

PROYECTO FINAL BIG DATA

Serrana Aguirregaray, Juan Alzugaray y Santiago
Torre

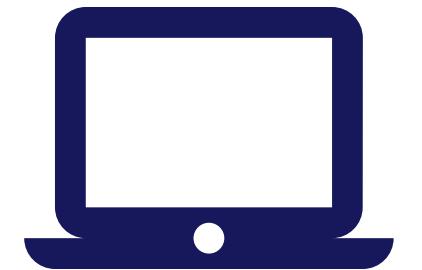
Elección del problema



Datos de NBA



Predicción de partidos



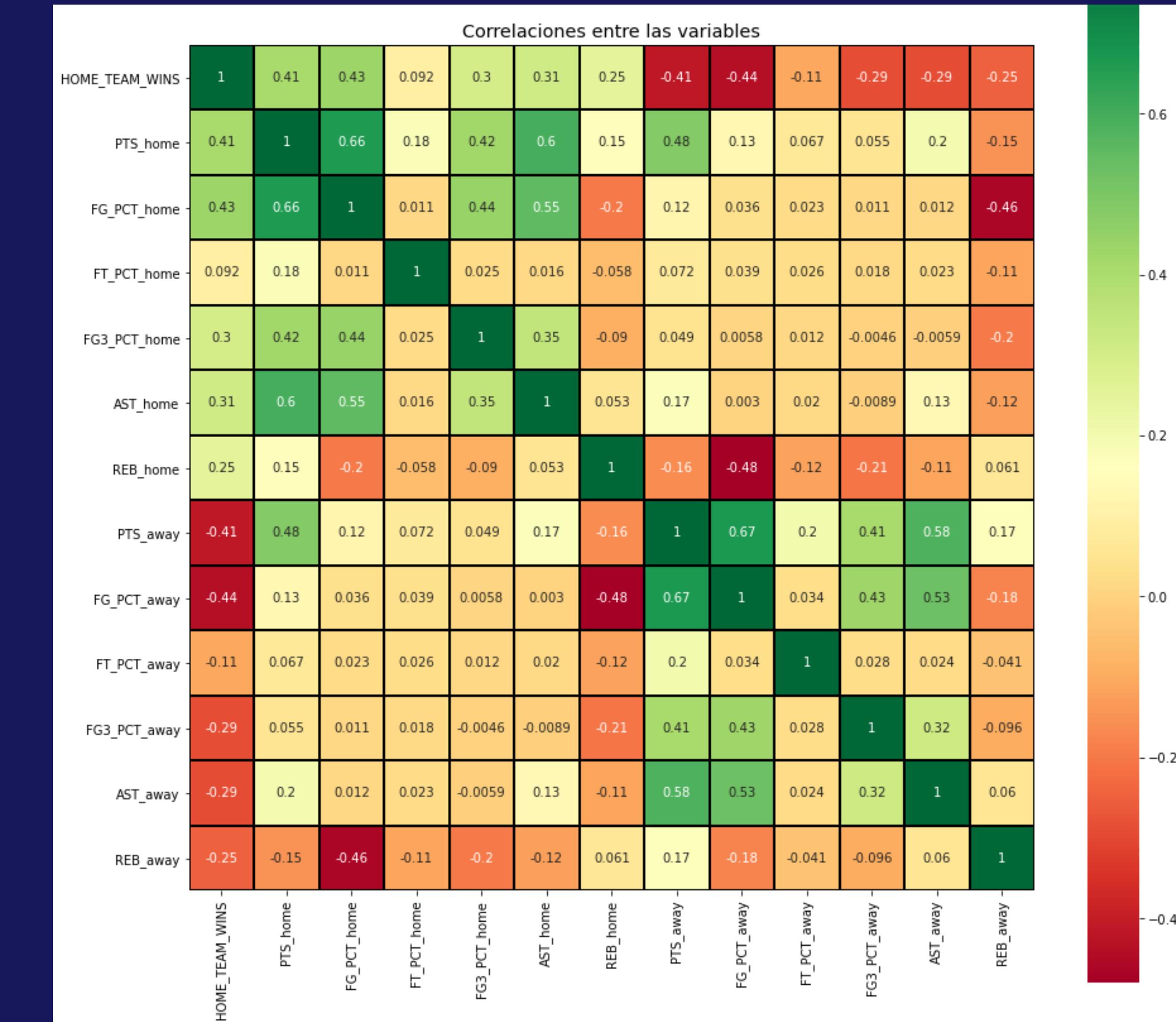
Simulación del próximo campeón



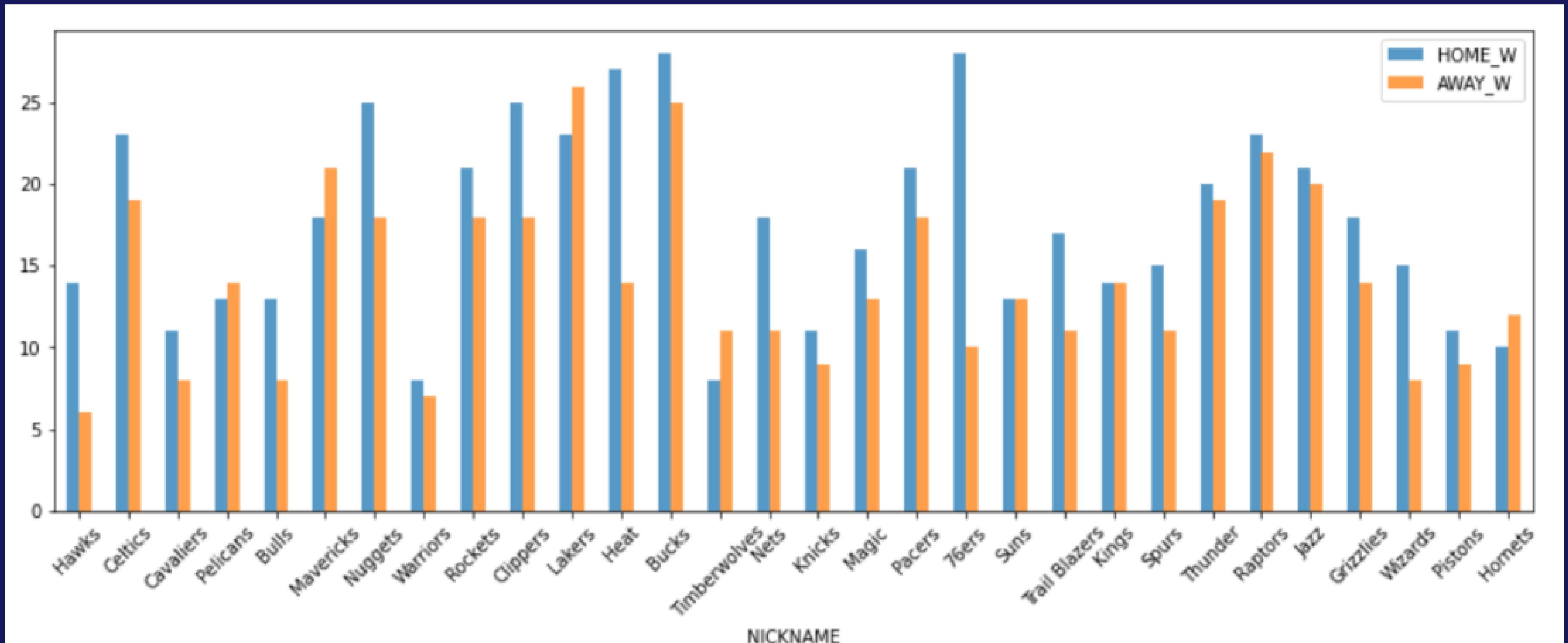
Ganar dinero

Data exploration

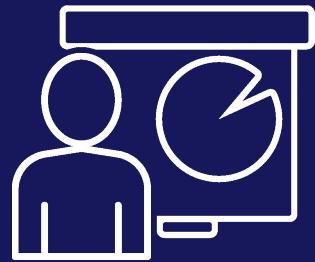
- Estudio previo de las variables del dataset
