

Crunching Metrics

from Public Data Sources

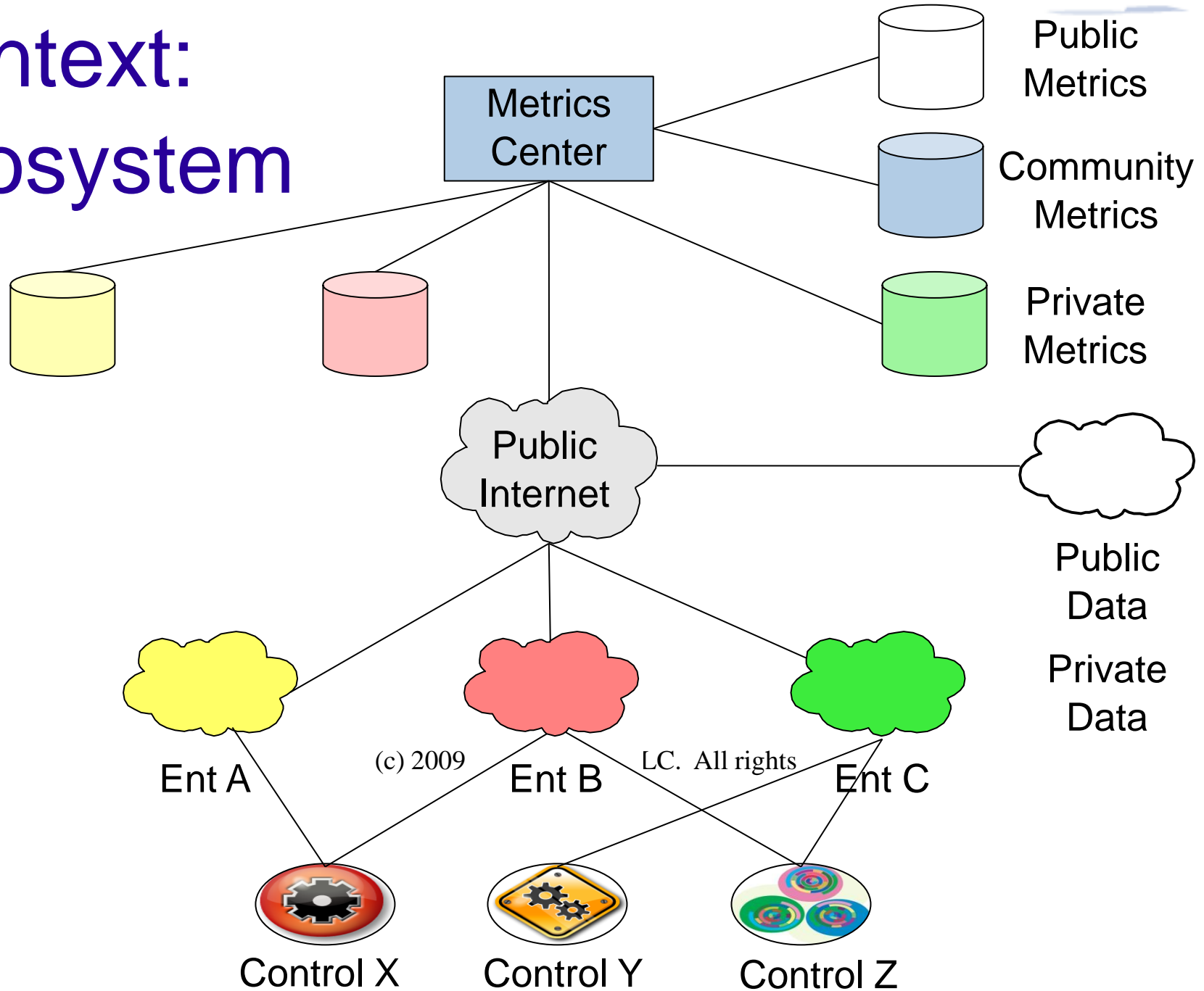
Metricon 4
11 Aug 2009

Elizabeth A. Nichols Ph.D. | CTO | PlexLogic

Outline

- Context(s)
- Questions
- Metrics + Models
- Conclusions

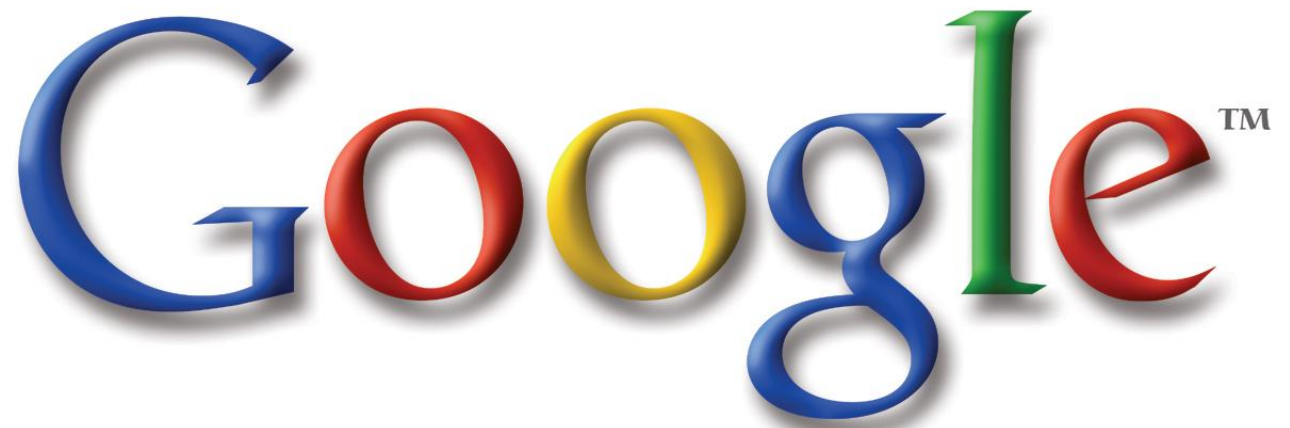
Context: Ecosystem



Context: Individual Business

- Breaches impact the business
- Total Impact = Impact(Expense) + Impact(MktCap)
 - Impact(expense) ~ Frequency + TotalAffected
 - Impact(mkt cap) ~ StockPrice
 - StockPrice ~ market movement + reputation
- Why should I care?
 - Stock price is clearly important. Is there any hard evidence that breaches do affect stock price, say within 60 days of breach?
 - Breach source: Inside vs Outside may imply differing controls or investments in people, process, technology

