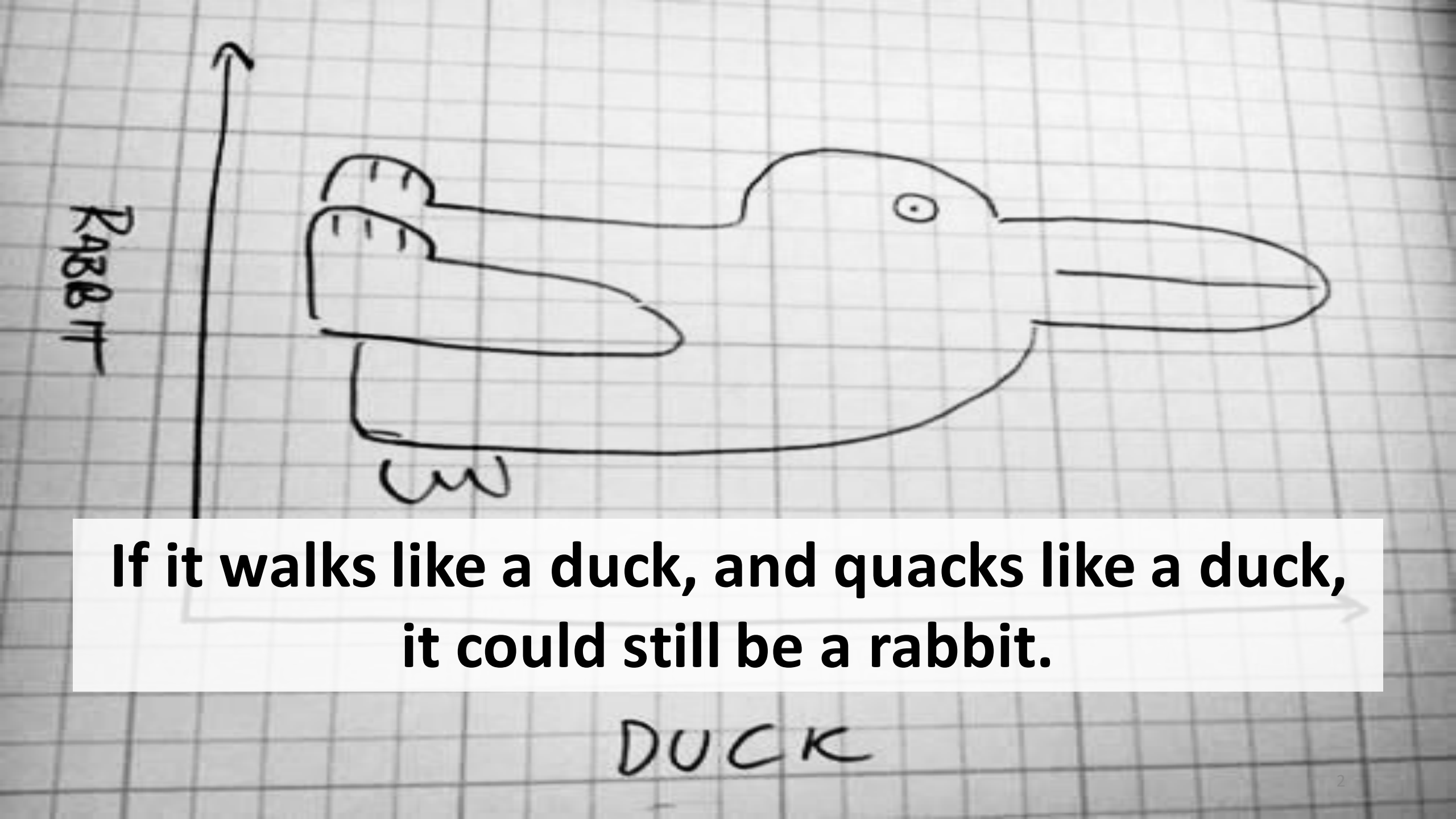
Data Driven Coaching

Safely turning team data into coaching insights (Troy Magennis)

These slides available here:
[link]
@t_magennis
troy.magennis@FocusedObjective.com



If it walks like a duck, and quacks like a duck, it could still be a rabbit.

Data is E V I L


Being judged
unfairly is
un-bearable…
Never coerce
Never embarrass


Make a difference,
not just make a point

Use data to tell a story…

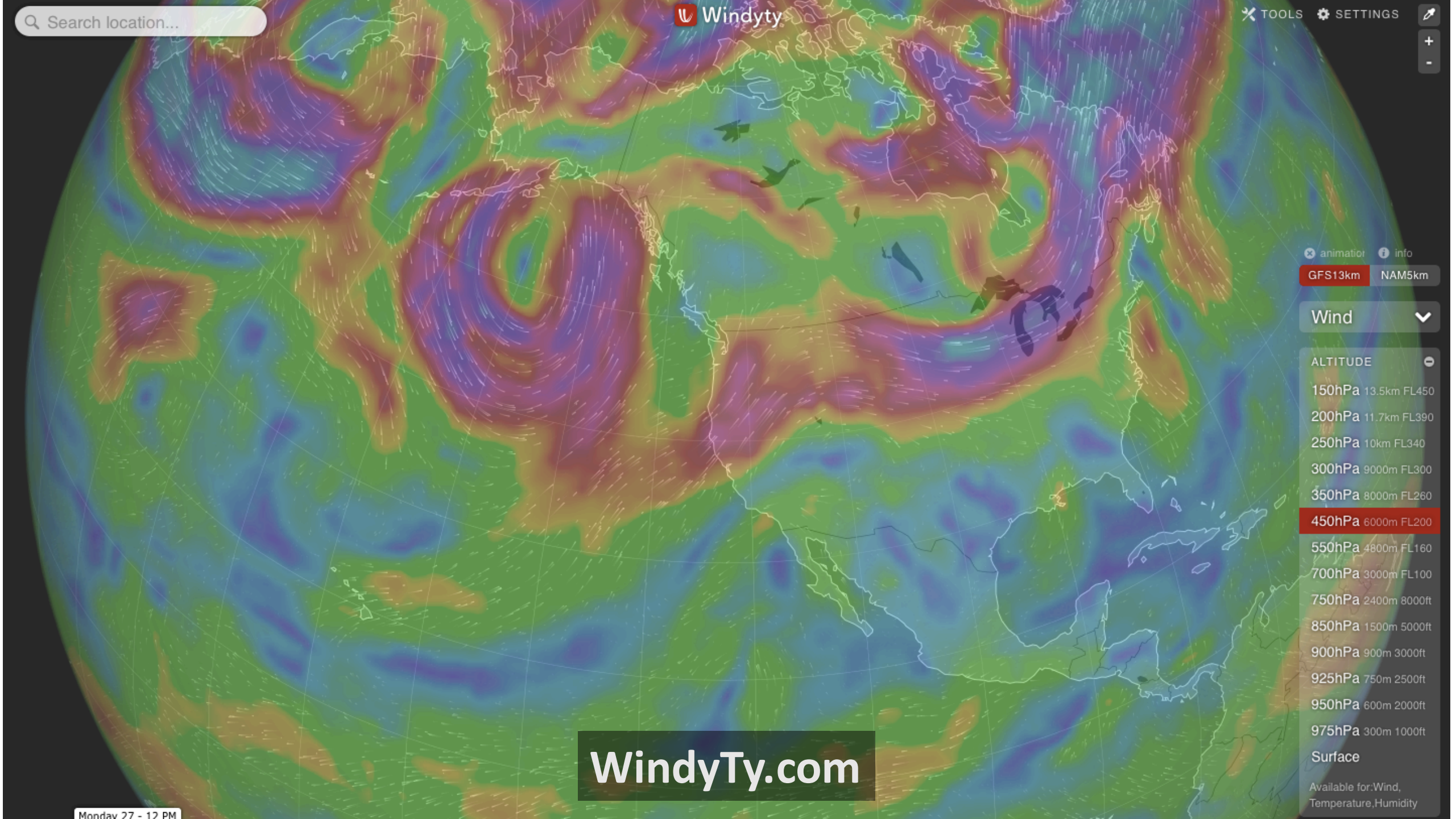

Without a
story, data
is boring…

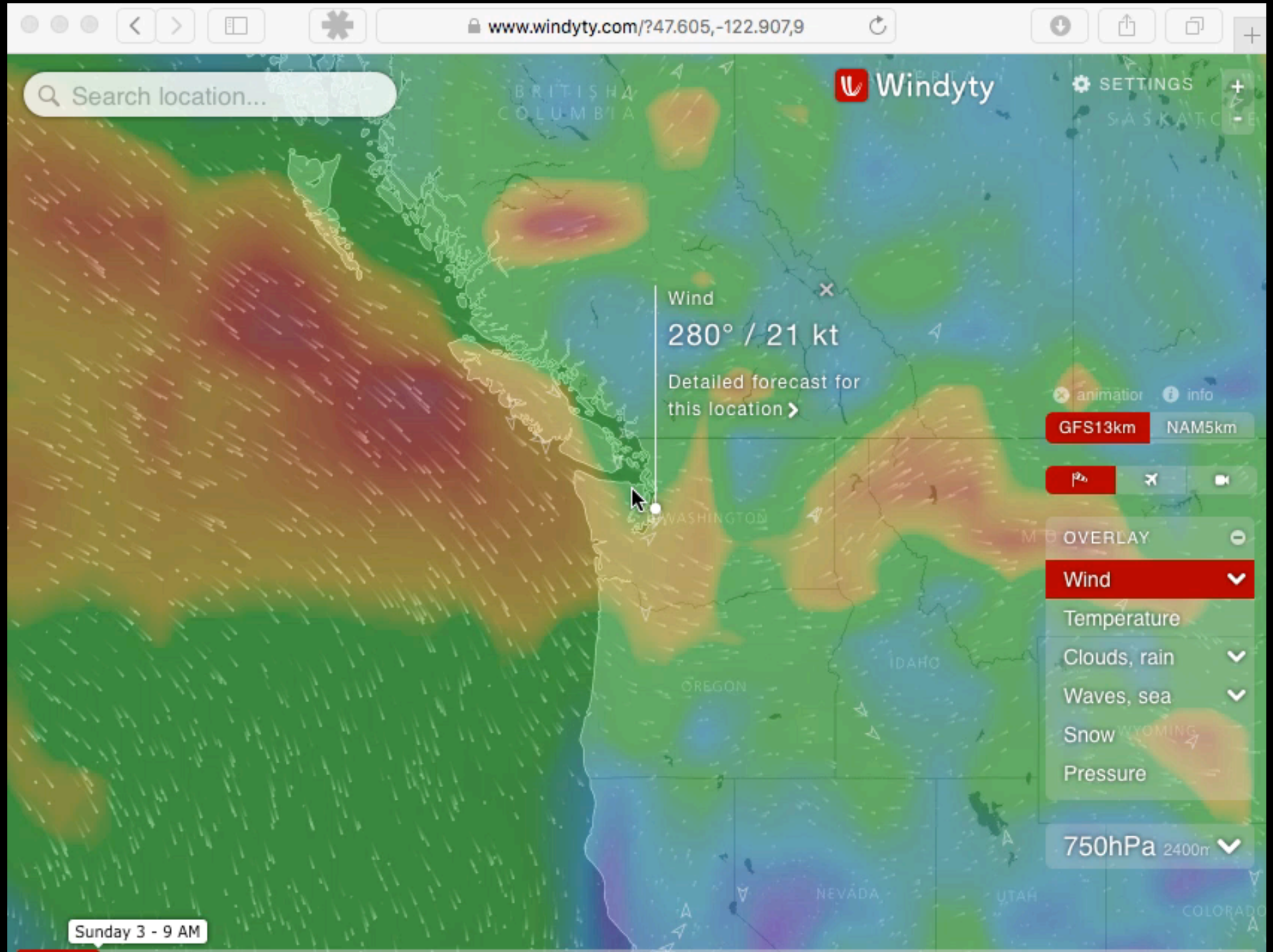
```
STATION,STATION_NAME,ELEVATION,LATITUDE,LONGITUDE,DATE,TMAX,TMIN,PRCP
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100101,-178,-311,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100102,-244,-322,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100103,-194,-289,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100104,-167,-200,15
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100105,-133,-167,999
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100106,-133,-172,999
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100107,-150,-278,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100108,-233,-328,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100109,-233,-322,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100110,-117,-244,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100111,-67,-128,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100112,-78,-122,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100113,-17,-89,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100114,39,-72,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100115,-67,-72,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100116,22,-50,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100117,33,-44,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100118,6,-172,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100119,-56,-183,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100120,-67,-139,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100121,-67,-94,25
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100122,-44,-67,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100123,-6,-44,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100124,0,-11,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100125,-11,-161,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100126,-161,-233,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100127,-167,-222,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100128,-167,-283,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100129,-189,-283,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100130,-156,-267,0
GHCND:USC00327027,PETERSBURG 2 N ND US,466.3,48.0355,-98.01,20100131,-150,-272,0
```

NOAA

Windyty

Ivo


Search location...
Windyty
TOOLS
SETTINGS
animation
info
GFS13km
NAM5km
Wind
ALTITUDE
150hPa 13.5km FL450
200hPa 11.7km FL390
250hPa 10km FL340
300hPa 9000m FL300
350hPa 8000m FL260
450hPa 6000m FL200
550hPa 4800m FL160
700hPa 3000m FL100
750hPa 2400m 8000ft
850hPa 1500m 5000ft
900hPa 900m 3000ft
925hPa 750m 2500ft
950hPa 600m 2000ft
975hPa 300m 1000ft
Surface
Available for:Wind, Temperature,Humidity
WindyTy.com
Monday 27 - 12 PM


www.windyty.com/?47.605,-122.907,9
Search location...
Windyty
SETTINGS
Wind
280° / 21 kt
Detailed forecast for this location
animation
info
GFS13km
NAM5km
OVERLAY
Wind
Temperature
Clouds, rain
Waves, sea
Snow
Pressure
750hPa 2400m
Sunday 3 - 9 AM
BRITISH COLUMBIA
WASHINGTON
OREGON
IDAHO
NEVADA
UTAH
COLORADO
WYOMING

WindyTy.com

	A	B	C	D	E	F	G
1	Weekly Poliomyelitis Case reports for Washington						
2	Data provided by Project Tycho, Data Version 1.0.0, released 28 November 2013.						
3	YEAR	WEEK	WASHINGTON				
855	1943	17	1				
856	1943	18	0				
857	1943	19	0				
858	1943	20	2				
859	1943	21	-				
860	1943	22	-				
861	1943	23	3				
862	1943	24	1				
863	1943	25	0				
864	1943	26	3				
865	1943	27	0				
866	1943	28	-				
867	1943	29	-				
868	1943	30	2				
869	1943	31	5				
870	1943	32	13				
871	1943	33	20				
872	1943	34	25				
873	1943	35	19				
874	1943	36	7				
875	1943	37	27				
876	1943	38	22				
877	1943	39	19				
878	1943	40	30				
879	1943	41	28				
880	1943	42	24				
881	1943	43	37				
882	1943	44	15				
883	1943	45	8				
884	1943	46	30				
885	1943	47	-				
886	1943	48	3				
887	1943	49	7				
888	1943	50	-				

University of Pittsburgh

PROJECT TYCHO
DATA FOR HEALTH

THE WALL STREET JOURNAL.

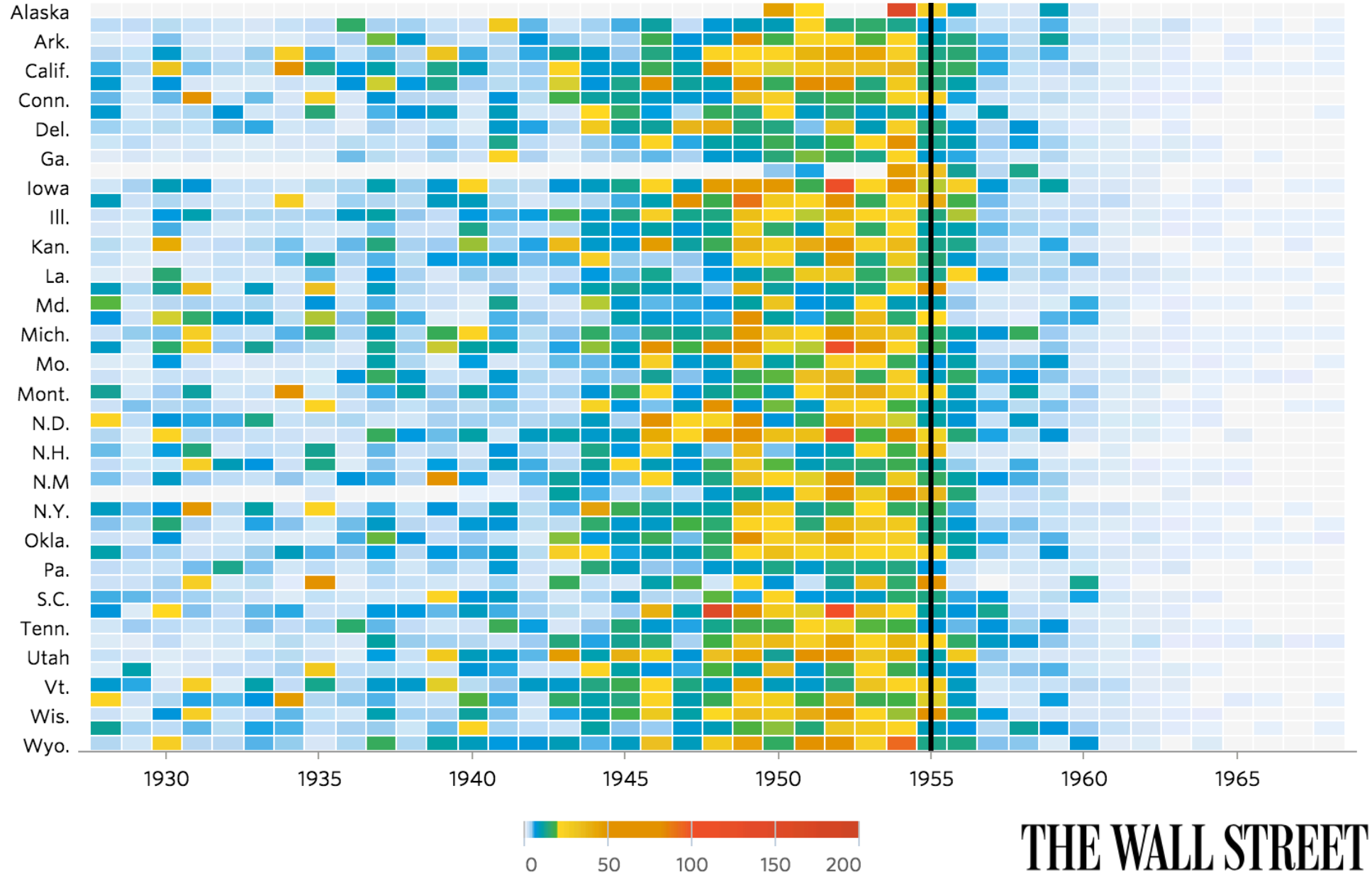
Battling Infectious Diseases in the 20th Century: The Impact of Vaccines

By Tynan DeBold and Dov Friedman
Published Feb. 11, 2015 at 3:45 p.m. ET

The number of infected people, measured over 70-some years and across all 50 states and the District of Columbia, generally declined after vaccines were introduced.

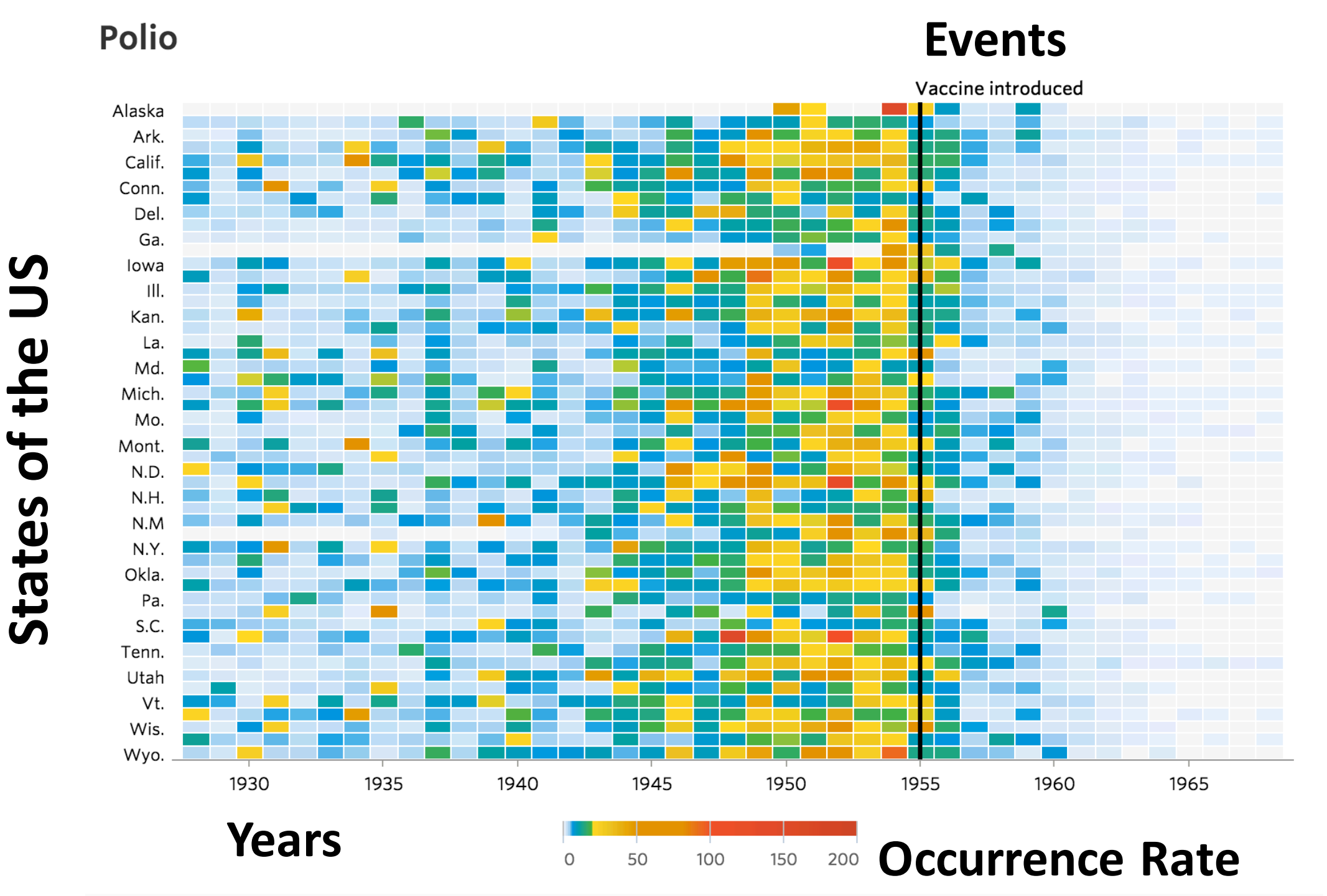
Polio

Polio Vaccine Introduced



THE WALL STREET JOURNAL.