- Necesario para una correcta elección de features a utilizar



Data exploration



Feature Engineering

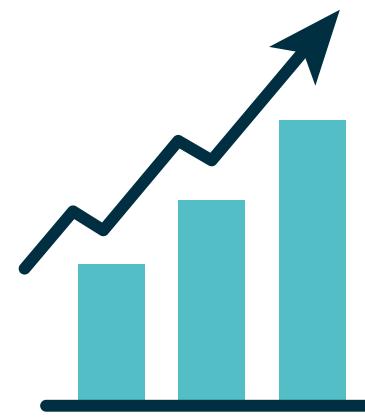


Objetivo: buenas predicciones

Desafío: hallar features significativas

Desafío: agregar nuevas features

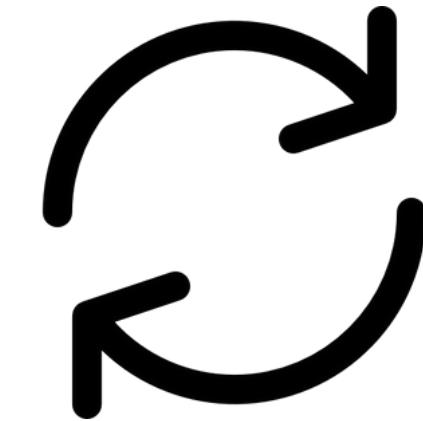
TEAM ELO



Ganar/perder hace que el ELO del equipo aumente/disminuya



Cuánto aumenta depende del margen de victoria y la expectativa previa del partido



