PRELIMARY RESEARCH:
How does an NBA point guard's regular season performance impact their salary through the 2020-21 NBA regular season?

Pages: 18

I. Introduction

From playing basketball and watching NBA (National Basketball Association) games, I have always been intrigued by the point guard position. The point guard is responsible for running a team's offense and initiating different strategies/plays to score points. It is also a position that requires a lot of split-second decision making, even more than other basketball positions. However, their playmaking ability is often overshadowed by flashy dunks and long-distance three pointers by athletic superstars like LeBron James. With this in mind, I wanted to learn how to identify successful NBA point guards through accurate statistics.

In the NBA, a player's salary is heavily based on their on-court basketball "performance". NBA salaries are often clear indicators to distinguish NBA stars from regular players. However, "performance" is a general term that can be represented through many measurements of an NBA player's ability to defend, score, and contribute to their team. Thus, within the research question, the investigation will also identify which measurement of an NBA point guard's performance has the greatest impact on their salary.

II. Which measurement of an NBA point guard's performance has the greatest impact on their salary?

To measure an NBA point guard's on-court performance, I will be using five different measurements and compare each measurement with their salaries.

- **Number of points per game (PTS)**
- **Number of assists per game (AST)**
- **Number of steals per game (STL)**
- **Efficiency (EFF)**

- This is a total performance statistic that measures individual player efficiency by including points, rebounds, assists, steals, and blocks, as well as missed field goals, missed free throws, and turnovers
- $EFF = PTS + REB + AST + STL + BLK - Missed\ FG - Missed\ FT - TO$

- **True shooting percentage (TS%)**
  - This measures a player's shooting efficiency by including field goal percentages, free throw percentages, and three-point field goal percentages.
  - $TS\% = \frac{PTS}{2(FGA+(0.44 \times FTA)}$

I believe these measurements are accurate statistics that define a point guard's ability on the basketball court, given that their role is to initiate offensive plays and guard the opposing point guard. The investigation will also use statistical visualizations like linear regression models and statistical calculations like the Pearson's product-moment correlation coefficient and the variance inflation factor (VIF) to determine the strength of the relationship between performance and salary as well as provide an insight on what aspects of a point guard's performance are highly wanted by NBA organizations.

III. <u>Hypothesis</u>

I predict that the assists statistic (AST) and the true shooting percentage statistic (TS%) will have the greatest impact on a player's salary because they measure a point guard's role to facilitate other players and become an offensive presence. With an exception to some defensively-minded point guards like Ben Simmons, point guards are generally smaller players that are positioned around the three-point line and dribble, pass, or shoot proficiently. Whilst they are a part of a team's defensive formation, the responsibility of a team's primary defensive presence is held by the center and forwards, larger players who are positioned closer to the basket. This is also why point guards generally average less steals (STL) and rebounds (REB) than centers and forwards.

I also believe that the efficiency statistic (EFF) would be greatly valued in determining player salary because it is a holistic evaluation of a player's capability on the basketball court. EFF calculates a player's effectiveness by measuring positive actions like points (PTS) and negative actions like turnovers (TO) and identifying which actions are done more frequently. However, I do not believe it will have impact similar to AST or TS% because EFF does not specifically measure a point guard's ability. By including measurements like REB, it measures aspects of performance that are more demanded by other positions, thus, not giving an accurate assessment of NBA point guards.

IV. Data Collection & Cleaning Process

Data collection:

From the 2020-2021 NBA regular season, I will collect each point guard's performance data through the official ESPN website [1]. From the same season, I will also collect each point guard's salary data through the HoopsHype website [2] and use both databases to create a dataset with all 93 NBA point guards (see Appendix). I understand that the NBA salaries are issued through player contracts which can last multiple years, however, the HoopsHype website automatically calculates an NBA player's salary for any particular year, making their data accurate.

Data cleaning:

I will create a sample dataset from the main dataset with a restriction due to the large number of data points from the main dataset. The restriction is that all NBA point guards in the sample must be a starter for their teams. With an exception to injuries, starters usually have more playing time which helps produce more reliable results since there is more data to extract and analyze. This creates a sample dataset with 30 data points because there are 30 teams in the NBA.

Here is the Python code I used to create a sample dataset through Anaconda (a Python programming platform):

```
"""
Created on Fri Jan 21 16:53:03 2022
@author: rigdenlama
"""
import pandas as pd

# The main dataset is called "ESPN database"
df = pd.read_csv(r'/Users/rigdenlama/Desktop/Math IA/ESPN
database.csv')

# Below is the list of all the starting players in the NBA
names = ['Trae Young',
'Kyrie Irving',
'LaMelo Ball',
'Kemba Walker',
'Stephen Curry',
```

```
'Luka Doncic',
'Jamal Murray',
'Ja Morant',
'Dejounte Murray',
'Chris Paul',
'Damian Lillard',
'John Wall',
'Russell Westbrook',
'Ben Simmons',
"De'Aaron Fox",
'Kyle Lowry',
'Jrue Holiday',
'Darius Garland',
'Coby White',
"D'Angelo Russell",
'Mike Conley',
'Malcolm Brogdon',
'Tyler Herro',
'Dennis Schroder',
'Eric Bledsoe',
'Cole Anthony',
'Patrick Beverley',
'Shai Gilgeous-Alexander',
'Elfrid Payton',
'Delon Wright']

# The sample dataset is called "sample_df"
sample_df = df[df.NAME.isin(names)]
print(sample_df)
```

