A nba player scoring per game statistics biology essay

Science, Biology

The main purpose of this report is to work out factors affecting the points players get per game in the NBA. The report comprises the best opencast and opencast & underground mining strategy, sensitivity analysis on the optimal strategy for a specific site and a generalised model for opencast mining with sites of similar properties.

2. 0 Background

When talking to the unbreakable records in the NBA history, it is unbelievable that Wilt Chamberlain got 100 points in a single game and average more than 50 points in a season. Thus, what affect points players score in each game? Here, I consider 8 factors: games played, playing time per game, field goals attempted per game, field goal percentage, 3-point field goals attempted per game , 3-point field goal percentage , free throws Attempted per game and free throws percentage that may affect points player get per game.

3. 0 Data

I take data of top 50 scores per game leaders in the NBA 2012-2013 regular season into consideration.

PLAYERGPMPGFGAFG%3PA3P%FTAFT%PTSKevin Durant4739. 518. 50. 5164. 70. 4149. 50. 90429. 6Carmelo Anthony3837. 8220. 4476. 60. 4097. 40. 82228. 5Kobe Bryant4738. 721. 10. 4665. 70. 3417. 50. 83827. 9LeBron James4338. 718. 70. 5473. 30. 4036. 40. 73426. 5James Harden4838. 317. 40. 445. 60. 32810. 10. 85925. 8Kyrie Irving3735. 618. 60. 4714. 80. 4125. 30. 85124Russell Westbrook4736. 318. 90. 4194. 10. 3256. 70. 80122. 6Stephen Curry433816. 70. 447. 10. 4573. 60. 90221. 1Dwyane

Wade393415. 40. 5081. 20. 3196. 30. 73820. 6LaMarcus Aldridge4538. 217. 60. 470. 20. 14. 90. 80120. 5Tony Parker4732. 715. 10. 5341. 10. 3964. 40. 80820. 1Jrue Holiday4238. 4170. 46330. 3543. 30. 77919. 4David Lee4637. 815. 60. 5140. 104. 20. 80219. 4Brandon Jennings4636. 816. 60. 4065. 70. 3743. 80. 82818. 7Brook Lopez4029. 414. 20. 526005. 10. 73418. 7Paul Pierce4633. 714. 80. 42250. 3465. 50. 78818. 6Monta Ellis4636. 417. 40. 43. 50. 2524. 70. 79918. 6Blake Griffin4832. 613. 90. 5310. 30. 1885. 60. 65818. 5Damian Lillard4738. 615. 40. 4236. 30. 3623. 60. 84518. 4O. J. Mayo4735. 913. 90. 4614. 80. 4273. 70. 84718Kemba Walker4635. 215. 30. 4323. 80. 3494. 30. 79718DeMar DeRozan4736. 7150. 441. 60. 284. 70. 82617. 4DeMarcus Cousins4431. 914. 70. 4440. 20. 25. 60. 76217. 4Luol Deng424014. 90. 4362. 90. 3364. 10. 8217. 3Paul George4637. 315. 10. 4275. 70. 3822. 80. 80817. 3Rudy Gay4336. 616. 40. 4113. 10. 3193. 70. 77217. 3Tim Duncan4329. 813. 70. 5050. 10. 440. 82817. 3Al Jefferson4732. 915. 40. 4770. 20. 22. 90. 83717. 1Chris Bosh4233. 912. 20. 540. 80. 254. 60. 81817. 1Danilo Gallinari4732. 913. 10. 4245. 40. 374. 90. 81117David West4733. 614. 50. 4850. 30. 2143. 90. 73917Joe Johnson4738150. 4255. 50. 3812. 60. 8217Ryan Anderson4831. 314. 10. 4347. 60. 3961. 90. 87816. 9Josh Smith4335. 515. 70. 4512. 20. 3024. 10. 49716. 9Deron Williams4636. 413. 50. 4155. 30. 344. 40. 85816. 8Klay Thompson4735. 314. 50. 41870. 3912. 10. 88816. 7Arron Afflalo4336. 714. 10. 4423. 80. 3463. 40. 85716. 7Jamal Crawford4629. 413. 50. 41750. 36240. 86316. 5Dwight Howard4334. 710. 30. 5770. 10. 259. 30. 49616. 5J. R. Smith4533. 415. 10. 4024. 90. 3383. 10. 79316. 3Al Horford4337. 313. 40. 532002. 90. 60216Nicolas Batum4638. 912. 50. 4256. 50. 3623. 50. 84915. 9Carlos Boozer4431. 214.

10. 475003. 50. 69915. 8Greg Monroe4732. 612. 80. 483004. 90. 68515.

7Zach Randolph4435. 213. 60. 4720. 40. 1253. 50. 7515. 5J. J.

Redick463211. 70. 4526. 20. 3992. 60. 89215. 3Thaddeus Young4636130.

5220. 10. 22. 60. 5715. 1Raymond Felton3333. 515. 30. 4014. 20. 3651. 70.

