A Picture is Worth a Thousand Words

Metrics With a Punch!

Beth Wisdom

Sr. QA Manager

Ulta Beauty

May 15, 2019

Agenda

- Software Quality Metrics – The Definition
- Leveraging Metrics
- Ulta Beauty Examples
- Getting Started
- Q&A

Objectives for Today

- ✓ Determine what metrics make sense for your challenges
- ✓ See real life examples of simple metrics that are easy to collect and can spur action and change perceptions
- ✓ Use data and metrics to show ROI in a meaningful way

Software Quality Metrics - Definition

There are **3** categories of software quality metrics:

- **Product**, which describe the characteristics of a product (e.g. size, complexity, performance, quality level)
- **Process**, which are leveraged to optimize/improve the software development and maintenance process (e.g. effectiveness of defect removal, defect opening/closing trends, defect aging)
- **Project**, which describe project characteristics and execution such as resources, cost, schedule, and productivity

Challenges - Feedback - Perception

What challenges have you experienced where you think metrics may help? What feedback or perception do you want to address?

1.

2.

3.

4.

Sound Familiar?

- “Testing Costs Too Much”
- “Too Much Time is Spent on Testing”

➡ ***QA Process is Inefficient!***

- “Initial Test Estimates Don’t Match Actuals”

➡ ***QA Estimation Process is Flawed***

- “QA Must Get ‘X’ Going – But There’s **No Money** for Tools/Resources”

Example: X=Automation

Real Examples

Example #1: Cost of Testing

Initial Assumptions:

- # test cycles
- # rounds of regression
- Defect retest intervals – based on historical trends

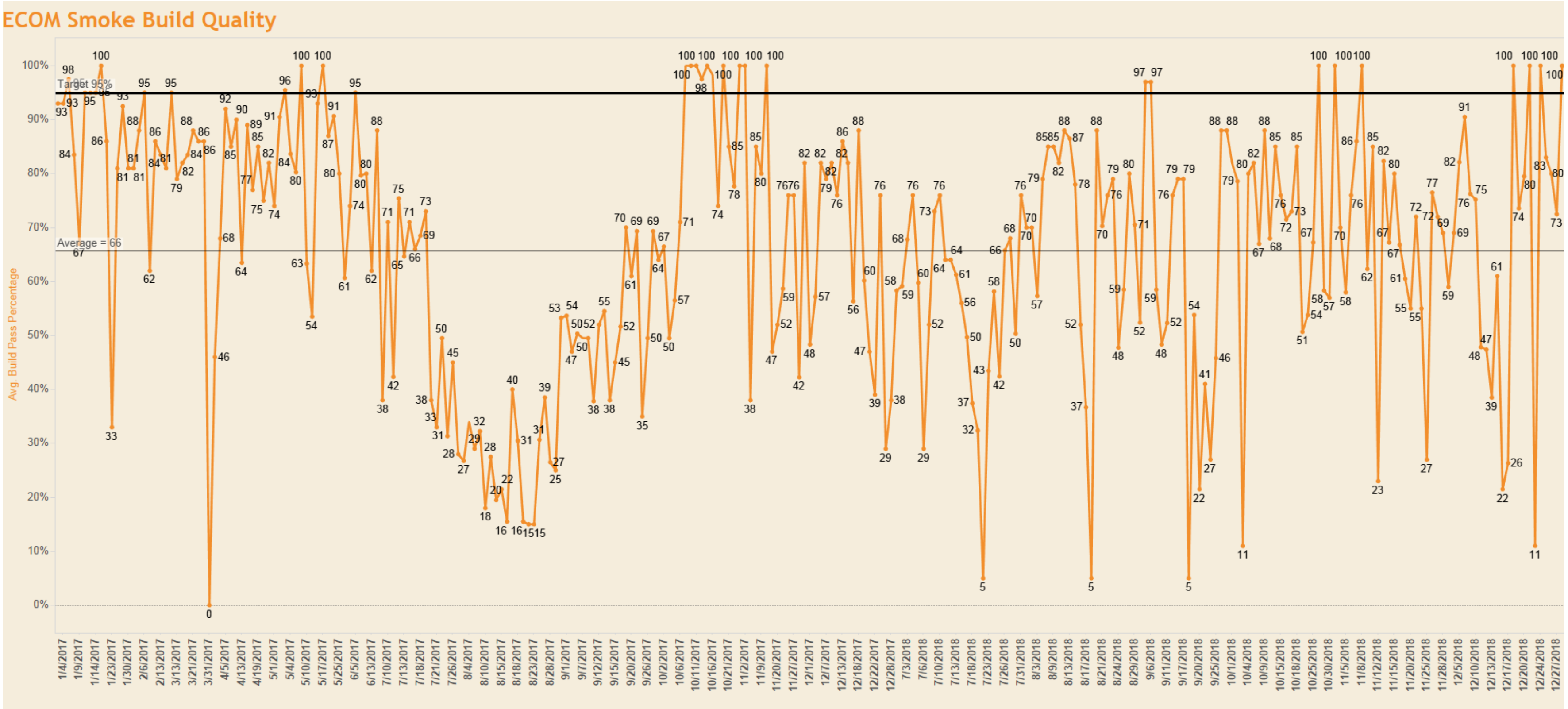
Keep a Change Log

- Log the events that change /affect the estimated effort
- Categorize the events and plot the trend over time, for example:
  - ○ Scope change
  - ○ Code quality/defects
  - ○ Environment Issue
  - ○ Late delivery to QA

Date	Event	Impact	Impact Category (scope/deliverable delay/env/quality/other)	Comments
8/23/18	CR added per business	Additional test cycles required	Scope	
9/1/2018	QA environment unavailable	Delay in completing test cycle 1	Env	Had to make up the time over the weekend to recover to keep end date.
9/4/2018	Defect fixes delayed to QA	Extension of resources and timeline to accommodate testing of remaining fixes	Deliverable delay	

