When Do NFL Running Backs Peak?

By Bogdan Gadidov and John Michael Croft

**Introduction**

National Football League (NFL) fantasy point production was utilized to model the age running backs (RBs) peak throughout their career. Data was collected for approximately 70 NFL RBs during the years 2008 – 2013. The primary focus was on players going through their “prime” at some point during 2008 – 2013, however some younger “rising star” players as well as older “aging” players were used for comparison purposes. For example, a player like Thomas Jones whose career was ending in 2008 was used in the model. A "rising" player such as Eddie Lacy, who entered the league in 2013, was also used in the model.

Individual player level fantasy point production was obtained from [www.espn.com](http://www.espn.com) while additional individual player level and team statistics were obtained from [www.nfl.com](http://www.nfl.com). Player level statistics included, but were not limited to: rush attempts, receptions, height, weight, and age. Team statistics included, but were not limited to: defensive rank and offensive rank. Touchdowns and rushing yards were not used in the model since fantasy points are calculated using these two variables, and would be highly significant in the model.

**Variables**

Below are the variables of interest attempted to model fantasy point production:

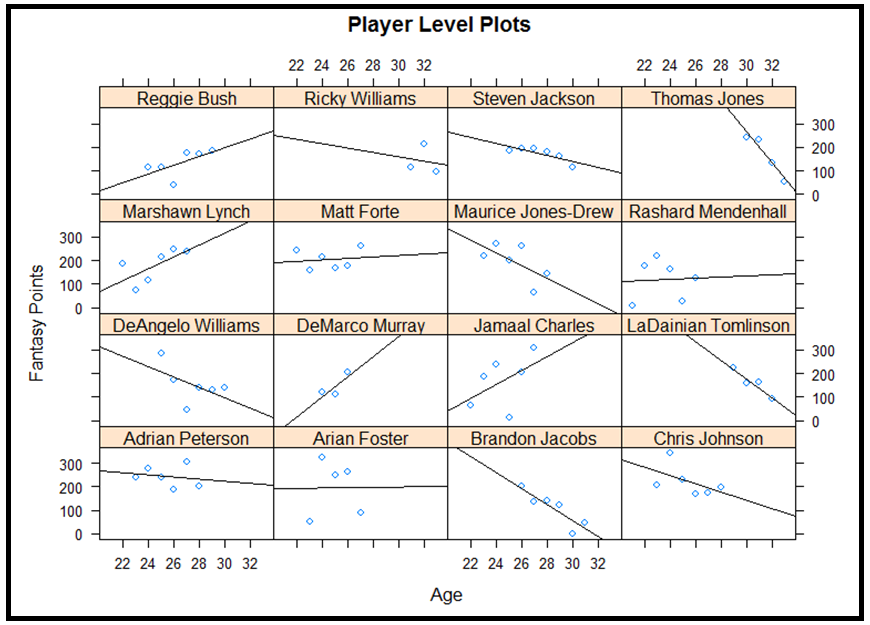
* Age – Age of player at the start of a given season.
* Ht – Player's height (in inches).
* Wt – Player's weight (in pounds).
* Def\_rank – Team defensive rank based on total yards allowed (1-32 ranking with 1 being the best).
* Pass\_rank – Team passing rank based on total passing yards (1-32 ranking with 1 being the best).
* Touches – Combination of total rushing attempts and receptions per player.
* Time – Player’s age minus 21 (created variable to allow for time 0 to be the reference point. 21 is the youngest age a player can enter the NFL).

**Methodology**

Looking at a player level plot as part of a preliminary analysis shows a general trend that players increase steadily in fantasy points during the earlier stages of their career, before experiencing a sharp decline towards the end of their career. For example, looking at figure 1 below, we can see the trend in fantasy points for some of the bigger name players from 2008 - 2013. Thomas Jones, who was 29 in 2008, has a very sharp decline in production once he reaches the age of 30. LaDanian Tomlinson has a similar decline around the age of 30. It can be seen, however, that Brandon Jacobs experiences his decline earlier in his career, around the age of 26-27 (which is more in line with when the majority of NFL running backs begin to decline). Younger players, such as, DeMarco Murray and Jamaal Charles, have increasing trends in their fantasy points, as they are entering or in their prime.