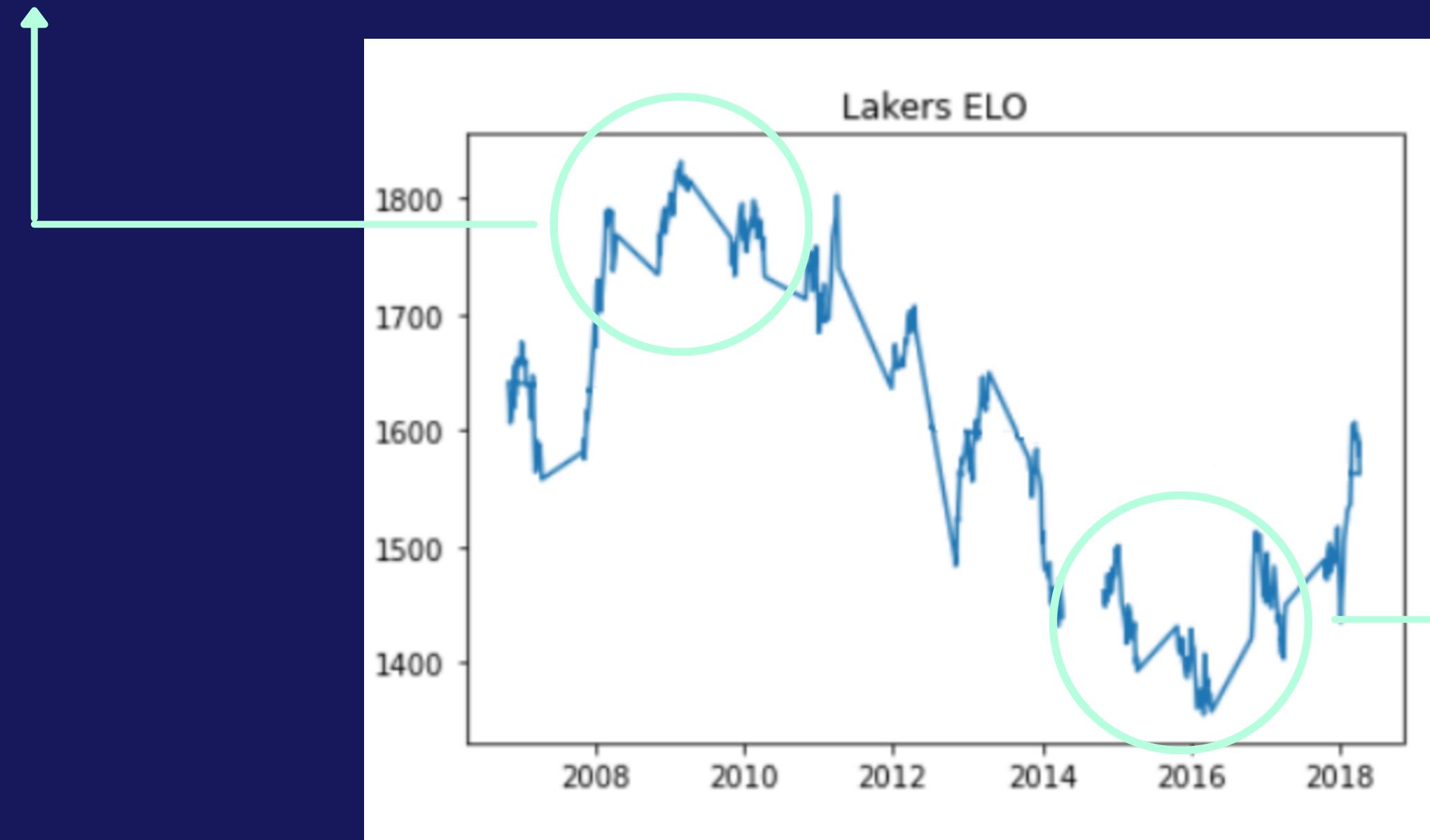
Se actualiza después de cada juego



Indicador de qué tan bueno un equipo es

Ejemplo

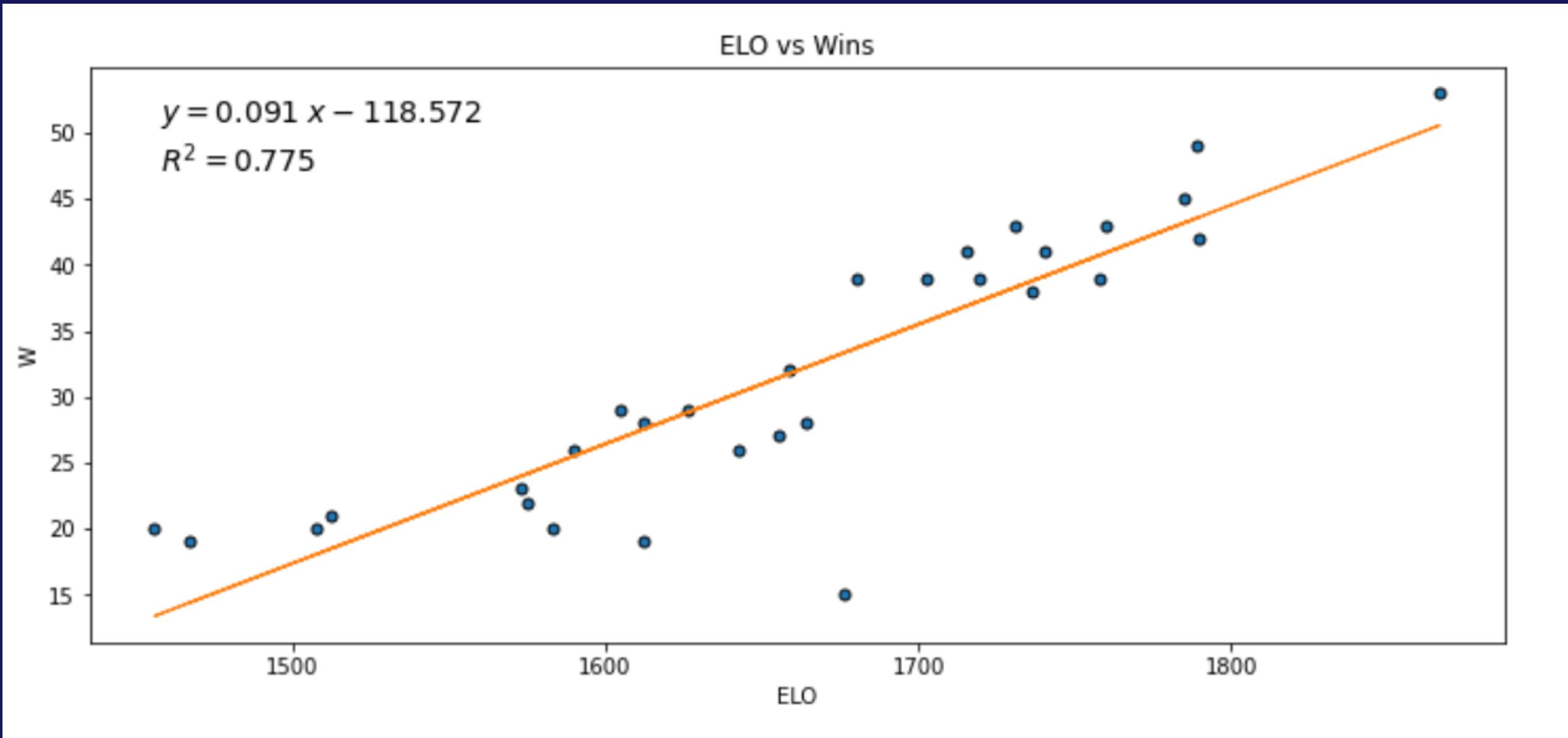
Campeones en 2009 y 2010



Últimos en 2016



ELO vs WINS



