sets of variables parsimoniously yet keeping factors meaningful.
- A good factor solution is both simple and interpretable.
- When factors can be interpreted, new insights are possible.

Fuzzy Analytics : Big Data Operations Recommendations – A Human Touch Factor

- Major executive involvement and vertical (up/down) communication is crucial to success
- Objectives (KPIs) / metrics/issues should drive the project
- BPR (Business Process Reengineering) should be done up front
- Have a Go / No Go decision point concerning operations
- Effective training is needed at all levels in improving the Data Quality.
- Major communications i.e. Thorough Conference Room Pilot , Frequent User Meetings / Handshakes Conference Calls , Involvement of users in the User testing, simulation are needed, including stress testing
- Have multiple Contingency Plans e.g. Resource Contingencies
- Scope Control during the implementation phase of the project
- Reward contributors any way you can—with praise, training, coveted team assignments, promotions, etc. Your good people are your finest asset.

Questions or Comments