Context: Public Data Sources

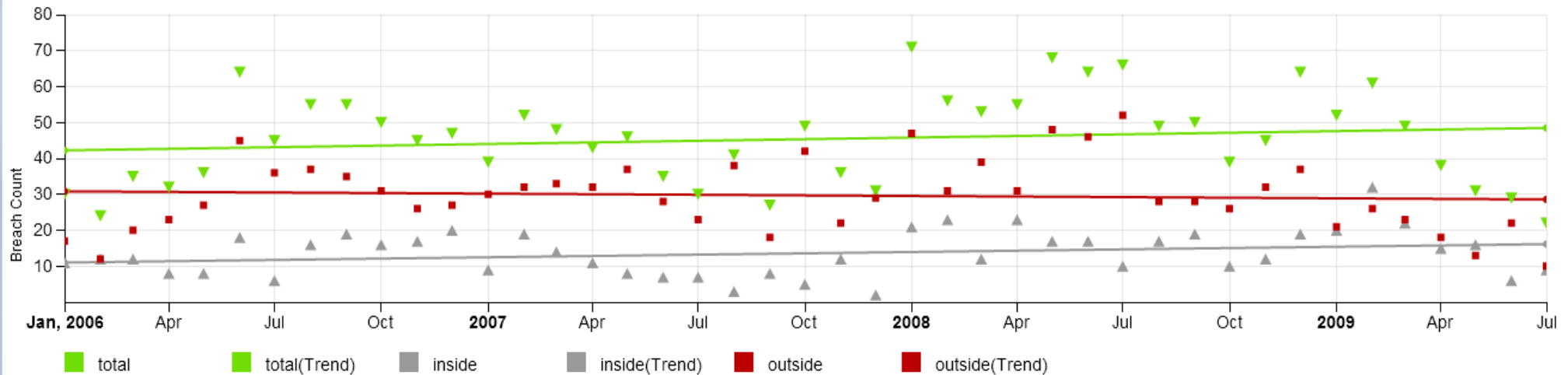


Questions

- Is Breach Frequency Increasing?
 - Frequency (Inside + Outside + Unk) is increasing
 - Frequency (Inside) is increasing
 - Frequency (Outside) is increasing
- How do Breach Frequency and Total Affected compare for Inside vs Outside sources?
 - Frequency (Inside) > Frequency (Outside)
 - TotalAffected(Inside) > TotalAffected(Outside)
- How does a breach affect stock price?
 - StockPrice(before) > StockPrice(after) for N days
 - StockPriceGradient changes sign from before(+) to after(-)

Breach Frequency

BreachCount/Month



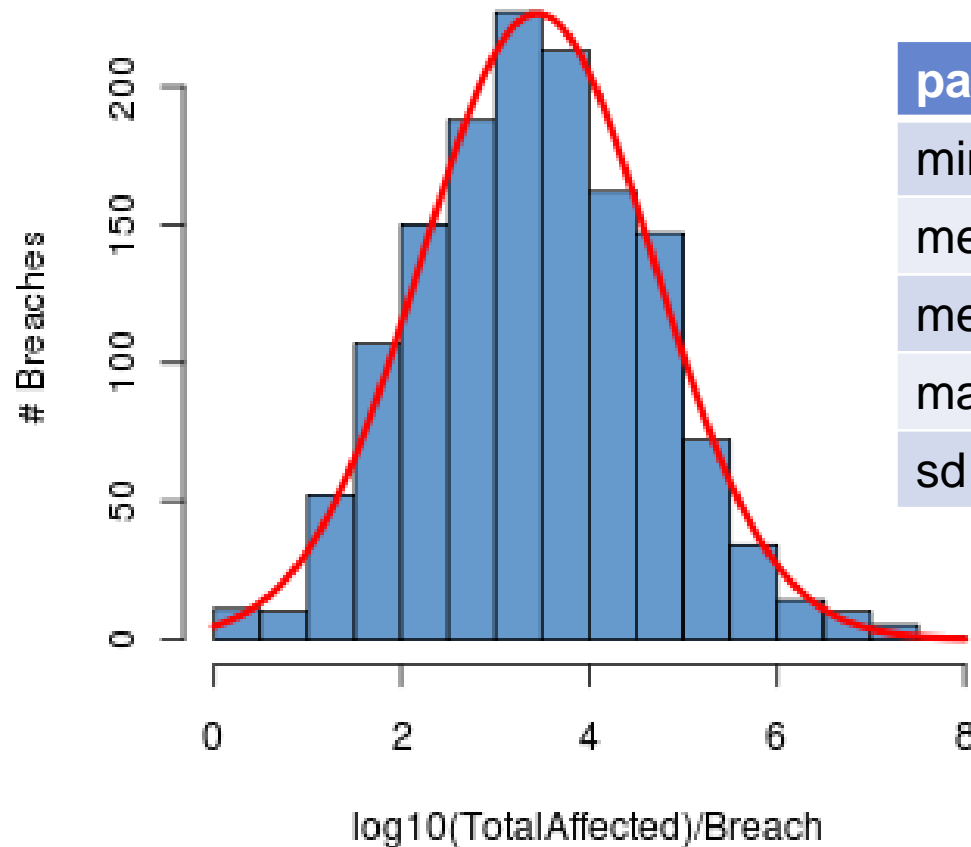
- ⊗ Frequency (Inside + Outside + Unk) is increasing
- ⊗ Frequency (Inside) is increasing
- ⊗ Frequency (Outside) is increasing
- ⊗ Frequency(Inside) > Frequency(Outside)

From the linear regression model:

Confidence > 80% that all of the above hypotheses are false.

