**Pineapple Project-An Exploration of Injuries in the NBA 2010-2018**

**Mark Bickerton, Justin Hobbs, Mychal Spivey, Katy Webb**

**Background**

NBA players can be very expensive (e.g. Kevin Durant’s 2017 contract worth ~$27 million), but injuries suffered on the court can have a huge impact on careers and teams. Using our Kevin Durant example, who missed the playoff finals and is expected to be sideline all next season, can we figure out which players may be more susceptible to being injured, and if so, can we use that information to guide the NBA?

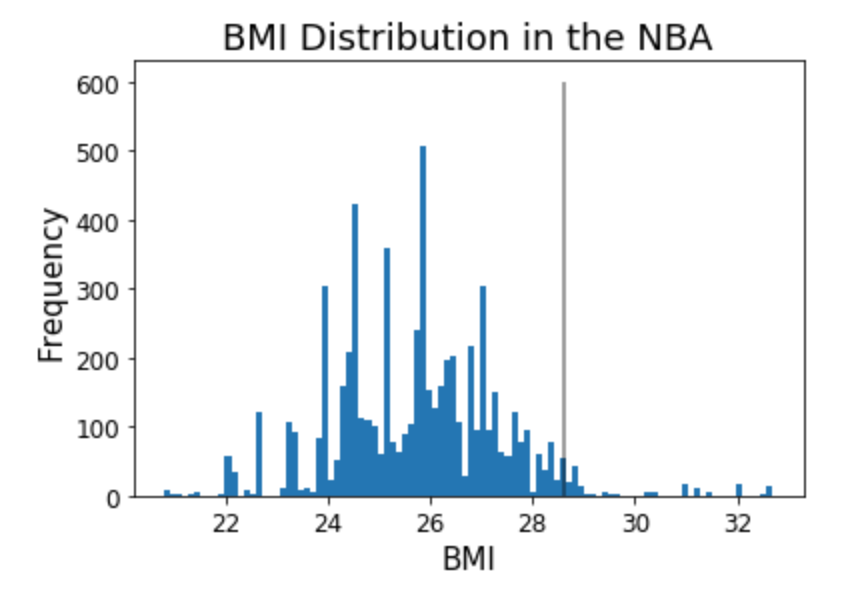
Using Kaggel as our data source, we pieced together two data sets (1[-NBA Injuries 2010-2018](https://www.kaggle.com/ghopkins/nba-injuries-2010-2018/kernels), and 2) [NBA Player Stats Since 1950](https://www.kaggle.com/drgilermo/nba-players-stats)) to get a look at player injuries between 2010-2018, combined with player stats including: height, weight, DOB, date of injury, injury notes, and college.

With this information we wanted to answer the following questions:

1. What are the most common injuries in the NBA?
2. Do certain player characteristics lend themselves to a greater frequency of being injured?
   1. Player age-Are younger players more rash or aggressive, and therefore injured more often? Or maybe older players who have years of injuries and rough game play more likely to be injured?
   2. Size-NBA players overall are tall, athletic guys, but some are still going ot be smaller and larger than others. Is there an effect of size as measured by BMI, or frequency in injury? Are smaller players more likely to get injured?
   3. Position-Some positions seem to be more aggressive than other, e.g. Forwards and Centers; does a more active position translate to more injuries, or are players whose positions place them on the receiving end of full-court push positions
   4. Collegiate training-Different college programs train their players differently and use different on-court strategies. Does this have an impact on the frequency of injuries?

**Findings**

Looking first at BMI as an indicator of player size, we see that BMI is distributed over a wide range, with a mean BMI of 25.7. The gray line shows the average US male’s BMI, just as a point of informal comparison.