Compared
To What?

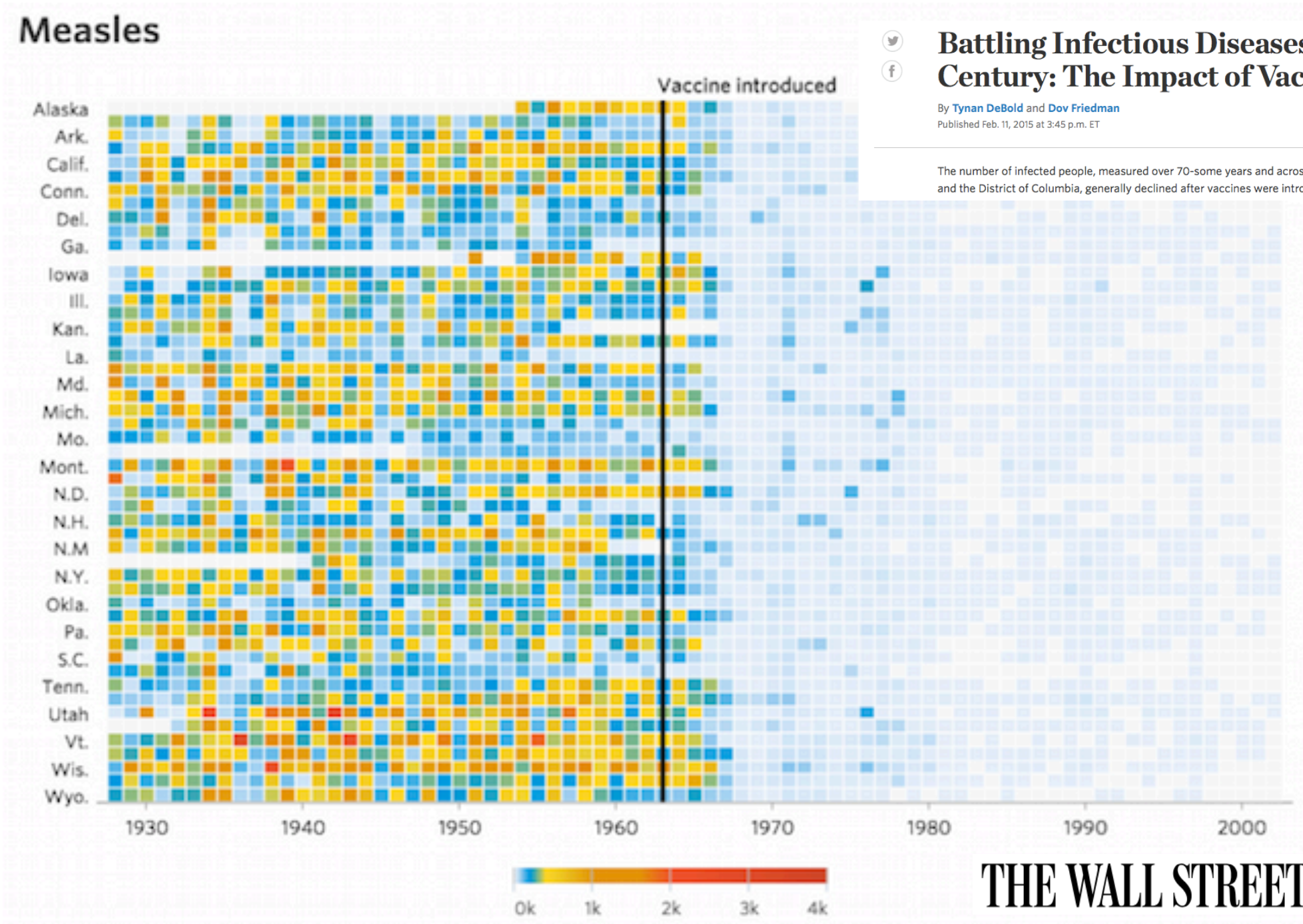

Polio
Events
Vaccine introduced
States of the US
Alaska
Ark.
Calif.
Conn.
Del.
Ga.
Iowa
Ill.
Kan.
La.
Md.
Mich.
Mo.
Mont.
N.D.
N.H.
N.M
N.Y.
Okla.
Pa.
S.C.
Tenn.
Utah
Vt.
Wis.
Wyo.
1930
1935
1940
1945
1950
1955
1960
1965
Years
0
50
100
150
200
Occurrence Rate

Battling Infectious Diseases in the 20th Century: The Impact of Vaccines

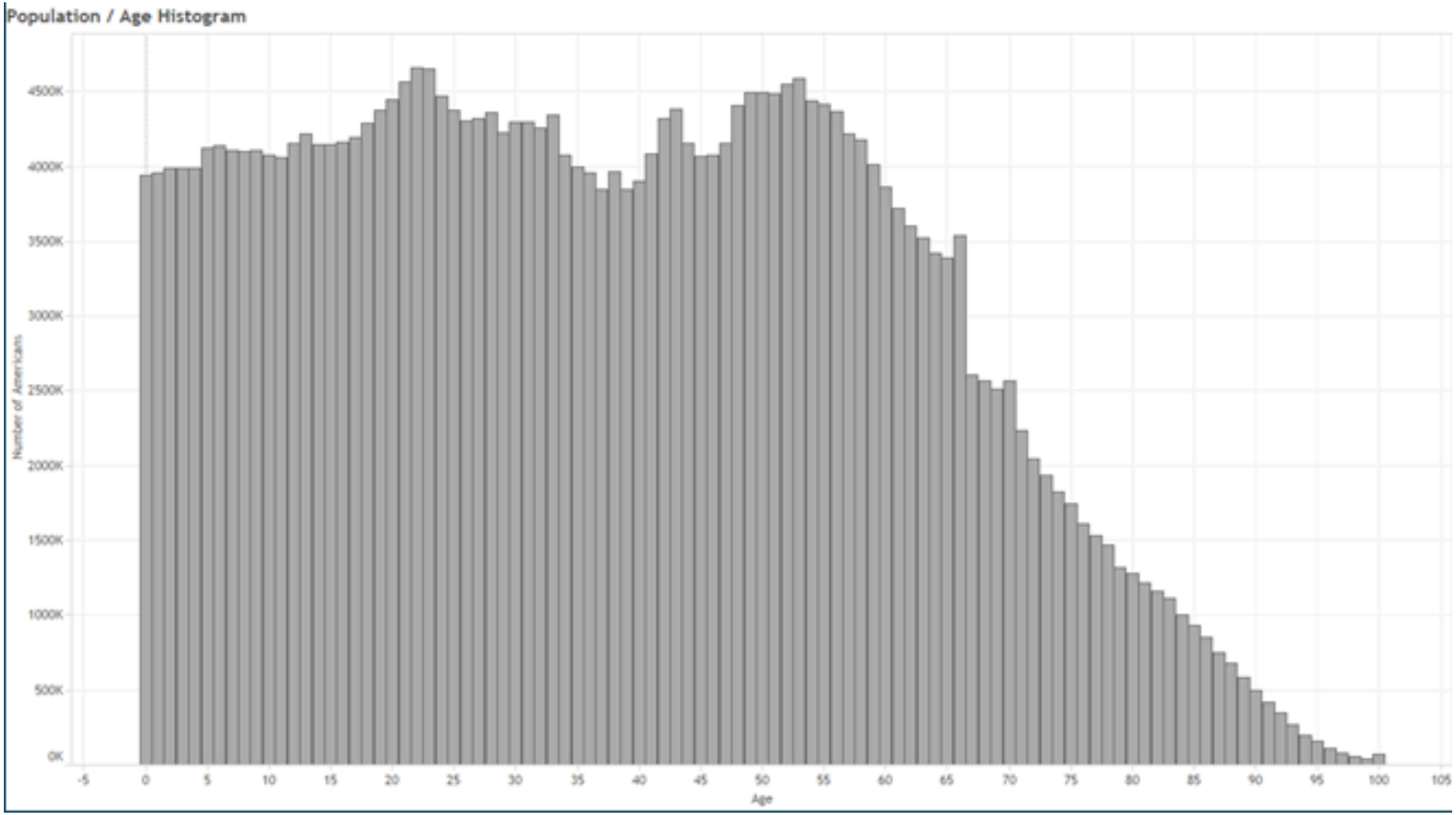
By Tynan DeBold and Dov Friedman

Published Feb. 11, 2015 at 3:45 p.m. ET

The number of infected people, measured over 70-some years and across all 50 states and the District of Columbia, generally declined after vaccines were introduced.



THE WALL STREET JOURNAL.

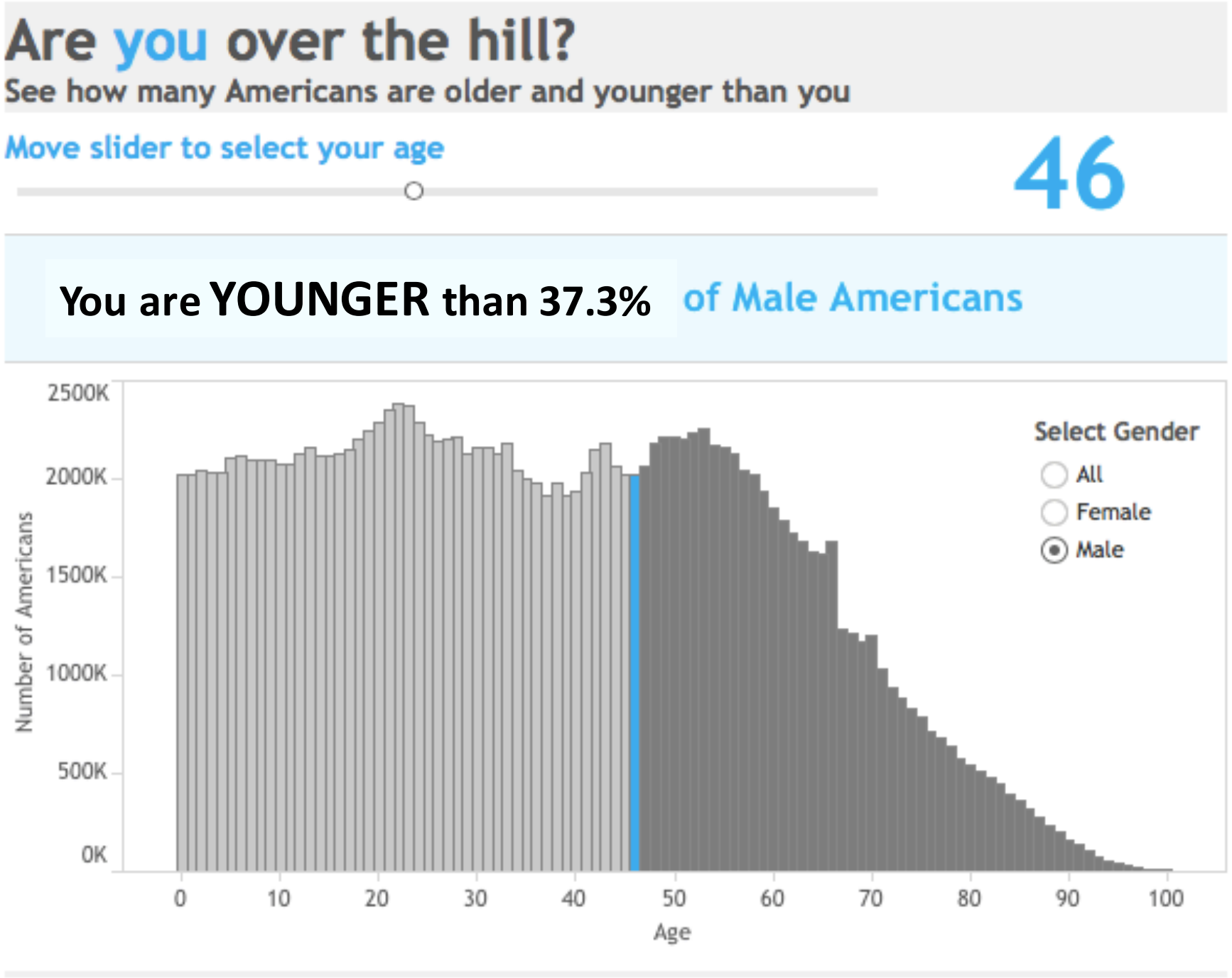


United States Census Bureau

DATA REVELATIONS

TURNING INFORMATION INTO INSIGHT

Not compared with anything meaningful (to me.) Boring.



"Me"
On the 7th August

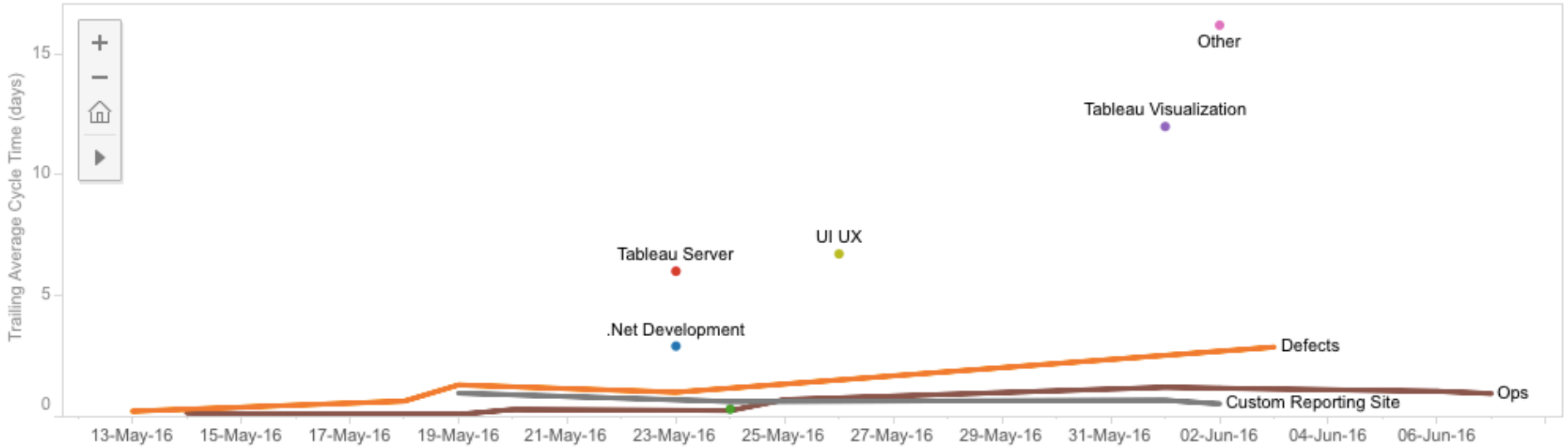
I WANT TO KNOW

But, I may not want you to know.

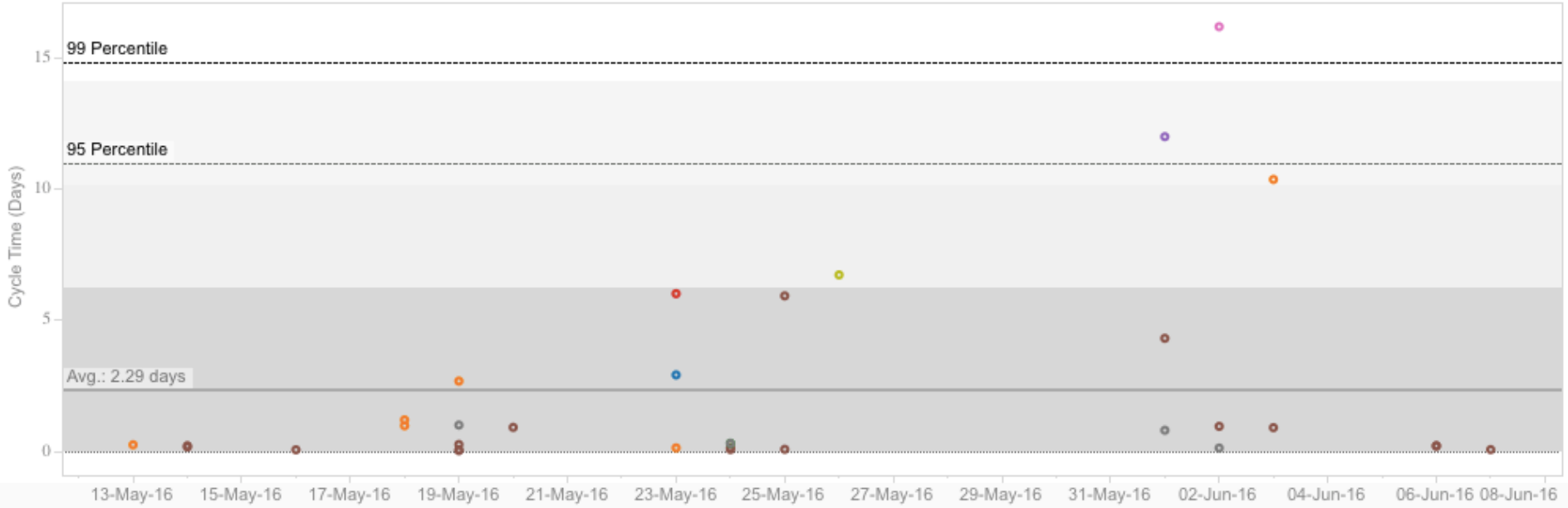
http://www.datarevelations.com/are-you-over-the-hill-in-the-usa

Makeover

How long does it take us to complete cards?



Where do we need to focus our improvement efforts?



DATE RANGE

05/11/2016 TO 06/10/2016

Card Event Data for Last 365 Days

TAGS

Calculate based on Card Size
(No card size = 1)

Reset Filters

Card Type

- .Net Development
- Tableau Visualization
- Node Development
- QA
- Tableau Server
- SQL Server
- Defects
- Other

Class of Service

Calculate Average By

2

Troy Magennis @t_magennis · Oct 2

"What types of work cause the most future variation in our process delivery rates and lead-time" - finding what work types cause chaos

4

2

Troy Magennis @t_magennis · Oct 2

"Is the age of work in process similar to the overall average age for similar types of work" - are we growing or shrinking our average

4

1

Troy Magennis @t_magennis · Oct 2

"Is the rate we are completing similar items of work sustainable?" - is the departure rate and arrival rate balanced, how transient.

6

3

Troy Magennis @t_magennis · Oct 2

"Are we completing similar work types at the same rate" - is our process getting to completion (done-done)

2

3

Troy Magennis @t_magennis · Oct 2

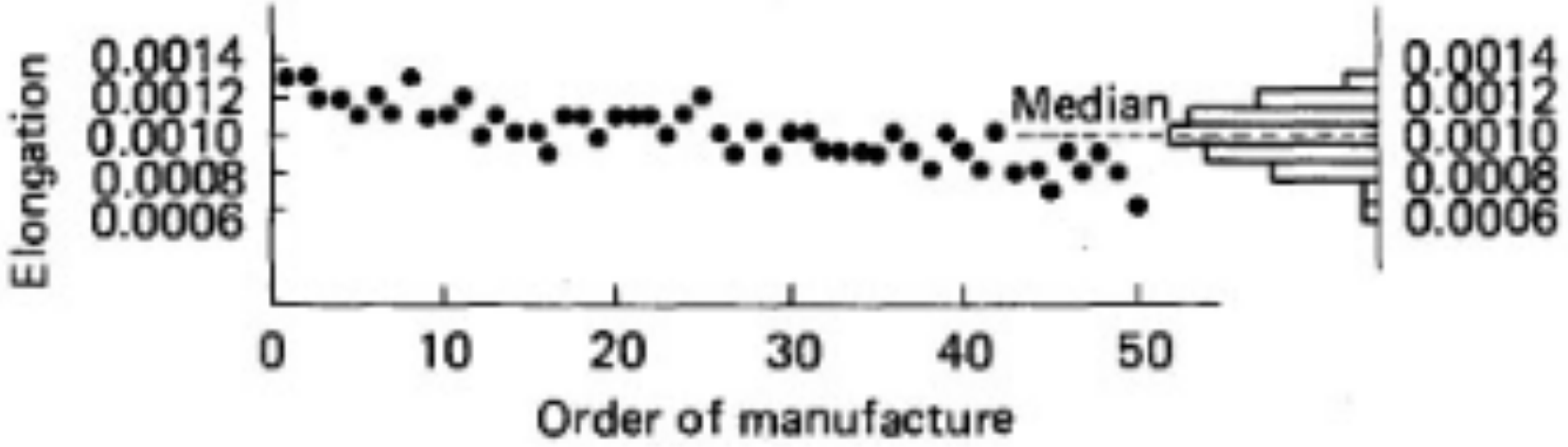
"Is similar work types taking longer to complete?" - is our process changing

Time and Pace related questions

1. Is it taking us longer to do the same type of work?
2. What is a good commitment cycle time to others? (SLA)
3. What is and how stable is our completed work rate?
4. Where should we focus improvement efforts?

