Development and Test Metrics 101

@catswetel | @everydaykanban
Software Development Performance Index (SDPI)
Larry Maccherone
QUALITY
RESPONSIVENESS
PRODUCTIVITY
PREDICTABILITY
What’s missing?
“The RIGHTER we do the WRONG thing, the WRONGER we become.”

-- Dr Russell Ackoff
Time in Process
Units of time per unit of work
When can we expect X?

90%

50%
How likely by X?

30d: 77%  21d: 50%
What’s the story?
QUALITY

RESPONSIVENESS

PRODUCTIVITY

PREDICTABILITY
This one happens more often

Groups are easier to see
WEIBULL DISTRIBUTION

@catswetel | @everydaykanban
WEIBULL DISTRIBUTION

...maybe
NOT Normal
Notice anything else?
Notice anything else?
90%
MULTI MODAL

@catswetel | @everydaykanban
QUALITY

RESPONSIVENESS

PRODUCTIVITY

PREDICTABILITY
What else can we learn?
THROUGHPUT

Units of work per unit of time

TIME IN PROCESS

DATE DELIVERED

JAN 2015
FEB 2015
MAR 2015
APR 2015
MAY 2015
JUN 2015
JUL 2015
AUG 2015
SEP 2015
OCT 2015
Boxes show variability of Time in Process, NOT Throughput
Does it always make sense to count all the dots?
THROUGHPUT
“We are forced to consider knowledge as something that changes as new evidence is provided by more data, or as new predictions are made from the same data by new theories.”
Lagging quality

@catswetel | @everydaykanban
ESCAPED DEFECTS
defects found in production

<table>
<thead>
<tr>
<th>Month</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feb</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mar</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apr</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>May</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jul</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aug</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sep</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

@catswetel | @everydaykanban
ESCAPED DEFECTS
defects found in production

How can we tell if we’re getting better at finding defects earlier?
DEFECTS
BY ENV
percentage

Key:

DEV
INT
STAGE
PROD

Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep

@catswetel | @everydaykanban
Predicting the movements of the **one bee** is nearly impossible without the context of the **swarm of bees**.

*Complex Adaptive Systems: An Introduction to Computational Models of Social Life*  
by John H. Miller, Scott E. Page
MORE FINISHED THAN STARTED

MORE STARTED THAN FINISHED

@catswetel | @everydaykanban
LITTLE’S LAW

Average # of items in a system = Average arrival rate * Average time spent in the system

Average time spent in a system = Average # of items in a system / Average throughput

@catswetel | @everydaykanban
It’s ok to make necessary trade-offs
Where should we focus improvement efforts?

@catswetel | @everydaykanban
FLOW EFFICIENCY

\[ \frac{\text{work}}{\text{wait} + \text{work}} \times 100\% = \text{flow efficiency} \]
...per state (incl. waiting states)
You can do any chart for any portion of the process!
Are we there yet?

@catswetel | @everydaykanban
#NOESTIMATES
“Close examination reveals that every meaningful interpretation involves a prediction.”

Statistical Method from the Viewpoint of Quality Control
Walter A Shewhart
3 Components of Knowledge:

1. Evidence
2. Prediction
3. Degree of Belief
<table>
<thead>
<tr>
<th>Category</th>
<th>Title</th>
<th>Day</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>EVIDENCE</td>
<td>Agile Quantified (Measuring the impact of Agility)</td>
<td>Wed</td>
<td>3:45</td>
</tr>
<tr>
<td></td>
<td>Larry Maccherone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREDICTION</td>
<td>Forecasting and Estimation Hands-On</td>
<td>Thu</td>
<td>10:45</td>
</tr>
<tr>
<td></td>
<td>Ozzie Yuce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEGREE OF BELIEF</td>
<td>Data Driven Coaching</td>
<td>Thu</td>
<td>9:00</td>
</tr>
<tr>
<td></td>
<td>Julia Wester &amp; Cheryl Hammond</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cat Swetel
Engineering Manager
Ticketmaster
cat@catswetel.com
@catswetel

Julia Wester
Co-Founder & Principal Consultant
Lagom Solutions
julia@lagom.solutions
@everydaykanban
“Computers are useless. They can only give you answers.”

Pablo Picasso
“STATISTICS are useless. They can only give you answers.”
LITTLE’S LAW

\[
\text{Average # of items in a system} = \text{Average arrival rate} \times \text{Average time spent in the system}
\]

Are your decisions making this more or less true?

\[
\text{Average time spent in a system} = \text{Average # of items in a system} / \text{Average throughput}
\]
Some joke about defining quality
TIME IN PROCESS

DATE DELIVERED

VARIABILITY

FEEL FREE TO CUT
SHEWHART: the cause of variability in an expected attribute. Requires JUDGEMENT to know which “important quality characteristics” (causes of undesirable variability) within the bounds of reasonable expected economic results.
In the 1970s, Deming's philosophy was summarized by some of his Japanese proponents with the following "a"-versus-"b" comparison:

(a) When people and organizations focus primarily on quality, defined by the following ratio,

\[
\text{Quality} = \frac{\text{Results of work efforts}}{\text{Total costs}}
\]

Quality tends to increase and costs fall over time.

(b) However, when people and organizations focus primarily on costs, costs tend to rise and quality declines over time.
CUMULATIVE FLOW DIAGRAM

AVERAGE ARRIVAL RATE
CUMULATIVE FLOW DIAGRAM

@catswetel at qconnyc
CUMULATIVE FLOW DIAGRAM

AVERAGE TIME IN PROCESS

@catswetel at qconnyc