

NBA GM: When to Extend or Let Go

Introduction:

The user story that inspired this report asked what factors would be useful to determine when a player has reached their peak performance. Additionally, what factors would lead the organization to extend a current contract or to let the player leave the organization.

The dataset included so much data that I decided on a method that could focus on 5 individuals and I would act as the General Manager(GM), making the decision on whether to extend the contract or let the player leave.

I have randomly selected 5 players from the dataset that each had at least 10 entries. I will look at each player and gather their;

Games Played(GP)
Minutes(MIN)
Turnovers(TOV)
Personal Fouls(PF)
Rebounds(REB)
Assists(AST)
Points(PTS)

As GM, these are metrics I value most. I will then analyze each player to determine their Peak, when to extend the contract or decline at let the player leave through Free Agency. I decided on a length of 3 years for metric to decide when it would a decision would need to be made whether to extend or decline a new contract.

NOTE:

When I originally ran this the 5 players who were returned were:
Glen Rice, Andrea Bargnani, Voshon Lenard, Kenny Anderson and B.J. Armstrong
These are the 5 players that I will use for this analysis.

Body:

Data:

The dataset included records beginning with the 1950 season through the 2021 season. There were numerous metrics included but as I've listed above, I only dealt with ones I felt most relevant to this analysis. While cleaning the data, I found there were no significant issues with the data. The only issues I ran into with the data involved 2 of the players I had selected having played for multiple teams in the same season and needed a way to combine that data.

Process:

As I was only working with 5 players, my goal was to create a data frame for each player and use that information for my analysis. Thereby I will relay each step of my analysis using only the first player, Glen Rice, as the extra details about each player would be redundant. Firstly I created a data frame with the desired categories mentioned above.

```
1 glen_rice_data[['Year', 'GP', 'MIN', 'TOV', 'PF', 'REB', 'AST', 'PTS']]
```

Initially I thought using the aggregate data would be fine but decided to look at averages because certain players playing less games wouldn't be as useful for comparison.

```
glen_rice_data['MIN/GP'] = glen_rice_data['MIN'] / glen_rice_data['GP']
glen_rice_data['TOV/GP'] = glen_rice_data['TOV'] / glen_rice_data['GP']
glen_rice_data['PF/GP'] = glen_rice_data['PF'] / glen_rice_data['GP']
glen_rice_data['REB/GP'] = glen_rice_data['REB'] / glen_rice_data['GP']
glen_rice_data['AST/GP'] = glen_rice_data['AST'] / glen_rice_data['GP']
glen_rice_data['PTS/GP'] = glen_rice_data['PTS'] / glen_rice_data['GP']

glen_rice_data[['Year', 'GP', 'MIN/GP', 'TOV/GP', 'PF/GP', 'REB/GP', 'AST/GP', 'PTS/GP']].round(2)
```

I used these averages to determine what year each player peaked by visualizing inspecting the data and creating GM Decision.

GM Decision

Peak

Rice hit his statistical peak during the 1997 season. His Minutes per Game(Min/GP), Points per Game(PTS/GP) and Games Played(GP) were all at or near their highest during his career. Though he did have higher averages in REB and AST in other seasons, the 1997 season averages were not too far off.

Extend?

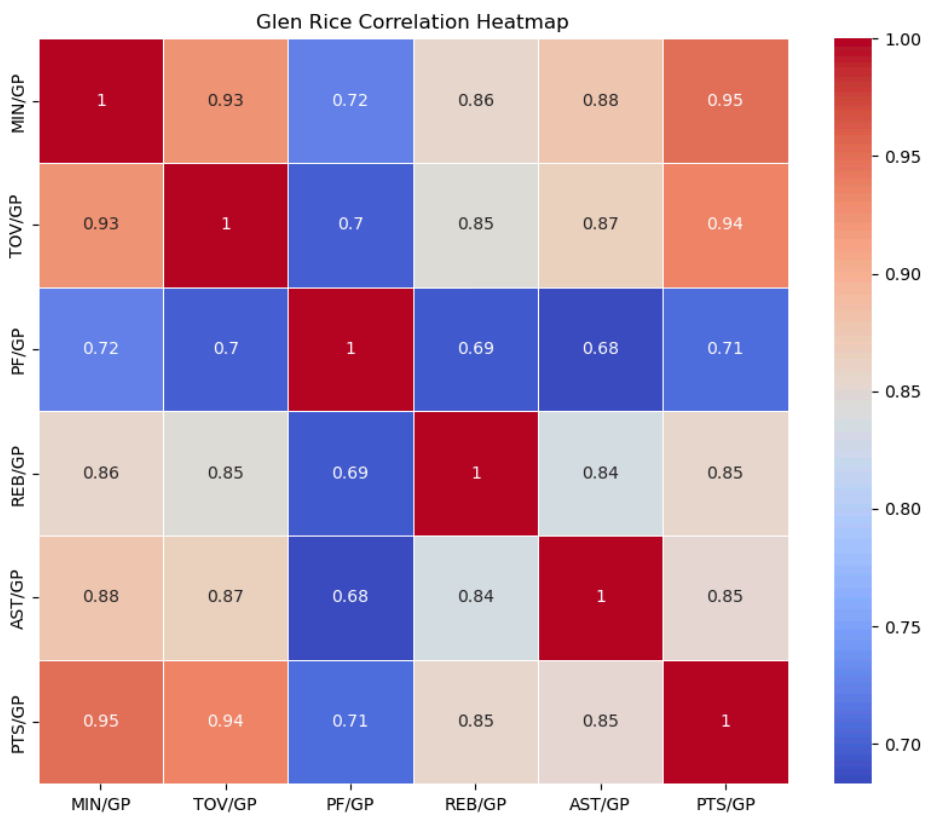
From his initial season in 1990 through 1992 there was significant positive increases across all my valued metrics. Happily, I'll extend him. Though there was a slight dip from 1992 to 1993, in 1994 and 1995 there was more positive increases across all categories. We will extend him, again. During the 1996 season Rice remained consistent and reached his statistical Peak during the 1997 season. 1998 was saw a slight decrease in his overall numbers but we will extend him again based the past three seasons. The decline for Rice began during the 1999 season, this looks likely due to an injury or other issue that led him to only play in 27 games where in all seasons prior he had played in at least 77. In 2000 and 2001 Rice continued his decline and sadly, his time with our team has come to an end.

Let Him Go?

Rice will be let go after the 2001 season. His GP remained good but all my other metrics showed decline. This decision was backed up by the remaining years of his career where he never came close to reaching his peak performance of the 1997 season.

Finally, I created a correlation matrix heat map that could give a great visual understating of whether or not certain metrics had a strong or weak correlation.

```
1 # Create a heatmap of the correlation matrix
2
3 plt.figure(figsize=(10, 8))
4 sns.heatmap(rice_corr_matrix, annot=True, cmap='coolwarm', linewidths=.5)
5 plt.title('Glen Rice Correlation Heatmap')
6 plt.show()
```



Results:

As you can see above, Glen Rice has a strong correlation between playing more minutes in each game and PTS/GP, AST/GP, REB/GP, TOV/ GP. These numbers make sense as a player playing more minutes has a better chance to produce.