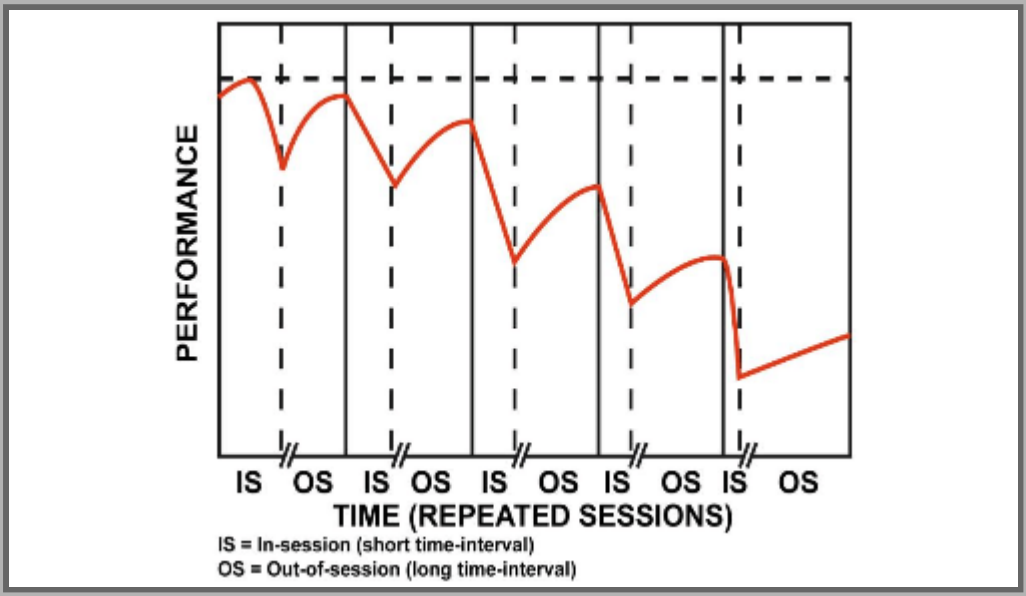
Measuring and Predicting Player Fatigue in the NBA

Austin Stephen, Matthew Yep and Grace Fain

7/13/2021

Why Fatigue?

- Injury Prevention
- Optimized Performance



Project Objectives

1. Measure player fatigue
2. Build a model that predict the impact of fatigue on a given game or multi-game sequence.

What is Fatigue

- Cannot use direct measures
- Identify a proxy

Proxy for Fatigue

- In-game distance
- Between game distance

Proxy for Fatigue

- In-game distance

Distance By Player Individual Games

Name	Date	Min	Distance (mi)	Distance Offense	Distance Defense	Speed (mph)	Speed Offense	Speed Defense
Al Horford	10/16/2018	29.95	2.15	1.21	0.94	3.91	4.15	3.64
Alex Abrines	10/16/2018	23.48	1.88	0.97	0.91	4.57	4.96	4.19
Alfonzo McKinnie	10/16/2018	2.35	0.18	0.11	0.07	4.79	4.99	4.61
Amir Johnson	10/16/2018	11.18	0.87	0.47	0.39	4.49	4.83	4.13