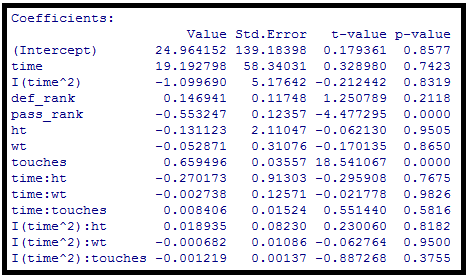
One important thing to note from the plots of the individual players is that some trajectories appear to be quadratic, such as Marshawn Lynch, Steven Jackson, and Jamaal Charles. Because of this apparent pattern, a quadratic effect for time (time^2) will be used in later modeling.

Figure 1. Individual Player Level Plots



In the first model that was run, the variable "touches" was used as one of the predictors. This variable is the sum of the rushing attempts and receptions that a running back received over the course of the season. In the models with this variable, this variable dominated the model. The output in figure 2 shows one of the models with the variable "touches" as one of the predictors. We can see in the preliminary model, the only significant variables are touches and pass rank. Even once the non significant variables were removed, these 2 variables were the only significant ones in the model. It was interesting to see that this model resulted in one of the team ranks being significant, but it was decided to not include touches in the model, as it was too dominant of a predictor for a fantasy running back's total fantasy points in a season.

Figure 2. Model with Variable "Touches"

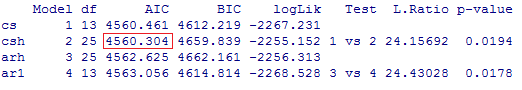


Several methods were considered for modeling the covariance structure. Specifically, a compound symmetry structure was compared to an autoregressive structure. First, a compound symmetry (CS) structure was compared to a compound symmetry heterogeneous (CSH) structure, with heterogeneous variances. In CS, there are 2 parameters, one for the correlation and one for the variance. The variance is constant for each time period in CS. In CSH, there is 1 parameter for the correlation again, but each time period is given its own variance, so there are parameters for the number of time periods in the model plus 1. A likelihood ratio test is used to determine whether to choose the CS or CSH structure. Since this likelihood ratio is comparing covariance structures, a restricted maximum likelihood (REML) must be used as this method gives variance estimates with less bias. The likelihood ratio test has a null hypothesis which states that the reduced model is as good as the full model. The alternate hypothesis to this test is that the reduced model is not as good as the full model. The full model is the model with more covariance parameters (the CSH model). The reduced model in this case is the CS model since it is a special case of the CSH model, where all the variances are equal for each time period. The p-value is .0194, so the null hypothesis is rejected, and the CSH structure is chosen.

An autoregressive structure (AR(1)) was also compared to an autoregressive heterogeneous (AR(1)H) structure. An AR(1) model has 2 covariance parameters, the correlation and variance. The difference between AR(1) and CS is that an AR(1) model has a correlation which diminishes to 0 as the time periods are further and further apart. The correlation is raised to a power determined by the distance between time periods, so as the time periods are further apart; this correlation will go to 0. An AR(1)H model is different from an AR(1) model in that it has a different variance for each time period. A likelihood ratio test can again be used to compare these 2 models. The AR(1) model is the reduced model, as it is a special case of the AR(1)H model where all the variances are equal for each time period. The null hypothesis in the likelihood ratio test is that the AR(1) structure is just as good as the AR(1)H structure. The p-value is .0178, so the null hypothesis is again rejected, and the AR(1)H structure is chosen.

The CSH and AR(1)H models are compared using the AIC as the criteria. We cannot use the likelihood ratio test because neither the CSH nor AR(1)H model is a special case of the other. After choosing the model with the lower AIC, the CSH model is chosen to use in further steps. Output from the likelihood tests described above is shown in figure 3 below.

Figure 3. Likelihood Ratio Test for Choosing Best Covariance Structure



The compound symmetry heterogeneous structure was then compared to a random intercepts model with heterogeneous variances. Since a random intercepts model without the heterogeneous variances yields the same conclusion when testing a CS versus CSH model, a random intercepts model with heterogeneous variances was also tested. The random intercepts model is a model in which each player is given their own personal adjustment (intercept) in addition to the equation of the model. This model will be outlined in more detail in the next section. This random intercepts model with heterogeneous variances was chosen for the final model based off the AIC. Output from the likelihood test is shown in figure 4 below.

