

CODEMASH



Answering the Question, *“When Will It Be Done?”*

Using Probabilistic Forecasting

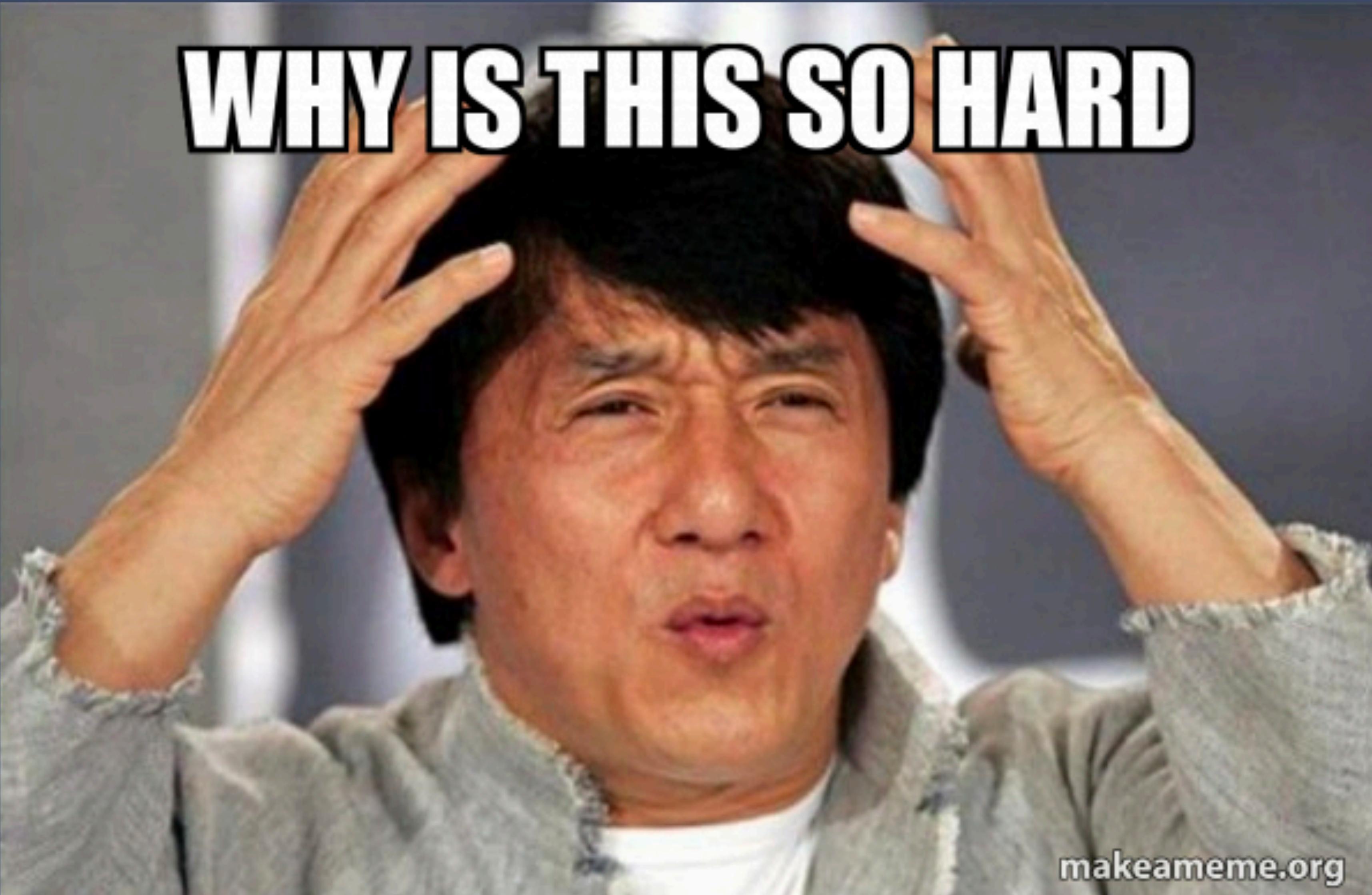


Chris Shinkle
Director of Innovation

**ARE WE
THERE
YET?**



When will it be done?



makeameme.org

MONTY HALL PROBLEM



**LET'S
MAKE
A
DEAL**



LET'S PLAY...

The Monty Hall Problem: A Study

Michael Mitzenmacher

Research Science Institute 1986

Abstract

The Monty Hall problem is based on apparent paradox that is commonly misunderstood, even by mathematicians. In this paper we define the Monty Hall problem and use a computer simulation to shed light on it. We then provide a mathematical explanation that fits the experimental results.

The Time Everyone “Corrected” the World’s Smartest Woman

Zachary Crockett



By all accounts, Marilyn vos Savant was a child prodigy.

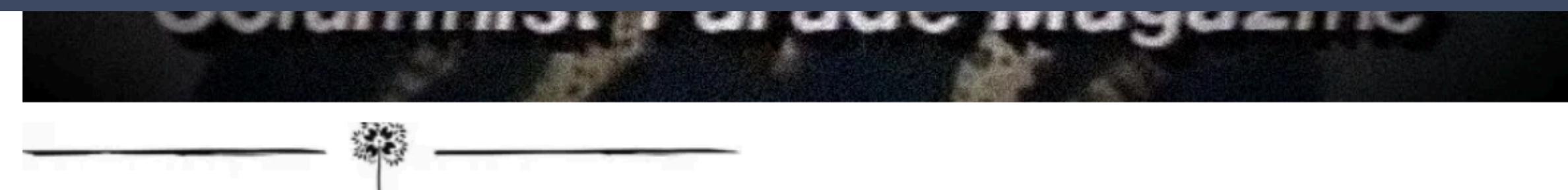
Born in St. Louis, Missouri in 1946, the young savant quickly developed an aptitude for math and science. At age 10, she was given two intelligence tests — the Stanford-Binet, and the Mega Test — both of which placed her mental capacity at that of a 23-year-old. She went on to be listed in the *Guinness Book of World Records* for having the “World’s Highest IQ,” and, as a

The Time Everyone “Corrected” the World’s Smartest Woman

Zachary Crockett

“Our brains are just not wired to do probability problems very well...”

- Persi Diaconis, Stanford Statistics Professor



By all accounts, Marilyn vos Savant was a child prodigy.

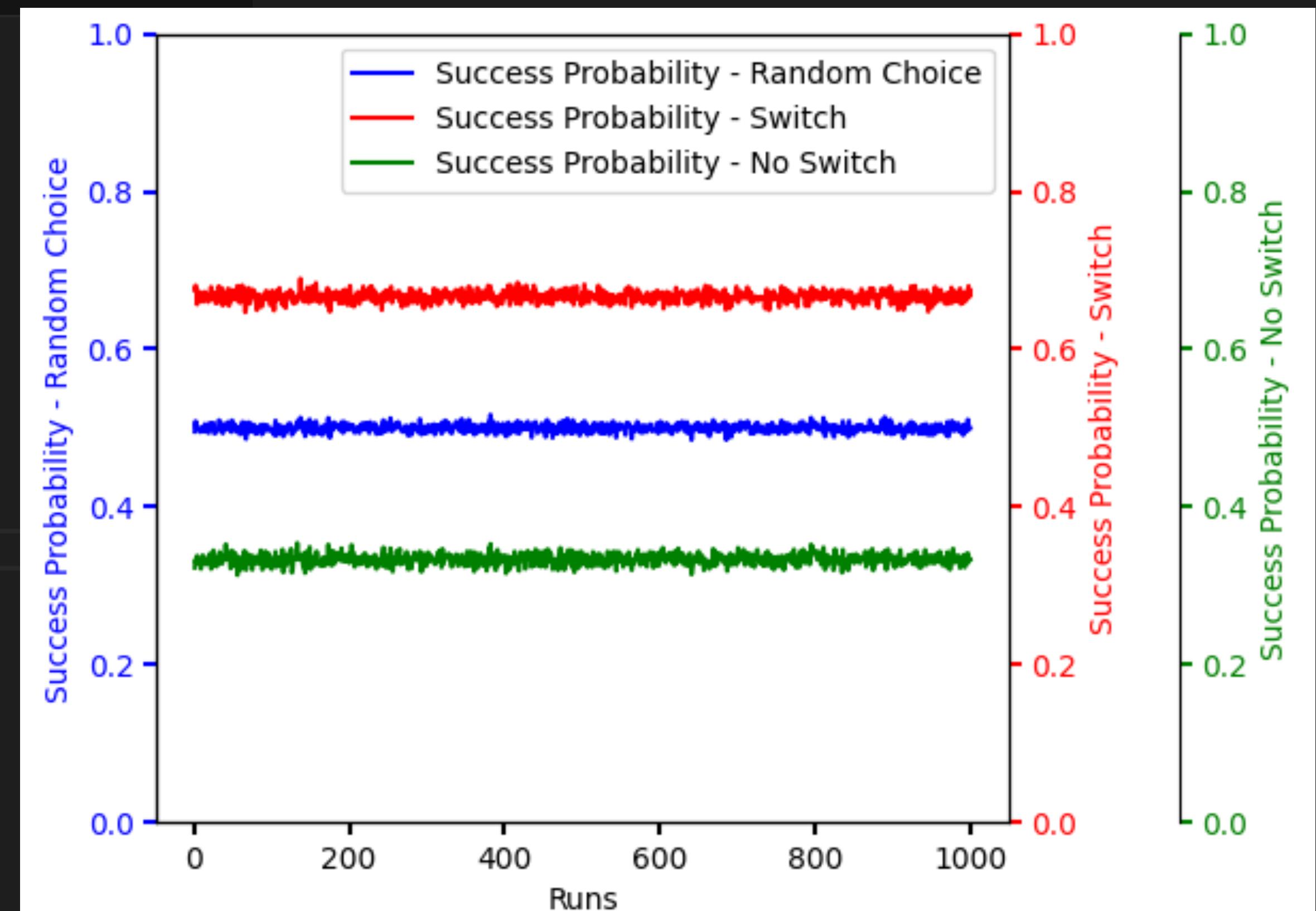
Born in St. Louis, Missouri in 1946, the young savant quickly developed an aptitude for math and science. At age 10, she was given two intelligence tests — the Stanford-Binet, and the Mega Test — both of which placed her mental capacity at that of a 23-year-old. She went on to be listed in the *Guinness Book of World Records* for having the “World’s Highest IQ,” and, as a