Data courtesy of nba.com and Second Spectrum

Proxy for Fatigue

- In-game distance

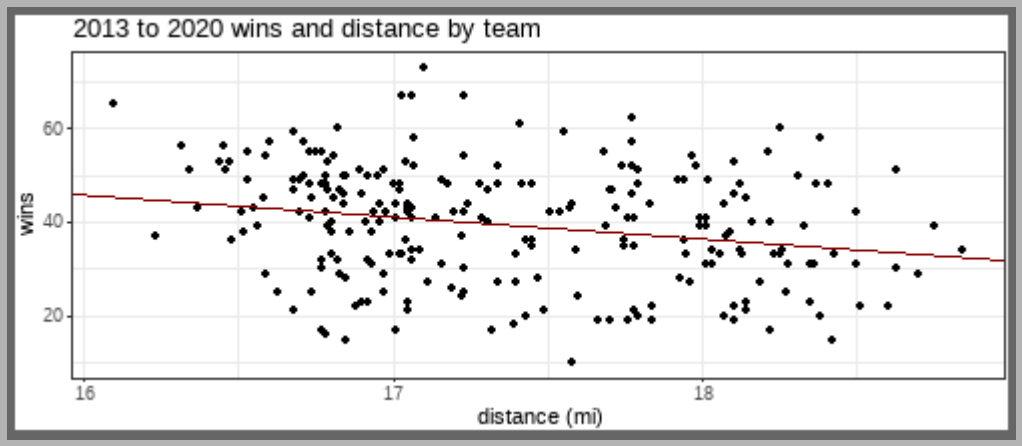
Season Distance By Team

Team	w	l	Distance (mi)	Distance Offense	Distance Defense	Speed (mph)	Speed Offense	Speed Defense
Atlanta Hawks	38	43	16.86	8.99	7.88	4.20	4.55	3.85
Boston Celtics	25	57	16.63	8.92	7.71	4.15	4.41	3.89
Brooklyn Nets	43	38	16.37	8.69	7.69	4.07	4.29	3.85
Charlotte Bobcats	42	39	17.20	9.51	7.69	4.26	4.61	3.89

Data courtesy of nba.com and Second Spectrum

Progress with in-game data

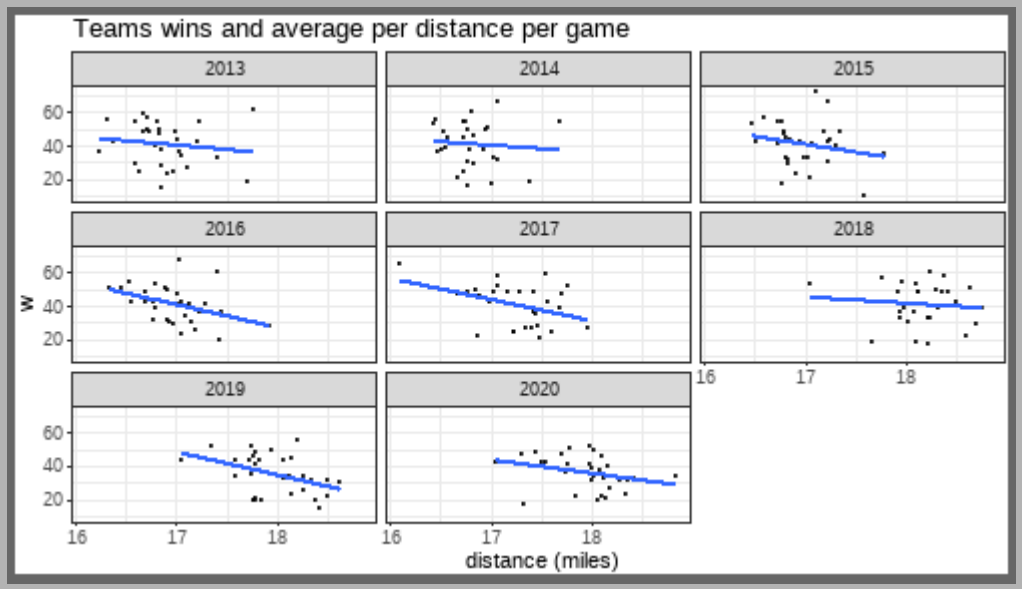
- Inverse correlation between wins and in-game distance



y = 121.6 + -4.73
p-value: .00018 and adj r-squared: .053

Progress with in-game data

- Consistent across all 7 seasons with data

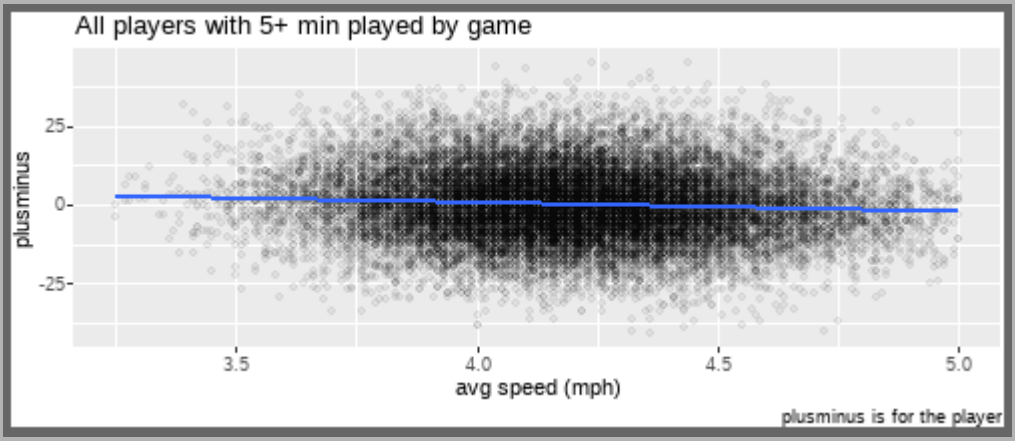


Progress with in-game data

- At team level *not* caused by
  - 3-point FG percentage
  - 2-point FG percentage
  - Turnovers
  - Assists
  - Offensive rebounds
  - Distribution of distance between bench and stars