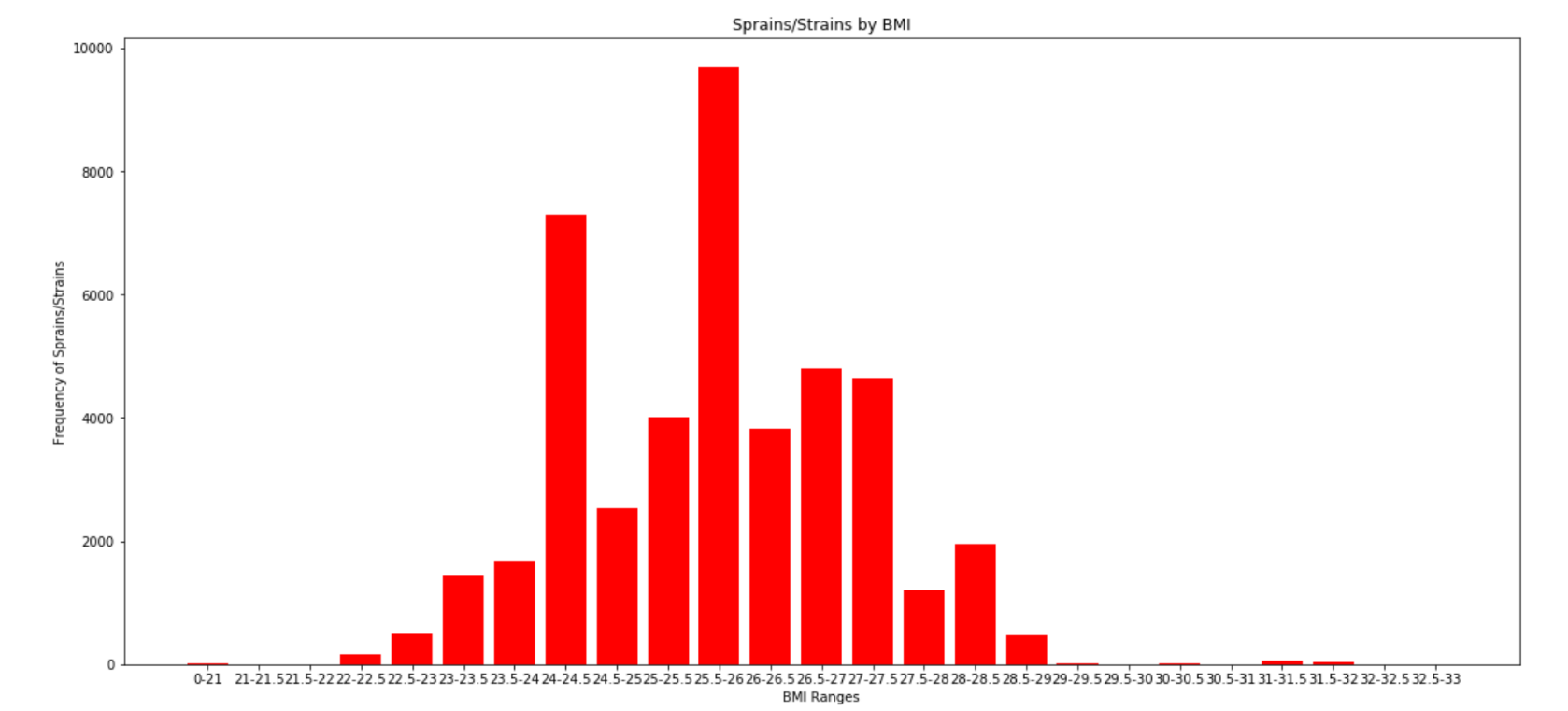


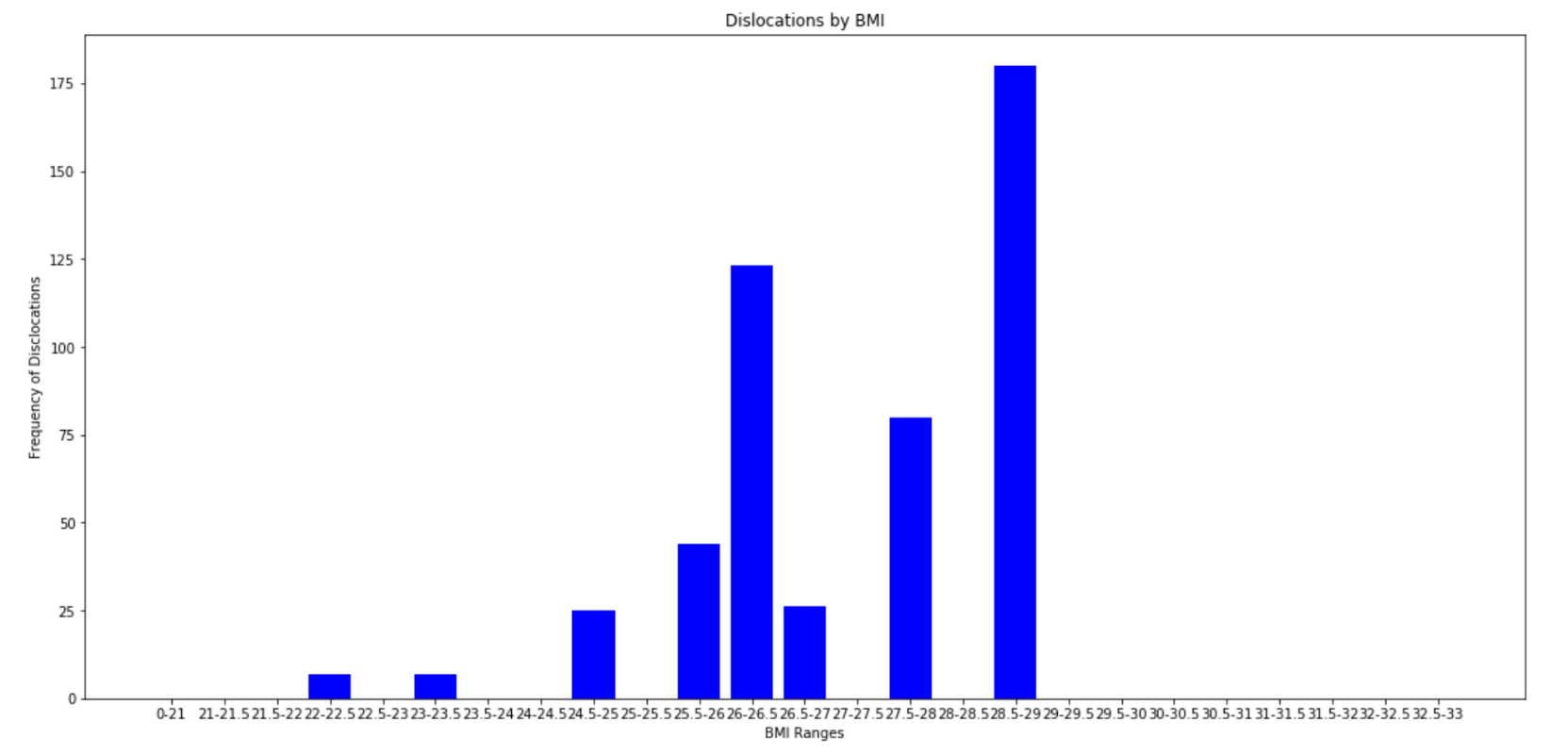
When we look at the frequency of injuries by BMI, we see three patterns emerge: 1) sprains and strains are by far the most common injuries, 2) for the most part the distribution of sprains and strains across the BMI bins looks very much like the distribution of BMIs, and 3) there does seem to be a higher frequency of arm fractures, leg fractures, and dislocations around BMIs of 27-27.5. It isn’t clear from our data if there is any correlation between fractures and dislocations, as dislocations could have occurred in either the upper or lower extremities.

