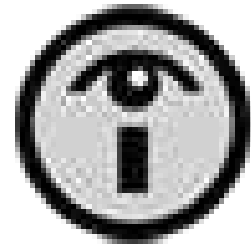


# Looking for Failure to Generate Success

Big Data and the Missing Pieces

Paul Edelblut

Vantage Labs



# VANTAGE



## MCCANN ASSOCIATES

An Innovative Leader in  
Assessment, Certification,  
Learning Management, and  
Business Intelligence  
Solutions.  
We Transform Organizations  
and Decision Making.

We Transform Organizations  
and Decision Making.



## VANTAGE LEARNING

Creating sophisticated  
knowledge solutions for K-12  
using award-winning AI and  
linguistic tools, providing  
immediate feedback to  
enhance teaching and  
student learning

student learning



## iseek.ai

Cognitive computing platform  
using natural language  
understanding and machine  
learning to solve Big Data



## VANTAGE LINGUISTICS

The leading provider of  
software solutions based on  
linguistic technology that  
improves business  
productivity, efficiency and  
organizational performance

organizational performance

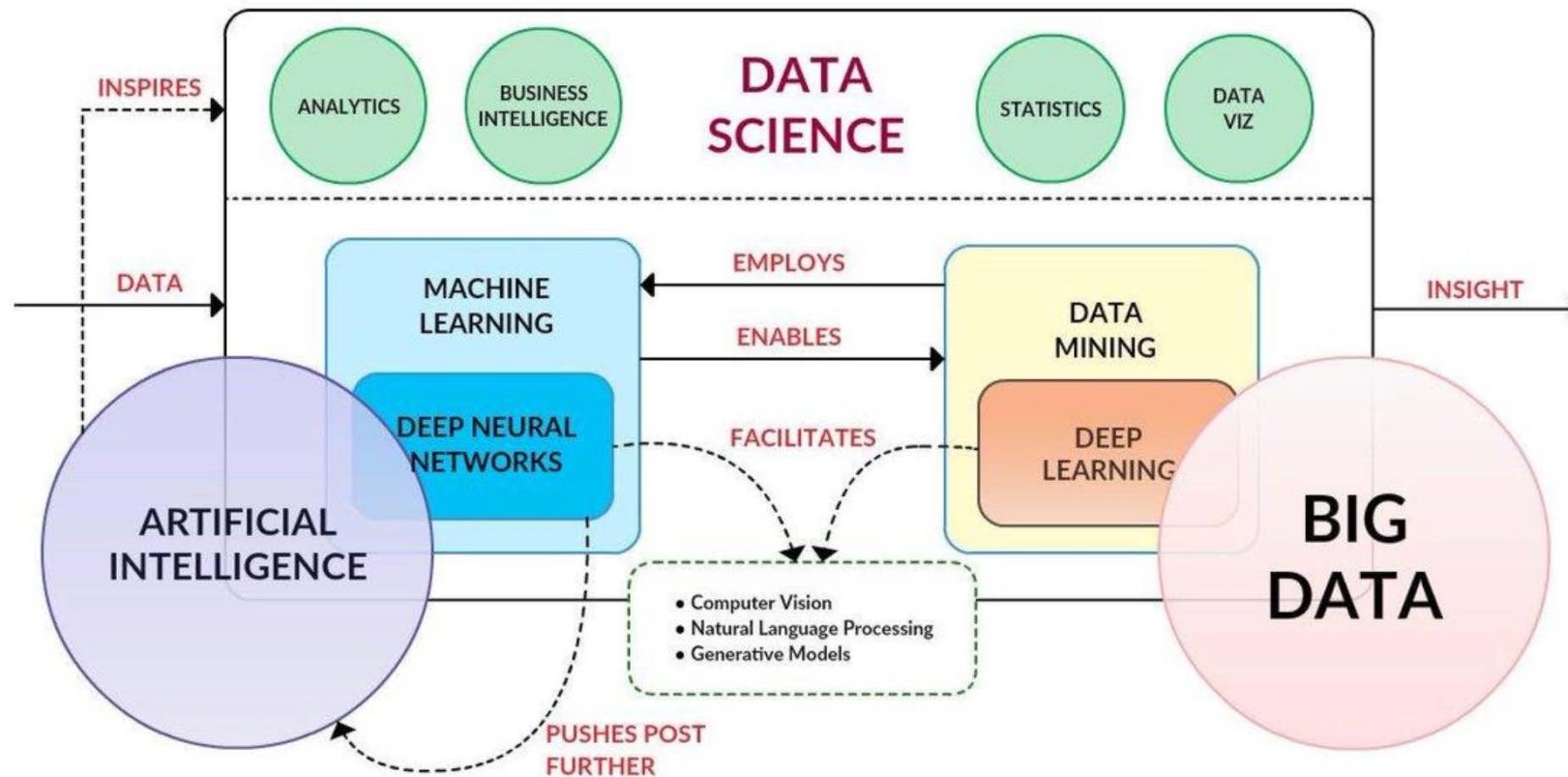


## VANTAGE COMMUNICATIONS

International leader in  
converged IP telephony  
communications solutions  
that deliver powerful and cost  
effective productivity gains to  
the desktop