- Compared to what?
  - Compared to the same type of work versus all work
  - Compared to the same time period last week/month/year
  - My work compares to others (only seen by me so I can improve)

Q. Is the process stable? First, do no harm.

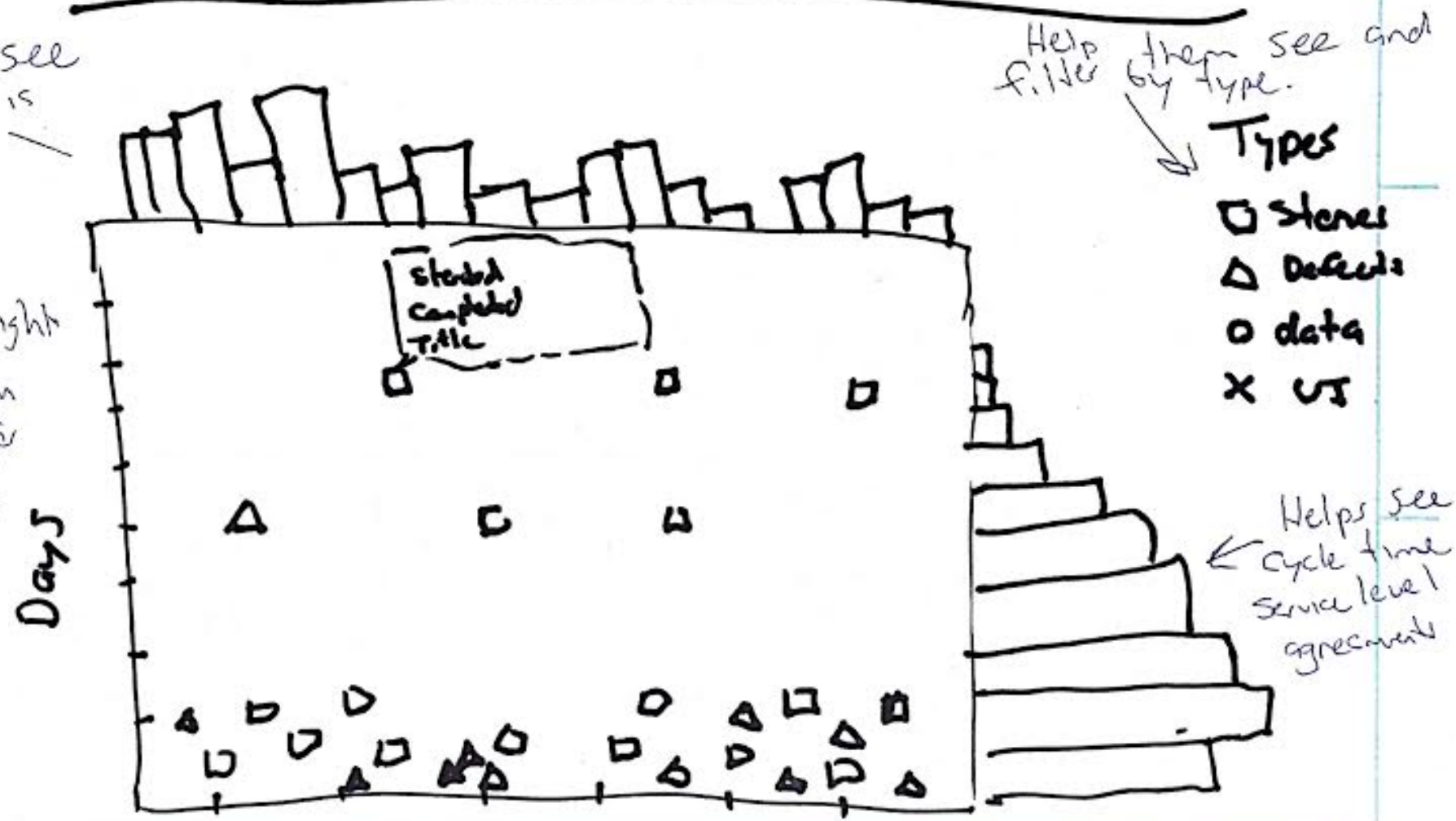


"If anyone adjusts a stable process, the output that follows will be worse than if (s)he had left the process alone"

Attributed to William J Latzko.

Source: Out of the Crisis. Deming.

Cycle time / Lead time Scatter plot -

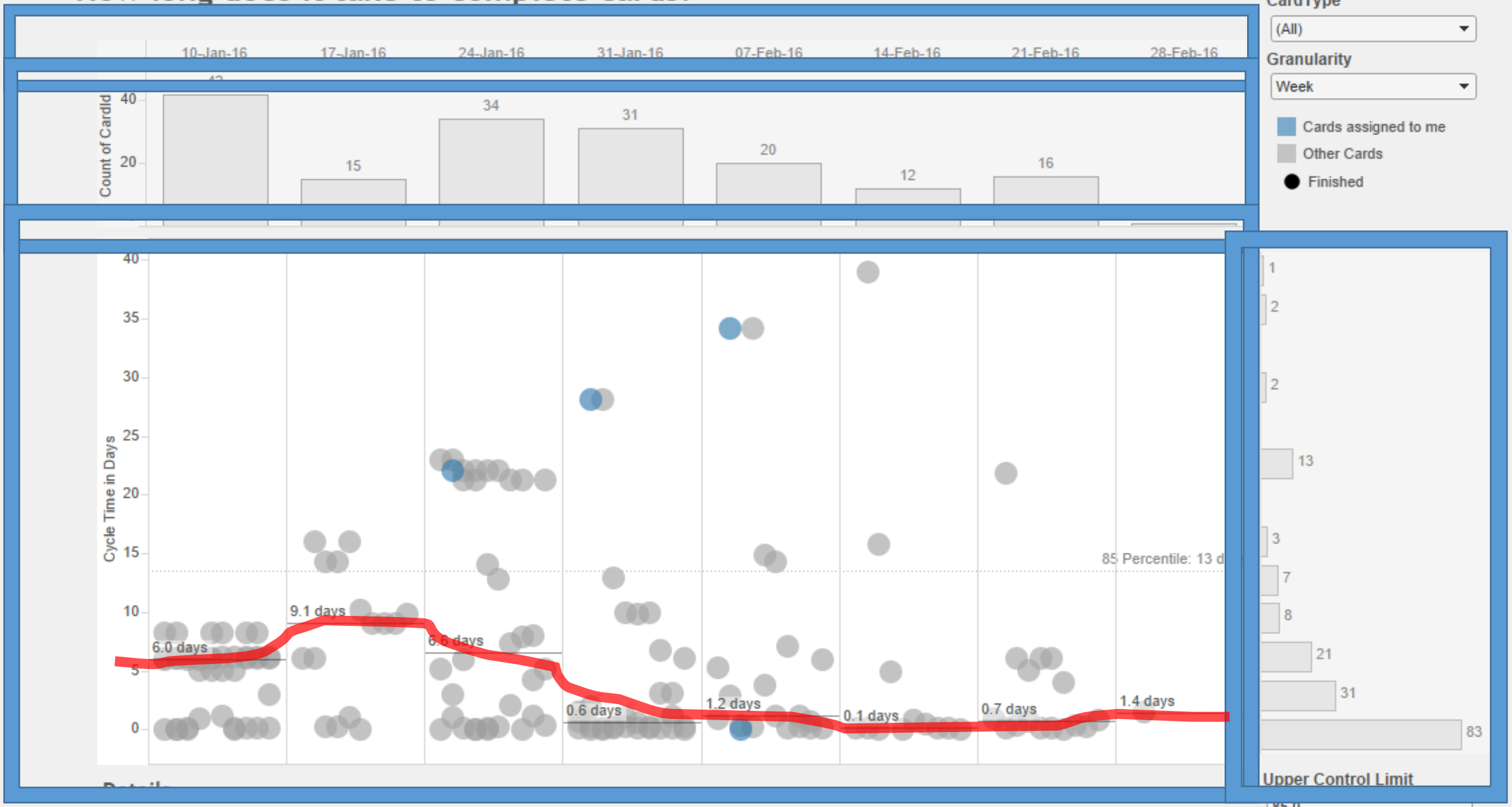


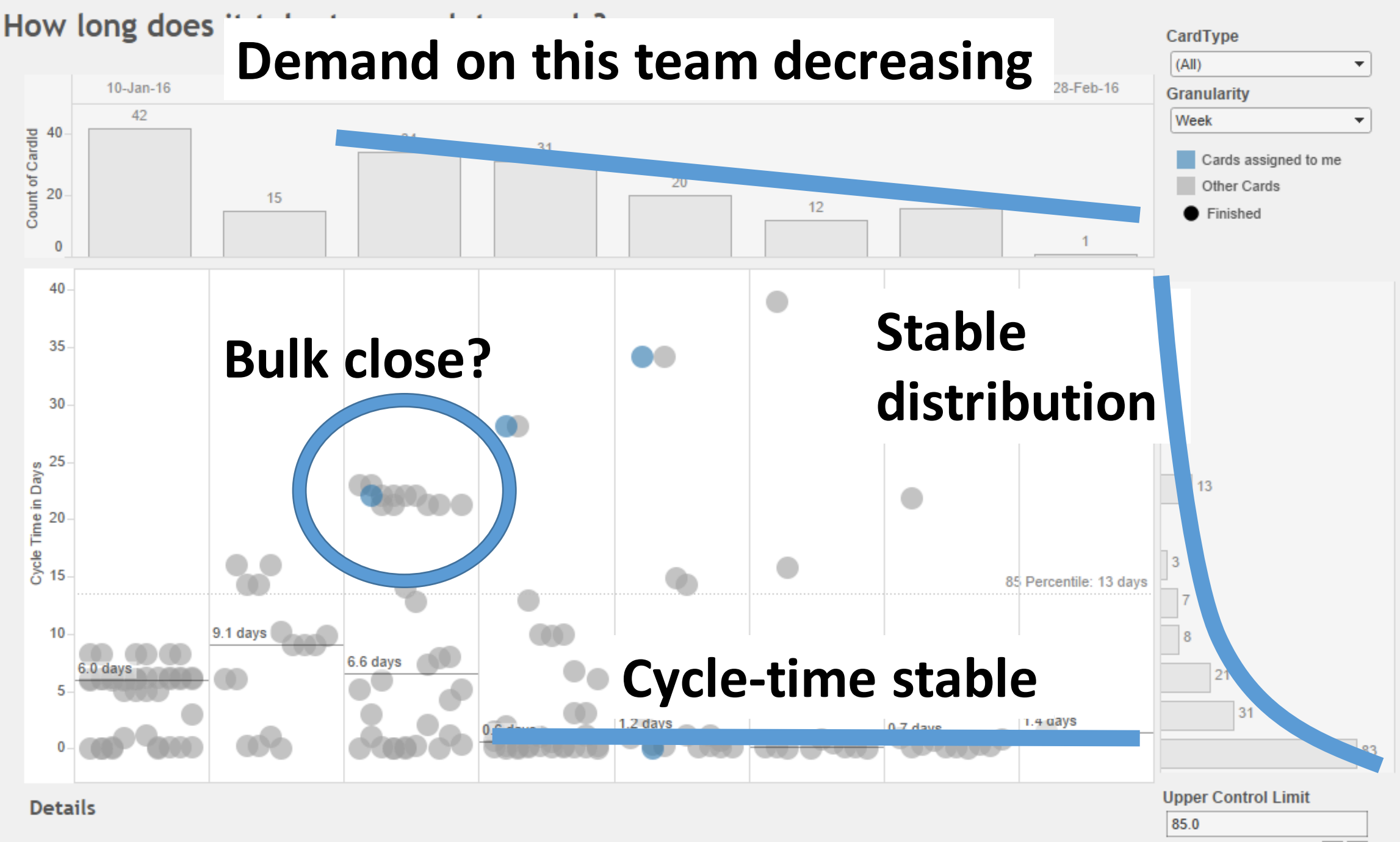
use color state for current state

completed Work in Prog. Assigned to you

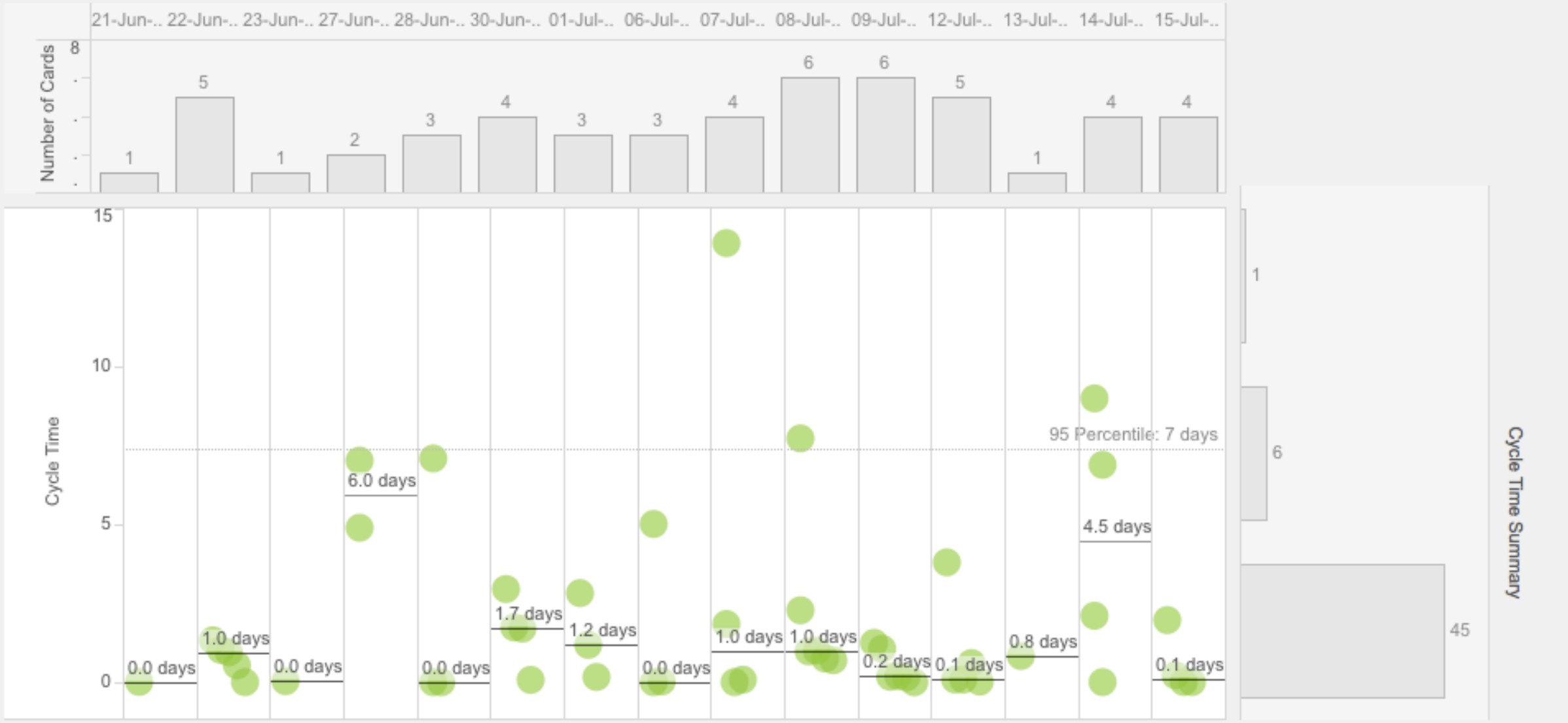
See yours at a glance

How long does it take to complete cards?




How long does
Demand on this team decreasing
CardType
(All)
Granularity
Week
Cards assigned to me
Other Cards
Finished
10-Jan-16
28-Feb-16
Count of CardId
42
15
31
20
12
1
Bulk close?
Stable distribution
Cycle Time in Days
85 Percentile: 13 days
6.0 days
9.1 days
6.6 days
1.2 days
13
3
7
8
21
31
Cycle-time stable
Details
Upper Control Limit
85.0

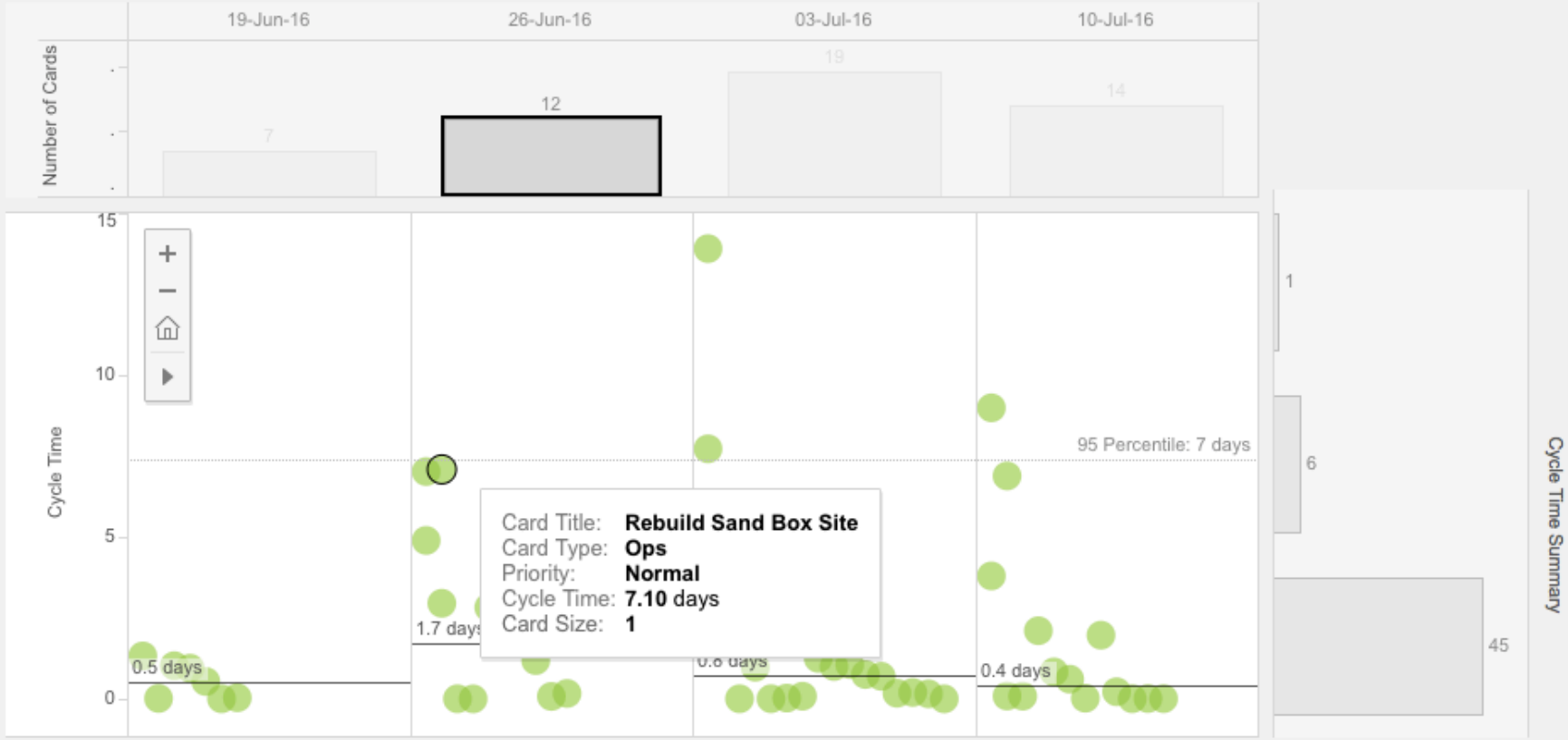
How long does it take us to complete cards?



To view card details, click on top bar chart area or scatter plot data points.

leankit

How long does it take us to complete cards?



Priority	Card Type	Card Title	Cycle Time
Normal	Ops	Fix SSL stuff	3.0
		Grant VPN Access to Scott/DL	7.0
		Make Tableau Dev Server Public IP and Access	4.9
		Rebuild Sand Box Site	7.1
	Other	Delete all boards except for the "USE ME"	0.0
		Uptime per quarter in 2015 for Nekotia	0.0

DATE RANGE

06/18/2016 TO 07/18/2016

Card Event Data for Last 365 Days

TIME PERIOD

Daily | Weekly | Monthly

REFERENCE LINE PERCENTAGE

95

TAGS

Calculate based on Card Size
(No card size = 1)

Reset Filters

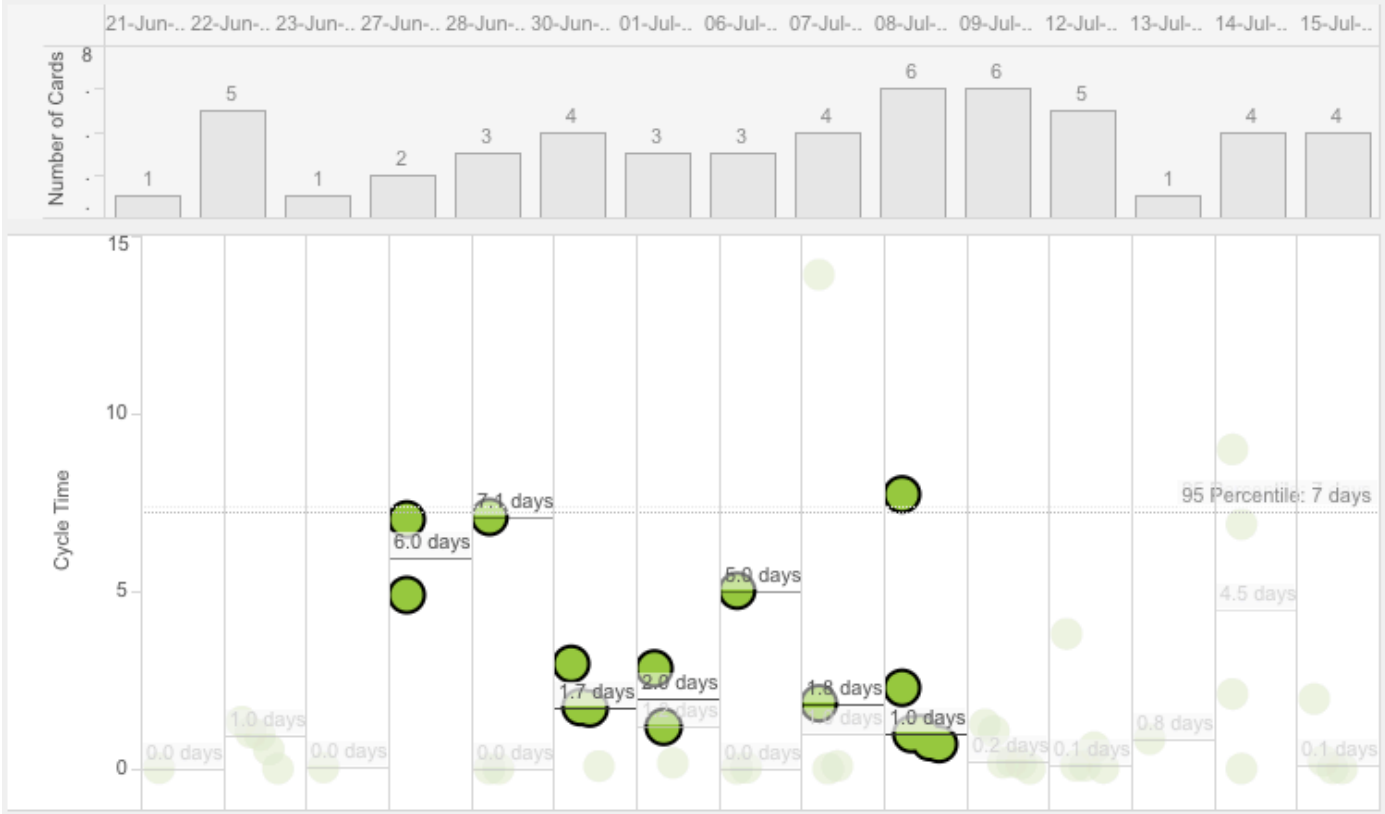
Card Type

- .Net Development
- Tableau Visualization
- Node Development
- QA
- Tableau Server
- SQL Server
- Defects
- Other

Class of Service

leankit

How long does it take us to complete cards?



