



# Forecasting using data

LEAN AGILE SCOTLAND 2017

Troy Magennis

@t\_magennis | [troy.magennis@focusedobjective.com](mailto:troy.magennis@focusedobjective.com)

All slides and spreadsheets: Bit.ly/SimResources

A road sign with two arrows pointing in opposite directions. The top arrow points straight ahead and is associated with the word "Forecast". The bottom arrow points to the left and is associated with the word "Reality".

Forecast

Reality

*“Remember that all models are wrong;  
the practical question is how wrong do  
they have to be to not be useful.”*

Statistician,  
George Box

@t\_magennis

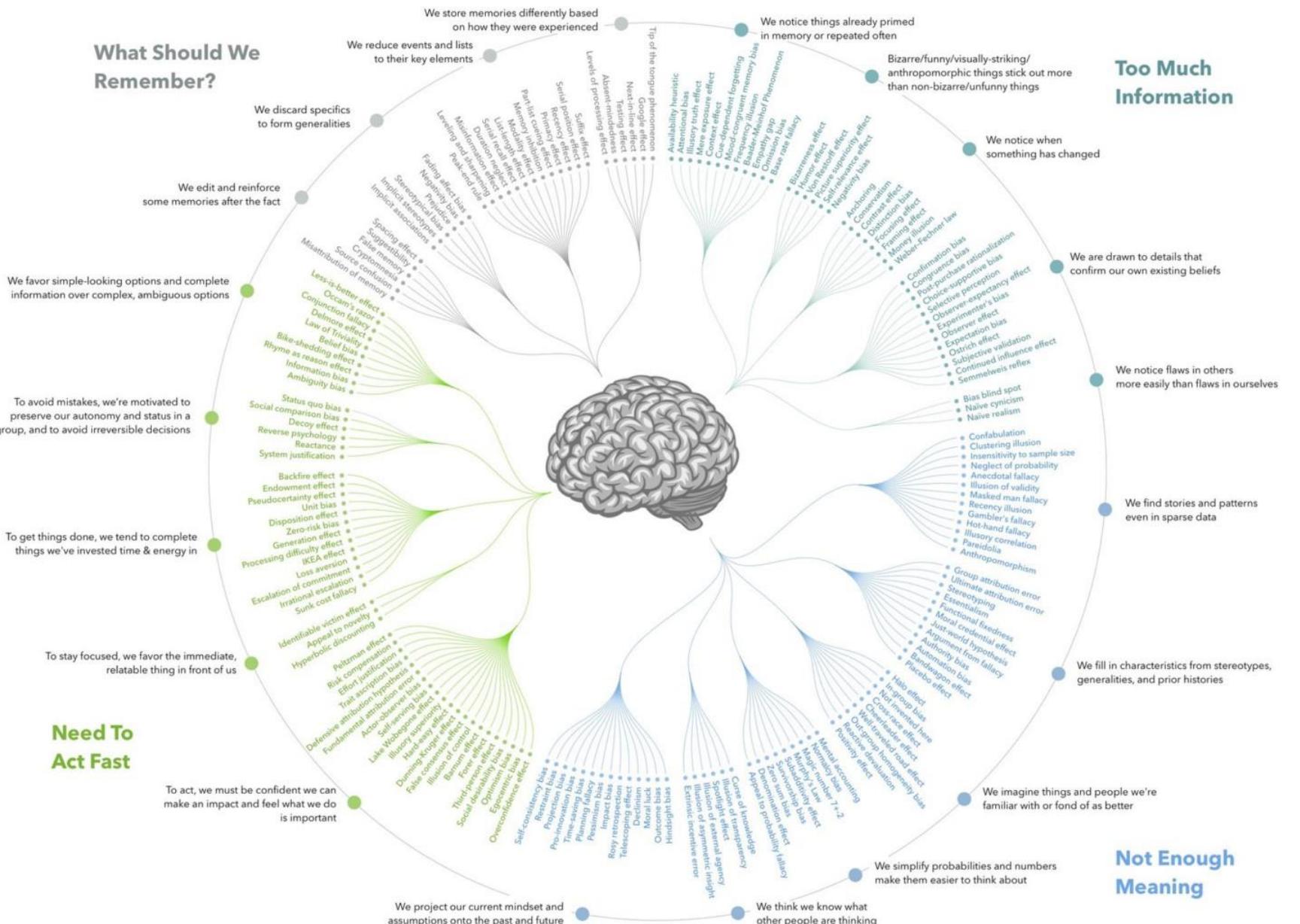


# A. Better than intuition

Not perfect. Not exact. Not always right.

Just better than what you do now, or even equal (just less expensive)

# COGNITIVE BIAS CODEX, 2016



# Too Much Information



- 1. Multiple Options – NOT one...**
- 2. Duration not ETA until commitment...**
- 3. Continuously updated once started...**