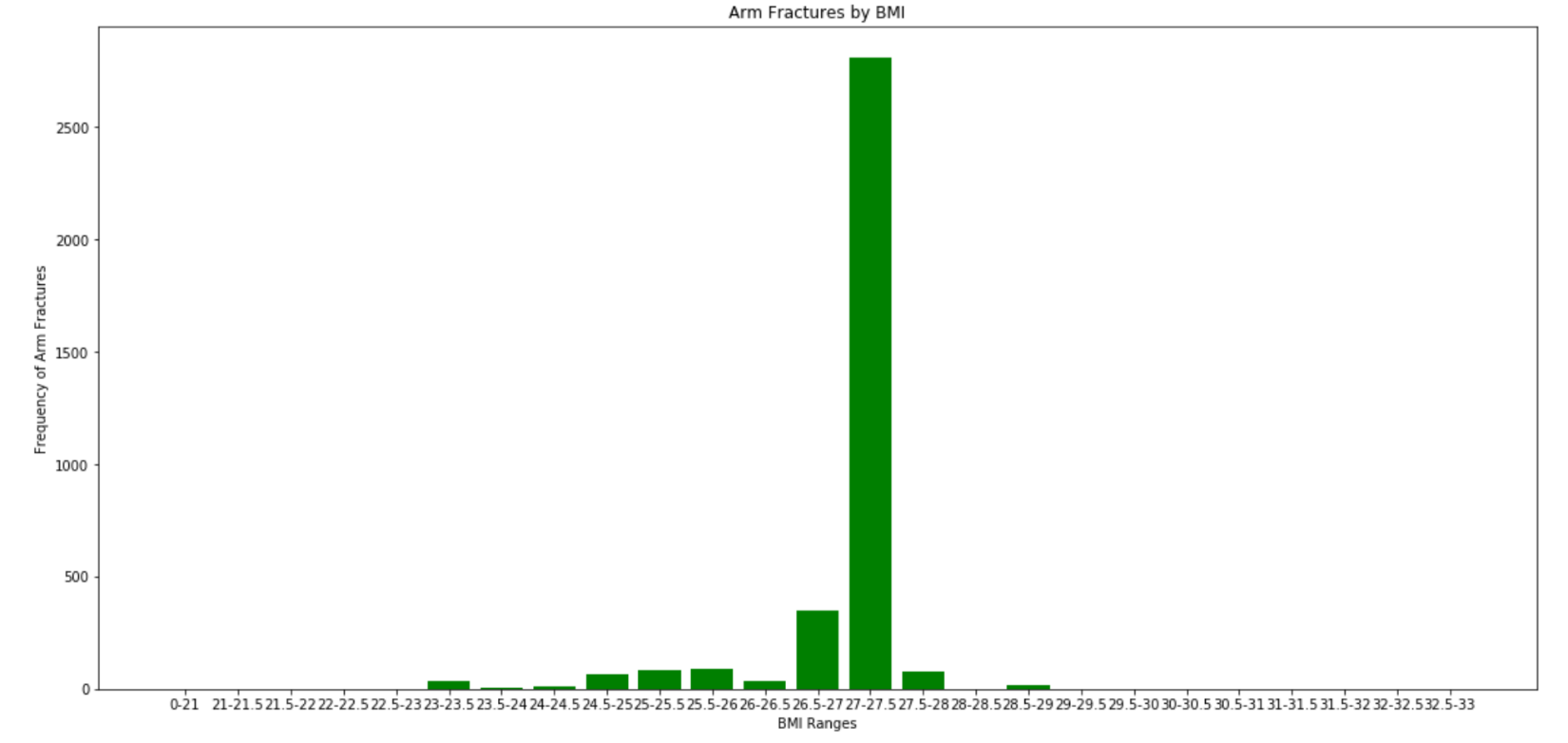
The data for sprains and strains doesn’t really lend itself to statistical analysis, as the distribution of datapoints is so wide and for some BMIs, so large, that a scatterplot of the data would have a shotgun effect.