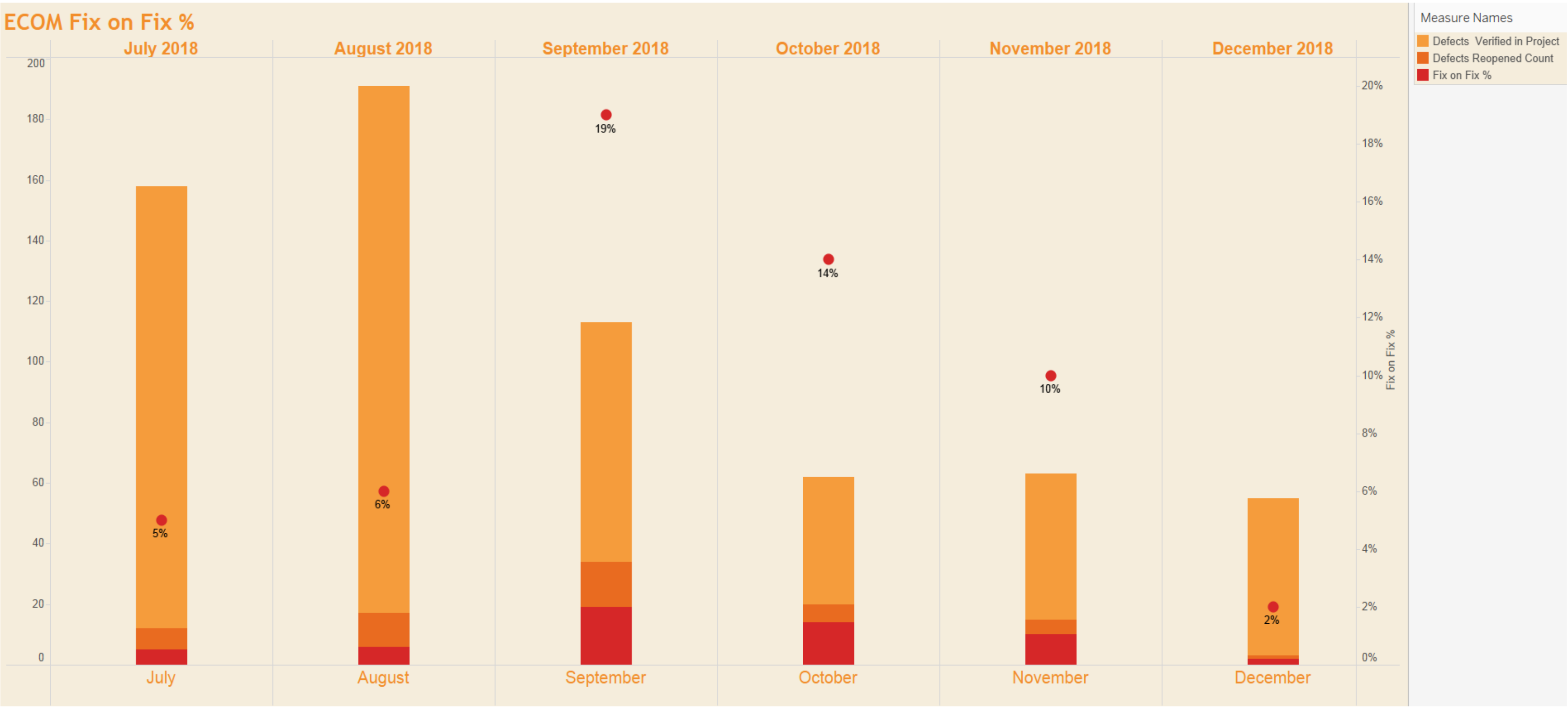
Capture Pass Rates – Of Everything

- Capture pass rates of all deliverables (fixes, features, etc.)
- Plot each pass rate and capture the trend over time
- Per project and/or per application



Capture Fix-on-Fix Rate

- If a defect fix fails retesting, count it
- Plot the rate as well to show how often the rework has occurred



Capture Cost of Defects

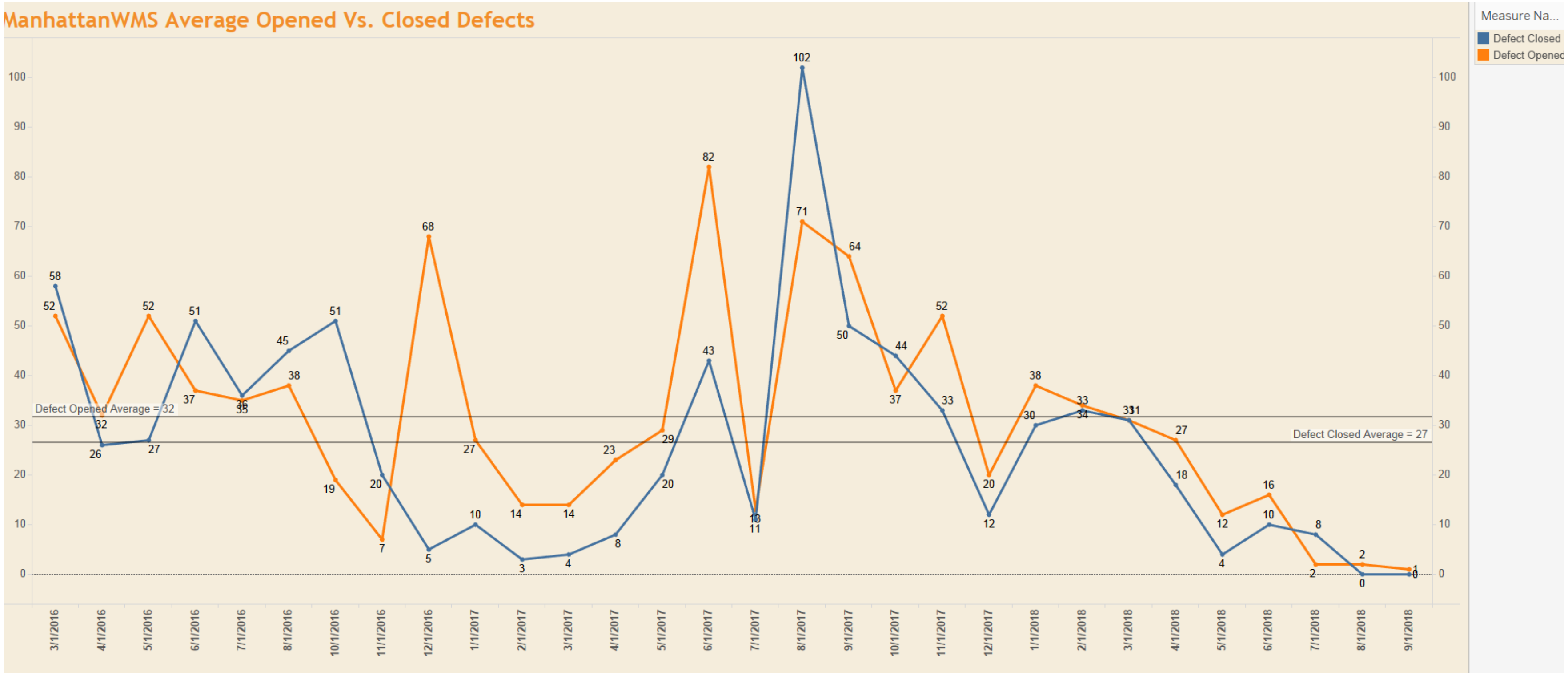
- Collect hours spent in development and QA on each defect
- Multiply totals by average hourly rate
- Show how rework hits the budget with actual $$$

Set (and Honor!) Acceptance Criteria

- Have a build stability test suite (BST) – critical business scenarios
- Execute with each code drop; leverage automation!
- Set acceptance pass rate % - do not proceed with testing unless it's met