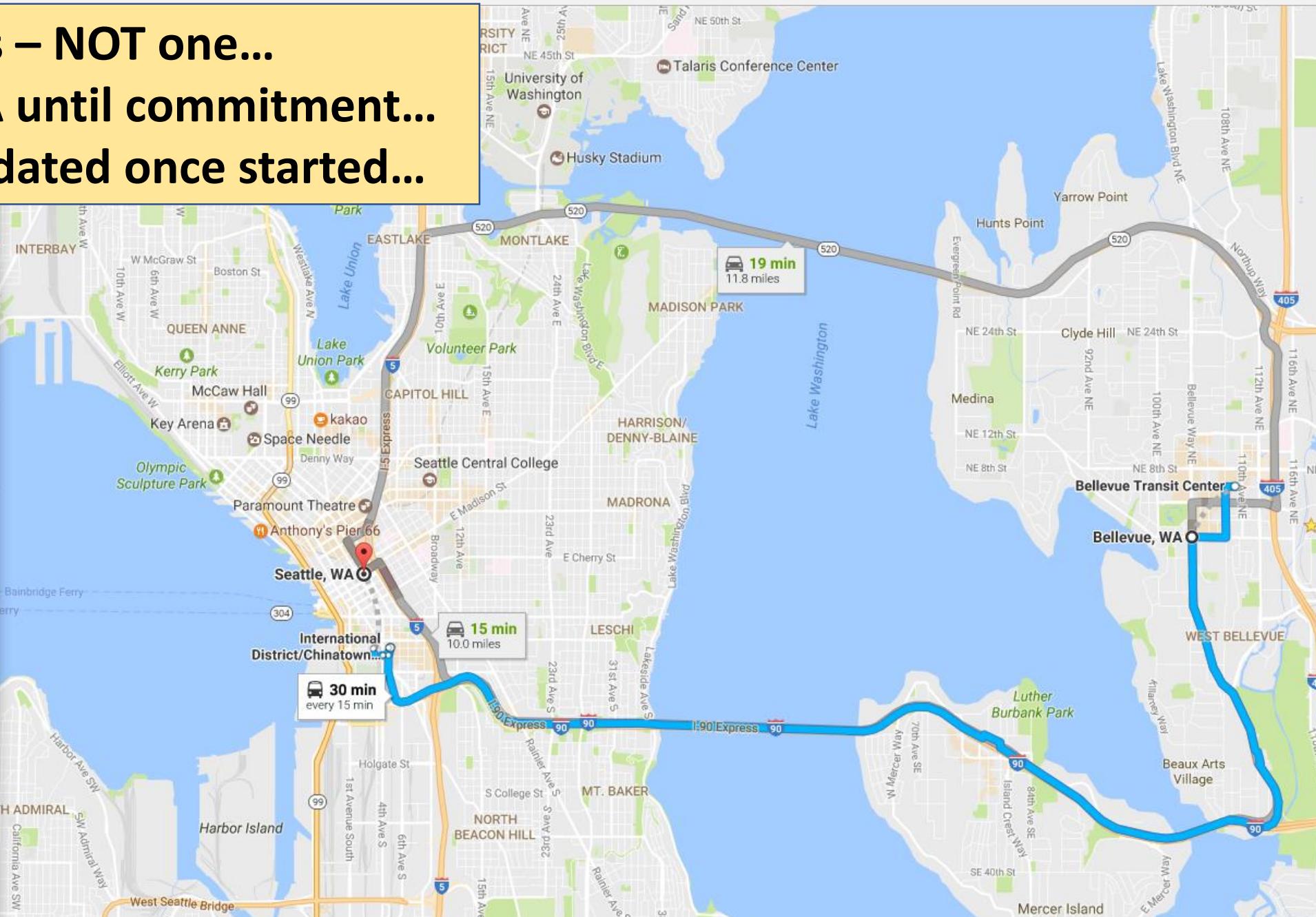
Leave now

OPTIONS

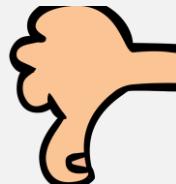
- Send directions to your phone**
- via I-90 W** **15 min**  
Fastest route, the usual traffic
- via WA-520 W** **19 min**  
11.8 miles
- 11:20 AM–11:50 AM** **30 min**  
550 ➔ ⚑  
11:20 AM from Bellevue Transit Center  
\$2.50 ⚑ 5 min every 15 min

DETAILS

SCHEDULE EXPLORER

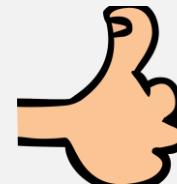


# Contrast Google Maps to Software Estimates



If you currently...

- Give one forecast even though multiple approaches considered
- Give a calendar date for undefined “complete” & “start”
- If the original date is in doubt we find out near the end
- Appear on-time until we are not. Measure progress from start.



Consider doing...

- Give multiple options of investment and implementation
- Give a duration and define what started & complete means
- If the original date is in doubt, know earlier and react faster
- Report remaining time to deliver not time since started

# Top Three Forecasting Questions....

## 1. How Big

- Understanding the size of a feature or project with less effort

## 2. How Long

- Understanding when a feature or project might be done
- Tracking progress

## 3. How Much

- OK, its too big,  
“what can I get by when...”
- Seeing options
- Making trade-off decisions earlier





# Q1: How Big?

Understanding the size of a feature or project with less effort

# How Big – Forecasting feature/project size

- First, can we avoid the question (quantitatively)
  - “When is it needed?” – perhaps it’s an easy binary answer of yes or no
  - “Compared to what?” – perhaps it’s an easy binary answer of bigger or smaller
- Then,
  - Do we know enough to perform a forecast?
    - Are the Knowns > Un-knowns
- If we do need a size forecast,
  - How can we forecast with the minimal effort possible:
    - Reference class forecasting
    - Sampling

Step 1

Feature 2  
3 Stories

Feature 3  
7-15 Stories

Feature 1  
15 Stories

Step 2

Feature 4  
?

Step 3

Feature 2  
3 Stories

Feature 3  
7-15 Stories

Feature 1  
15 Stories



Feature 4  
10-15 Stories

Known as Reference Class Forecasting

# Forecasting Total Story Count

- Question: How can I estimate the size of a feature or project without analyzing every piece of work?
- Theory: The “size” patterns of randomly sample epics, will persist through all other epics. Analyze a few and compute for the many...

<http://bit.ly/StoryCountForecaster>

Sampling based Monte Carlo story count forecasting Excel spreadsheet

Feature or Epic Name	Estimated # Stories or points (before starting)
Feature 1	5
Feature 2	3
Feature 3	8
Feature 4	4
Feature 5	2
Feature 6	7
Feature 7	
Feature 8	
Feature 9	
Feature 10	
Feature 11	
Feature 12	
Feature 13	
Feature 14	
Feature 15	



Process to estimate total size –

1. Pick a 5-10 features at random
2. Build sets of 15 re-samples (say 1000 times)
3. The number of sets that reach certain story count levels give probability

### 1. How many total features do you want to forecast?

15

total features entered on input sheet:

Enter the total number of features or epics you wish to forecast. The patterns exhibited by the story count breakdown of the samples fatures and epics will be extrapolated to this many total features.

### 2. What rate do you expect work to split?

low guess

1

high guess

1

actual

Work often splits into smaller pieces when started by the team. Also, new work gets discovered through defects and learning. Account for that here.

1 no change, 2 means every one item might be split into two, 3 means every item might become three items, etc. Most common range I've seen is 1 to 3.

