

# “Lies, Damned Lies, and Teens Who Smoke”

Daniel S. Vacanti  
@danvacanti



**ProKanban.org**

# Thank you to Dr. Donald Wheeler



MALCOLM  
GLADWELL  
Revisionist  
History

“The Big Man  
Can’t Shoot”

June 2016

# “The Greatest Basketball Game Ever Played” ...



Players Who Scored 60 or More Points					
Points	Player	Team	Season	W/L	Game
100	Wilt Chamberlain	Philadelphia Warriors	1961-62	W	169-147 vs. Knicks
81	Kobe Bryant	Los Angeles Lakers	2005-06	W	122-104 vs. Raptors
78	Wilt Chamberlain	Philadelphia Warriors	1961-62	L	147-151 3OT vs. Lakers
73	David Thompson	Denver Nuggets	1977-78	L	137-139 @ Pistons
73	Wilt Chamberlain	San Francisco Warriors	1962-63	W	127-111 @ Knicks
73	Wilt Chamberlain	Philadelphia Warriors	1961-62	W	135-117 vs. Packers
72	Wilt Chamberlain	San Francisco Warriors	1962-63	L	115-127 @ Lakers
71	David Robinson	San Antonio Spurs	1993-94	W	112-97 @ Clippers

Hershey, Pennsylvania March 1962

# About Wilt...



“That night in Hershey, PA he’s an incredible free throw shooter.”

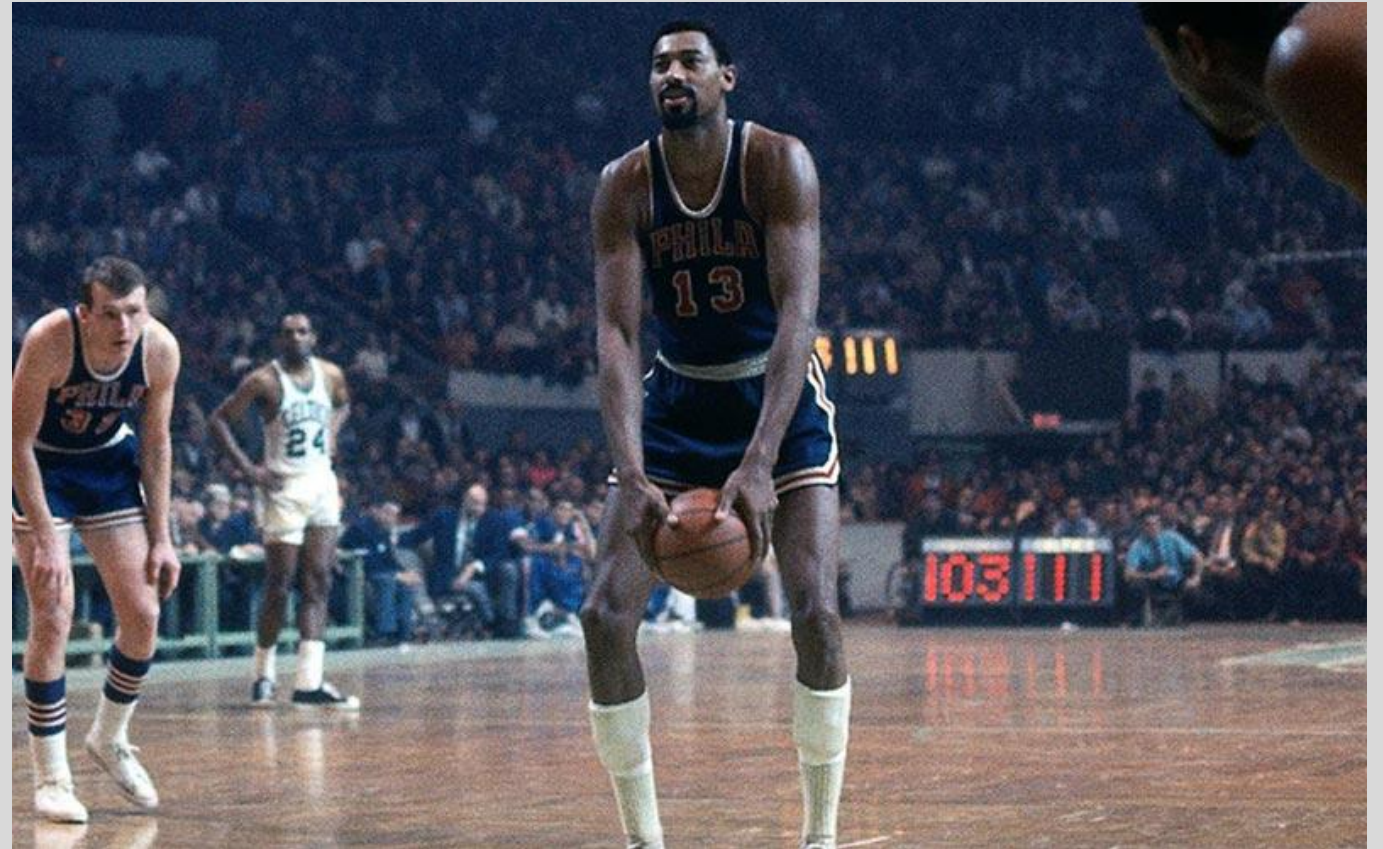
Career Free Throw Rate -- 51%

100-point game -- 87.5%

(someone remind me to talk about the problem with this comparison later)



But in 1962...



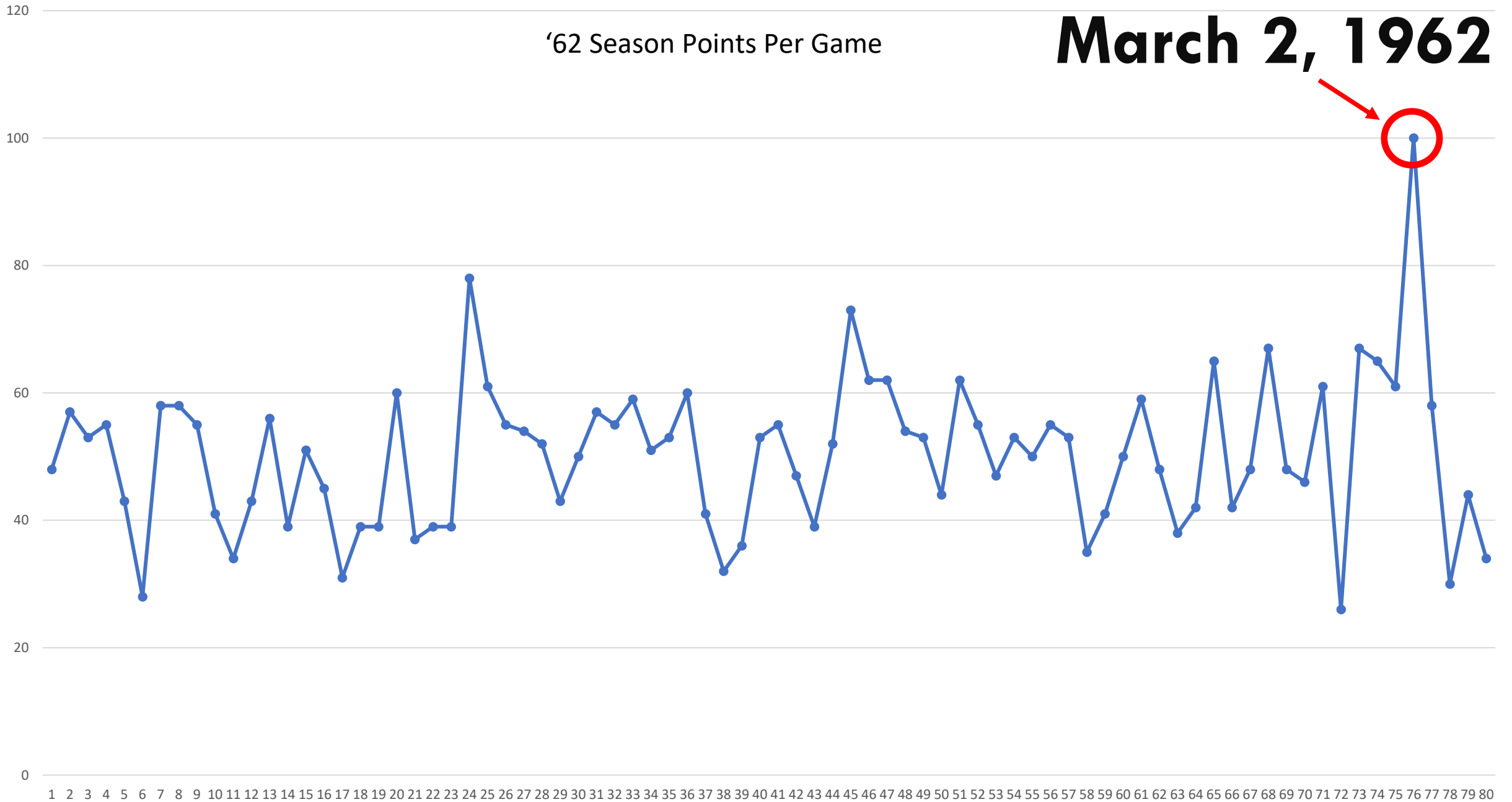
“...Wilt Chamberlain switches to a better [free throw] shooting technique. It pays off in the greatest basketball game ever played.”

But were free throws really what made the difference?



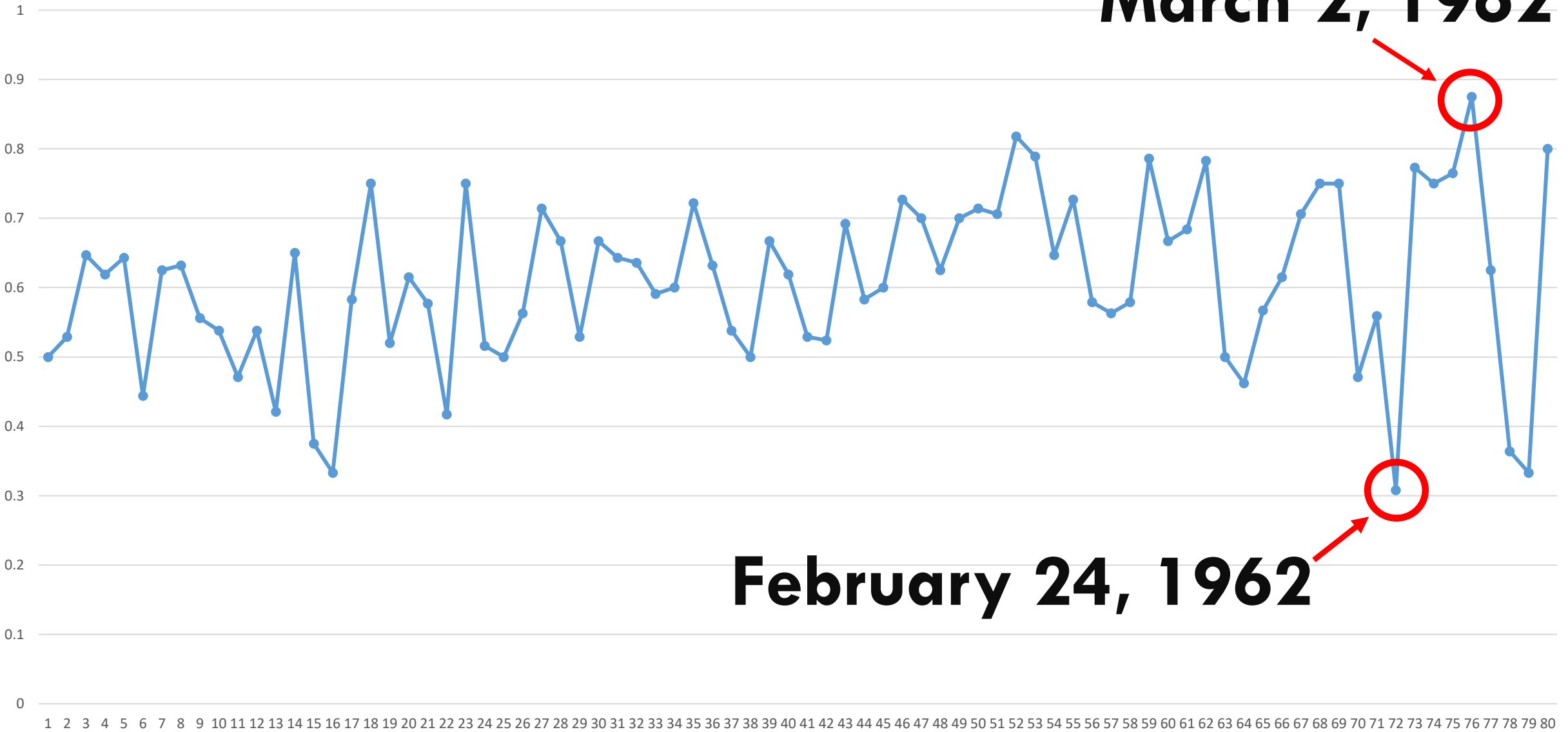
'62 Season Points Per Game

**March 2, 1962**



'62 Season FT% - Shooting Underhand

**March 2, 1962**



**February 24, 1962**

Source: <https://www.basketball-reference.com/players/c/chambwi01/gamelog/1962>

All data have noise.

Some data have signal(s).

The primary goal of data analysis  
is to separate  
the possible signal  
from  
the probable noise.

Mistake noise for signal.

Mistake signal for noise.

The trick is to separate

the possible signal

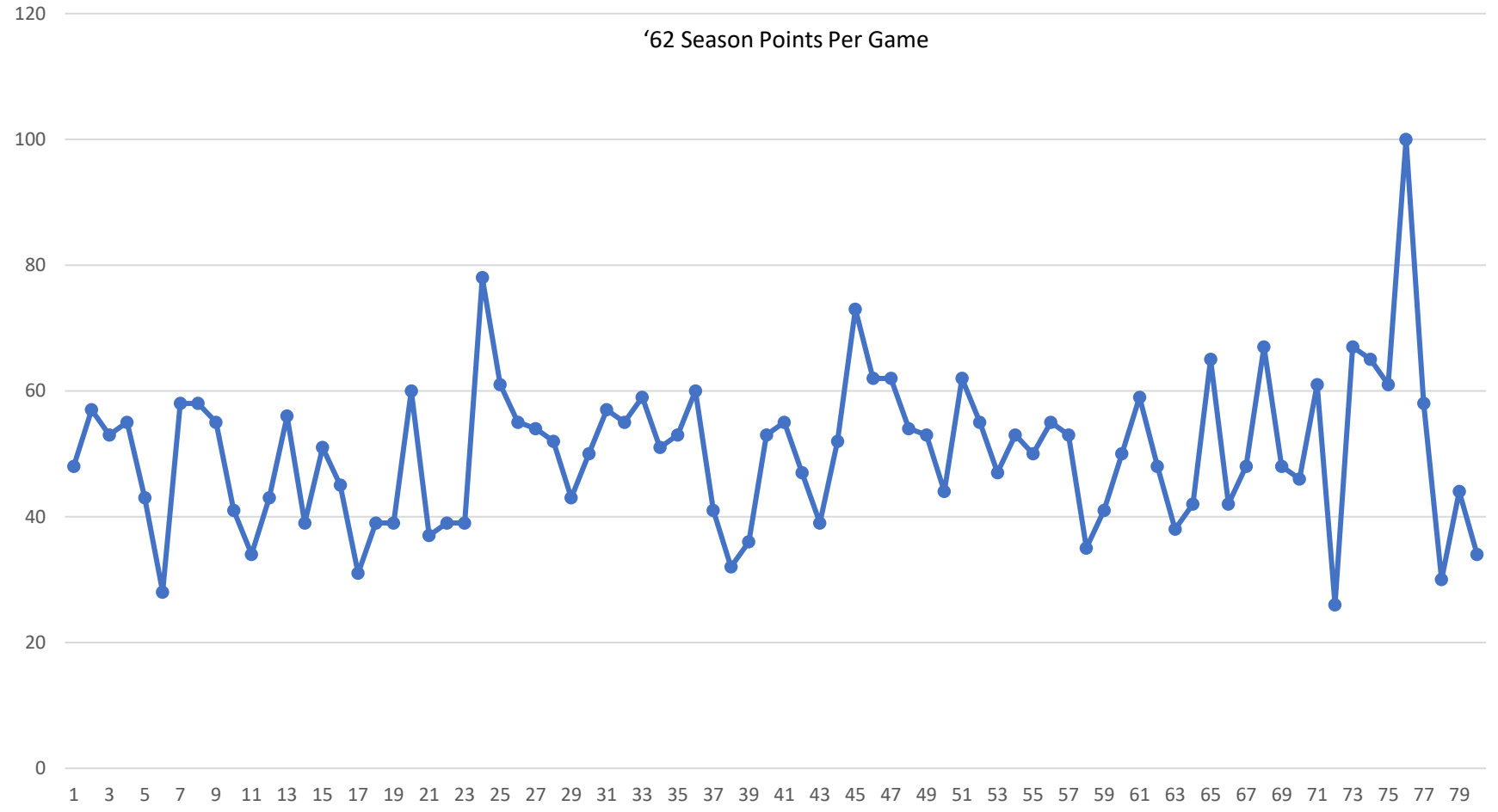
from

the probable noise

in such a way that we minimize the economic impact of making either or both mistakes.