El ELO explica más del 75% de la variación de WINS de un equipo.

Efficiency



01 Combina todas las estadísticas de un jugador

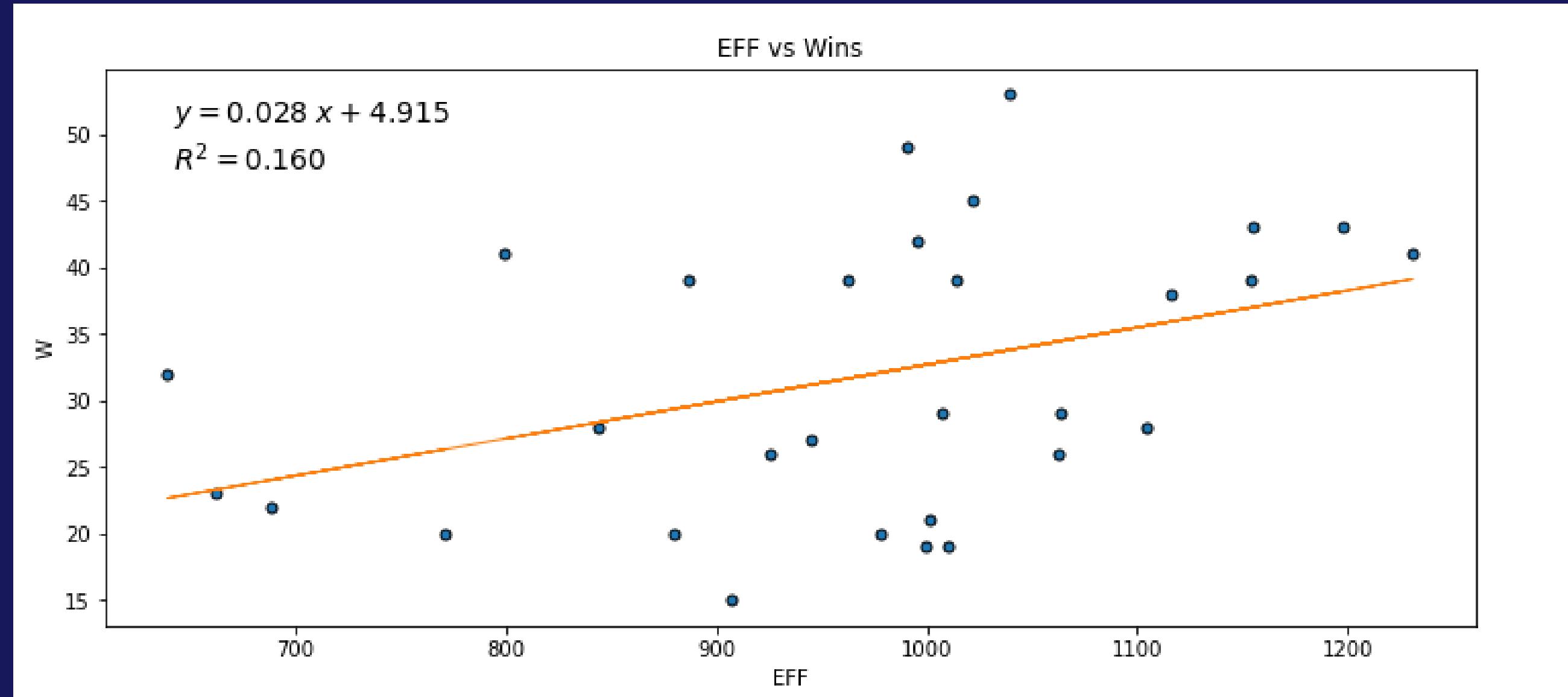
02 Se calcula una efficiency para cada jugador