### 3. Result: Forecast total story count or total story points

Likelihood	Total Story Count/points
50%	73
85%	81
95%	85

#### Odds in english

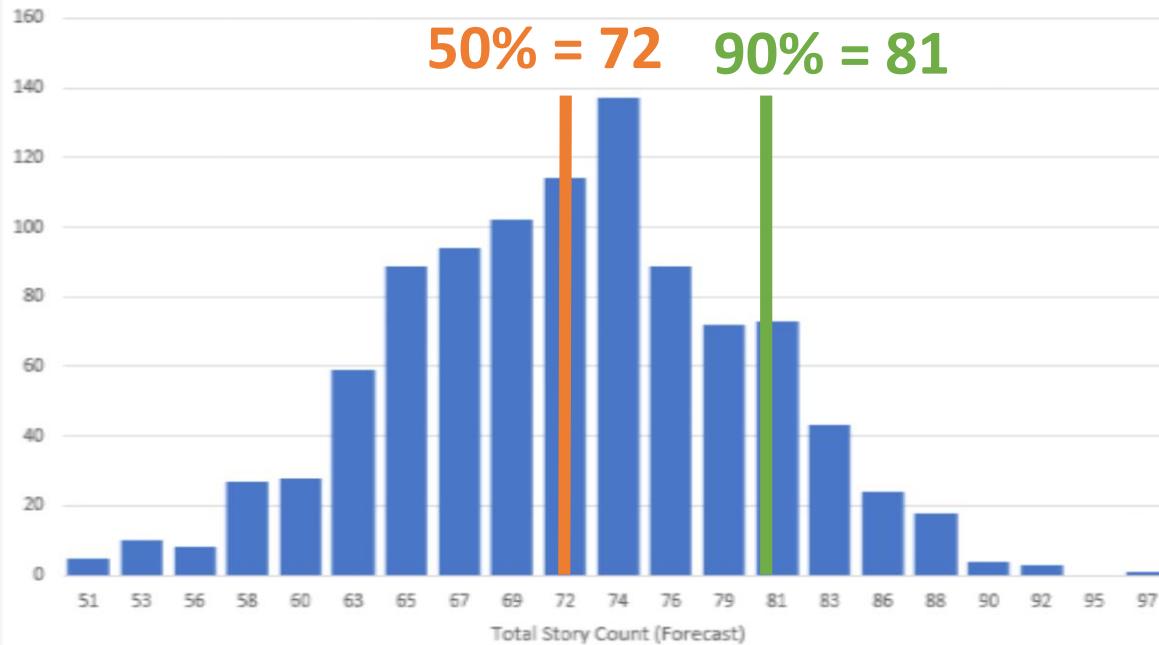
50% = Coin toss odds. Same chance being above or below this story count

85% = Pretty sure to be equal or less than this story count.

95% = Almost certain to be equal or less than this story count.

Histogram - Total Story Count for 15 features

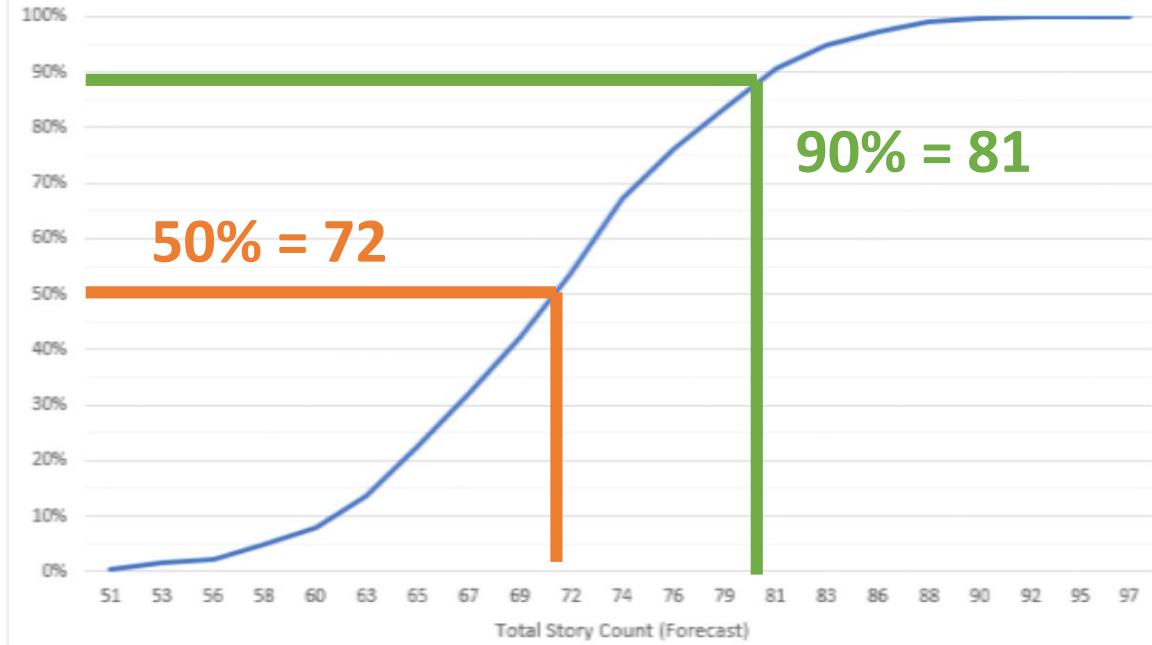
**50% = 72    90% = 81**



Probability - Total Story Count for 15 features

**50% = 72**

**90% = 81**



This chart plots how likely each total story count result occurred in the Monte Carlo simulation. The higher the bar, the more often that count of stories occurred. It is used to understand the pattern of the results to see how wide the tail values extend (those either side of the peak). If the range is too wide to make forecasting useful, the only solution is to make the features more consistent in size to avoid the story count estimate range being so huge.

This chart plots how likely each total story count result is as cumulative probability the result is less than or equal to a total story count. If you want to know what total story count is 80% likely, look at where the 80% on the Y-Axis intersects the blue line, and read off the total story count on the X-Axis.