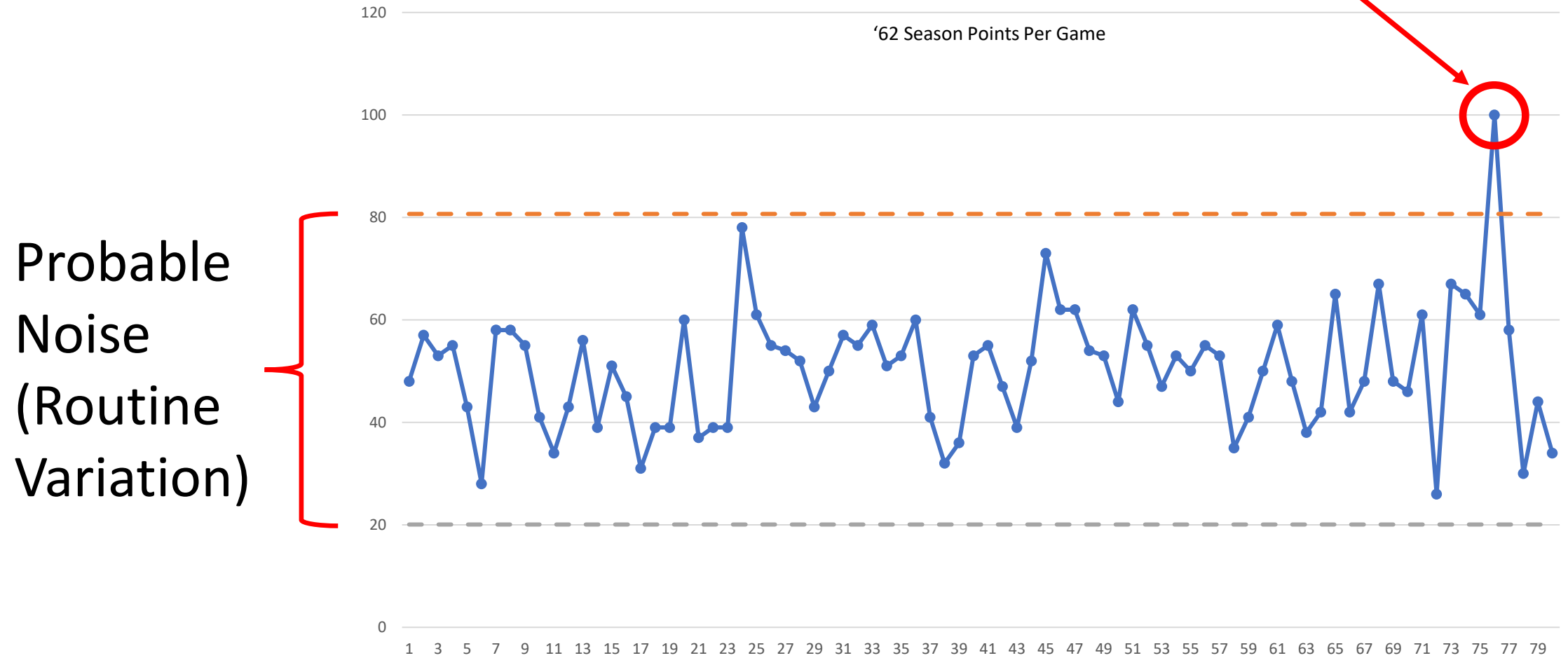


Source: <http://magazine.amstat.org/blog/2009/09/01/waltershewhartsep09/>





# Possible Signal (Assignable Cause)



# Process Behaviour Chart (PBC)

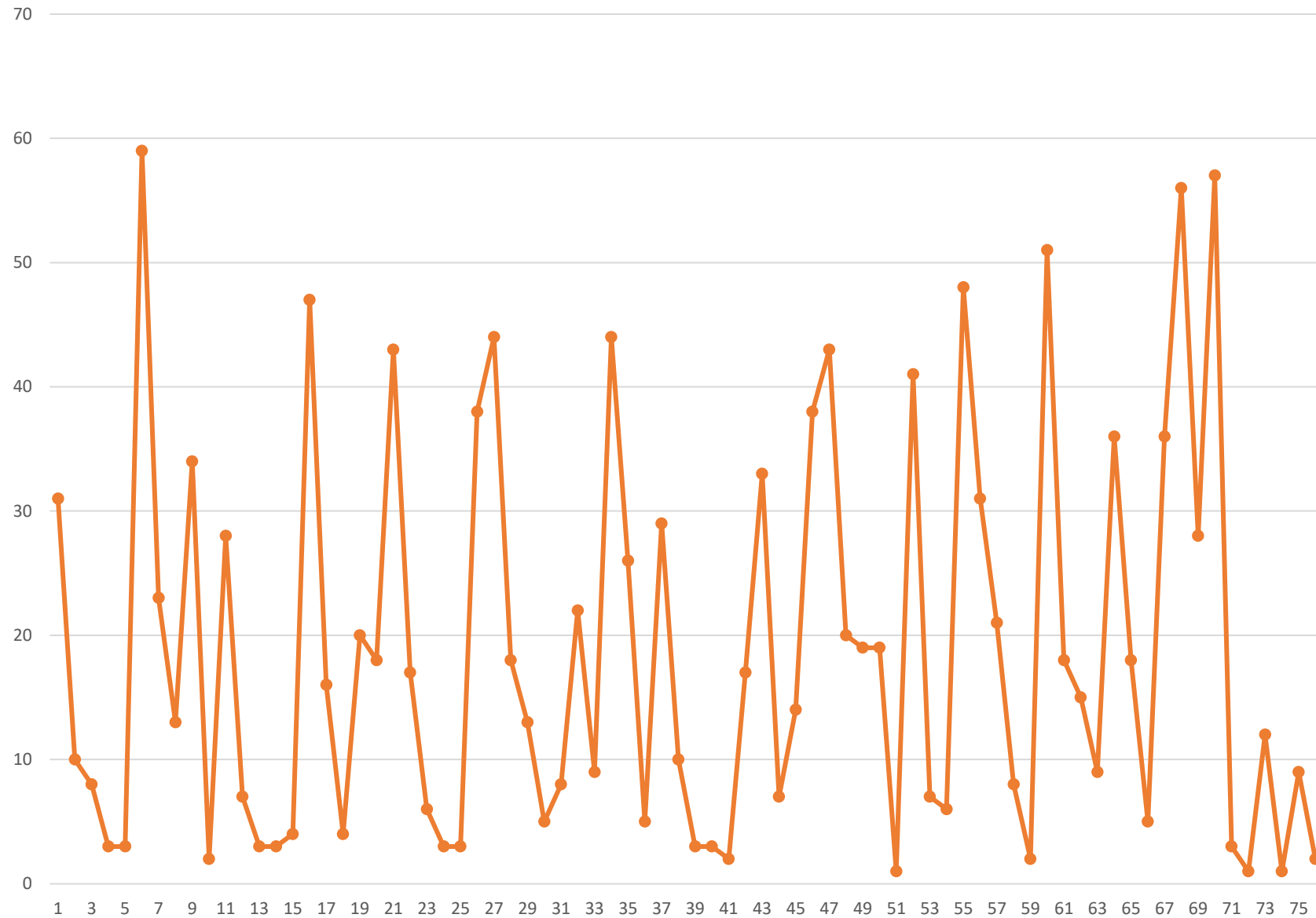
# The idea is simple:

Any process that exhibits assignable cause is said to be changing (or have changed) and is therefore **unpredictable**\*. In these cases, the most economically sound course of action is to identify and address the assignable cause(s) first.

Any process improvement made before assignable causes have been dealt with is **premature**.

If all assignable causes have been dealt with, then process improvement can only come from systemic change.

Scrum Team Doing Two Week Sprints

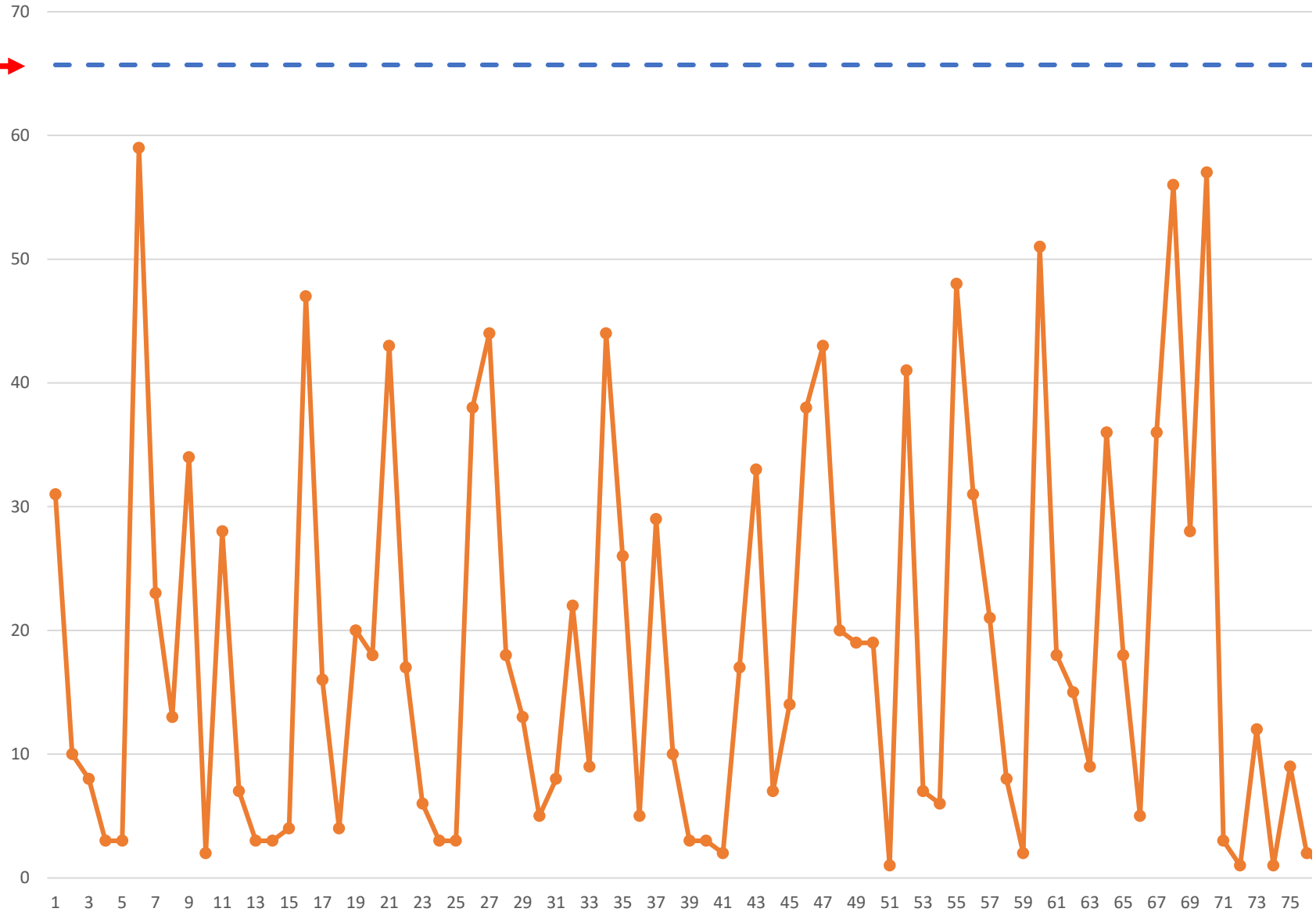


Cycle Time  
for Scrum  
Team doing  
two-week  
Sprints

# Scrum Team Doing Two Week Sprints

65 days

Believe it or not, this process is predictable



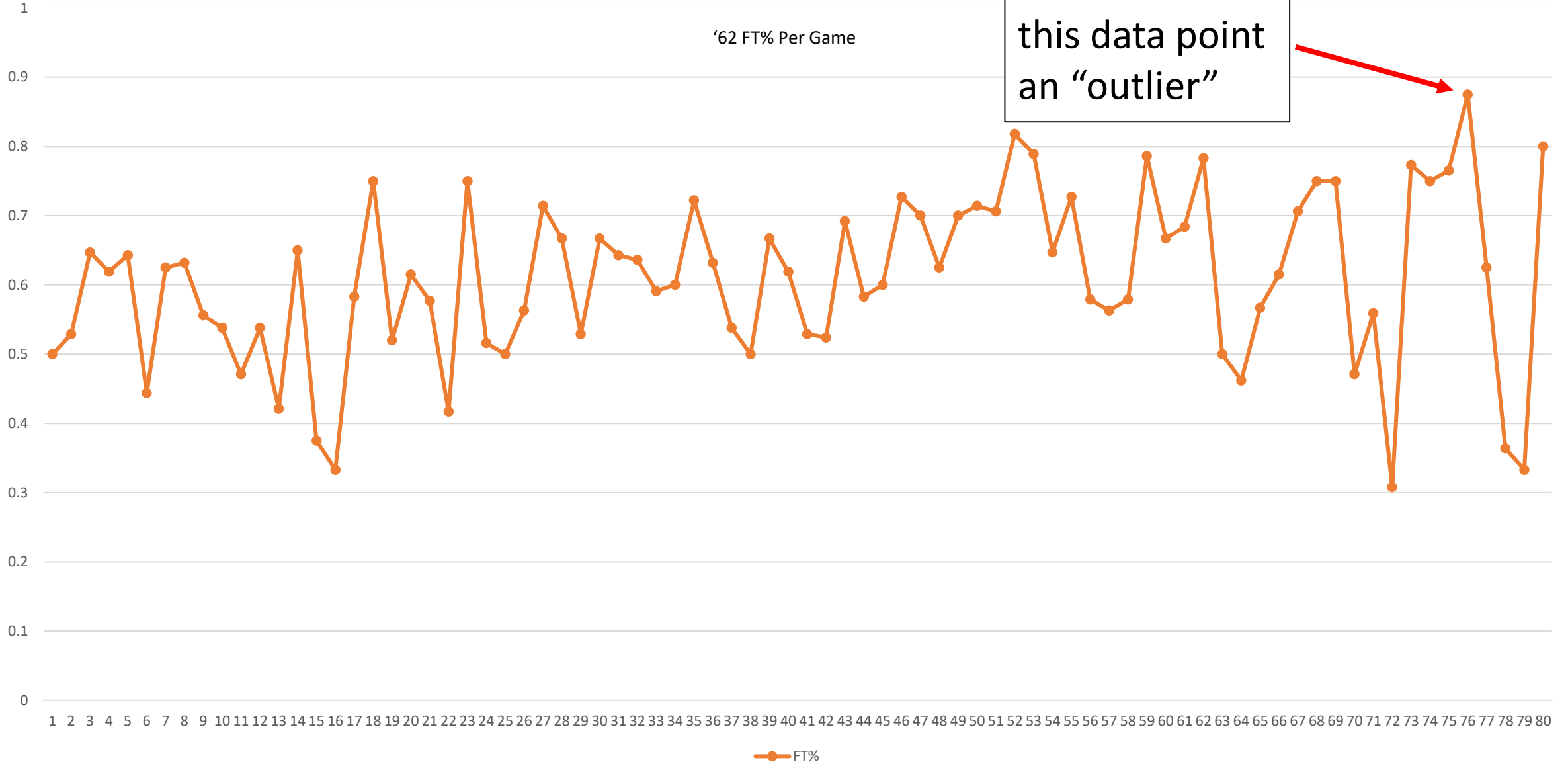
What does this mean in terms of  
data analysis?

