Modeling NBA Career Length

Austin Miles, Nathan Hwangbo, Wentao Liao
Group 19

Research Questions

- How long do NBA Careers last?
- To what degree does the skill (or perceived skill) of the player impact career length?
- To what degree does the position of the player impact career length?
- Has career length in the league changed over time?

Data

- Data comes from basketball-reference.com
- Included the following draft classes:
  - 1998 - Vince Carter
  - 2003 - Lebron James, Carmelo Anthony, Dwayne Wade
  - 2009 - Steph Curry, James Harden
- Using draft round number (1 or 2) as a proxy for skill
- Consolidated positions into Smalls (guards) and Bigs (everyone else)

```
  Year       Rd          Pos             Yrs
1998:56    1:88      Big  :89     Min.   : 1.000
2003:54    2:77      Small:76     1st Qu.: 3.000
2009:55                           Median : 8.000
                                  Mean   : 7.727
                                  3rd Qu.:11.000
                                  Max.   :22.000
```
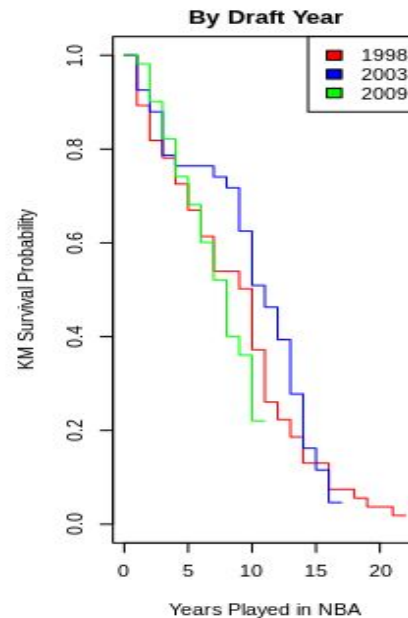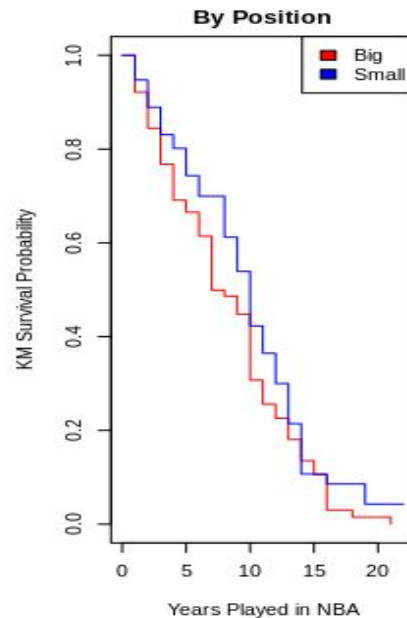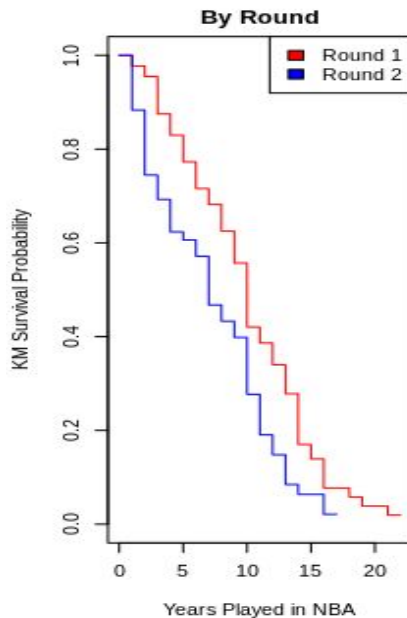
Distribution Of Career Lengths
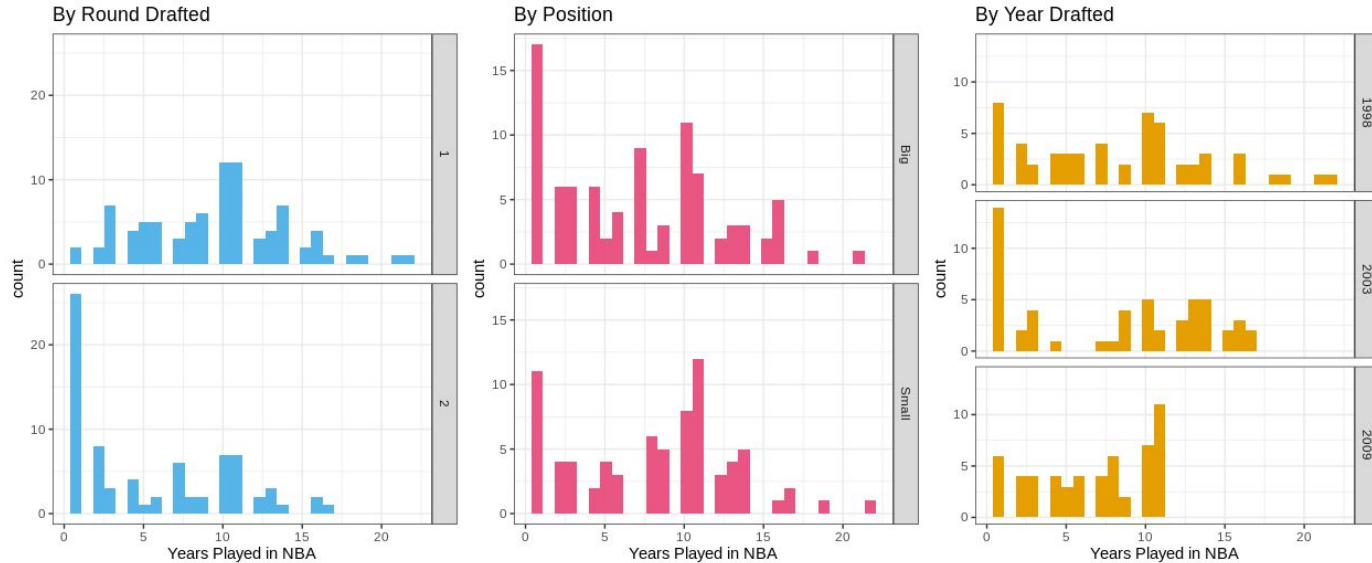
KM Survival Curves