Total Affected Frequency Histogram

DatalossDB: log10(TotalAffected) Histogram
ds.logTA: 1Q06 to Thu Aug 6 14:23:51 2009

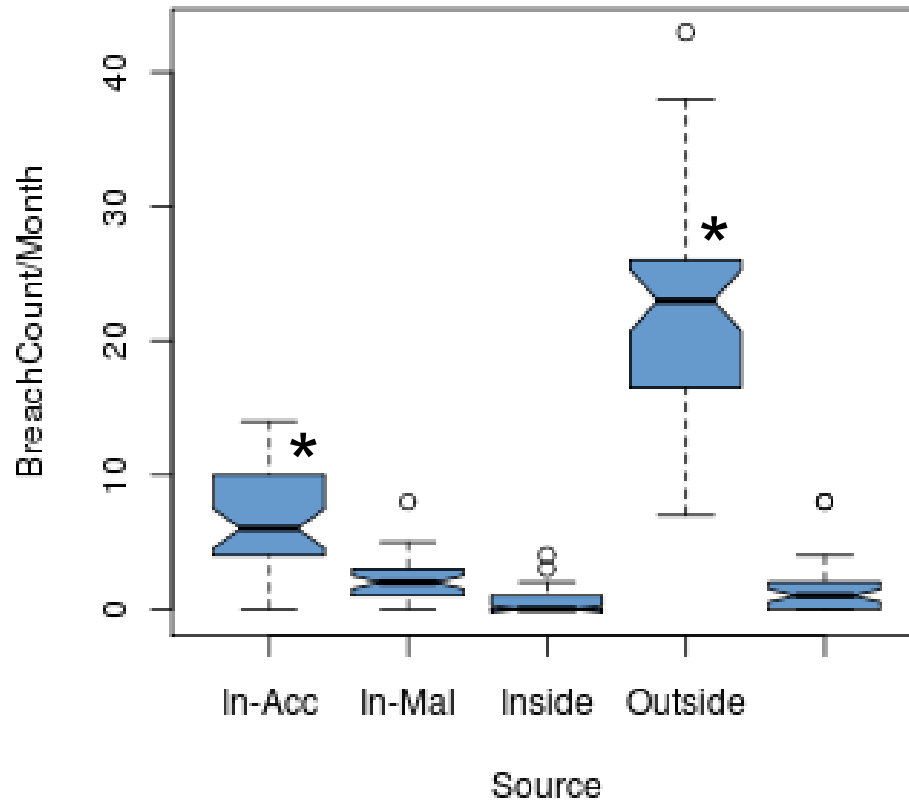


parm	TotalAffected(TA)	Log10(TA)
min	0	1
median	663	3.48
mean	185,800	3.51
max	94,000,000	7.97
sd	2,614,917	1.23