What advice have you been given that  
does not help us to  
separate signal from noise?

# Outliers

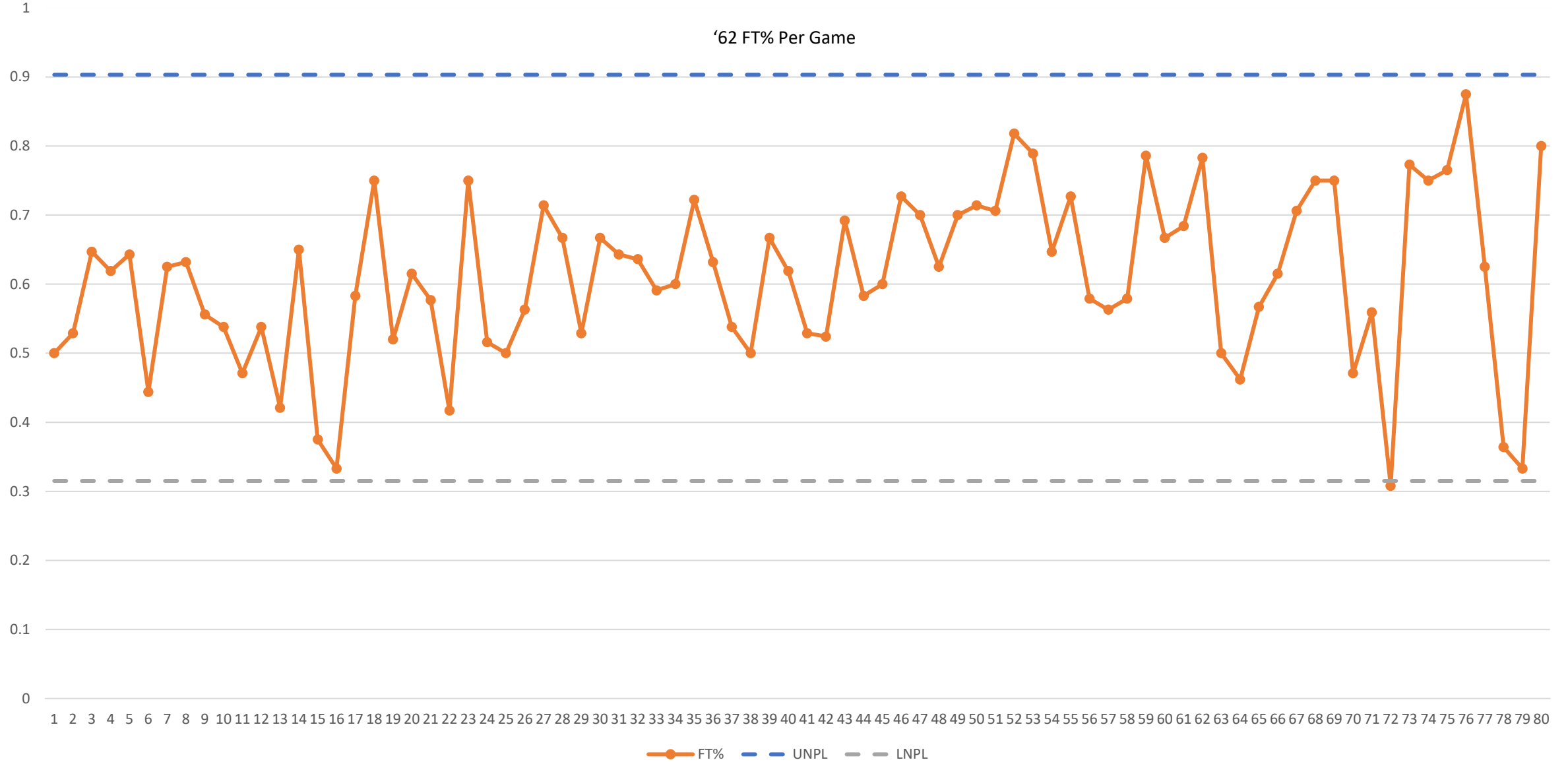
# What is an outlier

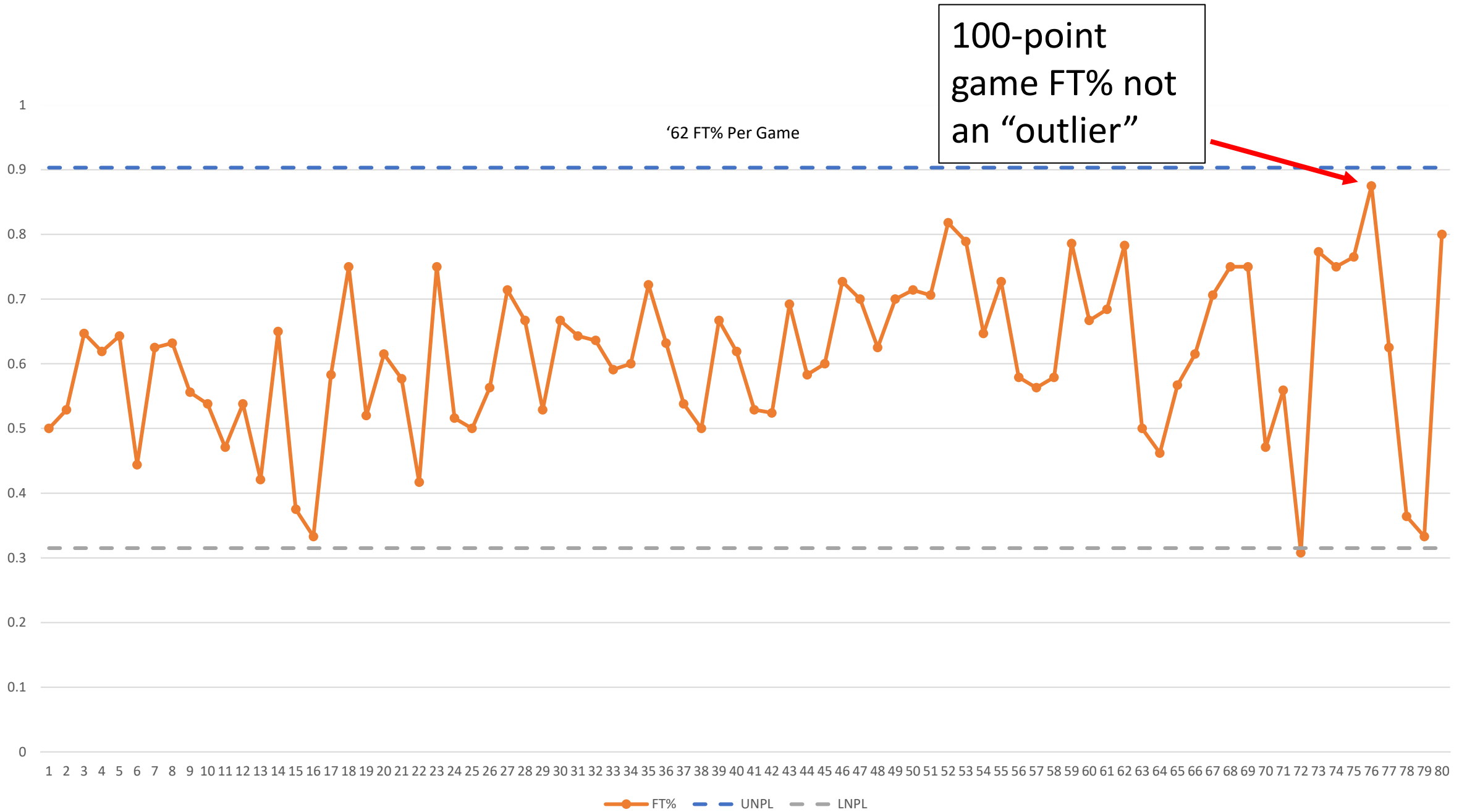
Mr. Gladwell wants to call this data point an "outlier"

