STA303 NBA Injury

David Wong

3/20/2021

```
glimpse(cleaning_injury_types)
```

```
## Rows: 8,378
## Columns: 8
## $ Date         <dttm> 2010-10-12, 2010-10-26, 2010-10-27, 2010-10-27, 2010-...
## $ Team         <chr> "Bucks", "Blazers", "Pistons", "Pistons", "Blazers", "...
## $ Relinquished <chr> "Darington Hobson", "Greg Oden", "Jonas Jerebko", "Ter...
## $ Notes        <chr> "surgery on left hip (out for season)", "placed on IL ...
## $ Status       <chr> "Season Ending", "Season Ending", "Season Ending", "Se...
## $ Rank         <dbl> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, ...
## $ type         <chr> "out for season", "out for season", "out for season", ...
## $ type_num     <dbl> 5, 1, 7, 4, 1, 1, 1, 4, 4, 4, 1, 1, 7, 1, 7, 1, 4, 4, ...
```

Looking at the cleaned data set, we have reclassified the relevant fields into dummy variable representations.

type	counts
DNP	4089
DTD	2722
out for season	985
out indefinitely	582

type	type_num	n	weight	weighted_num
DNP	1	971	0.2374664	230.5798484
DNP	2	224	0.0547811	12.2709709
DNP	3	336	0.0821717	27.6096845
DNP	4	1143	0.2795304	319.5033015
DNP	5	251	0.0613842	15.4074346
DNP	6	253	0.0618733	15.6539496
DNP	7	647	0.1582294	102.3744192
DNP	8	178	0.0435314	7.7485938
DNP	9	86	0.0210320	1.8087552
DTD	1	534	0.1961793	104.7597355
DTD	2	149	0.0547392	8.1561352
DTD	3	256	0.0940485	24.0764144
DTD	4	760	0.2792065	212.1969140
DTD	5	217	0.0797208	17.2994122
DTD	6	157	0.0576782	9.0554739
DTD	7	450	0.1653196	74.3938281

type	type_num	n	weight	weighted_num
DTD	8	156	0.0573108	8.9404849
DTD	9	43	0.0157972	0.6792799
out for season	1	308	0.3126904	96.3086294
out for season	2	63	0.0639594	4.0294416
out for season	3	41	0.0416244	1.7065990
out for season	4	217	0.2203046	47.8060914
out for season	5	59	0.0598985	3.5340102
out for season	6	89	0.0903553	8.0416244
out for season	7	151	0.1532995	23.1482234
out for season	8	30	0.0304569	0.9137056
out for season	9	27	0.0274112	0.7401015
out indefinitely	1	168	0.2886598	48.4948454
out indefinitely	2	115	0.1975945	22.7233677
out indefinitely	3	16	0.0274914	0.4398625
out indefinitely	4	116	0.1993127	23.1202749
out indefinitely	5	23	0.0395189	0.9089347
out indefinitely	6	48	0.0824742	3.9587629
out indefinitely	7	52	0.0893471	4.6460481
out indefinitely	8	21	0.0360825	0.7577320
out indefinitely	9	23	0.0395189	0.9089347

From looking at our distribution of types of injuries: DTP (Did not play), DTD (Day to day), Out indefinitely, Out for season, these were labeled 1 to 4 respectively from least severe to most severe. Furthermore, our injury type classifications ranged from numbers 1 to 9 where the various types of injuries were classified under the localized area they affect:

knee - 1

hand - 2

back - 3

foot - 4

hip - 5

arm - 6

leg - 7

head - 8

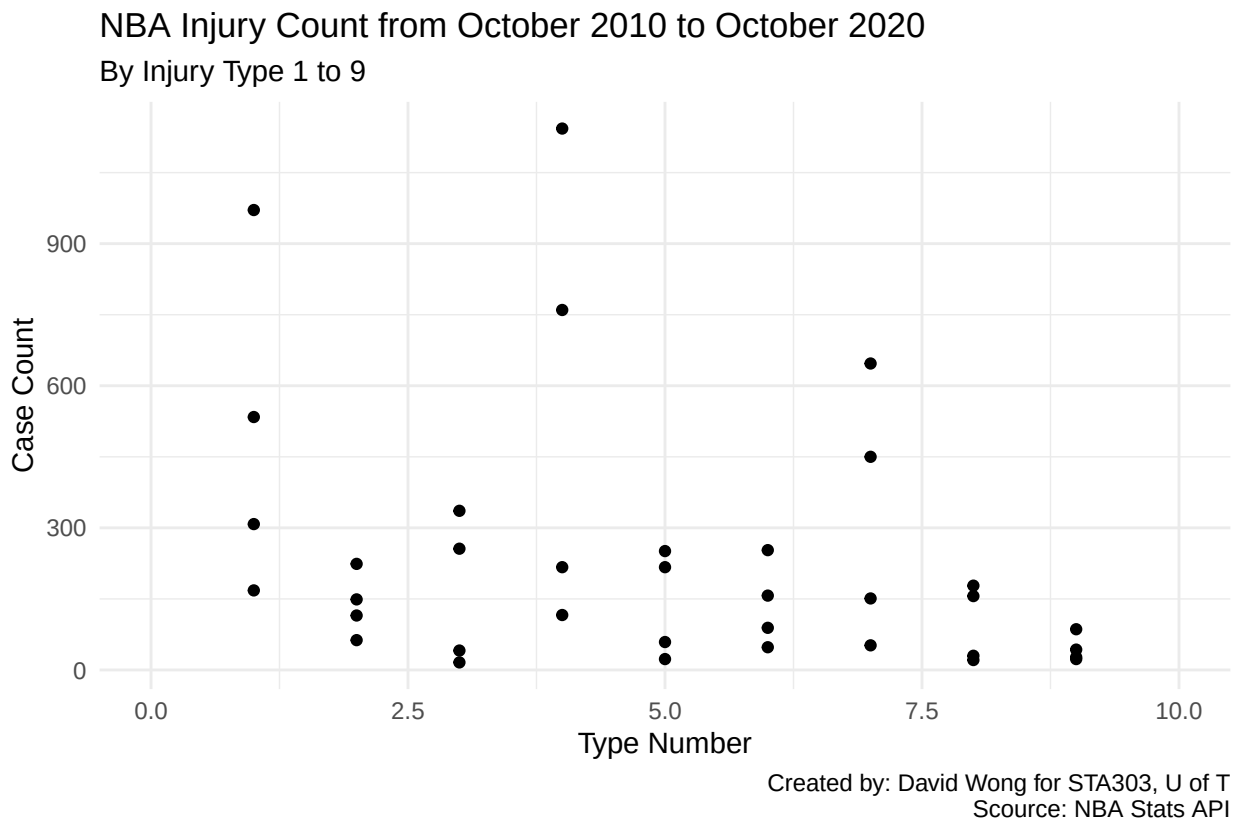
torso - 9

Thus, we can see that for DNP: type 4 (foot), type 1 (knee), and type 7 (leg) injuries are the most common and have had players sit out for a game.

Similarly, for DTD: type 4 (foot), type 1 (knee), and type 7 (leg) injuries are also the most common that had players sit out for multiple games.