Progress with in-game data

- Teams are an abstraction
- 15 player summary statistic



plusminus = 11.429 - 2.33(avg_speed)
p-value: <2e-16

Progress with in-game data 2

- Ideas?

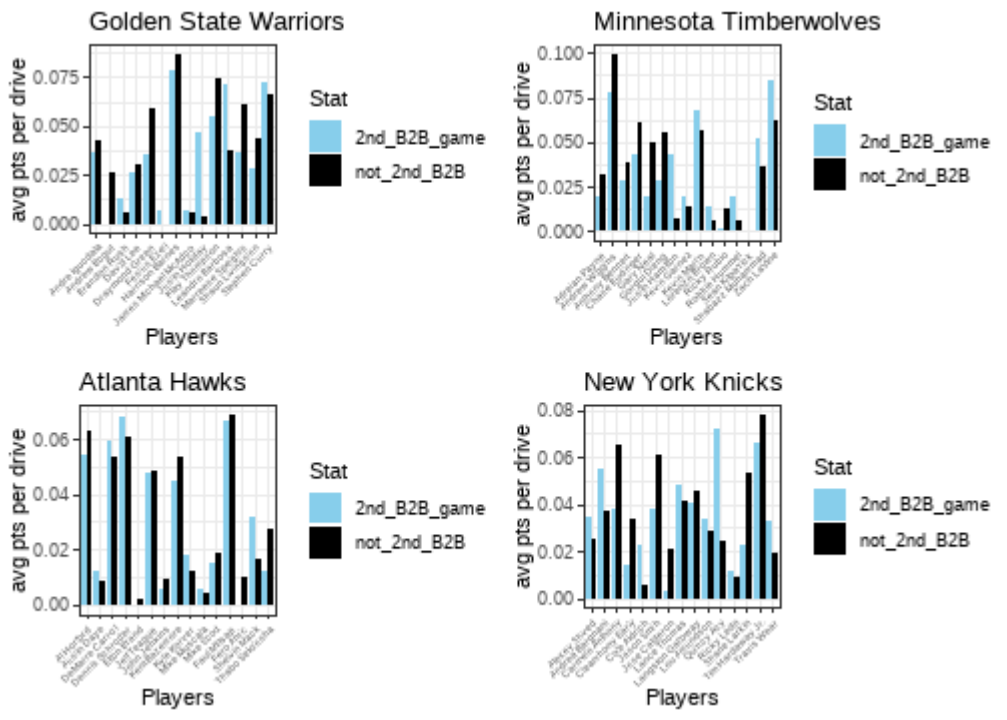
Drives and schedule density data

Drives, Distance, and Schedule Density by Player

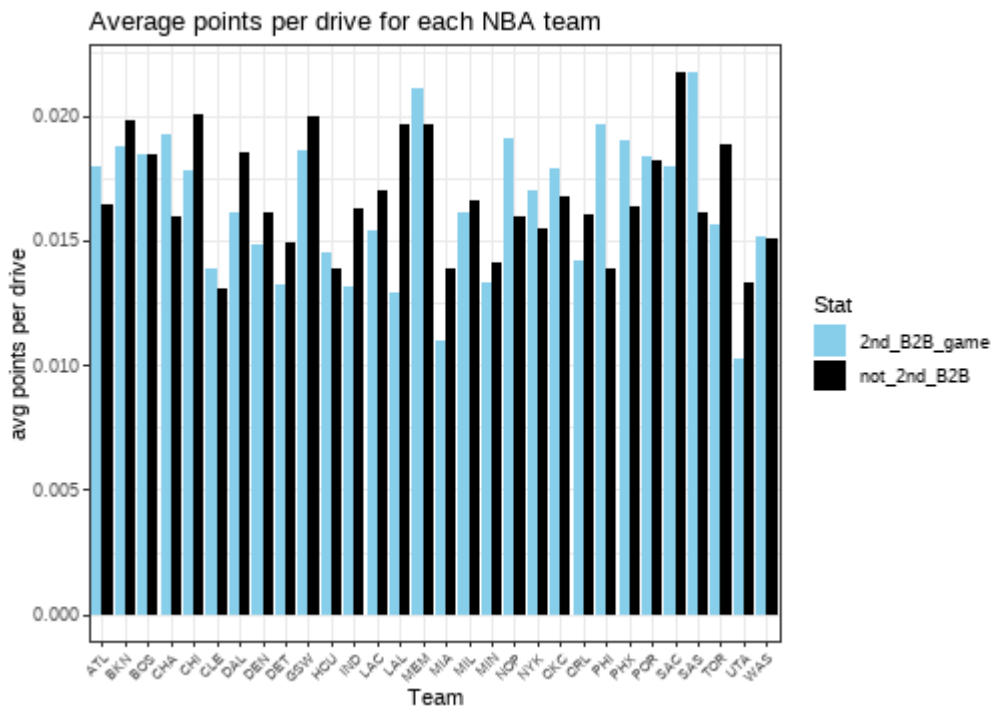
Team	Player	w	l	# of Drives	drive_pts	points per drive	passes per drive	turnovers per drive	Distance (mi)	2nd of a B2B