# Why should I believe this forecast anyway?

1. Sample Count: Keep cutting data and compare the result
2. Random groups: Split data into random groups and compare

Total for 100 Features using	Total Count 85% Likelihood
36 samples	506
10 samples	494
3 samples	504

## Should I believe this forecast?

Number of samples:

8	Good
---	------

Error of average in two random groups:

13%
-----



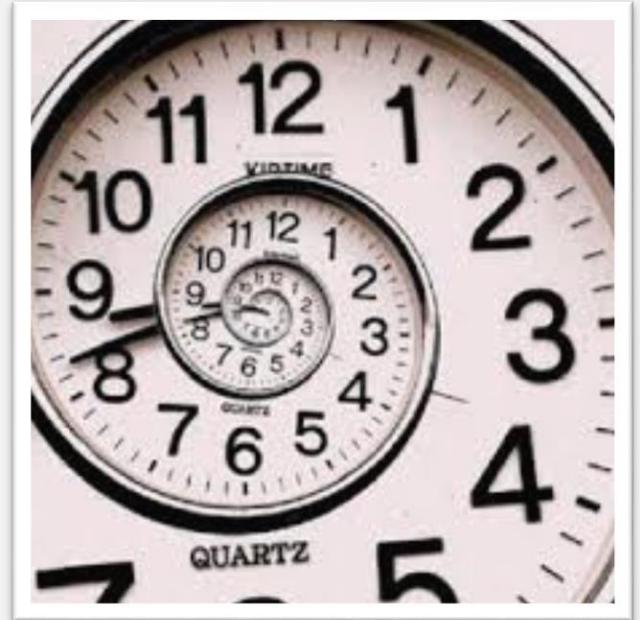
(note: with less than 7 samples, error is often 'unstable,'

hit F9 a few times to see how this changes (I use best of 5!).

0-25% good, 25-75% fair, >75% then too unstable to forecast)

### Average Error calculation –

1. Split the samples into 2 groups
2. Calculate the average of both groups
3. Compare the difference as a % of range  
 $\text{error \%} = \text{error of avg} / (\text{max-min})$



# Q2: How Long?

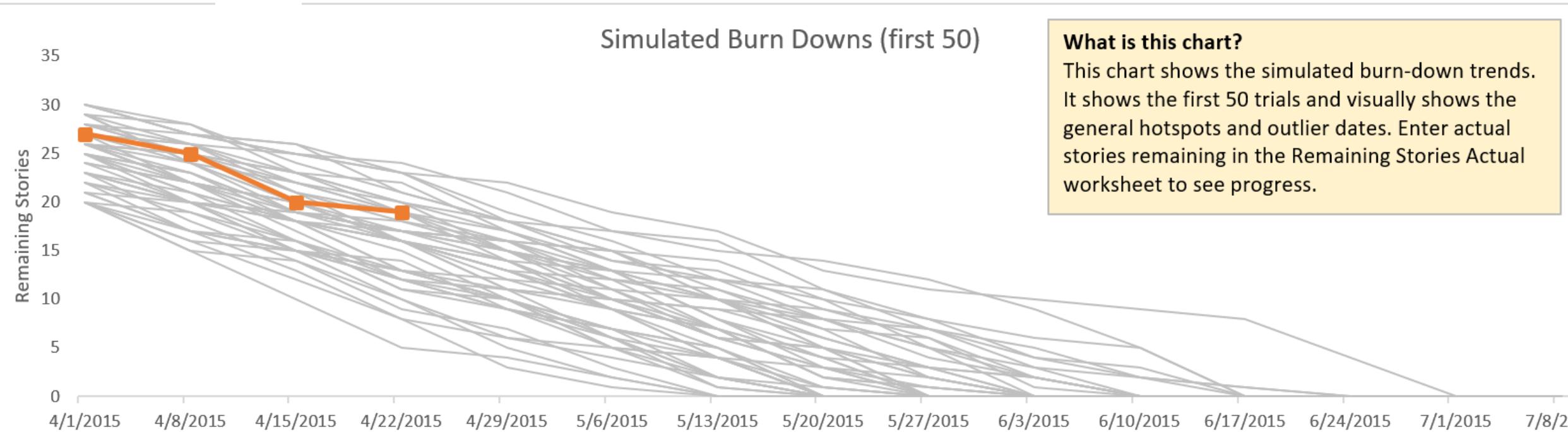
Forecasting duration if nothing else was done...

# Forecasting Duration (and delivery date)

- Question: How can I estimate the amount of time it will take to deliver a feature or project?
- Theory: Using a range estimate or actual team delivery rate data, calculate how many of those periods of time to complete delivery

<http://bit.ly/ThroughputForecast>

Estimate or Sampling based Monte Carlo  
duration and date forecasting Excel spreadsheet



## Forecast Completion Date

1. Start Date

4/1/2015

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

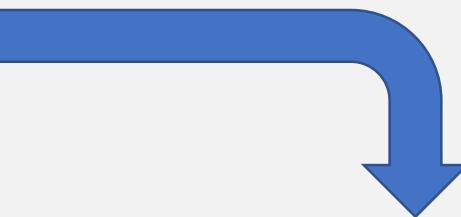
1

Highest guess

5

### Can I use velocity rather than throughput?

Yes. If you do have estimates in story points, then you can sum 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 completed stories per week/sprint) is that the individual stories don't require estimation in story points.



## Results

Likelihood	Duration in Week's	Date
100%	14	7/8/2015
95%	12	6/24/2015
90%	11	6/17/2015
85%	11	6/17/2015
80%	10	6/10/2015
75%	10	6/10/2015
70%	10	6/10/2015
65%	9	6/3/2015
60%	9	6/3/2015
55%	9	6/3/2015
50%	9	6/3/2015
45%	8	5/27/2015
40%	8	5/27/2015
35%	8	5/27/2015
30%	8	5/27/2015
25%	7	5/20/2015
20%	7	5/20/2015
15%	7	5/20/2015
10%	7	5/20/2015
5%	6	5/13/2015
0%	5	5/6/2015

Almost certain

Somewhat certain

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

# Using range estimates or data?

## None

- Use a uniform range estimate
- Use data to adjust uniform range estimate boundaries

## < 11 samples (in context)

## 11+ samples (in context)

- Use data.
- Monitor distribution weekly for changes

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

Throughput estimate/samples are per

days

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

Low guess

Highest guess



A	Enter Samples Below
1	3
2	5
3	2
4	1
5	5
6	0
7	0
8	
9	

Select “Data” rather than “Estimate” to use samples



# Q3: How Much?

OK, what can we get?