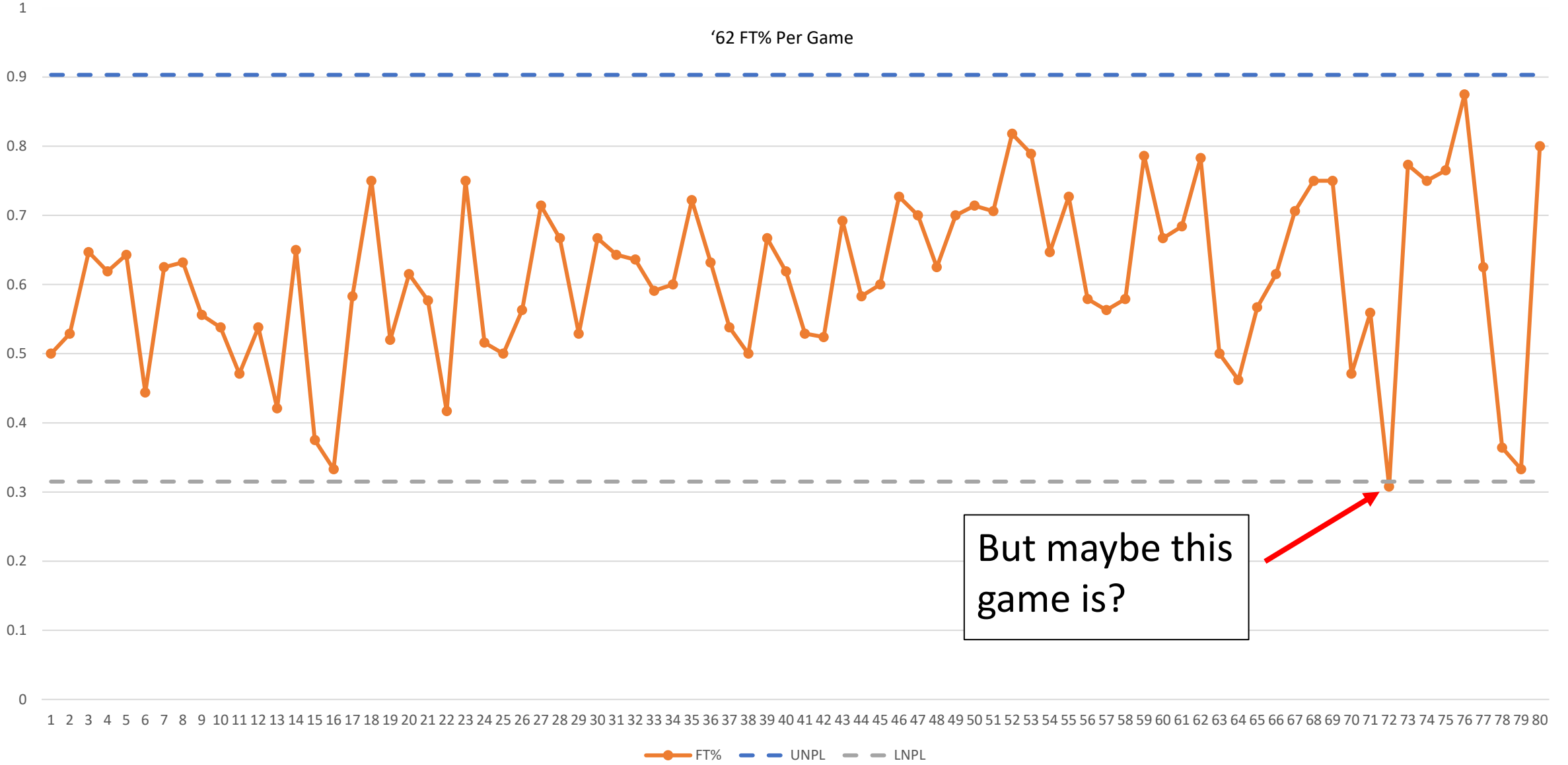




# What is an outlier?



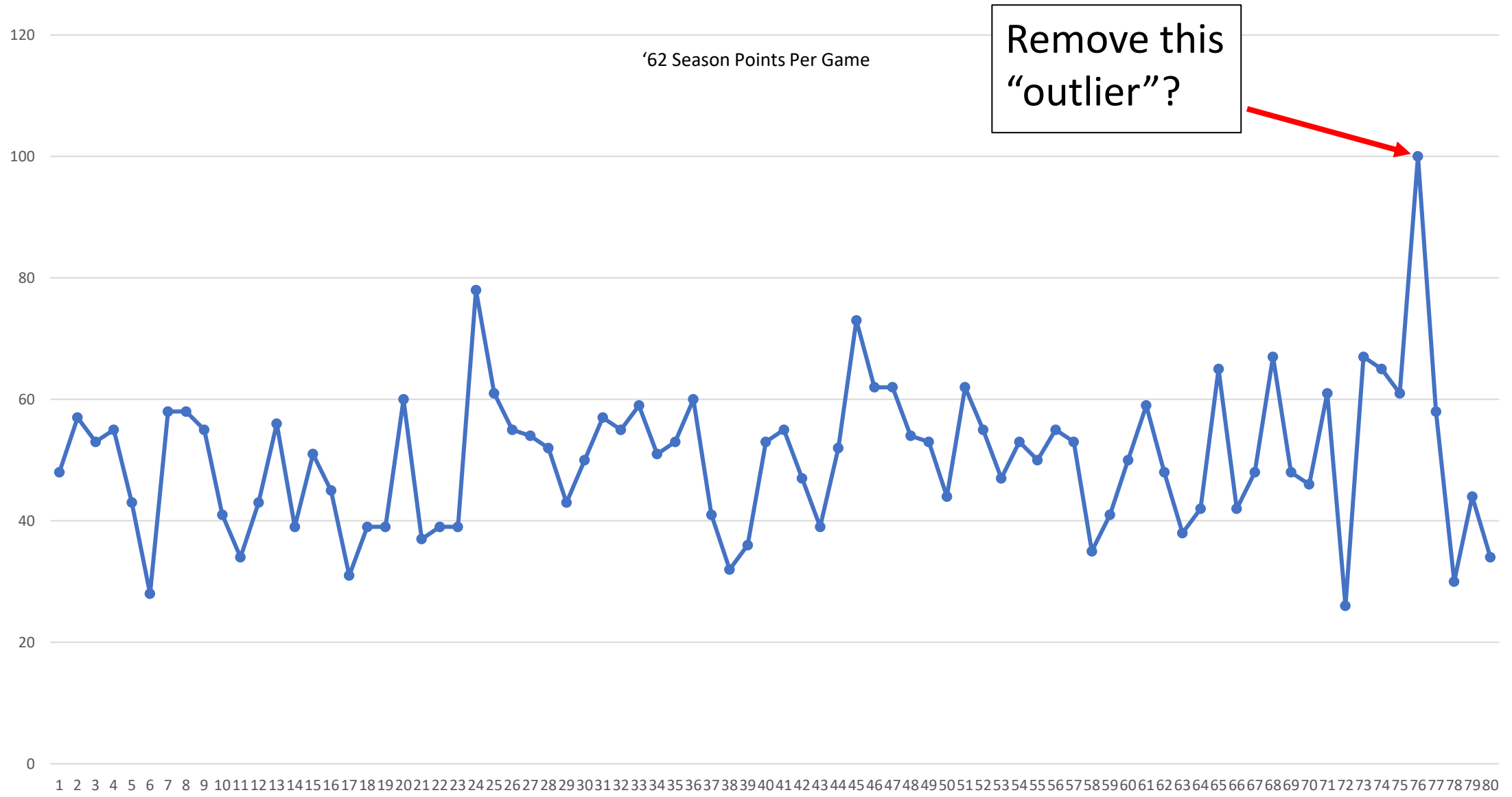




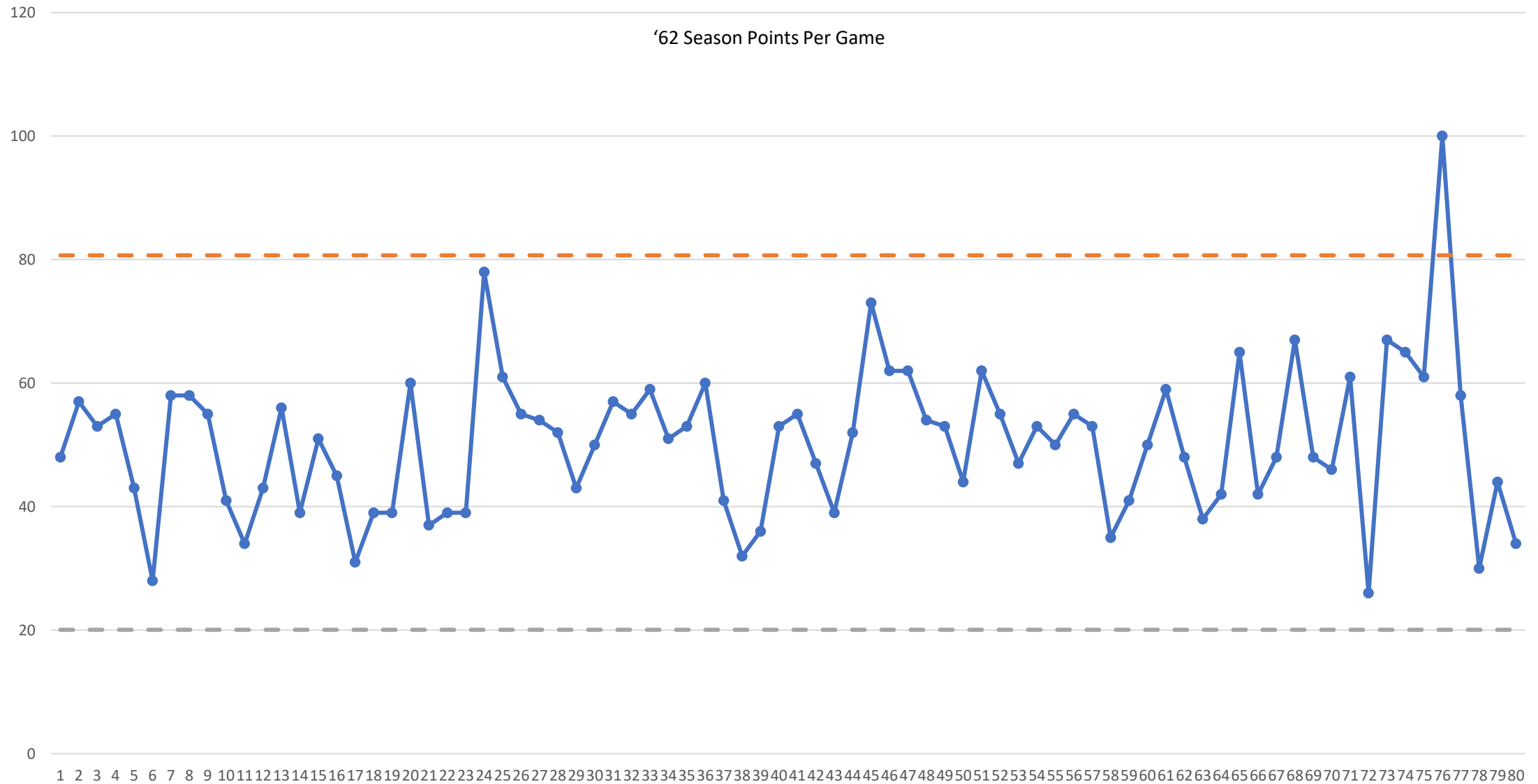
But maybe this game is?



