THE MINIMUM DAILY ADULT

The Right Metrics and The Wrong Metrics

CMG Canada

Denise P. Kalm

Director, Product Marketing

Denise.Kalm@ca.com

CMG 2008

Write a paper

Get a mentor

Deadline: Coming soon – June 13

Get out there!

"If you want to inspire confidence, give plenty of statistics. It does not matter that they should be accurate, or even intelligible, as long as there are enough of them."

Lewis Carroll

The Minimum Daily Adult

Agenda

- Introduction
- When a Rule of Thumb = ROT
  - The Tyranny of Average Response Time
  - Standard deviation – the watch-dog
  - Other Averages to Watch
  - Adding Rabbits and Tortoises
  - Avg(Avg) = Garbage
  - Percent of What?
  - The Tyranny of Correlation

Agenda - 2

- The Right Metrics
  - Consistency and σ
  - Percentages – the Right Way
  - Displaying Data
- Summary

Why Me?

Introduction

What metrics do you use?

- What do they mean?
- Are they useful to you (or anyone else)?
- Are they misleading?
- Were they just easy to get?

"Statistics are like women; mirrors of purest virtue and truth, or like whores to use as one pleases."

Theodor Billroth

ROT & Average Response Time

RT (Tran Code A) = 0.2

Actual response times: 0.1, 0.1, 0.3, 0.3, 0.2, 0.2, 0.2, 0.2

RT (Tran Code B) = 0.2

Actual response times: 0.1, 0.1, 0.1, 0.6, 0.2, 0.2, 0.2. 0.3

Are both equally good?

"Then there was the man who drowned crossing a stream with an average depth of six inches."