$$EFF = PTS + REB + AST + STL + BLK - (TO + FGMISSES + FTMISS)$$

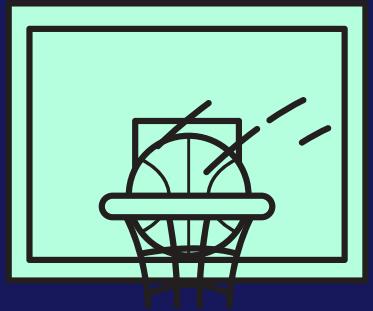
03 Se determina una efficiency por equipo como el promedio de la de sus jugadores en la temporada



EFFICIENCY vs WINS

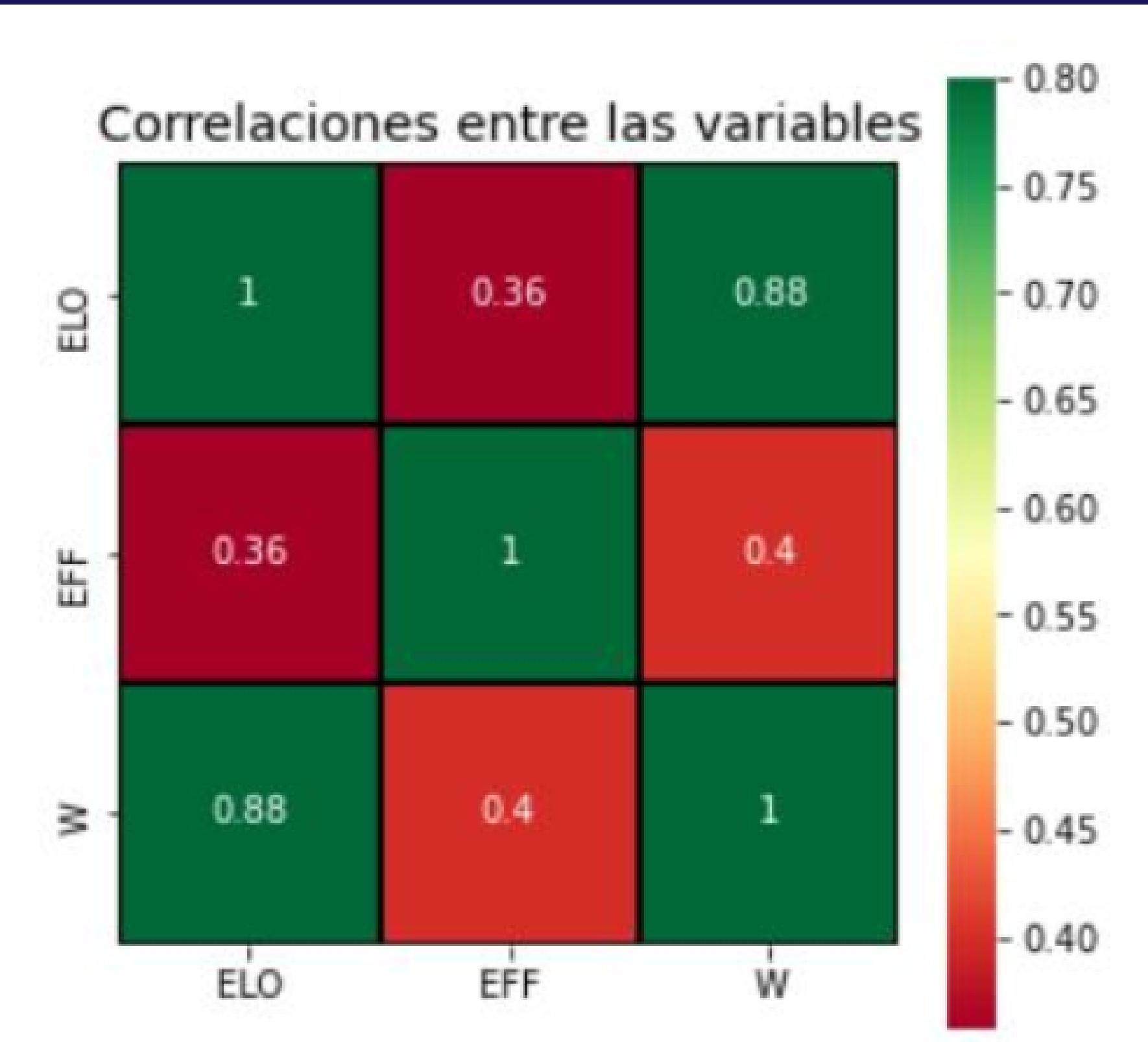


Aunque se vea una gran dispersión, hay una tendencia positiva. Explica un 16% de la variación de las WINS de un equipo.



Correlación

- Alta relación entre ELO y WINS.
- Relación considerable entre Efficiency y Wins.
- Relación menor entre Efficiency y ELO.





PREDICTION

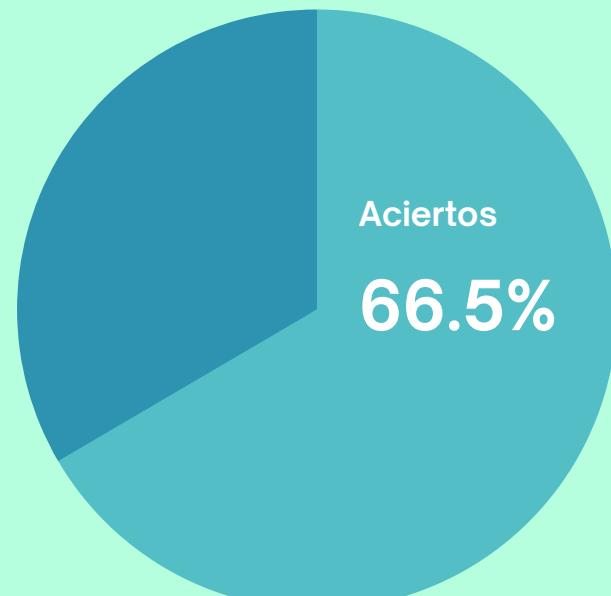


Spark RANDOM FOREST

ROC



PORCENTAJE DE ACIERTOS



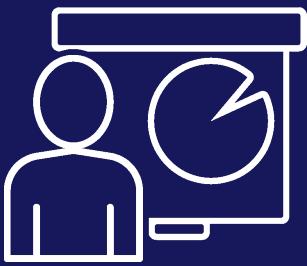
Features

- Season
- Team ELO
- Team efficiency
- Wins (total, as local, as visitor)
- Loses (total, as local, as visitor)

