Time Series Analysis on QPR's performances in the EFL Championship since 2015

George McNamara

This project is to use statistical models and time series analysis to measure the success of Queens Park Rangers since the football club were relegated from the Premier League to the Championship in the summer of 2015.

In the few years preceding this relegation, QPR were as dramatic on the pitch as off it having seen two relegations, survival on the final day coupled with Manchester City's first title in 46 years and the infamous late Sergio Aguero winner and a playoff victory to see the club return to the Premier League at the first time of asking due to a late winner from Bobby Zamora. Off the pitch, QPR spent riches to rival the biggest clubs in England, spending large wages on former Champions League winners such as Julio Cesar and Jose Bosingwa and a massive €15 million on defender Christopher Samba - only to sell him at the next opportunity for a €3.4 million loss. Despite all the spending by the club's hierarchy and management, the club were demoted back to the Championship during the 2014/2015 season picking up a grim 30 points seeing the club finish bottom of the league. Nine years on, QPR still find themselves in the English second tier.

In this time, QPR have had nine different managers take charge of the club on the pitch all with large variations in success - from the brink of relegation to the English third tier to being close to returning to the Premier League and then back to the threaten of relegation.

Ensuing will be a comparison of the QPR bosses via contrasting basic football statistics and using time series analysis to determine overall performance of QPR on the pitch whilst a certain manager was or is in charge.

Currently, the Spaniard Marti Cifuentes is the QPR manager having been named in that role on the 30th of November 2023 but has his naming of manager been successful in the eyes of statisticians or have the club made a mistake in the past letting a previous manager go? Should QPR have done more to keep Michael Beale from the clutches of Glasgow Rangers or should the club have given more time to Chris Ramsey following their relegation? And should Neil Critchley have been given more than eleven games in charge? Hopefully the following work should help answer these questions.

Loading in relevant libraries and the data

```
library(tseries)
library(astsa)
library(forecast)
```

```
library(readxl)
QPR_data <- read_excel("QPR_data.xlsx")
```

The first three lines of code allows us to load in the necessary libraries that will be used in the analysis of the data. Following that, the excel file which includes the data is loaded in and stored under the variable name 'QPR_data' which can be viewed in the table.

Creating basic linear models and graphs to compare the managers of QPR

```
ManagerPPG <- c(1.27,1.26,1.17,1.13,1.39,1.48,0.73,0.7,1.5)
ManagerWinPercentage <- c(0.333,0.279,0.329,0.308,0.391,0.429,0.091,0.185,0.406)
ManagerCSpg <- c(0.2,0.302,0.132,0.308,0.232,0.238,0.182,0.148,0.406)
Moneyspent <- c(13130000,13640000,4220000,0,9010000,0,0,0,0)
Moneyreceived <- c(600000,17830000,4250000,4000000,26986000,900000,0,3110000,0)
```

When comparing the last nine managers of Queens Park Rangers, many statistics can be used. In this project, the manager's Points per game (abbreviated to PPG), Win Percentage, Clean sheets per game, Money spent and money received were all used to compare manager's to determine their level of success. The first data point in each vector corresponds to QPR's first manager since relegation to the Championship for the 2015/2016 football season which is Chris Ramsey, the second data point corresponds to QPR's second manager being Jimmy Floyd Hasselbaink and so on till Marti Cifuentes' appointment as manager who is still in charge at the time of writing.

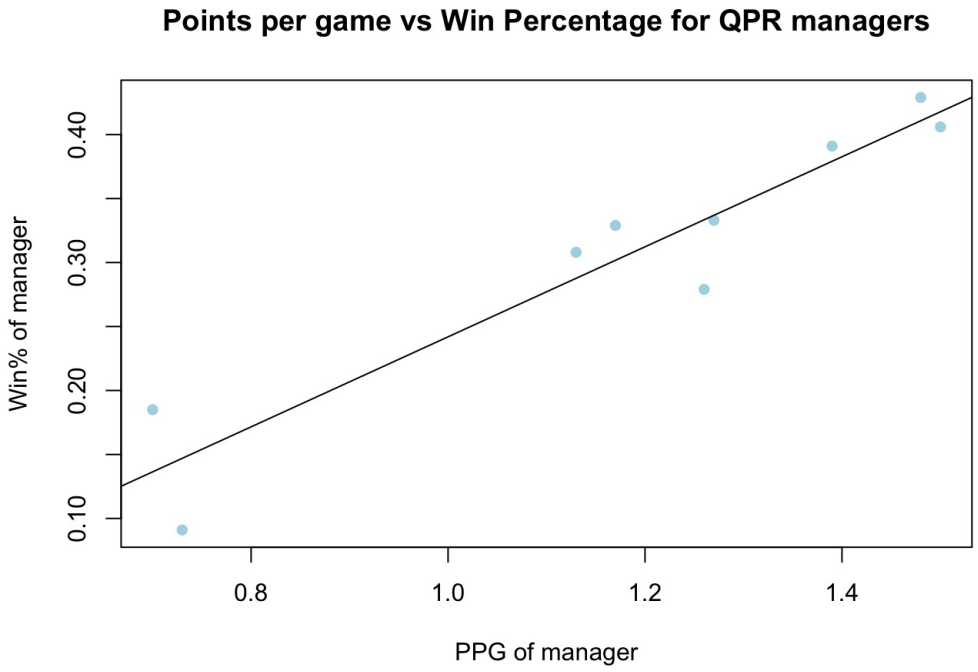
Firstly, it would be useful to check whether PPG, Win percentage and Clean sheets per game are well correlated to potentially show that a high score in one statistic could correlate to a high score in another.

Running Linear models

```
PPGvsWin <- lm(ManagerPPG~ManagerWinPercentage)
summary(PPGvsWin)
```

```
##
## Call:
## lm(formula = ManagerPPG ~ ManagerWinPercentage)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17501 -0.05703 -0.00758  0.06437  0.14653
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           0.4057     0.1073   3.780 0.006890 **
## ManagerWinPercentage  2.5367     0.3328   7.623 0.000124 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1027 on 7 degrees of freedom
## Multiple R-squared:  0.8925, Adjusted R-squared:  0.8771
## F-statistic: 58.11 on 1 and 7 DF,  p-value: 0.0001239
```

```
plot(x=ManagerPPG,y=ManagerWinPercentage,xlab='PPG of manager',ylab='Win% of manager',main='Points per game vs Wi
n Percentage for QPR managers',pch=16,col="lightblue")
abline(lm(ManagerWinPercentage~ManagerPPG))
```




```
summary(PPGvsWin)$r.squared
```