# Remove outliers?

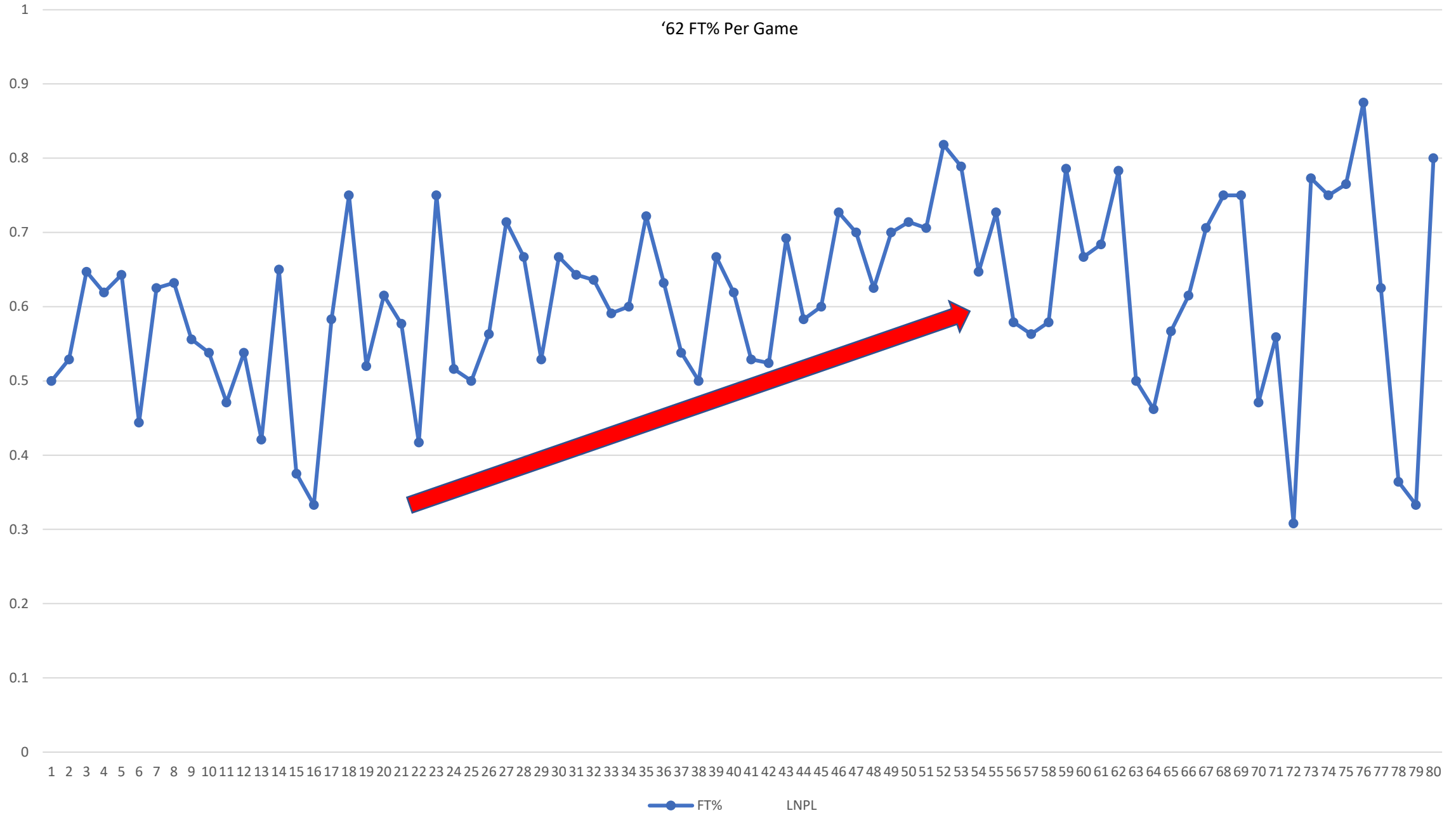


# Remove outliers?



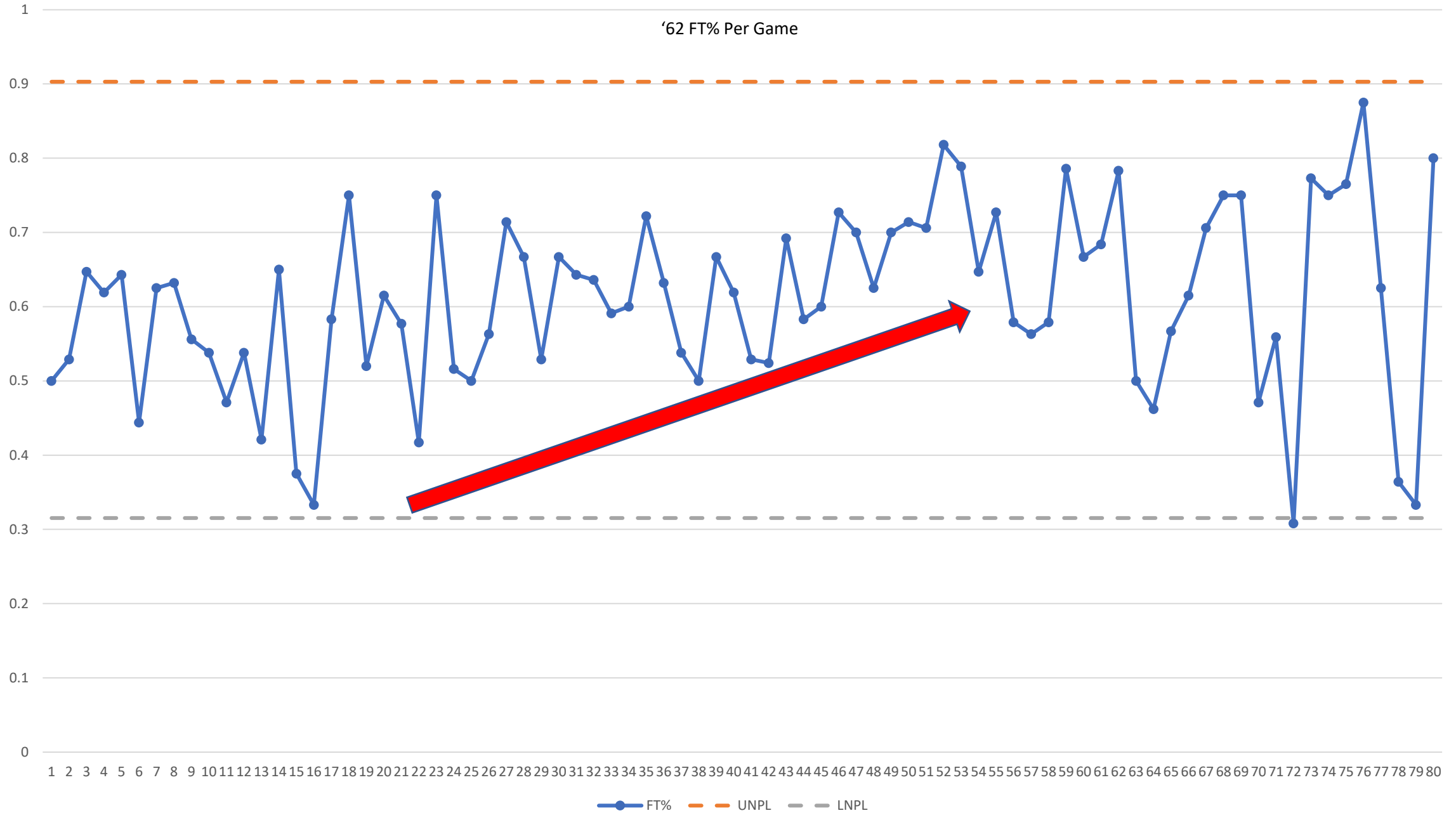
# Trends

'62 FT% Per Game



FT% LNPL

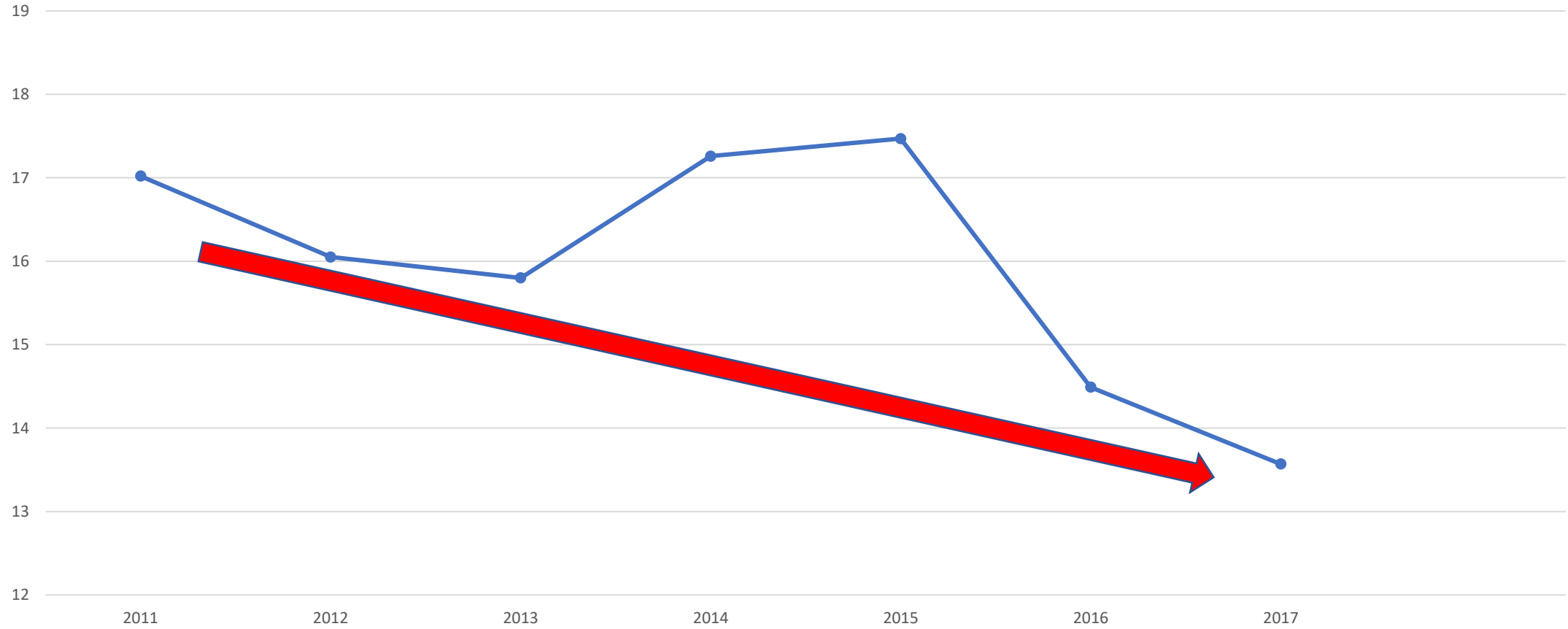
'62 FT% Per Game



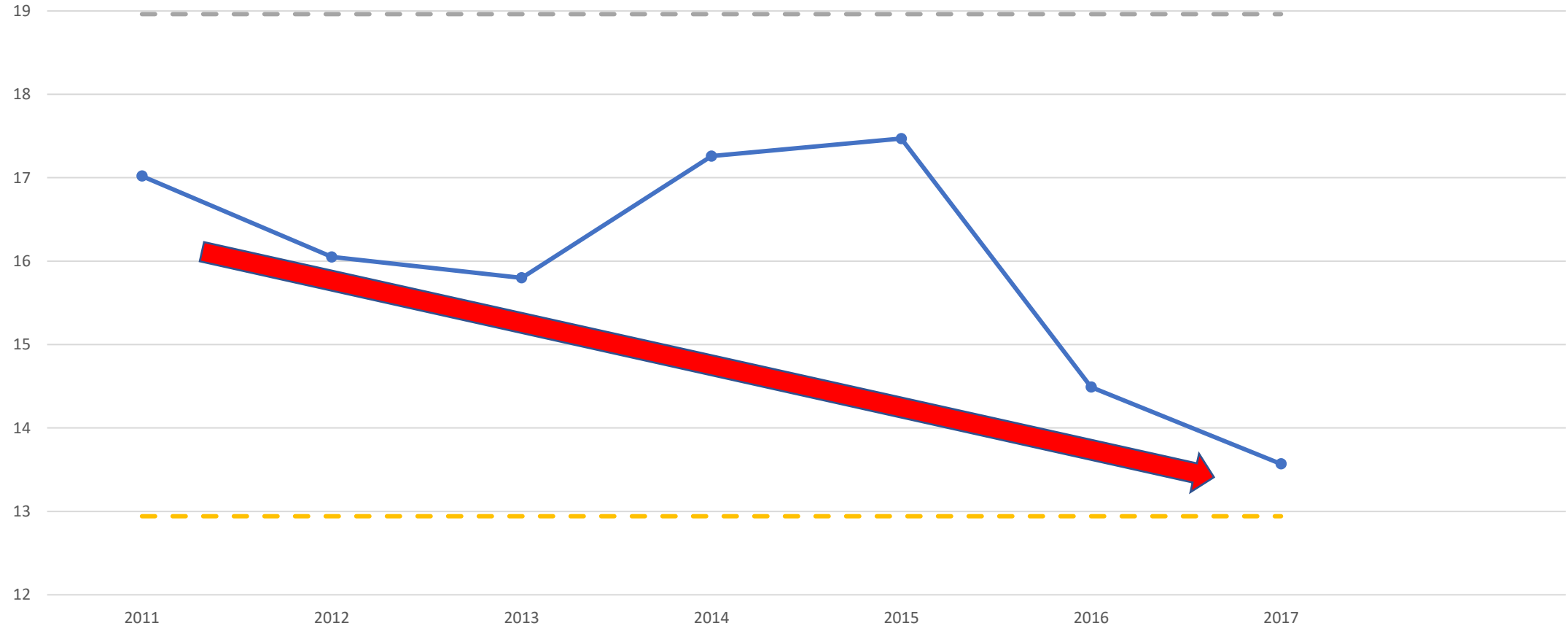
FT% UNPL LNPL



# Teen Tobacco Use Percentage

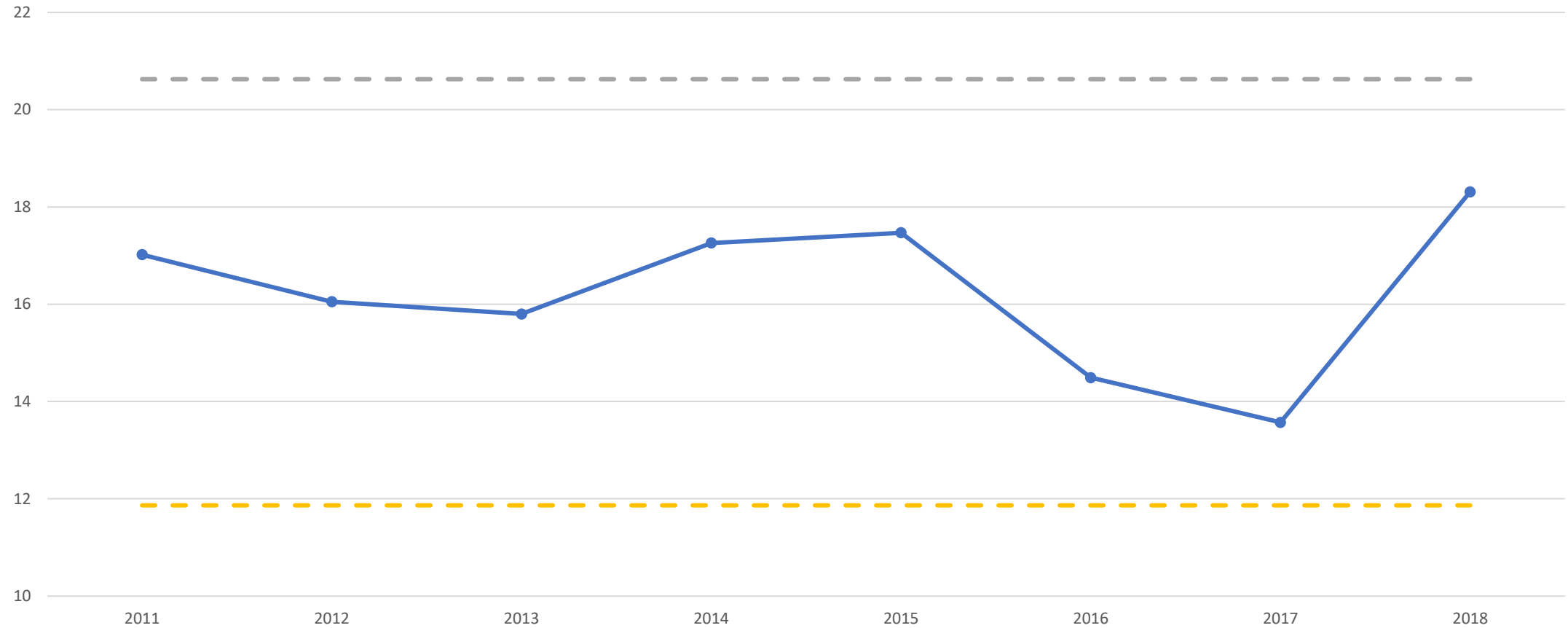


# Teen Tobacco Use Percentage



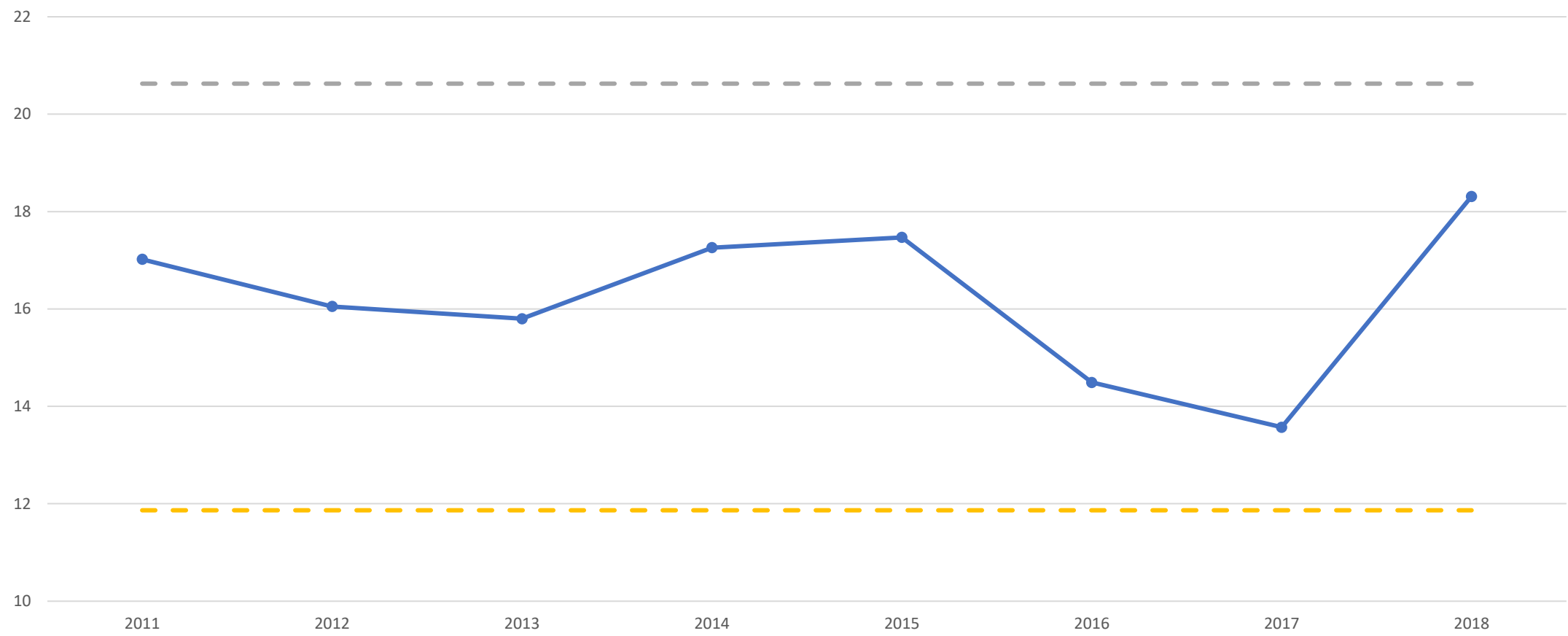
# Adding in data for 2018...