Log-rank Results:

- Rd: p = 0.002
- Pos: p = 0.1
- Year: p = 0.1

Cox PH

- Round is the only significant variable in the model
- 2nd rounders have ~73% hazard increase

```
Call:
coxph(formula = Surv(Yrs, event) ~ Pos + Year + Rd, data = nba_draft)

               coef exp(coef) se(coef)       z      p
PosSmall   -0.2805    0.7554   0.1783 -1.573 0.1157
Year2003   -0.1645    0.8483   0.2126 -0.774 0.4391
Year2009    0.3453    1.4124   0.2279  1.515 0.1297
Rd2         0.5501    1.7334   0.1788  3.076 0.0021

Likelihood ratio test=15.6  on 4 df, p=0.00361
n= 165, number of events= 134
```

Forward and Backward Selection

```
Start:  AIC=1104.59
Surv(Yrs, event) ~ 1

        Df   AIC
+ Rd     1 1097.9
+ Pos    1 1104.3
<none>     1104.6
+ Year   2 1104.8

Step:  AIC=1097.88
Surv(Yrs, event) ~ Rd

        Df   AIC
+ Year   2 1097.5
+ Pos    1 1097.6
<none>     1097.9

Step:  AIC=1097.49
Surv(Yrs, event) ~ Rd + Year

        Df   AIC
+ Pos    1 1097.0
<none>     1097.5

Step:  AIC=1097
Surv(Yrs, event) ~ Rd + Year + Pos

Call:
coxph(formula = Surv(Yrs, event) ~ Rd + Year + Pos, data = nba_draft)

            coef exp(coef) se(coef)     z        p
Rd2        0.550    1.733    0.179  3.08 0.0021
Year2003  -0.165    0.848    0.213 -0.77 0.4391
Year2009   0.345    1.412    0.228  1.52 0.1297
PosSmall  -0.280    0.755    0.178 -1.57 0.1157

Likelihood ratio test=15.6  on 4 df, p=0.004
n= 165, number of events= 134
```

```
Start:  AIC=1097
Surv(Yrs, event) ~ Rd + Pos + Year

        Df   AIC
<none>     1097.0
- Pos    1 1097.5
- Year   2 1097.6
- Rd     1 1104.2
Call:
coxph(formula = Surv(Yrs, event) ~ Rd + Pos + Year, data = nba_draft)

            coef exp(coef) se(coef)     z        p
Rd2        0.550    1.733    0.179  3.08 0.0021
PosSmall  -0.280    0.755    0.178 -1.57 0.1157
Year2003  -0.165    0.848    0.213 -0.77 0.4391
Year2009   0.345    1.412    0.228  1.52 0.1297

Likelihood ratio test=15.6  on 4 df, p=0.004
n= 165, number of events= 134
```
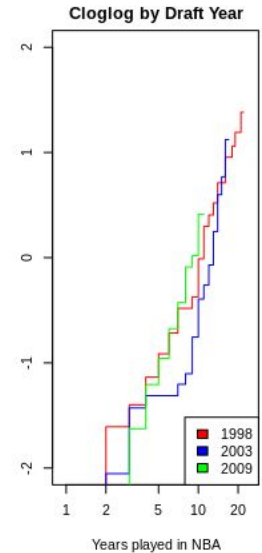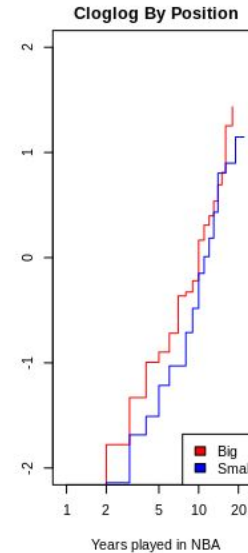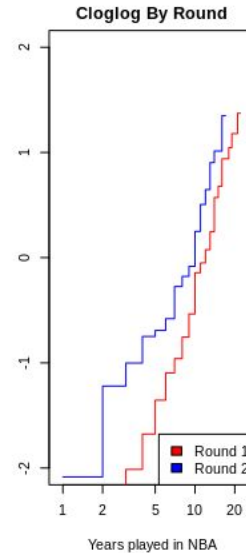
Validating PH assumption

- Slight crossing in Position and Draft Year
- Schoenfeld residual tests indicate that all three satisfy proportional hazard.

Key Takeaways

- Players drafted in the first round tend to have longer careers
- We lack evidence to conclude that draft year has a significant effect on career length, at least for the years 1998, 2003, and 2009
- We lack sufficient evidence to conclude that Position has a significant effect on career length