Motivating Precalculus and Calculus Concepts with Data

Christopher Hayes
Associate Professor, Mathematics & Data Science Program
Gallaudet University

Before we start

Suppose you ask a student six months after Precalculus what a quadratic function is *used for*.

Before we start

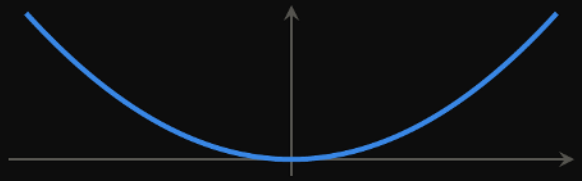
Suppose you ask a student six months after Precalculus what a quadratic function is *used for*.

What do you think they'd say?

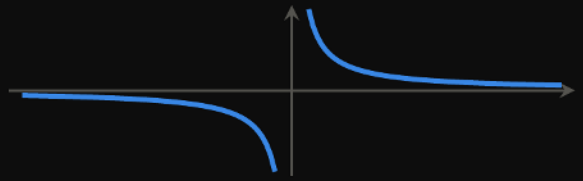
Review: Parent Functions

Remember, here are some parent functions.

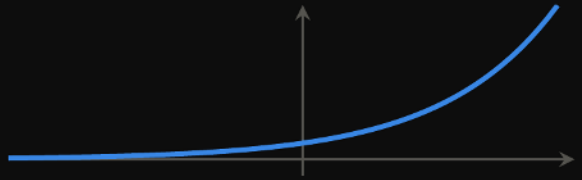
$f(x) = x^2$



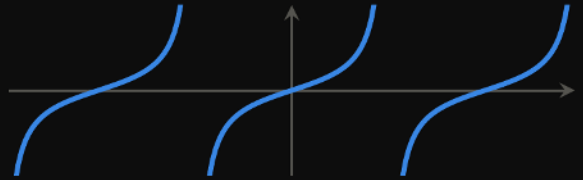
$f(x) = 1/x$



$f(x) = e^x$

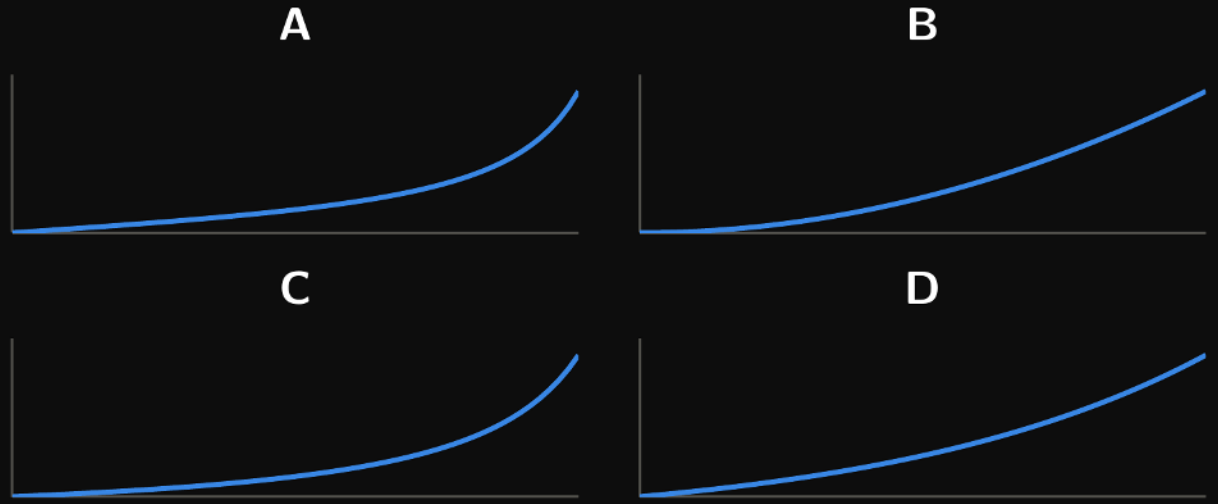


$f(x) = \tan x$



Same General Shape?

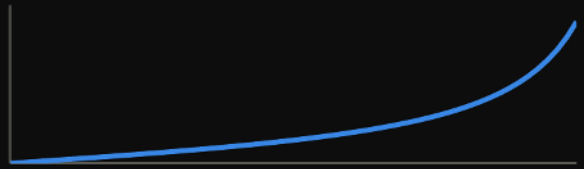
Here they are, randomized, transformed, and cropped. Which is which?



