

NBA Investigation

Edmund Lo, Zexi Lv

Domain

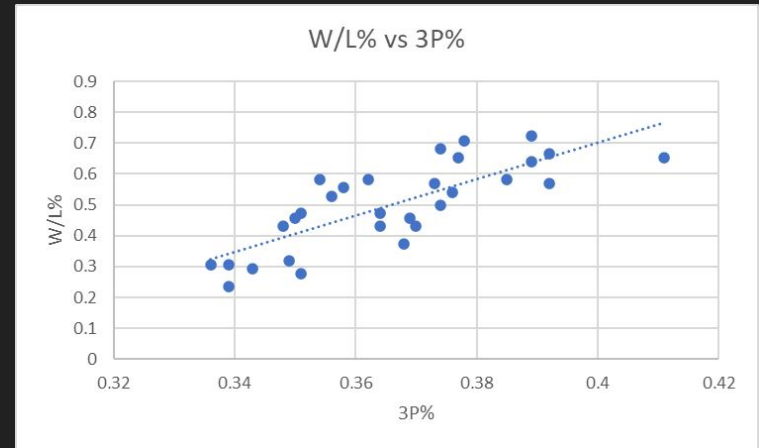
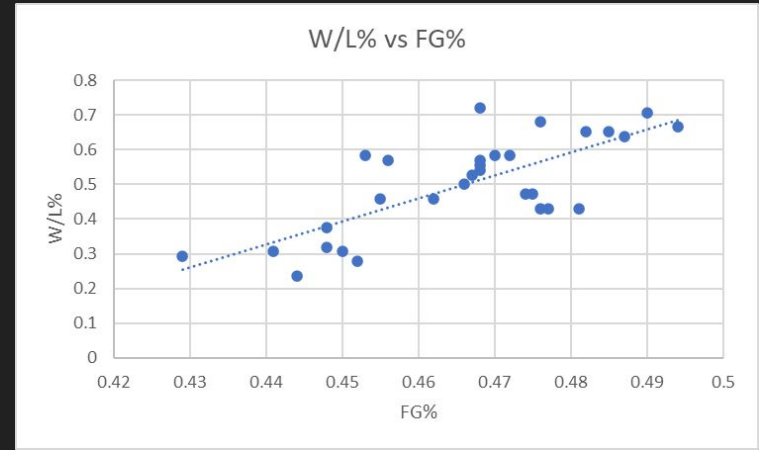
- NBA 2020-2021 regular and playoff season
- Statistics for teams, players, coaches and games

Investigative Questions

1. What are the stats or factors that have the greatest effect on a team's win/loss percentage in the 2020-2021 NBA season?
2. What are the stats or factors that have the greatest effect on a player's points per game in the 2020-2021 NBA season?
3. What is the relationship between a coach's experience and their team's success during the 2020-2021 NBA season?

Results - 1. Team Winrate

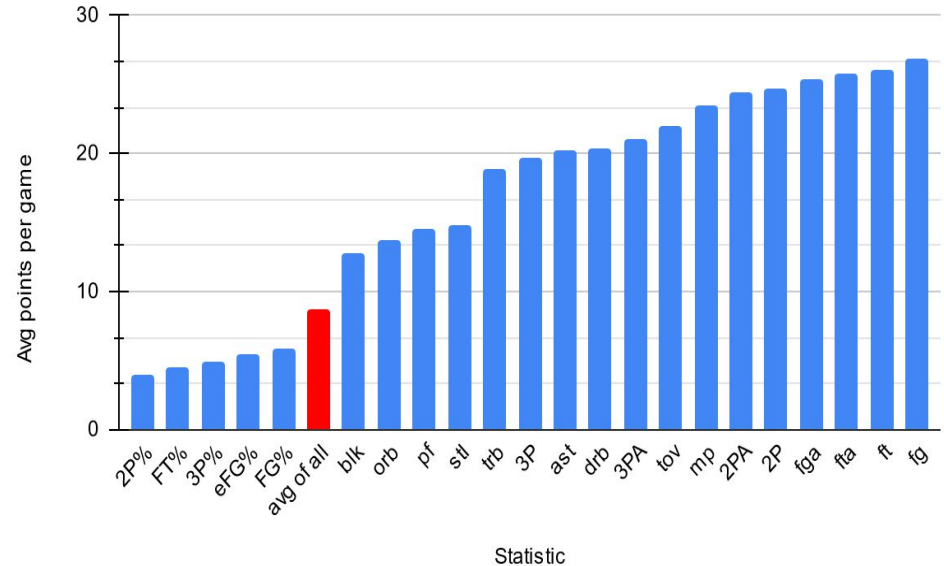
- Top ranked teams had much higher FG% and 3P%
- Accuracy was a great indicator of success
- 8 of the top 10 teams also had a top 15 scorer
- Experienced and older teams had higher winrate