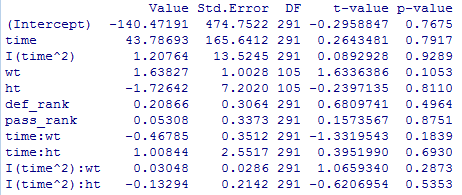
Figure 4. Comparison of CSH to Random Intercepts Model



**Results**

The random intercepts model with heterogeneous variances was run with the following predictors: height, weight, def\_rank, and pass\_rank. Several interaction terms were also included: time\*weight, time\*height, time^2\*weight, and time^2\*height. Figure 5 shows the model with all these terms, along with their respective p-values.

Figure 5. Full Random Intercepts Model



Terms in the above model are removed in a hierarchical fashion, such that the non significant quadratic terms are removed first, followed by lower order interaction terms next, and so forth. The final model that was chosen is shown in figure 6 below. The final terms in the model were time, time^2, weight, and time\*weight. A significance level of .1 was chosen. Notice that even though the weight term has a p-value greater than .1, since it is used in a higher order term (time\*weight), it should be left in the model. Once the final model in figure 6 was obtained, it was compared to the original model in figure 5 to assure that after removing the non significant variables, the final model was as good as the original model. This could again be done using a log likelihood test, but instead of doing a REML method, a maximum likelihood method was used in this test since the way that the mean was modeled was being analyzed instead of the way the covariance was being modeled. The null hypothesis of the test is that the reduced model, in figure 6, is as good as the full model, in figure 5. The reduced model in this test is the model in figure 6 since this model only has 4 of the original terms used in the figure 5 model. The p-value from this test is .9415, so the null hypothesis cannot be rejected, and the conclusion is that the model in figure 6 is just as good as the original model in figure 5.

Figure 6. Final Random Intercepts Model after Removing

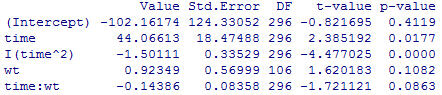
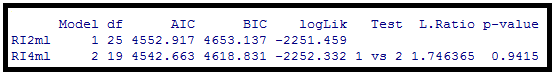


Figure 7. Comparison of Final Random Intercepts Model to Full Random Intercepts Model



Residual plots were analyzed and are shown in figures 8 and 9. Figure 8 has a plot of the Cholesky residuals vs. age. Cholesky residuals are good for longitudinal data because data points tend to be highly correlated within subjects. In the plot in figure 8, one looks for constant variance in the residuals for each age group; ideally, there should be a horizontal band of the residuals. The residuals appear fairly constant in the plot, as the band widths of residuals are approximately the same height for each age group. No age group appears to have significantly higher or lower residuals than any other age group. Figure 9 shows a normal probability plot of the Cholesky residuals. In looking at this plot, we can see that there are no concerns for non normality of the residuals as the residuals lie close to the line on the probability plot.