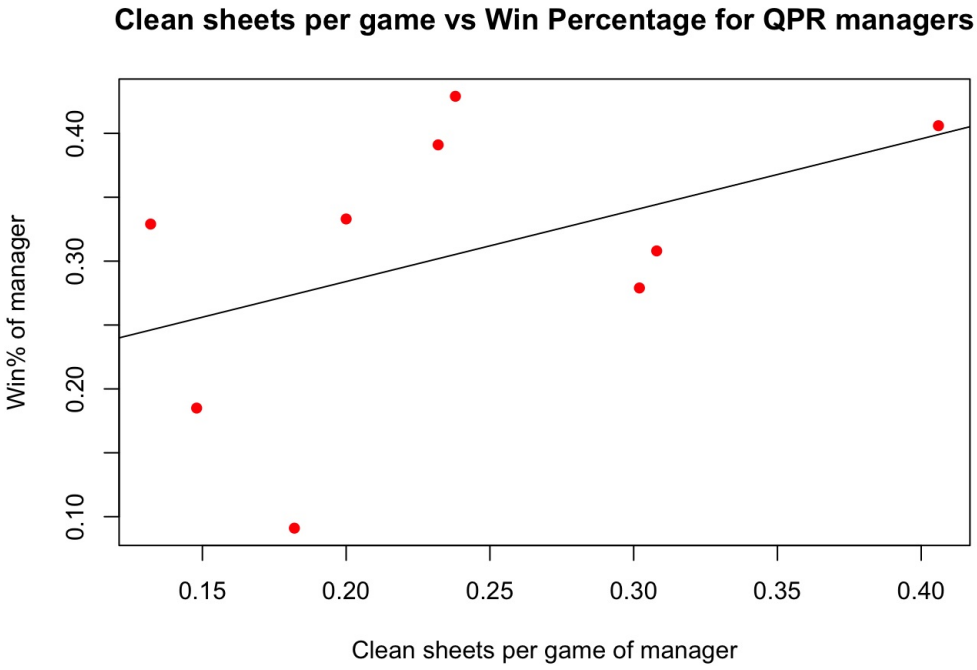
```
## [1] 0.8924874
```

The following linear model shows that there is a positive correlation between PPG and Win percentage due to the upward sloping regression line shown on the graph - which has a formula of y=0.3518x-0.1099 as shown in ExCel. Additionally, the coefficient of determination is very close to one since it is equal to 0.8924874 which means that PPG and Win percentage are well correlated which indicates a good fit for the model.

```
WinvsCS <- lm(ManagerWinPercentage~ManagerCSpg)
summary(WinvsCS)
```

```
##
## Call:
## lm(formula = ManagerWinPercentage ~ ManagerCSpg)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.183011 -0.062046  0.006857  0.082920  0.123706
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.1723     0.1065   1.619    0.150
## ManagerCSpg   0.5586     0.4216   1.325    0.227
##
## Residual standard error: 0.1043 on 7 degrees of freedom
## Multiple R-squared:  0.2005, Adjusted R-squared:  0.08627
## F-statistic: 1.755 on 1 and 7 DF,  p-value: 0.2268
```

```
plot(x=ManagerCSpg,y=ManagerWinPercentage,xlab='Clean sheets per game of manager',ylab='Win% of manager',main='Cl
ean sheets per game vs Win Percentage for QPR managers',pch=16,col="red")
abline(lm(ManagerWinPercentage~ManagerCSpg))
```




```
summary(WinvsCS)$r.squared
```

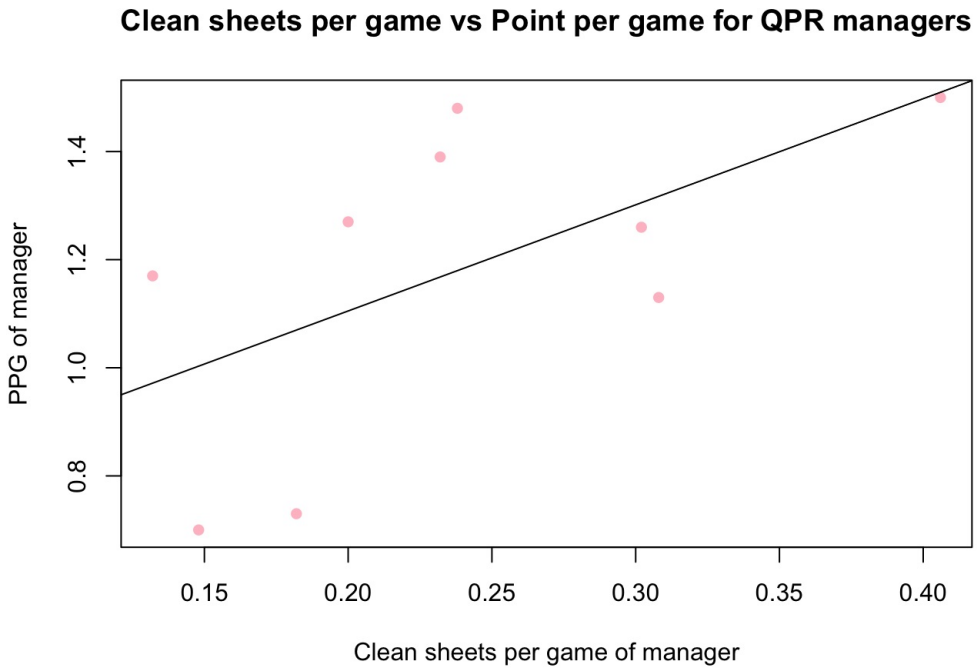
```
## [1] 0.2004856
```

The next linear model shows the relationship between Clean sheets per game and Win percentage, demonstrating a positive correlation due to the upward sloping regression line, although with a weaker gradient compared to the previous linear model. Also, the coefficient of determination is significantly smaller at 0.2004856, revealing a weak fit for the model. This indicates that keeping a clean sheet does not necessarily lead to a win and vice versa.

```
PPGvsCS <- lm(ManagerPPG~ManagerCSpg)
summary(PPGvsCS)
```

```
##
## Call:
## lm(formula = ManagerPPG ~ ManagerCSpg)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33980 -0.18730 -0.00981  0.19842  0.30020
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.7123     0.2590   2.750   0.0285 *
## ManagerCSpg   1.9643     1.0257   1.915   0.0970 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2537 on 7 degrees of freedom
## Multiple R-squared:  0.3438, Adjusted R-squared:  0.2501
## F-statistic: 3.668 on 1 and 7 DF,  p-value: 0.09703
```

```
plot(x=ManagerCSpg,y=ManagerPPG,xlab='Clean sheets per game of manager',ylab='PPG of manager',main='Clean sheets per game vs Point per game for QPR managers',pch=16,col="pink")
abline(lm(ManagerPPG~ManagerCSpg))
```




```
summary(PPGvsCS)$r.squared
```

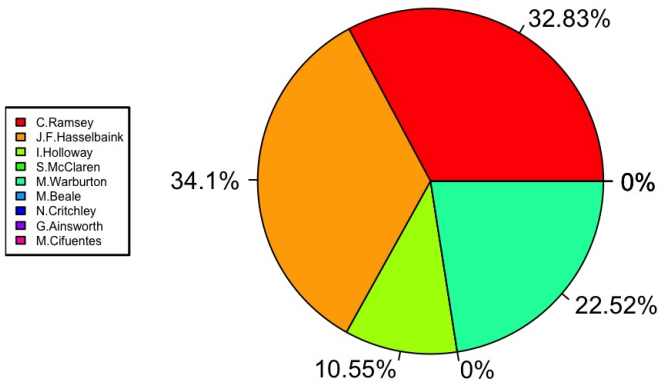
```
## [1] 0.3438063
```

The final linear model shows the positive correlations between Clean sheets per game and PPG for the QPR managers. This is due to the upward sloping regression line, meaning that keeping a clean sheet is more likely to result in points (this is true since keeping a clean sheet in football would result in minimum a draw and therefore one point at worst).Although, the coefficient of determination is low at 0.3438063 signalling a weak fit between the two variables but stronger than that between Clean sheets per game and Win percentage (which is to be expected).