You Pick	Prize Door	Don't Switch	Switch
1	1	Win	Lose
1	2	Lose	Win
1	3	Lose	Win
2	1	Lose	Win
2	2	Win	Lose
2	3	Lose	Win
3	1	Lose	Win
3	2	Lose	Win
3	3	Win	Lose
		3 Wins (33%)	6 Wins (66%)

monty_hall.py X

monty_hall.py > experiment

```
1 import random
2 import matplotlib.pyplot as plt
3
4 def experiment(switch_door):
5     # compute the correct door randomly
6     correct_door = random.choice([1, 2, 3])
7     # choose a door randomly
8     door = random.choice([1, 2, 3])
9
10    # Among two remaining door, get a random incorrect door
11    doors = [1,2,3]
12    try:
13        doors.remove(door)
14        doors.remove(correct_door)
15    except:
16        pass
17
18    random_incorrect_door = random.choice(doors)
19
20    # Remove the random incorrect door from the options available to you
21    doors = [1, 2, 3]
22    doors.remove(random_incorrect_door)
23
24    # Now among your original choice of door and the new set of options, switch your choice
25    # remove your original choice from the options
26    doors.remove(door)
27    # Now as only one option is there within
28    if switch_door:
29        final_choice = doors[0]
30    else:
31        final_choice = door
32
33    # If the final choice is the correct door, then return 1, else return 0
34    if final_choice == correct_door:
35        return 1
```



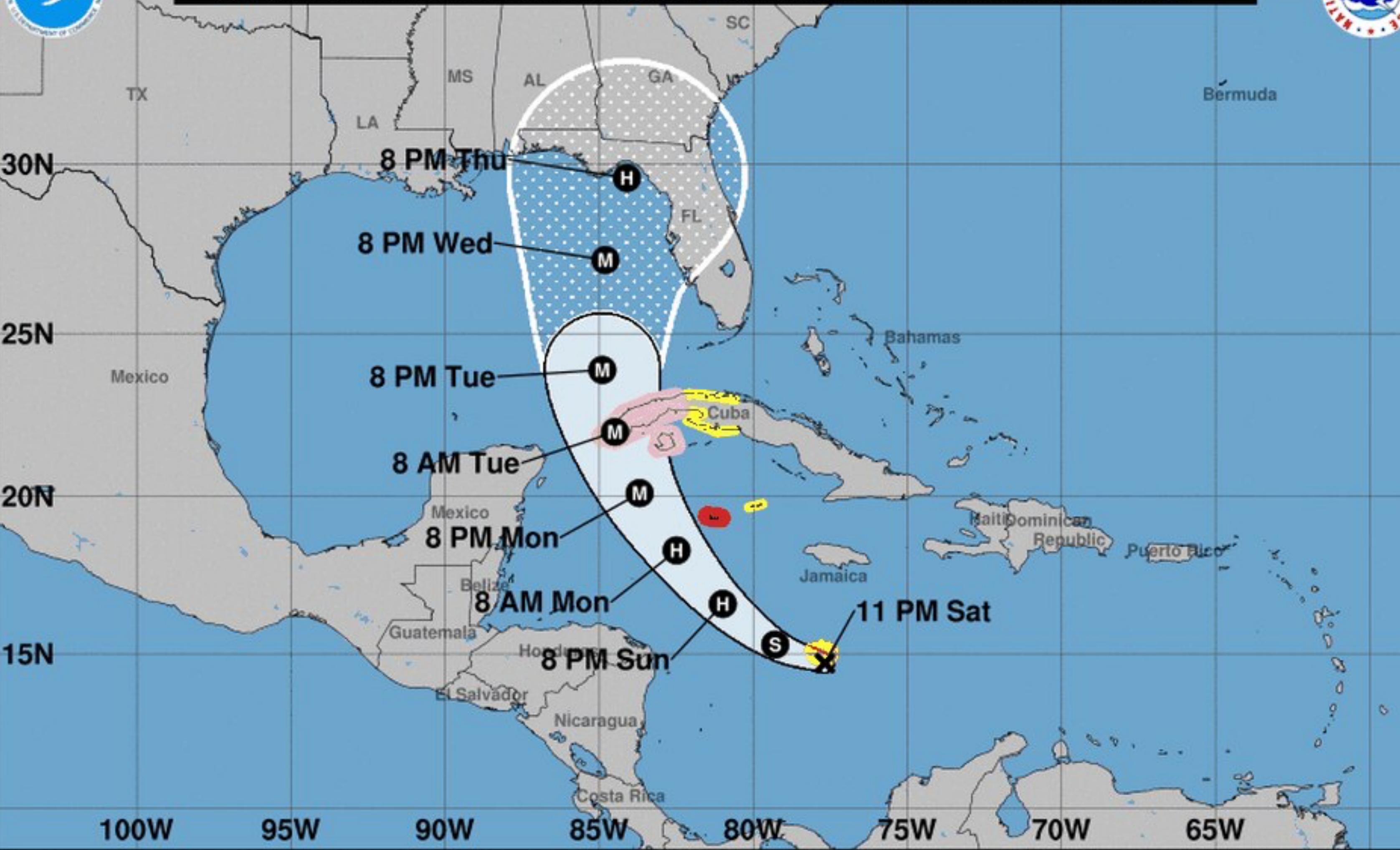
Giving an accurate answer to "When will it be done?" is hard.
Humans struggle...

- 1. THINKING PROBABILISTICALLY**
- 2. FILTERING EVIDENCE: CHERRY-PICKING, BIAS AND INTERESTS**
- 3. ACCOUNTING FOR COMPOUND PROBABILITIES**
- 4. OVERLOOKING DEPENDENCIES**

**Every project is unique and we can't
predict the future!**



Note: The cone contains the probable path of the storm center but does not show the size of the storm. Hazardous conditions can occur outside of the cone.



Tropical Storm Ian

Saturday September 24, 2022

11 PM EDT Advisory 8

NWS National Hurricane Center

Current information: x

Center location 14.7 N 77.7 W
Maximum sustained wind 50 mph
Movement W at 13 mph

Forecast positions:

● Tropical Cyclone ○ Post/Potential TC
Sustained winds: D < 39 mph
S 39-73 mph H 74-110 mph M > 110 mph

Potential track area:

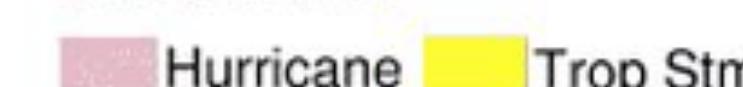


Day 1-3



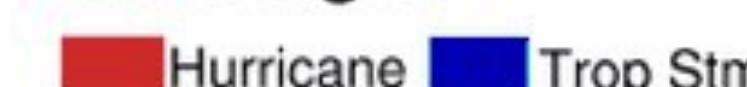
Day 4-5

Watches:



Hurricane Trop Stm

Warnings:



Hurricane Trop Stm

Current wind extent:



Hurricane Trop Stm

How does the NWS National Hurricane Center predict the future?

ECMWF: 09/04 00Z

NCEP: 09/04 00Z

Statistical: 09/04 06Z

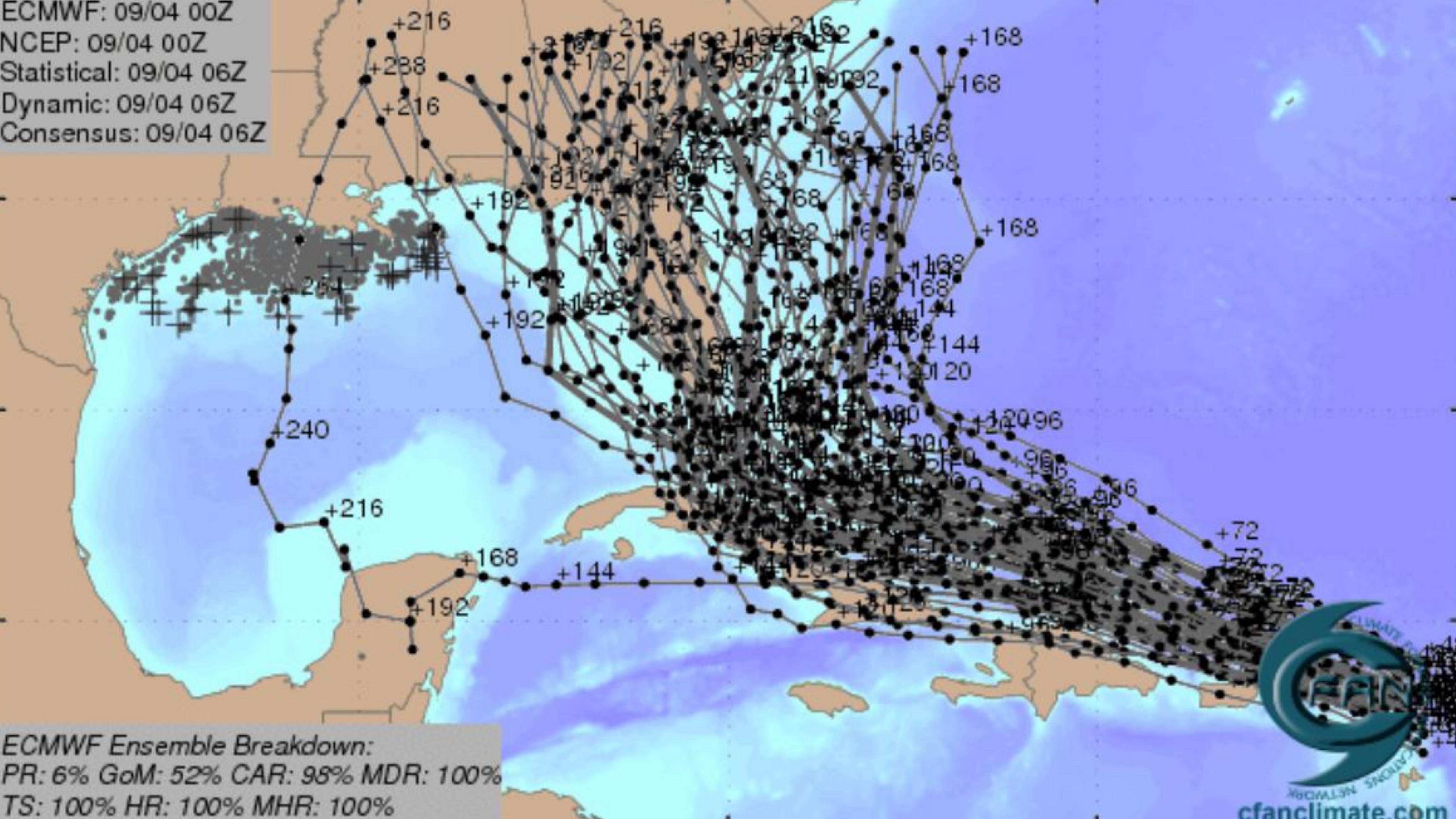
Dynamic: 09/04 06Z

Consensus: 09/04 06Z

ECMWF Ensemble Breakdown:

PR: 6% GoM: 52% CAR: 98% MDR: 100%

TS: 100% HR: 100% MHR: 100%



cfanclimate.com

What can we learn from the NHC?