V. Sample Dataset

NAME	POS	GP	MIN	PTS	FGM	FGA	FG%	FTM	FTA	FT%	REB	AST	STL	BLK	TO	SALARY (USD)
Stephen Curry	PG	63	34.2	32	10.4	21.7	48.2	5.7	6.3	91.6	5.5	5.8	1.2	0.1	3.4	43006362
Damian Lillard	PG	67	35.8	28.8	9	19.9	45.1	6.7	7.2	92.8	4.2	7.5	0.9	0.3	3	31626953
Luka Doncic	PG	66	34.3	27.7	9.8	20.5	47.9	5.2	7.1	73	8	8.6	1	0.5	4.3	8049360
Kyrie Irving	PG	54	34.9	26.9	10.2	20.1	50.6	3.7	4	92.2	4.8	6	1.4	0.7	2.4	33722850
Trae Young	PG	63	33.7	25.3	7.7	17.7	43.8	7.7	8.7	88.6	3.9	9.4	0.8	0.2	4.1	6571800
De'Aaron Fox	PG	58	35.1	25.2	9.1	19.1	47.7	5.2	7.2	71.9	3.5	7.2	1.5	0.5	3	8099627
Shai Gilgeous-Alexander	PG	35	33.7	23.7	8.2	16.1	50.8	5.3	6.5	80.8	4.7	5.9	0.8	0.7	3	4141320
Russell Westbrook	PG	65	36.4	22.2	8.4	19	43.9	4.2	6.4	65.6	11.5	11.7	1.4	0.4	4.8	41358814
Jamal Murray	PG	48	35.5	21.2	7.9	16.5	47.7	2.8	3.2	86.9	4	4.8	1.3	0.3	2.3	27285000
Malcolm Brogdon	PG	56	34.5	21.2	7.9	17.5	45.3	2.7	3.2	86.4	5.3	5.9	0.9	0.3	2.1	20700000
John Wall	PG	40	32.2	20.6	7.3	18.2	40.4	4	5.3	74.9	3.2	6.9	1.1	0.8	3.5	41254920
Kemba Walker	PG	43	31.8	19.3	6.6	15.7	42	3.1	3.5	89.9	4	4.9	1.1	0.3	2	34379100

Ja Morant	PG	63	32.6	19.1	6.8	15.2	44.9	4.3	5.9	72.8	4	7.4	0.9	0.2	3.2	9166800
Jrue Holiday	PG	59	32.3	17.7	7	13.9	50.3	1.8	2.3	78.7	4.5	6.1	1.6	0.6	2.2	25110000
Darius Garland	PG	54	33.1	17.4	6.7	14.9	45.1	2.1	2.4	84.8	2.4	6.1	1.2	0.1	3	6720720
Kyle Lowry	PG	46	34.8	17.2	5.7	13	43.6	3	3.5	87.5	5.4	7.3	1	0.3	2.7	30500000
Chris Paul	PG	70	31.4	16.4	6.3	12.6	49.9	2.4	2.6	93.4	4.5	8.9	1.4	0.3	2.2	41358814
Mike Conley	PG	51	29.4	16.2	5.6	12.5	44.4	2.4	2.8	85.2	3.5	6	1.4	0.2	1.9	34502132
LaMelo Ball	PG	51	28.8	15.7	5.7	13.2	43.6	2.5	3.2	75.8	5.9	6.1	1.6	0.4	2.8	7839960
Tyler Herro	PG	54	30.3	15.1	5.7	12.9	43.9	1.7	2.2	80.3	5	3.4	0.6	0.3	1.9	3822240
Coby White	PG	69	31.2	15.1	5.4	13.1	41.6	1.9	2.1	90.1	4.1	4.8	0.6	0.2	2.3	5572680
Ben Simmons	PG	58	32.4	14.3	5.6	10.1	55.7	3	4.9	61.3	7.2	6.9	1.6	0.6	3	30559200
Cole Anthony	PG	47	27.1	12.9	4.7	11.7	39.7	2.3	2.8	83.2	4.7	4.1	0.6	0.4	2.3	3285120
Elfrid Payton	PG	63	23.6	10.1	4.3	9.9	43.2	1.2	1.7	68.2	3.4	3.2	0.7	0.1	1.6	4767000
Patrick Beverley	PG	37	22.5	7.5	2.5	5.9	42.3	1	1.2	80	3.2	2.1	0.8	0.8	0.9	13333333
Delon Wright	PG	63	27.7	10.2	3.8	8.2	46.3	1.6	2	80.2	4.3	4.4	1.6	0.5	1.3	9000000
Eric Bledsoe	PG	71	29.7	12.2	4.3	10.3	42.1	1.9	2.7	68.7	3.4	3.8	0.8	0.3	1.6	16875000
Dennis Schroder	PG	61	32.1	15.4	5.4	12.5	43.7	3.4	4	84.8	3.5	5.8	1.1	0.2	2.7	16000000
Dejounte Murray	PG	67	31.9	15.7	6.6	14.5	45.3	1.6	2	79.1	7.1	5.4	1.5	0.1	1.7	14286000
D'Angelo Russell	PG	42	28.5	19	6.7	15.5	43.1	2.7	3.5	76.5	2.6	5.8	1.1	0.4	2.7	28649250

VI. Mathematical Concepts

Below, I will describe the mathematical concepts used to analyze the data of NBA point guards and my reasons for choosing these concepts in this investigation.