Overall,this shows that keeping clean sheets and building a strong defense does not automatically lead to more points and/or more wins. Furthermore, PPG and Win Percentage indicate a strong relationship - which is to be predicted since winning equals three points (the highest amount of points possible in a game of football).

Graphs to compare Manager's performance

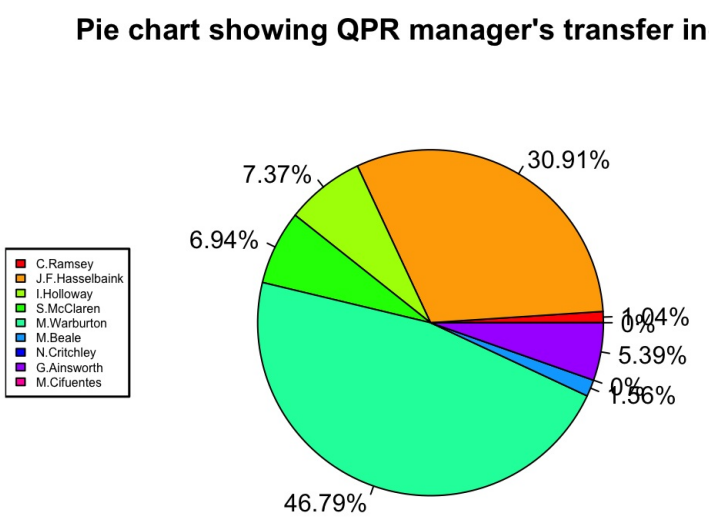
```
lbls <- c("C.Ramsey","J.F.Hasselbaink","I.Holloway","S.McClaren","M.Warburton","M.Beale","N.Critchley","G.Ainsworth","M.Cifuentes")
pctspend <- paste0(round(100*Moneyspent/sum(Moneyspent),2),"%")
pie(Moneyspent,labels=pctspend,main="Pie chart showing QPR manager's transfer expenditure",col=rainbow(length(Moneyspent)))
legend(x="left",box.lwd=1.5,legend=lbls,cex=0.5,col=rainbow(length(Moneyspent)),fill=rainbow(length(Moneyspent)))
```



The first pie chart shows the two biggest spending managers are the first two QPR managers since returning to the Championship Chris Ramsey and Jimmy Floyd Hasselbaink having spent 32.83% and 34.1% of QPR's spending since their relegation respectively, this is because of the enormous parachute payments from the Premier League following relegation which gives QPR more money to potentially spend on transfers and a greater ambition to return to the top league in England sooner rather than later. This is then followed by Mark Warburton, possibly due to him being the longest serving QPR manager in this time (resulting in more time to spend) and his high transfer income permitting more money to be spent.

During this nine year period five managers have not spent a penny on transfers, this may be due to QPR's dire financial situation due to not being in the Premier League or lack of time in the job (in Neil Critchley's case, having only two months and one transfer window as QPR manager resulting in little time to spend).

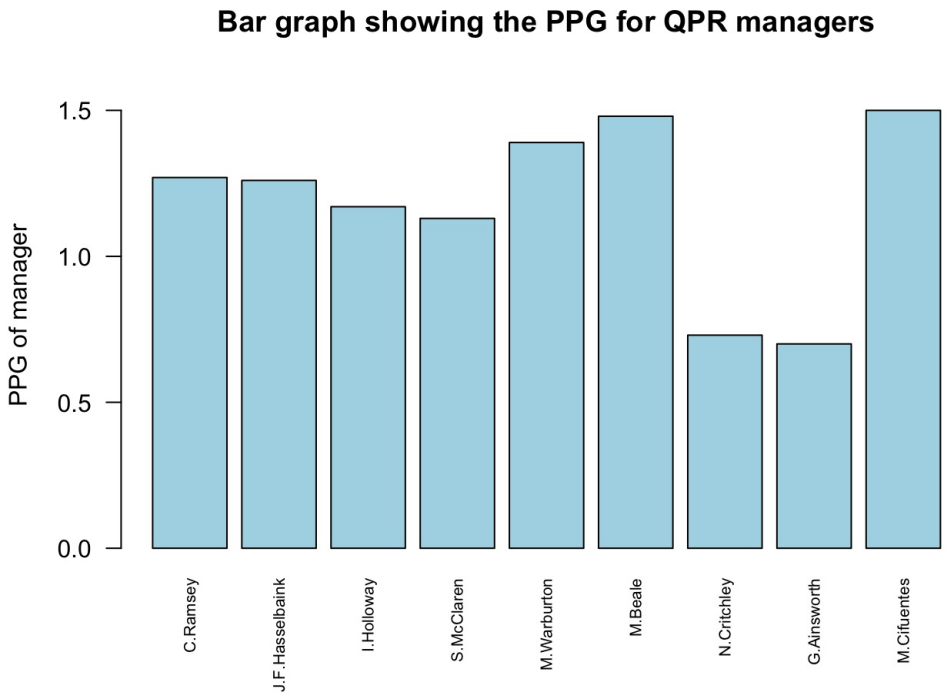
```
pctreceived <- paste0(round(100*Moneyreceived/sum(Moneyreceived),2),"%")
pie(Moneyreceived,labels=pctreceived,main="Pie chart showing QPR manager's transfer income",col=rainbow(length(Mo
neyreceived)))
legend(x="left",box.lwd=1.5,legend=lbls,cex=0.5,col=rainbow(length(Moneyreceived)),fill=rainbow(length(Moneyrecei
ved)))
```



The second pie chart shows the managers to bring in the most transfer income are Mark Warburton and Jimmy Floyd Hasselbaink. For Mark Warburton, this is on account of large sales of Eberechi Eze to Crystal Palace for €17.8 million and Luke Freeman to Sheffield United for €5.6 million - signifying an amazing transfer profit since the two players combined were bought for €350,000.Furthermore, Mark Warburton has generated nearly half of QPR's transfer income in the nine years the club have spent in the division despite being in charge for only a third of that time. In terms of J.F.Hasselbaink's time in charge, he took over a newly relegated Premier league team with Premier League quality players, meaning large fees could be fetched for said players from top flight clubs (such as the sales of Matt Phillips to West Bromwich Albion for €6.5 million, Leroy Fer to Swansea City for €5.6 million and Charlie Austin to Southampton FC for €5.2 million).