Priority	Card Type	Card Title	Cycle Time
Normal	Custom Reporting Site	stand up a new "LKLD" ARS for training	0.7
		Training ARS (CRS) needs	1.0
	Defects	ZD #80096: Error loading Tableau data in https://d1-d09-analytics.leankit.io/	5.0
		ZD #80125: Can't save new workbook/view in https://d1-d09-analytics.leankit.io/	2.3
	Ops	Distribute New Tableau Desktop Keys	7.7
		Fix D09	0.8

DATE RANGE

06/18/2016 TO 07/18/2016

Card Event Data for Last 365 Days

TIME PERIOD

Daily | Weekly | Monthly

REFERENCE LINE PERCENTAGE

95

TAGS

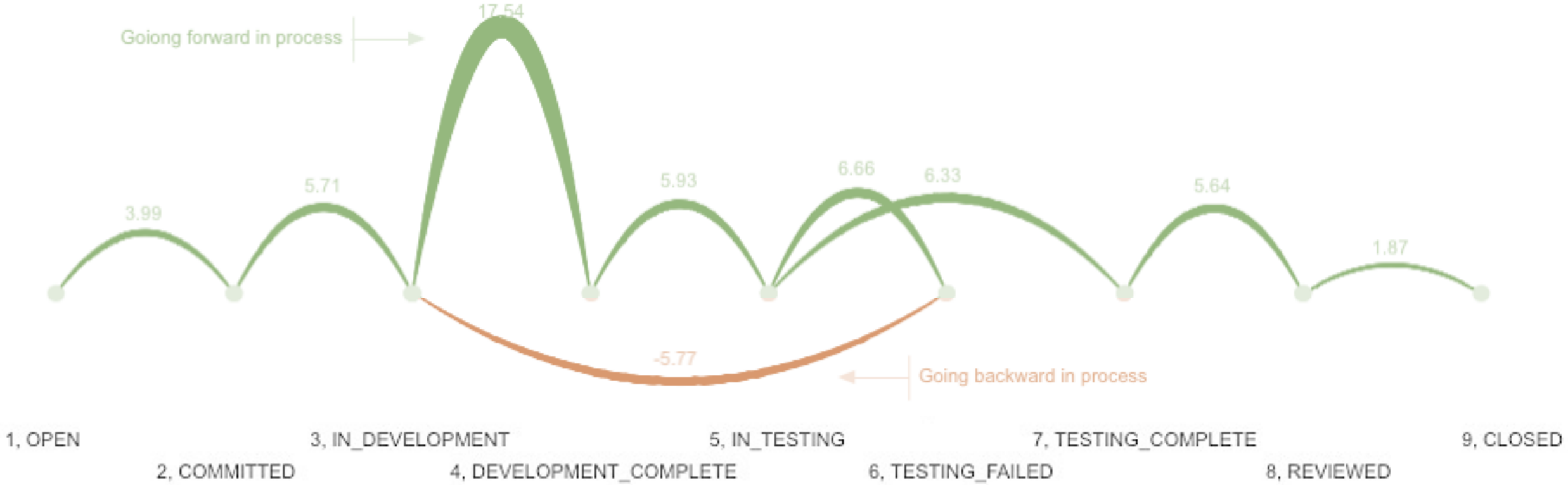
☐ Calculate based on Card Size
(No card size = 1)

Reset Filters

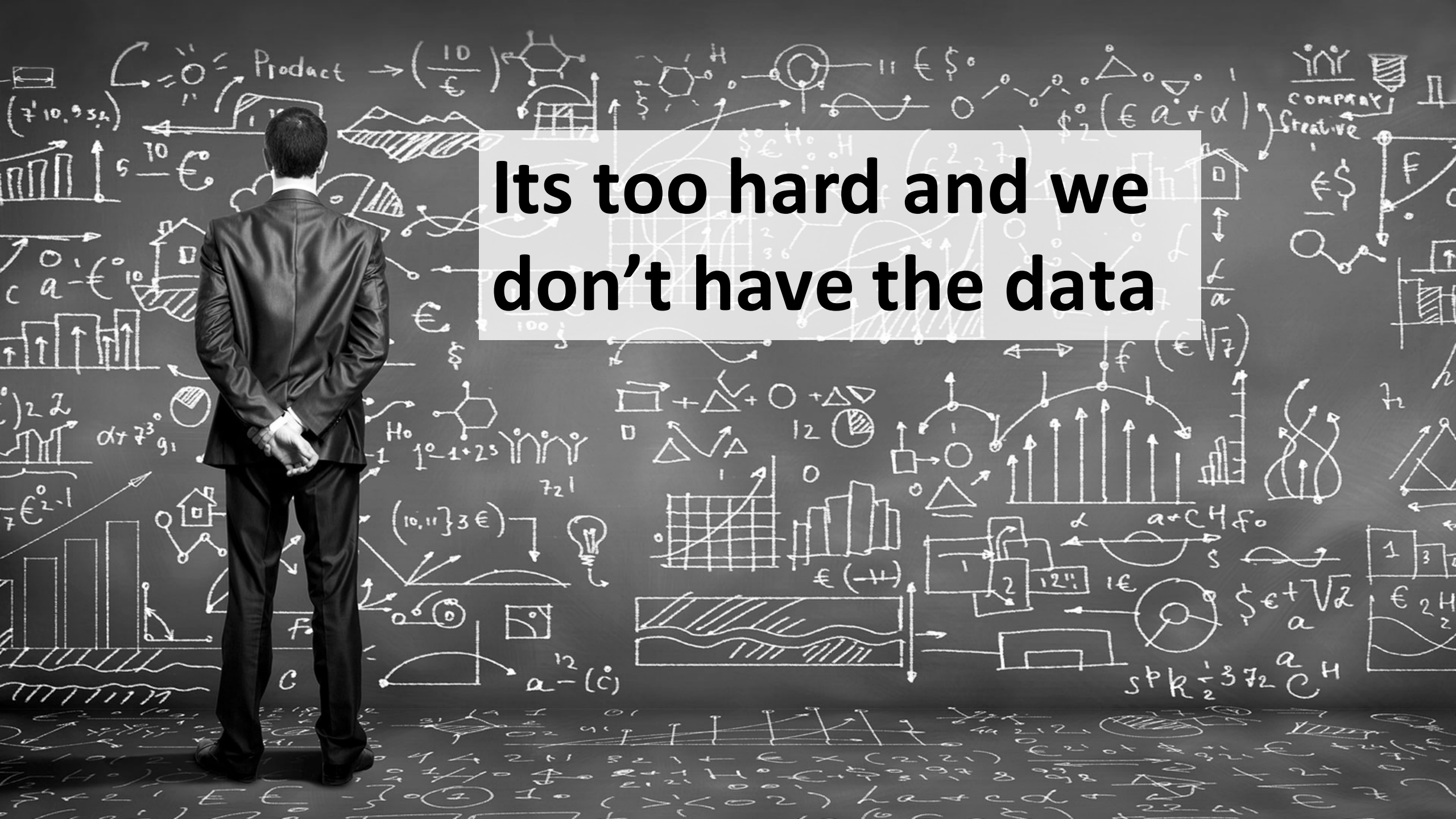
Card Type

- .Net Development
- Tableau Visualization
- Node Development
- QA
- Tableau Server
- SQL Server
- Defects
- Other

Class of Service



Source: JumpPlot.com (Tom VanBuskirk and Chris DeMartini)


Its too hard and we don’t have the data

Q. What could I do with just start and completed date?

http://bit.ly/Throughput

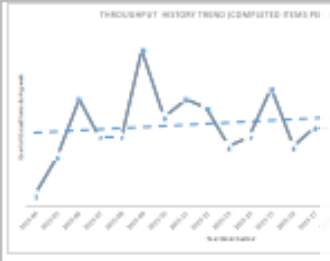
Or follow @t_magennis

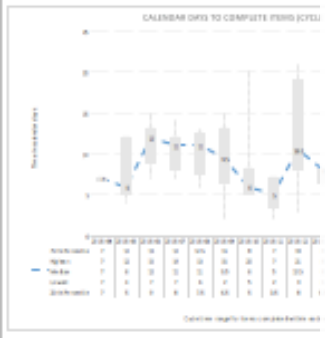
Format

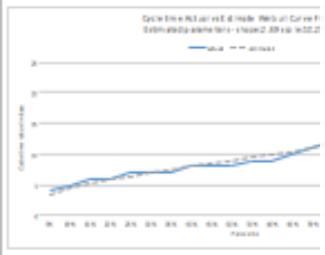
A2

http://bit.ly/Throughput

	A		
1	Completed Date	Start Date (optional)	Type (optional)
2	1/21/15	1/14/15	
3	1/26/15	1/14/15	Story
4	1/26/15	1/14/15	Defect
5	1/26/15	1/21/15	Story
6	1/26/15	1/22/15	Story
7	1/29/15	1/23/15	Story
8	2/2/15	1/23/15	Story
9	2/2/15	1/20/15	Defect
10	2/2/15	1/20/15	Defect
11	2/4/15	1/20/15	
12	2/4/15	1/26/15	
13	2/4/15	1/23/15	
14	2/4/15	1/22/15	





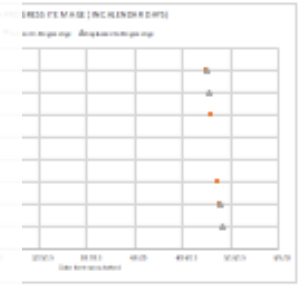


17 charts so far…

Throughput (planned & un-planned)

Throughput Histogram(s)

Cycle Time (planned & un-planed)



Cycle Time Histogram(s)

Work In Process

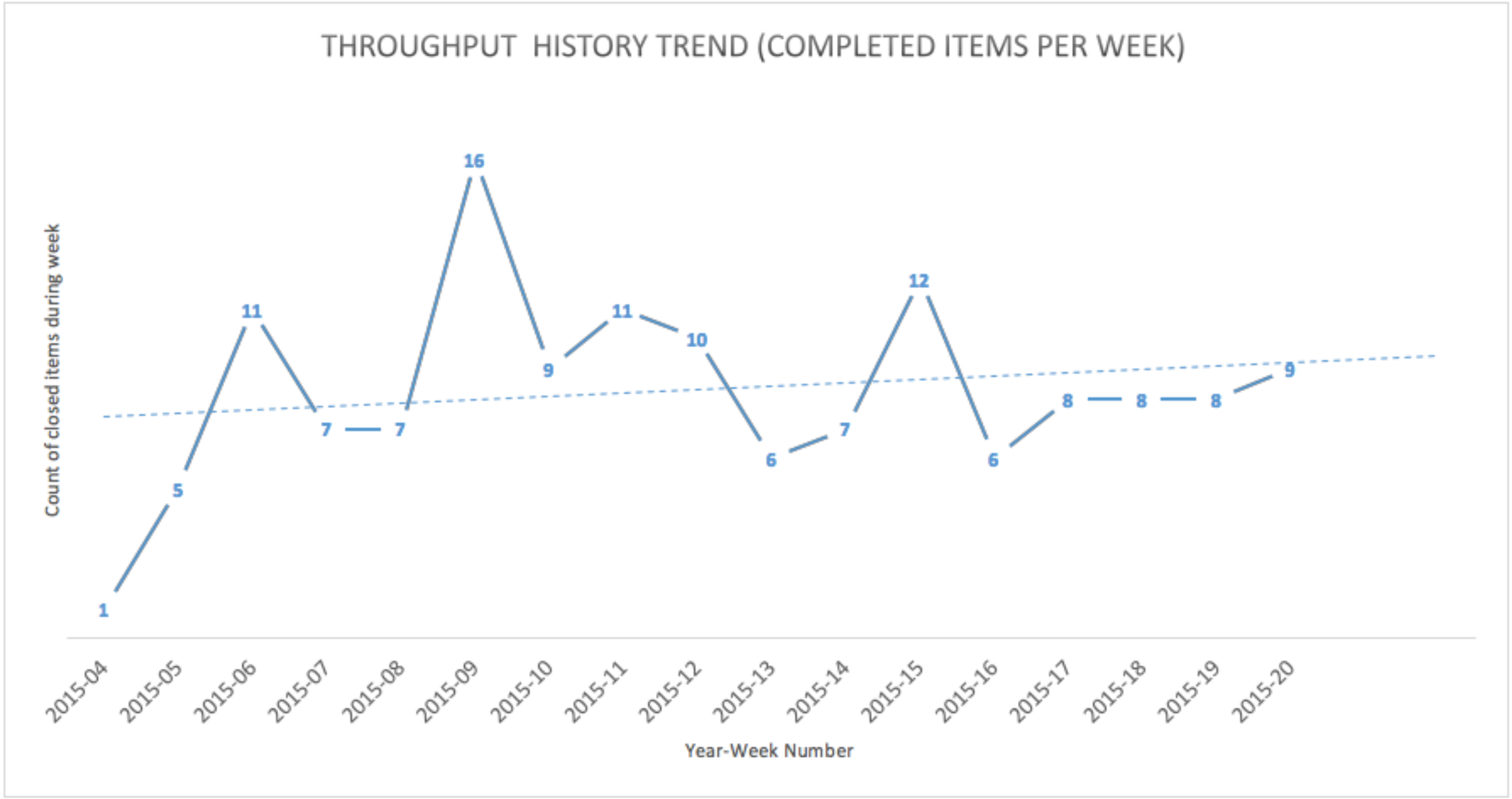
Cumulative Flow

Arrival vs Departure Rate

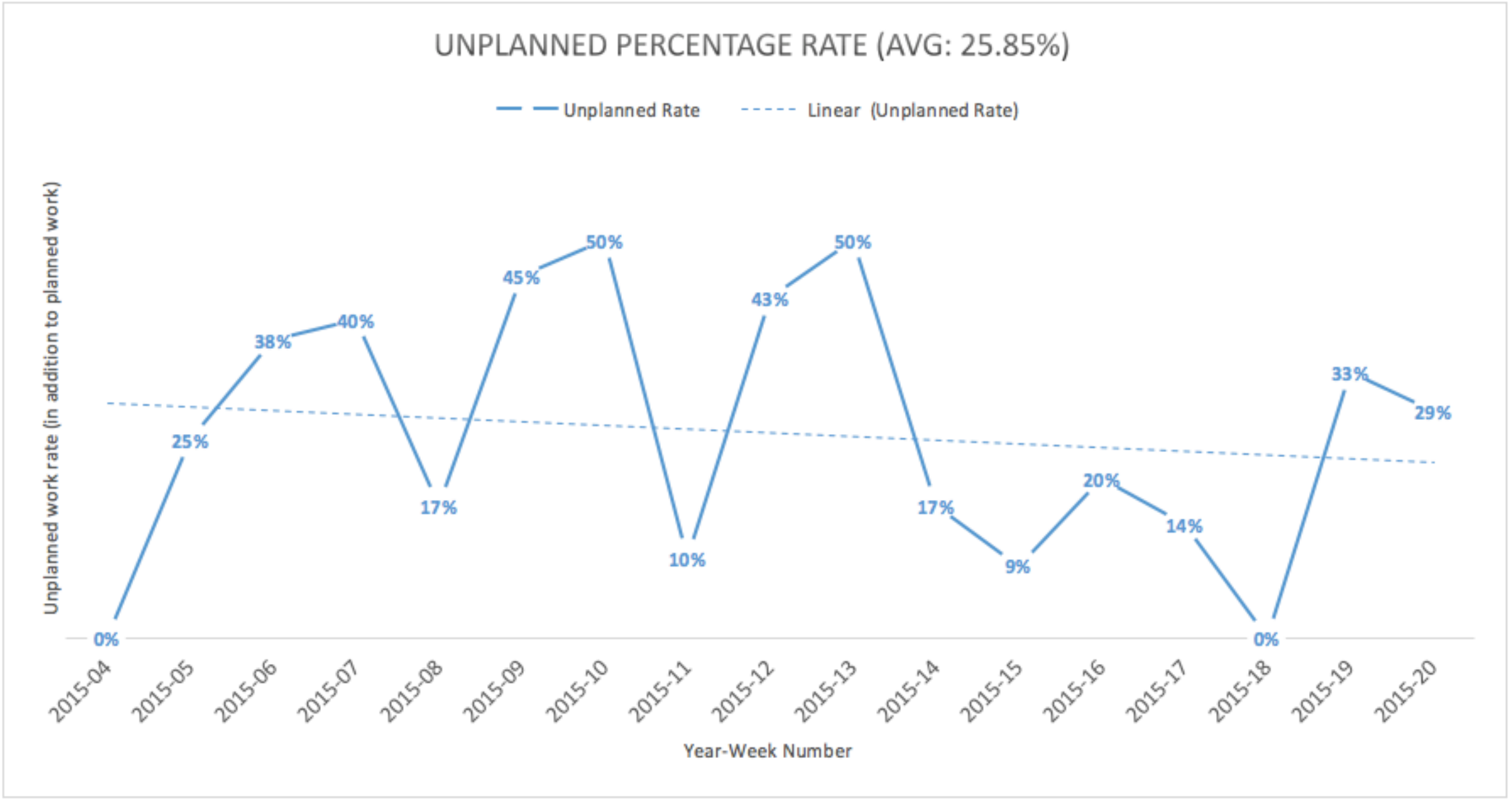
Un-planned work Percentage

Cycle Time Distribution Fitting

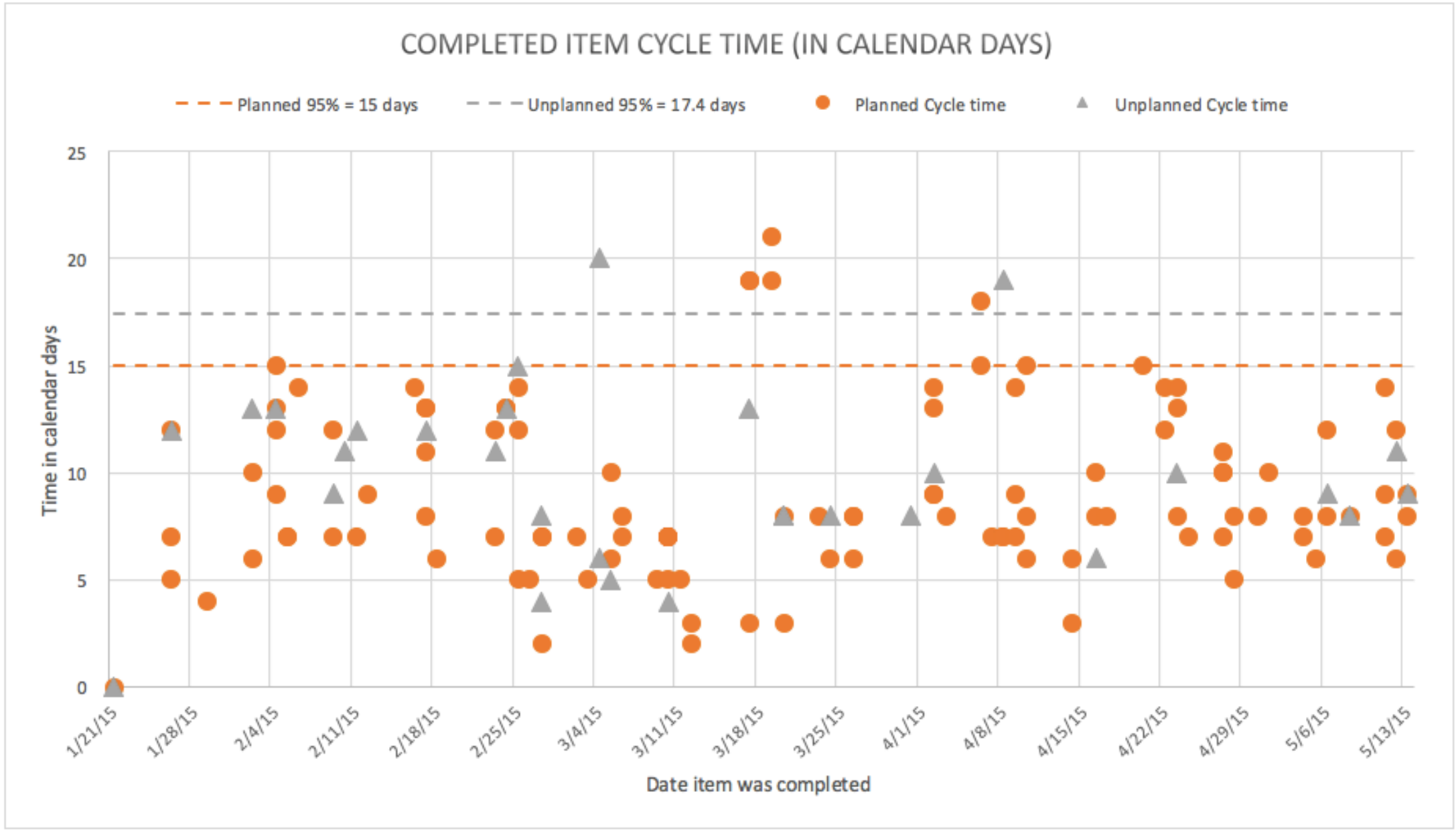
http://bit.ly/Throughput



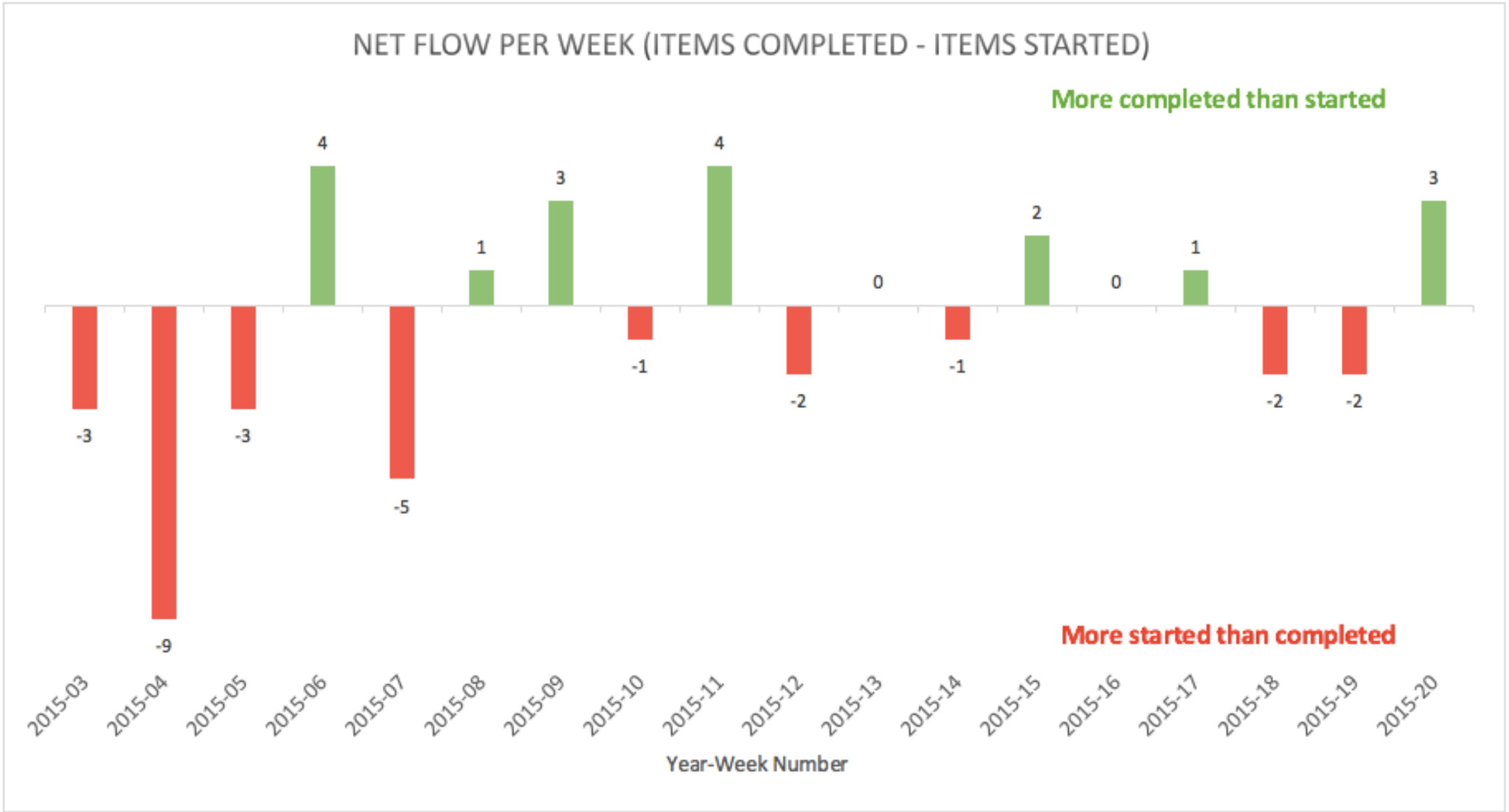
http://bit.ly/Throughput



http://bit.ly/Throughput



http://bit.ly/Throughput



1930 to 2012
National League MVP
23% = 19 out of 82 (last time 1988 - Carston)