Teen Tobacco Use Percentage

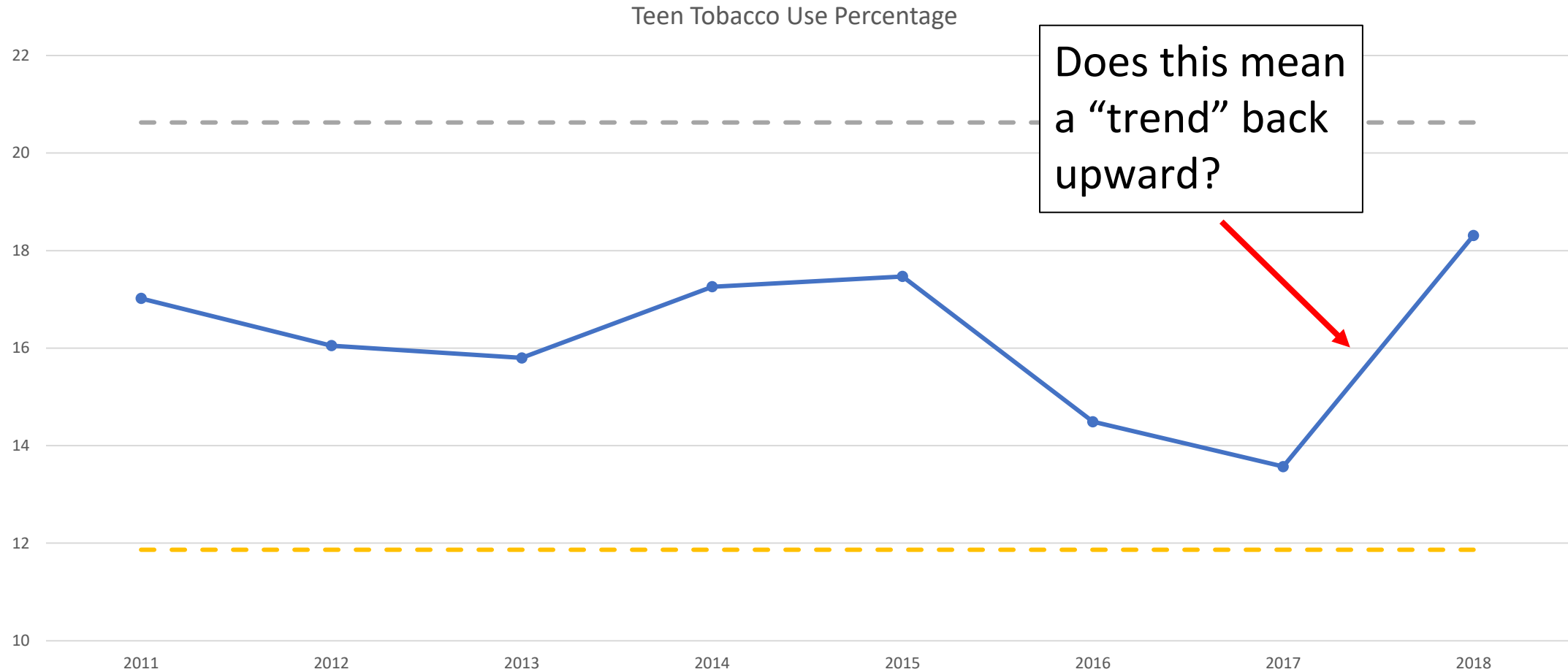


# “Teen use turns upward” -- USA Today

Teen Tobacco Use Percentage



# “Teen use trends upward” -- USA Today



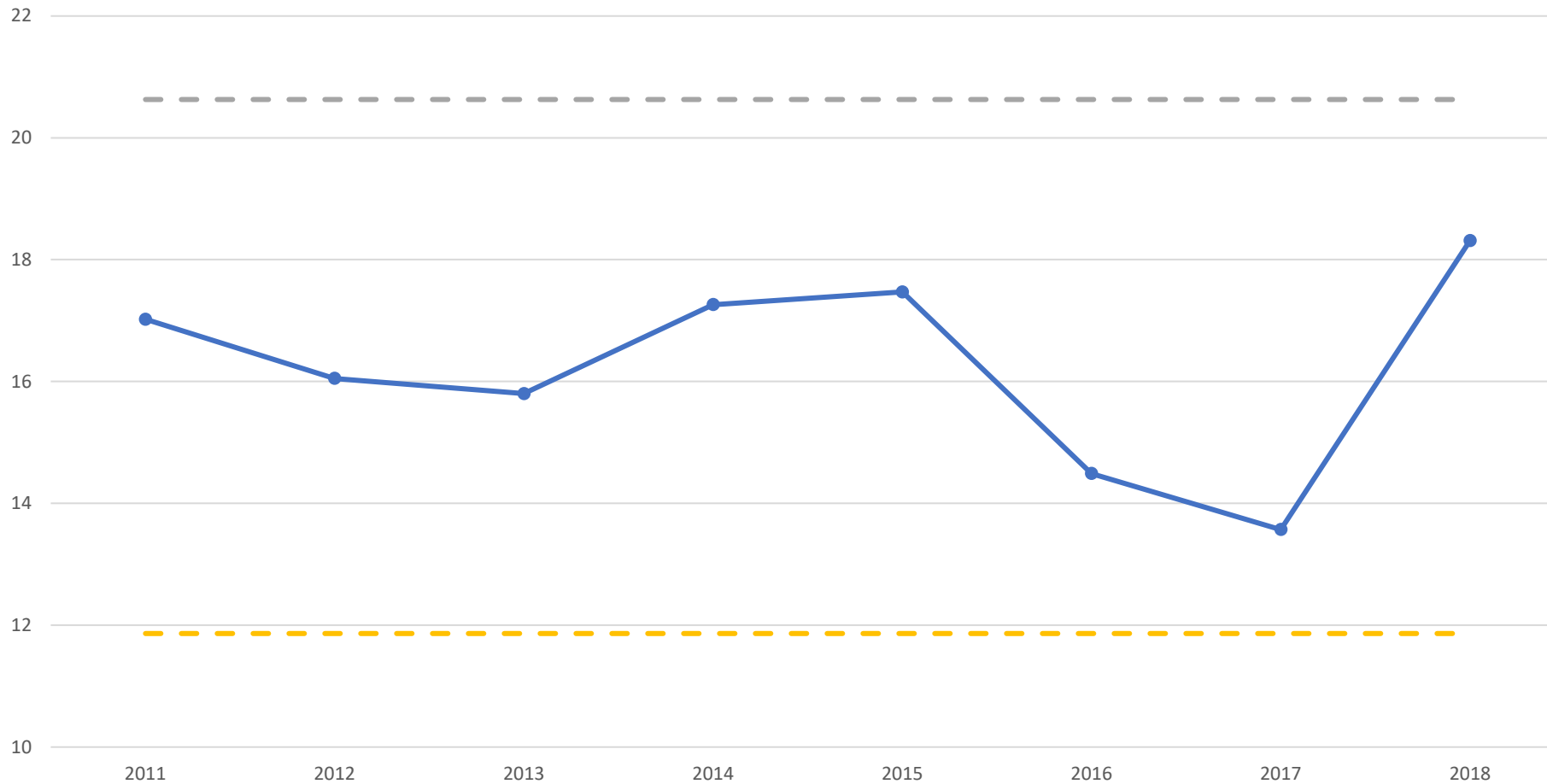
There are no trends in noise!!

(Just for fun)

Comparing a single point  
to an average

# “Teen use trends upward” -- USA Today

Teen Tobacco Use Percentage



2018 Value:

**18.3%**

Average 2011-2017:

**16.2%**

# **“Experts Predict Above Average 2022 Hurricane Season”**

<https://www.forbes.com/advisor/homeowners-insurance/2022-hurricane-season/>

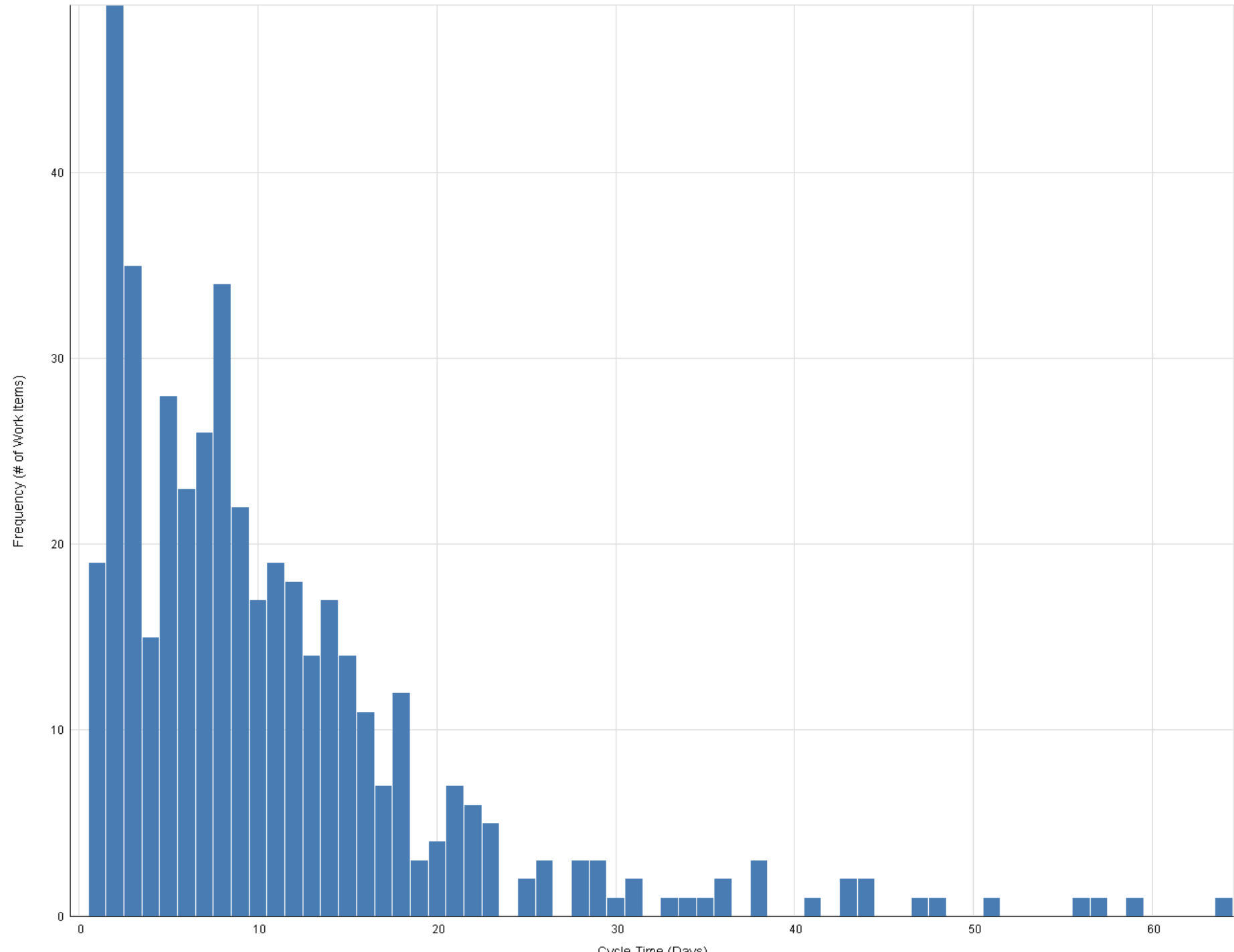


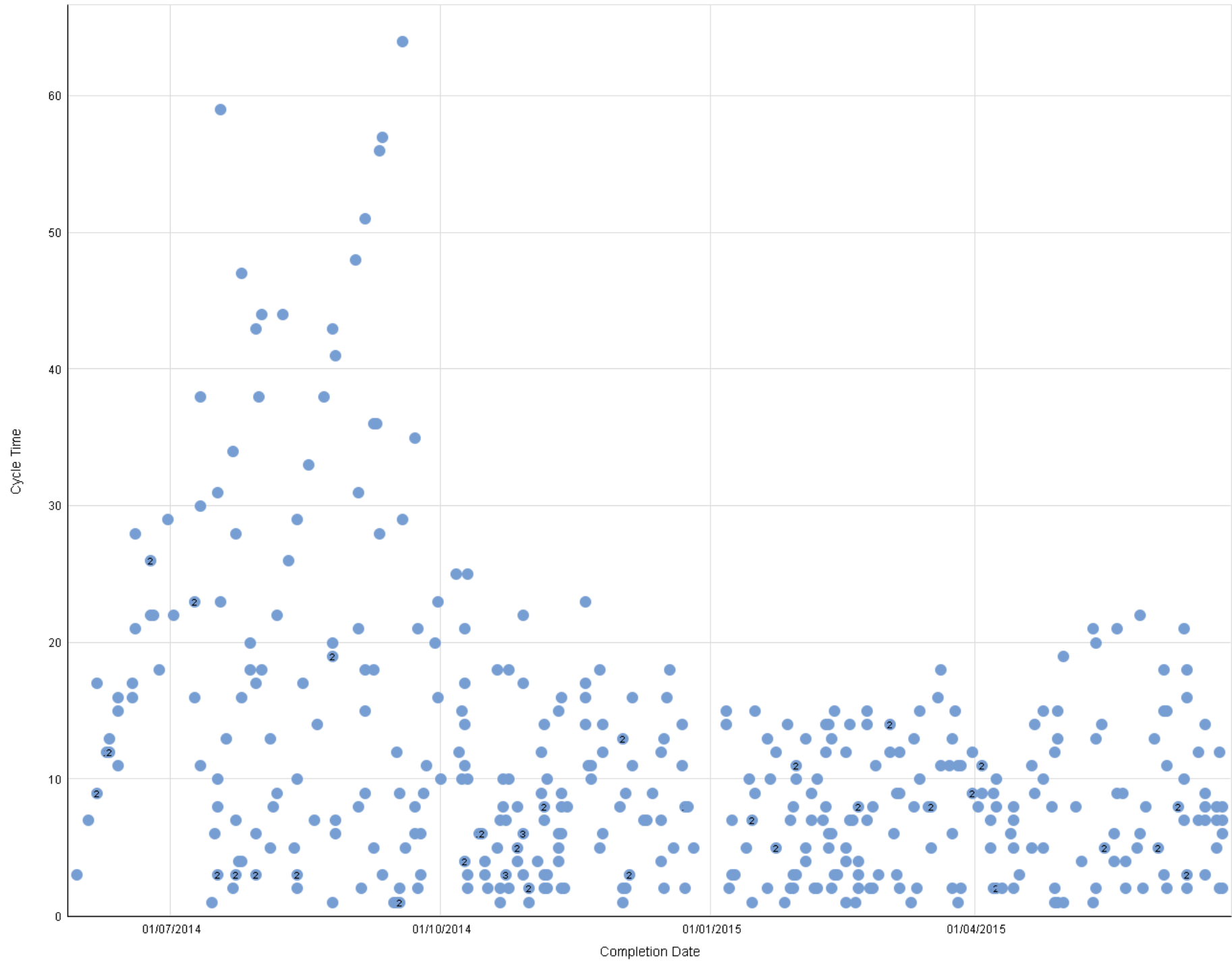
Where have we seen  
this false comparison  
before?

Career Free Throw Rate -- 51%

100-point game -- 87.5%

# Histograms





For extra credit:  
**NEVER** fit a probability  
distribution to your data

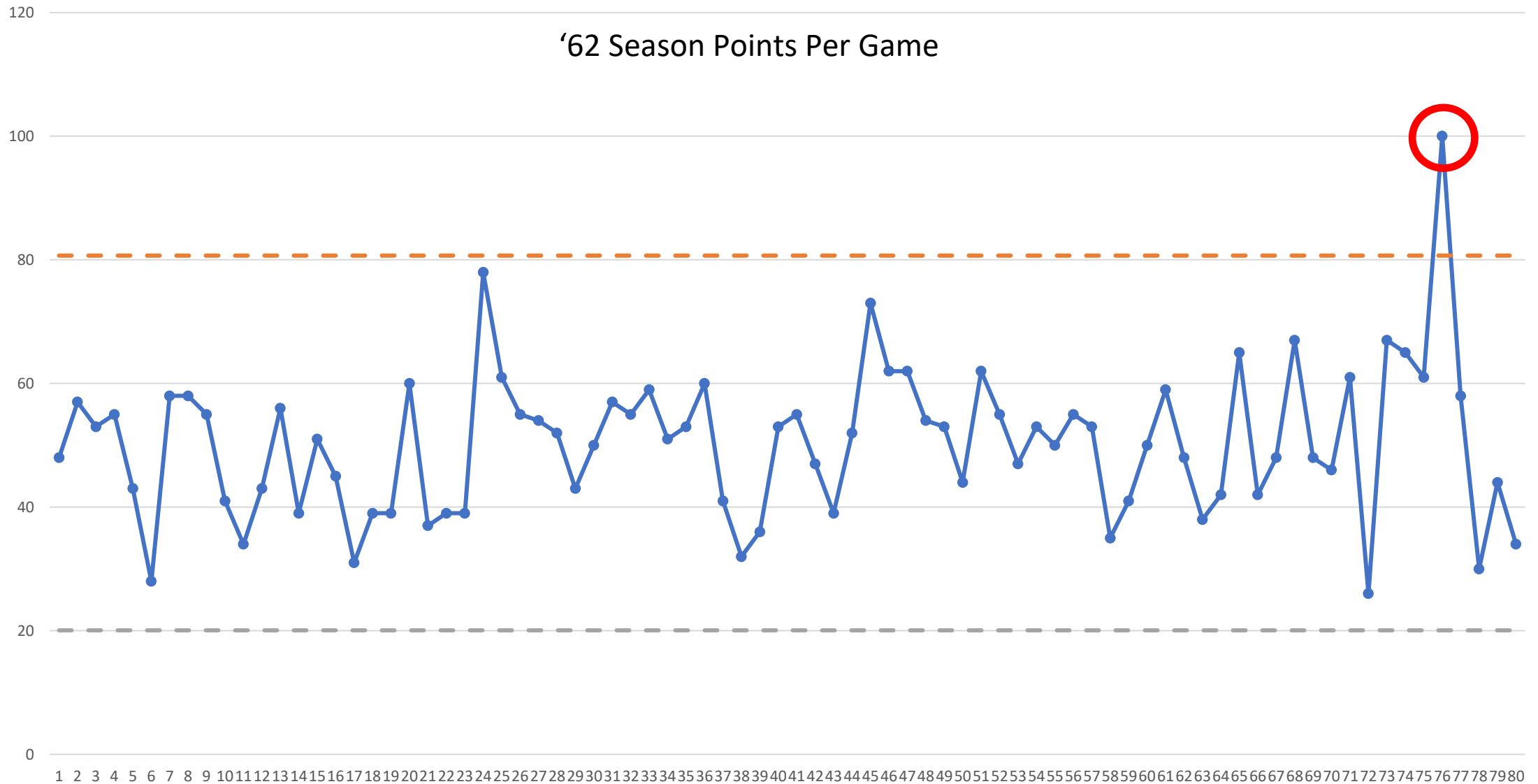
Lastly, don't forget your get out  
of jail free card:

“Data have no meaning apart  
from their context.”