Same General Shape?

All four pass through the points $(0, 0)$ and $(5, 5)$.

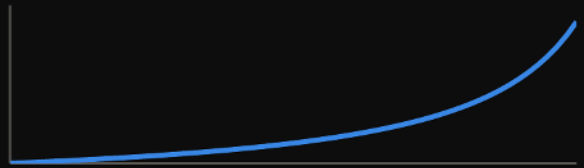
A: $\tan(0.275x)$



B: $0.2x^2$



C: $x/(6-x)$

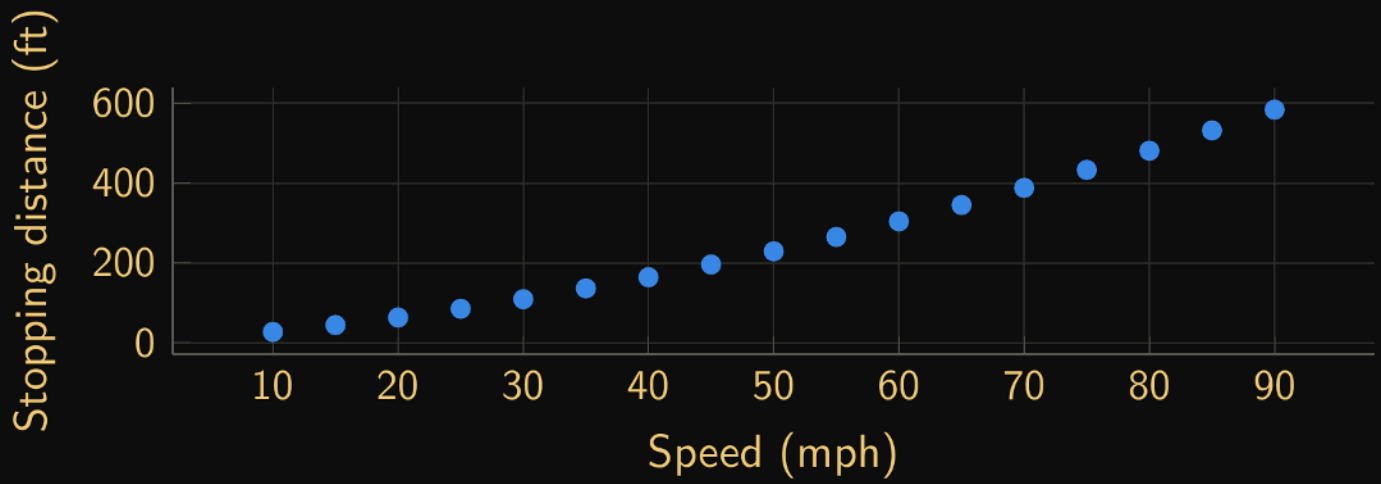


D: $e^{0.358x} - 1$



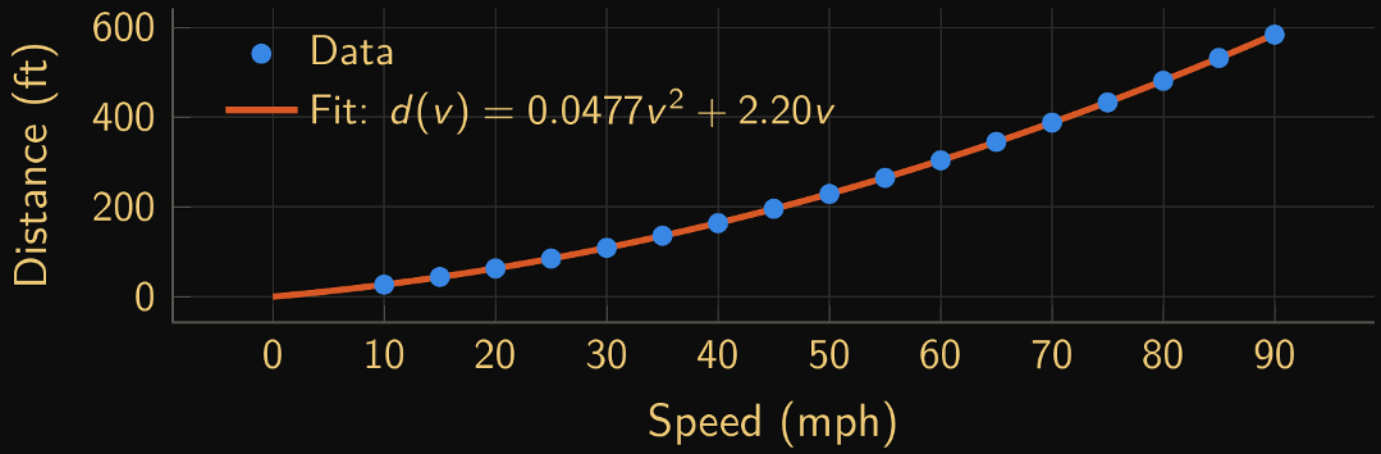
Fitting Data & Picking Models

Real data: total stopping distance for a car at different speeds.



Which parent function makes the most sense here? Why that one?

Fitting Data & Picking Models



A quadratic fits well. But most of our parent functions could look fine too.

Is quadratic really the best?

“All models are wrong, some models are useful"

What are we doing in precalculus and calculus?

- Several curves can fit one dataset. Fit alone doesn’t settle the question.
- We could justify a certain parent function with an intuitive model. This often looks to differential equations (physics).
- Another priority may be to value computational speed.

What Do Precalculus and Calculus I cover?

- Polynomials, exponentials, logarithms, trig, limits, derivatives, integrals.
- Sometimes sequences, series, vector and matrix algebra.