It should be said that a manager who spends more money should result in more success (shown in high PPG and high Win Percentage) since higher transfer expenditure should lead to a better assembled squad leading to more success on the pitch. However, this cannot be said as will be shown by comparing the manager's statistics.

```
barplot(ManagerPPG,ylab="PPG of manager",names.arg = lbls,main="Bar graph showing the PPG for QPR managers",col="
lightblue",las=2,ylim=c(0,1.6),cex.names = 0.65)
```

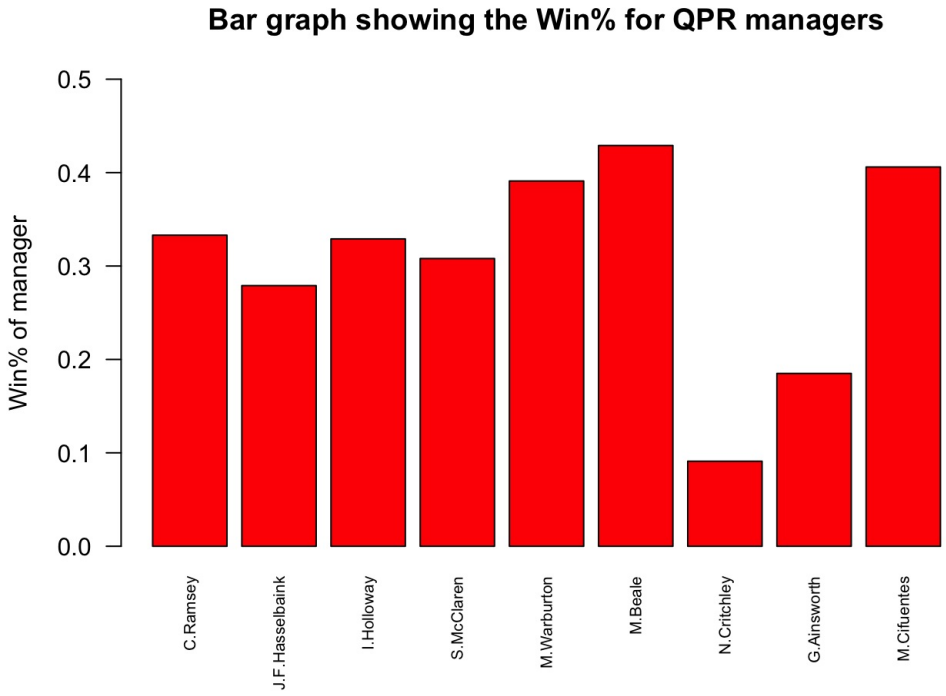


This bar graph shows the PPG of each QPR manager in the last nine years. The highest PPG are Marti Cifuentes followed by Michael Beale demonstrating the great success of QPR - especially recently since M.Cifuentes is the current manager and has turned the club around from the previous manager Gareth Ainsworth who happened to have the lowest PPG of 0.7 less than half of his successor's PPG of 1.5.

However, M.Cifuentes and M.Beale have only managed 32 and 21 games as manager respectively - so a smaller sample space would result in a large variation for PPG - similarly the two smallest PPG stats belong to Neil Critchley and Gareth Ainsworth who only managed 11 and 27 games in charge respectively. What is impressive is Mark Warburton's 138 games managed as QPR manager whilst maintaining a high and respectable PPG of 1.39.

Overall though, Marti Cifuentes can be seen to be the best manager of the club's era in the English second tier in terms of Points per game.

```
barplot(ManagerWinPercentage,ylab="Win% of manager",names.arg = lbls,main="Bar graph showing the Win% for QPR man
agers",col="red",las=2,ylim=c(0,0.5),cex.names = 0.65)
```

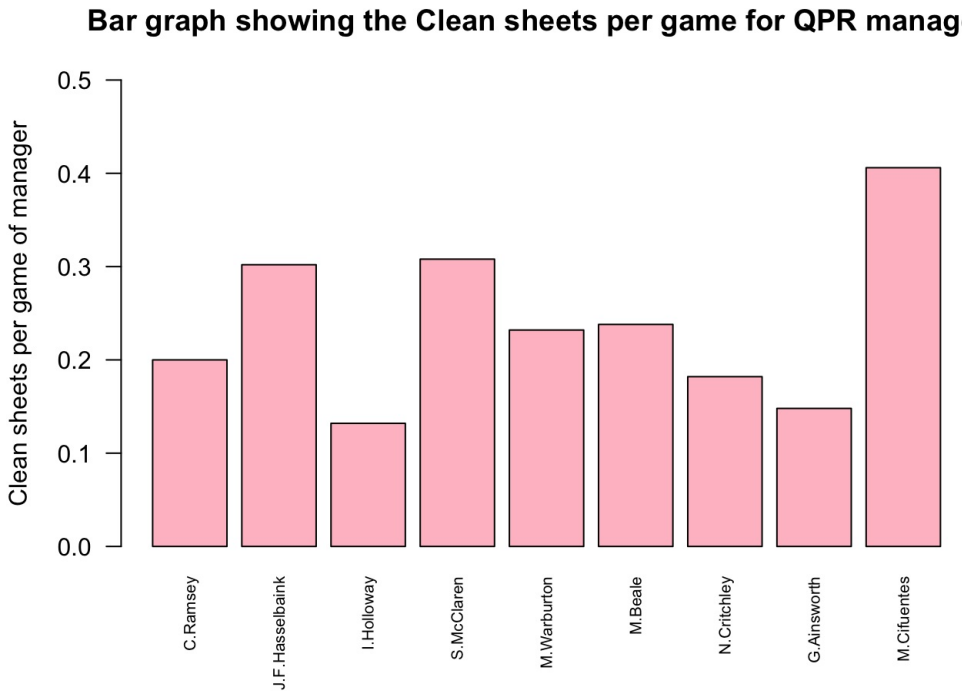


The next bar graph shows the Win percentage of each QPR manager. Once again the highest two win percentages belong to Michael Beale and Marti Cifuentes with 42.9% and 40.6% respectively - however the same limitations of lack of games managed still apply which in some eyes could hinder their level of success.

On the opposite end of the spectrum, Neil Critchley has a win percentage of only 9.1% not only making him have the lowest win percentage of the nine managers in question but also having the lowest win percentage of the 56 people to manage QPR in their 142 year history. Although, he was only in charge for eleven games.