Frequency by Source

Whisker Plot

DatalossDB: BreachCount/Mon by Source
ds.logTA: 1Q06 to Thu Aug 6 14:29:02 2009

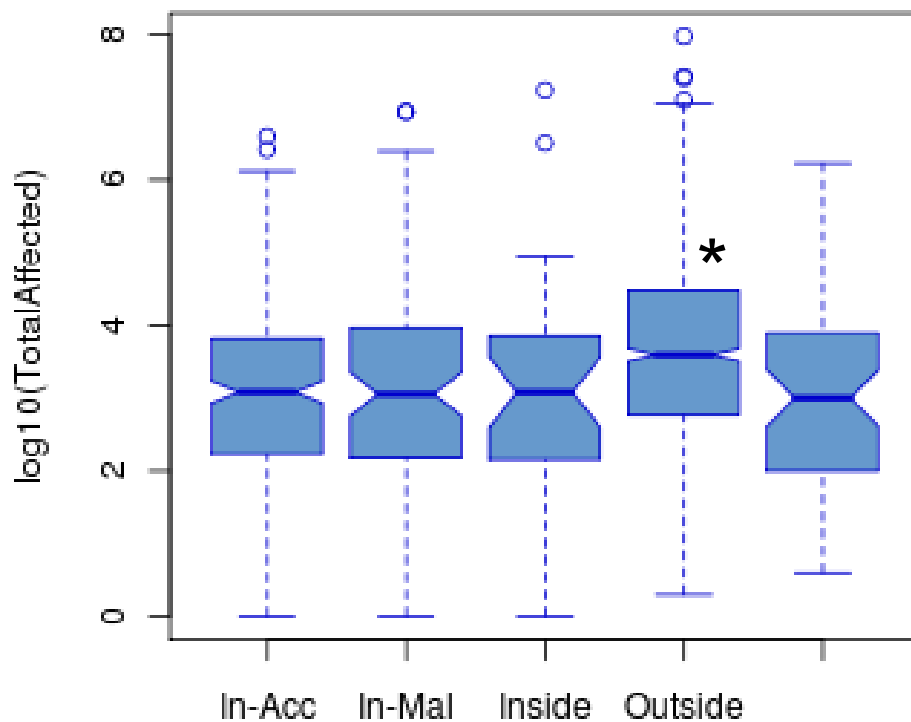


- * Statistically significant differences
- Outside breaches occur much more frequently than inside ones
- Inside-Accidental breaches occur much more frequently than Inside-Malicious ones.

Total Affected by Source

Whisker Plot

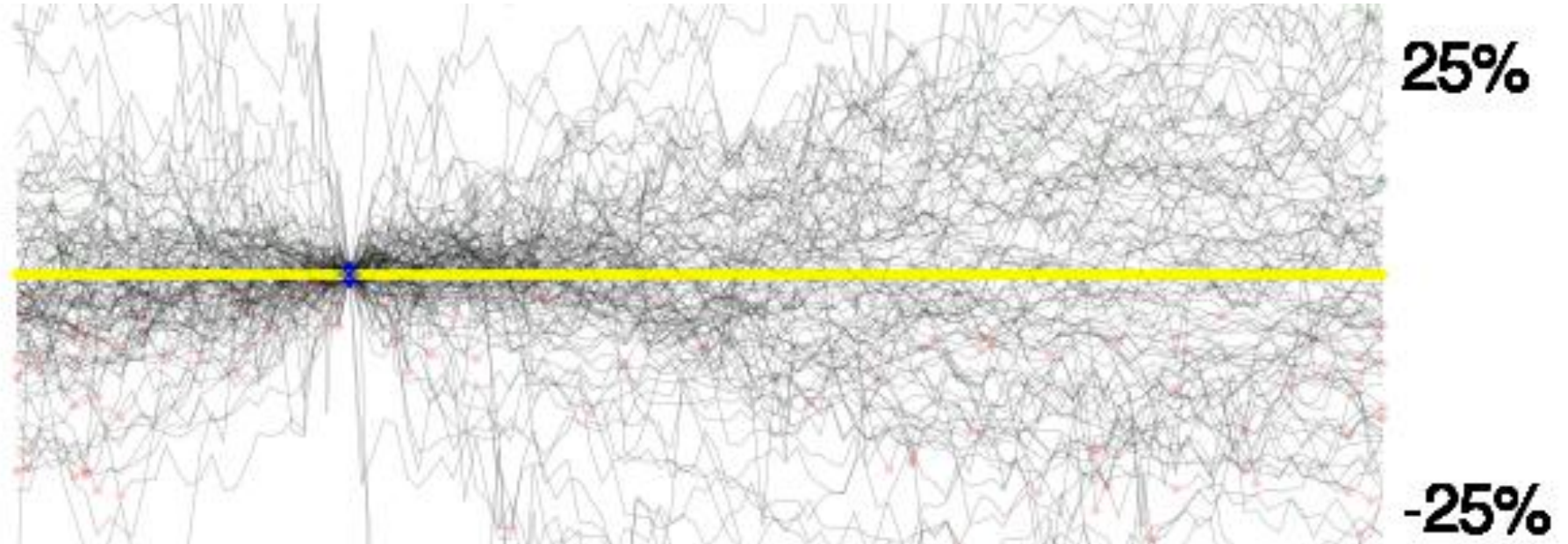
DatalossDB: BreachImpact/Mon by Source
ds.logTA: 1Q06 to Thu Aug 6 14:29:41 2009