W. I. E. Gates

Median Can Be Better

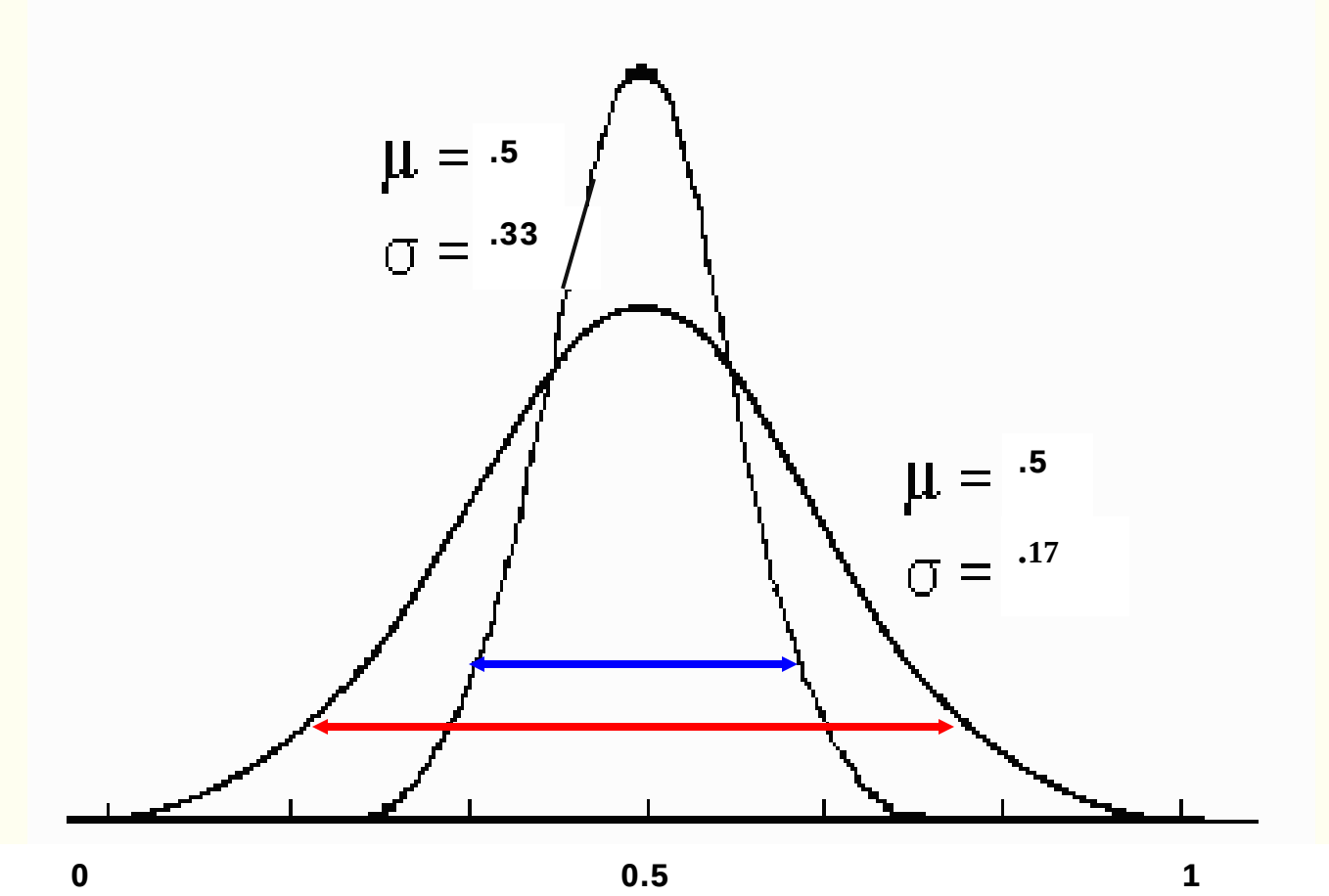
Median – that va
observation

RT: 0.1, 0

Mean = .22

Median = 0.1 (mo perience)

Measuring Consistency



2σ = 69% 3σ = 95% of data

Standard Deviation & Variance

Standard Deviation – square root of the variance. For normal data, 2/3 of the data points are within 1 SD of the mean on either side.

Variance – amount of "spread" of the data around the mean:

$S^2 = ((x_1-X)^2 + (x_2-X)^2 + \ldots (x_n-X)^2 \ldots /n-1$

Where x=mean [illegible] data point, n is the number of sa[illegible]s

Standard Deviation of a Sample

If the SD is large, you need to inspect your sampling method. This may indicate:

- Suspect data
- Poor interval choices
- Mixed workload
- Problems…

"I abhor averages. I like the individual case. A man may have six meals one day and none the next, making an average of three meals per day, but that is not a good way to live."

Louis D. Brandeis

It isn’t just response time

Any averages can present problems:

Avg CPU busy or CPU/tran can be problematic if variance is high.

- Track variance
- Understand outliers
- Apply knowledge to capacity planning for “worst case” scenarios

"The average human has one breast and one testicle."

Des McHale

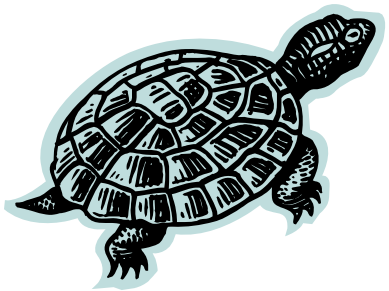
Adding Rabbits & Tortoises Together

What work can be combined (and what cannot)?

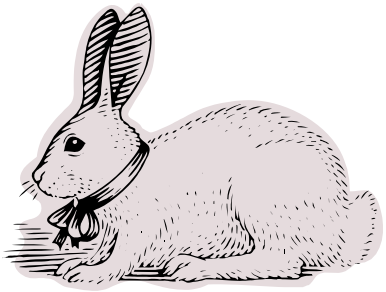
Do not combine:

- Background trans with foreground
- Response time critical work with lower criticality
- Daemons with anything else

Understand your data!



+



= ?

"Not everything that can be counted counts; and not everything that counts can be counted."

George Gallup

Averages of Averages

- Most collected data is already averaged
- When you then average across a larger interval, you average averages
- Avg(Avg) smoothes out variability

8	9	10	11	12	13	14	15	16
0.4	0.2	0.2	0.2	.1	1.0	.3	.3	.2

Avg=0.32

Percent of What?

Breast Cancer Risk Reduction

Women who exercise can reduce their risk of breast cancer by 18%!

The Rest of the Story

If lifetime risk is only 3%, a 18% reduction in risk only improves your risk down to 2.5%

What is the baseline data?

Percent of What?

HRT Scare Story

Women on HRT for 15 years increased their risk of cancer by 48%.