In general, M.Beale does categorically have the highest Win percentage of the nine managers although both M.Cifuentes and M.Warburton are not far behind with win percentages of 40.6% and 39.1% respectively.

```
barplot(ManagerCSpg,ylab="Clean sheets per game of manager",names.arg = lbls,main="Bar graph showing the Clean sh
eets per game for QPR managers",col="pink",las=2,ylim=c(0,0.5),cex.names = 0.65)
```



The concluding bar graph shows the Clean sheets per game of each QPR manager. It is Marti Cifuentes who has by far the highest clean sheets per game statistic at 0.406 - significantly higher than any QPR manager in this time showing how M.Cifuentes is the best defensive manager in recent years for the football club. Next by some distance is Steve McClaren at 0.308 clean sheets per game followed by Jimmy Floyd Hasselbaink at 0.302 clean sheets per game - showing both managers to be good at structuring a defense but not a getting points or winning football matches.

Contrarily, the lowest clean sheets per game statistic surprisingly does not belong to Neil Critchley or Gareth Ainsworth but to Ian Holloway - that is despite spending €4.22 million on transfers compared to the aforementioned two managers not spending any money in the way of new signings.

In conclusion, this part of the study indicates that Marti Cifuentes has been the superior manager of QPR in the nine seasons that the club has played in the Championship. This is due to Cifuentes having the highest PPG and Clean sheets per game of any manager and having the second highest win percentage narrowly falling short to Michael Beale. This has all been completed despite Cifuentes not spending any money on transfers since arriving from Hammarby in Sweden to become the QPR manager in late 2023.

Although, it may be short-sighted to judge his performances and statistics having only managed 32 games in the Championship as QPR boss (which is lower than Chris Ramsey, Jimmy Floyd Hasselbaink, Ian Holloway, Steve McClaren and Mark Warburton all managed) his early success as manager having taken QPR from a potential relegation to England's third tier to a comfortable 18th place finish in the second tier shows that QPR are once again on an upward trajectory and could possibly challenge for promotion in the upcoming 2024/2025 Championship season as long as Cifuentes stays on as the manager.

Time Series Analysis: Processing the data

```
PPG <- QPR_data$PPG
TimeseriesPPG <- ts(data=PPG,frequency = 12)
frequency(TimeseriesPPG)
```

```
## [1] 12
```

```
deltat(TimeseriesPPG)
```

```
## [1] 0.08333333
```

From the ExCel file is the monthly statistics for PPG and Win Percentage, in this project we will look at the monthly PPG since in my opinion it is the best method in determining the success of the club in said month. Extracting the data and storing in under the variable 'PPG' followed by transforming it into a time series with yearly seasonality since the 12 month period encompasses a season in the English second tier. Some statistics that can be checked are the frequency of the time series (which is 12) and the time in between each observation.

However, in a year, the football season only takes places for 10 months (even 9 months during the 2020/2021 season as a result on the backlog from the previous Covid-19 interrupted season). This means that there are multiple blanks in the data that need to be removed in order to properly analyse the time series.

Removing NAs

```
RemovedPPG <- na.omit(PPG)
NewTSPPG <- ts(data=RemovedPPG,frequency=10)
frequency(NewTSPPG)
```

```
## [1] 10
```

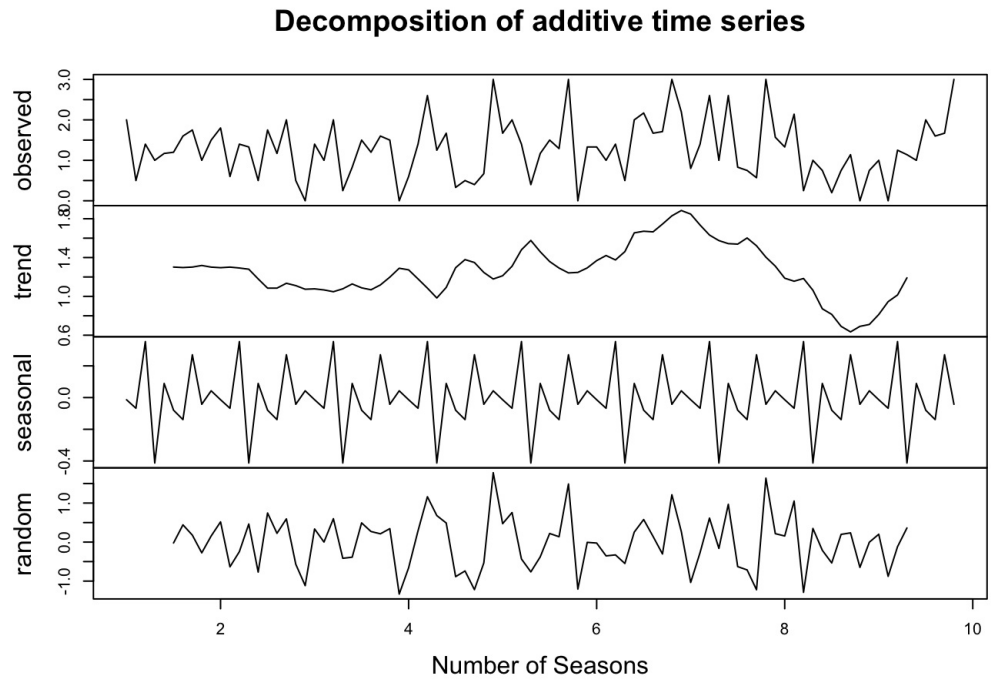
```
deltat(NewTSPPG)
```

```
## [1] 0.1
```

The following chunk of code removes all the blank observations corresponding to the time when games are not taking place. This lowers the number of observations from 108 to 89. Furthermore, the seasonality has now been reduced to repeating itself every 10 observations since in the last 8 of 9 seasons, there have been 10 months in the football calendar.

Time Series Analysis:Splitting the observed data

```
plot(decompose(NewTSPPG),xlab="Number of Seasons")
```



The Time Series is an additive model that splits the observed data into three parts, trend, seasonality and residual error. The above code decomposes our time series 'NewTSPPG' into the three parts and plots it.