- * Statistically significant differences
- Outside breaches affect more individuals than inside ones
- The number affected from Inside-Accidental, Inside-Malicious, and Unknown do not differ significantly

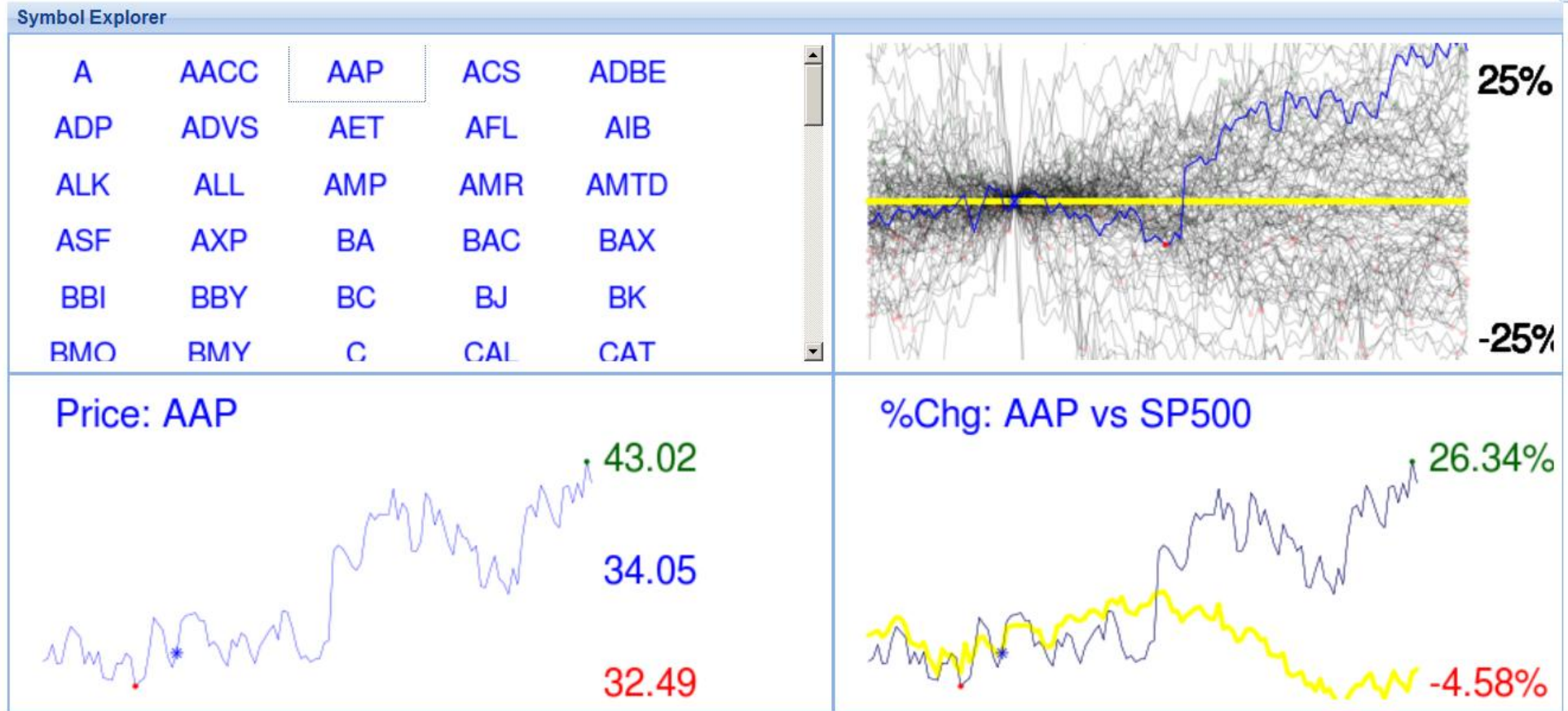
Stock Price Impact

Overlaid Sparklines: %ChgVsMkt



- 116 public companies with available, consistent ticker data
- Time Interval: [breachDate – 30 days, breachDate + 90 days]
- $s = \text{symbol}; \text{Price}(s) = \text{StockPrice}(s);$
- $\text{BreachPrice}(s) = \text{Price}(s) @ \text{BreachDate}$
- $\% \text{Chg}(s) = (\text{Price}(s) - \text{BreachPrice}(s)) / \text{BreachPrice}(s)$
- $\% \text{ChgVsMkt}(s) = \% \text{Chg}(s) - \% \text{Chg}(\text{SP500})$

Interactive Symbol Picker



- Selector in upper left cell triggers updates to other three cells
- Price is sourced from www.google.com/finance
- Proxy for SP500 is SPY

Stock Price Gradient Analysis

16 possibilities: 116 observations

Before (s) After(s) : Before(mkt) After(mkt)