The Rest of the Story

If lifetime risk is only 1%, a 48% increased risk only raises your risk to 1.48%

What is the baseline data?

In Capacity Planning

Reduce CPU busy of transaction by 50%

- How much was it using before?
- How many of them are there?
- How much spare capacity do you have?
- What do those CPU seconds cost?

If CPU/tran was 0.01 sec and you only run 1000/day, you have now saved 5 seconds of CPU

Raw Percentage

What does 50% CPU busy really mean?

- How many engines?
- Is it 50% of one or of total capacity?
- How busy can the system run?
- What is the practical capacity?

"Say you were standing with one foot in the oven and one foot in an ice bucket. According to the percentage people, you should be perfectly comfortable."

Bobby Bragan

"USA Today has come out with a new survey – apparently, three out of every four people make up 75% of the population."

David Letterman

Correlation

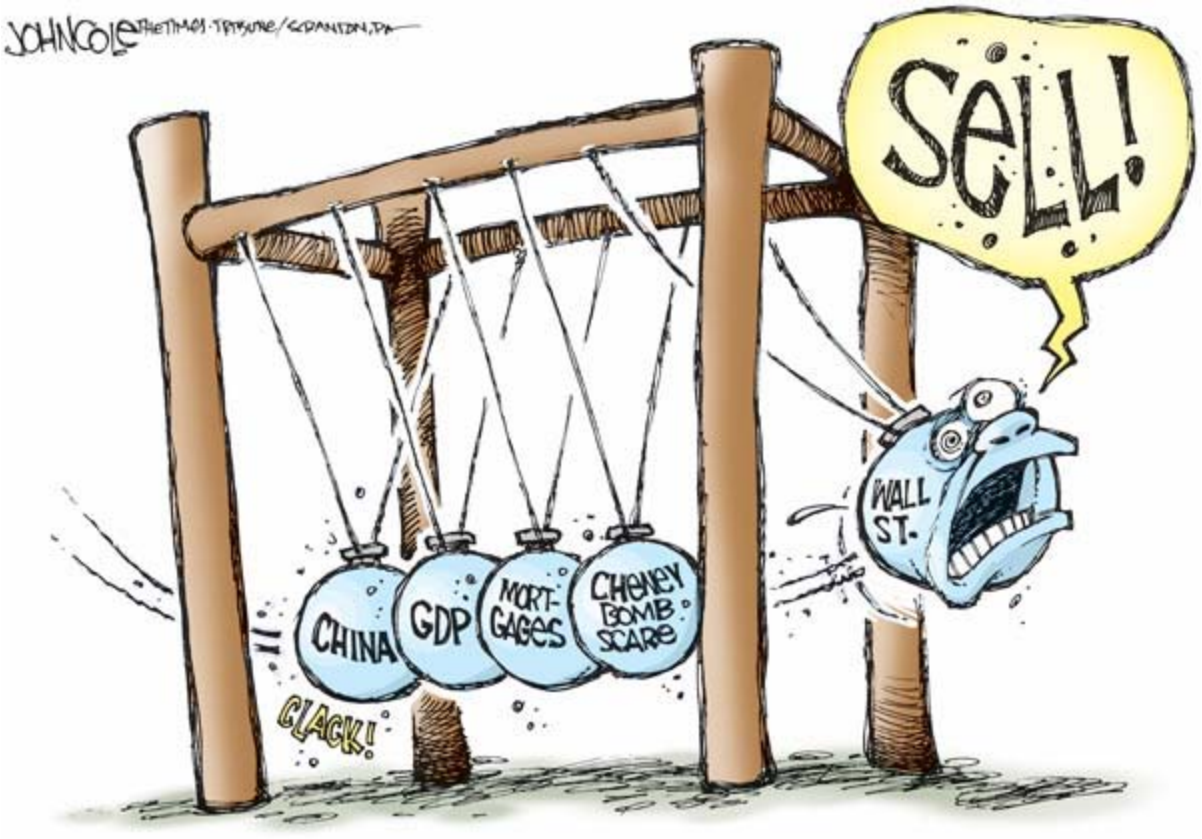
Correlation coefficient - ($R^{2)}$ measures the degree of relationship(and direction) between two variables.