Figure 8. Plot of Cholesky Residuals vs. Age

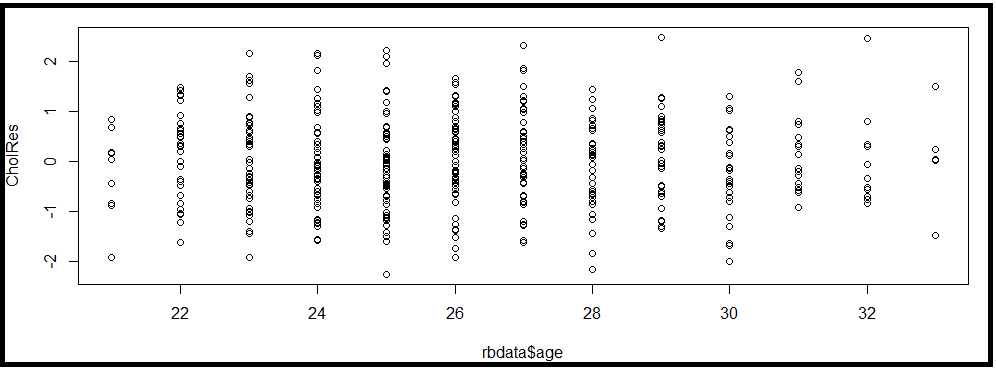
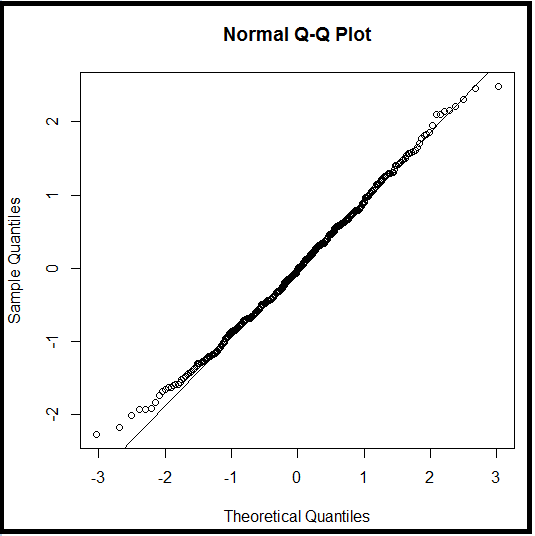
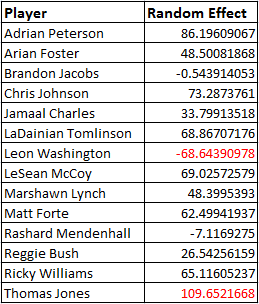


Figure 9. Normal Probability Plot of the Residuals



As stated earlier, the random intercepts model gives each player an individual adjustment. That is, when predicting a player's fantasy points based off the player's age and time, that player receives an additional adjustment based on the value of this random effect. Figure 7 shows these random effects for some of the bigger name running backs in the NFL. If Thomas Jones were compared to a similar running back with the same weight and age, he would receive an additional 109.65 point adjustment to his fantasy point’s prediction. Leon Washington, however, would receive a negative 68.64 point adjustment to his fantasy total when compared to a similar player at the same age and weight level.

Figure 7. Random Intercepts for 14 Players in Dataset

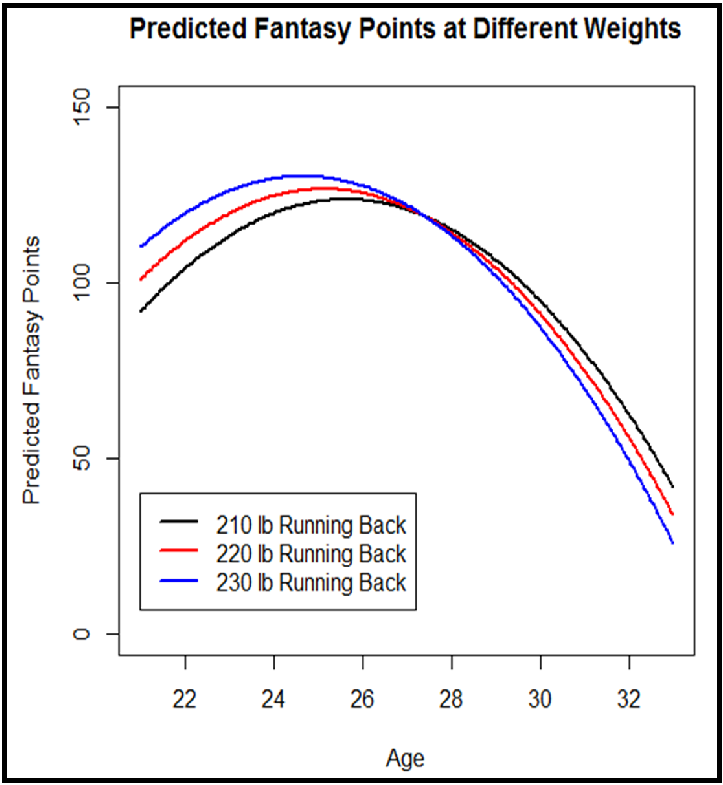


Since weight and time were both significant in the model, predicted fantasy points were graphed at three different weight classes (210 lbs, 220 lbs, and 230 lbs). These weight limits were chosen as 210 pounds was the 25th percentile of running back weights and 230 pounds was the 80th percentile of running back weights. The predicted fantasy points were created using the coefficients from the final model:

Fantasy Points= -102.16174 + 44.06613\*time -1.50111\*time^2 +.92349\*weight - .14386\*weight\*time

To create the graph in figure 8, values for age were entered into the equation above in increments of .05 for ages from 21-33. There are 3 distinct trajectories for the fantasy point projections at each weight level. One can see that heavier running backs have the highest peak, but also peak at the earliest age around 24. Lighter running backs have the smallest peak, but peak around 26. Lighter running backs also have more productive seasons in terms of fantasy point productions throughout the later stages of their careers. Heavier running backs have the fastest decline in production after they reach their peaks. From the plot, it appears that by the age of 27, regardless of the weight of the running back, the productivity of the running back is declining at a fairly rapid rate. Running backs are in the prime of their careers between the ages of 22 to 27, and peak between 24 to 26.

Figure 8. Predicted Fantasy Points at Different Weights



**Conclusion**

Overall, it appears RB NFL production peaks around the ages 24 – 27 with lighter RBs peaking later and remaining marginally more productive after age 28 due to assumed different running styles (i.e. “ground and pound” between the tackles for heavier backs). NFL teams should focus on heavier RBs out of college while looking to lighter RBs over the age of 27. However, given the recent history of the NFL – teams may overpay a premium for RBs around the ages of 26 / 27 based on past performance which is not indicative of future performance base on the data.

Recommend NFL teams consider these findings when negotiating RB contracts. Since the 2011 CBA was put in place, rookie contracts are now all four years in length with predetermined dollar ranges available to be negotiated (no longer is it a pure market negotiation between two parties as the CBA has placed constraints on Max and Min contract amounts). Teams would be better off focusing on heavier RBs out of college and looking to sign / re-sign, 4 years later, but for a much shorter contract terms at higher dollar values. For example, while DeMarco Murray has been a top performer over the past couple of years under his rookie contract, the data suggest he has only another year or two left as a top producer – so it may not be in a team’s best interest to re-sign to a long term contract for at higher dollar value base on past performance. If Murray is re-signed at a higher dollar value, recommend it be for 2-3 years with a team option in year 4. It may be more productive to drafter a younger, heavier RB for significantly less money with only a marginal drop in production.

**References**

* "Statistics." *NFL.* 26 Oct. 2014. <http://www.nfl.com/stats/categorystats?tabSeq=2&offe nsiveStatisticCategory=TEAM\_PASSING&conference=ALL&role=TM&season =2013&seasonType=REG>
* "NFL Players By Position - Running Back." *ESPN NFL*. 1 Nov 2014. <http://espn.go.c om/nfl/players/\_/position/rb>

**R Code**

rbdata <- read.csv("C:/Users/Bogdan/Word Documents/STAT 8225/rbdata.csv")

rbdata<- rbdata[rbdata$time<= 12,]

names(rbdata)

head(rbdata, 411)

library(lattice)

library(nlme)

mean\_fantasy <-aggregate(rbdata$fantasy, by=list(rbdata$age), FUN=mean)

mean\_fantasy

# average pts by age plot

plot(0,0, xlim=c(21,33), ylim=c(50,150), type="n", main="Average Fantasy Points by Age",

xlab="Age", ylab="Fantasy Points")

lines(mean\_fantasy$Group.1,mean\_fantasy$x,lwd=3,col="blue")

axis(1,at=c(21:33))

# plot of players

xyplot(fantasy~age|player, data=rbdata[rbdata$player %in%

c("Adrian Peterson", "Arian Foster", "Marshawn Lynch", "DeMarco Murray", "Jamaal Charles",

"Maurice Jones-Drew", "Matt Forte", "Chris Johnson", "LaDainian Tomlinson", "DeAngelo Williams",

"Ricky Williams", "Reggie Bush", "Rashard Mendenhall", "Thomas Jones", "Brandon Jacobs", "Steven Jackson"),],

panel = function(x, y) {

panel.xyplot(x, y)

panel.lmline(x, y)