the desktop

# The Landscape

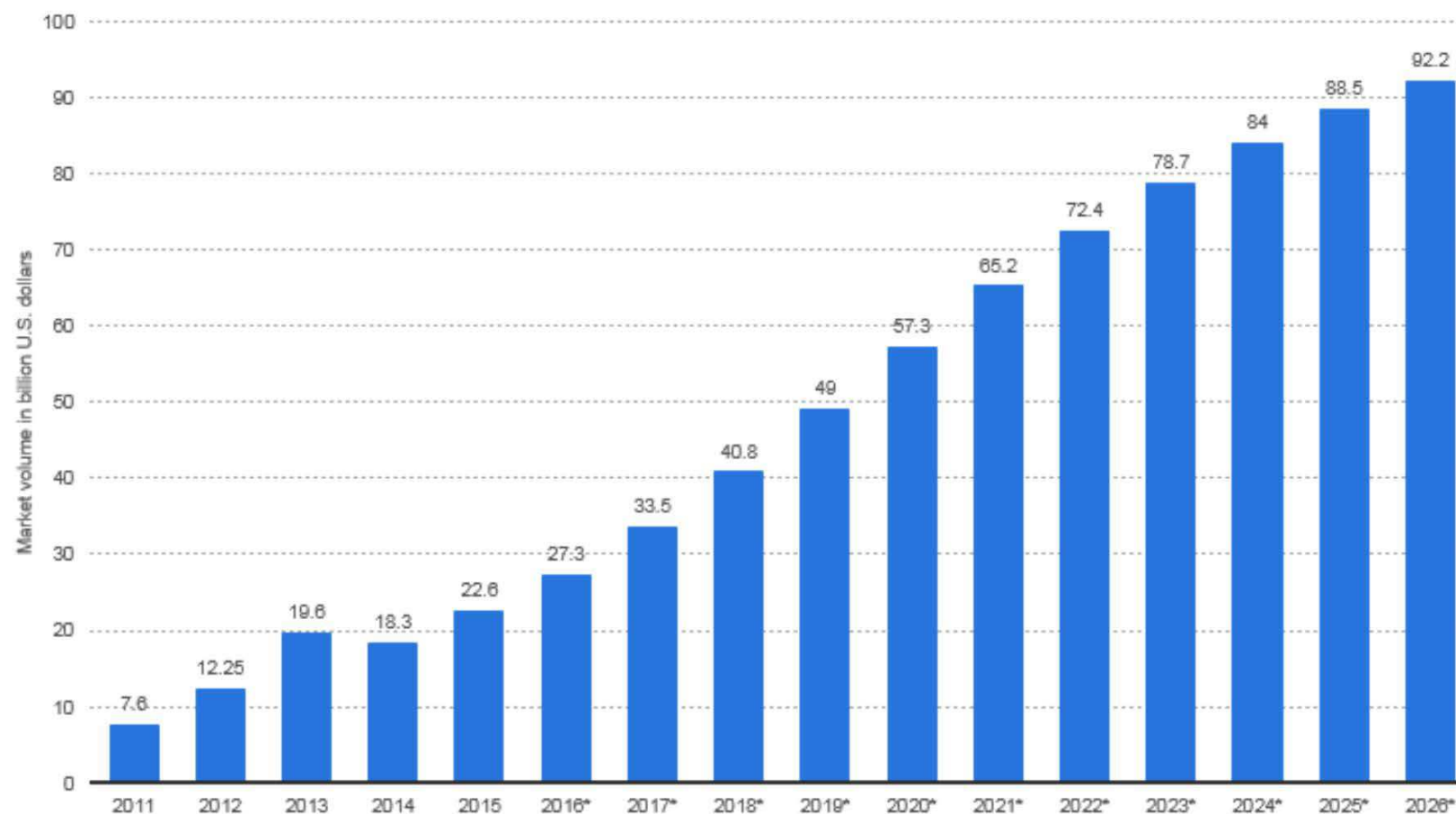


# We are teens all over again

- 2016—Big Data Failure Rates ~60%--Gartner
- 2017—Oops, actually, Big Data Failure Rates ~85%--Gartner
- Everyone is still learning

***Big Data is like teenage sex;  
everyone talks about it,  
nobody really knows how to do it,  
everyone thinks everyone else is doing it,  
So everyone claims they are doing it.  
--Dan Ariely, Duke University***