R^2 =1.00 indicates a perfect correlation

R^2 = 0.0 means there is no relationship at all.

R^2 = a negative number means that as one variable increases, the other decreases.

It is NOT cause and effect



The Stork Correlation

In a small Welsh town, there was a .95 correlation between the arrival of storks and the arrival of babies.

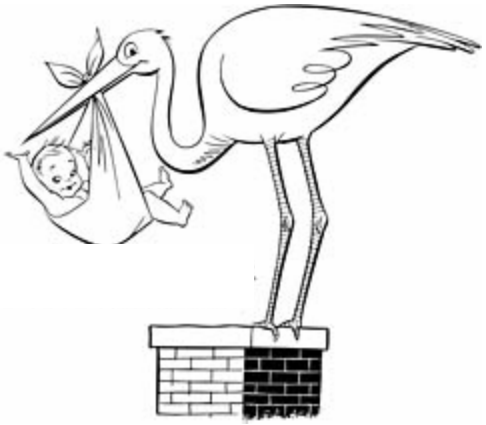
Why?

The Stork Correlation

There was also a 1.0 correlation between the dates fishermen were home from the sea and the likely dates of conception.

Key Points to Remember

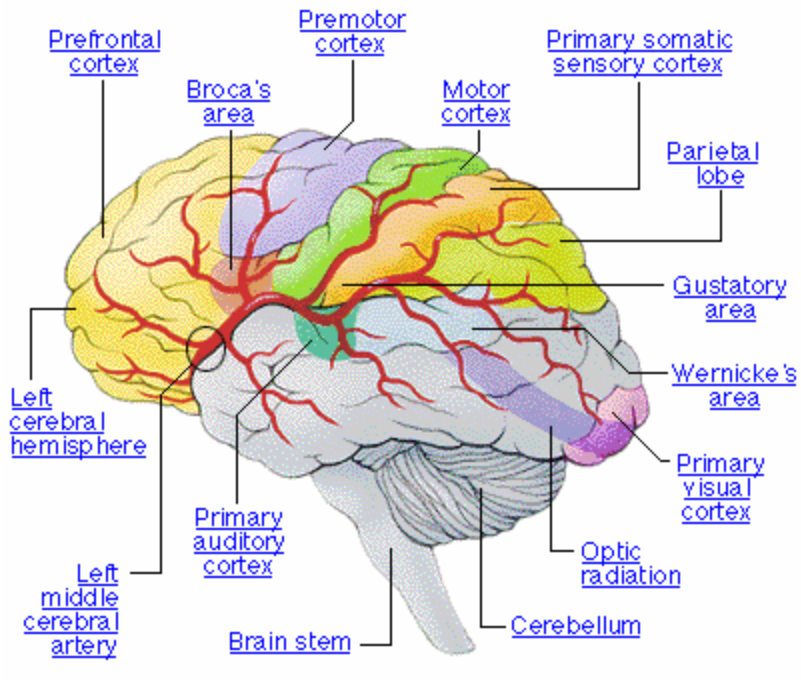
- Correlation is NOT cause and effect.
- Though there may be a causal relationship between two variables, you cannot infer it from a correlation analysis.
- A third factor may really be causing the correlation.



