* A note specifying which dataset you analyzed
* A statement of the question(s) you posed
* A description of what you did to investigate those questions
* Documentation of any data wrangling you did
* Summary statistics and plots communicating your final results

**A note specifying which dataset you analyzed**

The required datasets are obtained from stats.nba.com. JSON files were downloaded, and the converted into CSVs. The dataset features full relevant data from the 2017-2018 NBA regular season.

**A statement of the question(s) you posed**

1. Does 3 point shot accuracy in the 1st and the 3rd quarters?

1.Paired C->Q t-test (does shot efficiency late in the game differs from early?)

Considering their drastic improvement from 1st quarter to 3rd Minnesota may consider doing the 3-point shooting more in the warmup. Doesn’t apply to Golden State, because they already shoot above average in the first quarter, but go superhuman in the 3rd.

2 Multiple C->Q ANOVA

Rookie-sophomore-veteran? Contested shots

Position – age?

count mean std min 25% 50% 75% max

PLAYER\_POSITION

C 63.0 27.111111 4.399821 20.0 24.0 27.0 30.5 37.0

C-F 23.0 26.217391 5.071816 19.0 22.5 26.0 28.5 40.0

F 150.0 26.426667 4.045623 20.0 23.0 25.0 29.0 38.0

F-C 31.0 25.935484 3.881026 20.0 23.5 25.0 28.5 37.0

F-G 23.0 26.782609 4.155572 20.0 23.0 26.0 29.5 34.0

G 205.0 26.321951 4.006582 19.0 24.0 26.0 28.0 40.0

G-F 39.0 27.717949 5.155216 20.0 24.0 27.0 31.5 41.0

The outliers did not influence the results

Do different positions run more?

3. C->C Chi-square test for independence

(in games where the lead is within 5 points at the half, does being ahead correlate with winning?)

At 4 doesn’t (don’t have evidence), at 5 it does significantly

Usa/not usa -> position?

Remove data entry outlier Walter Lemon Jr. He's not the only "pure" pg, other pure PGs were classified as guards, so we’ll consider this data entry error

Origin International USA

PLAYER\_POSITION

C 0.476190 0.523810

C-F 0.347826 0.652174

F 0.180000 0.820000

F-C 0.322581 0.677419

F-G 0.173913 0.826087

G 0.136585 0.863415

G-F 0.179487 0.820513

All 0.213483 0.786517

p-value practically zero

shows that international basketball more likely to be centers or big men, and the trend in US basketball towards guards (“smallball”)

4. Q->Q regression t-test for the slope

Miles def -> def rebounds?

Miles def -> deflections? There is some, by eye test

Miles def -> contested shots?

Pace true shooting

Draft number height? Not linear

Height/wingspan -> deflections/ steals?

We’ll choose only players who had played more than 15 games.

Mostly not significant, although for guards is close at pvalue pvalue=0.05779801365934409. But with different minimum cutoffs for games (i.e. not 15 games, but 20, 30, 40) the p-value is even bigger, so it doesn’t confirm)

Extra.

C- > Q. Do international run more?