

QB-iQ

Gene Woodstock
Data Scientist

Football, in a snap

Objective

- Touchdown (6 points): possessing the football in the opponent's end zone.
 - Bonus play after is worth 1-2 additional points, if successful.
- Field Goal (3 points): kicking the ball through the “uprights” located in the opponent's end zone.

Process

- Move the ball via “forward pass” or by “running” with the ball.
- 4 opportunities to gain 10 yards at a time. Referred to as “downs.”
- If attempts to gain 10+ yards was a success, the number of downs is reset.
- A “Drive” starts when a team gets the ball and ends when they fail to score or convert a 1st down.

A Brief History of Football Analytics

Traditional Analytics

Summation of Performance

1. Pass yards
2. Yards per play
3. Points per game
4. Total 1st downs
5. 3rd down conversion rate
6. Interception rate

Advanced Analytics

Total QB Rating (QBR)

Expected Points Added (EPA)

Advanced Analytics

Total QB Rating (QBR)

Expected Points Added (EPA)

A metric used to measure the impact a Quarterback makes independently from the performance of teammates.

For example,

Yards gained in an end of game comeback are weighted more heavily than yards gained while leading by 24+ points in the first half.

Advanced Analytics

Total QB Rating (QBR)

Expected Points Added (EPA)

Not all yards are valued equally.

(Combined point value) x (Probability of all potential outcomes)

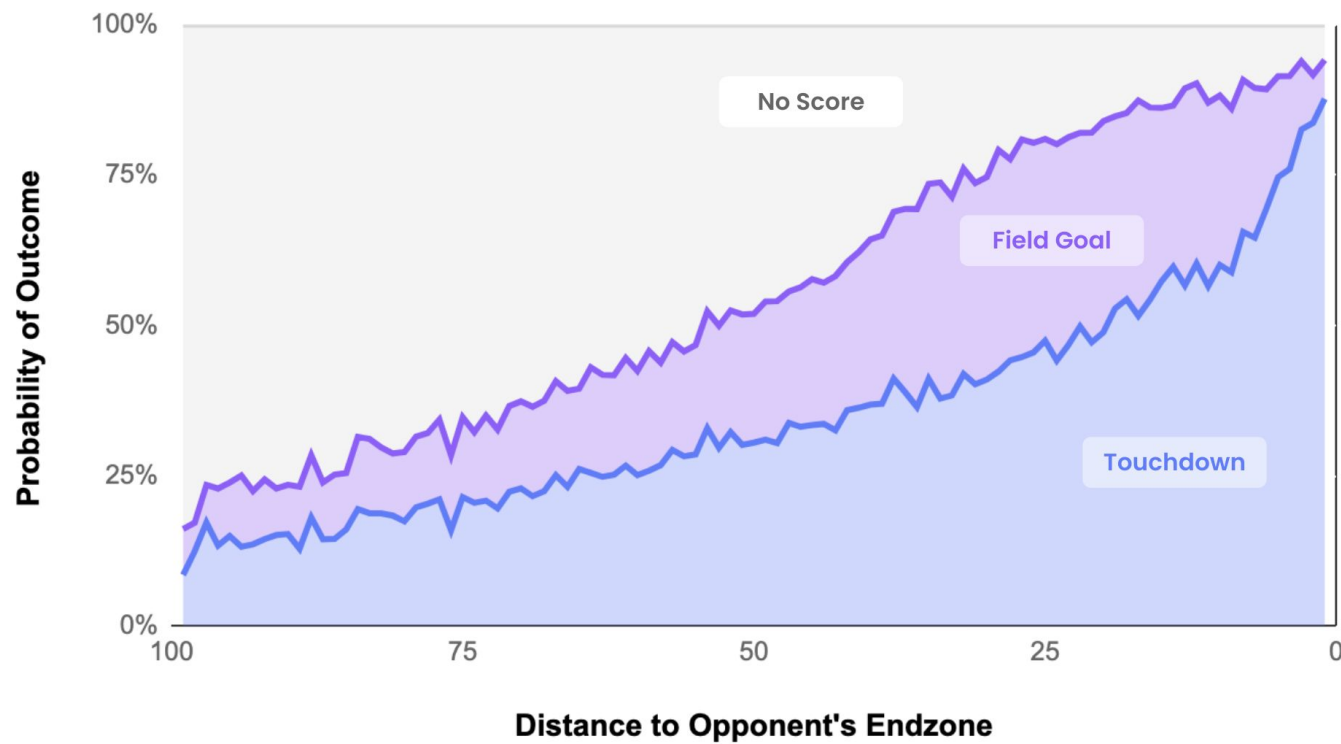
As teams approach their opponent's end zone, the probability of scoring increases.