DAL	Monta Ellis	0	1	14	6	0.429	0.143	0.143	2.68	No
DAL	Chandler Parsons	0	1	6	2	0.333	0.333	0.000	2.44	No
DAL	Tyson Chandler	0	1	0	0	0.000	0.000	0.000	1.70	No
DAL	Devin Harris	0	1	9	5	0.556	0.333	0.000	2.04	No
HOU	James Harden	1	0	7	8	1.143	0.286	0.000	1.93	No
HOU	Trevor Ariza	1	0	2	0	0.000	0.500	0.000	1.84	No

data courtesy of nba.com, Second Spectrum, and the airball r package

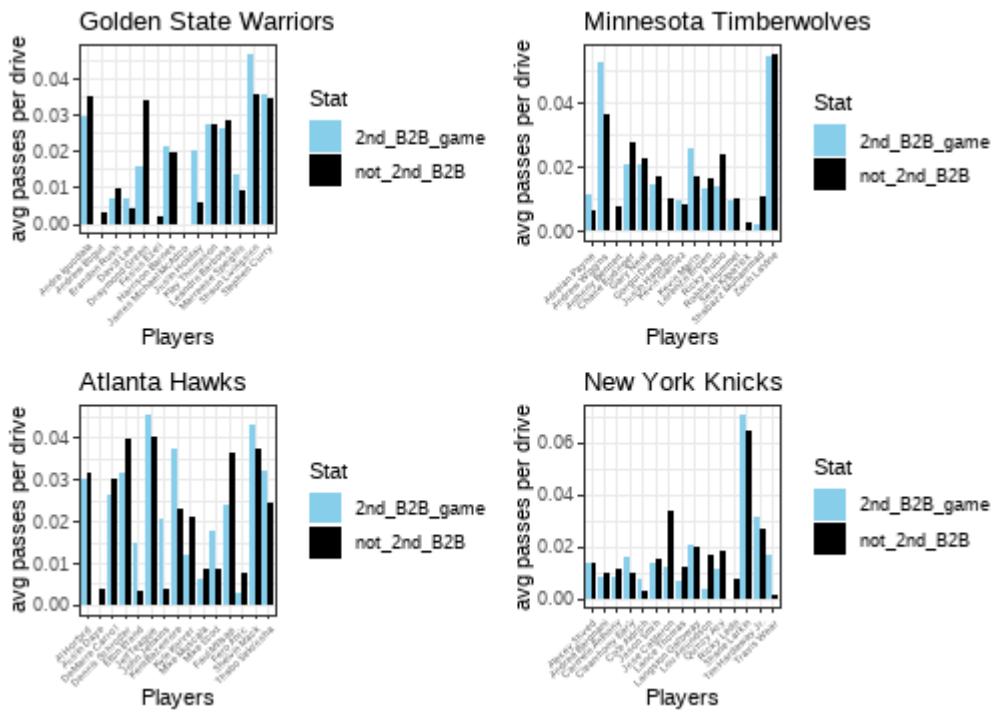
Average Points Per Drive and Back-To-Back Games