SIMULATION



Objetivo: simular playoffs de la temporada 2021



Desafío: resultados que concuerden con la realidad



Desafío: predicciones de campeones para apostar y ganar dinero

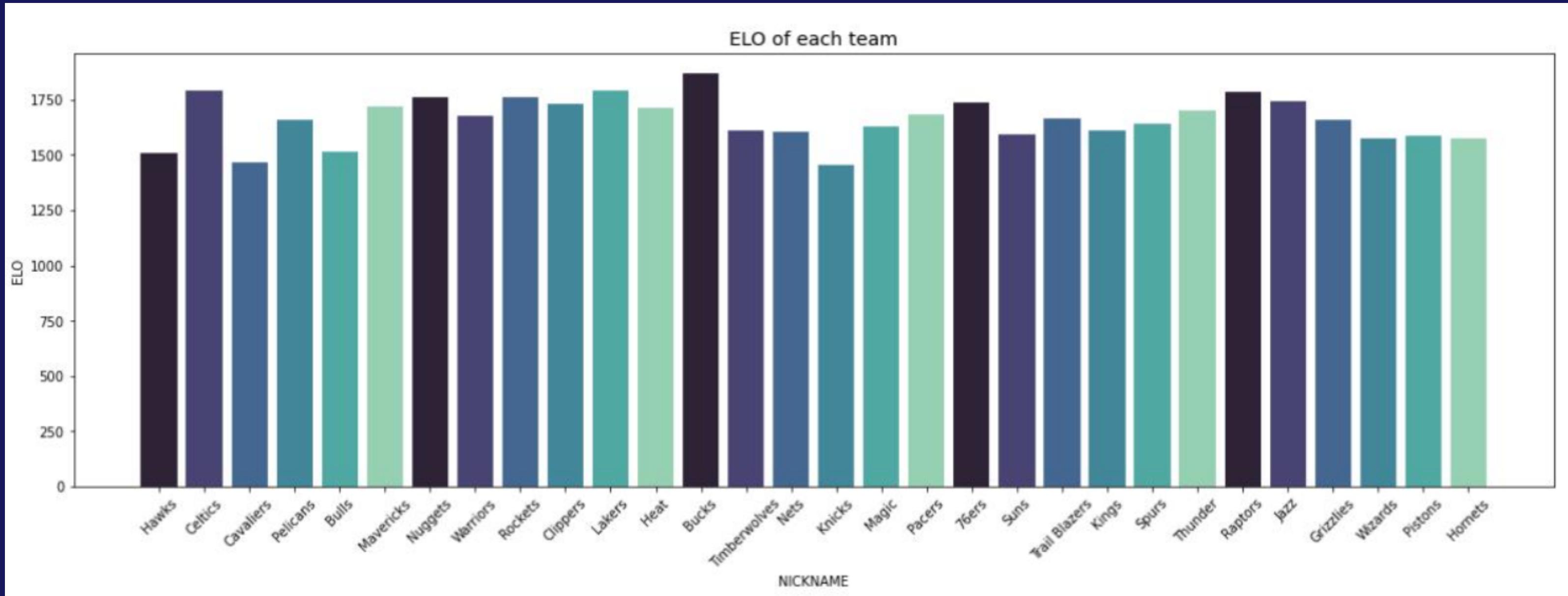
NBA playoffs structure

- Cruces predeterminados por sus posiciones en la temporada regular
- Pasa la serie el primero en ganar 4 partidos
- 8 equipos de cada conferencia disputan entre ellos un lugar en la final.
- Los ganadores de cada conferencia juegan entre sí para determinar el campeón de la NBA