## Big data market size revenue forecast worldwide from 2011 to 2026 (in billion U.S. dollars)



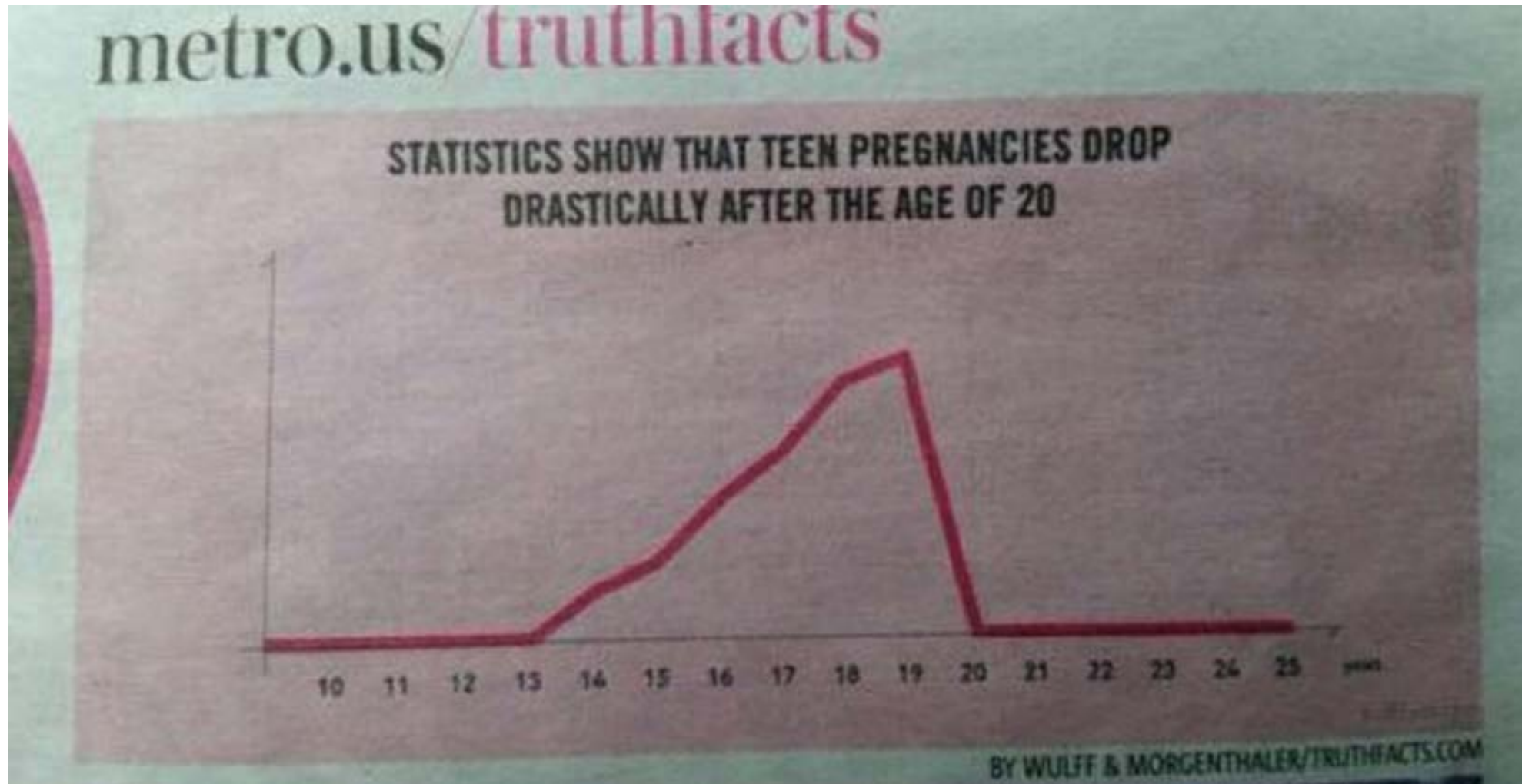
source statista via @mikequindazzi

Note: Worldwide; 2014 to 2016

Further information regarding this statistic can be found on [page 100](#).

Source: Wikibon; [ID 254266](#)

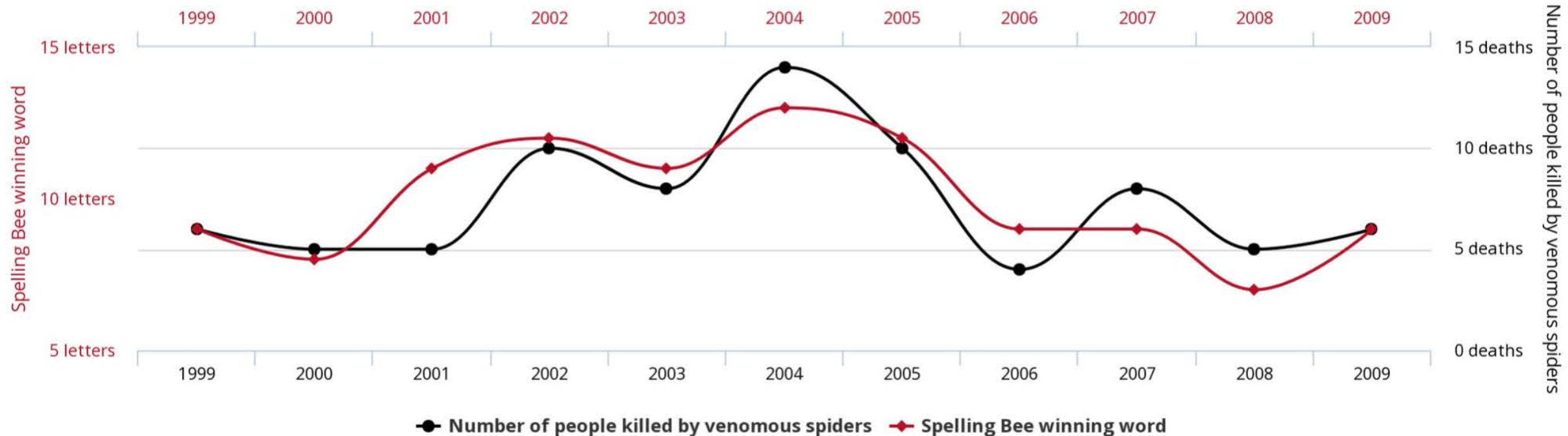
# There are Lies, Damn Lies and Big Data Lies



# If you look long enough...

**Letters in Winning Word of Scripps National Spelling Bee**  
correlates with

**Number of people killed by venomous spiders**





# So, what is the problem?

- Technology
- Data Hygiene
- Data Governance
- People



# What if.....

Fox built an algorithm for hiring

- based on the last 20 years of hiring data (fair)
- weighted in favor of those who stayed a minimum of 4 years (fair)
- weighted in favor of those who were promoted at least once (fair)



Uh Oh!

- 50% of the population would largely be filtered out
- At Fox, women traditionally didn't meet the definitions of success as noted above
- Big Data doesn't automatically bring equity and yet it is built on data! Data may not lie but it can be wrong.
- Big Data can accelerate bad decisions!