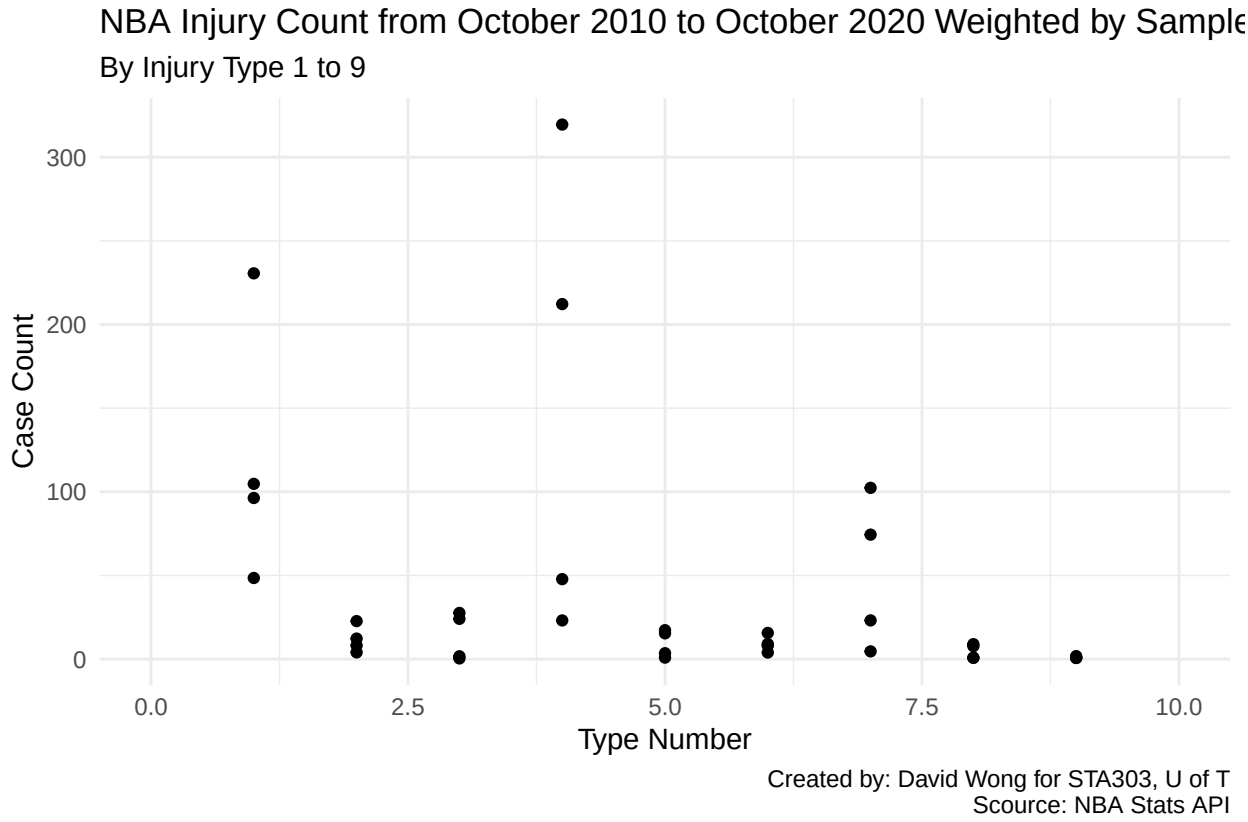
For injury type to be severe, I assume that any injury rendering a player to be out for the season (4) or indefinitely (3) would be severe. For out indefinitely: type 4 (foot), type 1 (knee) and type 2 (hand) injuries cause players to be out indefinitely the most frequent.

Lastly, for out for season: type 4 (foot), type 1 (knee), and type 7 (leg) injuries are the most common for season ending incidences for NBA players.



Looking at this from a different perspective, we can map out the frequencies of cases of each injury type (1 to 9) where it seems type 1, type 4, and type 7 are by far the most frequent injuries in the NBA.

```
table_injury_tally_weight %>%
  ggplot(aes(x = type_num, y = weighted_num)) +
  geom_point() +
  theme_minimal() +
  labs(title = "NBA Injury Count from October 2010 to October 2020 Weighted by Sample Proportion",
       subtitle = "By Injury Type 1 to 9",
       x = "Type Number",
       y = "Case Count",
       caption = str_c("Created by: David Wong for STA303, U of T\n Scource: NBA Stats API\n"
                       )) +
  scale_fill_manual(values = c("#86BCB6", "#B9CA5D")) +
  theme(legend.title = element_blank(), legend.position = c(0.15,0.8)) +
  scale_x_continuous(limits = c(0, 10))
```