- The content is largely set by norms (frozen curriculum).
- Exams and courses create an emphasis on rote practice.
- Math first, models as supplements, data almost never.

What Do Students Remember?

QUICK DISCUSSION

- A semester or a year later, what do you believe your students tend to remember from mathematics courses (not just precalculus and calculus)?
- What do they best connect it to, in your experience?

My Experience

- Brief motivation, concept, and definition, then worked examples, then word problems.
- Word problems are not the central focus.
- Concepts ‘feel’ understood in the classroom but are not retained.
- Students remember data sets and real-world experiences.

The Idea, Part 1: Data First

Start the lesson with a real dataset.

You can find it online, generate one, or have students create it themselves.

The Idea, Part 2: Concept

Organize the unit around the idea the data calls for: rate of change, accumulation, periodicity, proportional growth.

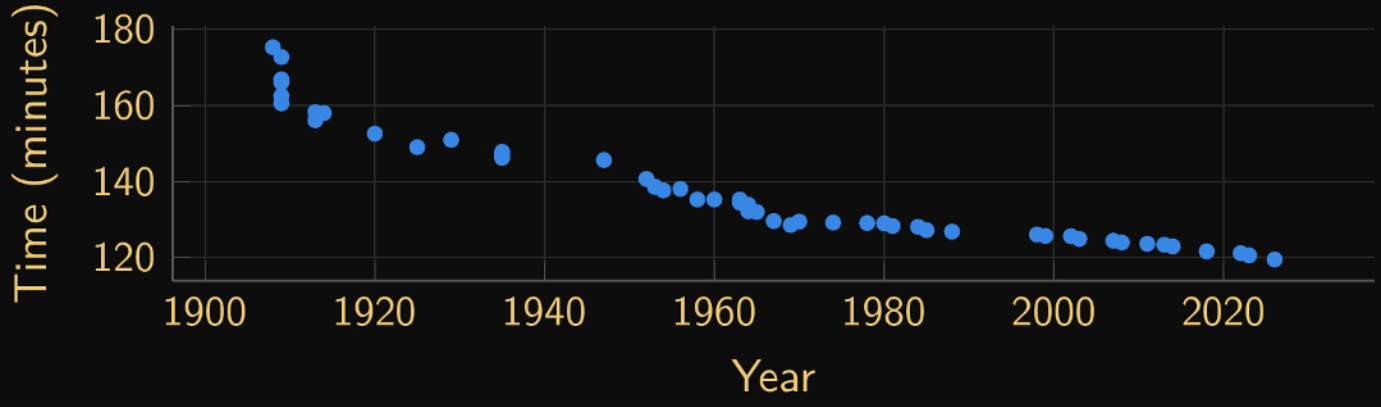
The procedure comes in as the tool for that idea.