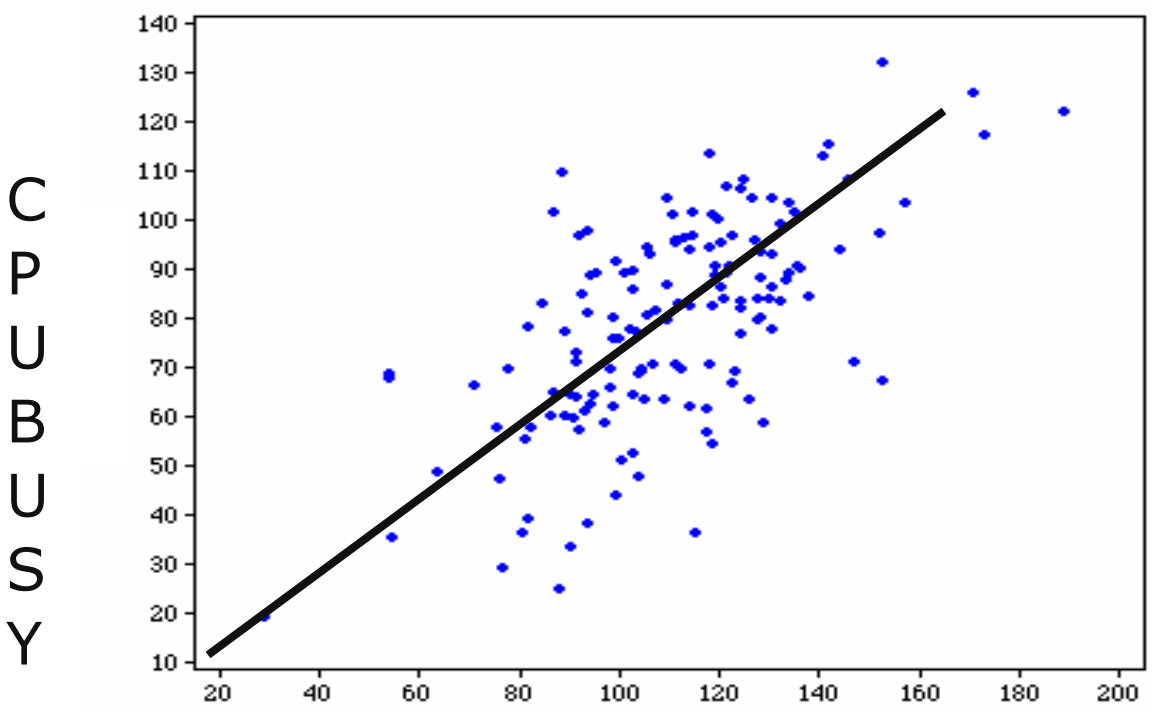
Key Points to Remember

- Don’t calculate it by hand – use a tool.

- Use your brain to interpret the results.



"A statistician is someone who is skilled at drawing a precise line from an unwarranted assumption to a foregone conclusion."



The Right Metrics

- Know what you are measuring – what is a business UOW?
- How much of a user-experienced response time can you measure?
- Understand your data
- Always ask – relative to what?
- Which data matters? Who cares?

Consistency

- Standard deviation (SAS, Excel…)
- Min and Max
- What is Six Sigma

"Consistency may be the hobgoblin of little minds, but it makes you very popular with your users."

Denise Kalm

Qualifying Percentages

- How many engines?
- How many can the workload use? (CICS)
- What is the usable threshold?
- What was the baseline value before the change?
- How much did the change actually matter?
- Who benefits?

60% of a 3-way processor, 50% of a device where 60% busy is the practical threshold, merger volumes increased CPU 50% from 48% to 72%; no issues expected.

Displaying Data

Performance Report

ABC Bank – 10/18/07

WORKLOAD	#/day	95%tile RT	95%tile at 10AM	95%tile last week
Credit Card Autho	10,000	0.5	0.5	0.5
Online Pay	15,000	0.4	0.7	0.4
ATM Trans	20,000	0.3	0.3	0.4

The Cost of Doing Business

Chargeback Report
ABC Bank
10/18/07

Workload	#/day	CPU per Unit	Cost/CPU sec	Total Cost
Credit Card Autho	10,000	0.05	0.05	$25.00
Online Bill Pay	15,000	0.1	0.05	$75.00
ATM Transactions	20,000	0.1	0.05	$100.00

CPU Busy? When Do I Care?

Processor	Type	#Engines	Usable %	Peak Busy Over all Engines	Queue length	Peak Busy on Busiest Engine
MainframeA	2084-102	2	100	98	1	100
SunboxB	E20K	30	55	40	6	65
AIXboxC	p5-570	4	60	45	2	50

When to be Extra Careful

> Virtualized servers

> Multi-core engines

> zIIPs and zAAPs and IFLs

What Does CPU Busy Mean?

"Without data, you are just another person with an opinion."

Unknown

"Do not put faith in what statistics say until you have carefully considered what they do not say."

William W. Watt

Questions?

"Statistics can be made to prove anything…even the truth."

Unknown

Summary

What are the wrong statistics for?

- To inspire national panic
- To mislead voters
- To gain unearned glory

Be proactive, notice trends earlier and develop a reputation for integrity and openness.

For More Information

Contact me at: Denise.Kalm@ca.com