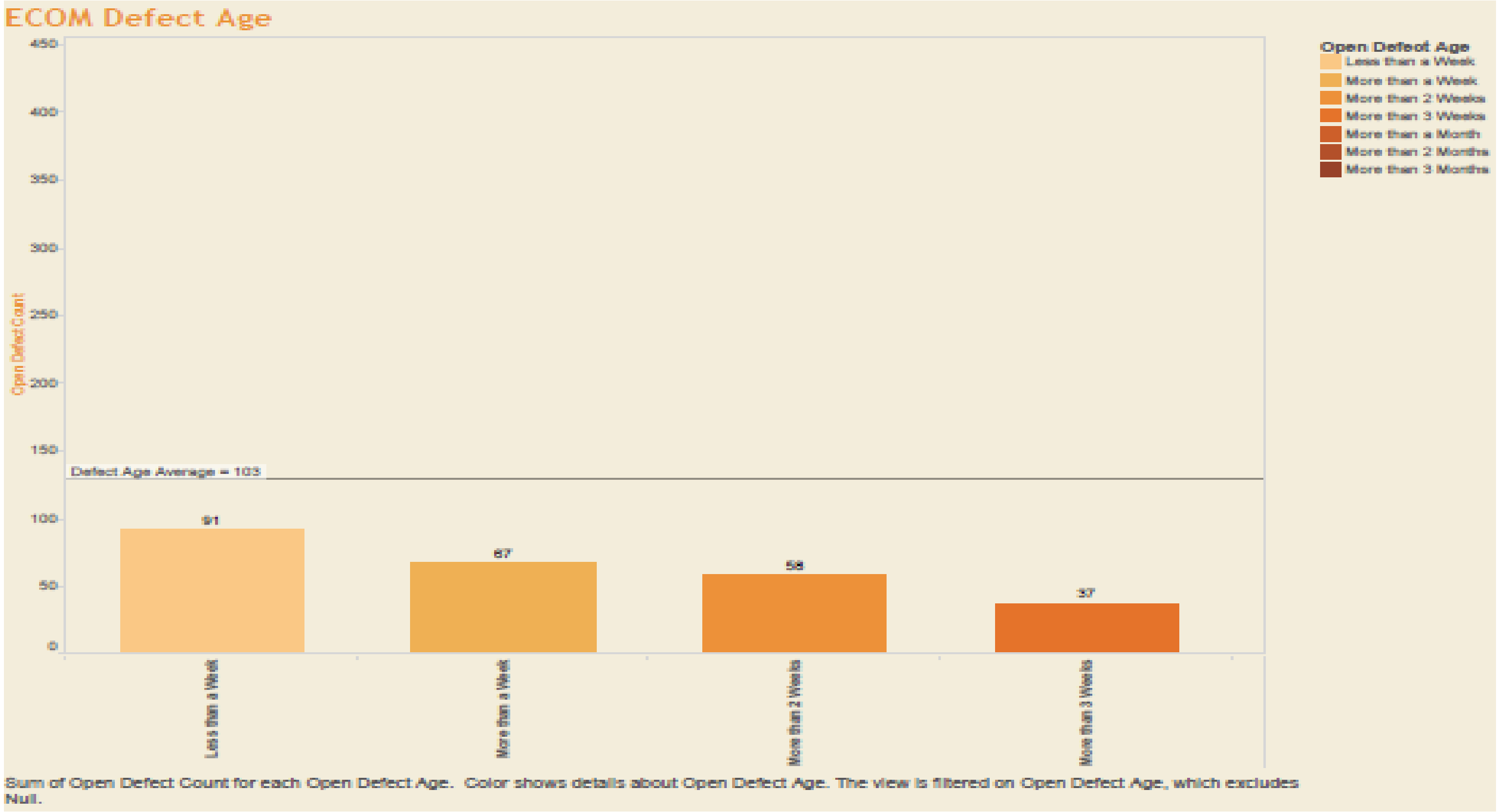
Defect Opening/Closing Trends

- Capture defect opening and closing rates; plot trend over time
- If opening rate > closing rate, it affects QA cost:
  - ○ extends QA interval waiting on fixes
  - ○ extra test cycles required



Defect Aging & Backlog Growth Trends

- Capture defect aging and backlog growth trends
- As with opening/closing trends, it could shed light on QA



Sum of Open Defect Count for each Open Defect Age. Color shows details about Open Defect Age. The view is filtered on Open Defect Age, which excludes Null.

Example #2: Planned vs Actual Test Effort

- Take Change Log in Example 1 and plot/bar graph #incidents per type, pictorially
- Have conversations based on incidents/events and start talking about how to prevent or minimize those in the future
- Continue to monitor and provide trend graph to confirm your organization is effectively controlling them through effective measures