Why Try It?

1. Usually, the origin of data can be explained straightforwardly.
2. Current tools make best-fit curves easy (Desmos or AI scripts).
3. Interpreting output and functions with complicated coefficients becomes more intuitive, and realistic.

Marathon World Records

The marathon is a 26.2 mile road race. Each point is a new world record in minutes, plotted by year.



What do you notice about the years after 1950?

Source: Wikipedia, marathon world record progression (men's records shown)

Activity 1: Pick a Model for the Marathon

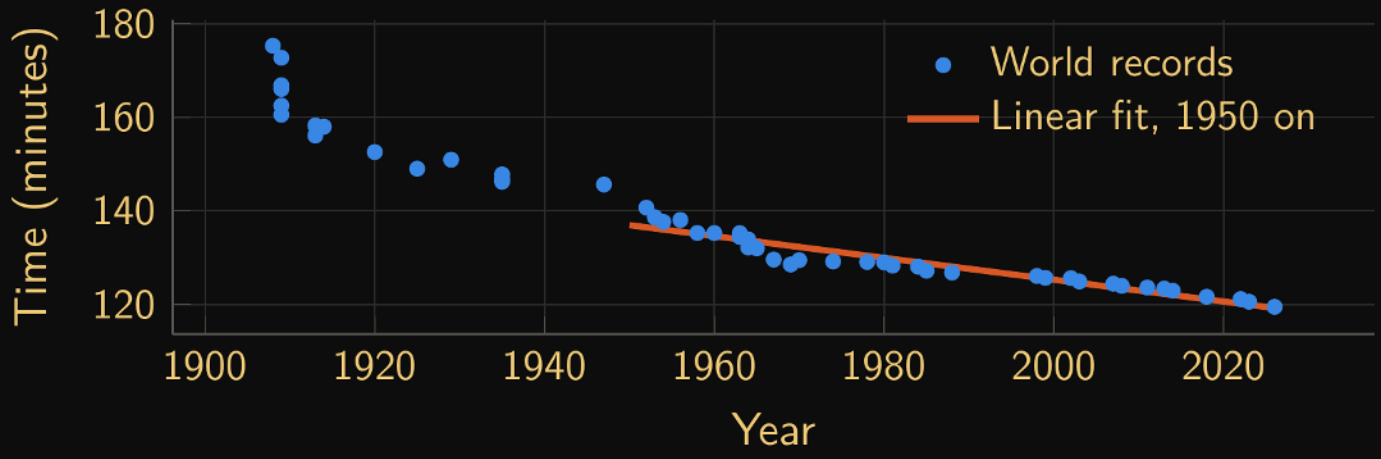
1. What model would you pick for this data? Why?
2. Find the best fit for your model.
3. Suppose your goal is to predict when the first 100-minute marathon will happen. How would you do this?
4. What is your prediction?



Scan for the data in Desmos

desmos.com/calculator/7fkxyv3sou

Marathon Records: A Linear Model



Since 1950: roughly 14 seconds faster per year, for 75 years.

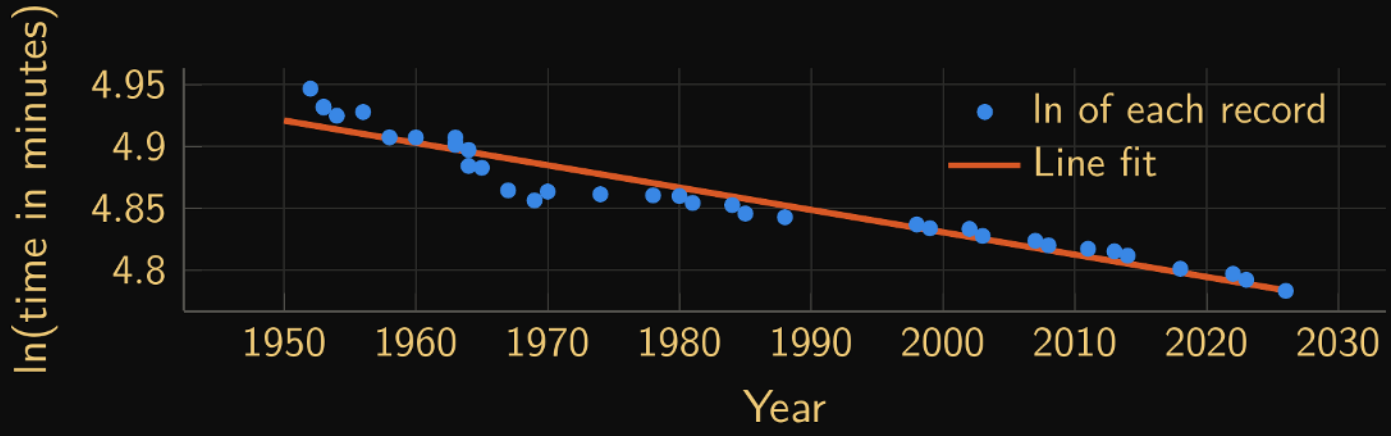
The slope has units: minutes per year.