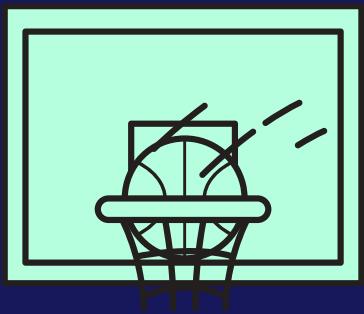




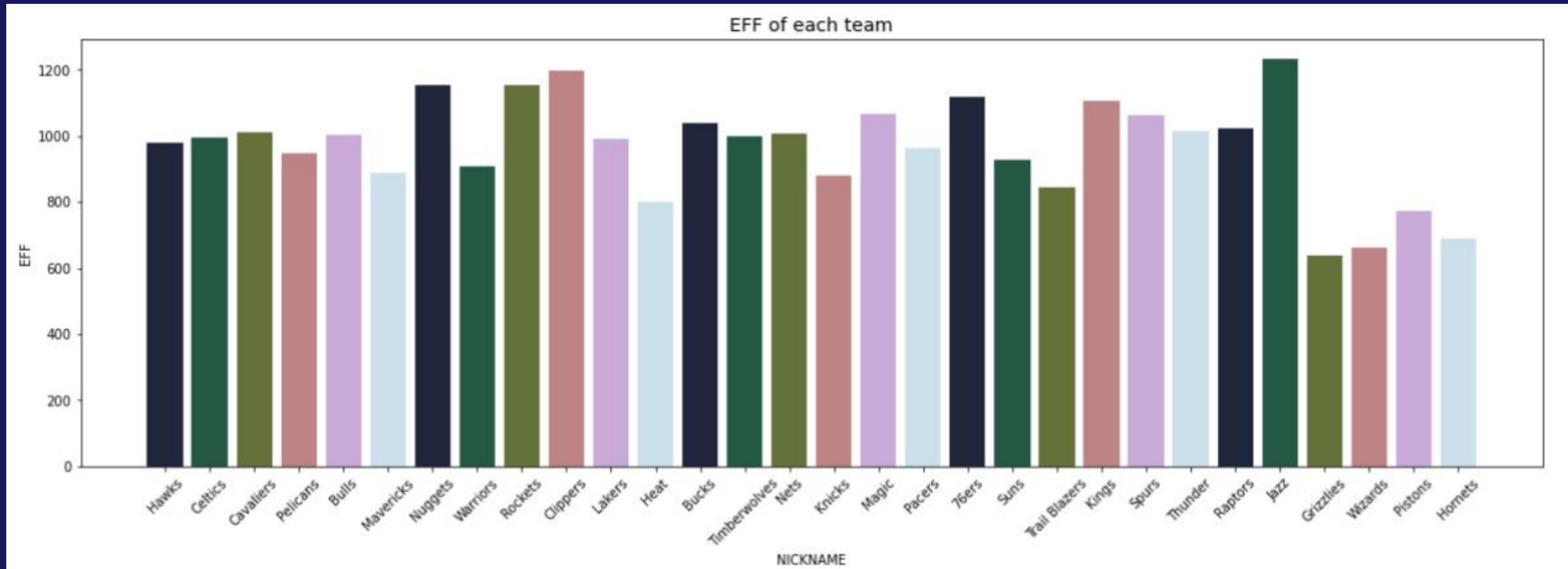
Season 2021



El ELO de cada equipo antes de comenzar los Playoffs



Season 2021



La Efficiency de cada equipo antes de comenzar los Playoffs

Simulation



- 01 Se simula quién gana un partido en función de la probabilidad de ganar según el modelo