# US Military Failure



Good Pilots

Better  
Planes

Worse  
Results

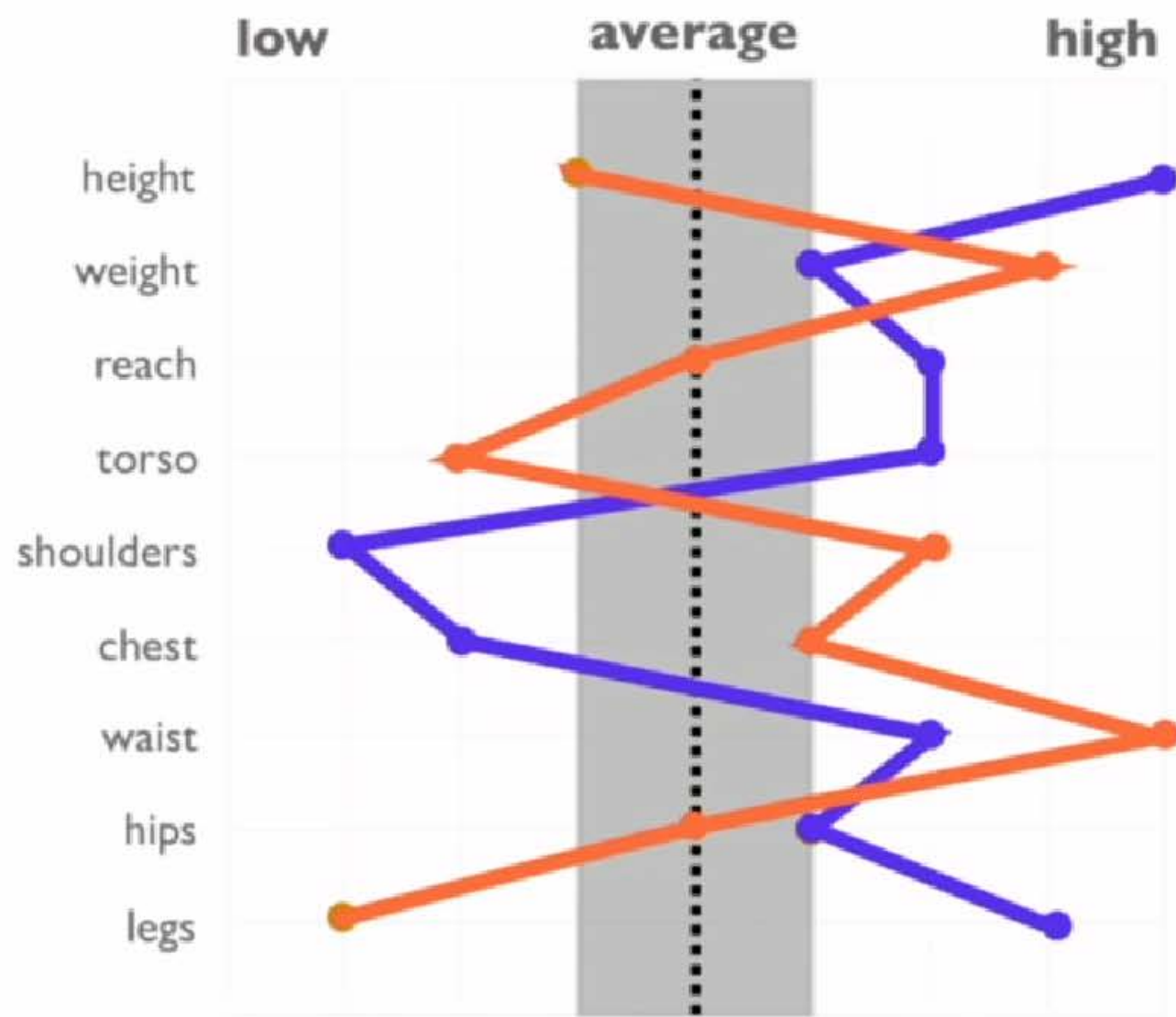
Blame the  
Pilots

Blame the  
Technology  
(planes)