1930 to 2013
All-American League MVP
19 out of 82 (last time 1984 - Hernandez)

Source: ESPN Playbook - SportsData (infographic at end of this deck)

1955-56 to 2015-16
NBA MVP
37% = 23 out of 62 (last time 2014 – Curry)

Source: NBA Most Valuable Player Award. (2016, June 24). In *Wikipedia, The Free Encyclopedia*. Retrieved 18:28, July 3, 2016, from https://en.wikipedia.org/w/index.php?title=NBA_Most_Valuable_Player_Award&oldid=726766319


Hope Solo
Team USA


You can’t WIN by just having the BEST goalkeeper
A 0-0 draw is the best you can HOPE
The goalkeeper can’t win SOLO
SOLO
1
OGIMI
17

Team versus individual improvement

- As professionals, we are expected to know our jobs
  - Just like in sports, NBA Kobe Bryant is expected to already know core skills.
  - We are expected to know our strengths and weaknesses

- Coaches and managers for professional teams deal more with
  - Balancing the skills available versus needed
  - Helping individuals work as a team in an effective way
  - Help the team focus on improvements based on recent performance

http://bit.ly/CapabilityMatrix

Team Name: _____ Your Name: _____

For each capability choose from the list of CURRENT skill level values. If in doubt, err low (left)!

	Know nothing	Can run and use the tools needed	Can tweak it or do easy bug fixes	Can start from nothing and create	Can teach others
CSS	☐	☐	☐	☐	☐
Javascript	☐	☐	☐	☐	☐
DB Backup/Restore	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐
	☐	☐	☐	☐	☐

Ready to Learn

Doers

Teachers

http://bit.ly/CapabilityMatrix

	Create Video Content	Create Written Content	Using Tableau	Penopto Admin	Penopto Content	Review & Consolidate Feedback	My Search	Sharepoint Admin	Sharepoint Content	Camtasia	Scrum-master'ing	M
Captains: Ability	5	11	5	3	5	9	4	2	8	3	4	
Players: Ability tc	9	11	9	6	8	9	8	6	8	5	6	
Bench: Ready to	1	0	2	0	0	1	1	1	2	3	2	
Person 1 - red	2	4	2	0	2	4	2	3	4	3	3	
Person 2 - blue	2	3	3	2	2	3	2	2	3	1	2	
Person 3	3	3	4	3	3	4	4	4	4	2	3	
Person 4	4	4	2	2	4	4	3	2	4	4	1	
Person 5 - blue	2	3	2	0	0	3	1	0	3	0	0	
Person 6	4	4	4	2	4	0	2	2	3	0	0	
Person 7 - red	4	3	4	4	4	3	3	1	1	4	4	
Person 8	1	4	3	0	0	3	3	0	1	1	2	
Person 9	4	4	2	4	4	4	2	2	3	1	3	
Person 10	2	3	1	0	2	3	0	0	3	2	1	
Person 11 - blue	0	4	1	0	0	1	0	0	0	0	0	
Person 12	0	0	0	0	0	0	0	0	0	0	0	
	28	39	28	17	25	32	22	16	29	18	19	

Urgency (Redder = more urgent)

	Teachers		
Doers	0	1	2+
0	9	7	3
1	8	5	2
2+	6	4	1

If skill is growing in demand, prepare the bench strength -

	Teachers		
Novices	**0**	**1**	**2+**
0	9	7	3
1	8	5	2
2+	6	4	1

Goals -

Have 2+ people who are Doer's for each skill on the team. If creating new innovations, have at least 1 teacher for each skill.

If a skill is in demand, have at least 1 (preferably 2) teachers on the team (or available), and know who is willing (or able) to be a novice in training to doer.

Know what skills might be needed elsewhere in the company, as your team members might be pulled off at short notice.

Know what skills might be needed to fix incoming defects or production issues when rolling to customer usage.

Know how long (and plan to reduce) the onboarding time from novice to doer levels, prioritized by the skills most anticipated in need for the future.

Its *not* a goal to have everyone at Teacher level for every skill: Your goal is to have a resiliant team given un-planned disruptions and the next feature demands


Find balance…
In changing conditions
And competing forces

Balanced competing metrics

- If you show just one metric, it will be hit…
  - At the expense of everything else

- Coaching is about seeing the bigger picture
- Coaching is about getting the team to recognize competing forces
- Coaching is about teach the team to make smart trades
- Coaching is about little adjustments
- Coaching is teaching how to adapt to changing conditions/pressures

1. Quality
(how well)

- Escaped defect counts
- Forecast to complete defects
- Measure of release “readiness”
- Test count (passing)

2. Productivity
(how much, delivery pace)

- Throughput
- Velocity
- Releases per day

3. Responsiveness
(how fast)

- Lead time
- Cycle time
- Defect resolution time

4. Predictability
(how repeatable)

- Coefficient of variation (SD/Mean)
- Standard deviation of the SD
- “Stability” of team & process

Quality (how well)

Productivity (how much)

Predictability Responsiveness (how repeatable) (how fast)

[redacted] Team Historical Agile Diagnostic Dashboard

Select your team...

Date range to display...
10/1/2014 12:00:00 AM to 3/31/2015 11:59:59 PM

ZBB - Days to Close Active Bugs

Dev Days to Zero Bugs: 138.1

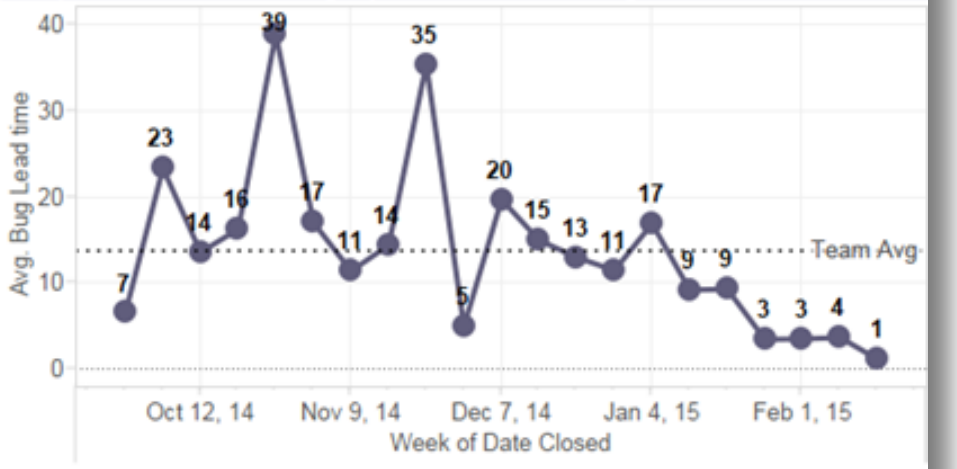
27 bugs would be resolved by 22 [redacted] developers in 6.28 days

Override Bug Count or 0
0

Override Dev Count or 0
0

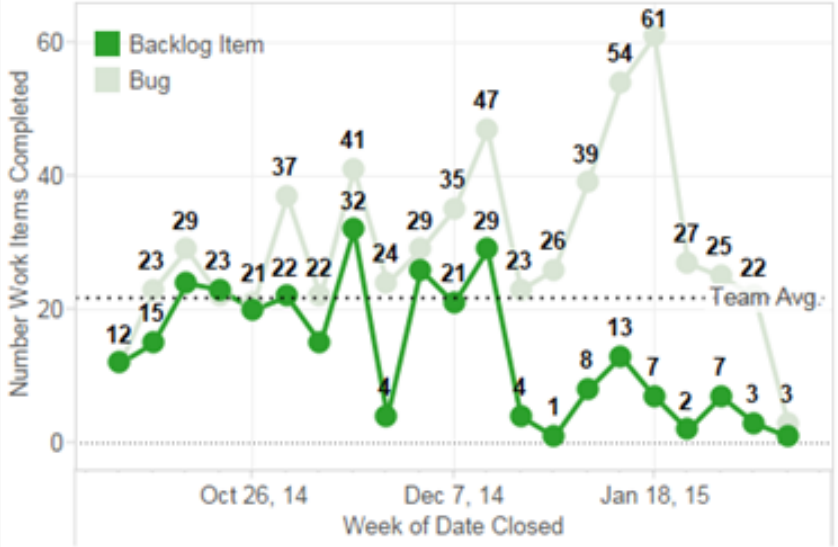
Avg forecast using recent bug cycle-time data. ***Lower is better.***

Responsiveness - Bug cycle-time average



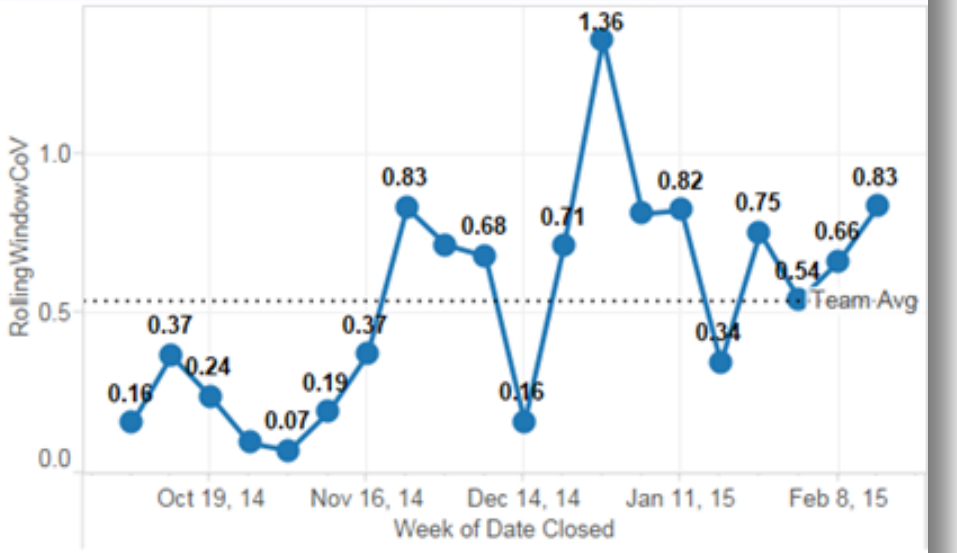
How long it takes from opened to resolved for bugs. ***Lower is better.***

Throughput - Close rate of work items



How many work items have been closed. ***Higher is better.***

Predictability - Consistency of delivery pace



How variable is work throughput. ***Lower is better.***

Quality

“If OUR entire TEAM did nothing else but fix bugs this sprint, at OUR historical rate, we would have x days of work”

- Goal is to keep the TEAMS within 10 days of releasable
- Forecast has to be personal for the team
- Days = $\frac{\text{Open Bugs x Avg(recent cycle time samples)}}{\text{Number of Devs on team}}$

Coaching Advice

Compare "my" team

All Teams Agile Dashboard - Your Trend vs Others

Select your team...

Other Teams
Selected Team

Other Teams
Selected Team

Show lines
True

Date range to display...
10/1/2014 12:00:00 AM to 3/31/2015 11:59:59 PM

Three ways to decrease bug counts and cycle time -
1. Triage bugs quickly. Set them to P1 (fix immediately), P2 (fix as soon as possible), or defer them.

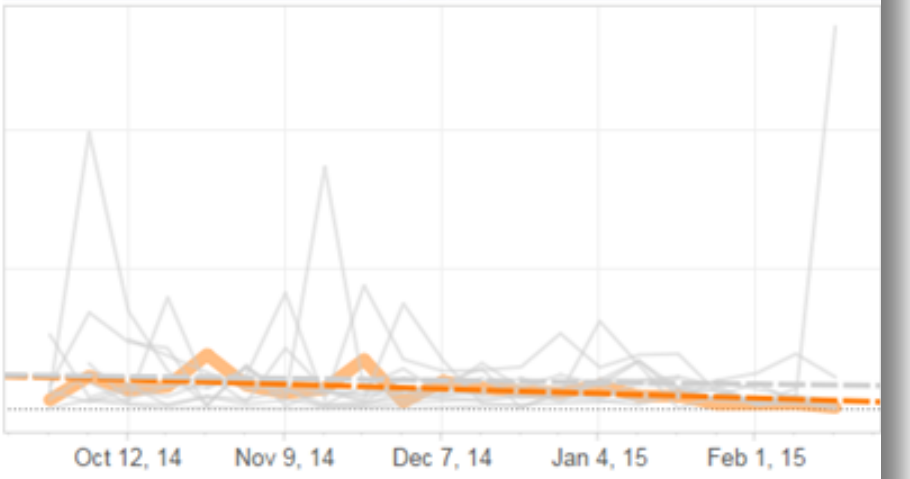
2. Share expert knowledge. Consider having the "expert" who would normally be assigned a defect in a code area lightly assist someone else – now you have an expert in training.

3. Before calling code complete, demo the software to the product owner and testers. This helps obvious defects being found later (and means you don't get disturbed six months from now).

Three ways to increase and stabilize work item throughput –
1. Stop starting, start finishing, stop starting. Avoid starting every story on day one of the sprint only to have everything ALMOST done at the end of the sprint.

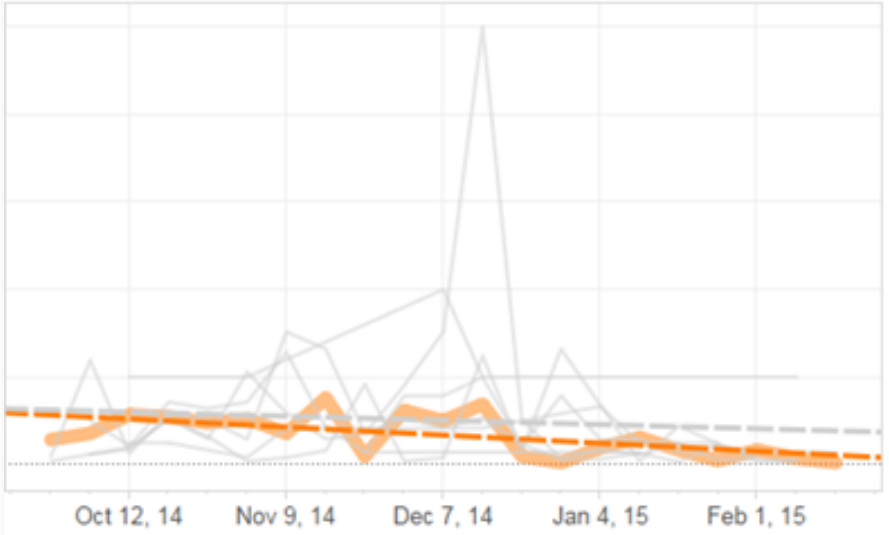
2. Get early feedback on your work from the product owner and testers. This early feedback will avoid bugs and mis-understandings that inhibit "Complete". ..

Responsiveness All - Bug cycle-time average



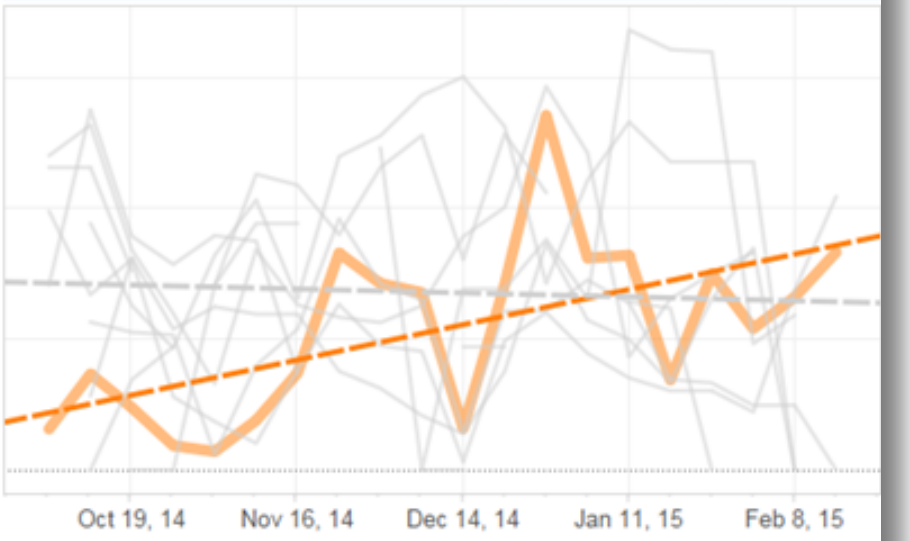
How long it takes from opened to resolved for bugs. ***Lower is better.***

Throughput All - Close rate of work items / Devs



How many work items have been closed. ***Higher is better.***

Predictability All - Consistency of delivery pace



How variable is work throughput. ***Lower is better.***

All Teams Agile Dashboard - Your Trend vs Others

Select your team...

Other Teams
Selected Team

Other Teams
Selected Team

Show lines
False

Date range to display...
10/1/2014 12:00:00 AM to 3/31/2015 11:59:59 PM

Three ways to decrease bug counts and cycle time -
1. Triage bugs quickly. Set them to P1 (fix immediately), P2 (fix as soon as possible), or defer them.

2. Share expert knowledge. Consider having the "expert" who would normally be assigned a defect in a code area lightly assist someone else – now you have an expert in training.

3. Before calling code complete, demo the software to the product owner and testers. This helps obvious defects being found later (and means you don't get disturbed six months from now).

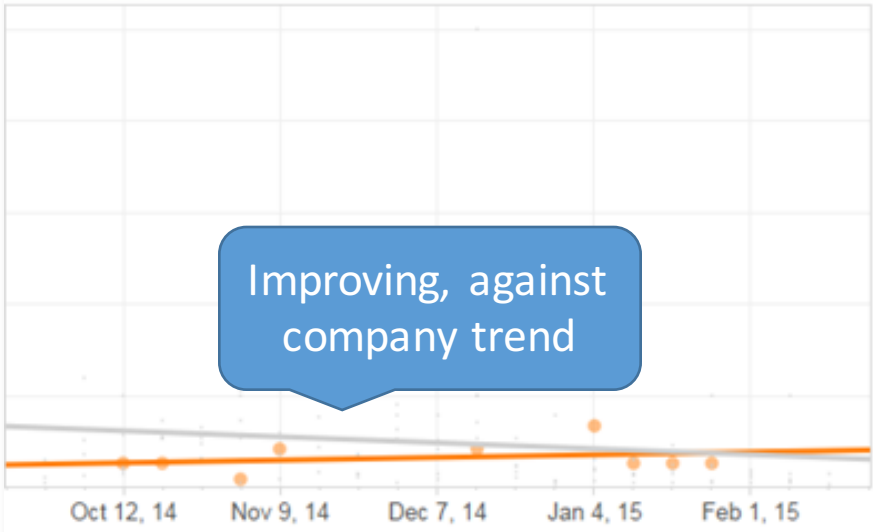
Three ways to increase and stabilize work item throughput –
1. Stop starting, start finishing, stop starting. Avoid starting every story on day one of the sprint only to have everything ALMOST done at the end of the sprint.

2. Get early feedback on your work from the product owner and testers. This early feedback will avoid bugs and mis-understandings that inhibit "Complete". ..