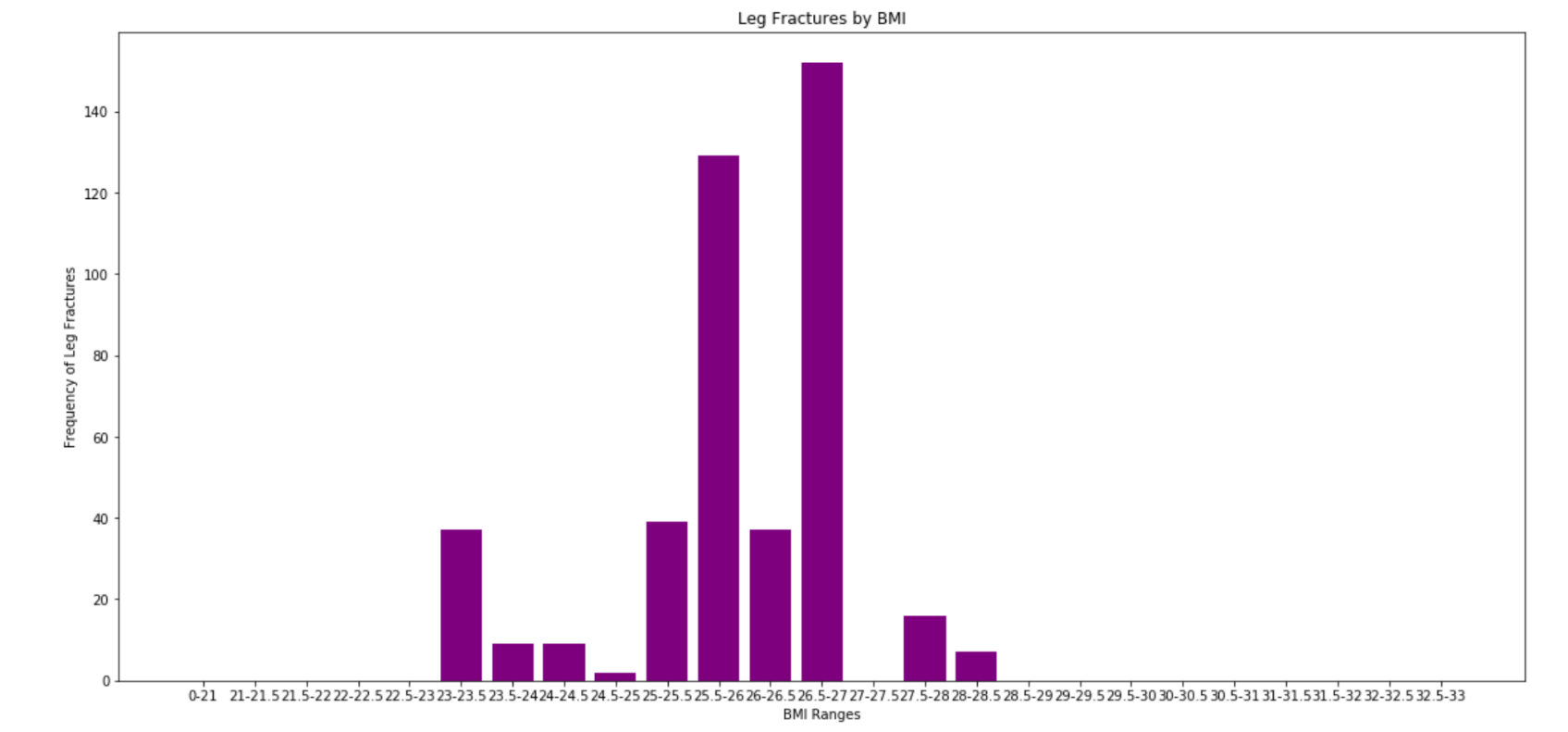
The data for arm fractures and leg fractures is interesting in that they seem to be clustered around a BMI of 27.5. With leg fractures, there are relatively few occurrences considering that our timeframe is 8 seasons, however had time allowed an ANOVA of arm/leg fractures may very well have showed a significant difference, particularly for arm fractures. As the dataset is essentially a collection of the number of games a player was out with a particular injury for each new injury that occurred, reliable ANOVA results are predicated on weighting for the number of games each injury had a player sidelined, otherwise the data is at risk of being confounded by some players taking longer to heal than others.

Dislocations seem to be centered around a BMI of 26-26.5 and 28.5-29; 26-26.5 overlaps with the mean BMI so it makes sense that there would be more injuries around that body mass, however for dislocations at a BMI of 28.5-29, there may be some suggestion that a heavier player is more prone to a dislocation. Again, this piece of data is limited by both relative infrequent occurrences, as well as a lack of knowing which joints are affected.