++:-- ++:-+ ++:+- ++:++

+-:-- +-:-+ +:+- ++:++

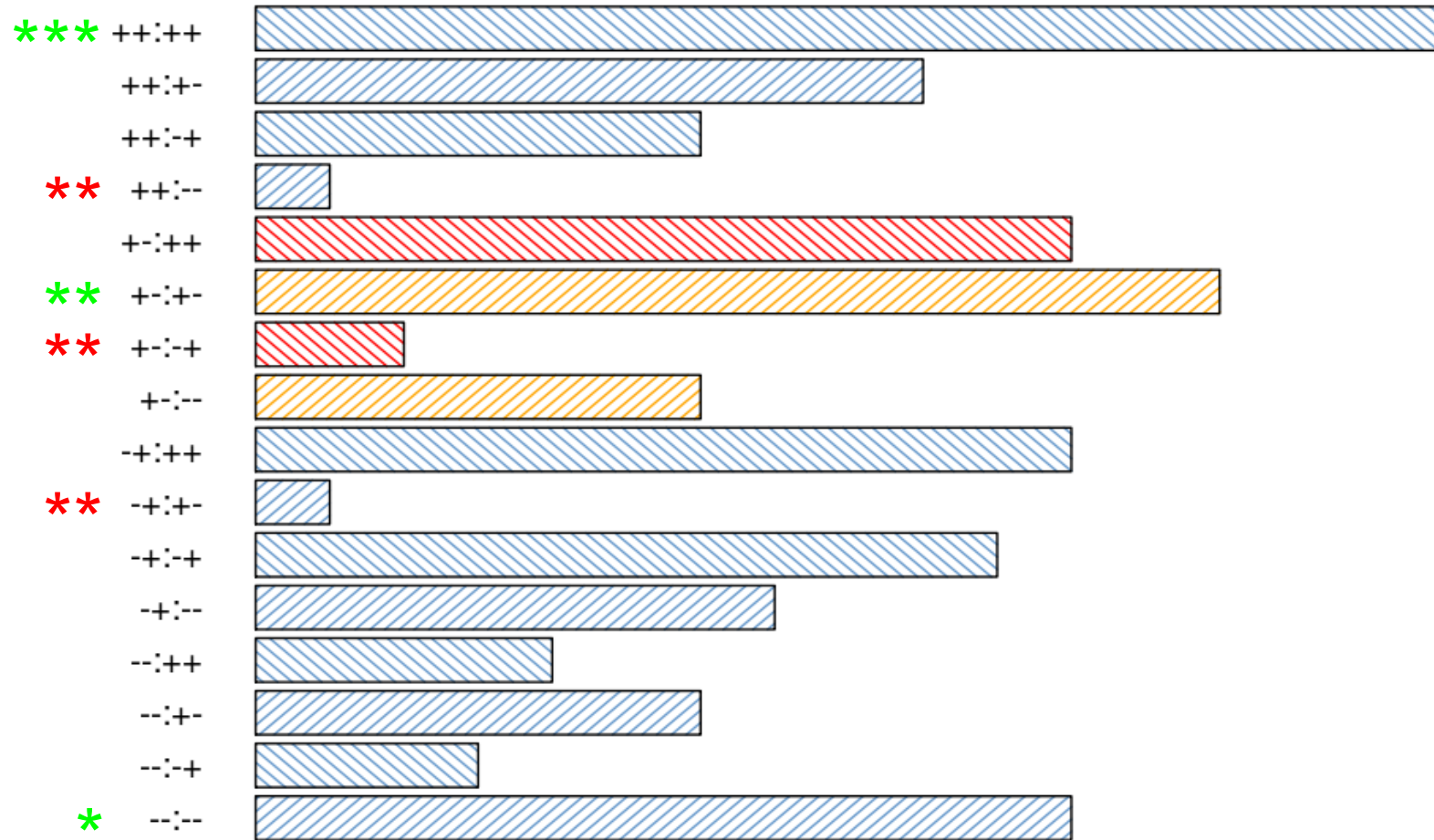
-+:-- -+:-+ -:+- -+:++

--:-- --:-+ --:+- --:++

Symbol stock price is increasing in the 30 day period before and decreasing in the 90 day period after the breach.

Stock Price Impact Analysis

Gradient Distribution Stock Price vs. Market



Statistically Significant High: * (>95%) ** (>99) *** (>99.5%)
 Statistically Significant Low: * (>95%) ** (>99) *** (>99.5%)

Conclusions: Company Context

- Is Breach Frequency Increasing?
 - Frequency (Inside + Outside + Unk) is increasing FALSE (80%)
 - Frequency (Inside) is increasing FALSE (80%)
 - Frequency (Outside) is increasing FALSE (80%)
- How do Breach Frequency and Total Affected compare for Inside vs Outside sources?
 - Frequency (Inside) > Frequency (Outside) FALSE (>90%)
 - TotalAffected(Inside) > TotalAffected(Outside) FALSE (>90%)
- How does a breach affect stock price?
 - StockPriceGradient changes sign from before(+) to after(-)
 - With market TRUE (>90%) with Market
 - Against market FALSE (>90%)

Conclusions: Ecosystem Context

- Public data can provide insight
- We are actively seeking additional data sets to analyze