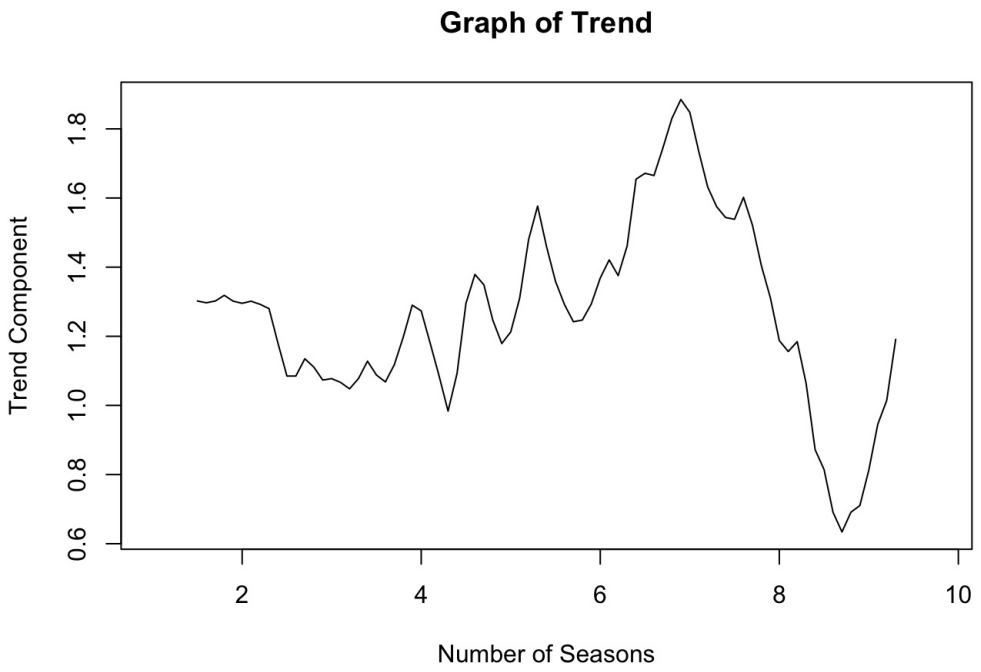
Reviewing the graphs shows stable trend that is slowly increasing and peaking at the end of the 7th season coinciding with M.Warburton's sacking and M.Beale's appointment (which is accurate since we saw that Michael Beale and Mark Warburton had the second and third highest PPG's respectively) before rapidly plummeting under N.Critchley before reaching its lowest point under G.Ainsworth. However, the Cifuentes era shows a recovery and an increase in trend (as shown by Marti Cifuentes having the highest PPG at 1.5 points per game).

The seasonality graph shows clear seasonality throughout QPR seasons - this can possibly be explained by May having only one fixture resulting in either the maximum or minimum (3 or 0) PPG given for that month. The graph also shows that QPR tend to do very well in August and therefore start the season off very well but have the worst month being September, halting any progress made in the previous month.

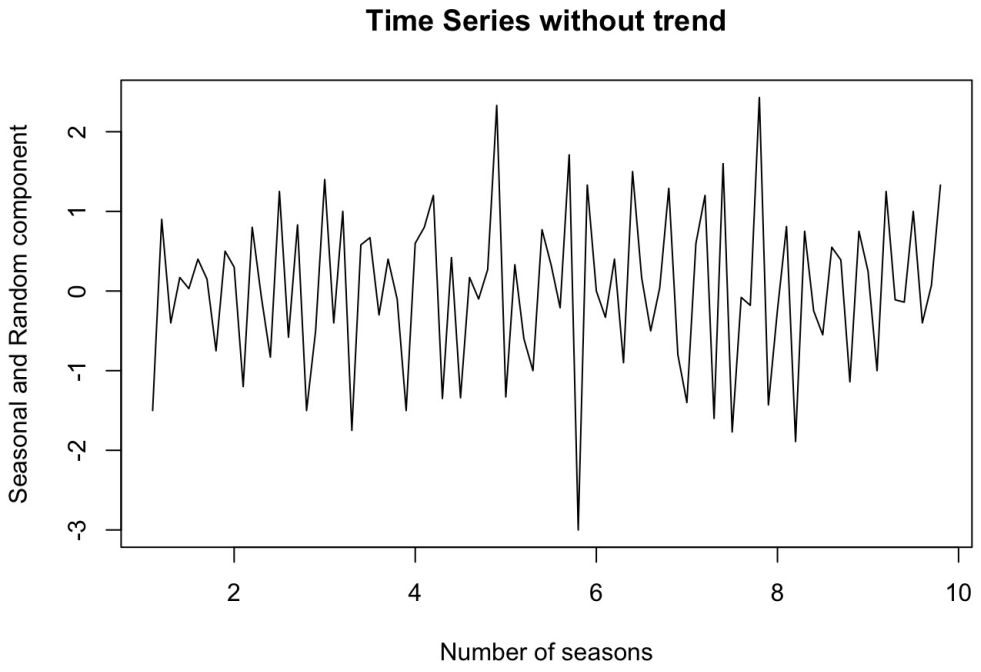
The graph showing residual error shows a large variance in PPG particularly around the middle of the graph demonstrating inconsistency in QPR's performance during S.McClaren's and M.Warburton's first few years as manager of QPR. QPR do start to become more consistent under Marti Cifuentes showing the club finally stabilizing.

Focusing on Trend

```
TrendPPG <- decompose(NewTSPPG)$trend
plot(TrendPPG,xlab="Number of Seasons",ylab="Trend Component",main="Graph of Trend")
```

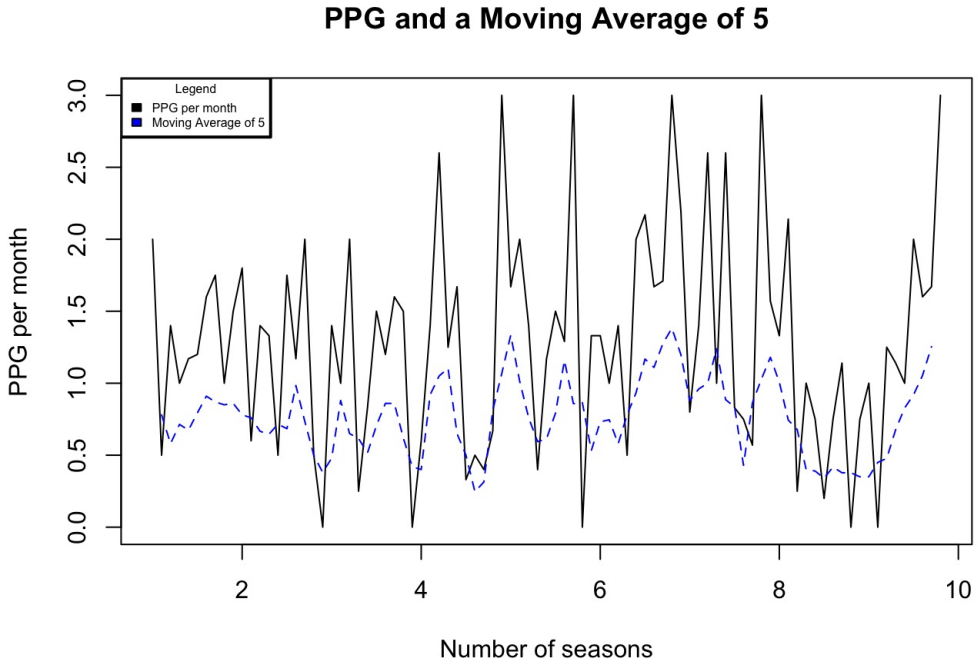



```
plot(diff(NewTSPPG),xlab="Number of seasons",ylab="Seasonal and Random component",main="Time Series without trend
")
```



The above graph shows the time series without any trend, which shows a large variance and exhibits randomness. Overall, this graph is not very helpful and it would be better to show a moving average to display a smoother trend and less variance.