Pearson's product-moment correlation coefficient (Pearson's r):

Pearson's r is a statistical test that measures the correlation strength and direction of the relationship between two variables (often denoted as x and y). The r value has a range between -1 to 1 that indicates the following:

- $0.7 < r \leqslant 1$ – positive and strong correlation
- $0.3 < r \leqslant 0.7$ – positive and weak to moderate correlation
- $-0.3 < r \leqslant 0.3$ – no correlation
- $-0.7 < r \leqslant -0.3$ – negative and moderate to weak correlation
- $-1 \leqslant r \leqslant -0.7$ – negative and strong correlation

For this investigation, using Pearson's r will be relevant because it analyzes the strength of the relationship between a factor of player performance and player salary as well as it provides an insight to what extent the variables influence each other.

Pearson's r formula:

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$$

Pearson's r is calculated by first finding the sum ($\sum$) of all the x values that were subtracted by $\bar{x}$ (the average of all the values) multiplied by all the y values that were subtracted by $\bar{y}$. After, the numerator value is divided by the square root ($\sqrt{}$) of the sum of $(x-\bar{x})^2$ multiplied by the sum of $(y-\bar{y})^2$ to produce the r value.

Variance Inflation Factor (VIF):

The VIF measures multicollinearity in a set of multiple regression variables. This means that it "measures how much behavior (variance) of an independent variable is influenced, or inflated, by its interaction/correlation with the other independent variables" [3]. To interpret a VIF value, here are the general value ranges that determine high or low correlation and multicollinearity:

- VIF > 5 – high (severe) correlation between the independent variables
- 5 < VIF < 1 – moderate correlation between the independent variables
- VIF = 1 – no correlation between the independent variables

Multicollinearity is the statistical concept of high intercorrelations between two or more independent variables. Often, high multicollinearity leads to skewed results from researchers determining correlations between an independent variable (e.g. points) and a dependent variable (salary). Hence, using a VIF will be relevant for this investigation because it provides a potential reason to why the correlations between performance and salary are determined as well as whether the correlations determined are due to the correlations between the other performance statistics.

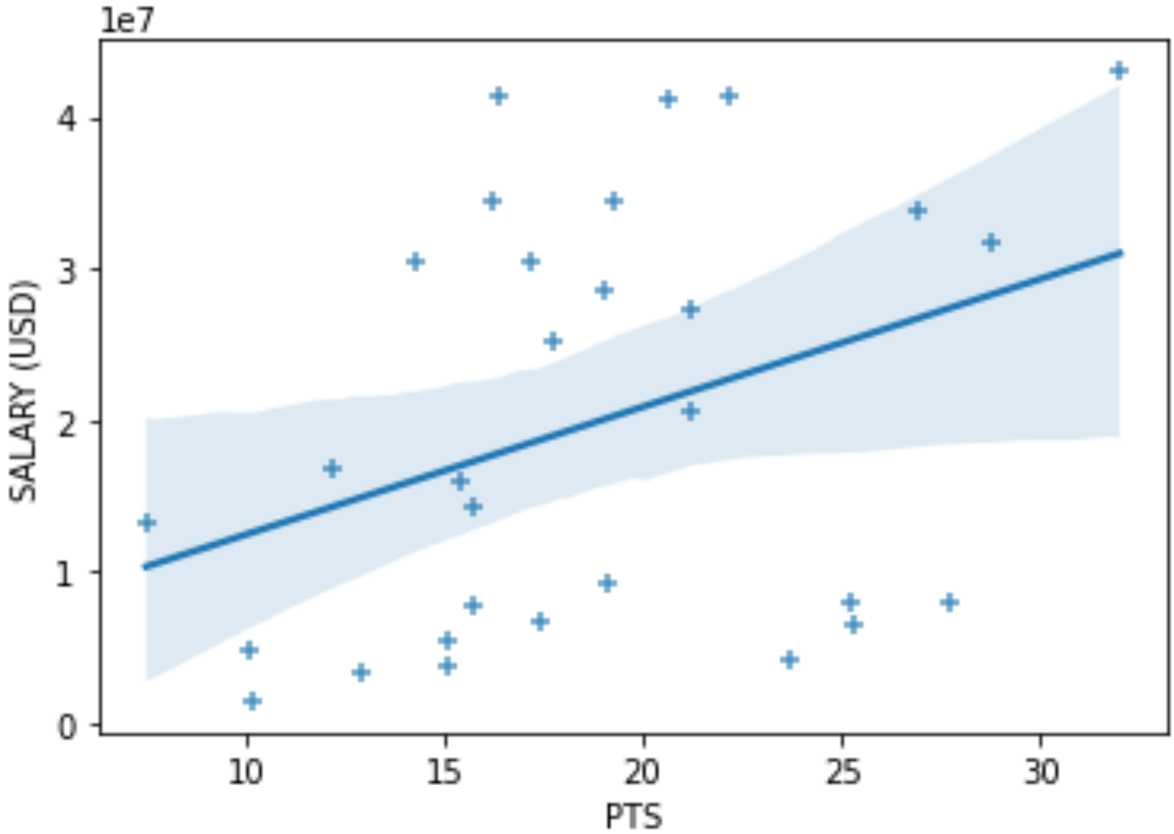
VIF formula:

$VIF = \frac{1}{1 - R_i^2}$

A VIF is calculated by first calculating the R^2 values from the multiple independent variables identified and then using the formula or performing a multiple linear regression with the chosen independent variable as the response variable and the other independent variables as the explanatory variables to calculate the response variable's VIF. After, the same process is repeated to find the VIFs for the other independent variables. However, VIFs are generally calculated through software so I will be using Python code to calculate the VIFs for all of the performance measurements.