Responsiveness All - Bug cycle-time average

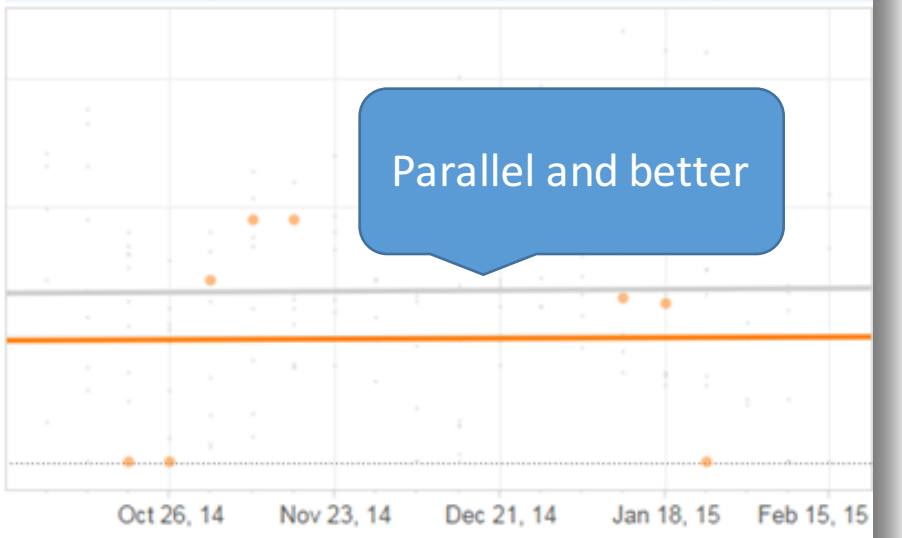
How long it takes from opened to resolved for bugs. ***Lower is better.***

Throughput All - Close rate of work items / Devs



How many work items have been closed. ***Higher is better.***

Predictability All - Consistency of delivery pace



How variable is work throughput. ***Lower is better.***

All Teams Agile Dashboard - Your Trend vs Others

Select your team...

Other Teams
Selected Team

Other Teams
Selected Team

Show lines
False

Date range to display...
10/1/2014 12:00:00 AM to 3/31/2015 11:59:59 PM

Three ways to decrease bug counts and cycle time -
1. Triage bugs quickly. Set them to P1 (fix immediately), P2 (fix as soon as possible), or defer them.

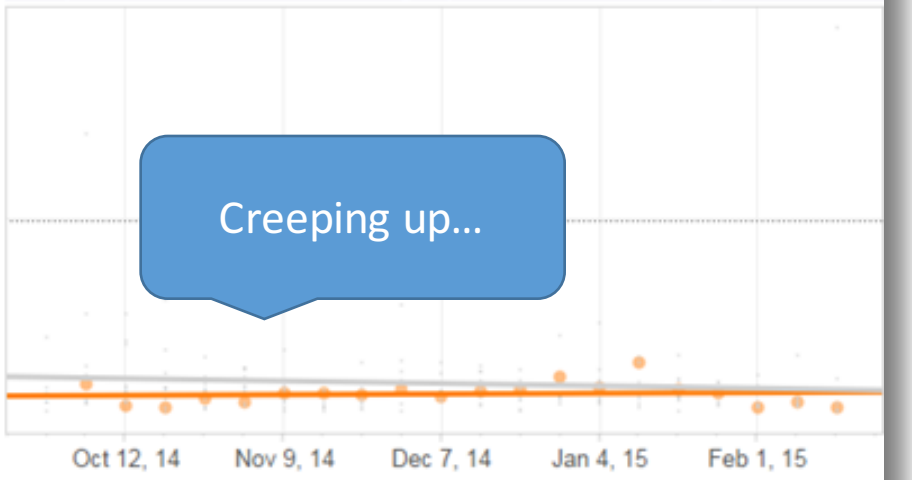
2. Share expert knowledge. Consider having the "expert" who would normally be assigned a defect in a code area lightly assist someone else – now you have an expert in training.

3. Before calling code complete, demo the software to the product owner and testers. This helps obvious defects being found later (and means you don't get disturbed six months from now).

Three ways to increase and stabilize work item throughput –
1. Stop starting, start finishing, stop starting. Avoid starting every story on day one of the sprint only to have everything ALMOST done at the end of the sprint.

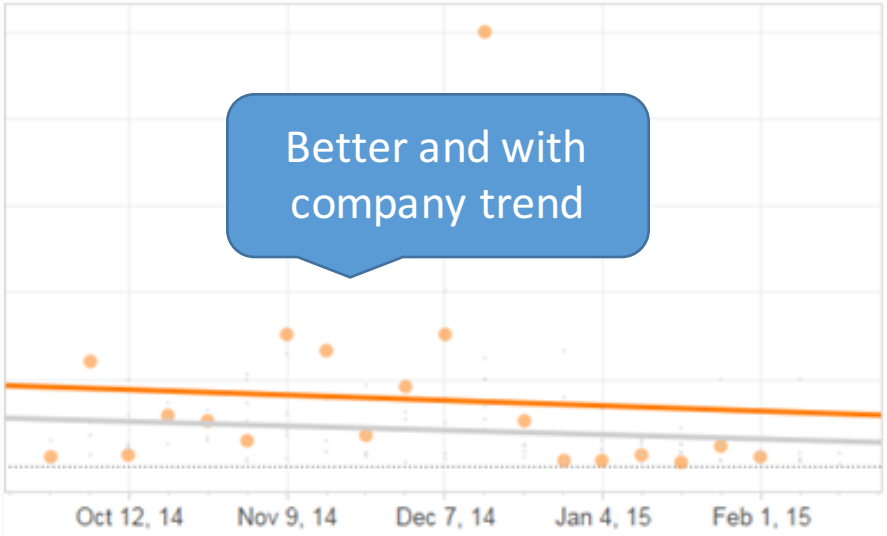
2. Get early feedback on your work from the product owner and testers. This early feedback will avoid bugs and mis-understandings that inhibit "Complete". ..

Responsiveness All - Bug cycle-time average



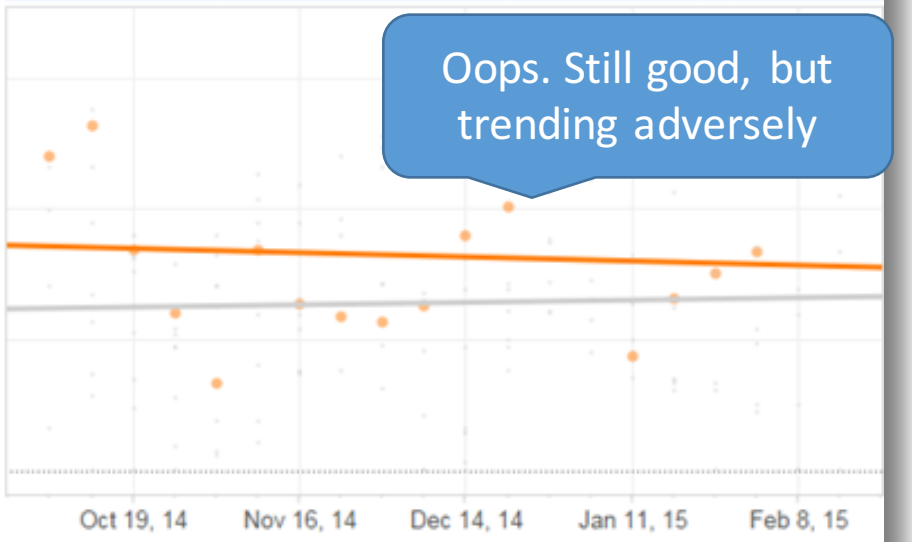
How long it takes from opened to resolved for bugs. ***Lower is better.***

Throughput All - Close rate of work items / Devs



How many work items have been closed. ***Higher is better.***

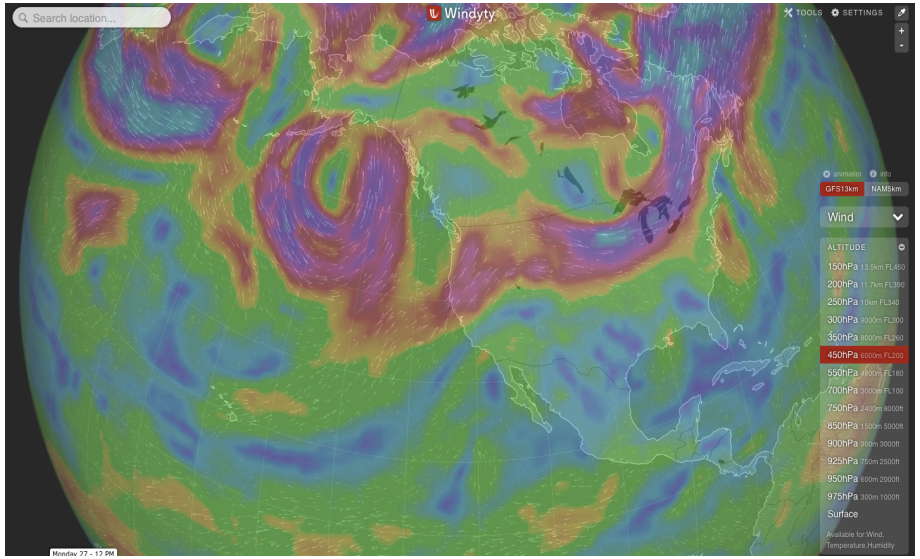
Predictability All - Consistency of delivery pace



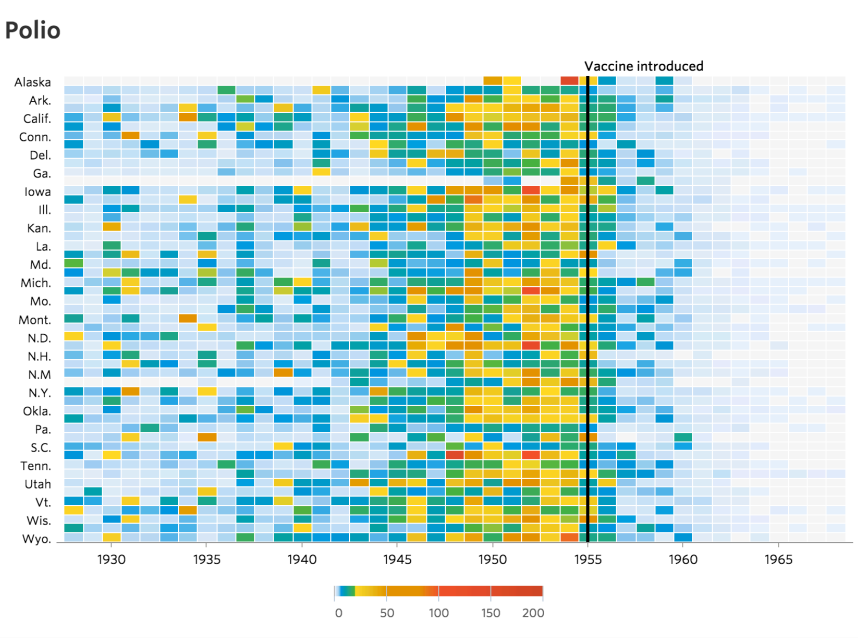
How variable is work throughput. ***Lower is better.***

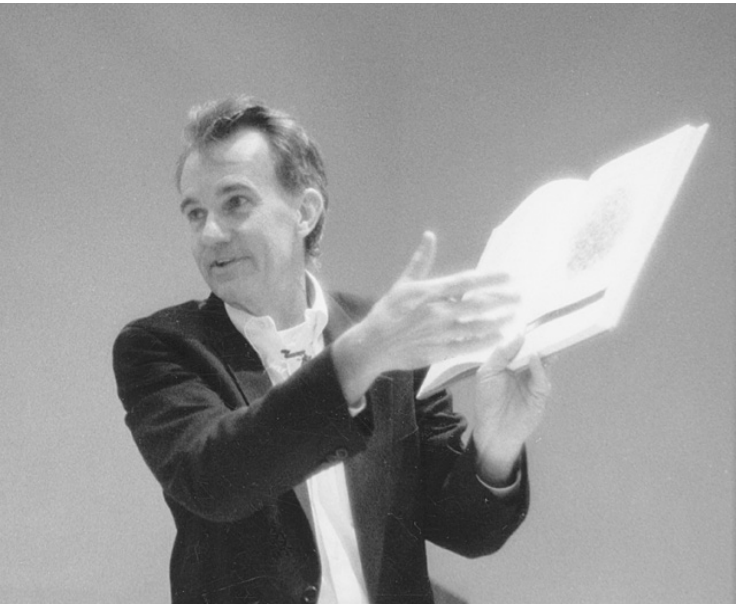
Don’t Make it Personal

Beautiful + Engaging

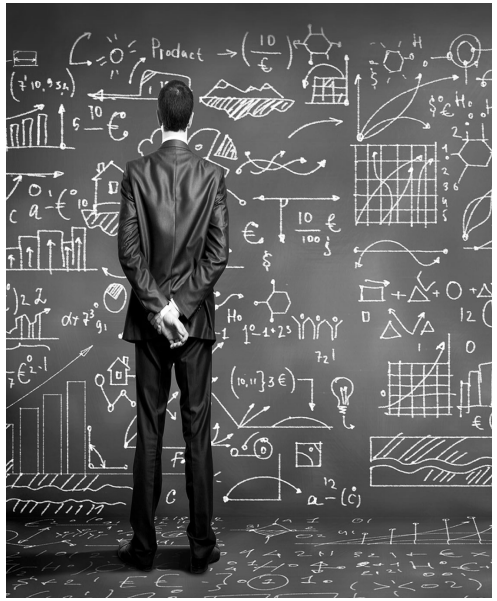


Tell a Story

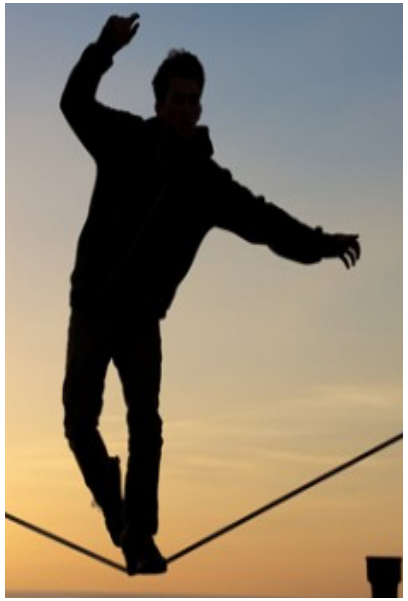




Compared to What



Keep it Simple



Balanced Metrics

1. Quality
(how well)

- Escaped defect counts
- Forecast to complete defects
- Measure of release “readiness”
- Test count (passing)

3. Responsiveness
(how fast)

- Lead time
- Cycle time
- Defect resolution time

2. Productivity
(how much, delivery pace)

- Throughput
- Velocity
- Releases per day

4. Predictability
(how repeatable)

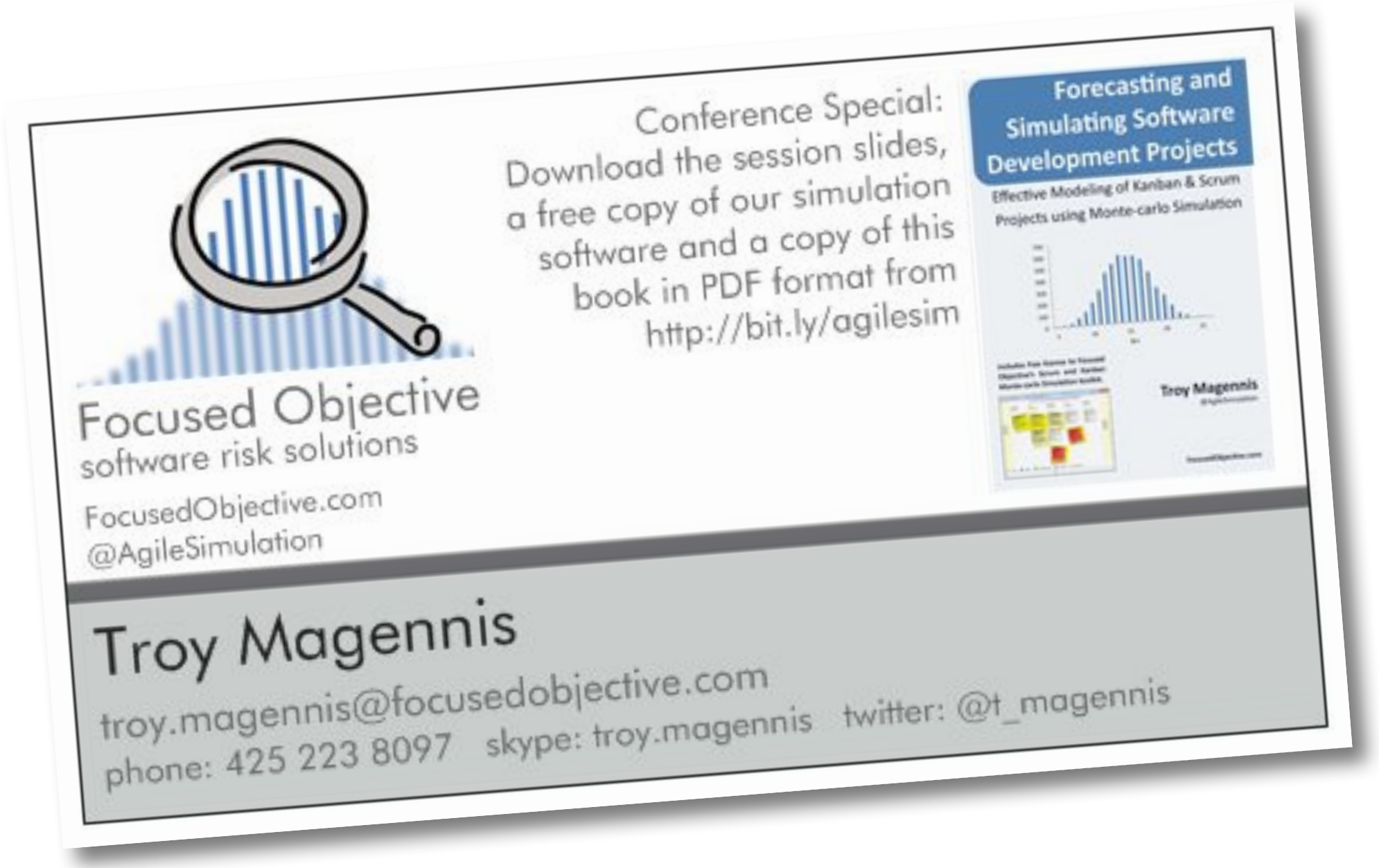
- Coefficient of variation (SD/Mean)
- Standard deviation of the SD
- “Stability” of team & process

Make GREAT tradeoff Decisions

@t_magennis

Troy.Magennis@FocusedObjective.com

Please consider doing the review


Focused Objective
software risk solutions
FocusedObjective.com
@AgileSimulation
Conference Special:
Download the session slides,
a free copy of our simulation
software and a copy of this
book in PDF format from
http://bit.ly/agilesim
Forecasting and Simulating Software Development Projects
Effective Modeling of Kanban & Scrum Projects using Monte-carlo Simulation
Troy Magennis
Troy Magennis
troy.magennis@focusedobjective.com
phone: 425 223 8097
skype: troy.magennis
twitter: @t_magennis

Forecasting and Risk

Helping teams see and understand risk impacts

Q. Could I make a simple forecast tool that worked?

Without macros or add-ins!

http://bit.ly/ThroughputForecast

Or follow @t_magennis

http://bit.ly/ThroughputForecast

Forecast Completion Date

1. Start Date 4/1/15

2. How many stories are remaining to be completed?

(enter the range estimate of stories. Tip: start wide and narrow as certainty increases)

Low guess 20 Highest guess 30

3. Stories are often split before and whilst being worked on. Estimate the split rate low and high bounds.

(often the throughput in the backlog is pre-split, but captured throughput post-split. Adjust for this here)

Low guess 1.00 Highest guess 1.00

4. Throughput. How many completed stories per week or sprint do you estimate low and high bounds?

Throughput estimate/samples are per Week 7 days

Use historical throughput data **OR** enter a low and high estimate below. Use: Estimate

Low guess 1 Highest guess 5

Can I use velocity rather than throughput?