- 1. WEATHER FORECASTING IS UNCERTAIN. THEREFORE, YOU CAN'T THINK DETERMINISTICALLY. YOU NEED TO THINK PROBABILISTICALLY.**
- 2. THINKING PROBABILISTICALLY MEANS ACKNOWLEDGING THERE IS MORE THAN ONE POSSIBLE FUTURE OUTCOME.**
- 3. UPDATE FORECASTS AS MORE INFO BECOMES AVAILABLE**
- 4. SHORTER TERM FORECASTS ARE BETTER THAN LONGER TERM ONES**

When will my software project be done?



**HOW ARE PEOPLE
ANSWERING THIS
TODAY?**

SOFTWARE PROJECT EXAMPLE

10 Stories* in week

80 Stories left to complete

When will it be done?

* Could also use story points, tasks, number of items to complete, etc.

**Humans, when
pressed for a
single number
or answer, will
often give an
average**

WHY WE UNDERESTIMATE RISK IN
THE FACE OF UNCERTAINTY

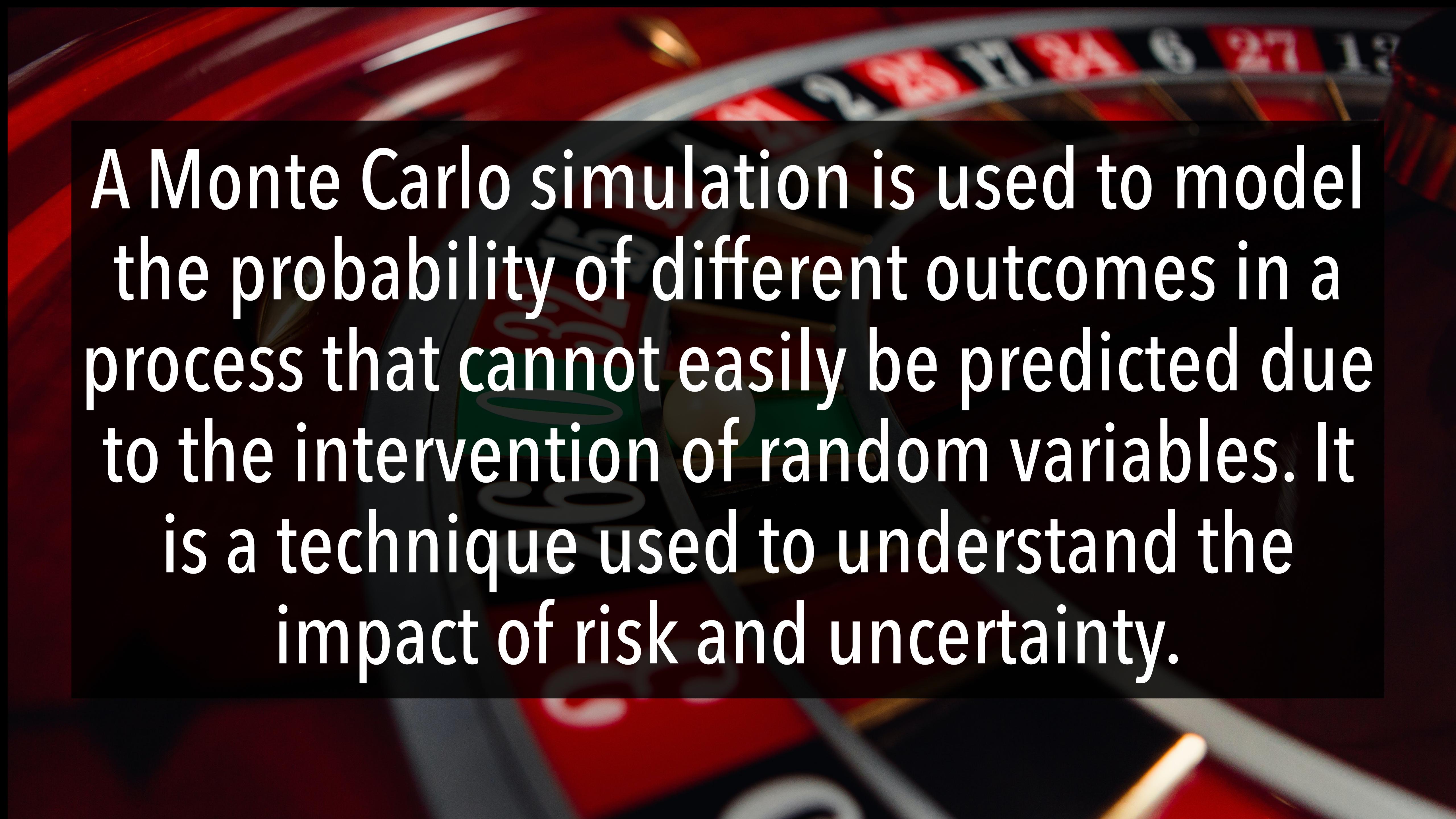
**THE
FLAW
OF
AVERAGES**

SAM L. SAVAGE
WITH ILLUSTRATIONS
BY JEFF DANZIGER



FORECASTING USING THE MONTE CARLO METHOD





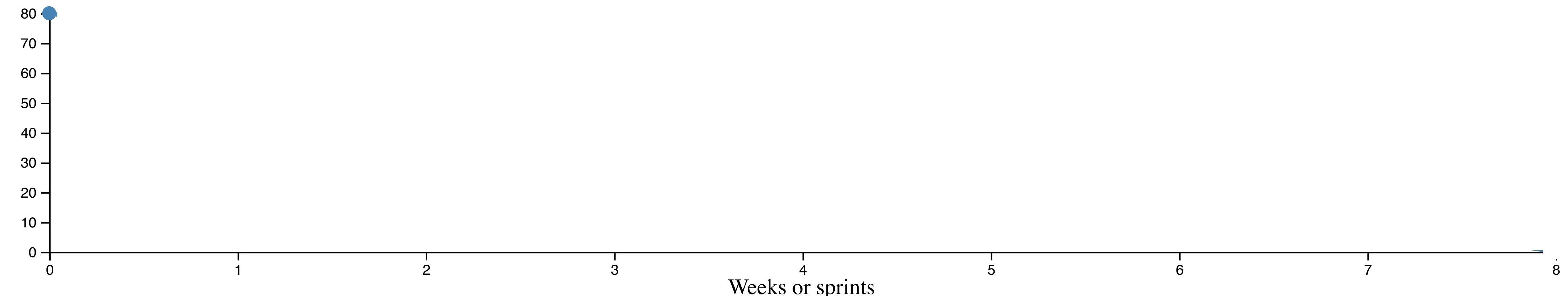
A Monte Carlo simulation is used to model the probability of different outcomes in a process that cannot easily be predicted due to the intervention of random variables. It is a technique used to understand the impact of risk and uncertainty.

FORECASTING MODEL

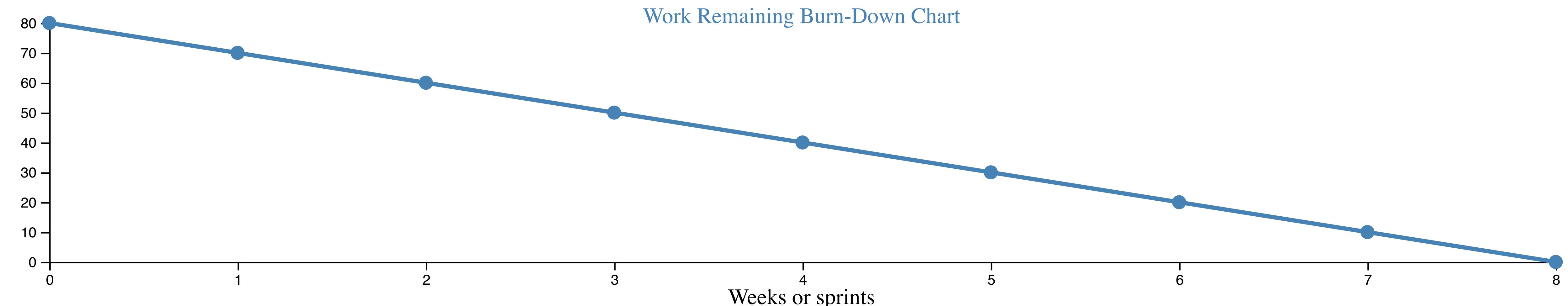
(size-growth-pace model)