VII. PTS & Salary

NUMBER OF POINTS PER GAME VS. PLAYER SALARY



Pearson's r value:

x values: PTS

y values: SALARY

$\bar{x} = 18.71$

$\bar{y} = 19801778.5$

$\sum(x - \bar{x})(y - \bar{y}) = 846832788.1$

$\sum(x - \bar{x})^2 = 1005.107$

$\sum(y - \bar{y})^2 = 5.59298 \times 10^{15}$

$$r = \frac{846832788.1}{\sqrt{1005.10 \times (5.59298 \times 10^{15})}}$$

$$r = \frac{846832788.1}{2.37 \times 10^9}$$

$r = 0.357\ (3\ s.f.)$

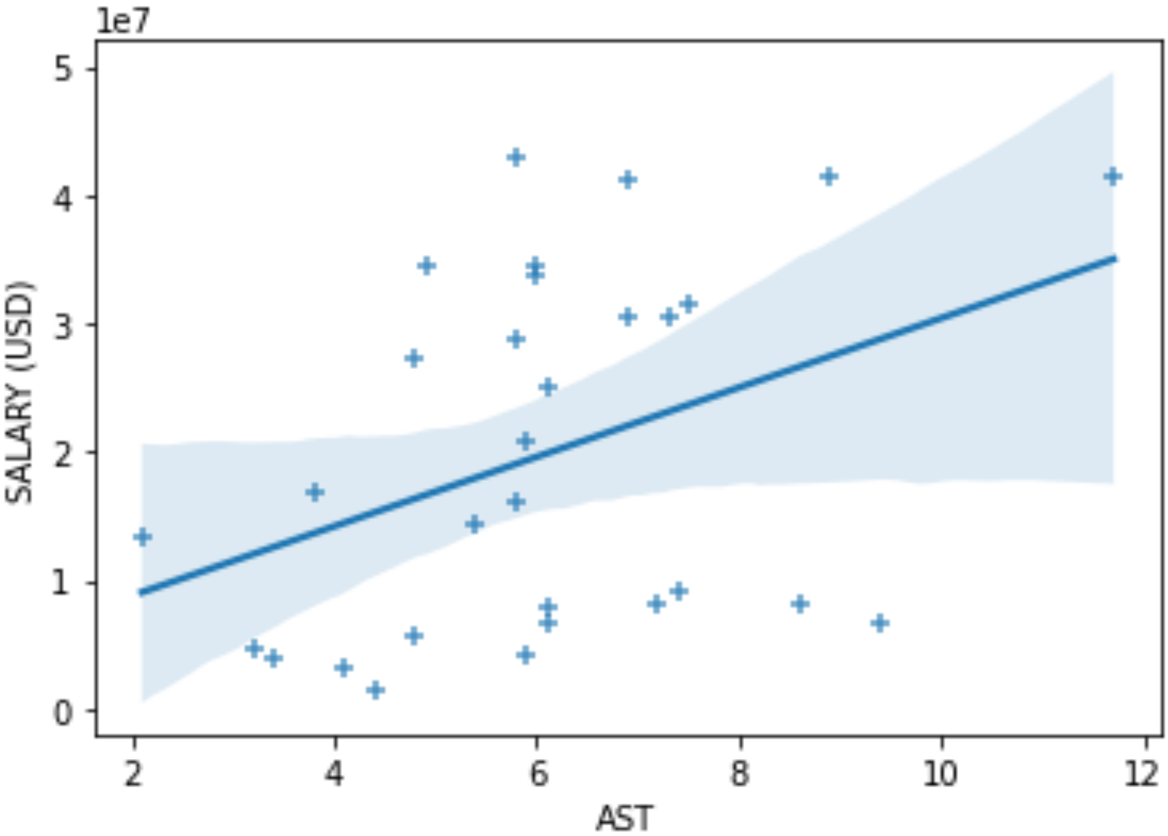
Here is the Python code I used to create the regression and recheck Pearson's r value:

```
"""
Created on Sat Jan 8 18:09:45 2022
@author: rigdenlama
"""
pts_salary = sample_df[['PTS', 'SALARY (USD)']]

import seaborn as sns
sns.regplot(data=sample_df, x='PTS', y='SALARY (USD)', marker='+')

import numpy as np
x = sample_df['PTS']
y = sample_df['SALARY (USD)']
correlation_matrix = np.corrcoef(x, y)
r_value = correlation_matrix[0,1]
print(r_value)
```

r = 0.3571659104507024

VIII. AST & Salary

NUMBER OF ASSISTS PER GAME VS. PLAYER SALARY



Pearson's r value:

x values: AST

y values: SALARY

$\bar{x} = 6.07\dot{3}$

$\bar{y} = 19801778.5$

$\sum(x - \bar{x})(y - \bar{y}) = 306737991$

$\sum(x - \bar{x})^2 = 113.458\dot{6}$

$\sum(y - \bar{y})^2 = 5.59298 \times 10^{15}$

$r = \frac{306737991}{\sqrt{113.458\dot{6} \times (5.59298 \times 10^{15})}}$

$r = \frac{306737991}{796600092.9}$

$r = 0.385\ (3\ s.f.)$

Here is the Python code I used to create the regression and recheck Pearson's r value:

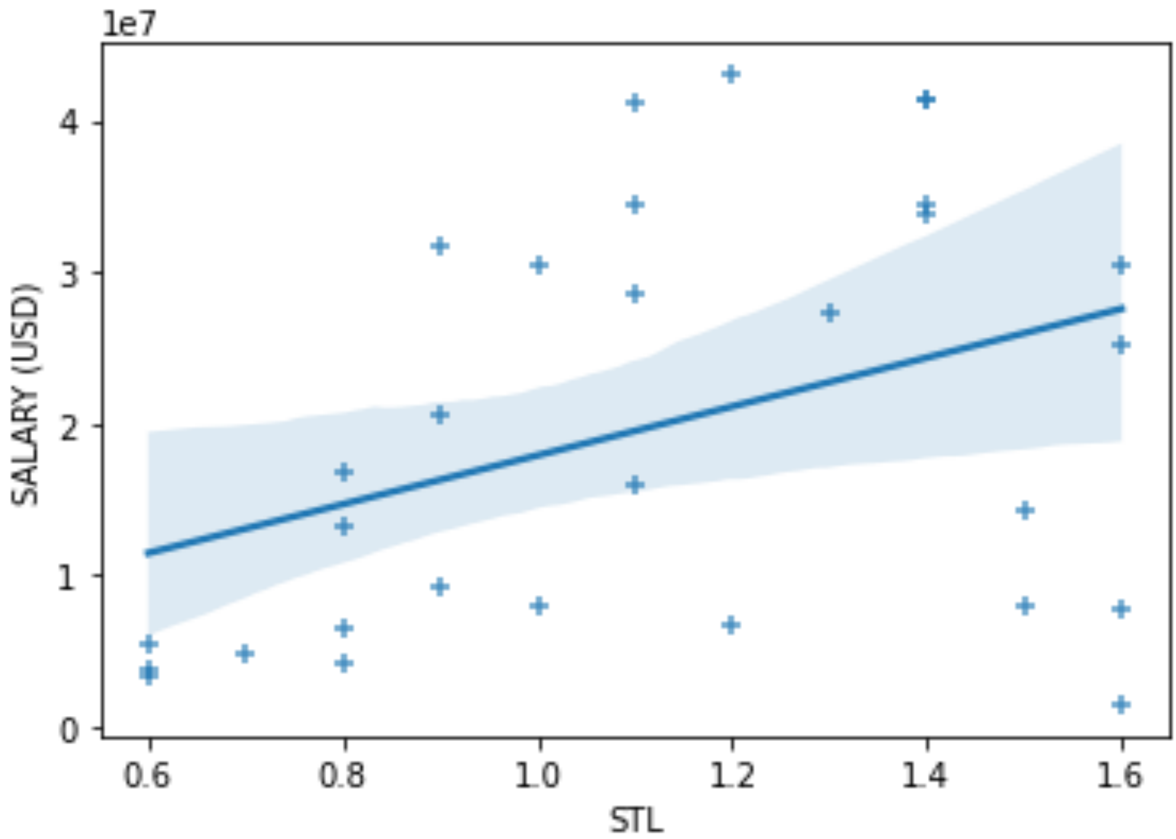
"""

```
Created on Sat Jan 8 18:11:32 2022
@author: rigdenlama
"""
ast_salary = sample_df[['AST', 'SALARY (USD)']]

import seaborn as sns
sns.regplot(data=sample_df, x='AST', y='SALARY (USD)', marker='+')

import numpy as np
x_values = sample_df['AST']
y_values = sample_df['SALARY (USD)']
correlation_matrix = np.corrcoef(x_values, y_values)
r_value = correlation_matrix[0,1]
print(r_value)
```

r = 0.3850589439309289

IX. STL & Salary

NUMBER OF STEALS PER GAME VS. PLAYER SALARY



Pearson's r value:

x values: STL

y values: SALARY

$\bar{x} = 1.11\dot{6}$

$\bar{y} = 19801778.5$

$\sum(x - \bar{x})(y - \bar{y}) = 50597990.3$

$\sum(x - \bar{x})^2 = 3.141\dot{6}$

$\sum(y - \bar{y})^2 = 5.59298 \times 10^{15}$

$$r = \frac{50597990.3}{\sqrt{3.141\dot{6} \times (5.59298 \times 10^{15})}}$$

$$r = \frac{50597990.3}{132556664.4}$$

$r = 0.382\ (3\ s.f.)$

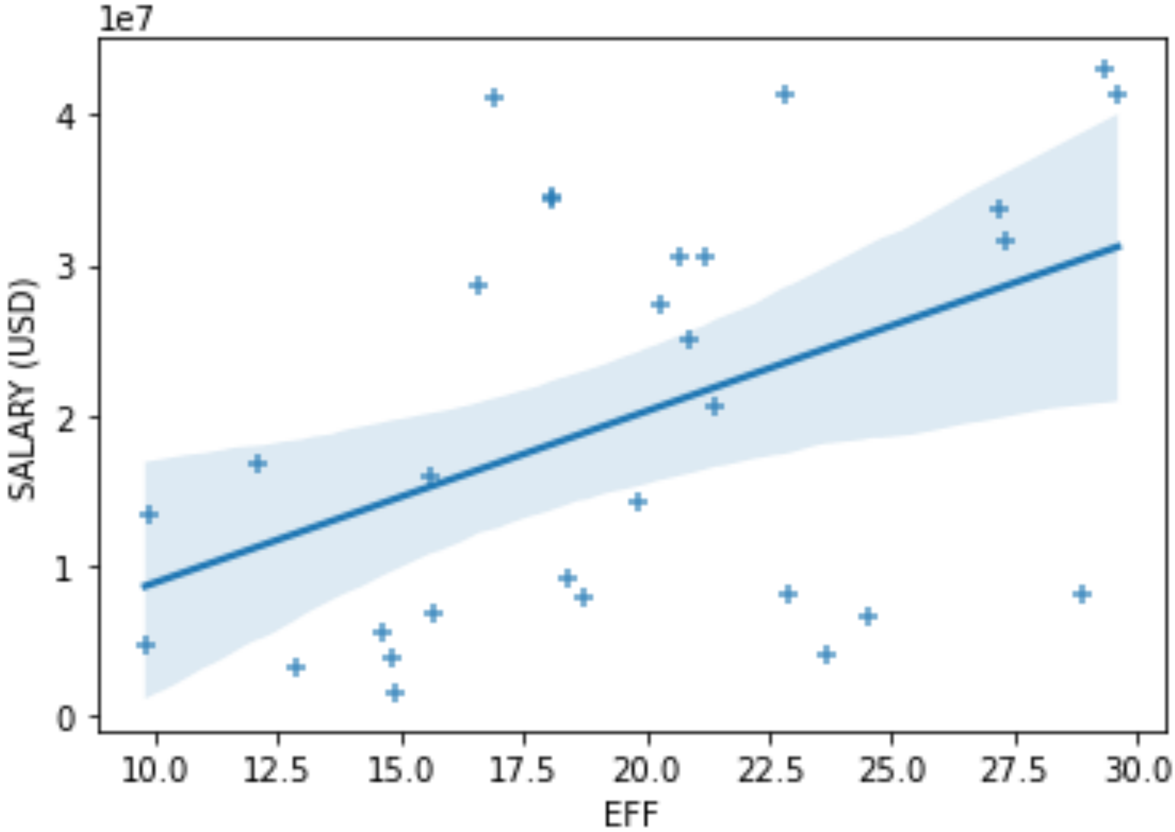
Here is the Python code I used to create the regression and recheck Pearson's r value:

```
"""
Created on Sat Jan 8 18:21:51 2022
@author: rigdenlama
"""
stl_salary = sample_df[['STL', 'SALARY (USD)']]

import seaborn as sns
sns.regplot(data=sample_df, x='STL', y='SALARY (USD)', marker='+')

import numpy as np
x_values = sample_df['STL']
y_values = sample_df['SALARY (USD)']
correlation_matrix = np.corrcoef(x_values, y_values)
r_value = correlation_matrix[0,1]
print(r_value)
```

r = 0.38170838482261044

X. EFF & Salary

EFFICIENCY VS. PLAYER SALARY



Pearson's r value:

x values: EFF

y values: SALARY

$\bar{x} = 19.58\dot{6}$

$\bar{y} = 19801778.5$

$\sum(x - \bar{x})(y - \bar{y}) = 1009063435$

$\sum(x - \bar{x})^2 = 884.454\dot{6}$

$\sum(y - \bar{y})^2 = 5.59298 \times 10^{15}$

$r = \frac{1009063435}{\sqrt{884.454\dot{6} \times (5.59298 \times 10^{15})}}$

$r = \frac{1009063435}{2224125574}$

$r = 0.454\ (3\ s.f.)$

Here is the Python code I used to create the regression and recheck Pearson's r value:

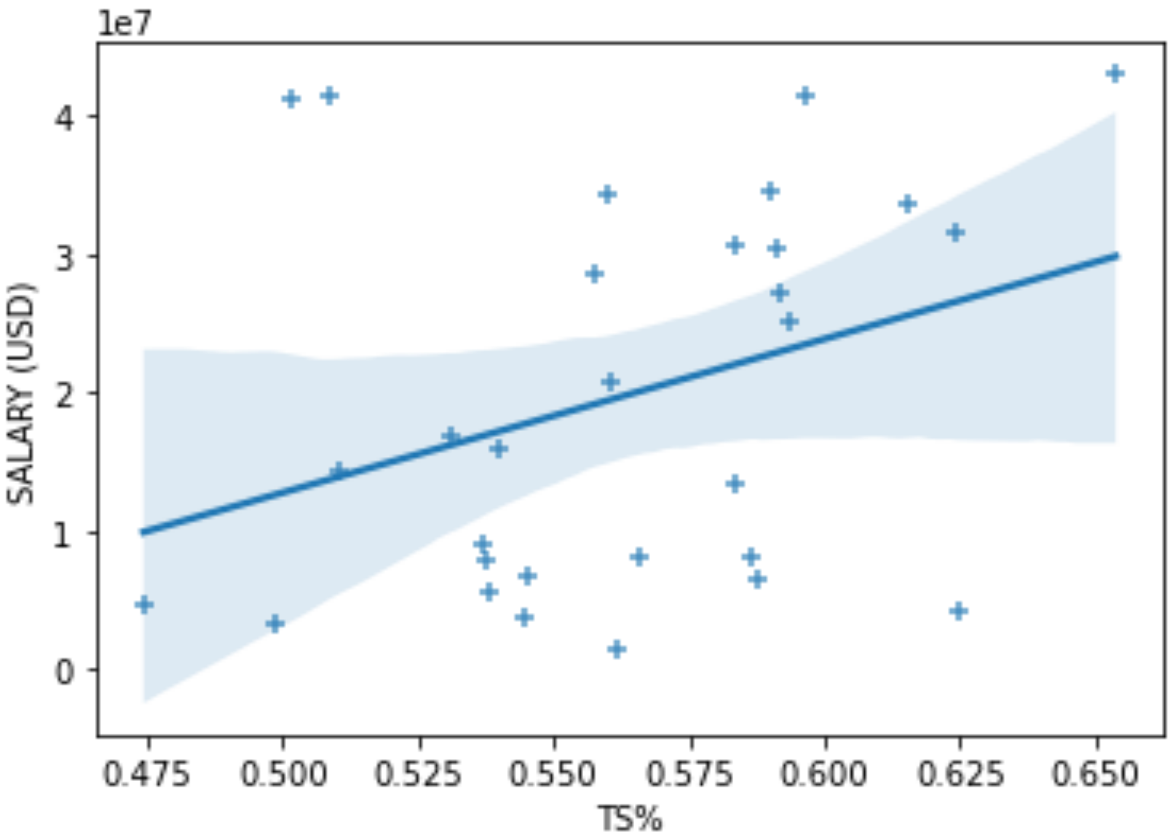
"""

```
Created on Sat Jan 8 18:25:55 2022
@author: rigdenlama
"""
sample_df['Missed FG'] = sample_df['FGA'] - sample_df['FGM']
sample_df['Missed FT'] = sample_df['FTA'] - sample_df['FTM']
sample_df['EFF'] = sample_df['PTS'] + sample_df['REB'] +
sample_df['AST'] + sample_df['STL'] + sample_df['BLK'] -
sample_df['Missed FG'] - sample_df['Missed FT'] - sample_df['TO']
eff_salary = sample_df[['EFF', 'SALARY (USD)']]

import seaborn as sns
sns.regplot(data=sample_df, x='EFF', y='SALARY (USD)', marker='+')

import numpy as np
x_values = sample_df['EFF']
y_values = sample_df['SALARY (USD)']
correlation_matrix = np.corrcoef(x_values, y_values)
r_value = correlation_matrix[0,1]
print(r_value)
```

r = 0.4536899564372864

XI. TS% & Salary

TRUE SHOOTING PERCENTAGE VS. PLAYER SALARY



Pearson's r value:

x values: TS%

y values: SALARY

$\bar{x} = 0.5631534$

$\bar{y} = 19801778.5$

$\sum(x - \bar{x})(y - \bar{y}) = 5620640.28$

$\sum(x - \bar{x})^2 = 0.05067625$

$\sum(y - \bar{y})^2 = 5.59298 \times 10^{15}$

$$r = \frac{5620640.28}{\sqrt{0.05067625 \times (5.59298 \times 10^{15})}}$$

$$r = \frac{5620640.28}{2224125574}$$

$r = 0.334\ (3\ s.f.)$

Here is the Python code I used to create the regression and recheck Pearson's r value:

```
"""
Created on Sun Feb 27 19:15:07 2022
@author: rigdenlama
"""
sample_df['TS%'] = sample_df['PTS'] / (2 * (sample_df['FGA']+(0.44 *
sample_df['FTA'])))
tspercentage_salary = sample_df[['TS%', 'SALARY (USD)']]

import seaborn as sns
sns.regplot(data=sample_df, x='TS%', y='SALARY (USD)', marker='+')

import numpy as np
x_values = sample_df['TS%']
y_values = sample_df['SALARY (USD)']
correlation_matrix = np.corrcoef(x_values, y_values)
r_value = correlation_matrix[0,1]
print(r_value)
```

r = 0.3338581795721239

XII. Pearson's Product Moment Correlation Coefficient Results

	Pearson's r value	Type of correlation
PTS & Salary	0.357	positive and weak to moderate correlation
AST & Salary	0.385	positive and weak to moderate correlation
STL & Salary	0.382	positive and weak to moderate correlation
EFF & Salary	0.454	positive and weak to moderate correlation
TS% & Salary	0.334	positive and weak to moderate correlation

XIII. Variance Inflation Factor (VIF) Results

Here is the Python code I used to calculate the VIF between the five independent variables (measurements of performance):

```
"""
Created on Tue Apr  5 14:51:14 2022

@author: rigdenlama
"""
From statsmodels.stats.outliers_influence import
variance_inflation_factor

X = sample_df[['PTS', 'AST', 'STL', 'EFF', 'TS%']]

vif_data = pd.DataFrame()
vif_data["Variables"] = X.columns

vif_data["VIF"] = [variance_inflation_factor(X.values, i)
                          for i in range(len(X.columns))]

print(vif_data)
```

INDEPENDENT VARIABLES	VIF
PTS	66.449913 or 66.4 (3 s.f.)
AST	27.187976 or 27.2 (3 s.f.)
STL	18.652628 or 18.7 (3 s.f.)
EFF	130.939064 or 131 (3 s.f.)
TS%	27.999154 or 30.0 (3 s.f.)