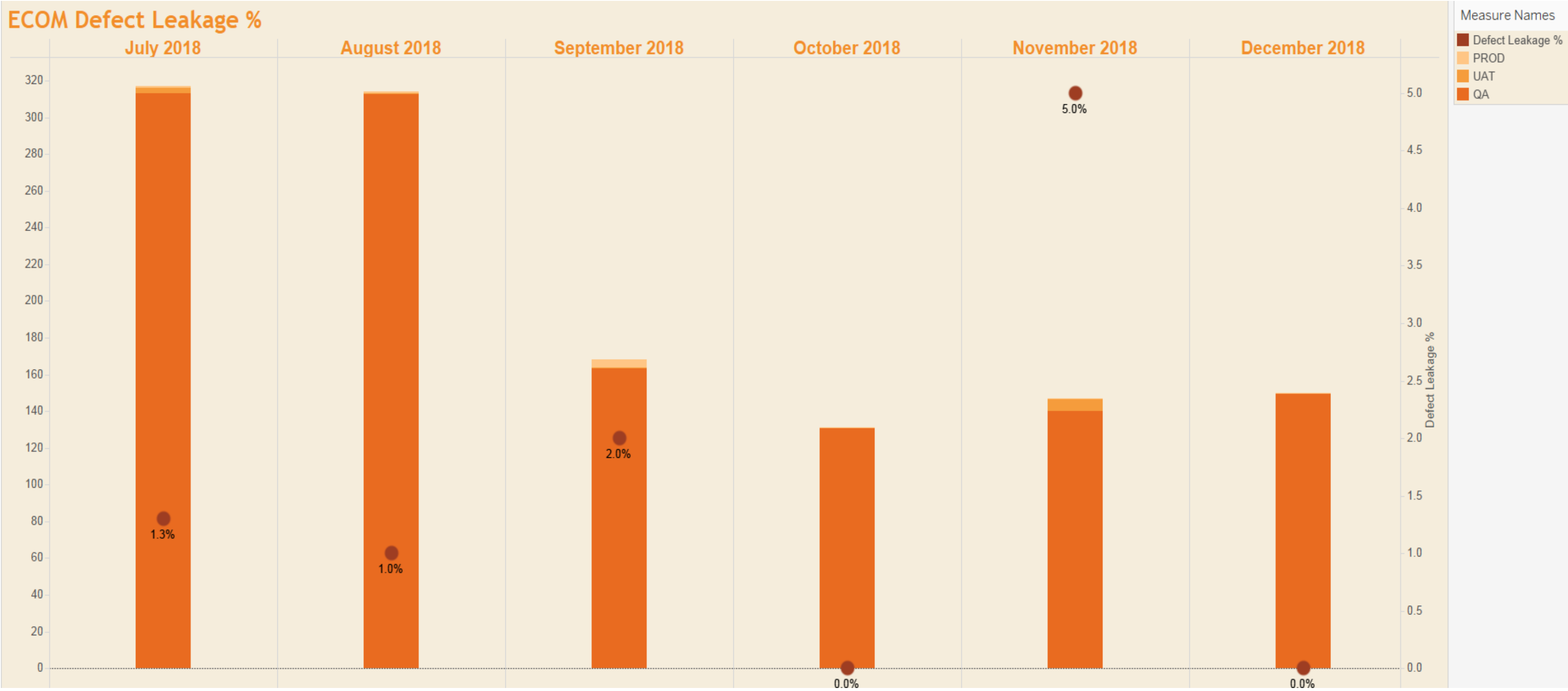
Example #3: QA Effectiveness

- Daily project metrics: % complete, pass rate, open defect count, etc. provides only a partial picture
- Collect # defects reported after QA phase (UAT, hyper-care, post production) and perform root cause analysis

Calculate Defect Removal Effectiveness (DRE):

$$DRE = \frac{\text{Defects removed during a development phase}}{\text{Defects latent in the product}} \times 100\%$$

Defect Removal Effectiveness



Leakage = (# defects found in QA cycles)/(# defects found in QA/UAT/Prod) – 1

Example #4: Capturing ROI for Cost Neutral Solutions

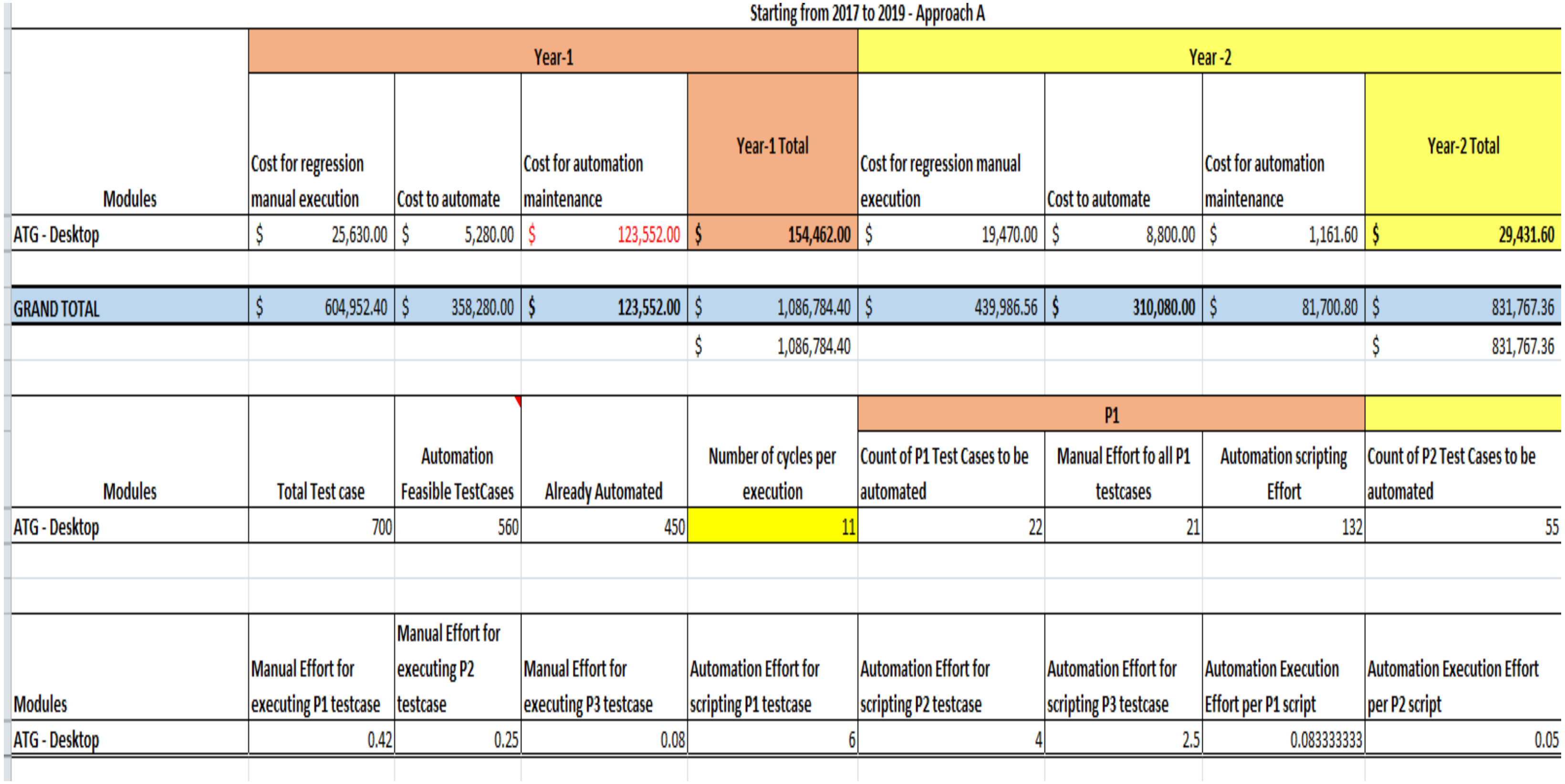
Challenge: Accomplish "X", but there's little/no budget for the resources and tools

Approach: Capture the effort savings and compare against the cost/investment

The point at which the savings > investment = ROI

Calculating ROI – Automation Example

- Calculate the costs based #test cases, which will be automated year 1, cost of required tools, #cycles expected/year, cost of automating each test case:



Starting from 2017 to 2019 - Approach A

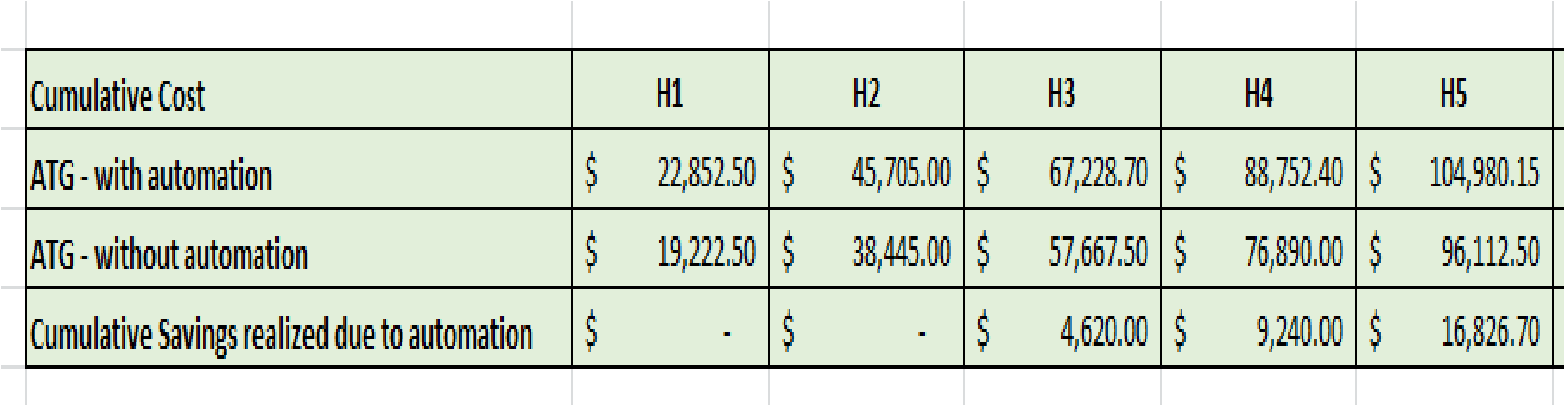
	Year-1				Year -2			
Modules	Cost for regression manual execution	Cost to automate	Cost for automation maintenance	Year-1 Total	Cost for regression manual execution	Cost to automate	Cost for automation maintenance	Year-2 Total
ATG - Desktop	$ 25,630.00	$ 5,280.00	$ 123,552.00	$ 154,462.00	$ 19,470.00	$ 8,800.00	$ 1,161.60	$ 29,431.60
GRAND TOTAL	$ 604,952.40	$ 358,280.00	$ 123,552.00	$ 1,086,784.40	$ 439,986.56	$ 310,080.00	$ 81,700.80	$ 831,767.36
				$ 1,086,784.40				$ 831,767.36

					P1			
Modules	Total Test case	Automation Feasible TestCases	Already Automated	Number of cycles per execution	Count of P1 Test Cases to be automated	Manual Effort fo all P1 testcases	Automation scripting Effort	Count of P2 Test Cases to be automated
ATG - Desktop	700	560	450	11	22	21	132	55

Modules	Manual Effort for executing P1 testcase	Manual Effort for executing P2 testcase	Manual Effort for executing P3 testcase	Automation Effort for scripting P1 testcase	Automation Effort for scripting P2 testcase	Automation Effort for scripting P3 testcase	Automation Execution Effort per P1 script	Automation Execution Effort per P2 script
ATG - Desktop	0.42	0.25	0.08	6	4	2.5	0.083333333	0.05

ROI example, continued

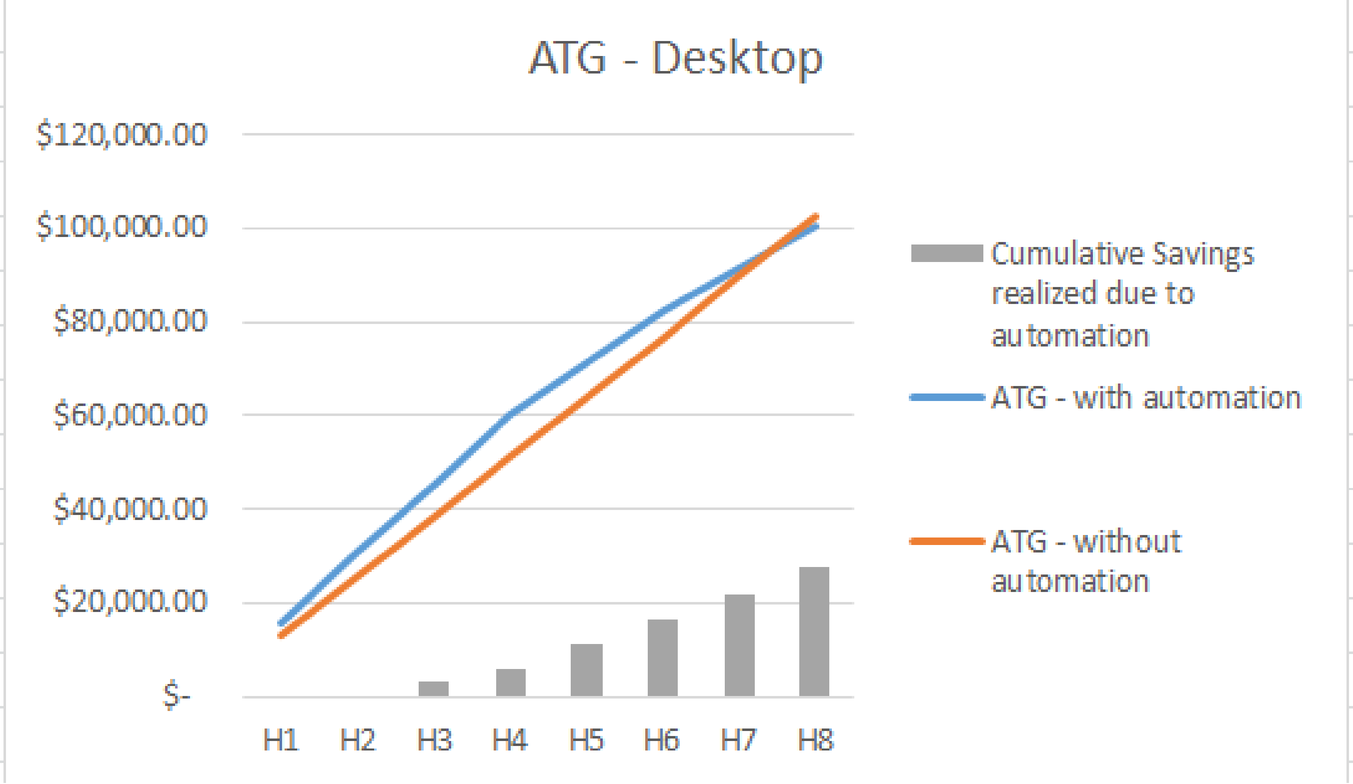
- Then compare today vs future and calculate savings year



Cumulative Cost	H1	H2	H3	H4	H5
ATG - with automation	$ 22,852.50	$ 45,705.00	$ 67,228.70	$ 88,752.40	$ 104,980.15
ATG - without automation	$ 19,222.50	$ 38,445.00	$ 57,667.50	$ 76,890.00	$ 96,112.50
Cumulative Savings realized due to automation	$ -	$ -	$ 4,620.00	$ 9,240.00	$ 16,826.70