To further examine the distribution of points, we weighted each injury type by its population proportions multiplied by the counts of the injury type occurrence. This shows the relative disparity between the frequencies of incidents better.

```
## Rows: 1,901
## Columns: 3
## Groups: Relinquished [898]
## $ Relinquished <chr> "A.J. Price", "Aaron Brooks", "Aaron Brooks", "Aaron G...
## $ Rank         <dbl> 1, 1, 2, 2, 3, 4, 2, 4, 2, 2, 4, 1, 2, 3, 1, 2, 4, 1, ...
## $ n            <int> 15, 2, 4, 11, 1, 1, 1, 2, 1, 2, 2, 7, 2, 2, 4, 10, 4, ...

## Rows: 776
## Columns: 3
## Groups: Relinquished [584]
## $ Relinquished <chr> "Aaron Gordon", "Aaron Gordon", "Acie Law", "Al-Farouq...
## $ Rank         <dbl> 3, 4, 4, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 3, 4, 4, 4, ...
## $ n            <int> 1, 1, 2, 2, 2, 4, 4, 2, 1, 2, 3, 4, 2, 3, 3, 3, 1, 2, ...
```

Now let's look at the jury type counts per NBA player during the past decade. In the first table, you can see that there are 1901 NBA players that have had various types of injuries within the past decade. Furthermore, in the following table, if we filter out the most severe injuries (rank 3 and 4 being out indefinitely and season ending respectively), 776 incidences among NBA players have had a severe injury during the past decade.

```
## Rows: 584
## Columns: 2
## $ Relinquished <chr> "Aaron Gordon", "Acie Law", "Al-Farouq Aminu", "Al Har...
## $ counts       <int> 2, 2, 2, 2, 4, 6, 3, 7, 5, 6, 1, 2, 4, 1, 1, 2, 2, 5, ...
```

Now if we look at this on a unique basis and count up all of the total severe injuries (rank 3 or 4) per NBA player, we can see how many severe injuries each NBA player has accumulated within the past decade. There are 584 unique NBA players that have had a severe injury whereas there are 776 occurrences of severe injuries happening among the 584 players. Now where is LeBron James?

Relinquished	counts
LeBron James	2

LeBron James has had a total of only 2 severe injuries in the past decade; and this was in 2019 only.

Date	Team	Relinquished	Notes	Status	Rank	type	type_num
2019-03-30	Lakers	LeBron James	left groin injury (out for season)	Season Ending	4	out for season	5
2019-03-31	Lakers	LeBron James	placed on IL with left groin injury (out for season)	Season Ending	4	out for season	5
2014-02-23	Heat	LeBron James	fractured nose (DTD)	DTD	2	DTD	8
2014-12-11	Cavaliers	LeBron James	sore left knee (DTD)	DTD	2	DTD	1
2015-01-28	Cavaliers	LeBron James	sprained right wrist (DTD)	DTD	2	DTD	2
2015-02-27	Cavaliers	LeBron James	back injury (DTD)	DTD	2	DTD	3
2018-12-26	Lakers	LeBron James	strained left groin (DTD)	DTD	2	DTD	5
2020-02-26	Lakers	LeBron James	left groin injury (DTD)	DTD	2	DTD	5
2011-01-13	Heat	LeBron James	sprained left ankle (DNP)	DNP	1	DNP	4
2012-01-07	Heat	LeBron James	sprained left ankle (DNP)	DNP	1	DNP	4
2013-03-31	Heat	LeBron James	strained right hamstring (DNP)	DNP	1	DNP	7
2013-04-02	Heat	LeBron James	strained right hamstring (DNP)	DNP	1	DNP	7
2013-04-05	Heat	LeBron James	strained right hamstring (DNP)	DNP	1	DNP	7
2013-04-10	Heat	LeBron James	strained right hamstring (DNP)	DNP	1	DNP	7
2013-04-15	Heat	LeBron James	strained right hamstring (DNP)	DNP	1	DNP	7

Taking a look at LeBron's total injury history in the past decade, there are only 15 incidences in the past decade where LeBron has sustained any form of injury, where 13/15 is not severe and 2/15 are severe.