# Forecasting How Much (OK, what can I get?)

- Question: I have a date in mind, what features will likely delivery given historical delivery pace?
- Theory: Using duration forecasts, discuss the start order of features that maximize value and likelihood of successful delivery

<http://bit.ly/MultipleFeatureForecast>

Estimate or Sampling based Monte Carlo  
duration and date forecasting Excel spreadsheet  
for multiple features at one time

**Feature Cut Line Forecaster and Explorer**

**Only edit orange input fields**

<b>1. Start Date</b>	<b>1/1/2015</b>	<b>2. Target Date</b>	<b>7/1/2015</b>	<b>3. Likelihood</b>	<b>85%</b>
----------------------	-----------------	-----------------------	-----------------	----------------------	------------

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

Low guess	1.00	Highest guess	3.00
-----------	------	---------------	------

**5. Throughput. How many PLANNED (post split) completed stories 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.

<b>Choose here:</b>	<b>Estimate</b>
---------------------	-----------------

Low guess	5	Highest guess	10
-----------	---	---------------	----

**7. Enter the features and story count estimates here...**

						Start date: 01/01/2015
<b>Start Order</b>	<b>Feature Name (just for reference)</b>	<b>Story Count</b>		<b>Feature Duration in Weeks</b>	<b>Forecast Completion Date (85% CI)</b>	
		<b>Low Guess</b>	<b>High Guess</b>			
1	Feature 1	5	10	3 ✓	1/22/2015	
2	Feature 2	8	15	5 ✓	2/26/2015	
3	Feature 3	15	25	8 ✓	4/23/2015	
4	Feature 4	20	30	10 !	7/2/2015	
5	Feature 5	10	40	11 ✗	9/17/2015	

# Give multiple options – discuss cuts early

						Start date: 01/01/2015
Start Order	Feature Name (just for reference)	Story Count Low Guess	Story Count High Guess	Feature Duration in Weeks	Forecast Completion Date (85% CI)	
1	Feature 1	5	10	3	✓ 1/22/2015	
2	Feature 2	8	15	5	✓ 2/26/2015	
3	Feature 3	15	25	8	✓ 4/23/2015	
4	Feature 4	20	30	10	! 7/2/2015	
5	Feature 5	10	40	11	✗ 9/17/2015	
6				0	✗ 9/17/2015	
7				0	✗ 9/17/2015	
8				0	✗ 9/17/2015	
9				0	✗ 9/17/2015	
10				0	✗ 9/17/2015	

**Legend**

- ✓ Forecast on or before the target date
- ! Forecast misses target date by one Week or less
- ✗ Forecast misses target date by MORE than one Week



# Top Three Forecasting Fail Reasons

Reasons you shouldn't have hired me five years ago

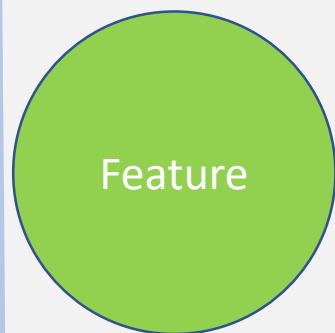
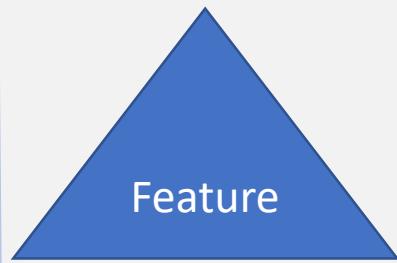
# Fail 1: Start Date On-Paper != Reality

- The assumed Start Date is often ONLY on paper
- Define what start means
  - Team is dedicated and in-place
  - They are trained and know how to do their work
  - They know and understand what work they need to deliver
  - Nothing inhibits them doing or delivering that work
- Team is never fully available on day one!

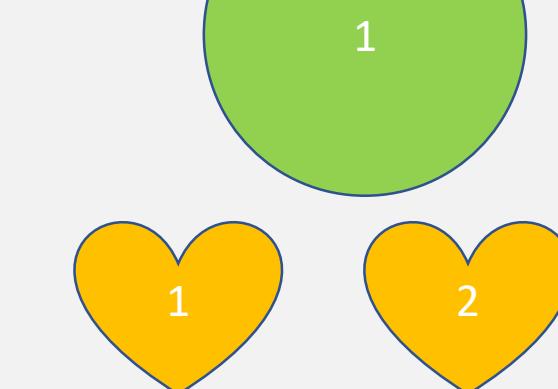
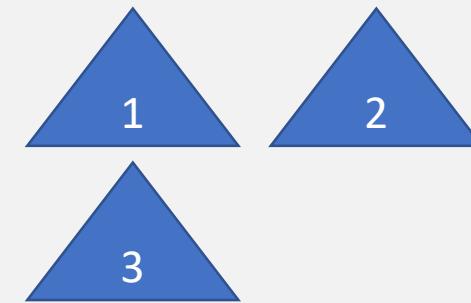
Start Date of Feature B  
is the finish date of  
Feature A  
What is the team doing now?

# Fail 2: Backlog Rate versus Delivery Rate

Features



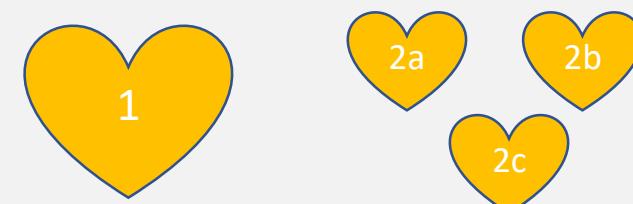
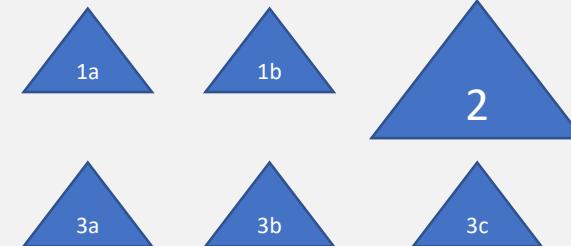
Estimated Stories



Actual Backlog

Rate Delivered = 6

Implemented Stories and Defects



Measured Throughput = 12

# Fail 2: Backlog Rate versus Delivery Rate

- Forecast using the “Completion rate” we may under-forecast
  - Backlog is Miles per Hour, Completion rate is Kilometers per Hour
- Normal split rates are between 1 to 3 times (most common seen)
- This means
  - If you don’t account for it, you will UNDER-FORECAST by 1 to 3 times!