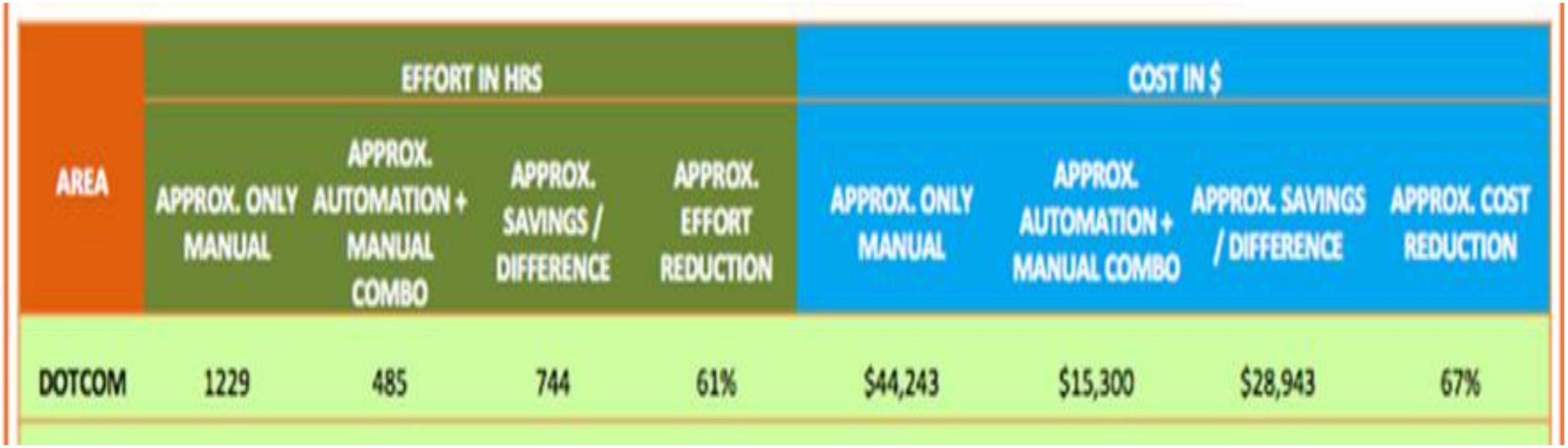
ROI example, continued

- **Depict the savings and ROI in a visual picture:**



ROI example, continued

- Keep track of the savings year over year to show the cost savings and to keep the investment going



AREA	EFFORT IN HRS				COST IN $			
	APPROX. ONLY MANUAL	APPROX. AUTOMATION + MANUAL COMBO	APPROX. SAVINGS / DIFFERENCE	APPROX. EFFORT REDUCTION	APPROX. ONLY MANUAL	APPROX. AUTOMATION + MANUAL COMBO	APPROX. SAVINGS / DIFFERENCE	APPROX. COST REDUCTION
DOTCOM	1229	485	744	61%	$44,243	$15,300	$28,943	67%

Where to Begin

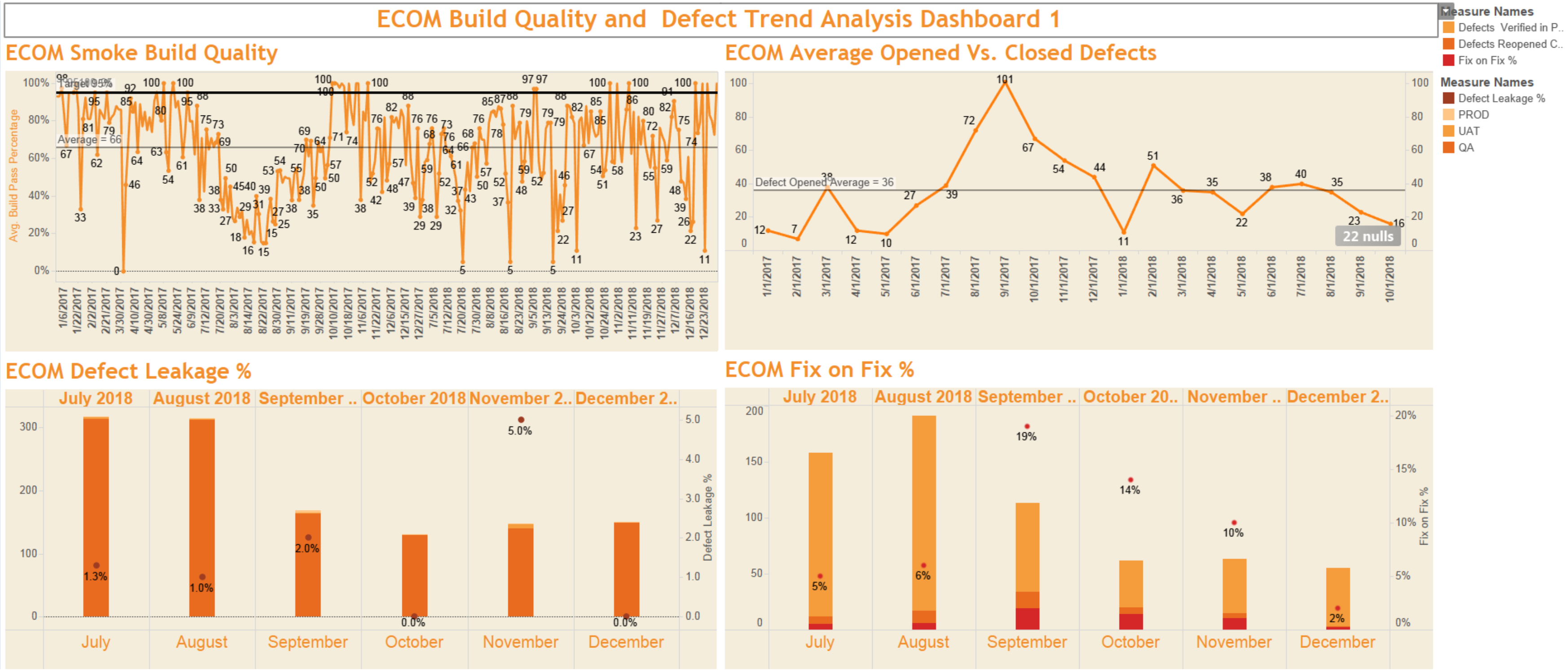
1. Solicit feedback – categorize it, prioritize it
   - ➢ Where is the team spending the most time?
   - ➢ What is the perception of QA that you need to address?
2. Target 2-3 categories initially

Collecting the Data

- Keep it simple! Do not collect data that will not be used or analyzed or acted upon
- Try to collect the data systematically (e.g. defect tracking system, project tracking system) and make it mandatory so you get the complete picture
- Train the team on why you are collecting the data and how it will be used. Provide the benefit statement(s). Reiterate quality is EVERYONE's job.
- Validate the data before publishing the report….is it complete and accurate?
- Make sure the data is in a db or a location that is backed up.

Presenting the Data

- Keep it clean and simple – like an executive dashboard



Presenting the Data, cont.

- Provide the results and analysis promptly, at regular intervals so that it can be used to drive continuous improvement
- Get a member of leadership to be the quality advocate and influence other leaders. Ensure your CIO is also aware of the metrics and any action (or lack thereof) as a result.