Blame the  
Flight  
Instructors



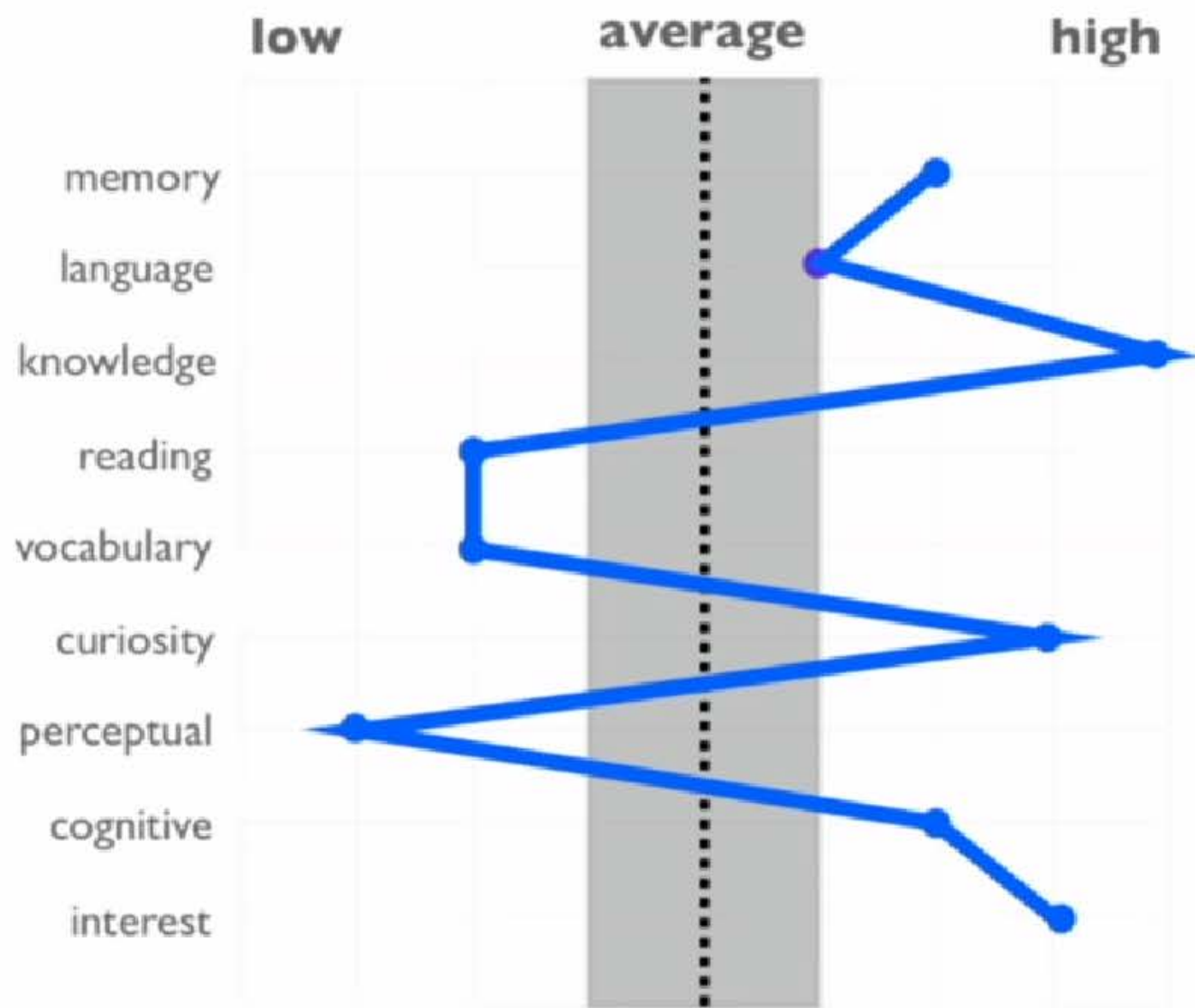
height  
shoulders  
chest  
waist  
hips  
legs  
reach  
torso  
neck  
thigh



