

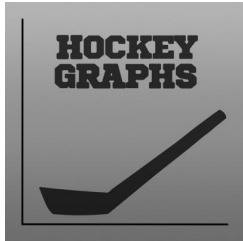


# Who Tracks the Trackers?

(Blame Josh Weissbock for the title)



Before



Now

# Who Are We?

Data tracking and analysis company

Specific and large scale tracking

Individual player to teams to leagues

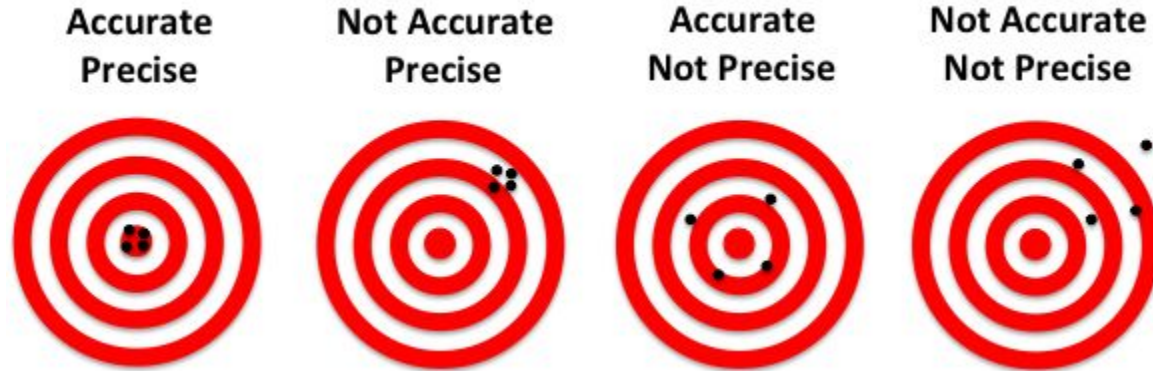
Clients range from a bantam kid to NHL teams

Track through video using combination of machine learning and individual tracking

Large team of trackers



# Three Potential Problems with Data



# Pt. 1) Accuracy error and bias

## The Technical

- Systematic error
- No reproducibility issue
- Not randomly or normally distributed
- Relative bias and absolute bias

## Source of the Problems

- Individual tracker bias
- Inherent human bias
- Measuring apparatus bias

## The Problem

No matter how much you increase sample size, you will not fix the error.

# Individual Tracker Bias

A WEIGHTED RANDOM NUMBER  
GENERATOR JUST PRODUCED  
A NEW BATCH OF NUMBERS.

LET'S USE THEM TO  
BUILD NARRATIVES!



ALL SPORTS COMMENTARY

# Individual Tracker Bias

## Solution:

- Diversify and randomize trackers
  - Different people will likely have different biases
- Introduce more opinions
  - Have more than one person measure - verify ----->
- Implement auto verification process
  - Use technology to create limitations on bias error
- Test performance versus others
  - Incentivize accuracy thresholds
  - Introduce “learning opportunities”
  - If you have to... adjust for in bias





# Inherent Human Bias

# Inherent Human Bias

Solutions... need to be a bit more creative...

- Technology prevent possibilities
- Test hypotheses in isolation of the potentially confounding forces
  - Home scorer's bias
  - M. Schuckers and J. Curro location arena adjustment
    - *(edit: apparently not very good by his opening remarks)*
- ....easier said than done

# Measuring Apparatus Bias

Solutions:

- Get better technology
- Boring meetings with developers

# Pt. 2) Precision error

## The Technical

- Random error
- Reproducibility issue
- Normally distributed

## Source of the Problems

- Usually: not being perfect

## The Problem

Sample size will give you the truth... but who wants to wait?

# Not perfect?

Solutions:

- Get better technology
- Minimize levels of human interaction with data
- Verify with other sources