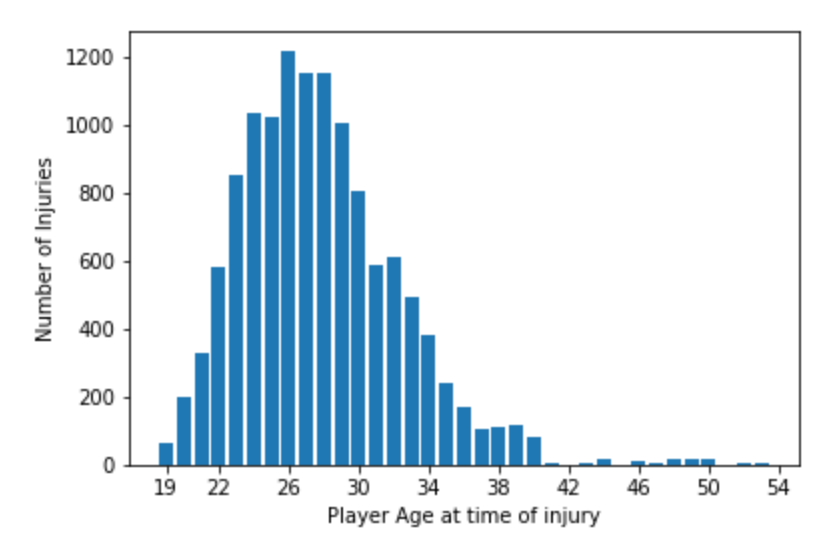




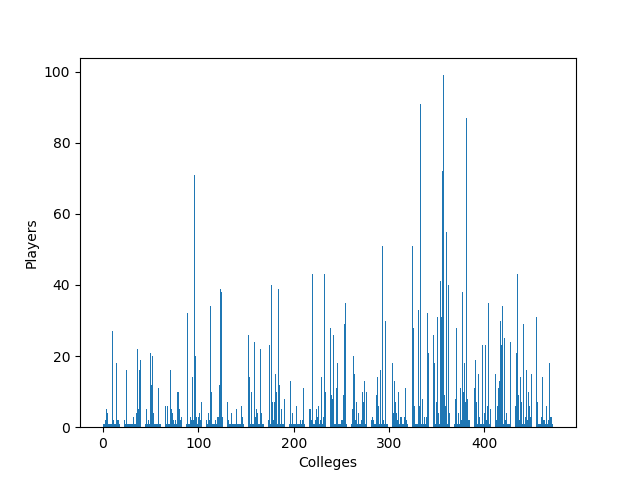
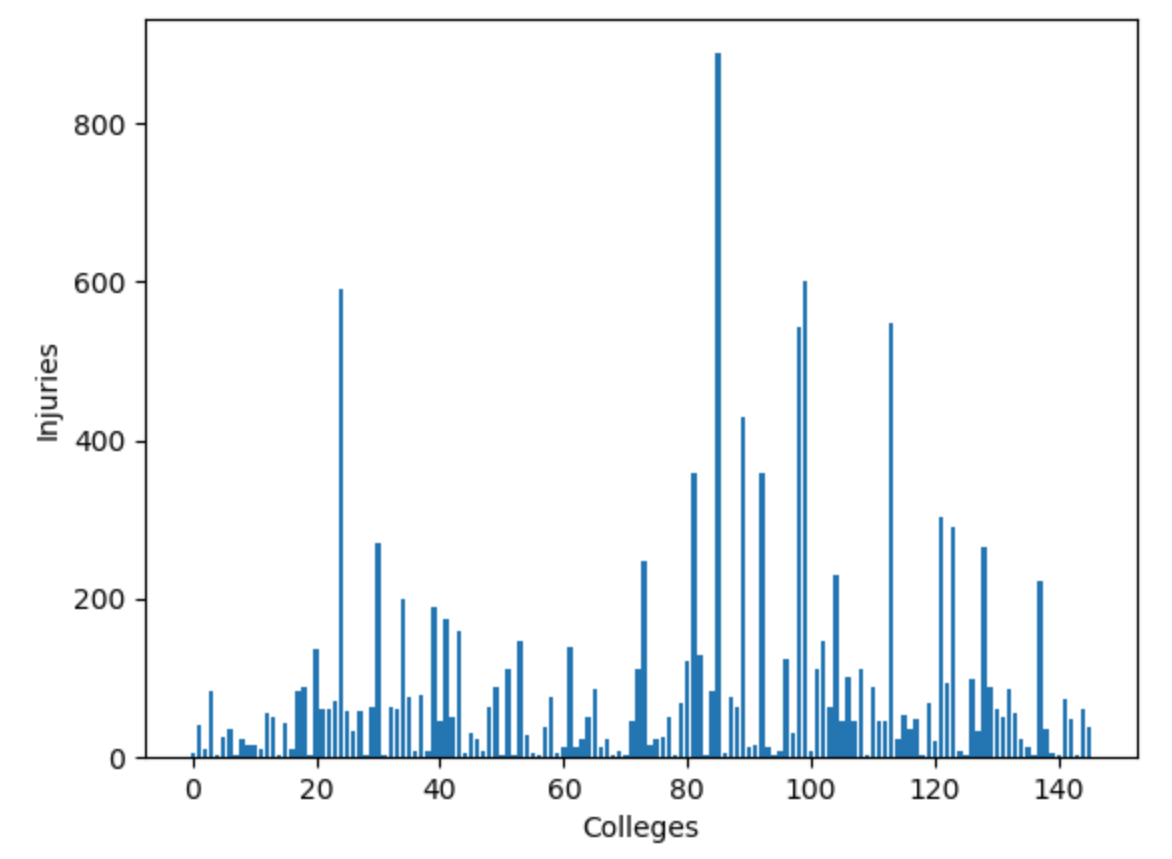