Size * Growth
How long = _____
Pace

10 Stories in week
80 Stories left to complete

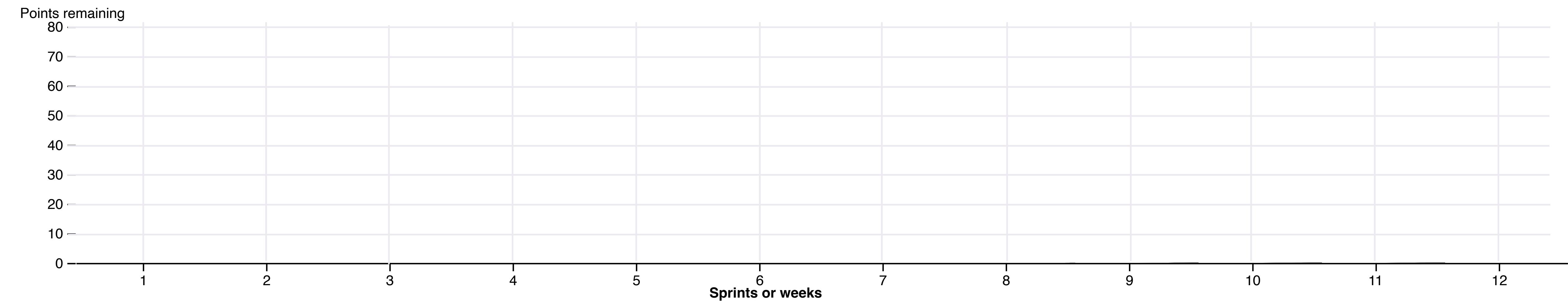


10 Stories in week
80 Stories left to complete



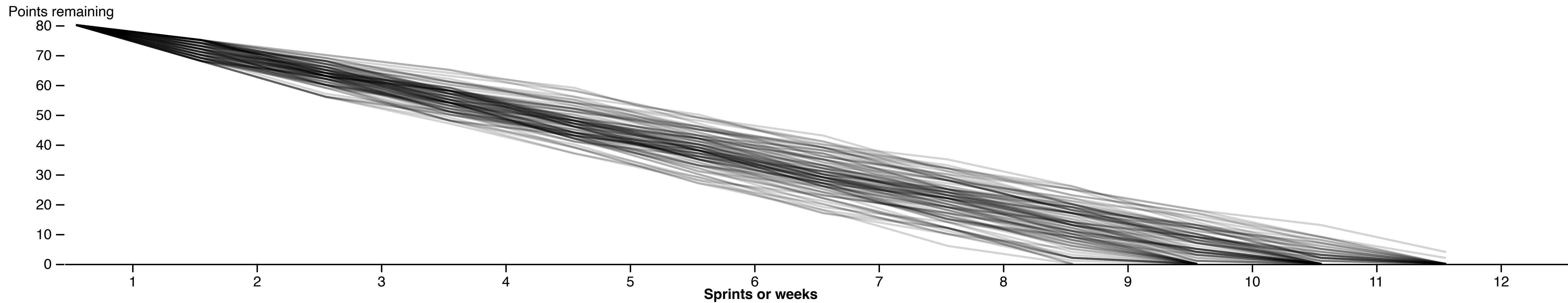
[5-12] Stories in week

80 Stories left to complete



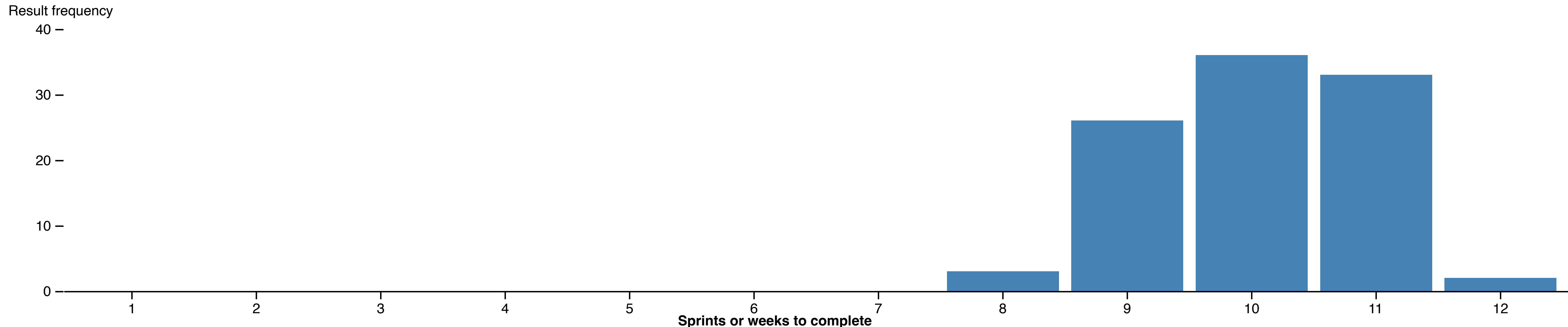
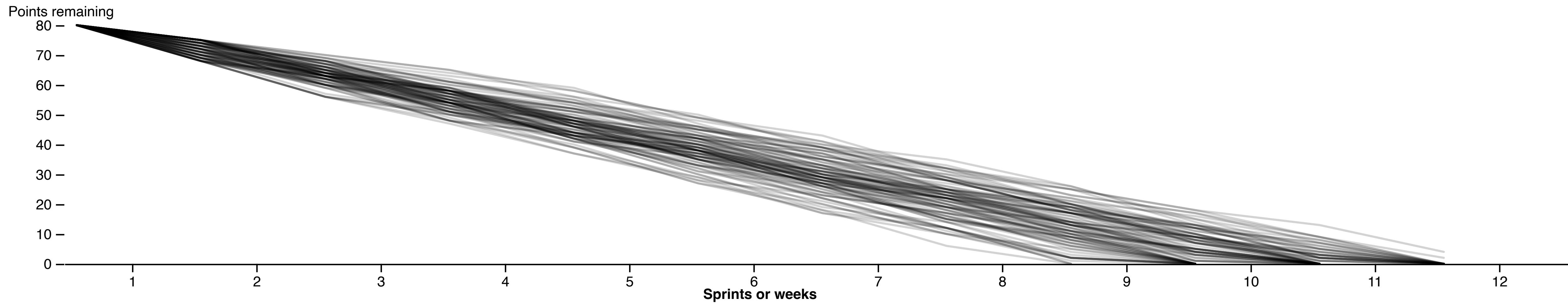
[5-12] Stories in week

80 Stories left to complete



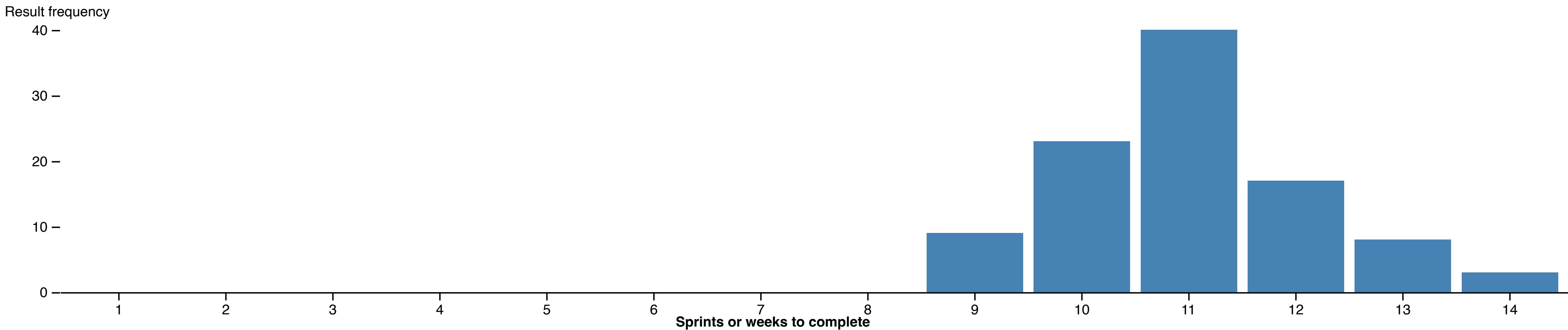
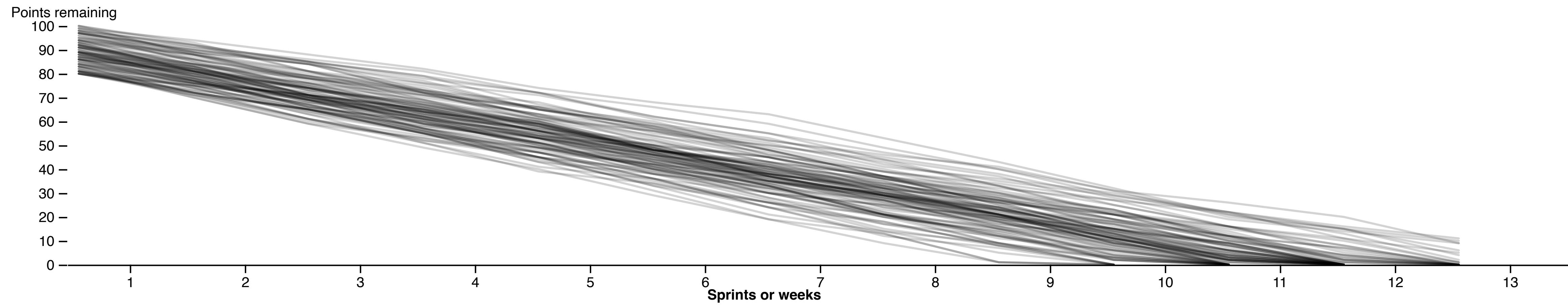
[5-12] Stories in week