Authors

Virgil Carter Cincinnati Bengals Quarterback in 1970

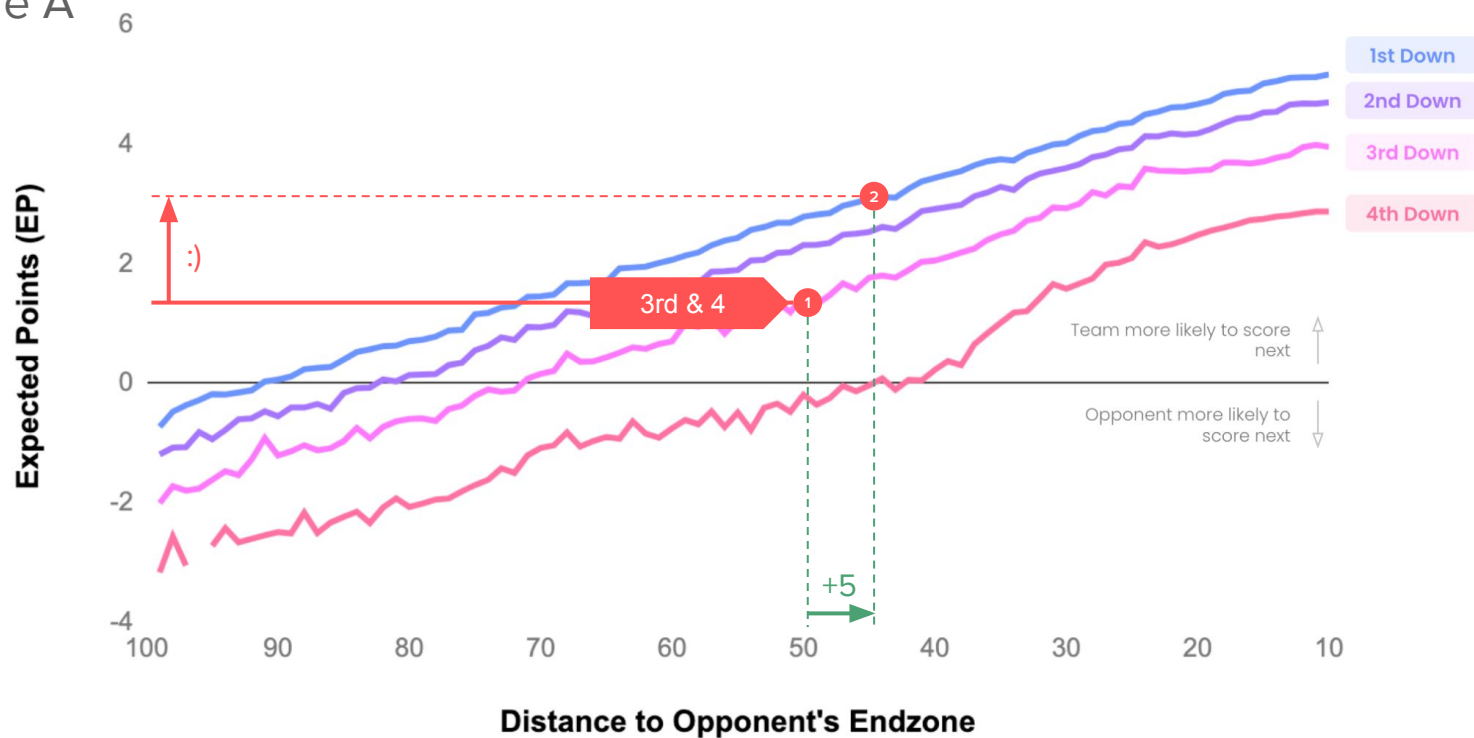
Prof. Robert Machol, Northwestern University