**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)

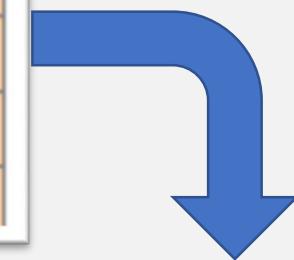
Low guess

1.00

Highest guess

3.00

Feature or Epic Name	Estimated # Stories or points (before starting)	Actual # Stories or points (after completed)	
Feature 1	5	8	$5 \times 1.6 = 8$
Feature 2	3	4	$3 \times 1.3 = 4$
Feature 3	8	10	$8 \times 1.25 = 10$
Feature 4	4		
Feature 5	2		
Feature 6	7		
Feature 7			



actual growth rate range seen:

re.

:1 to 3.

1.25

to

1.60

(enter actual count after work is completed in the input sheet)

This split-rate adjust the rate that the team appears to be completing work with the original backlog items are started. This isn't only growth/creep, its adjusting for the way work splits when the team has a closer look at it.

# Fail 3: Ignoring Risks

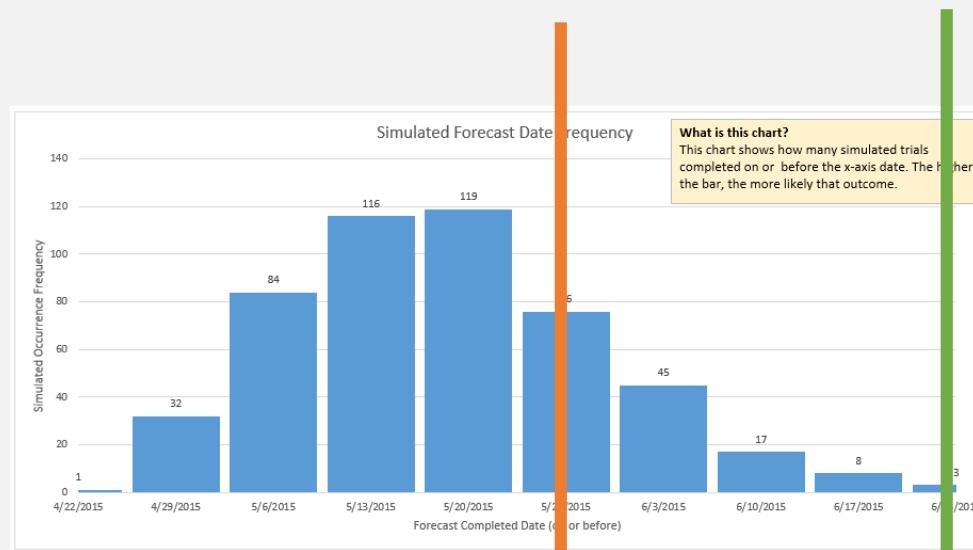
- **Risk = Work that “might” need to be done but we don’t know yet**
- Some samples
  - Fails on Internet Explorer 6, or now Safari on phones
  - Fails performance testing under load, or uses too much memory
  - CSS alignment issues with German text translations, things wrap
  - Production network security blocks traffic, awaiting vendor to fix
  - Fails on real customer data (we designed for 50 items, they have 500)

**WITHOUT  
RISKS INCLUDED**

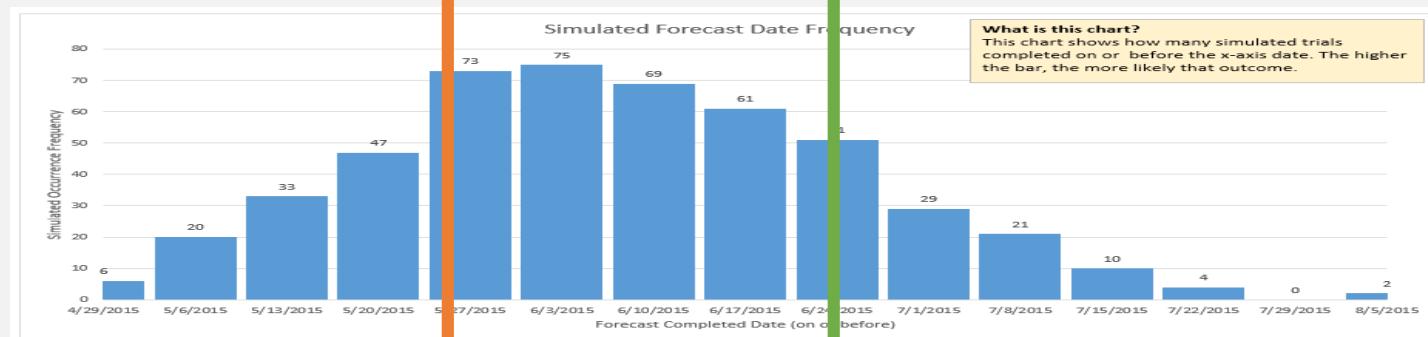
**27<sup>th</sup> May**  
(highest late June)