80 Stories left to complete



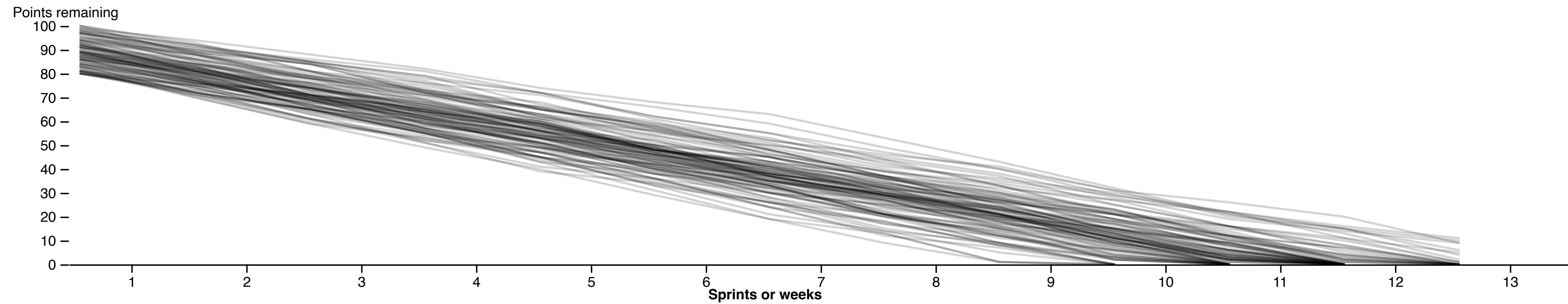
[5-12] Stories in week

[80-100] Stories left to complete

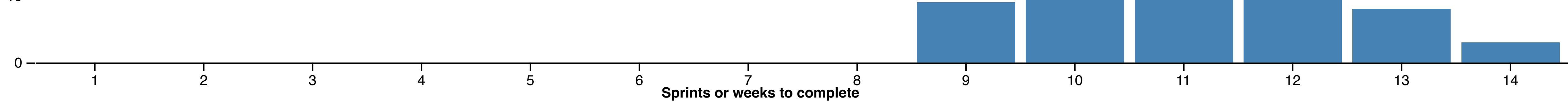


[5-12] Stories in week

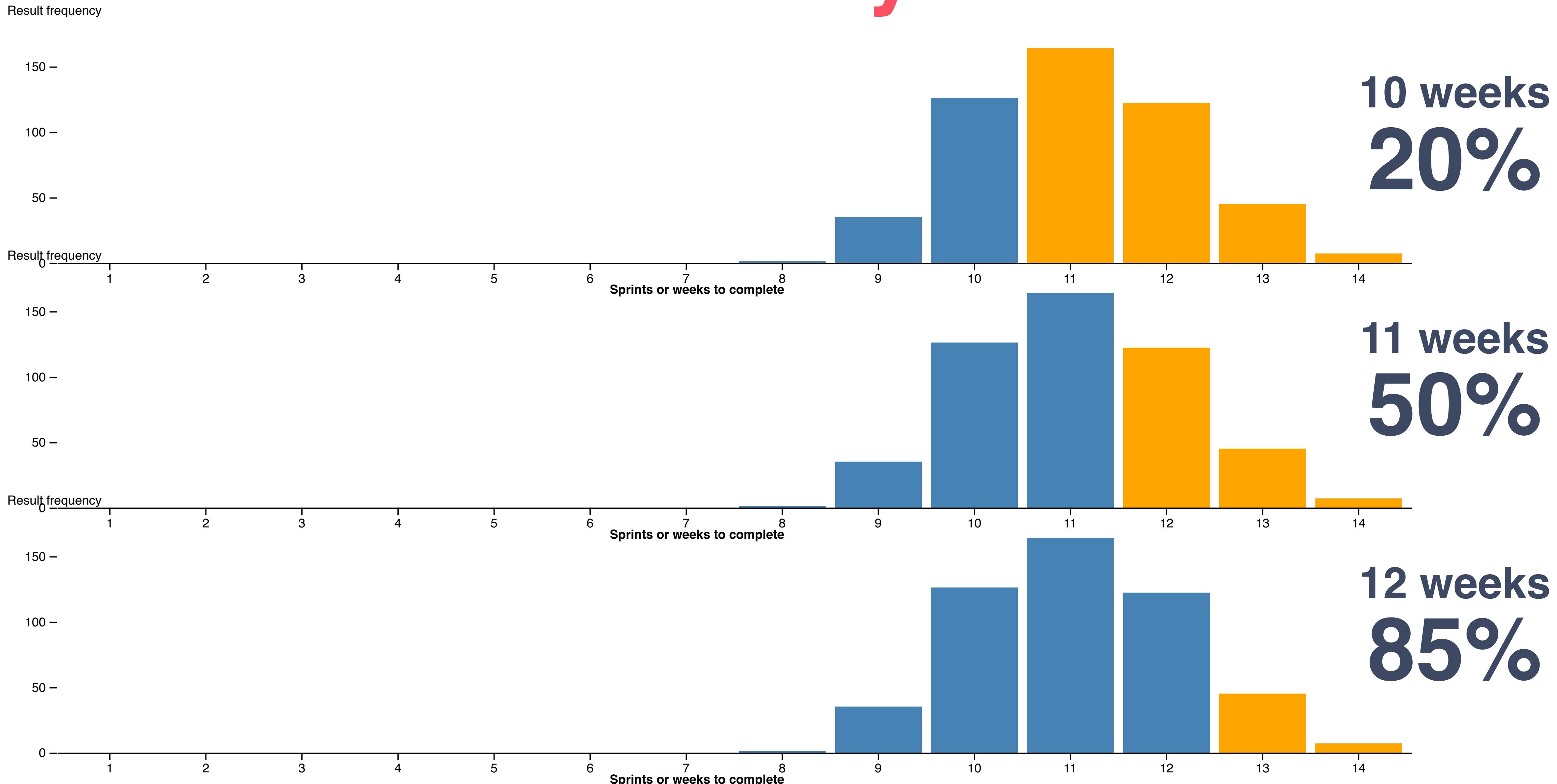
[80-100] Stories left to complete



What date do we communicate?



How likely?



FORECASTING MODEL

How long = $\frac{\text{Size} * \text{Growth}}{\text{Pace}}$

DETERMINISTIC APPROACH

$$80 * 1$$

How long = $\frac{80 * 1}{10}$

PROBABILISTIC APPROACH

(math using ranges giving multiple results)

$$[80-100] * [1-3]$$

How long = $\frac{[80-100] * [1-3]}{[5-12]}$

this is why we need to use Monte Carlo simulations

Use Excel to run simulations...

THROUGHPUT FORECASTER



Troy Magennis

Focused Objective

<https://www.focusedobjective.com>

Forecast Duration and Completion Date

1. Start Date

8/11/03 (optional)

enter the values in
orange cells only

2. How many stories are remaining to be completed?

low guess 80
scope complexity Clear and understand
adjusted scope 80

high guess 100
(change this list and growth factors in the "Settings" worksheet)
100

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

Often the throughput/velocity in the backlog is pre-split, but captured completed stories post-splitting by the dev team making forecasts optimistic.

low split guess 1.00

highest split guess 1.00

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

Throughput/velocity data or estimate is for 1 week || 7 days

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

worst case 5 most often 8 (optional) best case 12

Team focus on THIS work

100% (only this work)

100%

Forecast Duration and Completion Date

1. Start Date

8/11/03 (optional)

enter the values in
orange cells only

2. How many stories are remaining to be completed?

low guess

80

high guess

100

scope complexity

Clear and understood

(change this list and growth factors in the "Settings" worksheet)

adjusted scope

80

100

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

Often the throughput/velocity in the backlog is pre-split, but captured completed stories post-splitting by the dev team making forecasts optimistic.

low split guess

1.00

highest split guess

1.00

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

Throughput/velocity data or estimate is for

1 week

7 days

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

Estimate

worst case

5

most often

8 (optional)

best case

12

Team focus on THIS work

100% (only this work)

100%

Results

No. of 1 week intervals

Likelihood	Duration	Date
100%	14	11/17/03
95%	13	11/10/03
90%	13	11/10/03
85%	12	11/3/03
80%	12	11/3/03
75%	12	11/3/03
70%	12	11/3/03
65%	12	11/3/03
60%	12	11/3/03
55%	11	10/27/03
50%	11	10/27/03
45%	11	10/27/03
40%	11	10/27/03
35%	11	10/27/03
30%	11	10/27/03
25%	11	10/27/03
20%	10	10/20/03
15%	10	10/20/03
10%	10	10/20/03
5%	10	10/20/03
0%	9	10/13/03

Almost certain

Somewhat certain

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

Using the Throughput Forecaster

SIZING THE BACKLOG

Methods for Sizing the Backlog

- ▶ Story Count Forecaster spreadsheet
- ▶ Reference class forecasting
- ▶ Story mapping
- ▶ Event storming
- ▶ Your current method i.e. how you do it today

Using the Throughput Forecaster

ACCOUNTING FOR GROWTH

Accounting for Growth

- ▶ Consider how work break downs. Ensure "units" between backlog and pace are the same.
- ▶ Ensure you document and account for risks
- ▶ Consider rate-based and time-based impacts
- ▶ How are you accounting for defects?
- ▶ Without any data, start with split rate of 1 to 3.

Using the Throughput Forecaster

ESTIMATING PACE

Estimating Pace

- ▶ Consider ramp-up and ramp down
- ▶ Need to see full range to be accurate, not just lower part of range
- ▶ Start with somewhere between 1x - 2x active developers
- ▶ Consider external dependencies and constraints

What about...?

USE THE MODEL TO
RUN DIFFERENT
SCENARIOS

Identifying and tracking these are key to success...

ASSUMPTIONS

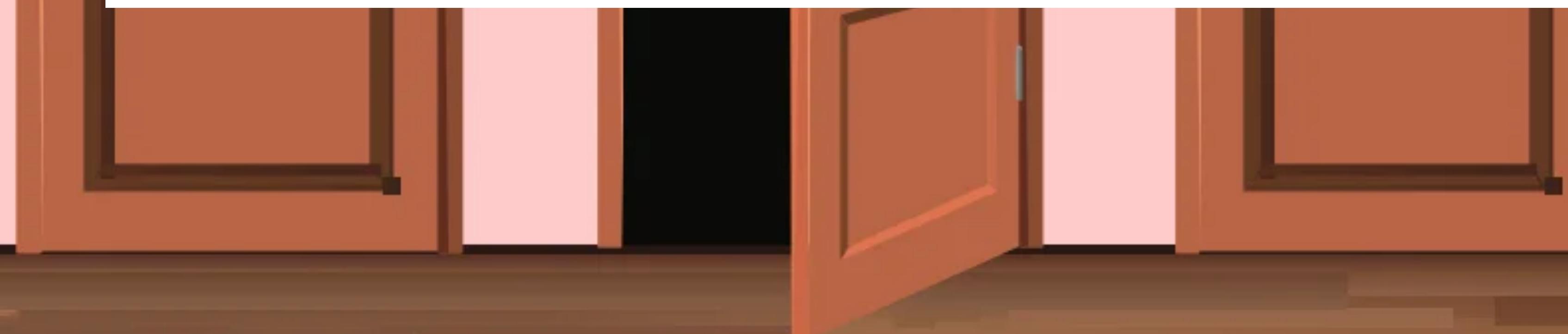
FORECASTS CONTAIN OUR ASSUMPTIONS



Forecasts aren't just numerical or date values. A forecast is often a **numerical value** along with the **assumptions** that allow that forecast to be reliable.

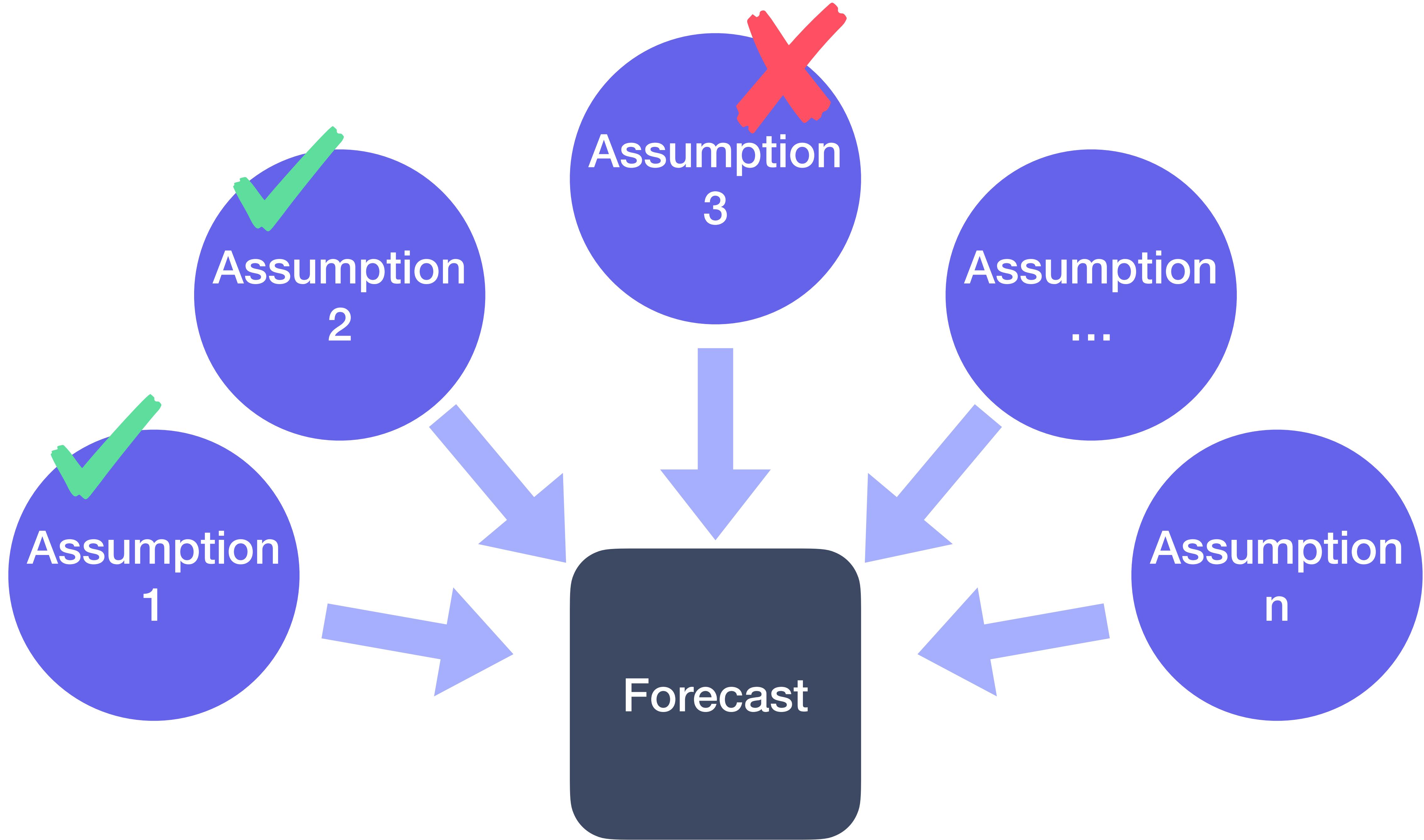
Don't Switch! Why Mathematicians' Answer to the Monty Hall Problem is Wrong

The Monty Hall problem is one of those rare curiosities – a mathematical problem that has made the front pages of national news. Everyone now knows, or thinks they know, the answer but a realistic look at the problem demonstrates that the standard mathematician's answer is wrong. The mathematics is fine, of course, but the assumptions are unrealistic in the context in which they are set. In fact, it is not clear that this problem can be appropriately addressed using the standard tools of probability theory and this raises questions about what we think probabilities are and the way we teach them.