EPA Visualized



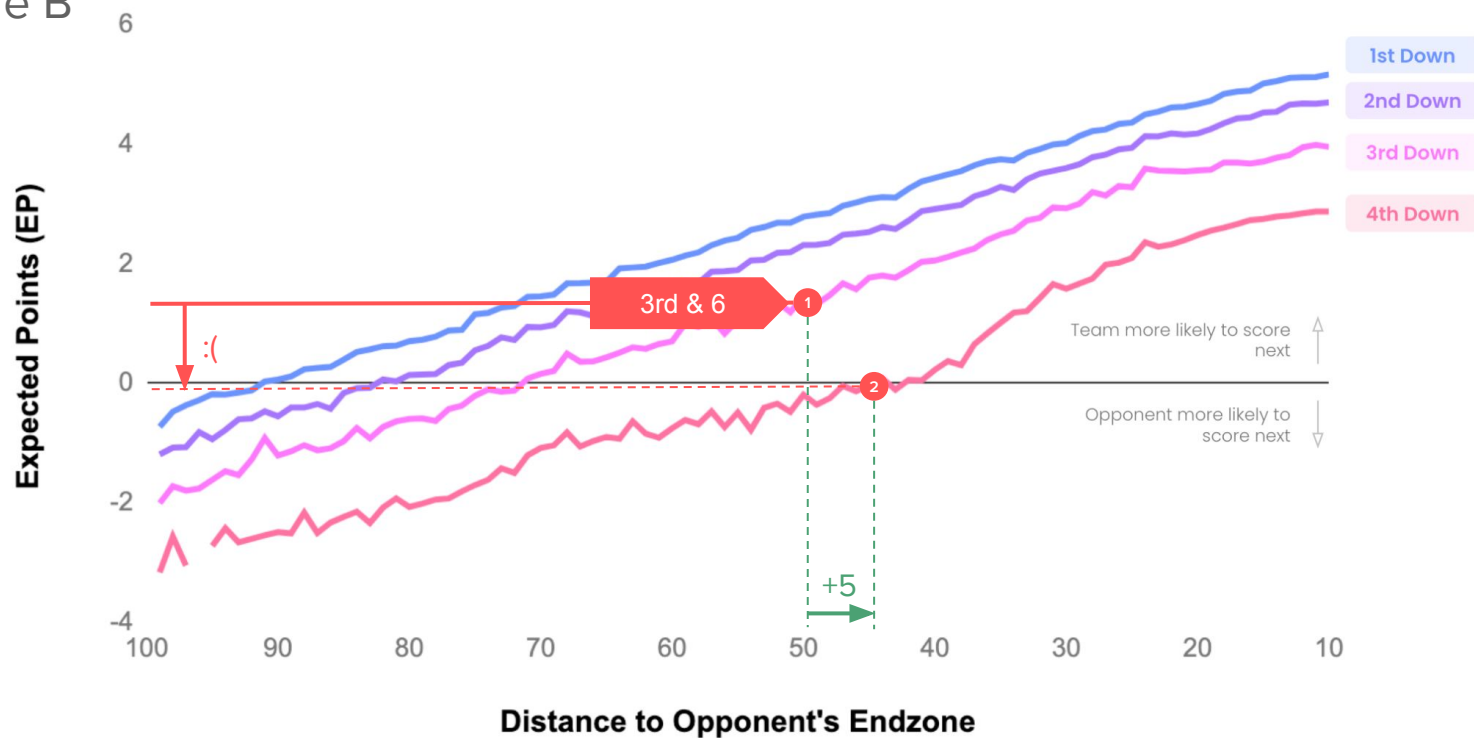
EPA Visualized

Example A



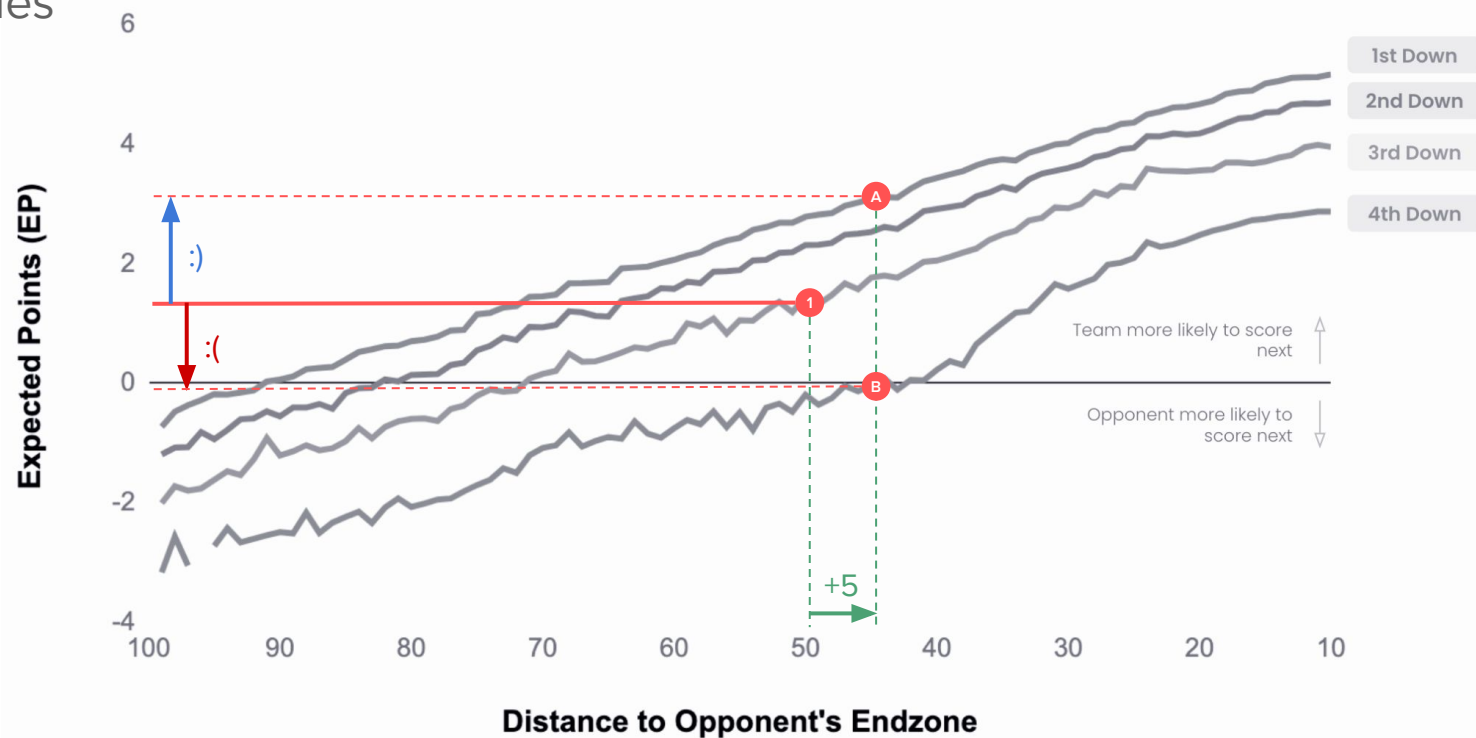
EPA Visualized

Example B

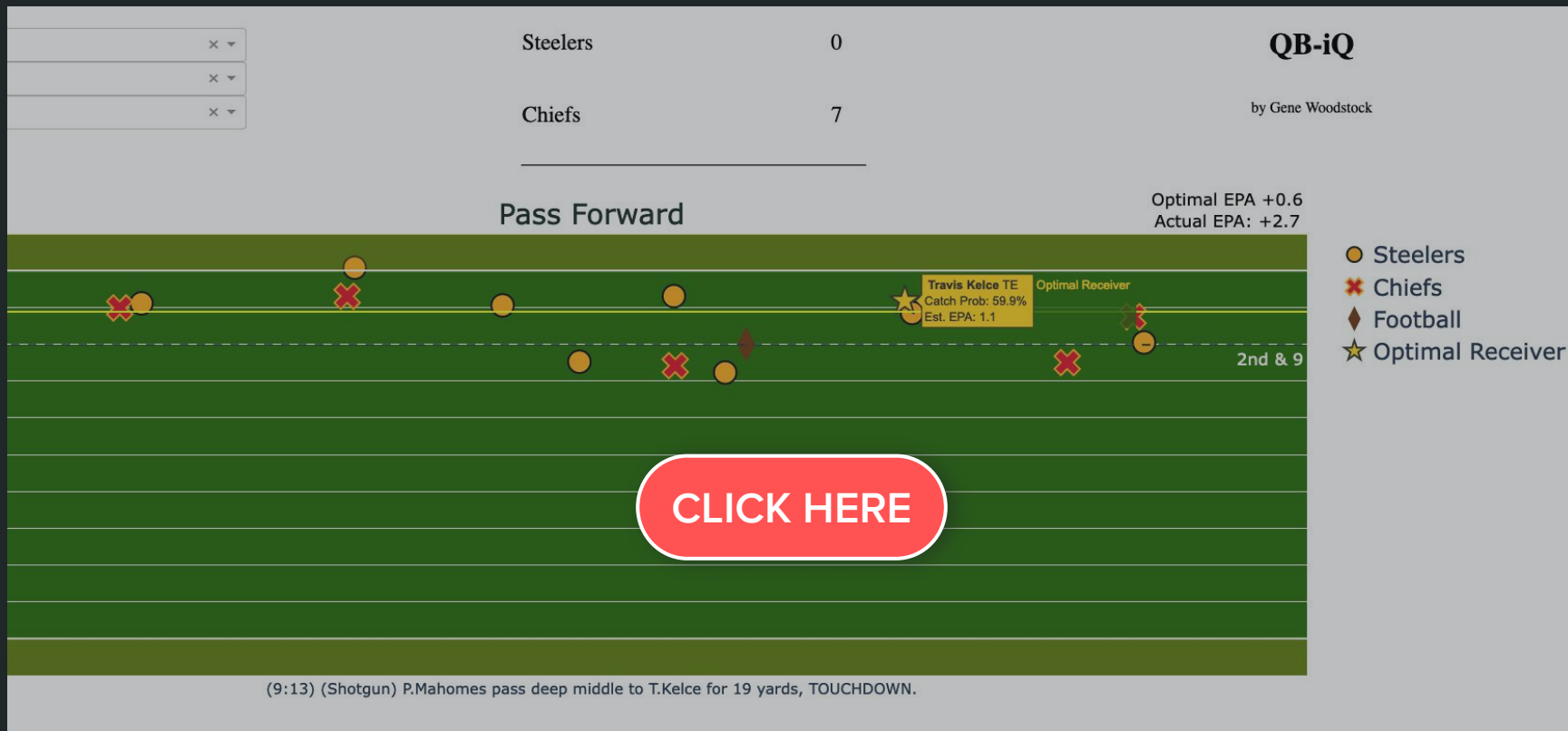


EPA Visualized

Examples



Interactive Dashboard



Problem Statement

Limited Measurements

Traditional football metrics only measure player's performance as it happened on the field – which can be measured and recorded in a “box score.”

However, no metrics currently exist that measure a quarterback's **decision making**.

Solution

QB-iQ

Pushing the boundaries of advanced football analytics