Yes. If you do have estimates in story points, then you can sume all of the estimates and use that for input 2 and estimate or use historical team velocity for input 4. The benefit of using throughput (count of

http://bit.ly/ThroughputForecast

Results

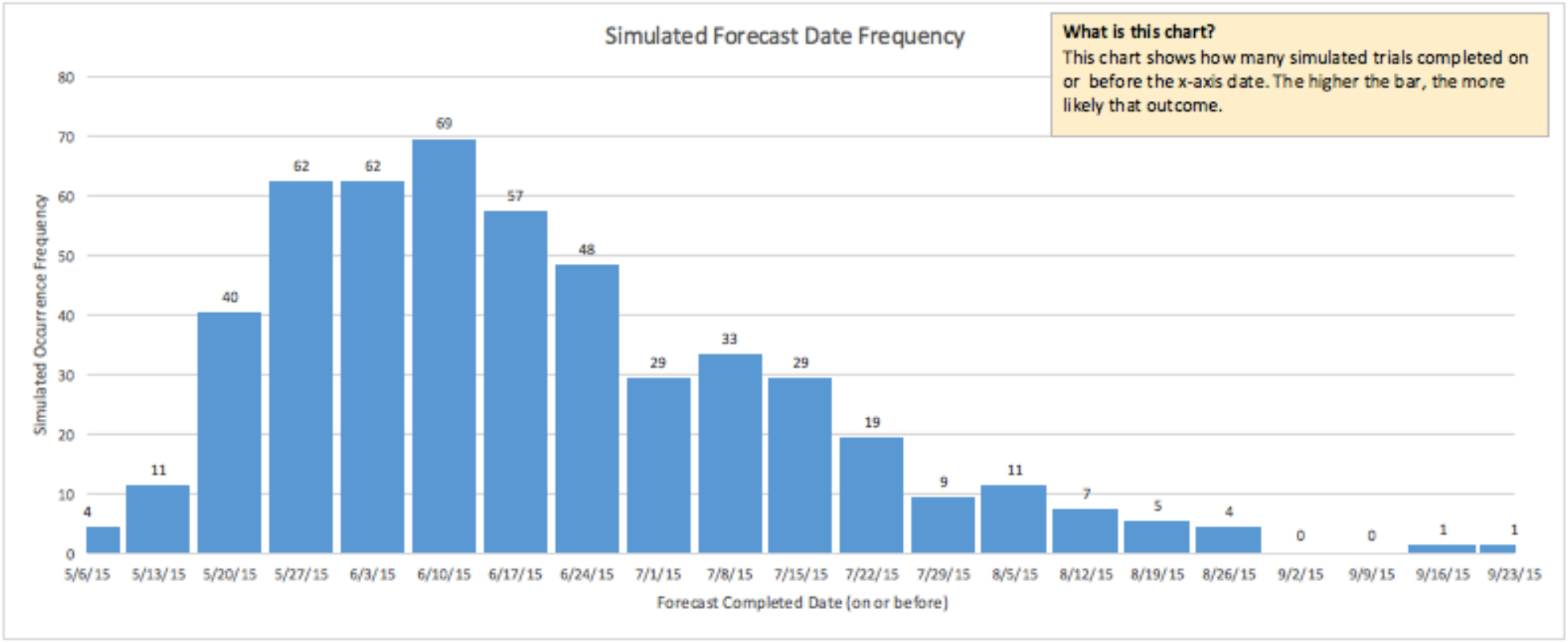
Likelihood	Duration in Week's	Date
100%	25	9/23/15
95%	18	8/5/15
90%	16	7/22/15
85%	15	7/15/15
80%	14	7/8/15
75%	13	7/1/15
70%	12	6/24/15
65%	12	6/24/15
60%	11	6/17/15
55%	11	6/17/15
50%	11	6/17/15
45%	10	6/10/15
40%	10	6/10/15
35%	9	6/3/15
30%	9	6/3/15
25%	9	6/3/15
20%	8	5/27/15
15%	8	5/27/15
10%	7	5/20/15
5%	7	5/20/15

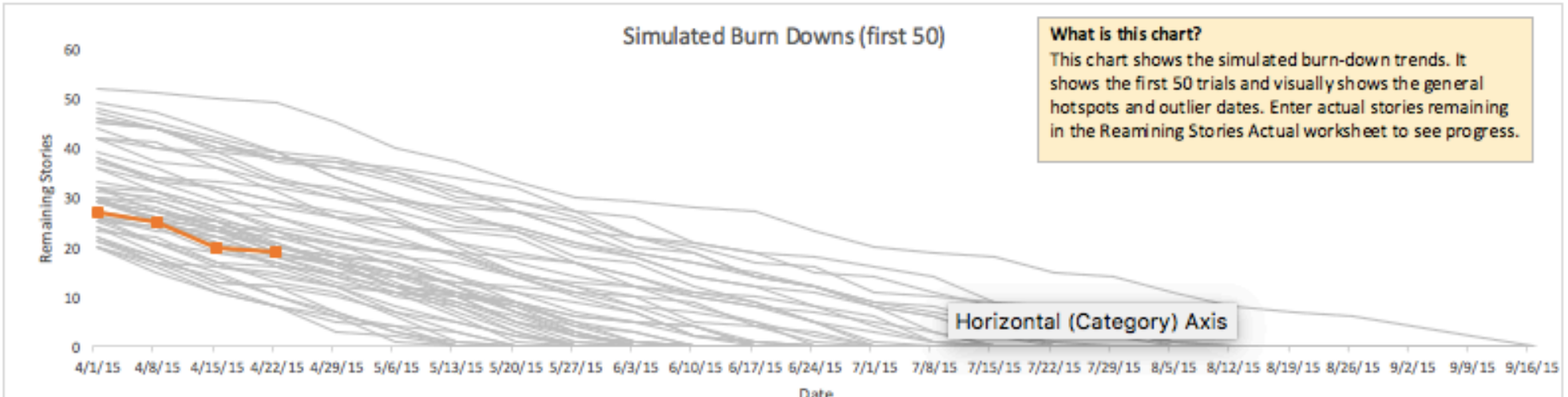
Almost certain

Somewhat certain

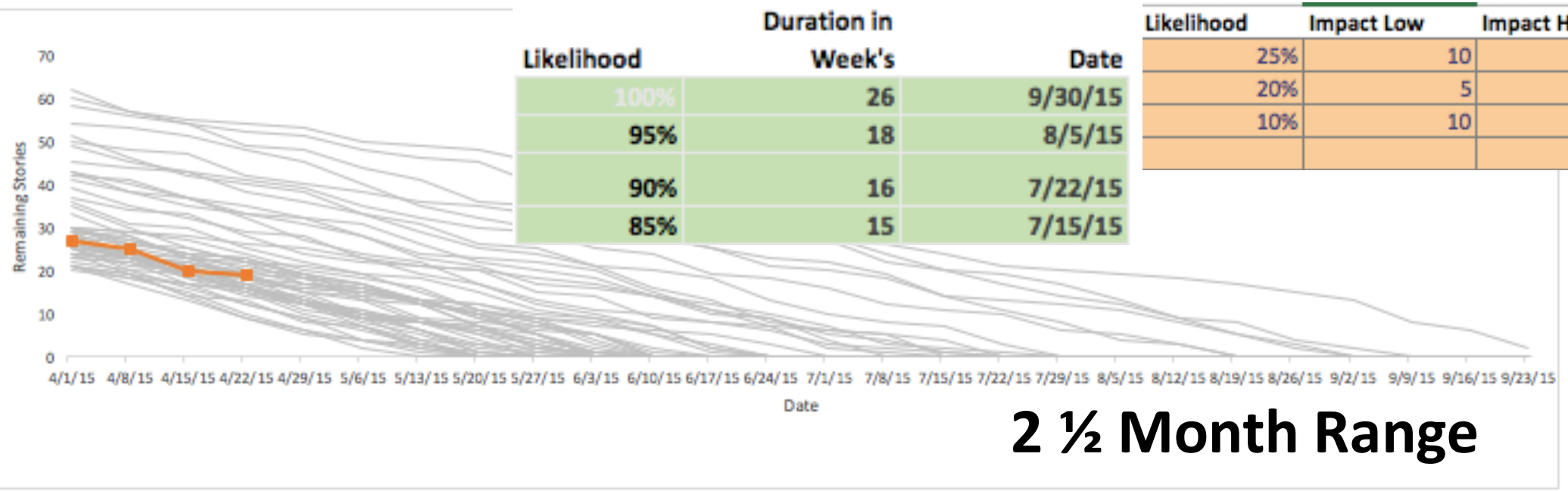
Less than coin-toss odds. But if you are game?

http://bit.ly/ThroughputForecast





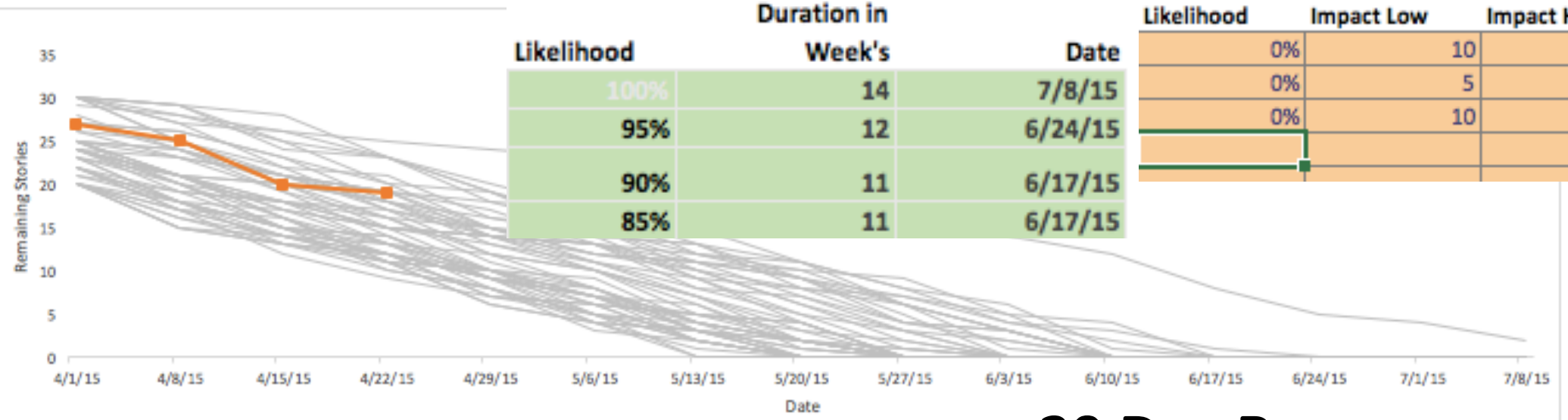
http://bit.ly/ThroughputForecast



Likelihood	Duration in Week's	Date
100%	26	9/30/15
95%	18	8/5/15
90%	16	7/22/15
85%	15	7/15/15

Likelihood	Impact Low	Impact High	Description
25%	10	20	Performance tuning if w
20%	5	10	Bitcoin support?
10%	10	20	PCI Audit remediation

2 ½ Month Range



Likelihood	Duration in Week's	Date
100%	14	7/8/15
95%	12	6/24/15
90%	11	6/17/15
85%	11	6/17/15

Likelihood	Impact Low	Impact High	Description
0%	10	20	Performance tuning if
0%	5	10	Bitcoin support?
0%	10	20	PCI Audit remediation

~ 20 Day Range

Cycle time analysis

How to interpret cycle time distributions in coaching

Q. Can historical cycle-time be used for coaching advice?

2015 48th Hawaii International Conference on System Sciences

The Economic Impact of Software Development Process Choice - Cycle-time Analysis and Monte Carlo Simulation Results

Troy Magennis
troy.magennis@focusedobjective.com

Abstract

IT executives initiate software development process methodology change with faith that it will lower development cost, decrease time-to-market and increase quality. Anecdotes and success stories from agile practitioners and vendors provide evidence that other companies have succeeded following a newly chosen doctrine. Quantitative evidence is scarcer than these stories, and when available, often unverifiable.

This paper introduces a quantitative approach to assess software process methodology change. It proposes working from the perspective of impact on cycle-time performance (the time from the start of individual pieces of work until their completion), before and after a process change.

This paper introduces the history and theoretical basis of this analysis, and then presents a commercial case study. The case study demonstrates how the economic value of a process change initiative was quantified to understand success and payoff.

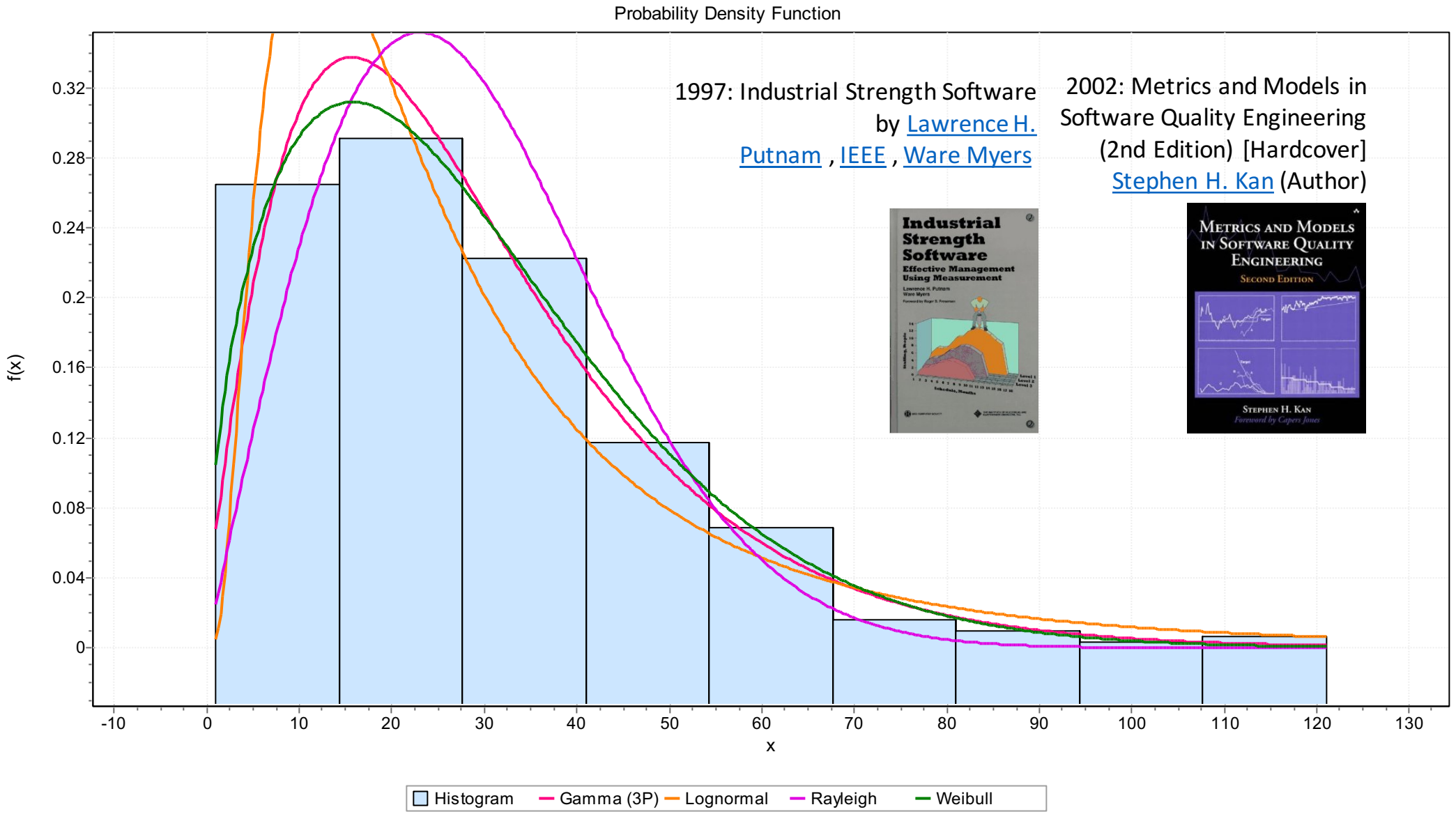
Cycle-time is a convenient metric for comparing proposed and ongoing process improvement due to its easy capture and applicability to all processes. Poor cycle-time analysis can lead to teams being held to erroneous service level expectations. Properly comparing the impact of proposed process change scenarios, modeled using historical or estimated cycle-time performance helps isolate the bottom line impact of process changes with quantitative rigor.

Scrum, and Kanban are some of the well-known processes that have risen to the top of the popularity charts, each with case studies (often just one) showing great impact when applied correctly by the inventors. The final choice appears to fall on faith based lines, with many organizations moving from one process to the next in search of nirvana. A quantitative framework for estimating and assessing true impact is needed for informed decisions.

Measuring the quantitative impact of a software development process change is hard. Measurable change takes weeks or months to evolve, and there is little in the way of control group – change is implemented and the outcome if that change wasn't performed isn't an interesting or easily discernable metric. This paper presents one technique for quantitatively estimating the potential economic outcomes both before and after a change has been implemented.

The basis for the method described here is probabilistically simulating the impact of changes in cycle-time samples from a prior project to a completed project using new methodology. To estimate the potential payoff for a new process, existing cycle-time samples can be discounted by fixed percentage amount to simulate the financial return for hypothetical reductions (10%, 25%, for example). Once change has occurred, actual results can be compared to the predicted data to validate the difference and improve modeling efforts on future initiatives.

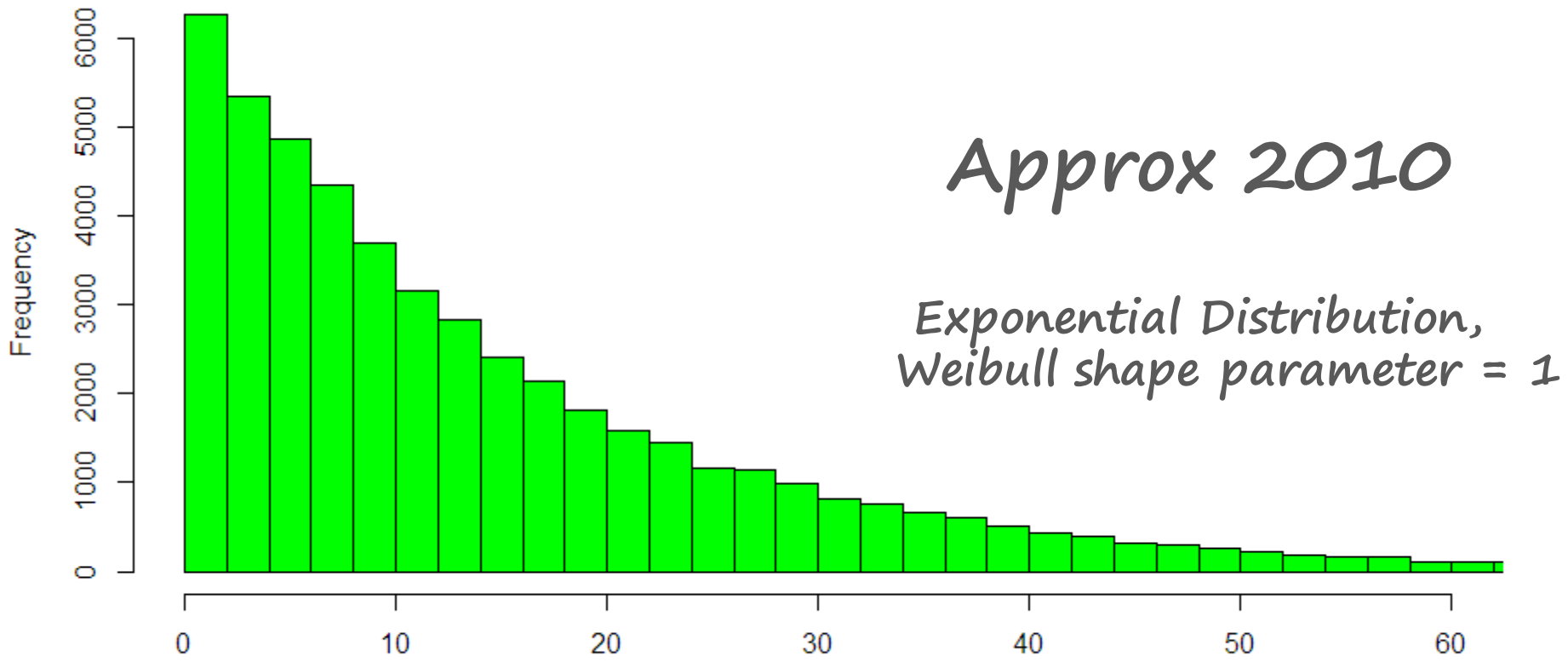
http://conferences.computer.org/hicss/2015/papers/7367f055.pdf



1997: Industrial Strength Software by Lawrence H. Putnam , IEEE , Ware Myers

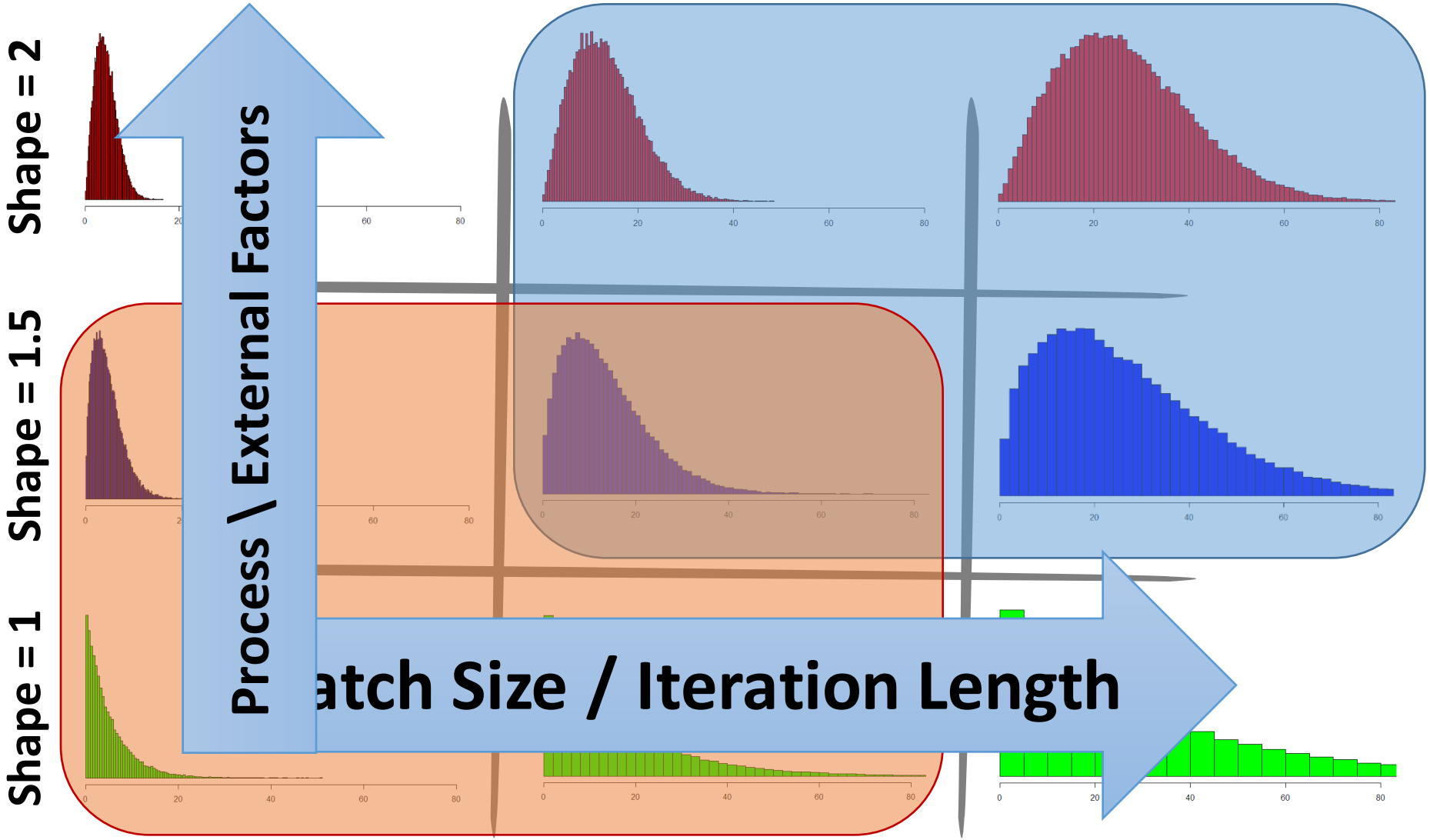
2002: Metrics and Models in Software Quality Engineering (2nd Edition) [Hardcover] Stephen H. Kan (Author)