Marathon Records: A Linear Model

- Inside the data, the line reads well: about -0.23 minutes per year.
- Extrapolate and the model predicts a zero-minute marathon around the year 2540!
- What function family would respect a lower limit?

Marathon Records: Log Scale

Let's replace each time T with its natural log, $\ln(T)$.



With t = years since 1950: $\ln T = 4.921 - 0.00181\,t$

What is the benefit of a log scale?

Marathon records: back to an exponential

Undo the log by raising e to both sides and simplifying:

$$T = e^{4.921 - 0.00181\,t} \approx 137.1\,(0.9982)^t$$

- 137.1 minutes: the model's record time in 1950.
- 0.9982: each year the record is multiplied by 0.9982.
- So the record drops about **0.18% per year**, not a fixed number of minutes.
- Right now 0.18% of 120 minutes is about 13 seconds. As times get faster, the yearly drop shrinks.

Marathon records: minutes vs. percent

	Linear	Exponential
Each year the record...	drops 14 seconds	drops 0.18%
Predicted record, 2100	102.0 min	104.6 min
Predicted record, 2540	0 min	47.3 min

The exponential model won't equal zero in finite time. But it still has a problem.

What might a better model look like?

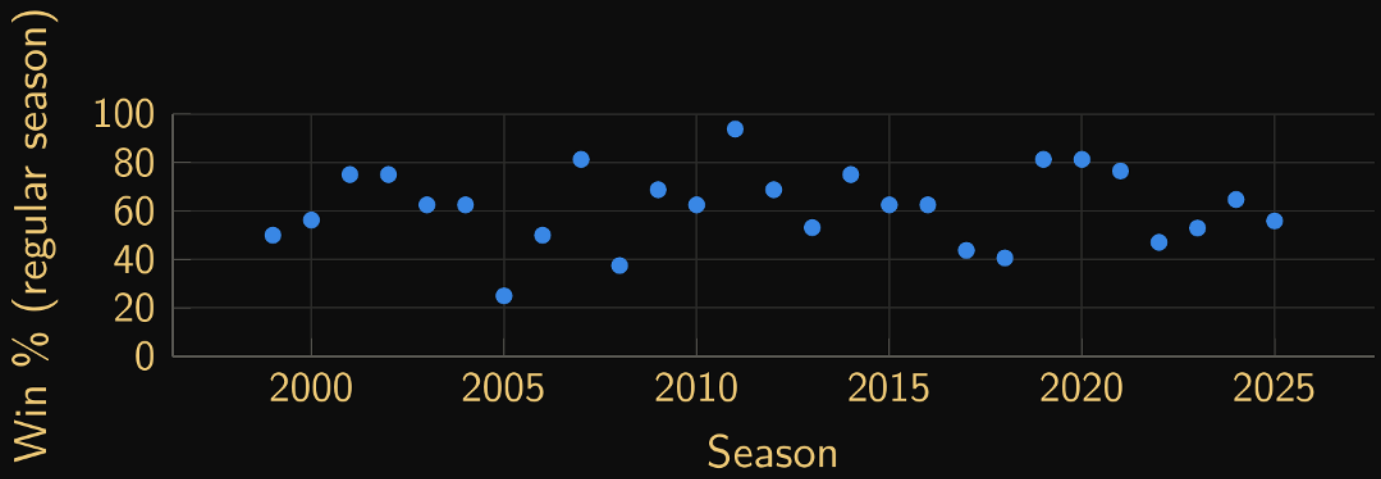
Back to Your 100-Minute Prediction

Model (fit to 1950 on)	First 100-minute marathon
Linear	about 2109
Exponential	about 2125
Your model	_____

- Did different models give very different years?
- Which prediction do you trust most, and why?