by quantifying **all possible outcomes**.

“NFL player tracking, *also known as Next Gen Stats,*
is the capture of real time location data, speed and acceleration
for every player, every play on every inch of the field.

Sensors throughout the stadium track tags on players' shoulder pads,
charting individual movements within inches.”

– NFL

QB-iQ Calculation

QB-iQ is the percentage of successful plays.

$$1 + (\text{Estimated Optimal EPA (opEPA)} - \text{Actual EPA (EPA)}) * \text{Actual EPA}$$

If $\text{EPA} \geq \text{opEPA}$, the play is scored as a 1 (full success)

If $\text{EPA} > 0$, the play is scored as a 0.5 (partial success)

If $\text{EPA} \leq 0$, the play is scored as a 0 (failure)

Top 5 QB-iQ from 2018

Patrick Mahomes	63.19%
Drew Brees	52.99%
Jared Goff	50.29%
Russell Wilson	48.06%
Matt Ryan	47.22%

Bottom 5 QB-iQ from 2018

Sam Darnold	-0.04%
Josh Allen	-0.09%
Josh Rosen	-0.14%
Blaine Gabbert	-0.32%
Josh McCown	-0.34%

Modeling Process

Baseline Model

Features	X, Y coordinates, Speed, Acceleration, Dist to Defender, Position, & Route
Metrics of Choice	Accuracy & balanced accuracy
Baseline Model	All QB pass completion rate ~ 63% Imbalanced data ~ 2:1 split