“The mathematics is correct, so you do indeed seem to double your chances by switching but only provided certain assumptions hold. ...there are actually a number of assumptions:

- 1. Monty will always open a door.**
- 2. Monty never opens the door you have chosen.**
- 3. Monty never opens the door with the car behind it.**
- 4. The car is equally likely to be behind any door.**
- 5. Given a choice of doors, Monty chooses at random.”**



A minimum set of assumptions should cover these:

1. The ability to start delivery
2. Initial scope and size.
3. Expected rework and scope growth.
4. Expected progress of scope delivery over time.
5. Acceptable quality to be able to deliver to customers.
6. Ability to deliver to customers (environments, process, logistics).



We need to account for new information...

RE-FORECASTING / CONTINUOUS FORECASTING

MonteCarlo

Team	Release	Code Freeze Date ▲	Stories Remaining	85% Completion Date	Completion Likelihood	Features Remaining
Hiring Integration	January 2017	01/10/2017	12	01/25/2017	37.42%	3
Identity Management	January 2017	01/31/2017	5	05/12/2017	24.56%	2
Compliance	January 2017	01/31/2017	1	01/10/2017	99.99%	0
Recruiting	January 2017	01/31/2017	35	01/03/2017	99.99%	5
Foundation Services	January 2017	01/31/2017	4	01/03/2017	99.90%	1
Onboarding	January 2017	01/31/2017	8	02/22/2017	28.46%	0
UltiPro TouchBase	January 2017	01/31/2017	27	01/17/2017	99.94%	5
Integration Services	January 2017	01/31/2017	5	02/03/2017	80.58%	0
People Domain	January 2017	01/31/2017	17	12/21/2016	99.99%	2
Bulk Data Export Services	February 2017	02/28/2017	47	03/24/2017	25.06%	8
Data Warehouse	February 2017	02/28/2017	5	12/21/2016	99.99%	0
Business Intelligence	February 2017	02/28/2017	12	01/17/2017	99.99%	4
Identity Management	V12.2.1 (R1 - Spring 2017)	03/15/2017	10	12/21/2017	0.74%	0
Foundation Services	V12.2.1 (R1 - Spring 2017)	03/15/2017	63	09/12/2017	0.01%	4
Hiring Integration	V12.2.1 (R1 - Spring 2017)	03/15/2017	6	02/08/2017	99.90%	0
UTM	V12.2.1 (R1 - Spring 2017)	03/15/2017	2	12/08/2016	99.99%	6
Payroll	V12.2.1 (R1 - Spring 2017)	03/15/2017	18	12/28/2016	99.99%	3
Talent Management	V12.2.1 (R1 - Spring 2017)	03/15/2017	57	02/17/2017	99.82%	6
UTA	V12.2.1 (R1 - Spring 2017)	03/15/2017	1	04/20/2017	48.62%	1
Compliance	V12.2.1 (R1 - Spring 2017)	03/15/2017	170	03/02/2017	98.30%	45
People Domain	V12.2.1 (R1 - Spring 2017)	03/15/2017	69	02/06/2017	99.99%	13
Tax Management Platform	TM-Beta	04/01/2017	111	11/15/2017	0.01%	25
Tax Management Integration	TM-Beta	04/01/2017	96	10/19/2017	0.01%	22
Tax Management	TM-Beta	04/01/2017	22	03/15/2017	96.96%	17
Identity Management	V12.1.2 (R2 - Fall/YE 2016)	09/29/2016	1	N/A	0.00%	0
SPS	SPS-CR	09/30/2017	40	02/13/2017	99.99%	8

Top 3 Reasons Your Forecasts May Fail

1. Start date incorrect

Teams are almost never fully available on day one.