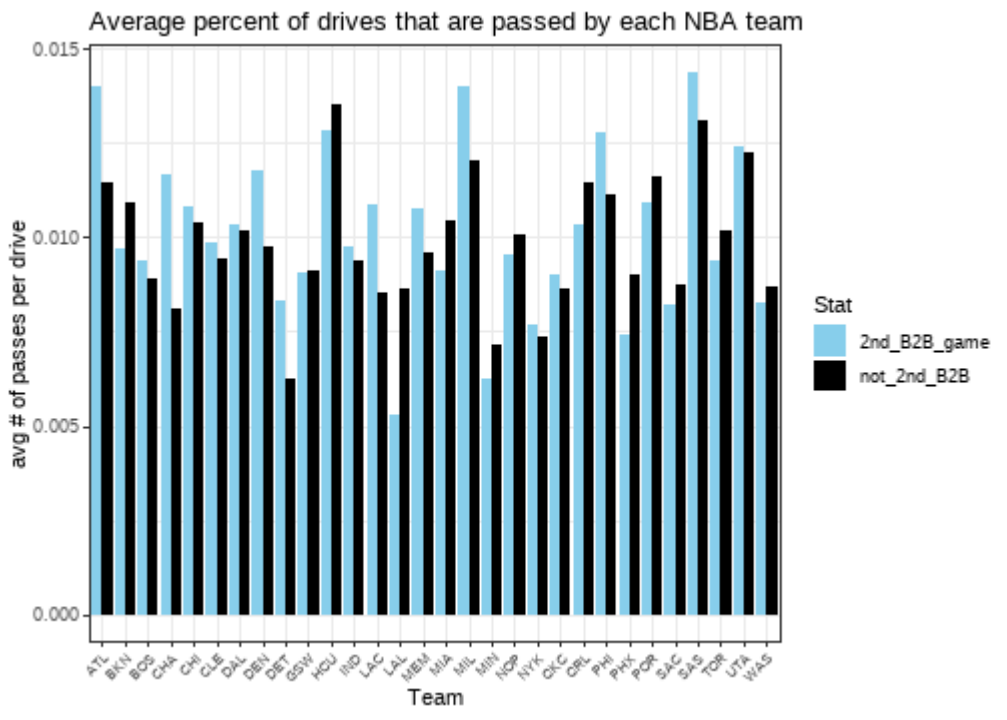
Over half of the NBA teams score less points per drive in the second game of a back-to-back



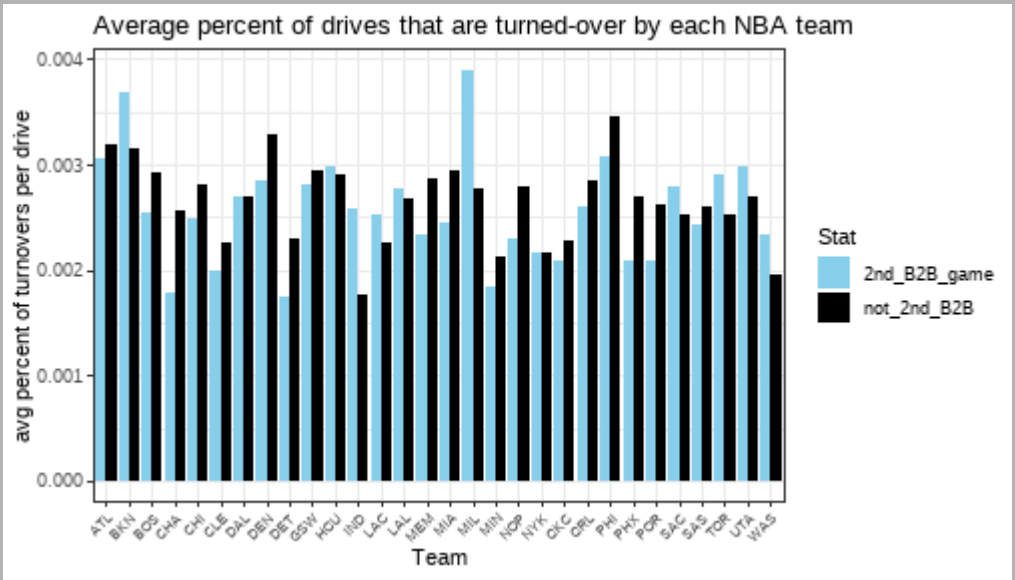
Average number of passes out of drives and back-to-back games