Models

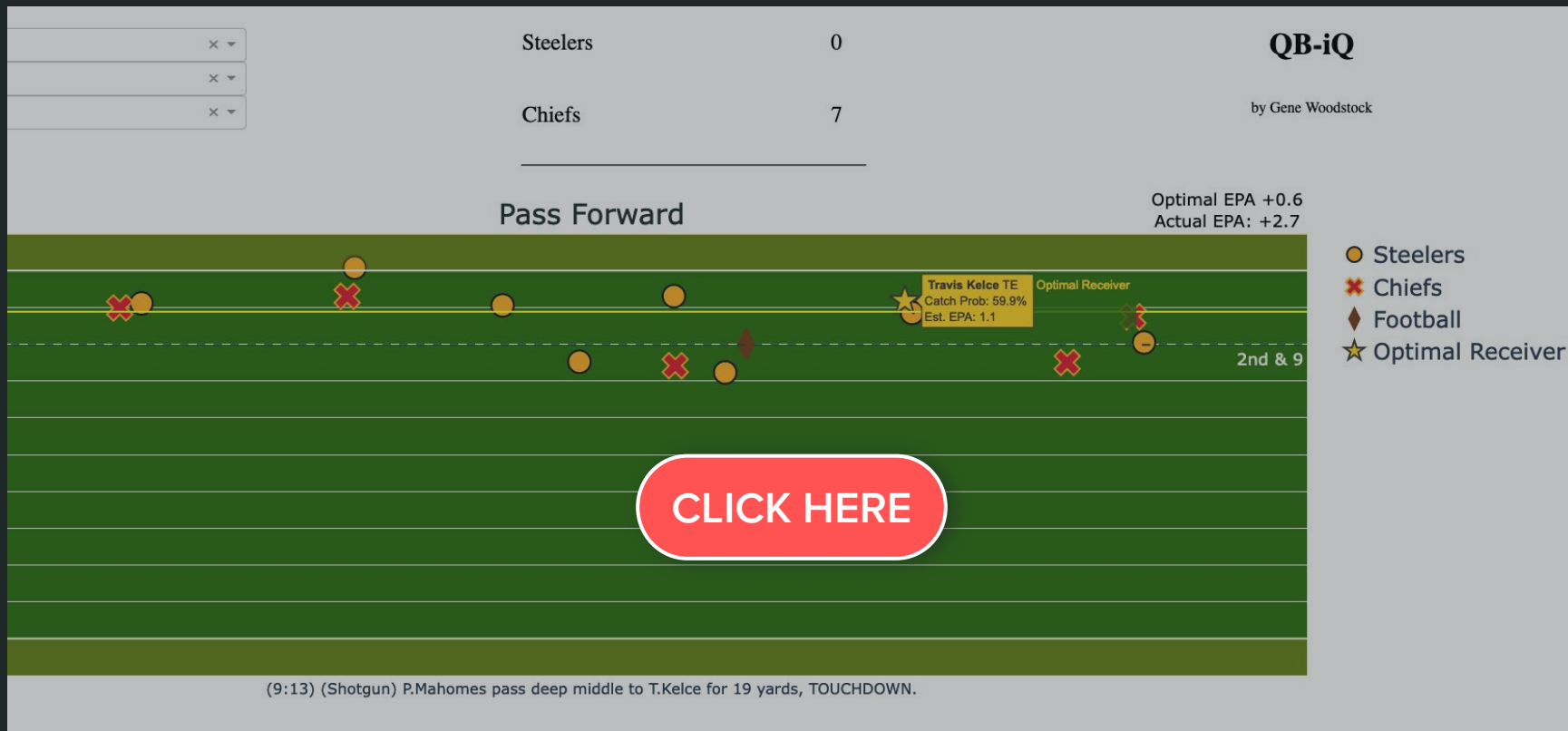
Simple Model Performance

- Logistic Regression
 - Accuracy Score: 61.2%
 - Balanced Accuracy: 61.9%
- Random Forest
 - Accuracy Score: 67.7%
 - Balanced Accuracy: 64.4%
- Extra Trees
 - Accuracy Score: 66.5%
 - Balanced Accuracy: 64.2%

Deep Learning

- Neural Network
 - Accuracy Score: 66.3%
 - Balanced Accuracy: 63.2%
- Ensemble
 - Accuracy Score: 67.9%
 - Balanced Accuracy: 64.8%

Interactive Dashboard



Strengths & Limitations

- **No access to offensive/defensive linemen coordinates.** Unable to know if a QB was under pressure when throwing the ball in many instances.
- **Missing Data.** Some plays were missing the opposing side of the ball. Could be other data inconsistencies.
- **Only 1 season.** 250+ games is a lot, but still a small sample size.
- **Weather** can affect pass accuracy and is not currently included.

Next Steps

1. Forecast future steps/locations for improved model performance. Where will the player's location relative to the defender be in 1, 2, 3, 5 frames from when the ball was thrown.
2. Model using RNN.
3. Build a model to gauge a WR's ability to get open/create separation.
4. Build a model to determine if the defense is in man coverage or zone coverage.
5. Add career stats as model inputs for players. Will help contextualize the priors.

Questions?

Gene Woodstock
Data Scientist