XIV. Conclusion: how does an NBA point guard's regular season performance impact their salary through the 2020-21 NBA regular season?

From the results of the Pearson's r values, an NBA point guard's regular season performance has a direct relationship with their salary and moves in tandem with each other. This is because the calculated r values show that all five performance measurements have positive correlations with point guard salaries, meaning a change in a performance statistic would result into a similar change in salary. However, the results of the Pearson's r values also show weak-to-moderate correlations between a point guard's performance and salary, meaning that a change in salary is not always caused by a change in performance. In other words, an NBA point guard's performance does have a direct impact on their salary but not to the extent that was predicted in the hypothesis (III). With highest r value being 0.454 (EFF vs. salary) and the lowest being 0.334 (TS% vs. salary), the range indicates there are also other external factors or different performance statistics that have a greater impact on the distribution of point guard salaries by NBA organizations instead of the PTS, AST, STL, EFF, and TS% statistics alone.

From the results of the VIFs, the high multicollinearity calculated between each performance statistic provides an explanation of the weak-to-moderate correlations identified in the Pearson's r values and introduces a potential insight of what NBA organizations value in their players when determining their salaries. With the highest VIF being 131 (EFF) and the lowest being 18.7 (STL), it is determined that all of the statistics have a high (severe) correlation with each other. This means that, rather than salary being dependent on performance statistics, performance statistics are more dependent on each other which is logical because an NBA point guard has many responsibilities that cannot be executed equally. Thus, the high multicollinearity lowers the r values calculated for performance and salary as the point guards must prioritize their responsibilities during their games (like whether to score more points or pass the ball more), thus increasing their numbers for a specific performance statistic and changing their numbers for the other statistics. However, this leads to another interesting conclusion that an NBA point guard's overall performance may have a greater correlation to salary rather than the specific performance statistics which are found to be more correlated with each other. For example, the EFF statistic has the highest Pearson's r value and includes points, rebounds, steals, blocks, turnovers, missed field goals, and missed free throws in its assessment of an NBA player.

Which measurement of an NBA point guard's performance has the greatest impact on their salary?

From the results of the investigation, the EFF statistic has the greatest impact of an NBA point guard's salary with a Pearson's r value of 0.454. This differs from the hypothesis (III) which predicted that the AST and TS% statistics would have a greater impact due to their specific assessments of a point guard's role and ability. However, through the VIF results, NBA organizations may value all-rounded players more than players with high numbers in certain performance statistics and therefore value statistics that have a holistic measurement of an NBA player's ability like EFF. In fact, the TS% statistic had the lowest impact with a Pearson's r value of 0.334 which I believe is due to the statistic being more catered for other basketball positions that demand more shooting and less decision-making like shooting guards.

XV. Reflection

From completing the investigation regarding the relationship between an NBA point guard's performance and salary in the 2020-21 NBA regular season, there are multiple limitations that can be observed. For example, I mainly used a Python programming language instead of the mathematical formula to calculate the VIFs for the performance measurements since I did not understand the calculation process well enough and found it very difficult. Also, calculating a VIF in the EFF statistic for multicollinearity is redundant because we already understand that EFF includes other performance statistics which would already cause high multicollinearity. In terms of the mathematical concepts employed in the investigation, I believe another limitation is that the data could not be used for statistical tests like the χ^2 test for independence because observed and expected values could not be calculated with the sample dataset.

However, I believe there are certain strengths to the investigation which has helped me increase my overall knowledge of the NBA. For example, the results of the investigation would be useful for NBA fans that are interested in how NBA player contracts are issued by NBA organizations without having any prior knowledge as well as how significant of a factor player performance is in their decision-making process.

Given the completion of this preliminary research, I have a greater understanding of the relationship between player performance and salary as well as which specific performance metrics are relevant to NBA point guards as I will transition the focus of the research to creating a prediction method to estimate an NBA point guard's salary by answering the following questions:

1) **Are NBA point guard salaries distributed accordingly to their performance data?**
2) **How can an NBA point guard's salary be predicted based on their performance data?**

Citations

1. ESPN. "ESPN: Serving Sports Fans. Anytime. Anywhere." *ESPN*, www.espn.com/nba/stats/player/_/season/2021/seasontype/2/position/point-guard.
2. HoopsHype. "These Are the Salaries of All NBA Teams." *HoopsHype*, hoopshype.com/salaries.
3. Investopedia. "Variance Inflation Factor (VIF)." *Investopedia*, 31 Oct. 2021, www.investopedia.com/terms/v/variance-inflation-factor.asp