Analyzing the Data – What Does it Show?

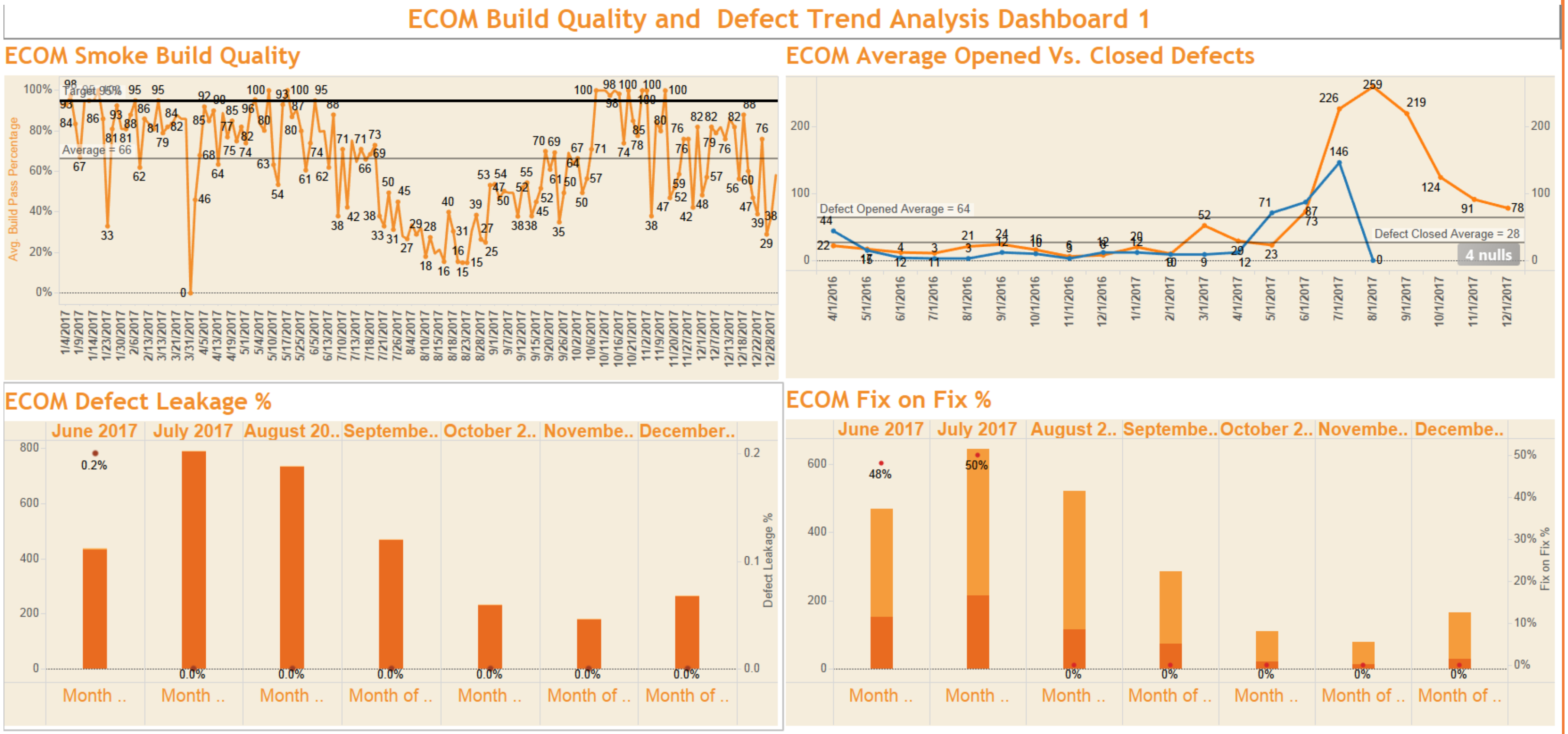
- If you are sharing the metrics and there's no action, then you are wasting your time!
- Proactively have conversations with your project/application teams around the results
  - Assign action items
  - Follow up on them
- Bring the metrics to the Lessons Learned sessions – ask for action plan and follow up!
- Track progress and validate the improvements made using the trend graphs
- Publish the results! Recognize teams that are demonstrating their investment in improving the quality!

The ULTA BEAUTY Example

Centralized metrics repository:

- Alignment across teams
- Governed and produced by QA Team
- Metrics targeted a few priority needs
- Action plans developed for top priorities based on data
- Results of actions plans measured