2. Work breakdown not considered

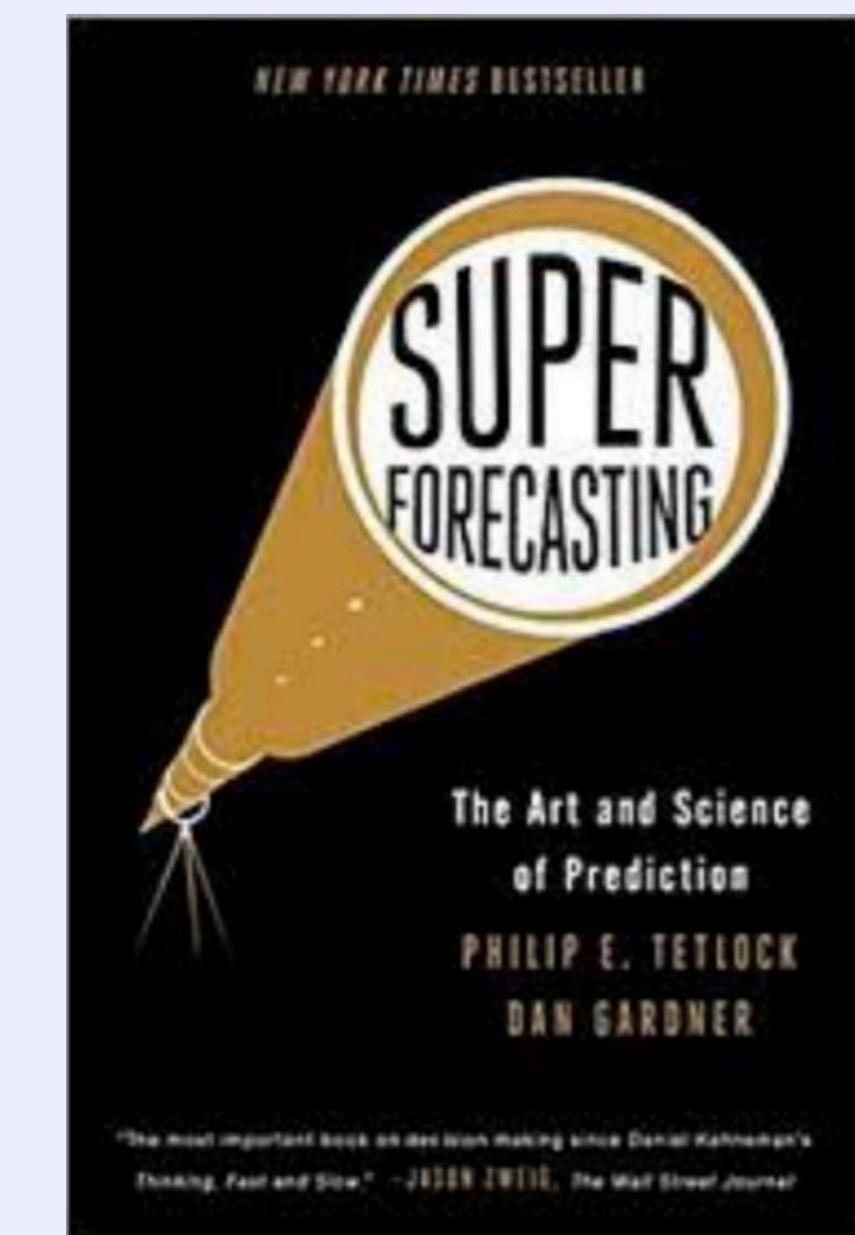
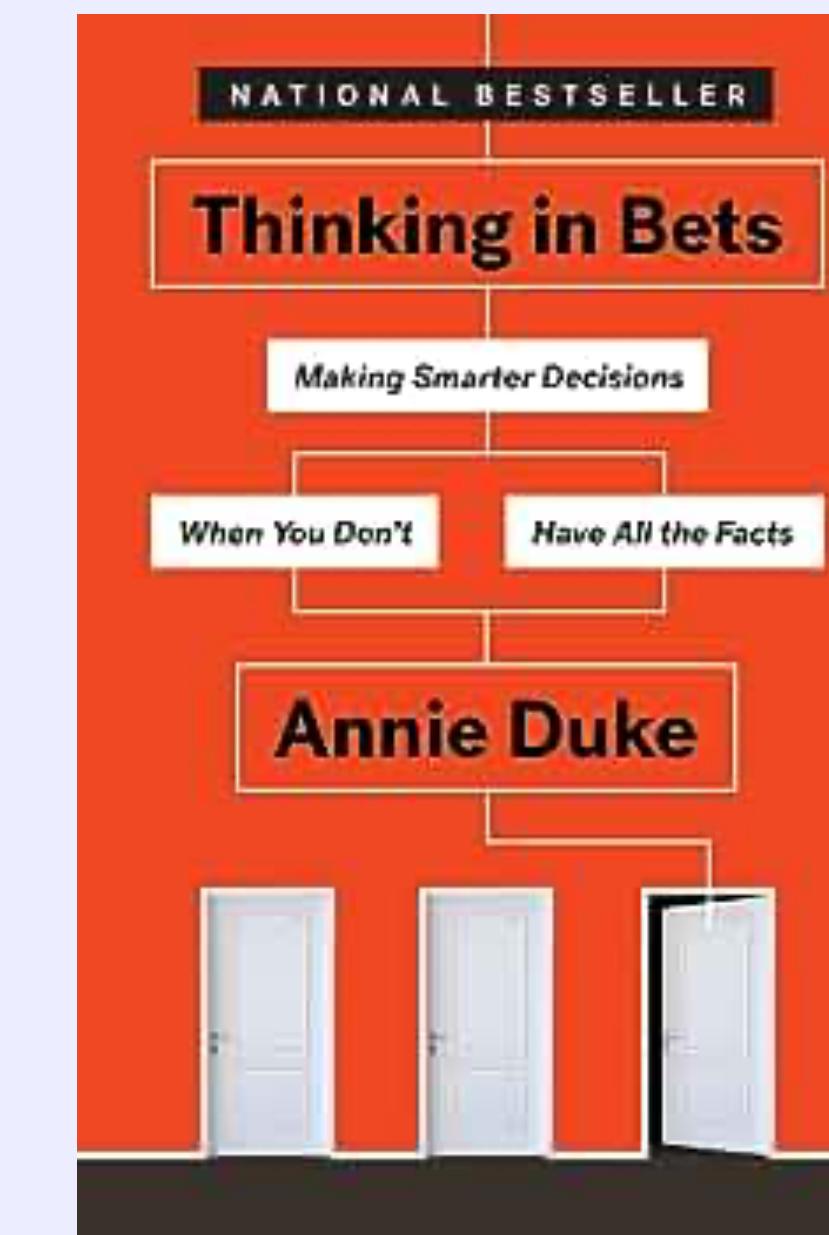
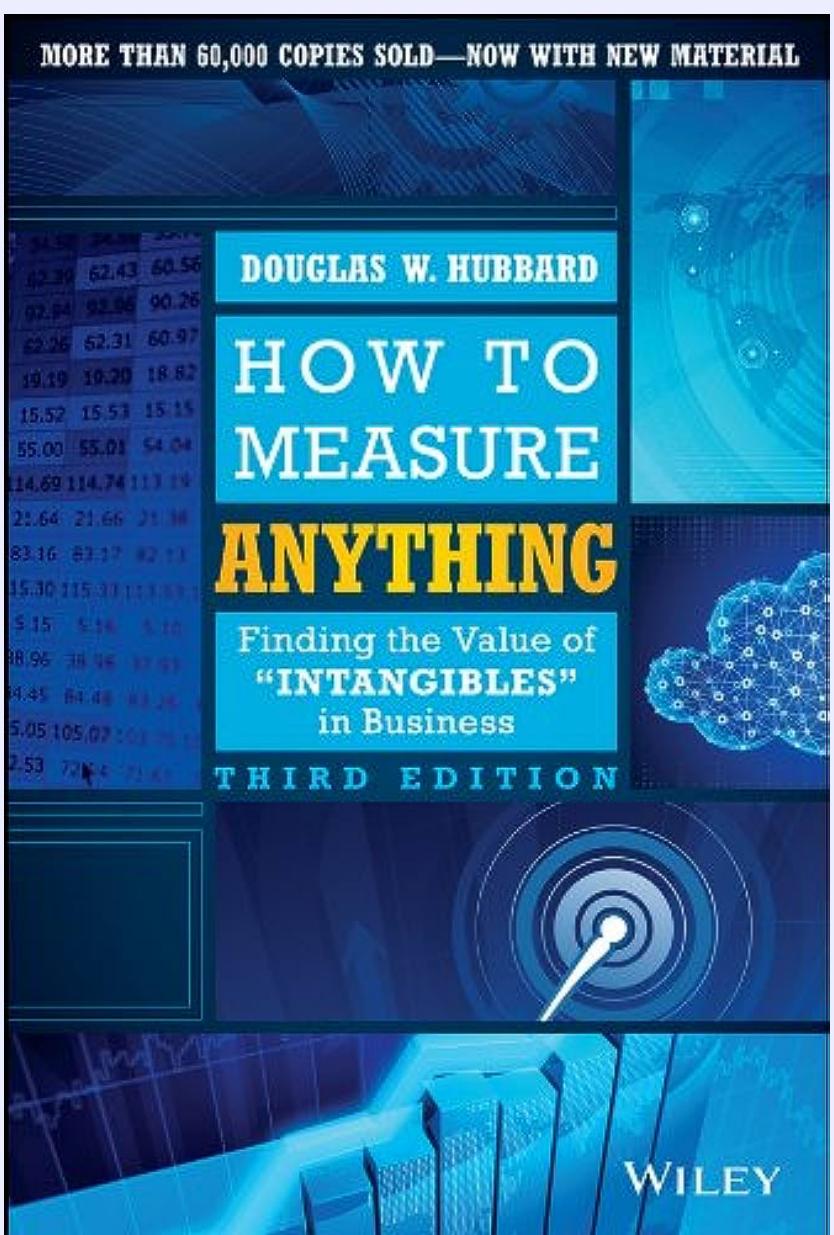
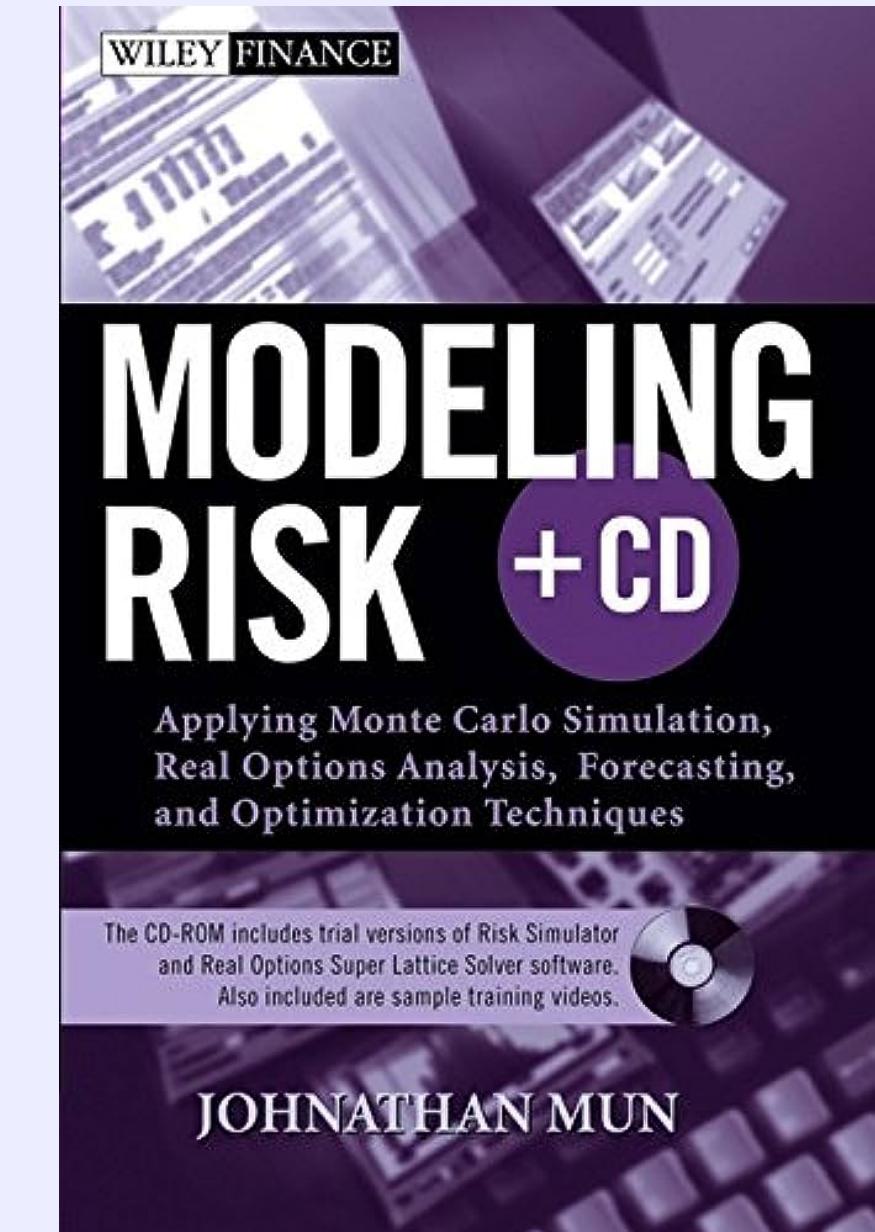
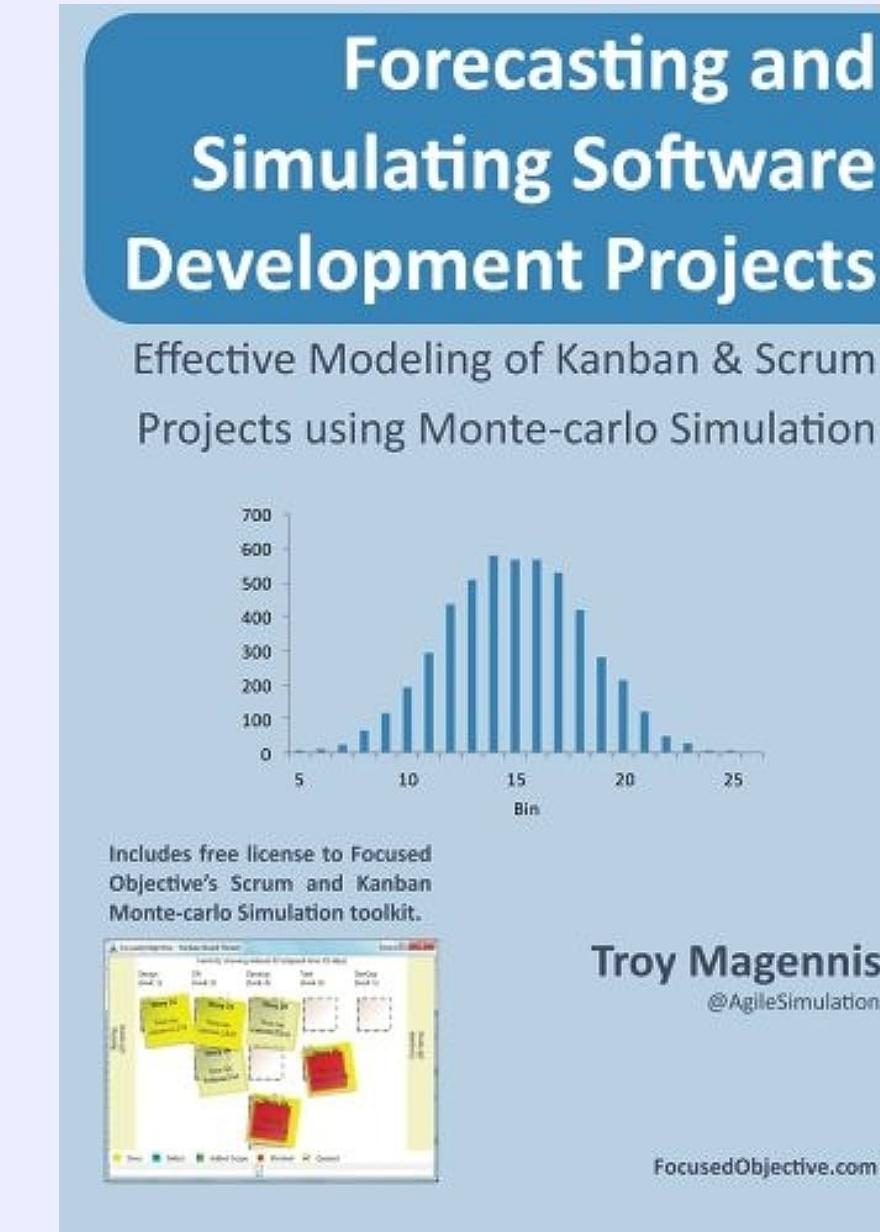
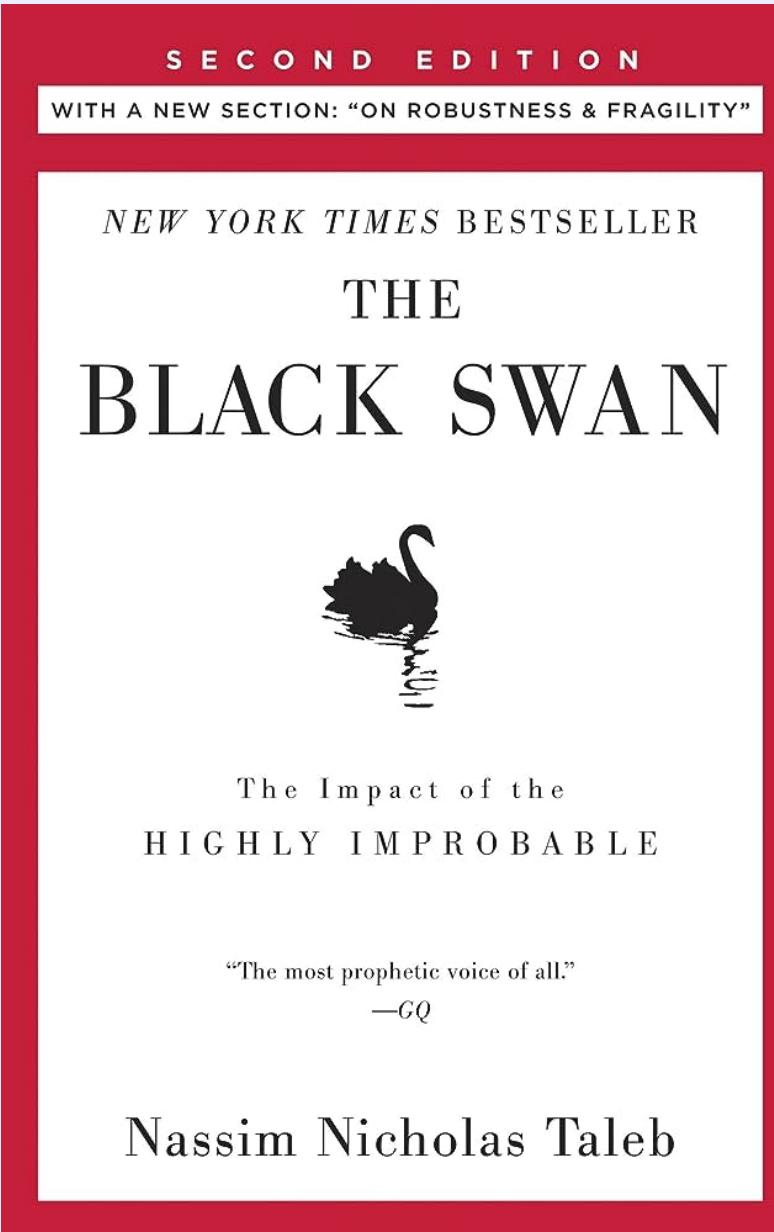
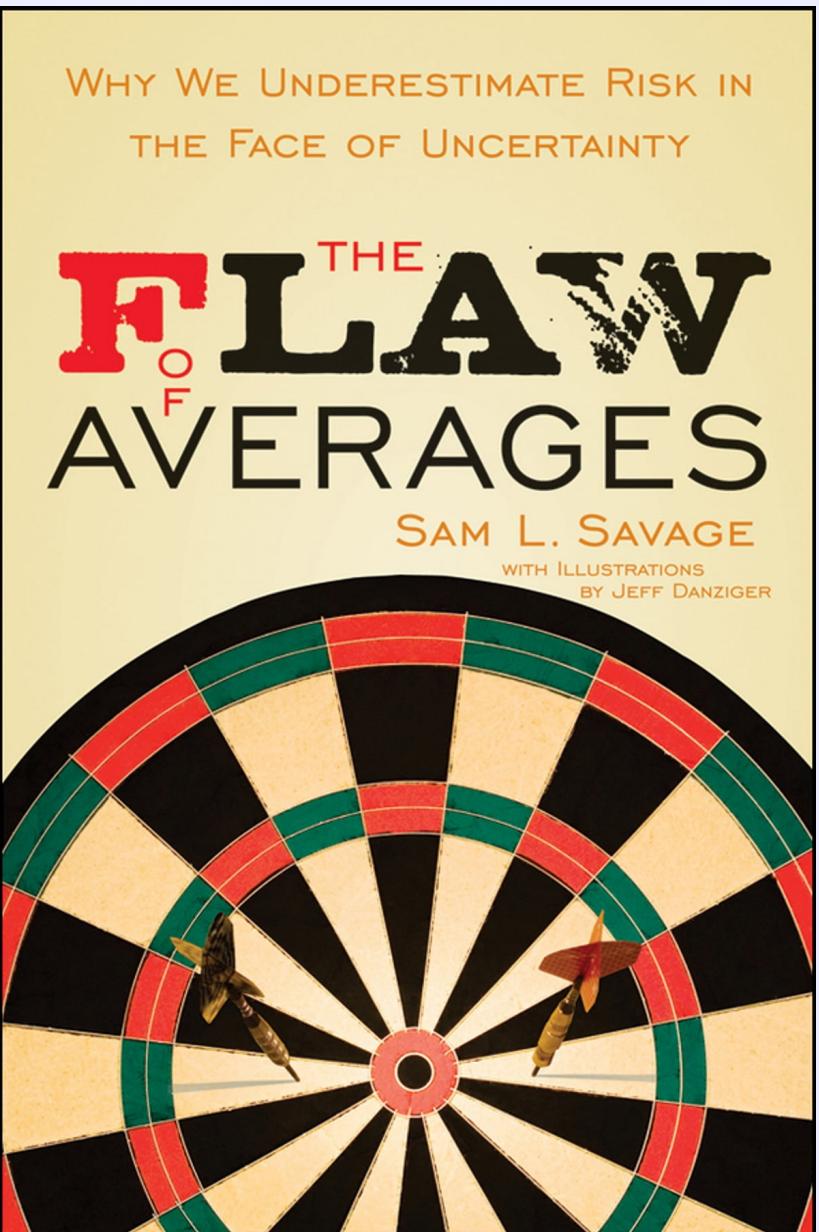
We're too optimistic if work breakdown isn't properly considered.