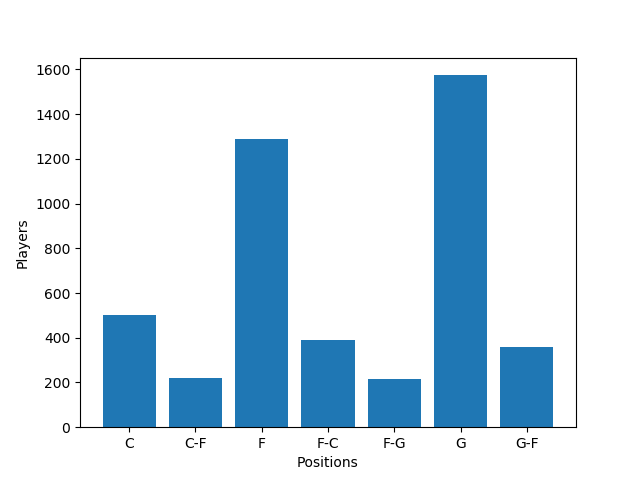
Our analysis of player age at the time of injury indicates that players 26-28 are more likely to be injured, with a peak at age 26. However, given that the mean age of players is 26.8, these findings are likely attributable to the characteristics of our population overall.



The following 2 graphs present the number of injuries based on the college from which a player was recruited (first image), and the second shows the overall number of players recruited from particular colleges. The greatest number of injuries overall was in players recruited from UCLA, however the second graph shows that the greatest number of players was recruited from UCLA, so we are unable to comment on whether there is some element(s) of the training program at UCLA that causes more injuries during an NBA career.



For our final analysis we examined the frequency of injuries based on a player’s position. The graph below shows that Guards are the most likely to be injured, which may be statistically confirmed with an ANOVA had time allowed.



**Conclusions**

We found that sprains and strain were by far the most common injuries overall, certain injuries such as leg fractures and dislocations may be more common in players around a BMI of 27-27.5, Guards are more likely to be injured than players in other positions, age at injury tracks with the mean age of NBA players, and although players who went to UCLA were injured more often, the NBA also recruits disproportionately from UCLA.