



Escaped defects

Time to feedback

Maintenance share

Flake rate

Baseline

DORA metrics

Test count

Pass rate

**J-curve of value
realization**

MAINTENANCE SHARE

The proportion of automation effort spent repairing existing tests rather than extending coverage.

Why it matters: The number that decides whether a program compounds or stalls - and the claim vendors make most often, with no standard definition behind it.

TIME TO FEEDBACK

How long after a change lands a developer learns whether it broke something.

Why it matters: A comprehensive but slow suite gets bypassed. This often predicts real quality better than coverage does.

ESCAPED DEFECTS

Defects that reached production despite a test activity meant to catch them.

Why it matters: The outcome measure. Everything else on this card deck is a leading indicator of this one.

DORA METRICS

Change lead time, deployment frequency, failed deployment recovery time, change fail rate and deployment rework rate.

Why it matters: DORA has retired MTTR, which expanded three different ways in three different standards. Correct this in any metrics deck you inherit.

BASELINE

What was true before you changed anything, written down, with the date on it.

Why it matters: Half an hour in week one. Without it you will improve things and be unable to prove it - and this is the step teams skip.

FLAKE RATE

The share of results that are non-deterministic for an unchanged version.

Why it matters: Above roughly 2%, people stop reading red. At that point the suite has stopped working regardless of what it covers.

J-CURVE OF VALUE REALIZATION

The dip before the gain: learning curve, verification tax, and adjusting downstream process to higher code volume.

Why it matters: Useful for setting expectations with a sponsor. The dip is the tuition, and it is predictable rather than a sign of failure.

PASS RATE

The share of executed tests that passed.

Why it matters: Confounded by the thing it is used to hide. With automatic retries it measures your retry policy as much as your product.

TEST COUNT

How many tests you have.

Why it matters: Not a result. It is the easiest number to move without improving anything, and agents make it trivially easy. Report what is proven, not what was written.



Hours displaced

Critical journey

Coverage of what matters

**COVERAGE OF WHAT MATTERS**

The share of critical journeys with a reliable automated check.

Why it matters: The question to ask of any suite: if every test passes, what do we now know is true?

CRITICAL JOURNEY

A path a customer takes to do the thing you charge them for.

Why it matters: Usually ten to twenty. If your list has sixty, you have listed features. Test it: if this broke at 2am Friday and nobody noticed until Monday, what would it cost?

HOURS DISPLACED

Manual test execution hours no longer being spent.

Why it matters: The number a finance conversation actually turns on. Capture it per release, not per month, so it survives a schedule change.