3. Failing to account for risks and dependencies:

Work that “might” need to be done but we don’t know for sure yet

TAKEAWAYS

- Be careful when using averages
- A forecast is a statement about a future outcome or unknown event AND a level of uncertainty
- Construct a model given your current understanding of reality
- Document assumptions and check often
- The math works... use the forecast to facilitate better and different conversations than you're having today

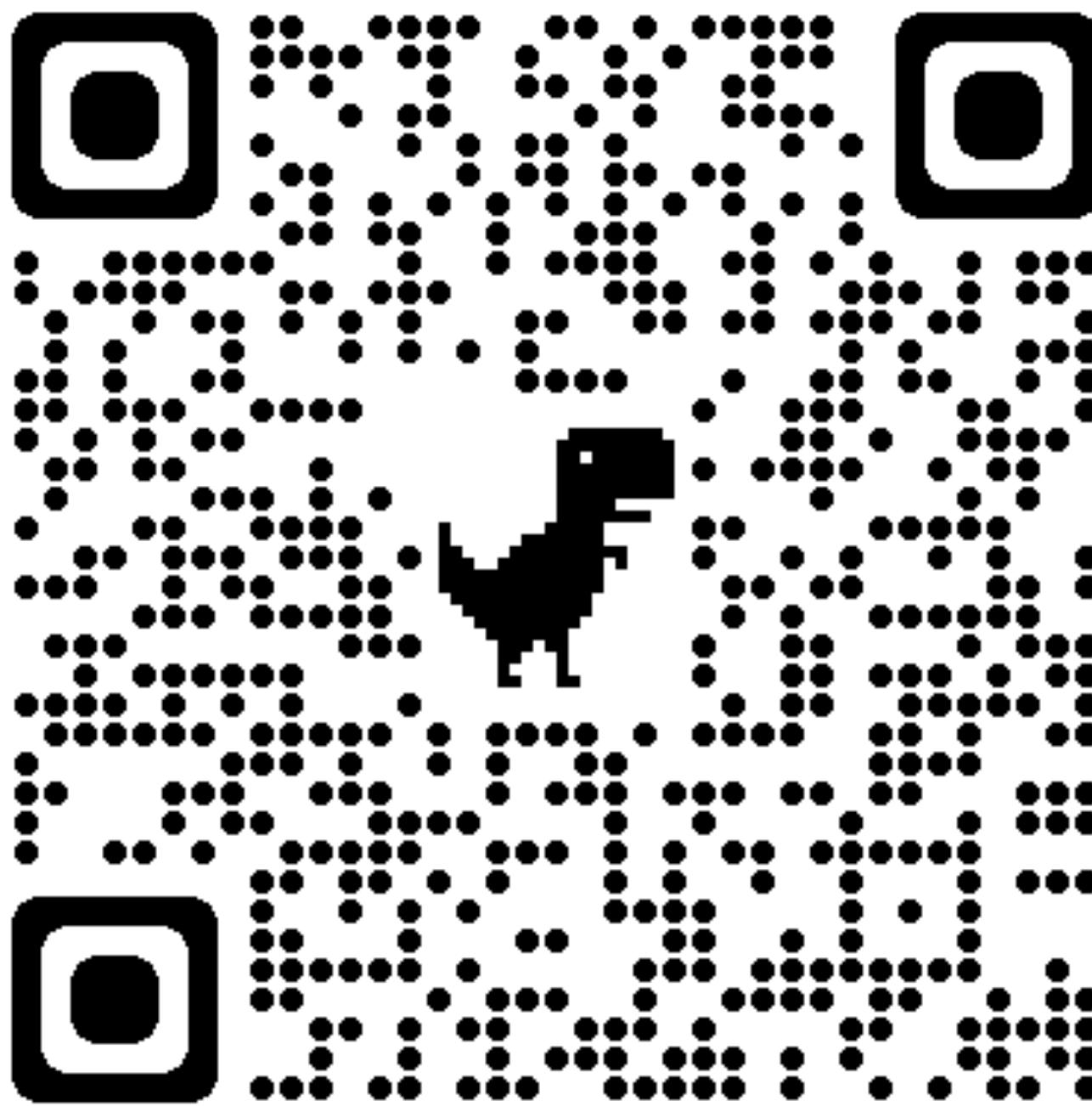


CHRIS SHINKLE

chrisshinkle@sep.com

LinkedIn: www.linkedin.com/in/chrisshinkle

Twitter: @chrisshinkle



slides and spreadsheets:
info.sep.com/codemash