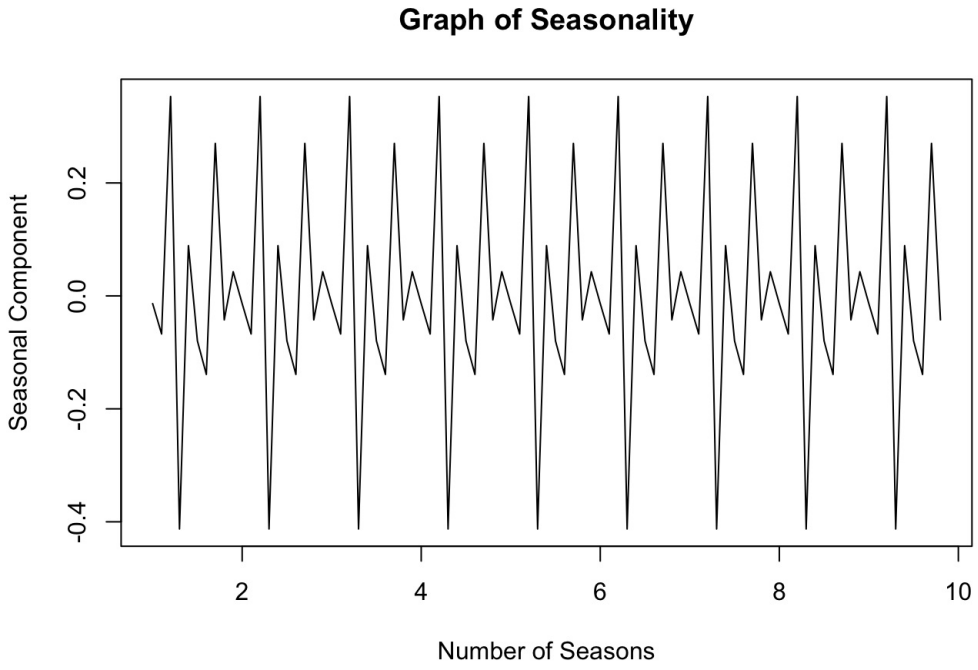
```
MA <- 5
MA_5 <- filter(NewTSPPG,c(1/MA,1/MA,1/MA))
plot(NewTSPPG,type="l",xlab="Number of seasons",ylab="PPG per month",main="PPG and a Moving Average of 5")
lines(MA_5,lty=2,lwd=1,col="blue")
legend(x="topleft",box.lwd = 2,title="Legend",legend=c("PPG per month","Moving Average of 5"),fill=c("black","blu
e"),cex=0.5)
```



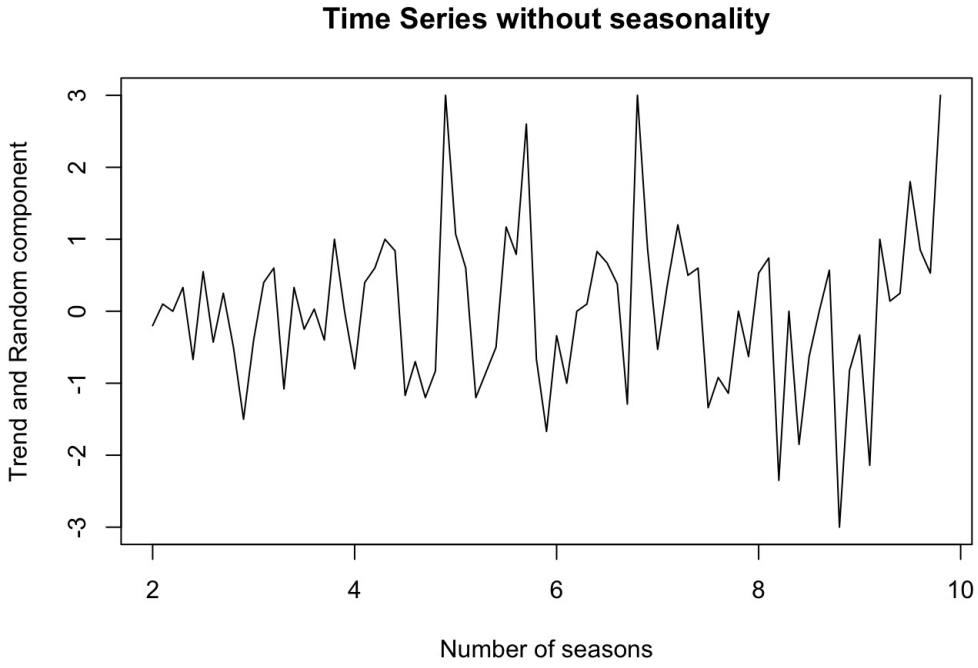
A moving average of 5 is created to reduce the variance of the trend to get a proper look at QPR's success and performance over this time period. As you can see, QPR's greatest success whilst maintaining consistency is under Mark Warburton's stewardship althought if Marti Cifuentes continues his form as the manager, it will show a new peak for this moving average. The worst performances are shown to be whilst N.Critchley was in charge (lowering the moving average at a steep gradient from Michael Beale's brief success) and under the final games whilst Ian Holloway was in charge at the club. In conclusion, the moving average is a much better demonstration of trend.

Focusing on Seasonality

```
SeasonalPPG <- decompose(NewTSPPG)$seasonal
plot(SeasonalPPG,xlab="Number of Seasons",ylab="Seasonal Component",main="Graph of Seasonality")
```



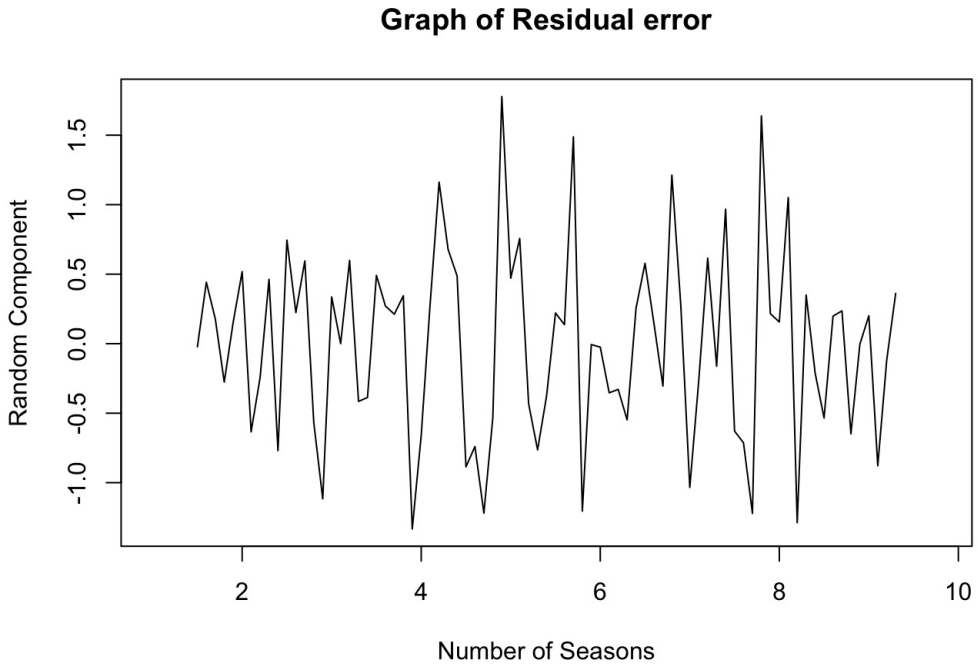

```
plot(diff(NewTSPPG,lag=10),xlab="Number of seasons",ylab="Trend and Random component",main="Time Series without s
easonality")
```