Relinquished	counts
Blake Griffin	16
Derrick Rose	16
Danilo Gallinari	15
Kyrie Irving	13
Rajon Rondo	12

Relinquished	counts
Dante Exum	11
Joel Embiid	11
Chandler Parsons	10
Michael Carter-Williams	10
Cody Zeller	9
Joakim Noah	9
Kevin Love	9
Kobe Bryant	9
Nikola Pekovic	9
Tyreke Evans	9
Anthony Davis	8
Avery Bradley	8
Brandon Knight	8
Chris Paul	8
DeMarcus Cousins	8
Gordon Hayward	8
Jabari Parker	8
Jerryd Bayless	8
John Wall	8
Kevin Durant	8
Kris Dunn	8
Kristaps Porzingis	8
Patrick Beverley	8
Alec Burks	7
Andrew Bogut	7
Brook Lopez	7
Carlos Delfino	7
Darrell Arthur	7
Enes Kanter	7
Eric Bledsoe	7
Festus Ezeli	7
Iman Shumpert	7
Leandro Barbosa	7
Quincy Pondexter	7
Rudy Gay	7
Al Jefferson	6
Alex Len	6
Ben Simmons	6
Brendan Haywood	6
David Lee	6
Ed Davis	6
Elliot Williams	6
Eric Gordon	6
Greg Oden	6
Kevon Looney	6
Lance Thomas	6
Martell Webster	6
Nene Hilario	6
Pau Gasol	6
Stephen Curry	6
Zach LaVine	6
Alex Ajinca	5

Relinquished	counts
Anderson Varejao	5
Andre Roberson	5
Andrea Bargnani	5
Brandan Wright	5
C.J. Miles	5
Chauncey Billups	5
Danny Granger	5
Denzel Valentine	5
Deron Williams	5
Devin Booker	5
Emmanuel Mudiay	5
Glen Davis	5
Glenn Robinson III	5
Hamidou Diallo	5
Isaiah Thomas	5
Jeremy Lin	5
Jonathan Isaac	5
Josh McRoberts	5
Jrue Holiday	5
Justise Winslow	5
Kawhi Leonard	5
Kent Bazemore	5
Lou Williams	5
Luol Deng	5
Marc Gasol	5
Michael Kidd-Gilchrist	5
Mike Conley Jr.	5
Nerlens Noel	5
Ramon Sessions	5
Russell Westbrook	5
Seth Curry	5
T.J. Warren	5
Thabo Sefolosha	5
Wilson Chandler	5
Xavier Henry	5

Next, we are going to filter out all of the NBA players that have had more than 5 occurrences of severe injuries happening in the past decade. We chose 5 people in a 10 year period, that would mean this player spent half of the decade essentially doing some form of long-term rehabilitation for their injuries. Unsurprisingly, Blake Griffin and Derek Rose place at the top of this list. There are 92 players with 5 or more severe injuries during the past decade in the NBA.

Given the probability to get an injury of rank 3 or 4 is 584/8478 and that of the 584 rank injuries, 92 players have had severe injuries at least 5 times within the past 10 years, how big of an outlier is Lebron James who only received an injury of rank 3 or 4 twice in the past decade?

```
##
## Call:
## glm(formula = ID ~ Rank + n, family = "poisson", data = filtered_player_injury_type)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
```

```
## -24.6266   -9.3041    0.2846    7.6356   18.5252
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.629951   0.015106 372.705   <2e-16 ***
## Rank         0.039266   0.004370   8.984   <2e-16 ***
## n           -0.050535   0.001585 -31.879   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 87956  on 775  degrees of freedom
## Residual deviance: 86889  on 773  degrees of freedom
## AIC: 92481
##
## Number of Fisher Scoring iterations: 5
```

Firstly, let's determine whether or not injury severity even is related to the type of player receiving them. We mapped a unique ID for each NBA player regardless of the injury type, only the severity of the injury (rank). I applied a Poisson Generalized Linear Model, aka a log-linear model, to understand the counts of cases (n) and severity of injury (rank) is related to the NBA player. We can see that all of the coefficients are statistically significant and thus the estimators have strong correlation to the NBA player (ID).

```
##                Estimate  Robust SE    Pr(>|z|)          LL          UL
## (Intercept)  5.62995134 0.14797568 0.000000000  5.33991900  5.91998367
## Rank         0.03926605 0.04294567 0.360549150 -0.04490746  0.12343956
## n           -0.05053542 0.01545721 0.001077854 -0.08083155 -0.02023929
```

Looking at the regression model outputs, these are the relevant summary statistics from the Poisson Generalized Linear Model.

```
##      res.deviance  df p
## [1,]     86889.33 773 0
```

With a chi-squared test, we are able to assess the significance of our results. From this, we can see that we have a significant p-value from our model output.