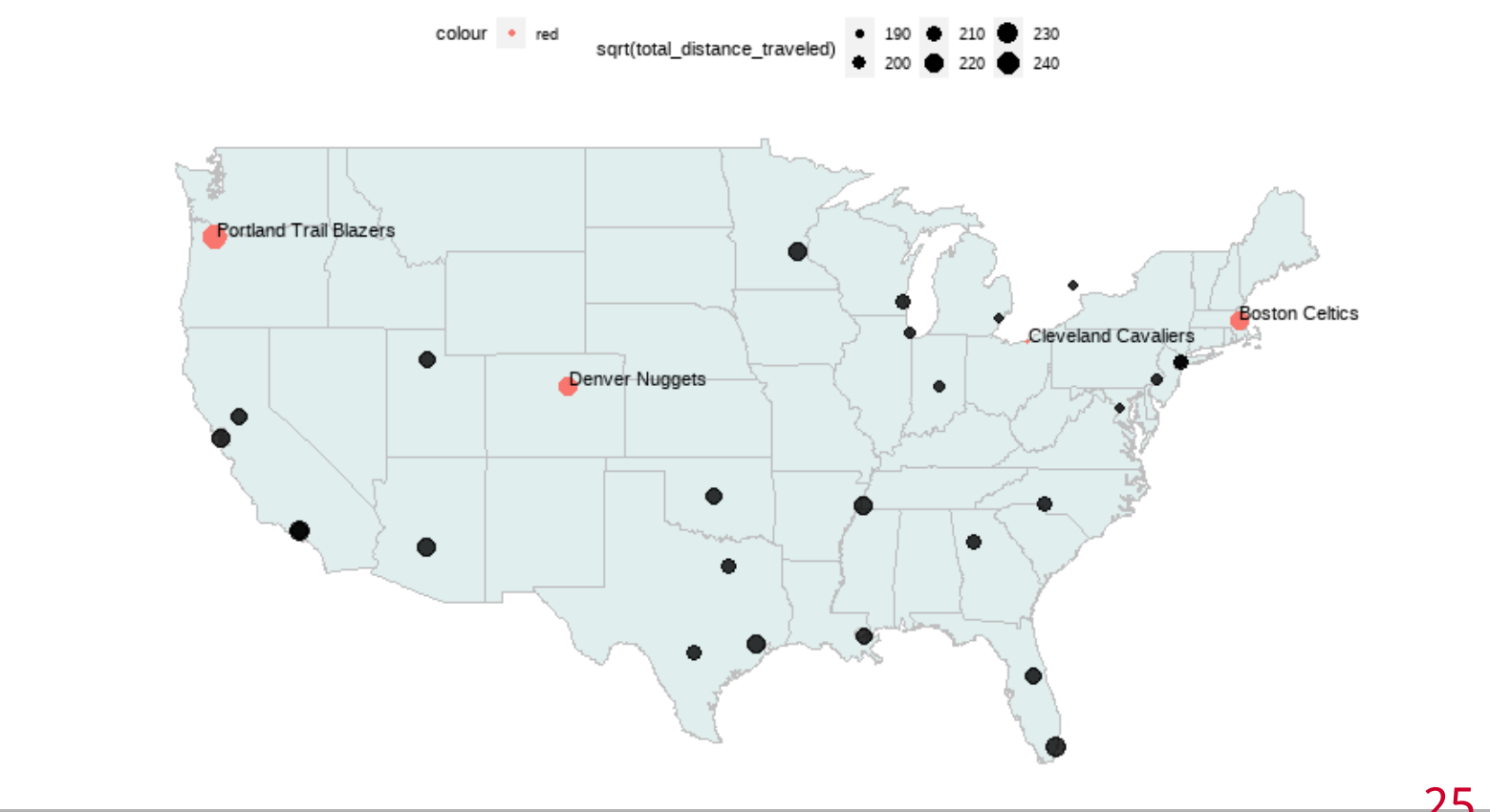
Over half of the NBA teams pass out of drives more in the second game of a back-to-back



Less than half of the NBA teams turnover the ball on drives more in the second game of a back-to-back game



How does travel between games affect team performance?