Predicting Football Win Percentages

The Green Bay Packers: share of games won each season.

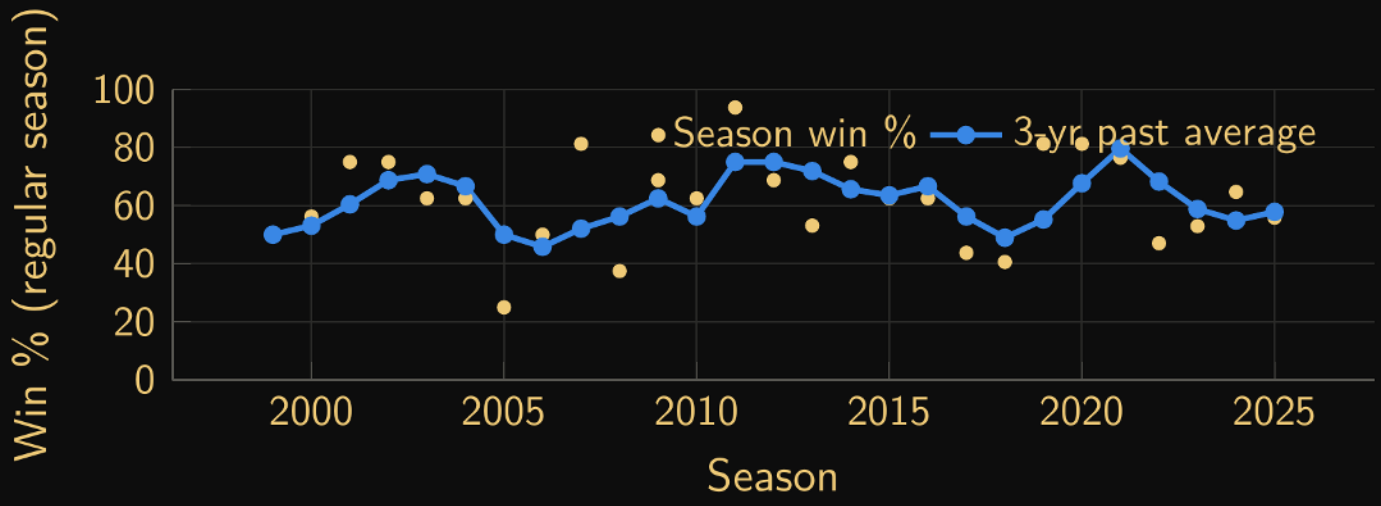


Fans talk about title runs and rebuilding years. Do you see cycles?

Source: nflverse, a public football data archive; regular seasons 1999–2025

Win percentage: 3-year past averages

I think it's helpful to replace each season with the average of the past three seasons.



Now do you see cycles?

