Regression in the NBA

Your turn to run a regression analysis on the 2017 NBA Season.

## Instructions

\* Run a regression analysis on the 2017 NBA season stats.

\* You will use `PTS` (points) as your Y variable, or dependent variable.

\* You will determine which X variables, or independent variables, to use.

\* After you have run regression, explain why your choice was or was not a good choice.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |  |  |
| Multiple R | 0.036156423 |  |  |  |  |  |  |  |
| R Square | 0.001307287 |  |  |  |  |  |  |  |
| Adjusted R Square | -0.000376849 |  |  |  |  |  |  |  |
| Standard Error | 470.21903 |  |  |  |  |  |  |  |
| Observations | 595 |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 171630.366 | 171630.366 | 0.776235903 | 0.378651755 |  |  |  |
| Residual | 593 | 131115820.1 | 221105.9362 |  |  |  |  |  |
| Total | 594 | 131287450.5 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | 369.4813779 | 121.0076694 | 3.053371574 | 0.00236428 | 131.8256453 | 607.1371105 | 131.8256453 | 607.1371105 |
| Age | 3.985780304 | 4.52393643 | 0.881042509 | 0.378651755 | -4.89910635 | 12.87066696 | -4.89910635 | 12.87066696 |

\* Do the results have a good `R Square` value?

The R Square value is too low that this regression model is not going to tell the correct information.

\* Is your choice statistically reliable?  
The value of F test is more than Alpha 0.05 value which , indicates that we fail to reject null hypothesis and this model of regression need to remodel . Similarly p value for Age is greater than alpha 0.05 so again this value is not statistically reliable.

\* Explain what your coefficients mean.

Coefficients value is 3 i.e positive but it is too high than 1 value which means that there is mistake in this calculation which means this model is wrong.

\* Finally produce a scatter plot and explain your overall analysis.

The scatter diagram doesn’t show any linear relation between PTS and Age . The selection of Age as X variable is wrong choice. The players are mostly selected from 20 to 40 years in most of the games and it is ovious that player 20 to 30 years are equaly energetic and may score similar no of PTS.