**memory**  
language  
**knowledge**  
reading  
**vocabulary**  
curiosity  
**perceptual**  
cognitive  
**interest**

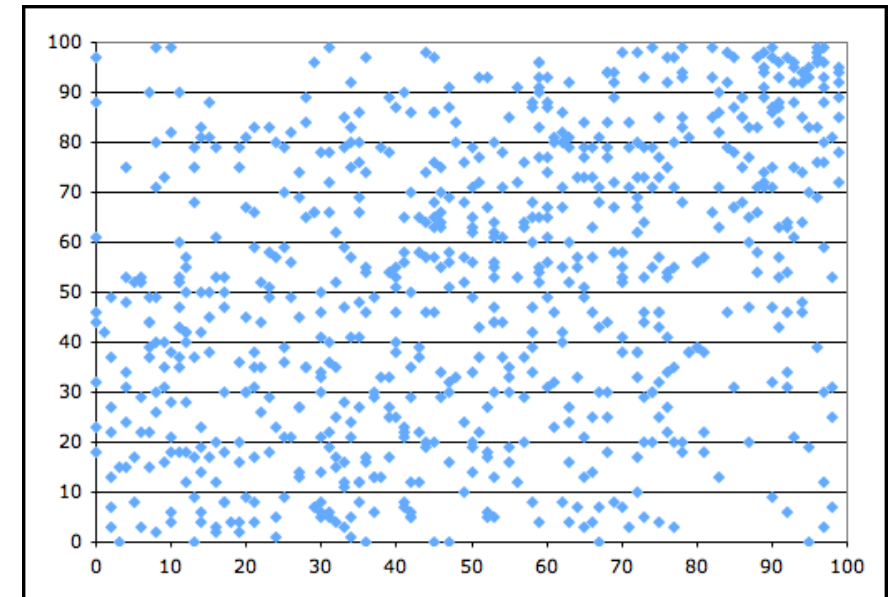
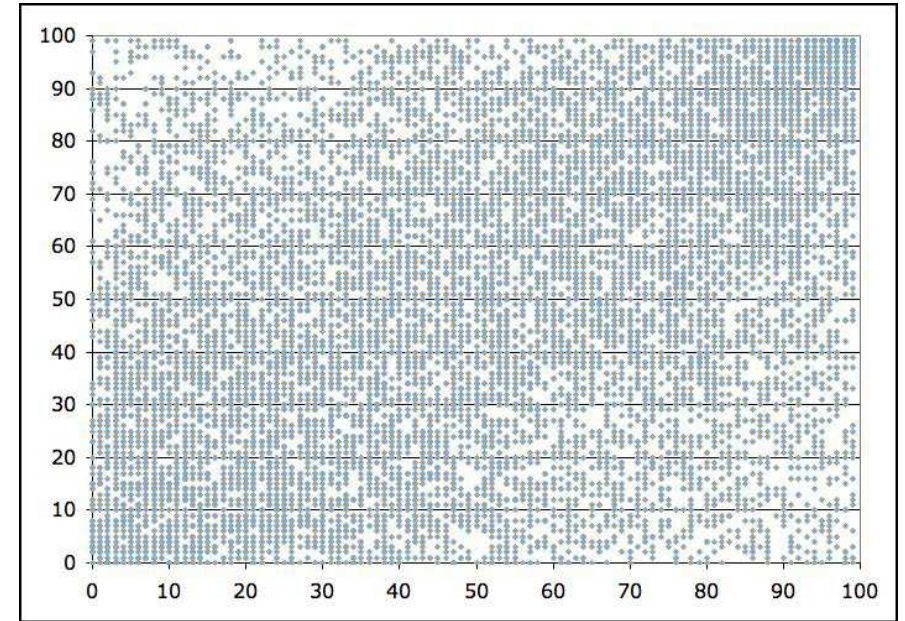






# New York City DOE-Fail

- Value Added Measure of 18,000 teachers
- Much like AI, the algorithm was never released
- The Data still tells a story
  - Scores produced for each T but against no scale or standard
  - Fig 1- 13K Ts 08-09 vs 09-10 with no discernable pattern—data hygiene
  - Fig 2- 707 Ts who had 1<sup>st</sup> and 2<sup>nd</sup> yr data—eye test