Results - 2. Player Performance

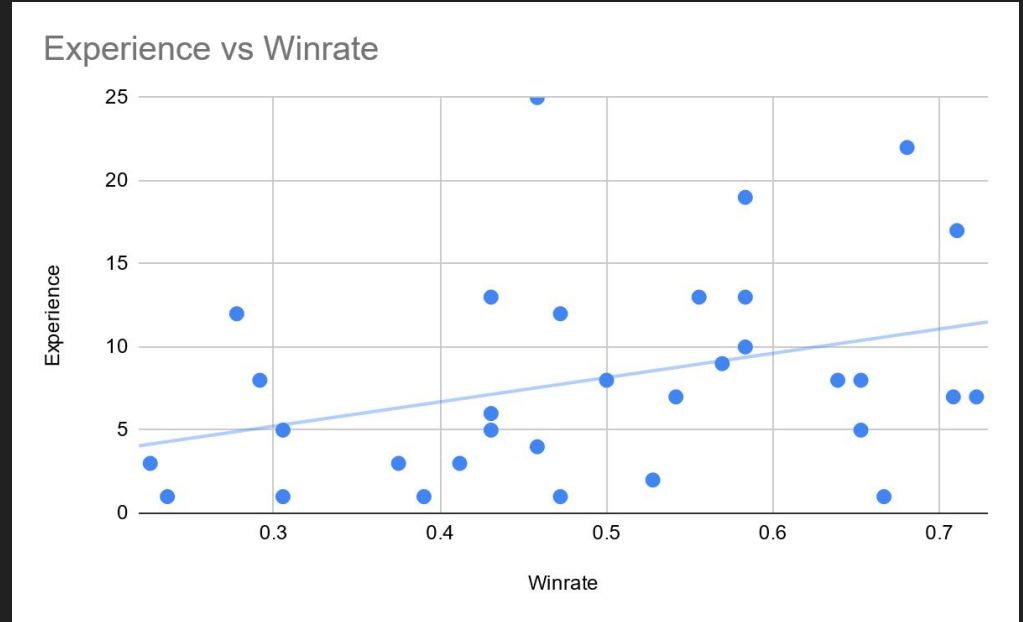
- Find score per game of top players per stat
- Scoring statistics had well-performing players
- Making many attempts lead to success
- Accuracy was not enough for success

Top Player Performance Per Stat



Results - 3. Coach Experience

- Coach experience was correlated with team success
- No observed impact on player success



Challenge - Picking a Domain

- Difficulty picking suitable domain
- Finding rich, high-quality data which was easy to use
- Interesting questions to ask
- Overcame by searching many places for data related to interests

Challenge - Design Decisions

- What to do with empty cells in our dataset
- A player's 3P% when they had never taken a 3P
- A coaches playoff wins if they never made it to the playoffs
- Solution was to substitute 0 to maintain our not null constraint

Lesson - Designing Schemas

- Best way to design a schema
- How to avoid redundancy
- Use of primary and foreign keys in each relation
- Data constraints

Lesson - Writing Queries

- Translating questions to queries
- Making observations about data
- Power of SQL

Questions?