**24<sup>th</sup> June**  
(highest early August)

**WITH  
RISKS INCLUDED**

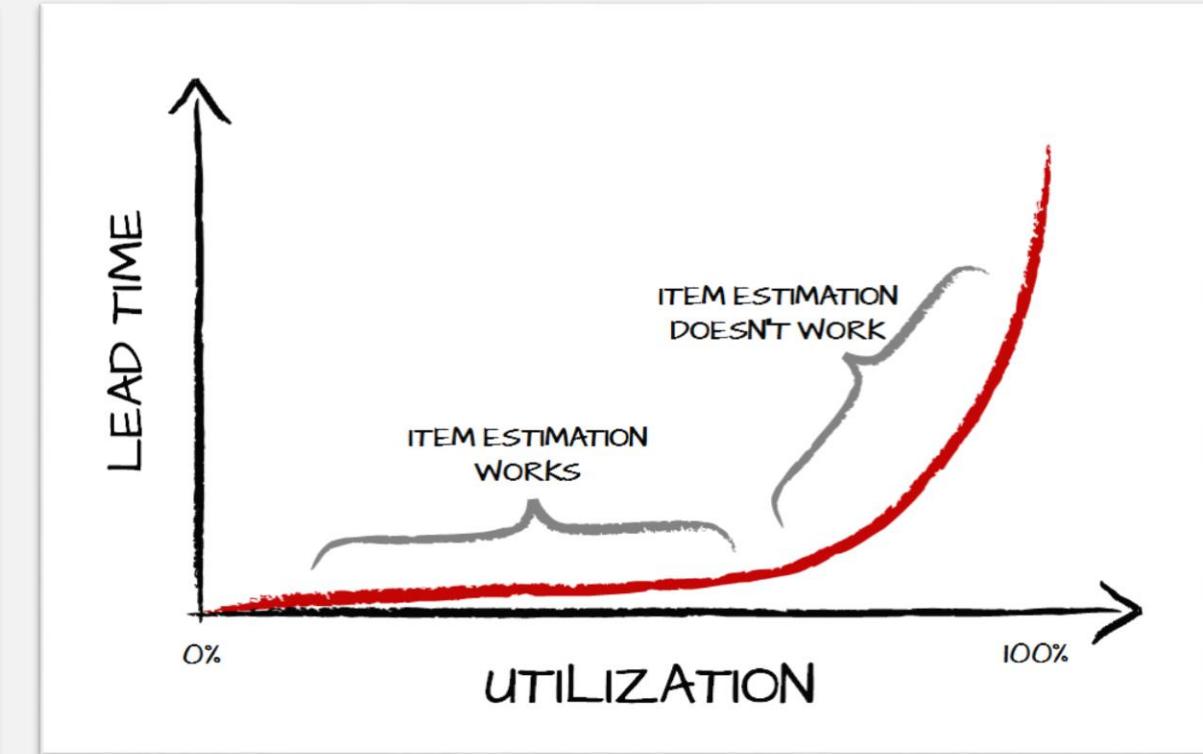


Likelihood	Impact Low	Impact High	Description
50%	7	10	Browser compatibility issues
40%	3	7	Performance under load
30%	5	10	Production configuration

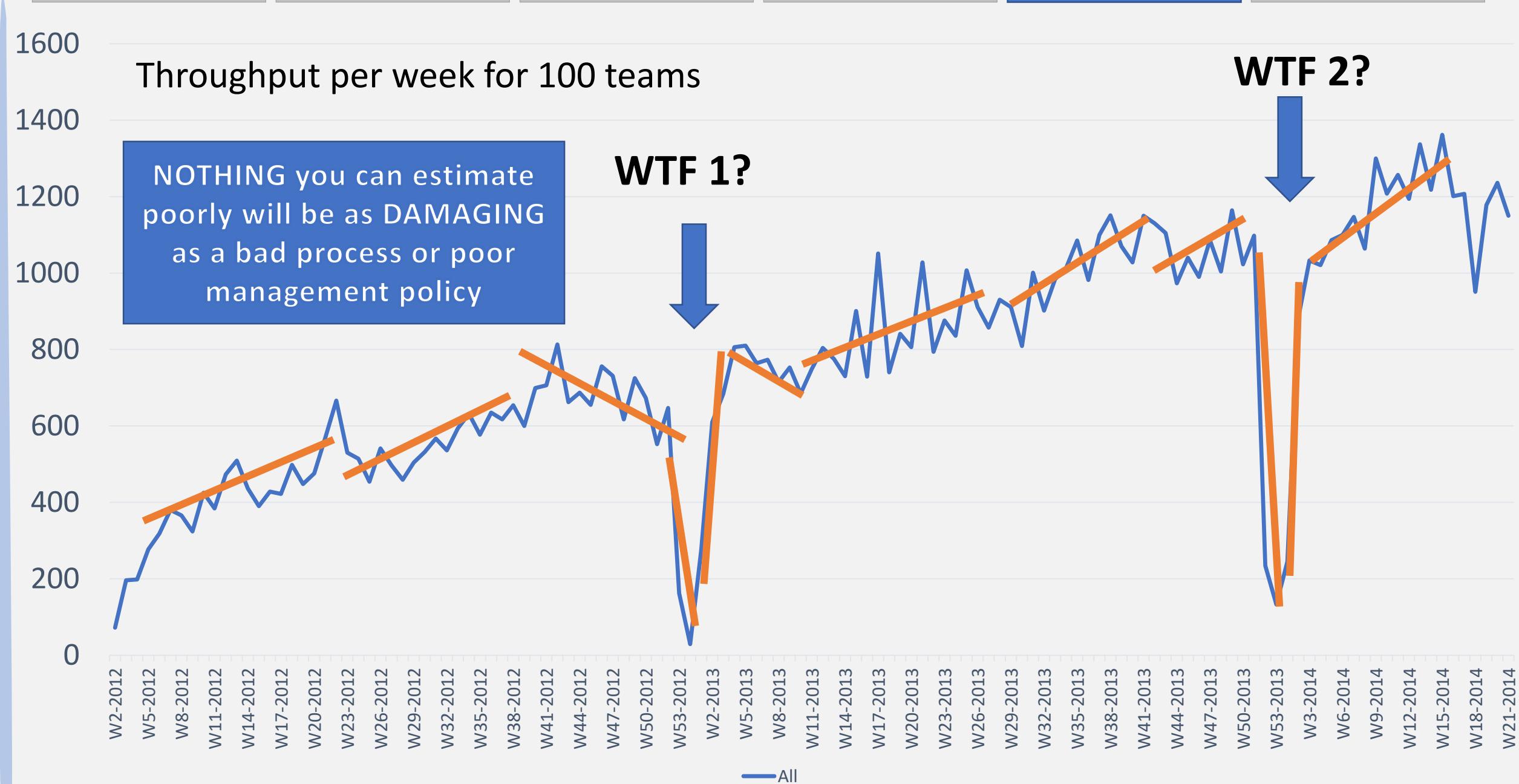


Forecasts shown at  
85<sup>th</sup> Percentile

# Bonus fail: High System Utilization



Can't forecast high utilization systems using item size...