Metrics and Code Quality



ECOM Build Quality and Defect Trend Analysis Dashboard 1

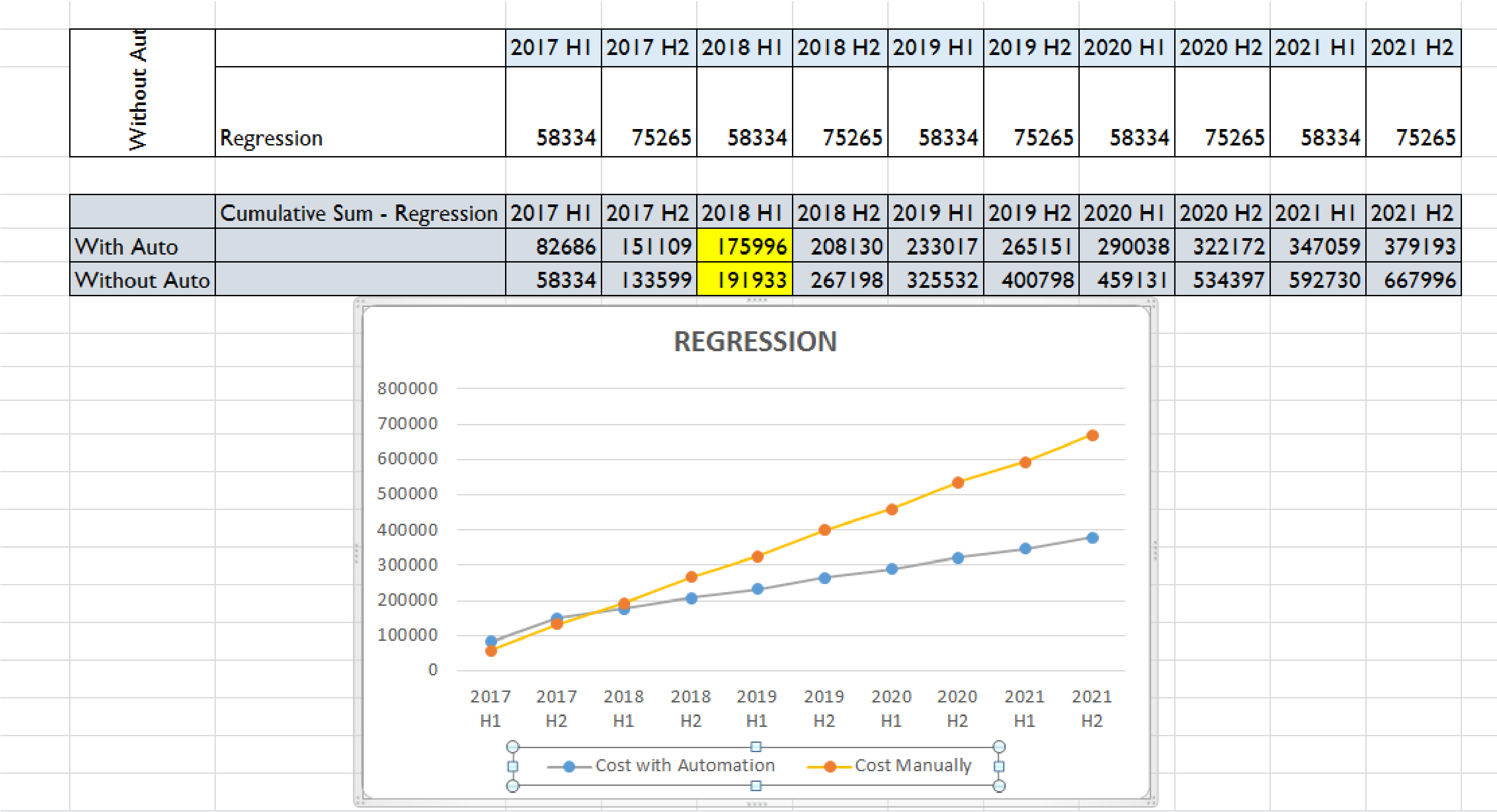
ECOM Smoke Build Quality

ECOM Average Opened Vs. Closed Defects

ECOM Defect Leakage %

ECOM Fix on Fix %

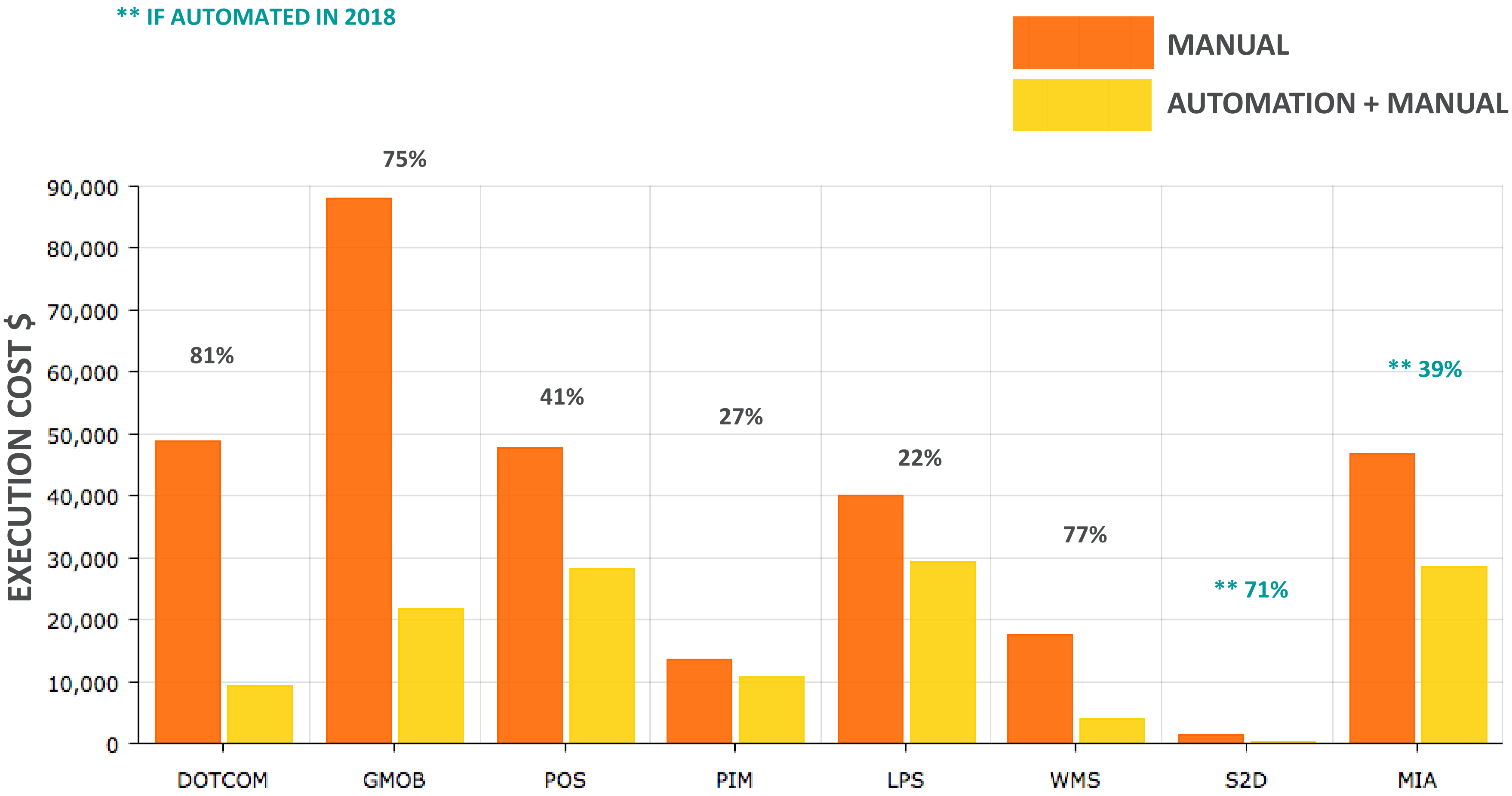
Automation Cost Estimate & ROI - Example



		2017 H1	2017 H2	2018 H1	2018 H2	2019 H1	2019 H2	2020 H1	2020 H2	2021 H1	2021 H2
Without Aut	Regression	58334	75265	58334	75265	58334	75265	58334	75265	58334	75265

	Cumulative Sum - Regression	2017 H1	2017 H2	2018 H1	2018 H2	2019 H1	2019 H2	2020 H1	2020 H2	2021 H1	2021 H2
With Auto		82686	151109	175996	208130	233017	265151	290038	322172	347059	379193
Without Auto		58334	133599	191933	267198	325532	400798	459131	534397	592730	667996

Cost Optimization – Test Automation



What Format Should I Use?

- **Comparison across categories**: column-based bar graph
- **Trends in data over a period of time**: line charts or line graphs
- **Comparison of parts over a whole**: pie chart

The ULTA BEAUTY Example: Lessons Learned

- ✓ Perception is often incorrect and highly subjective – turn it around with data that is objective and factual
- ✓ Clarify and align expectations for acceptable quality levels
- ✓ Combine short-term returns (quick hits) with long-term commitment
- ✓ Champions and leaders make the difference in initiating change – keep them looped in on progress and trends
- ✓ Use data to show the ROI - a compelling value proposition is essential
- ✓ Pictures say what you don't need to

Q&A