78215. 1Kevin Martin4629. 810. 60. 455. 20. 4353. 60. 90415. 1Ty

Lawson4734. 313. 10. 4312. 90. 363. 60. 73715

http://espn. go. com/nba/statistics/player/_/stat/scoring-per-game

GP: Games PlayedMPG: Minutes Per GamePTS: Points Per GameFGA: Field

Goals Attempted Per GameFG%: Field Goal Percentage3PA: 3-Point Field

Goals Attempted Per Game3P%: 3-Point Field Goal PercentageFTA: Free

Throws Attempted Per GameFT%: Free Throws Percentage

4. 0 Analysis

4. 1 CorrelationFirstly, I use scatterplot with regression to show the links

between points per game and the 8 factors respectively. From the picture,

we can observe FGA has the strongest relationship with PTS as the data

points are closest to the line. At the same time, GP have a weakest

relationship with PTS. In addition, I use correlation coefficient to show the

relationships between average points and the other 8 factors. Below is a

correlation matrix for all variables in the model. Numbers are Pearson

correlation coefficients, go from -1 to 1. Closer to 1 means strong correlation.

A negative value indicates an inverse relationship (roughly, when one goes

up the other goes down). From the table above, we can observe field goals

attempted per game(FGA) and free throws attempted per game(FTA) have a

great influence on the points per game players get as the correlations are 0. 840 and 0. 727 respectively. 4. 2 Simple linear regressionI take independent variables FGA and GP as the typical examples to show the linear relationship with PTS. Therefore, I build the simple regression model using FGA as the independent variable and PTS as the dependent variable and get results below. The t-values test the hypothesis that the coefficient is different from 0. To reject this, we need a t-value greater than 1. 96 (for 95% confidence). In this case, t-value of FGA is 10. 72, which indicates there is a linear relationship between PTS and FGA. Two-tail p-values test the hypothesis that each coefficient is different from 0. To reject this, the p-value has to be lower than 0. 05 (you could choose also an alpha of 0. 10). In this case, with a p-value of 0. 000, there is very strong evidence to suggest that the simple linear regression model is useful for PTS. The r2 value listed on the output is 70. 5%, which is implies that about 70. 5% of the sample variation in points per game(PTS) is explained by field goals attempted per game(FGA) in a straight-line model. There are some unusual observations and thus there are likely other variables that affect PTS. Moreover, the most important part of the ANOVA table is the probability. The probability is calculated by assuming that the independent variable in question has no effect and then gauging the likelihood that the outcome you observed would occur. The effect of the variable is called statistically significant if the P value is less than 0. 05 or 0. 01, with smaller numbers indicating higher significance. Since the P value above is well below 0. 01, we can reasonably say that the tested factor(PGA) has a real impact on the response variable（PTS）.. The regression equation is PTS = -0. 84 +1. 29*FGA. For each one-point increase in FGA, scores

increase by 1. 29 pointsA typical assumption in regression is that the random errors () are normally distributed. The normality assumption is important when conducting hypothesis tests of the estimates of the coefficients (). Fortunately, even when the random errors are not normally distributed, the test results are usually reliable when the sample is large enough. In this case, it is a well-behaved residual. Then let me show another typical relationship by using PTS as a response and GP as a predictor. In this case, t-value of GP is -0. 67, greater than -1. 96 (for 95% confidence), which indicates we have insufficient evidence to conclude that a statistically significant relationship between PTS and FGA exists. Alternatively, with the p-value of 0. 509, greater than 0. 05, we can also obtain there is no significant relationship between PTS and FGA. The r2 value listed on the output is only 0. 9%, which is implies that nearly no sample variation in points per game(PTS) is explained by games player(GP) in a straight-line model. Moreover, the P value of the ANOVA table above is 0. 509 far greater than 0. 05 and there are many unusual observations, we can reasonably say that the tested factor(GP) has no impact on the response variable（PTS）. The picture illustrates that the random errors are not normally distributed, there the test results are not reliable. Overall, there is no significant relationship between PTS and FGA. 4. 3 Multiple Linear RegressionTo improve the results obtained above, we can use the multiple linear regression to get the relationship between the points per game and other 8 factors. The t-values test the hypothesis that the coefficient is different from 0. To reject this, you need a t-value greater than 1. 96 (at 0. 05 confidence). The t-values also show the importance of a variable in the model. In this case, FGA is the most

important. Alternatively, two-tail p-values test the hypothesis that each coefficient is differentfrom 0. To reject this, the p-value has to be lower than 0. 05 (you could choose also an alpha of 0. 10). In this case, GP, MPG, and 3P % are not statistically significant in explaining PTS. FGA, FG%, 3PA, FTA, FT% are variables that have some significant impact on PTS. Moreover, the model explains 99. 0% of variances on PTS. The P value of the ANOVA table above is 0. 000 far less than 0. 05, we can reasonably say that the tested factors has great impact on the response variable（PTS）. Overall, the model describes the variation in data well, however, we can still improve it as there are some factors that are not important in explaining PTS. 4. 4 ImprovementAs discussed before, the three factors GP, MPG and 3P% are not statistically significant in explaining PTS, therefore, I exclude the three factors and build a new multiple regression model with other 5 factors. Similarly, we can conclude the five facotrs have some influence on PTS. Among them, FGA have the largest impact.