120000

## Cumulative Throughput per week for 100 teams

100000

Avoid cumulative charts for short term forecasting, they over-smooth short term trends

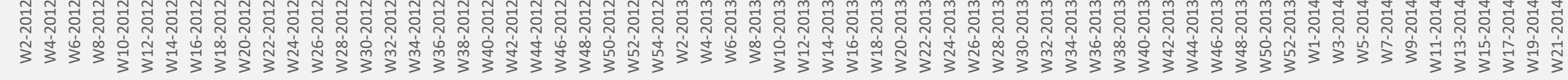
80000

60000

40000

20000

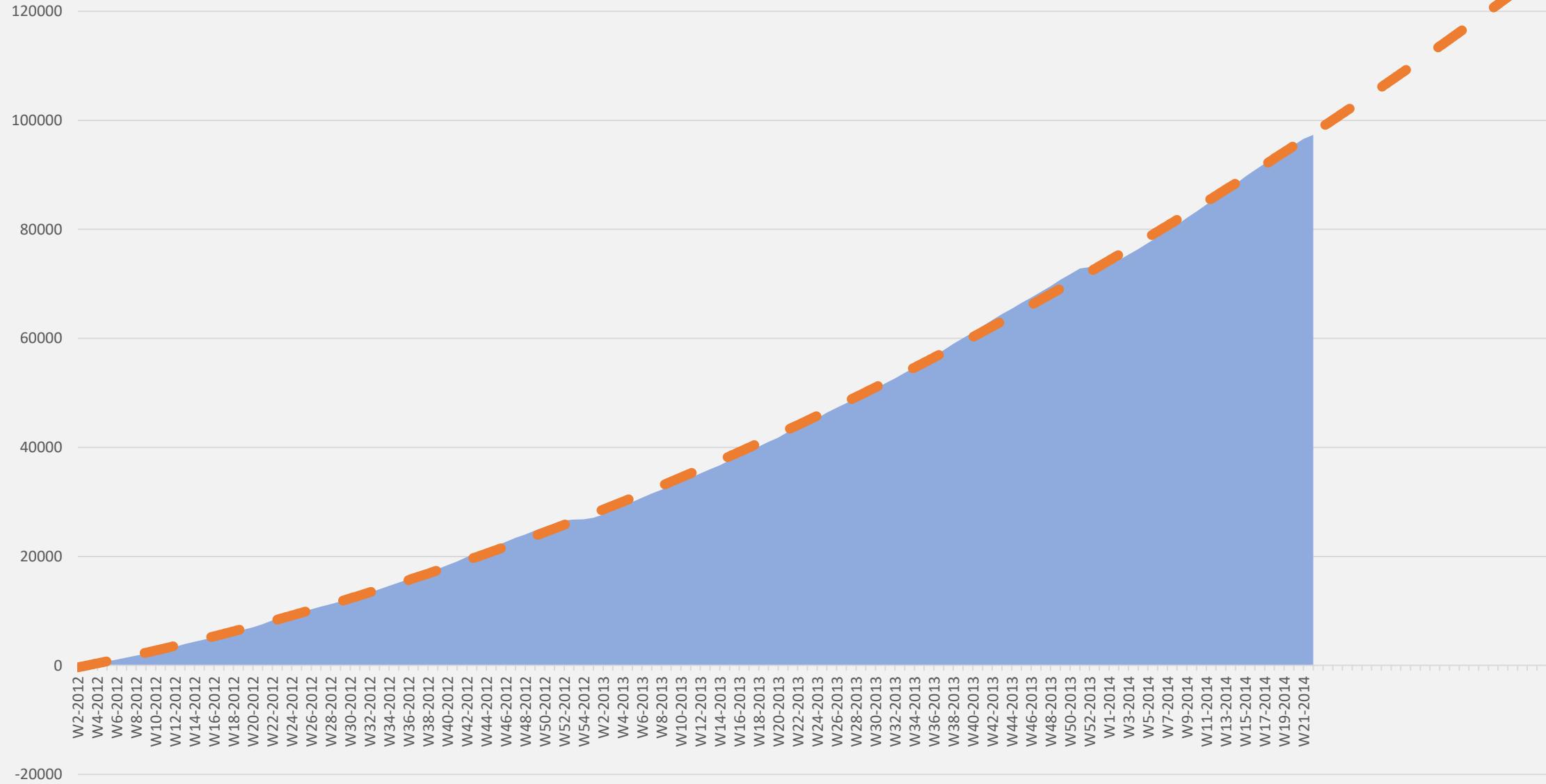
0



WTF 2?

WTF 1?

## Cumulative Throughput per week for 100 teams



# Key Take-aways and Resources

- Forecasting requires a system view,
- Three samples will outperform intuition (use most recent 7 samples)
- Give multiple options, not just one
- Forecast duration NOT date until “Start Conditions” are defined
- Track actual progress versus planned, and update the model continuously
- Get everything here: Slides and tools:

**Bit.ly/SimResources**

Get everything here: Slides and tools:

[Bit.ly/SimResources](https://Bit.ly/SimResources)



Me on Twitter

@t\_magennis

Slides, spreadsheets, and other stuff

Bit.ly/SimResources

Everything you see is freely available

# About me...

- What I do
  - Teach how to use data for forecasting
  - Teach simple math to executives, especially “demand > supply”
  - Teach how to know (earlier) that you are on the wrong side of an expectation
- What I did
  - Started in software 1986. I actually liked Assembler & Cobol
  - Have worked at senior exec level, and now beside them for major corporations so I have some insight into what passes their decision filters
- How to reach me
  - Twitter: @t\_magennis or email: [troy.magennis@focusedobjective.com](mailto:troy.magennis@focusedobjective.com)
  - Lots of free spreadsheets and stuff at FocusedObjective.com