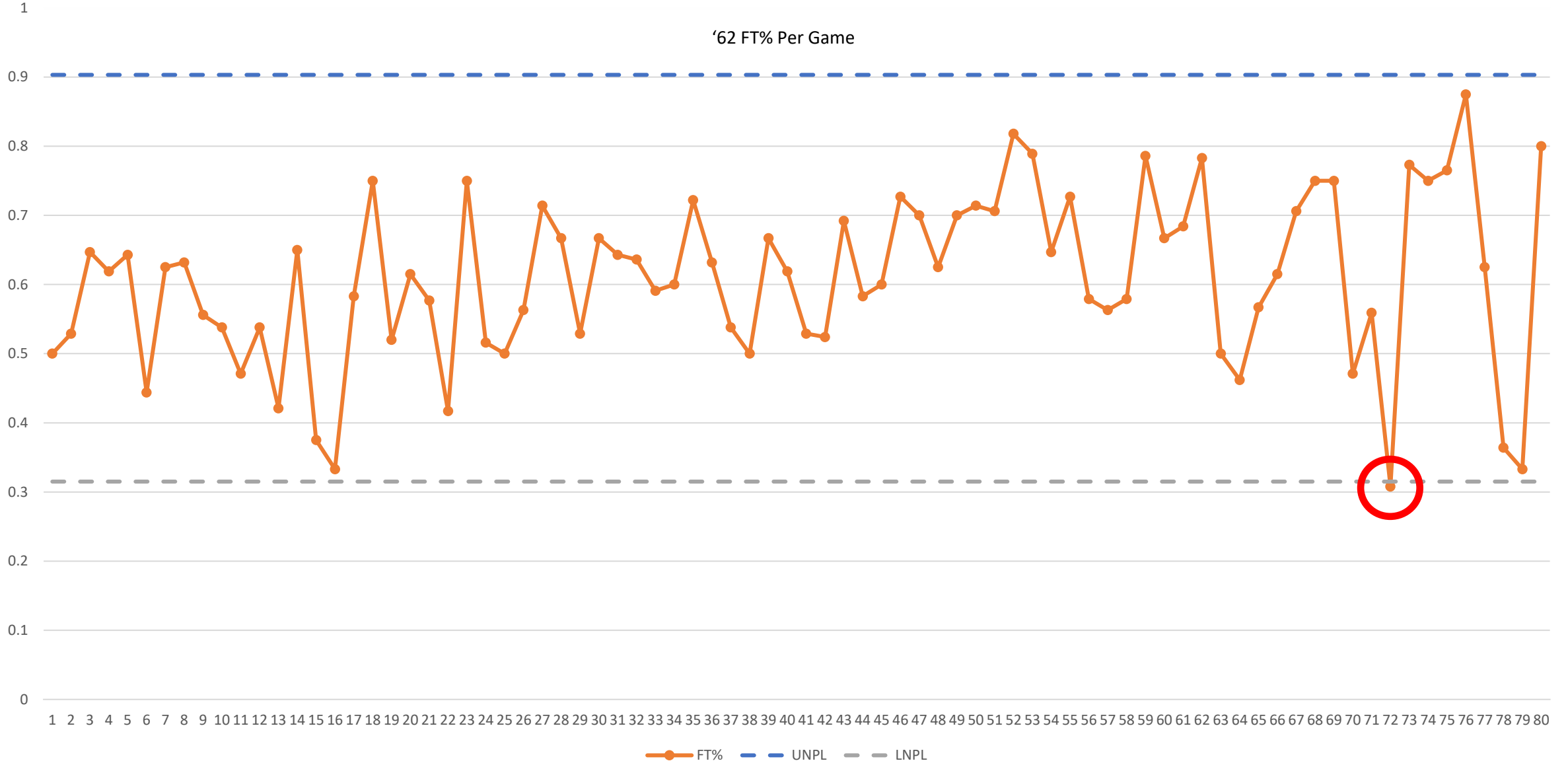
# Back to Wilt...



# There is possible signal here:



# But that it is due to underhand free throws is dubious



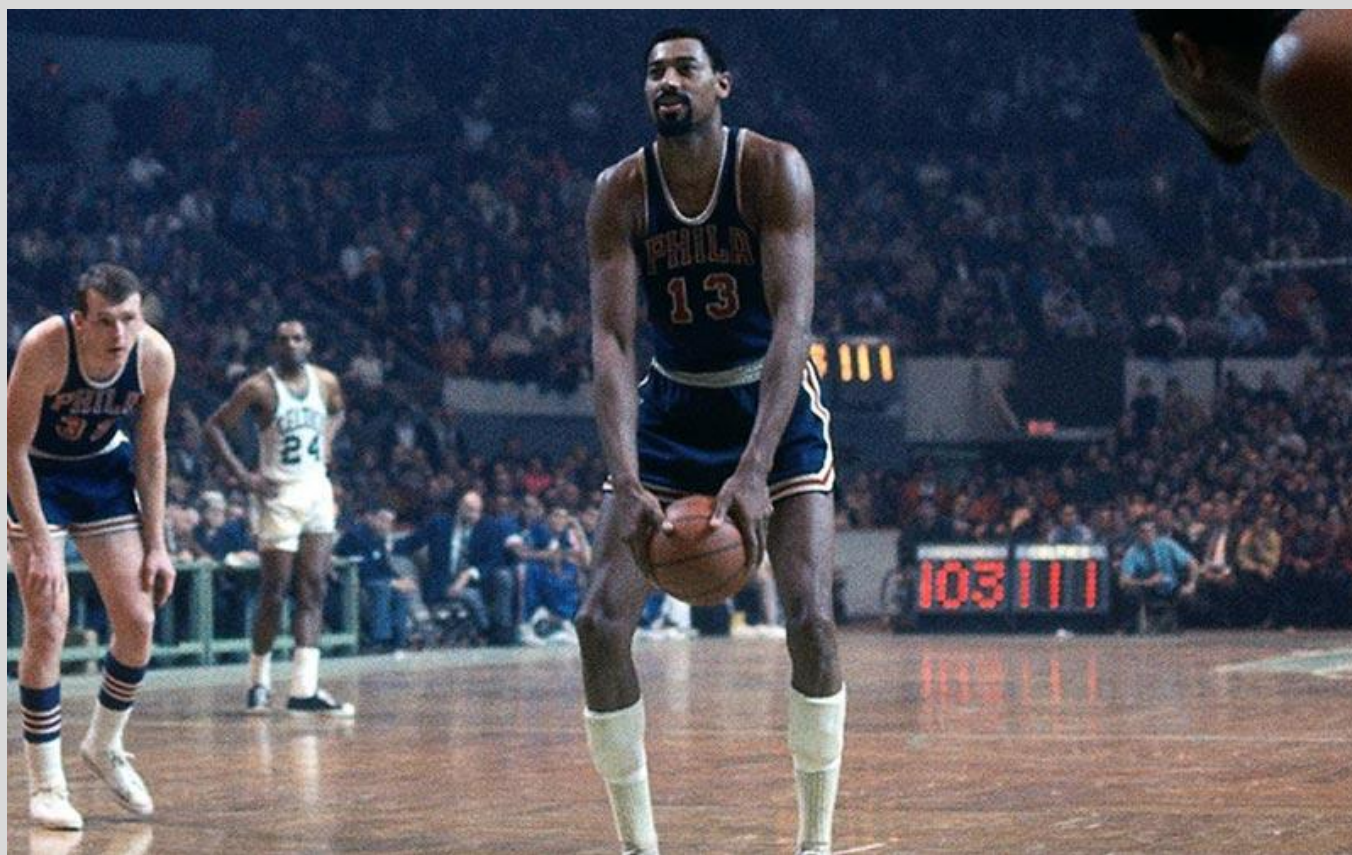
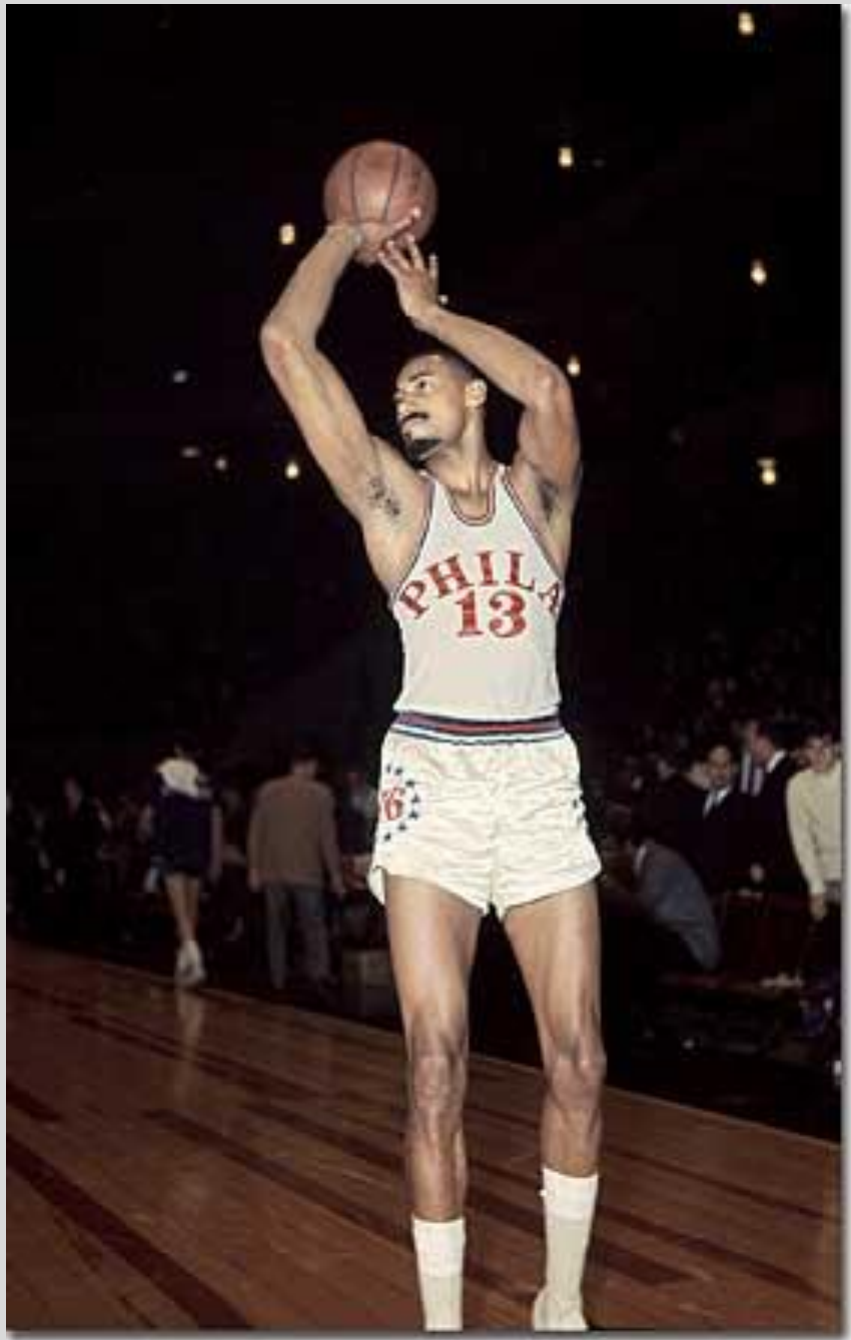
Therefore, it is reasonable to  
conclude that something else  
happened that night:

[https://en.wikipedia.org/wiki/Wilt\\_Chamberlain%27s\\_100-point\\_game](https://en.wikipedia.org/wiki/Wilt_Chamberlain%27s_100-point_game)

“...Wilt Chamberlain switches to a better [free throw] shooting technique. It pays off in the greatest basketball game ever played.”

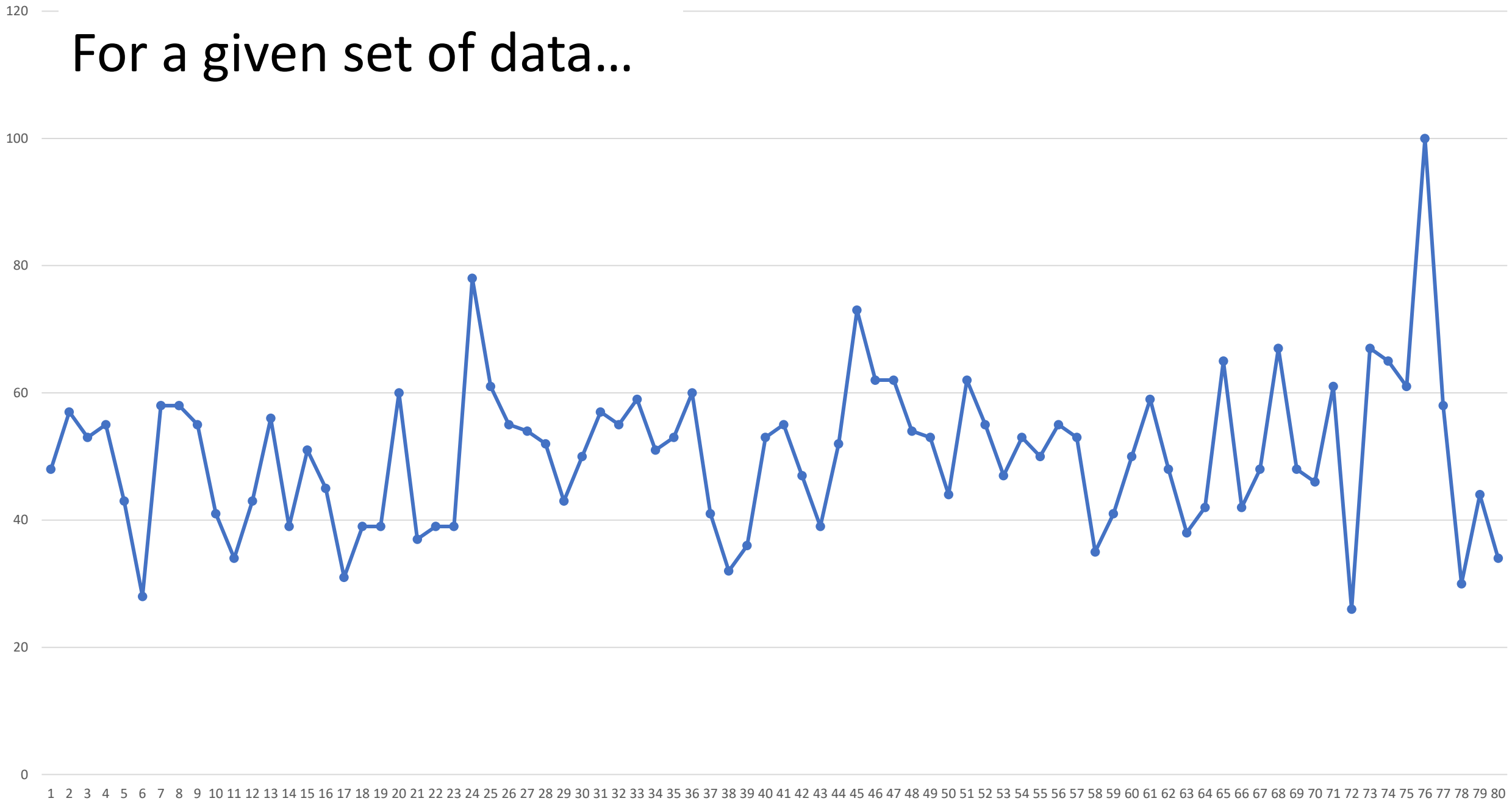
Mr. Gladwell is mistaking **noise** for **signal**.