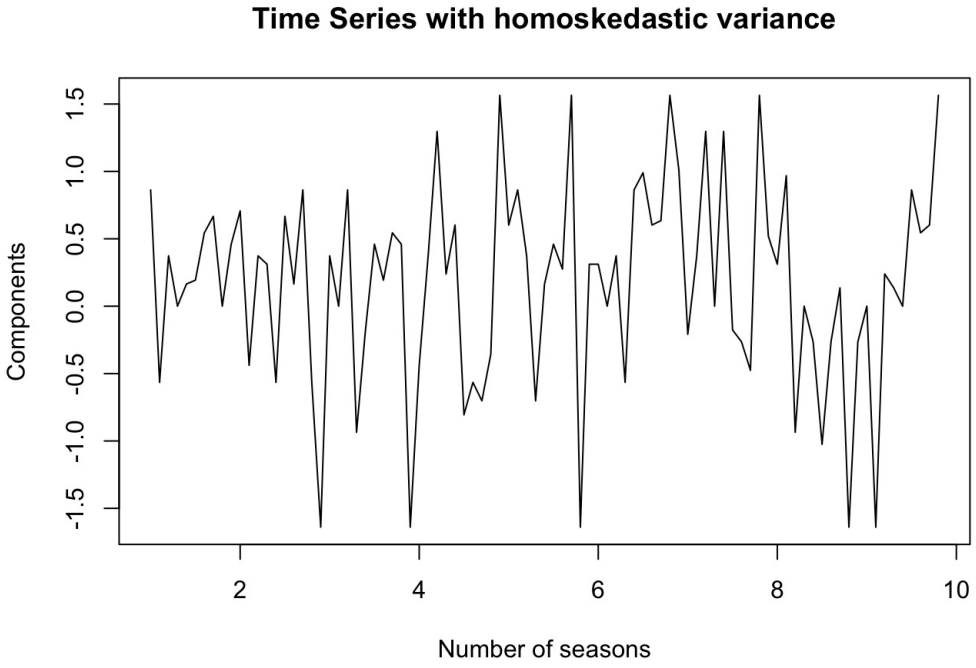
This graph shows the time series without any seasonality. This indicates that QPR's greatest successes have come whilst Mark Warburton, Michael Beale or Marti Cifuentes have been at the helm. Also, the most dire times have come under Gareth Ainsworth. Other than that, the graph represents how QPR have been very average since returning to the Championship in 2015, backed up by QPR neither getting promoted or relegated and being one of the three longest serving clubs in the English second tier along with Preston North End and Bristol City as of the start of the 2024/2025 season.

Focusing on Residual Error

```
RandomPPG <- decompose(NewTSPPG)$random
plot(RandomPPG,xlab="Number of Seasons",ylab="Random Component",main="Graph of Residual error")
```




```
lambda <- BoxCox.lambda(NewTSPPG)
BCtransformation <- BoxCox(NewTSPPG,lambda)
plot(BCtransformation,type="l",xlab="Number of seasons",ylab="Components" ,main="Time Series with homoskedastic v
ariance")
```



The following is a Box-Cox transformation that can change a time series with possible heteroskedastic variance to homoskedastic variance - meaning that the variance becomes stationary and stays the same throughout that shows a more consistent QPR side.

The graph with stationary variance enforces that the worst performances have come during N.Critchley or G.Ainsworth have been in charge of the club whilst the best performances and most consistent results have been under M.Beale or M.Warburton.

Testing for a stationary variance

```
adf.test(NewTSPPG)
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  NewTSPPG
## Dickey-Fuller = -3.4236, Lag order = 4, p-value = 0.05629
## alternative hypothesis: stationary
```

```
kpss.test(NewTSPPG)
```

```
## Warning in kpss.test(NewTSPPG): p-value greater than printed p-value
```

```
##
##  KPSS Test for Level Stationarity
##
## data:  NewTSPPG
## KPSS Level = 0.1009, Truncation lag parameter = 3, p-value = 0.1
```

```
Box.test(RandomPPG,lag=10,type="Ljung")
```

```
##
##  Box-Ljung test
##
## data:  RandomPPG
## X-squared = 26.888, df = 10, p-value = 0.002712
```

The Augmented Dickey-Fuller test and Kwiatkowski-Phillips-Schmidt-Shin test are to solve whether a Time Series has a stationary variance. Both tests determine that the time series show a stationary variance - indicating that over this selected nine year period, QPR have been overall consistent in terms of their points per game.

The Ljung-Box test is to govern whether the residuals of a time series are white noise and therefore have no correlations or if the residuals are not white noise and have some non-trivial correlation. The output of this test shows that the residuals are not white noise and do have some correlations - enforcing the idea of consistency.

Conclusion

To conclude, QPR have had a rollercoaster of a ride during their nine year stay in the Championship with the highs of Michael Beale's short but sweet reign taking the club as far as top of the league on the 22nd of October 2022 to not so long after winning 2 of 22 matches after Michael Beale's sudden departure to north of the border which easily could have resulted to relegation to the English third tier under Gareth Ainsworth's time in charge. It is safe to say QPR's time in this division has been exciting to the neutral - but not so much to the QPR fans.

Since Marti Cifuentes' appointment at the back end of 2023, he has seen a massive uptick in form with Marti having the highest PPG and clean sheets per game of any manager since August 2015. Additionally, both trend and moving average graphs clearly indicate an improvement in QPR's fortunes made even more impressive by inheriting a QPR squad who under Gareth Ainsworth would have certainly been relegated had it not been for Cifuentes taking the managerial job.

It would be much easier and be able to reach a conclusion about Marti Cifuentes' QPR if he stays on for a few more years in the job but if early signs in his six months at the helm specify anything it is that Marti Cifuentes is the best manager QPR have had in charge since Neil Warnock took the club up to the Premier League in 2011 and if his record continues we could see QPR become a Premier League club once again in the not to distant future.