- Community metrics can provide significant value via benchmarking and peer-to-peer comparisons
- Enterprise metrics programs can benefit
 - By incorporating insights from public data
 - By selective sharing of internal results

Appendix: Statistical Analysis

- 16 possible observations; 116 companies
- Hypotheses:
 - $H_0: \Pr(\text{observation}) \leq 1/16$
 - $H_1: \Pr(\text{observation}) > 1/16$
- $n = \# \text{ obs} = 116$; $p_0 = 1/16$, $q_0 = (1-p) = 15/16$; $npq = 6.9$
- $Npq > 5 \rightarrow$ Use normal distribution as an approximation to the binomial distribution for hypothesis testing
- $Z = \frac{p - p_0}{\sqrt{p_0 q_0 / n}}$ yields the # sd's from mean for each obs.

Appendix: p-values

Gradients	# Obs	p-value
--:--	11	.08
--:++	3	.95
--:+-	6	.69
--:++	4	.54
-+:--	7	.15
-+:--+	10	.08
-+:+-	1	.99
-+:++	11	.69

Gradients	# Obs	p-value
+-:--	6	.69
+-:--+	2	.98
+-:+-	13	.01
+-:++	11	.08
++:--	1	.99
++:--+	6	.69
++:+-	9	.26
++:++	16	.0005

The p-value indicates the probability of rejecting a true H0, namely that the probability of any of the above possible observations is 1/16 or less.

Elizabeth A. Nichols, Ph.D.
CTO for Metrics, PlexLogic
betsy.nichols@plexlogic.com
+1 (703) 963-7202

<http://www.metricscenter.org>