Paper: http://bit.ly/14eYFM2



Work Item Cycle Time or Lead Time Distribution Through the Ages

Paper: http://bit.ly/14eYFM2

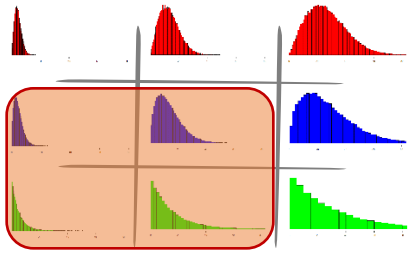


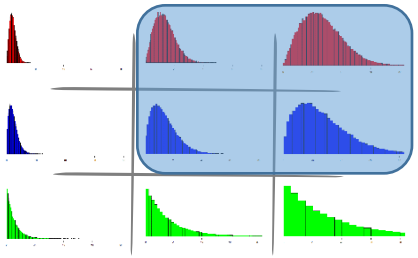
Scale = 5
< 1 week

Scale = 15
~ 2 week sprint

Scale = 30
~ 1 month

Work Item Cycle Time or Lead Time





Lean, Few dependencies

- Higher work item count
- More granular work items
- Lower WIP
- Team Self Sufficient
- Internal Impediments

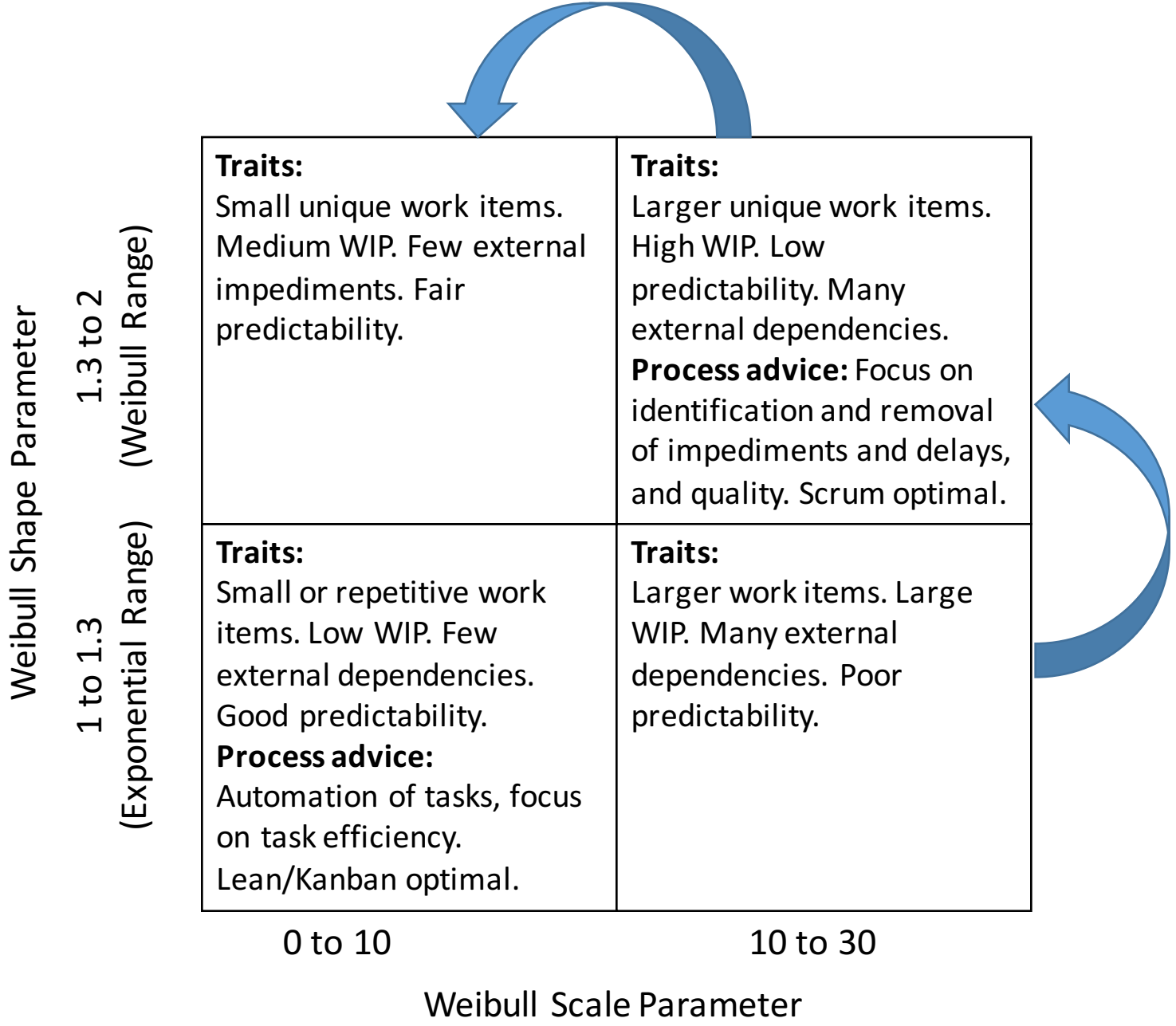
- Do: Automation
- Do: Task Efficiency

Sprint, Many dependencies

- Lower work item count
- Chunkier work items
- Higher WIP
- External Dependencies
- External Impediments

- Do: Collapse Teams
- Do: Impediment analysis

Paper: http://bit.ly/14eYFM2


Weibull Shape Parameter
1.3 to 2 (Weibull Range)
1 to 1.3 (Exponential Range)
Traits:
Small unique work items. Medium WIP. Few external impediments. Fair predictability.
Traits:
Larger unique work items. High WIP. Low predictability. Many external dependencies.
Process advice: Focus on identification and removal of impediments and delays, and quality. Scrum optimal.
Traits:
Small or repetitive work items. Low WIP. Few external dependencies. Good predictability.
Process advice:
Automation of tasks, focus on task efficiency. Lean/Kanban optimal.
Traits:
Larger work items. Large WIP. Many external dependencies. Poor predictability.
0 to 10
10 to 30
Weibull Scale Parameter

References, Sources and Links

Tools

- Excel or Google Sheets Spreadsheets (all free)
  - General metrics spreadsheet (17 charts) –
  - Team Capability Matrix -
  - Forecasting –
  - 10+ other spreadsheets tools all free -
- Visualization Tools
  - Tableau ($995-$1995) – Tableau.com
  - PowerBI (free) –
  - Plotly (free) –
- Online Lean/Kanban Tool
  - Leankit.com

Cool Visualization Resources and Websites

- My blog – FocusedObject.com/blog
- WindyTy.com – weather
- NY Times
- Tableau Public

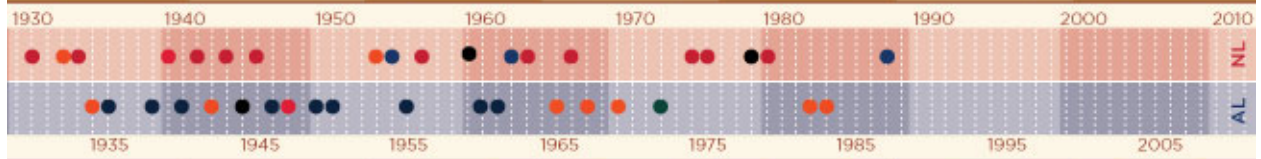
- Books
  - Tufty
  - Few

FALL CLASSICS

MVPs whose teams won the World Series the same year

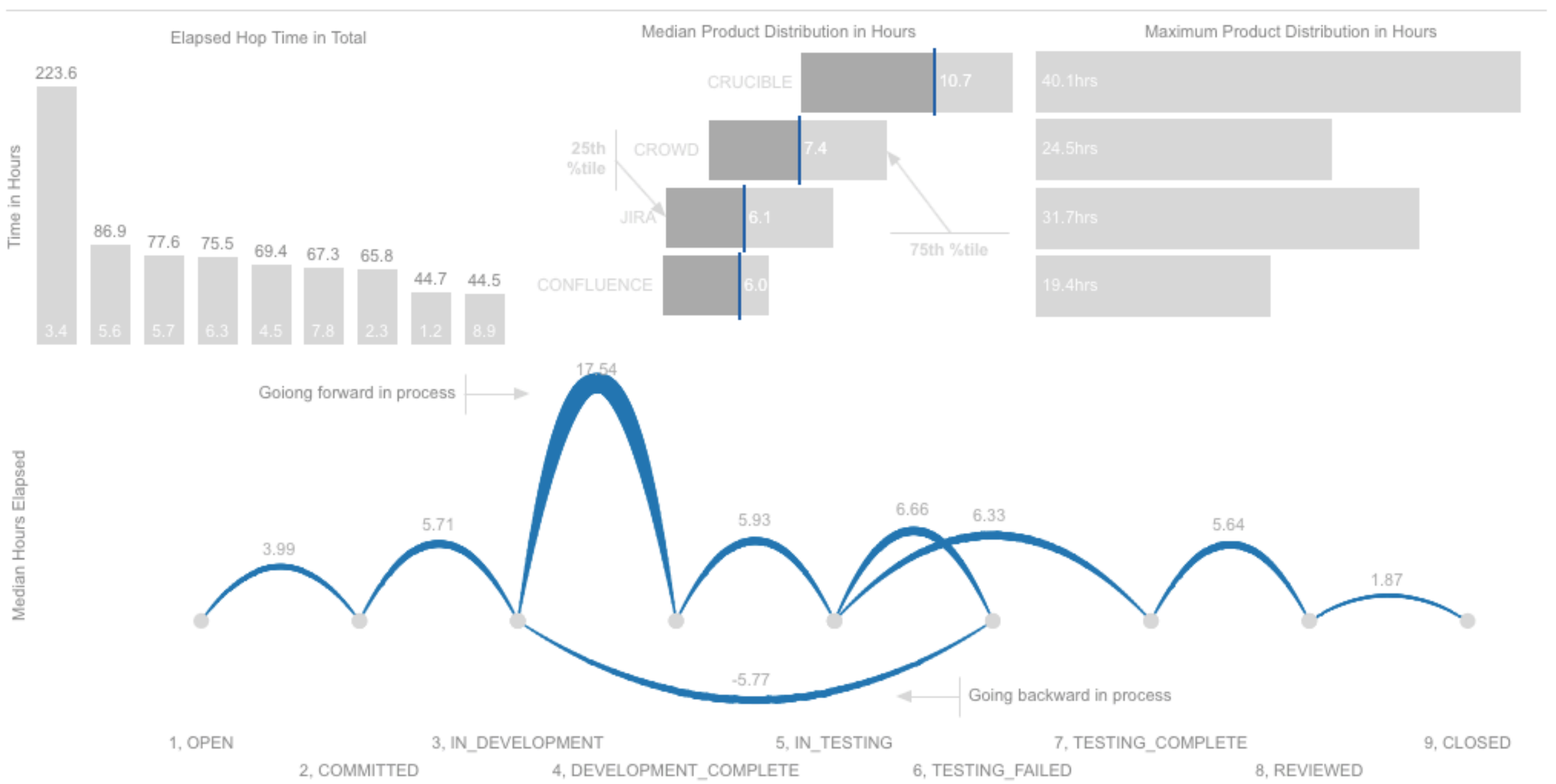
ESPN PLAYBOOK

NATIONAL LEAGUE MVPS



AMERICAN LEAGUE MVPS

SD SPORTSDATA



Source: JumpPlot.com (total kudos to Tom VanBuskirk and Chris DeMartini)

Coaching professional teams

- Is about team performance, not individual
  - If they don’t know it by now, they self improve it
- http://www.landofbasketball.com/awards/nba_season_mvps_year.htm
- 23 championships + MVP / 60 = ~ 1/3
- http://www.nba.com/2011/news/features/04/08/race-to-the-mvp-final-rankings/index.html
- http://national.suntimes.com/nba/7/72/1237030/lebron-james-stephen-curry-nba-finals-mvp

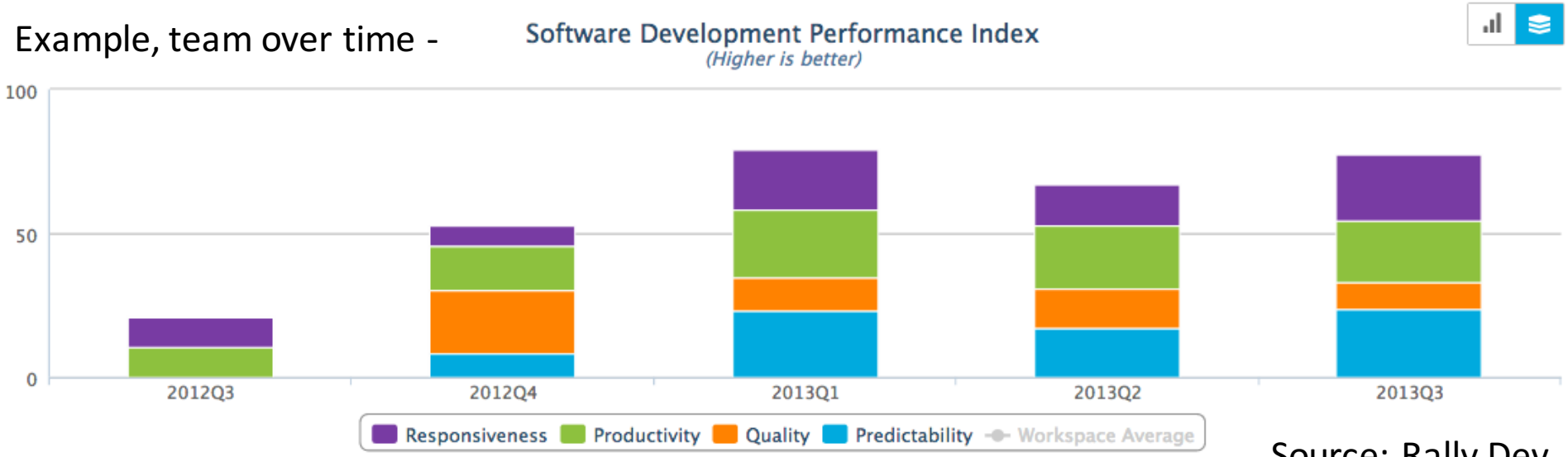
SDPI Dimensions

Responsiveness Productivity Quality Predictability

- Productivity = throughput avg / team size
- Predictability = variability of throughput / size
- Responsiveness = time in process average
- Quality = released defect density / throughput

The Software Development Performance Index
The SDPI framework includes a balanced set of outcome measures. These fall along the dimensions of Responsiveness, Quality, Productivity, Predictability, …

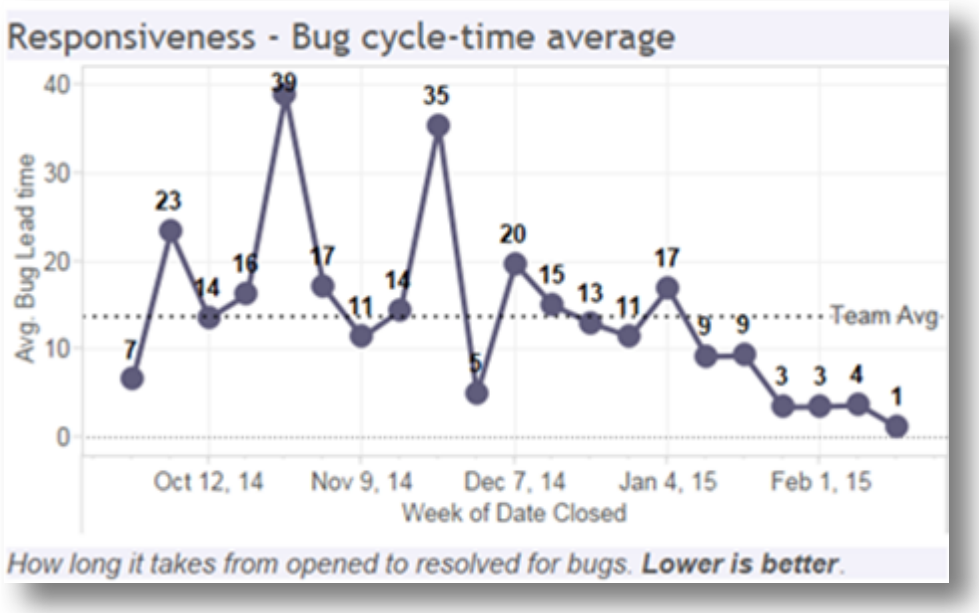
Example, team over time -



Source: Rally Dev.

Responsiveness

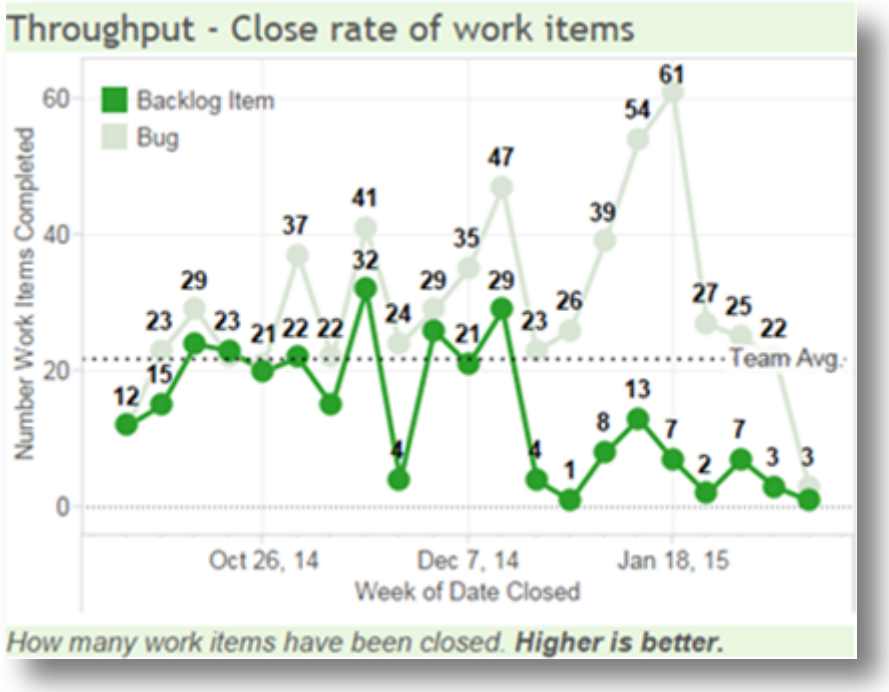
“If something urgent comes along, how fast can we turn that around”



- Average or median of the number of days between two dates for items closed within a period
- Cycle time or Lead time of ???
  - If reliable first touch date, use that
  - If just created date, then use P1 and P2 bug

Completion Rate

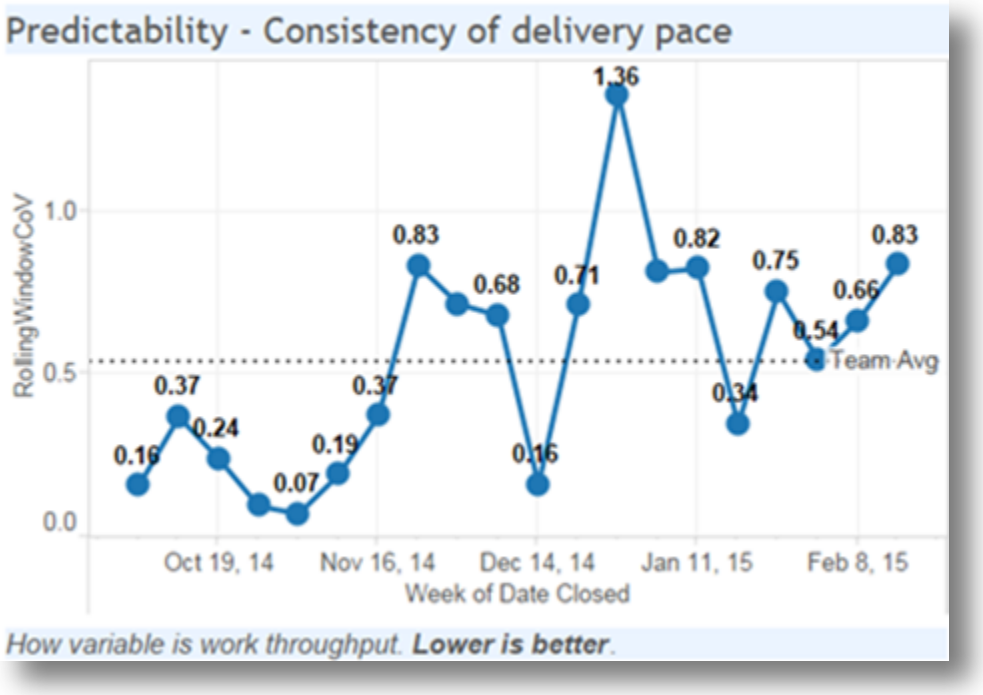
“What is holding us back on completing more. Lets discuss dependencies and blockers in the retrospective”



- Team goal is to maximize number of COMPLETED items, not started items
- Count of items completed each period
- Don’t celebrate bug throughput (as much)

Predictability

“How consistently do we deliver value?”



- How much variation there is each week in throughput, normalized by “team size” in a rough way
- Coefficient of Variation = Mean/SD

Data is evil, but it doesn’t have to be

- Manipulate behavior by
  - Embarrass
  - Co-erce
- Make a point rather than make a difference
  - No action,
  - just data to classify someone is good and someone is bad

- They can tell a story that helps balance and improve where time and energy is best spent
  - Metrics tell a story
  - We learn from the story and make actions
  - Through these action we improve

Purpose of coaching dashboards

- **Improvement – what to change**
  - To help teams identify their weakest area comparable to other teams in similar circumstances
  - To confirm improvement has been achieved after a process change experiment
  - To identify what was traded to achieve that improvement
- **Avoidance – what to watch (sense)**
  - To identify what internal team factors most disrupt team momentum
  - To identify what external factors most disrupt team momentum

Data is un-necessary…

- When there is unlimited time and money, or the journey to a destination is well known, perhaps.
  - But this isn't the most common case
  - And even when data isn't captured on paper, its assumed in peoples heads

- There is always more demand than supply
- There are always insights that are missed
- There is always room for improvement

- You can try and guess. You may be often right, but how do you know?

What makes a good metric?

- **Is relevent to the team or individual** - personalized
  - To compare against others, and to see progress
- **Is within the teams ability to move** (or get moved)
  - Has value in being diagnostic
- **Passively captured** (low cost and effort)
  - Look for cheapest correlated metric to a more costly metric where possible
- **Balances another metric** – demonstrates tradeoff and impacts – trends adversely to another metric
  - Look for cheapest metric that will likely be negatively impacted by movement of another metric