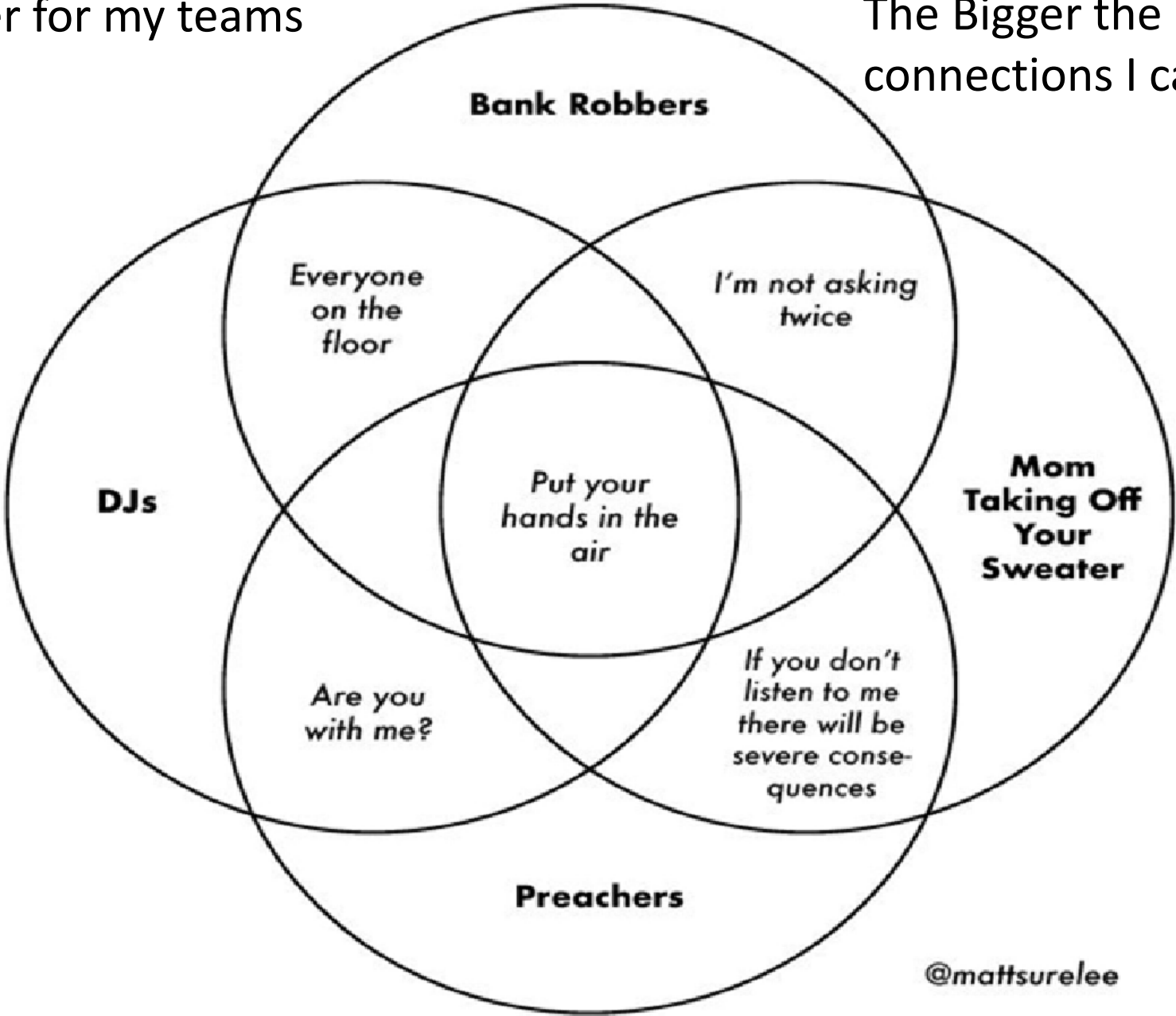


# Big Data in Hiring

- Embrace the learning that comes from failures, don't just follow the success, a real culture shift on this
- Flip the model, let the machines work then ask the research questions
- Flip the model, does data predict success in the job or it can predict the need to change the job
- Add distinct and un-expected data source(s)
- Always do the eye test
- Essential entire organization of open to the transparency this brings
- Start SMALL but commit BIG
- Design to the edges, not the middle, for diversity

A simple reminder for my teams

The Bigger the Data the more connections I can make



@mattsurelee

# Thank You!

Paul Edelblut

[pedelblut@vantage.com](mailto:pedelblut@vantage.com)

267-756-1159

Twitter: PaulEdelblut