# So What?



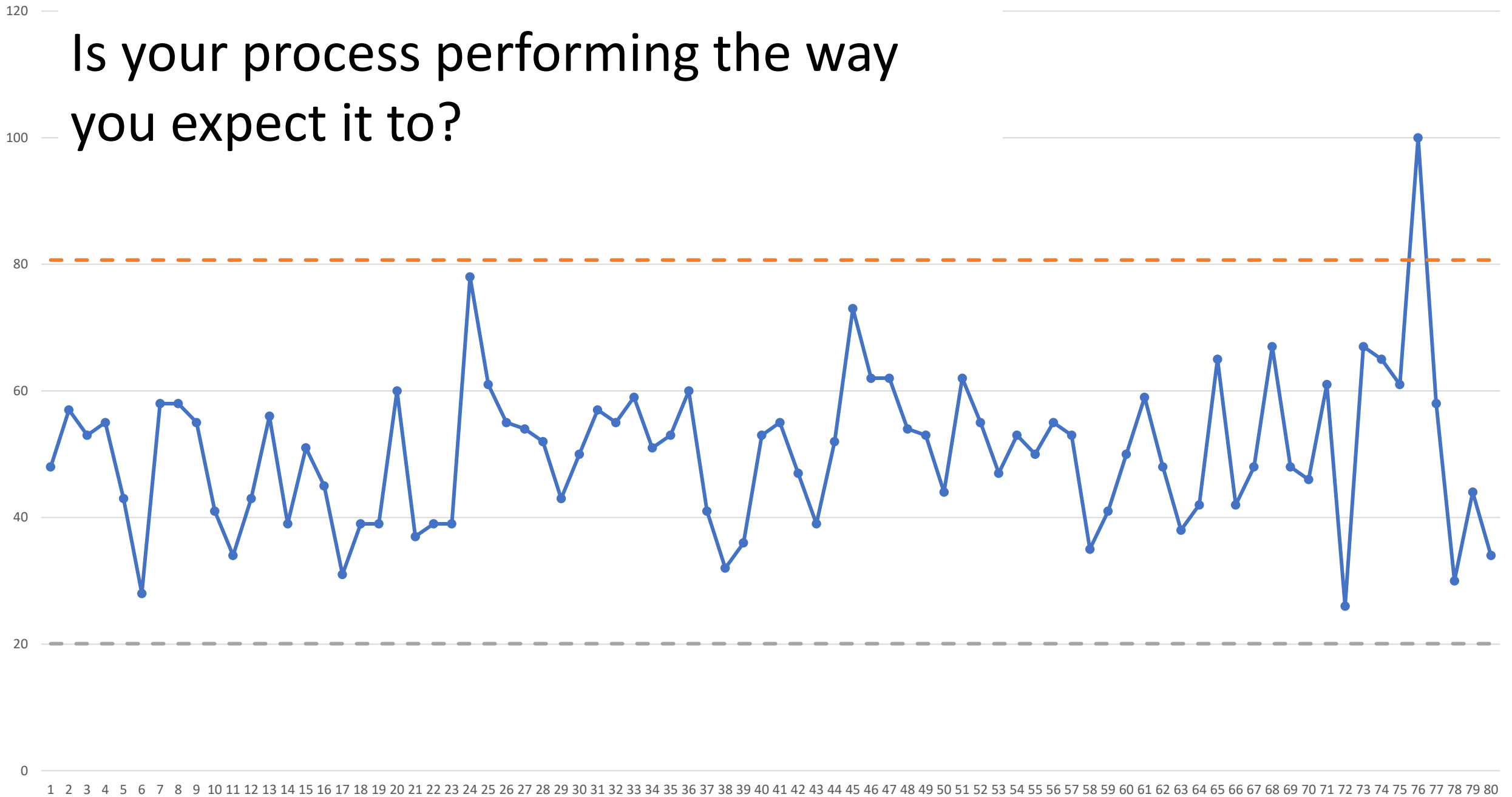
Our process  
world is  
dominated by  
randomness

# For a given set of data...

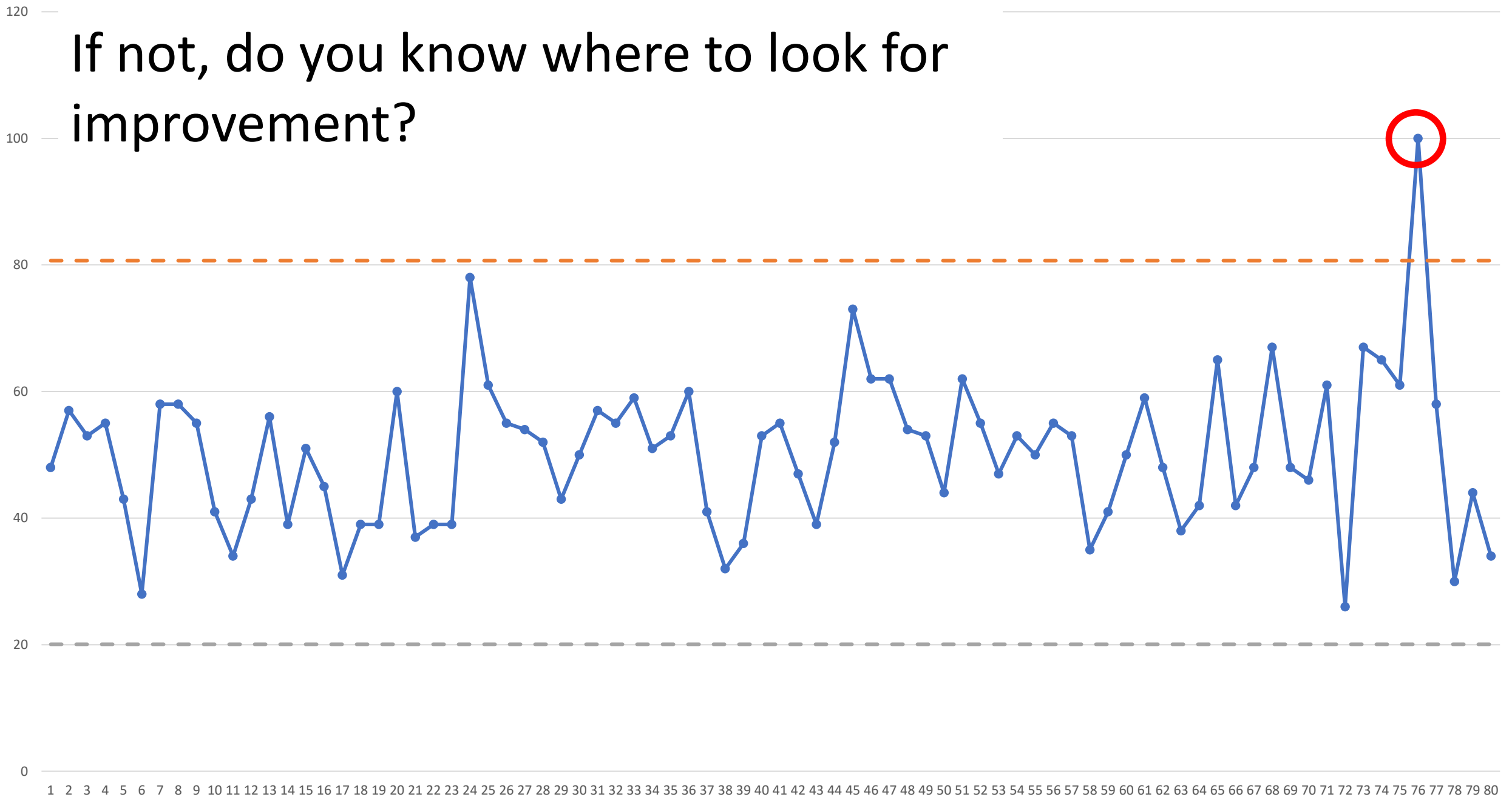




# Is your process performing the way you expect it to?



# If not, do you know where to look for improvement?



# Once you make a change, how do you know if it worked or not?

22

20

18

16

14

12

10

2011

2012

2013

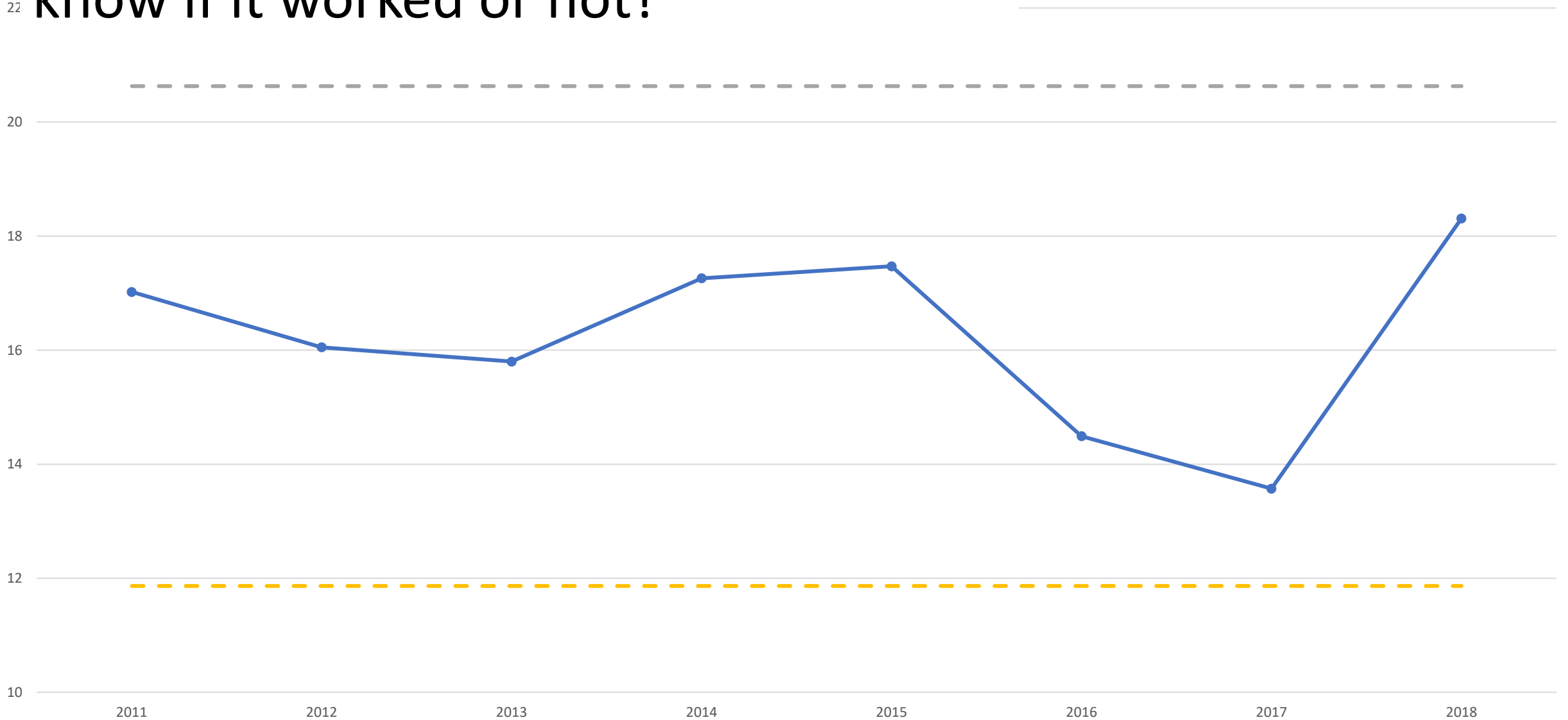
2014

2015

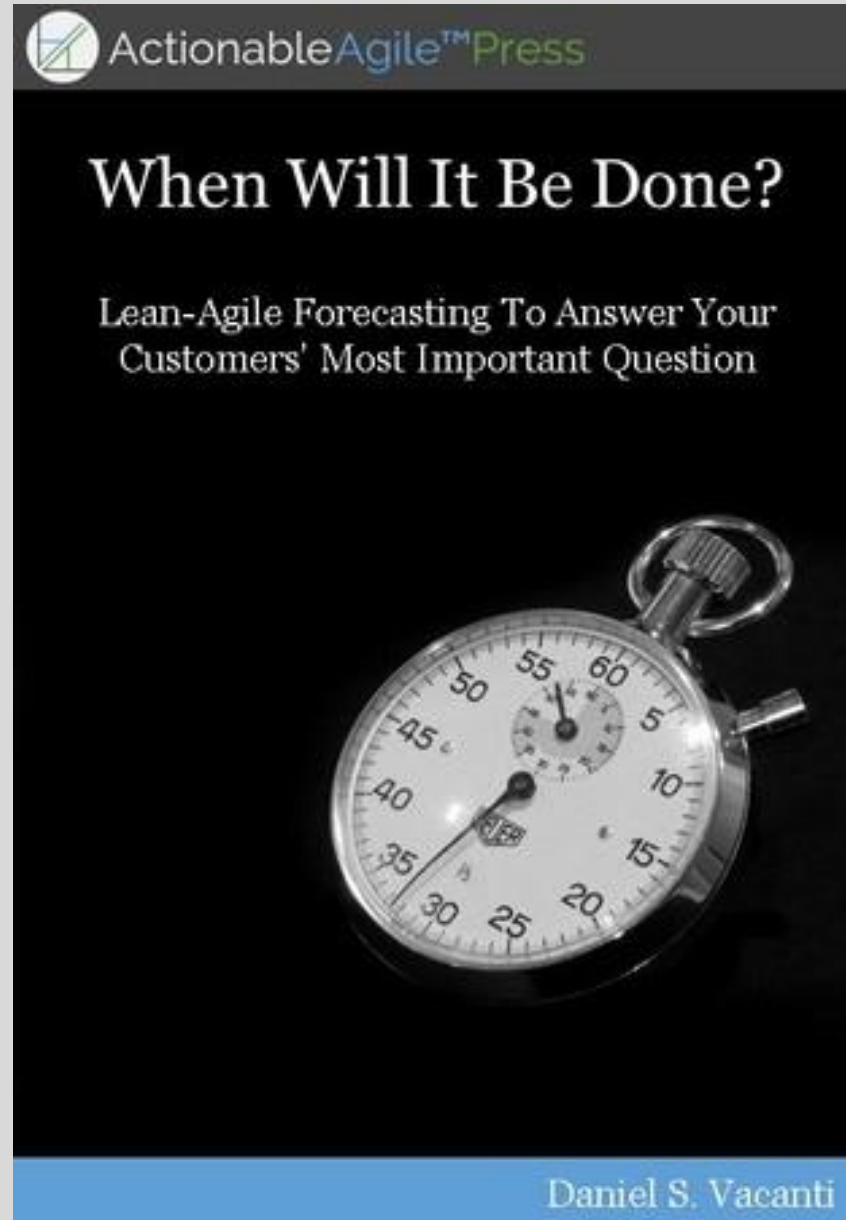
2016

2017

2018



Maybe most importantly, have you designed a process that can answer...



# Another special thanks to...



For next time:

Yes, it really is output over  
outcomes

P.S.

Don't Use Story Points!

# To Sum Up



Think about your data in  
terms of  
signal and noise

If you detect signal, forget trying to make predictions and/or making process improvements and focus on eliminating assignable causes.

If you do not detect any possible signal, you are now in a place where you can make predictions and/or focus on other process improvements.

For more information...

<https://spcpress.com>

<https://actionableagile.com>

<https://prokanban.org>

Drunk Agile YouTube channel

# QUESTIONS?

# THANK-YOU!

**Daniel S. Vacanti**

**<https://prokanban.org>**

**@danvacanti**



**ProKanban.org**