The Cavaliers also cross far less timezones than the coastal teams

Travel, Schedule and Density Data

Date	Team	Opponent	Distance	Rest	shift	b2b_2nd	traveling_west	win_percent_diff	net_rating_diff	score_diff
2011-01-11	Cleveland Cavaliers	Los Angeles Lakers	366	1	-1	FALSE	TRUE	-50.7	11.9	-55
2014-11-13	Philadelphia 76ers	Dallas Mavericks	1299	3	-1	TRUE	TRUE	-66.7	-12.2	-53
2015-11-02	Memphis Grizzlies	Golden State Warriors	1804	1	-2	TRUE	TRUE	-50.0	-13.0	-50
2018-11-14	Utah Jazz	Dallas Mavericks	419	1	0	FALSE	FALSE	7.1	6.5	-50
2016-03-14	Memphis Grizzlies	Houston Rockets	701	1	-1	FALSE	TRUE	7.5	-2.4	-49
2017-01-22	Los Angeles Lakers	Dallas Mavericks	1249	1	2	FALSE	FALSE	-0.8	-3.7	-49
2017-11-24	Chicago Bulls	Golden State Warriors	600	1	-1	FALSE	TRUE	-56.1	-12.9	-49
2014-11-14	Minnesota Timberwolves	New Orleans Pelicans	1045	1	0	TRUE	FALSE	-37.5	-9.9	-48
2016-03-28	Los Angeles Lakers	Utah Jazz	582	0	1	FALSE	FALSE	-29.7	-11.5	-48
2017-04-10	Orlando Magic	Chicago Bulls	988	1	-1	FALSE	TRUE	-14.8	-6.8	-47
2017-12-04	Washington Wizards	Utah Jazz	1845	2	-2	TRUE	TRUE	-2.0	-4.0	-47
2017-11-15	Sacramento Kings	Atlanta Hawks	544	1	0	FALSE	FALSE	1.4	-1.6	-46
2018-02-12	Phoenix Suns	Golden State Warriors	652	1	-1	FALSE	TRUE	-46.2	-14.9	-46
2010-12-21	Philadelphia 76ers	Chicago Bulls	988	2	-1	TRUE	TRUE	-26.1	-12.0	-45
2014-02-09	Philadelphia 76ers	Los Angeles Clippers	2398	1	-3	TRUE	TRUE	-37.9	-15.9	-45
2016-12-17	Portland Trail Blazers	Golden State Warriors	956	1	-1	FALSE	TRUE	-40.9	-12.0	-45
2018-11-15	Atlanta Hawks	Denver Nuggets	956	1	1	FALSE	FALSE	-46.7	-9.5	-45

- **Granularity:** Each row represents a game. The stats are relative to the visiting team in the *Team* column

Linear Model: Distance doesn't have a significant effect on the score diff

score_diff = β_0 + β_1 win_percent_diff + $\beta_2$distance + $\beta_3$rest + β_4b2b_2nd + $\beta_5$three_in_four + β's for each shift

```
Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.5590778  0.3128822  -8.179 3.22e-16 ***
win_percent_diff  0.2219714  0.0042123  52.696  < 2e-16 ***
Distance          0.0003376  0.0004250   0.794   0.4270
Rest              0.4383415  0.0714643   6.134 8.92e-10 ***
hours_shift-3    -3.1328084  1.3759289  -2.277   0.0228 *
hours_shift-2    -0.6153907  0.8174023  -0.753   0.4516
hours_shift-1    -0.2674273  0.3539147  -0.756   0.4499
hours_shift1     -0.4271058  0.3486236  -1.225   0.2206
hours_shift2     -0.3637441  0.7936980  -0.458   0.6468
hours_shift3     -1.1895248  1.3829682  -0.860   0.3897
three_in_fourTRUE -1.1371529  0.2792383  -4.072 4.69e-05 ***
b2b_2ndTRUE      -0.5111220  0.2829811  -1.806   0.0709 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.81 on 9576 degrees of freedom
Multiple R-squared:  0.2275,    Adjusted R-squared:  0.2266
F-statistic: 256.4 on 11 and 9576 DF,  p-value: < 2.2e-16
```

What's on deck

- Find stronger proxies to account for difference in team vs opponent strength
- Logistic regression model predicting win or loss instead of score difference
- Test the models on Bubble and 2021 data (where there was none of significantly less travel/shift between timezones)