- 02 Se simula la anterior 7 veces y devuelve el ganador de la serie
- 03 Se simula cada conferencia y la final, retornando el campeón
- 04 Se realiza N veces, determinando probabilidades de salir campeón para cada equipo

Season 2021 - Real

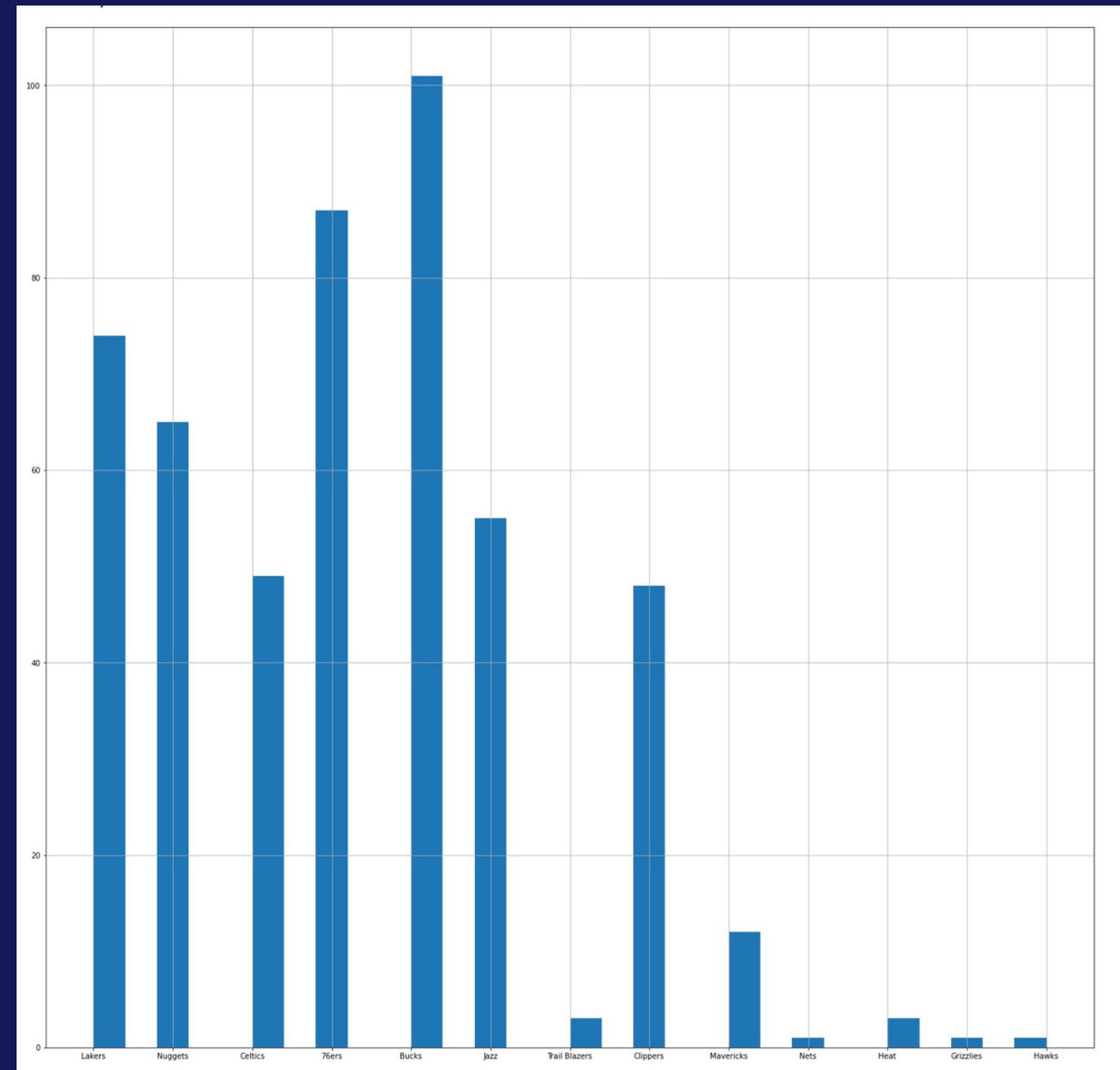


Season 2021 -

Simulación de Sportbetting



Season 2021 - Simulation



¡Gracias!

¿Preguntas?