}, ylab="Fantasy Points",xlab="Age",main="Player Level Plots")

# including touches in the model

csht <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+touches+

time:ht+time:wt+touches:time+

I(time^2):ht+I(time^2):wt+touches:I(time^2),

correlation=corCompSymm(form=~time|player),

weights=varIdent(form=~1|as.factor(time)),

data=rbdata)

summary(csht)

csht2 <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+touches,

correlation=corCompSymm(form=~time|player),

weights=varIdent(form=~1|as.factor(time)),

data=rbdata)

summary(csht2)

# choosing best cov without touches

cs <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

correlation=corCompSymm(form=~time|player),

data=rbdata)

csh <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

correlation=corCompSymm(form=~time|player),

weights=varIdent(form=~1|as.factor(time)),

data=rbdata)

summary(csh)

# compare cs to csh

anova(cs, csh)

ar1 <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

correlation=corAR1(form=~time|player),

data=rbdata)

arh <- gls(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

correlation=corAR1(form=~time|player),

weights=varIdent(form=~1|as.factor(time)),

data=rbdata)

anova(ar1, arh)

# compare ar1 to arh

anova(cs,csh,arh,ar1)

#Testing RI v RIAS

RIAS <-lme(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

data=rbdata, random=~time|player, na.action=na.omit)

random.effects(RIAS)

RI<- lme(fantasy~time+I(time^2)+def\_rank+pass\_rank+ht+wt+

time:ht+time:wt+

I(time^2):ht+I(time^2):wt,

data=rbdata, random=~1|player, na.action=na.omit)

summary(RI)

anova(RI,csh) #choose csh over RI

RI2<- lme(fantasy~time+I(time^2)+wt+ht+def\_rank+pass\_rank+

time:wt+time:ht

+I(time^2):wt+I(time^2):ht,

weights=varIdent(form=~1|time),

data=rbdata, random=~1|player, na.action=na.omit)

anova(RI2,csh)

random.effects(RI2)

anova(RI,csh) # choose RI

anova(RIAS,RI)

random.effects(RI)

RI3<- lme(fantasy~time+I(time^2)+wt+ht+def\_rank+pass\_rank+

time:wt+time:ht,

weights=varIdent(form=~1|time),

data=rbdata, random=~1|player, na.action=na.omit)

summary(RI3)

RI4<- lme(fantasy~time+I(time^2)+wt+

time:wt,

weights=varIdent(form=~1|time),

data=rbdata, random=~1|player, na.action=na.omit)

summary(RI4)

random.effects(RI4)

write.table(random.effects(RI4),"C:/Users/Bogdan/Word Documents/STAT 8225/RI.csv",sep=",",quote=FALSE)

CholRes <- residuals(RI4,type="normalized")

plot(rbdata$age, CholRes)

qqnorm(CholRes)

qqline(CholRes)

# compare RI2 to RI4

RI2ml <- update(RI2, method="ML")

RI4ml <- update(RI4, method="ML")

anova(RI2ml, RI4ml)

summary(rbdata$wt)

quantile(rbdata$wt,c(.25,.6,.8))

# 210lbs - 25% 220lbs - 60% 230lbs - 80% creating plot with predictions

tp <- seq(0,12,by=.05)

y210 <- -102.16174 + 44.06613\*tp -1.50111\*tp^2 +.92349\*210 - .14386\*210\*tp

y220 <- -102.16174 + 44.06613\*tp -1.50111\*tp^2 +.92349\*220 - .14386\*220\*tp

y230 <- -102.16174 + 44.06613\*tp -1.50111\*tp^2 +.92349\*230 - .14386\*230\*tp

tp2=tp+21 # to make x-axis reflect age

plot(y210~tp2,type="l", ylim=c(0,150),lwd=2,xlab="Age"

,ylab="Predicted Fantasy Points",main="Predicted Fantasy Points at Different Weights")

lines(tp2,y220,col="red",lwd=2)

lines(tp2,y230,col="blue",lwd=2)

legend(21,40,c("210 lb Running Back","220 lb Running Back","230 lb Running Back")

,col=c("black","red","blue"),

lty=1,lwd=2)