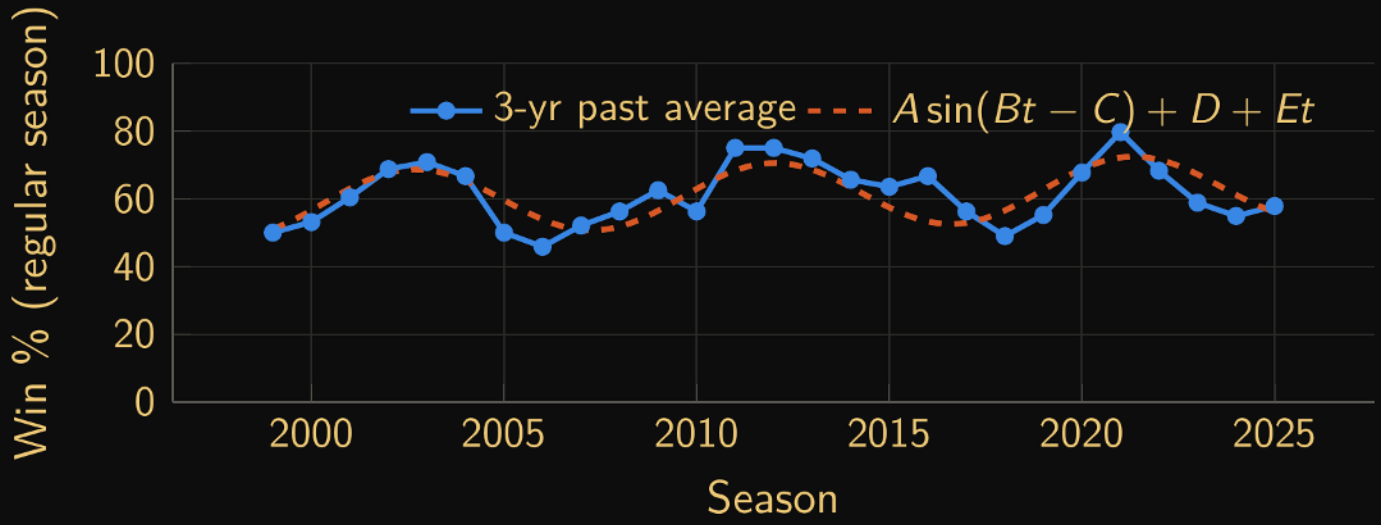
Activity 2: The Packers' Next Peak

1. What model would you pick for this data? Why?
2. Find the best fit for your model. Will you fit the season win % or the 3-year averages?
3. Suppose your goal is to predict when the Packers will be at their best again. What does "at their best" mean in your model?
4. What is your prediction? Which season, and what win %?

Scan for the data in Desmos

desmos.com/calculator/5olx09tryv

A Model: Try a Line + Sine



Best fit: amplitude 9 points, period 9.3 seasons, trend +0.2 point per season.

Win percentage: the fitted numbers

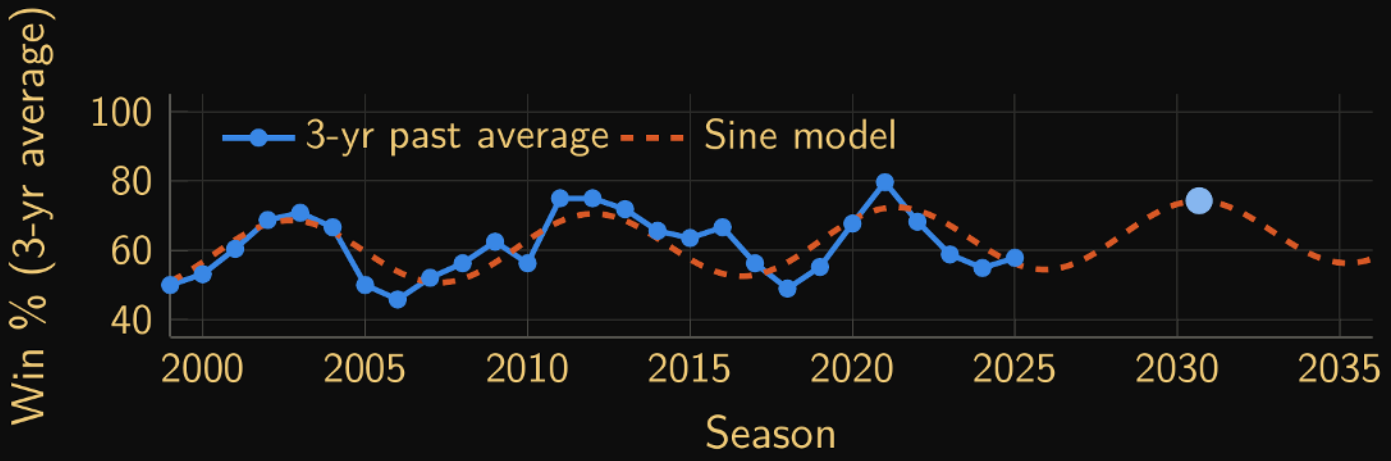
With $t =$ seasons since 1999 and W in percent:

$$W(t) = 9.5 \sin(0.674\, t - 0.90) + 58.5 + 0.20\, t$$

- $A = 9.5$
- $B = 0.674$ per season, so the period is $2\pi/B \approx 9.3$ seasons
- $C = 0.90$
- $D = 58.5$
- $E = +0.20$ per season

What would A, B, C, D, E each mean here?

Back to Your Packers Prediction



The model's next peak is around the **2030–31 season**, at about **74%** (12–13 wins). Its past peaks lined up with 2002, 2012, and 2021.

Would you bet on it? What could break the cycle?

Thank You!
Any Questions?