Now that we can conclude that our variables are statistically significant and show strong correlation between NBa player and their injury severity and frequency, we can ask ourselves, "How often do you see a LeBron James?".

```
## Rows: 367
## Columns: 2
## $ Relinquished <chr> "Aaron Gordon", "Acie Law", "Al-Farouq Aminu", "Al Har...
## $ counts       <int> 2, 2, 2, 2, 1, 2, 1, 1, 2, 2, 1, 1, 1, 2, 2, 1, 1, 2, ...
```

Of the total of 584 players in the NBa during the past decade, 367 have had 2 or less severe injuries. However, this is likely because of players entering the NBA much earlier than LeBron as well as players nearing the end of their careers. Let's narrow the scope down to players who have had an all-star appearances at least 5 times. To achieve a level where you're selected for all-star awards means recognition from all levels of the game from spectators to players themselves. The fan vote accounts for 50 percent of the total vote to determine the starters for the NBA All-Star Game. Current players and media members split the other 50 percent.

```
## Rows: 157
## Columns: 5
## $ Rk           <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,...
## $ Relinquished <chr> "Kareem Abdul-Jabbar", "Kobe Bryant", "LeBron James", ...
## $ Tot          <dbl> 19, 18, 17, 16, 15, 15, 15, 14, 14, 14, 14, 13, 13, 13...
## $ NBA          <dbl> 19, 18, 17, 11, 15, 15, 15, 14, 14, 14, 14, 13, 13, 13...
## $ ABA          <dbl> 0, 0, 0, 5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 4, 0, ...
```

This is a list of all-star selections of all time where players have been selected at least 5 times for either all-star or all-aba games (before 1976).

Relinquished	counts	Rk	Tot	NBA	ABA
LeBron James	2	3	17	17	0
Tim Duncan	1	5	15	15	0
Dwyane Wade	2	16	13	13	0
Chris Bosh	2	29	11	11	0
Ray Allen	1	36	10	10	0
Jason Kidd	1	41	10	10	0
James Harden	1	45	9	9	0
Yao Ming	2	59	8	8	0
Paul George	2	70	7	7	0
Joe Johnson	1	74	7	7	0
Damian Lillard	1	104	6	6	0
Jermaine O'Neal	2	108	6	6	0
Amar'e Stoudemire	2	111	6	6	0
Klay Thompson	2	151	5	5	0

Looking at all NBA players who have been selected for at least 5 all-star games and filtering for those who have played during the past decade, this is what we get. 'counts' tells you how many severe injuries they've sustained within the past decade, 'Rk' represents their ranking of total all-star selections where 'NBA' and 'ABA' tells you what basketball league they achieved this in. In the table above, we are ranking it by total all-star appearances with LeBron James at the top with 17.

Thus, we can conclude that out of all NBA players who have achieved great success (at least 5 all star selections), only 14 players in total can boast similar achievements while having experienced less than 2 severe injuries in the past decade. Considering that LeBron is in year 18 of the NBA, he has been an all-star for literally 17 of the 18 (94.4%) seasons he has been playing professionally. As a close second, Tim Duncan was in the NBA for 19 seasons, 15 of which he was an all-star as well.

So we get it, LeBron James is statistically someone who has very frequently achieved a high level of success in the NBA while being widely admired from NBA player and spectators alike. With a career as long as his, an severe injury count of only 2 in a current 18 and on-going year journey, no one really is close to LeBron when considering all factors of longevity, success, and physical invincibility.

Debates of the Greatest of All Time (GOAT) has always been between players like LeBron, Duncan, Kobe, Jordan, Olajuwon etc... but statistically, LeBron is a prime specimen of what it means to be invincible.