Marketing Mix Assignment

Introduction

The following models were prepared using the 'MLB' dataset for the 1993 season salaries of Major League Baseball players . This dataset contained a number of variables affecting MLB player's *'salaries'* which was our outcome variable. Our independent variables were split into dichotomous and continuous moderators, where *'frstbase', scndbase', 'thrdbase', 'shrtstop', 'outfield' and 'catcher'* represent the dichotomous variables and *'hruns'* and *'games'* represent the continuous variables. The goal of our report was to analyze how a MLB manager might determine a player's salary.

Spotlight Analysis- Dichotomous Moderator

We ran several regressions relating salaries and the dichotomous variables mentioned above to *hruns*. These regressions are shown in the appendix below under section 1.

All of our dichotomous variables are position based, so these regressions answered the question of whether the position a player played affected their salary given an interaction of the number of career home runs they had with their position. As these position variables all came within the dataset as dichotomous, where 1 means that is their position and 0 means it is not there was no need to create any further variables beyond non variables for each position. Given the magic rule of 0's these non position variables are essential to see the actual effect of homeruns interaction with position on salary.

As can be seen in Appendix 1 there aren't significant interactions between *frstbase* and *hruns* as well as between *thirdbase* and *hruns*. This means that *hrun* is independent of these two positions. The remaining positions such as *scndbase, 'shrtstop', 'outfield' and 'catcher'* are significant with respect to *hurns*. As we can see from Appendix 1F, players playing in the *catcher* position earned $12479 more than players playing other positions for each additional career home run (*hruns*). This makes the position of catcher most dependent on *hruns* compared to the other positions. .

Spotlight Analysis- Continuous Moderator

Here we first ran a mean regression on the effect between *hruns* and *games* (career games played) and the effect they have on *salary* (1993 salary). From appendix 2A it can be seen that this effect is highly significant at the 0.001 level. Thus we decided to run further regressions for High and Low standard deviations from the mean to understand which direction this effect is highest.

From looking at appendix 2B and 2C it can be seen that both interactions of high and low produce significant results. Therefore, we must examine the Beta values of both interactions.

We can see that if a player played a low no. of games his salary increased by $19860.3409 for every additional hrun he scored. While if he played a high no. of games, the player's salary increased by $10694.3042 per home run. There is a difference of $9166.0367 for a player's salary per home run in relation to the number of games they played. This is due to the law of diminishing returns, as a player's ratio of home runs to games decreases as more number of games are played.

Floodlight Analysis

We used the Johnson-Neyman Plot as seen in Appendix Table 3 to analyze the effect of continuous moderator no. of *games* played on *hruns* scored in their career. And it's evident from the graph that there's a decline in the no. of *hruns* as their career progresses towards playing more no. of games. This interaction between *hruns* and *games* remains significant until the no. of games played reaches 2000 and it turns insignificant as most of the players retire after that point in their career.

Conclusion

Based on the models, we can interpret that players playing at the *catcher* position are the ones who earn the most compared to other players when compared to their number of career *hruns* as they carry two very important responsibilities throughout the game giving them a dual role. Given the graph for floodlight analysis no. of *games* played and total *hruns* scored throughout their career has an inversely proportional relationship. Players tend to score more *hruns* in their prime rather than through mid or end of their career. So MLB managers definitely look at the no. of games played interacted with career home runs to make the determination that they are likely better off drafting younger rookie players and developing them rather than trading for older players who have been in the league for many years.

Appendix

Table 1A:

```
Output_frstbase <- lm(salary~hruns*frstbase,data)
summary(Output_frstbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * frstbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5128281  -614857  -388753   499932  4164510
##
## Coefficients:
##                 Estimate Std. Error t value            Pr(>|t|)
## (Intercept)     720717.8    78516.7   9.179 <0.0000000000000002 ***
## hruns            11639.6      909.8  12.794 <0.0000000000000002 ***
## frstbase       -179182.3   240801.4  -0.744               0.457
## hruns:frstbase     670.3     2044.5   0.328               0.743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1115000 on 349 degrees of freedom
## Multiple R-squared:  0.3771, Adjusted R-squared:  0.3718
## F-statistic: 70.43 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_frstbase <- data$frstbase - 1
Output_non_frstbase <- lm(salary~hruns*non_frstbase,data)
summary(Output_non_frstbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_frstbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5128281  -614857  -388753   499932  4164510
##
## Coefficients:
##                     Estimate Std. Error t value        Pr(>|t|)
## (Intercept)         541535.5   227641.0   2.379          0.0179 *
## hruns                12309.9     1830.9   6.723 0.0000000000726 ***
## non_frstbase       -179182.3   240801.4  -0.744          0.4573
## hruns:non_frstbase     670.3     2044.5   0.328          0.7432
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1115000 on 349 degrees of freedom
## Multiple R-squared:  0.3771, Adjusted R-squared:  0.3718
## F-statistic: 70.43 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 1B:

Appendix 1B:

salary given interaction of hruns on scndbase

```
Output_scndbase <- lm(salary~hruns*scndbase,data)
summary(Output_scndbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * scndbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4993560  -597135  -402805   474501  4208572
##
## Coefficients:
##                Estimate Std. Error t value            Pr(>|t|)
## (Intercept)    696003.2    78172.3   8.903 <0.0000000000000002 ***
## hruns           11363.2      822.8  13.811 <0.0000000000000002 ***
## scndbase       -49499.8   231550.9  -0.214               0.831
## hruns:scndbase   8392.1     3675.5   2.283               0.023 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1106000 on 349 degrees of freedom
## Multiple R-squared:  0.3879, Adjusted R-squared:  0.3827
## F-statistic: 73.73 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_scndbase <- data$scndbase - 1
Output_non_scndbase <- lm(salary~hruns*non_scndbase,data)
summary(Output_non_scndbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_scndbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4993560  -597135  -402805   474501  4208572
##
## Coefficients:
##                    Estimate Std. Error t value    Pr(>|t|)
## (Intercept)          646503     217956   2.966     0.00322 **
## hruns                 19755       3582   5.515 0.000000068 ***
## non_scndbase         -49500     231551  -0.214     0.83085
## hruns:non_scndbase     8392       3676   2.283     0.02301 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1106000 on 349 degrees of freedom
## Multiple R-squared:  0.3879, Adjusted R-squared:  0.3827
## F-statistic: 73.73 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 1C:

Appendix 1C:

salary given interaction of hruns on shrtstop

```
Output_shrtstop <- lm(salary~hruns*shrtstop,data)
summary(Output_shrtstop)
```

```
##
## Call:
## lm(formula = salary ~ hruns * shrtstop, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4994864  -614800  -367975   486099  3956318
##
## Coefficients:
##                 Estimate Std. Error t value            Pr(>|t|)
## (Intercept)     715984.2    80715.3   8.870 <0.0000000000000002 ***
## hruns            11316.3      833.7  13.574 <0.0000000000000002 ***
## shrtstop       -148310.9   200172.8  -0.741              0.4592
## hruns:shrtstop    7400.7     3528.8   2.097              0.0367 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1109000 on 349 degrees of freedom
## Multiple R-squared:  0.384,  Adjusted R-squared:  0.3787
## F-statistic: 72.53 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_shrtstop <- data$shrtstop - 1
Output_non_shrtstop <- lm(salary~hruns*non_shrtstop,data)
summary(Output_non_shrtstop)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_shrtstop, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4994864  -614800  -367975   486099  3956318
##
## Coefficients:
##                    Estimate Std. Error t value     Pr(>|t|)
## (Intercept)          567673     183178   3.099       0.0021 **
## hruns                 18717       3429   5.459 0.0000000912 ***
## non_shrtstop        -148311     200173  -0.741       0.4592
## hruns:non_shrtstop     7401       3529   2.097       0.0367 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1109000 on 349 degrees of freedom
## Multiple R-squared:  0.384,  Adjusted R-squared:  0.3787
## F-statistic: 72.53 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 1D:

Appendix 1D:

salary given interaction of hruns on thrdbase

```
Output_thrdbase <- lm(salary~hruns*thrdbase,data)
summary(Output_thrdbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * thrdbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5092884  -611435  -404937   518721  4177620
##
## Coefficients:
##                 Estimate Std. Error t value            Pr(>|t|)
## (Intercept)     712363.7    77135.6   9.235 <0.0000000000000002 ***
## hruns            11571.7      831.7  13.913 <0.0000000000000002 ***
## thrdbase       -150305.5   285608.7  -0.526               0.599
## hruns:thrdbase    1867.7     3339.5   0.559               0.576
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1116000 on 349 degrees of freedom
## Multiple R-squared:  0.3766, Adjusted R-squared:  0.3713
## F-statistic: 70.29 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_thrdbase <- data$thrdbase - 1
Output_non_thrdbase <- lm(salary~hruns*non_thrdbase,data)
summary(Output_non_thrdbase)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_thrdbase, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5092884  -611435  -404937   518721  4177620
##
## Coefficients:
##                    Estimate Std. Error t value  Pr(>|t|)
## (Intercept)          562058     274995   2.044    0.0417 *
## hruns                 13439       3234   4.155 0.0000409 ***
## non_thrdbase        -150306     285609  -0.526    0.5990
## hruns:non_thrdbase     1868       3340   0.559    0.5763
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1116000 on 349 degrees of freedom
## Multiple R-squared:  0.3766, Adjusted R-squared:  0.3713
## F-statistic: 70.29 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 1E:

Appendix 1E:

salary given interaction of hruns on outfield

```
Output_outfield <- lm(salary~hruns*outfield,data)
summary(Output_outfield)
```

```
##
## Call:
## lm(formula = salary ~ hruns * outfield, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4382208  -587850  -393337   489172  4120517
##
## Coefficients:
##                Estimate Std. Error t value             Pr(>|t|)
## (Intercept)      574613      92130   6.237        0.00000000129 ***
## hruns             14355       1189  12.072 < 0.0000000000000002 ***
## outfield         302874     153006   1.979              0.04855 *
## hruns:outfield    -4984       1618  -3.081              0.00223 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1102000 on 349 degrees of freedom
## Multiple R-squared:  0.3926, Adjusted R-squared:  0.3874
## F-statistic: 75.19 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_outfield <- data$outfield - 1
Output_non_outfield <- lm(salary~hruns*non_outfield,data)
summary(Output_non_outfield)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_outfield, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4382208  -587850  -393337   489172  4120517
##
## Coefficients:
##                    Estimate Std. Error t value             Pr(>|t|)
## (Intercept)          877486     122159   7.183 0.000000000004146059 ***
## hruns                  9371       1097   8.544 0.000000000000000409 ***
## non_outfield         302874     153006   1.979              0.04855 *
## hruns:non_outfield    -4984       1618  -3.081              0.00223 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1102000 on 349 degrees of freedom
## Multiple R-squared:  0.3926, Adjusted R-squared:  0.3874
## F-statistic: 75.19 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 1F:

Appendix 1F:

salary given interaction of hruns on catcher

```
Output_catcher<- lm(salary~hruns*catcher,data)
summary(Output_catcher)
```

```
##
## Call:
## lm(formula = salary ~ hruns * catcher, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5001906  -639010  -346512   455946  4163420
##
## Coefficients:
##                Estimate Std. Error t value            Pr(>|t|)
## (Intercept)    749010.3    80490.9   9.306 <0.0000000000000002 ***
## hruns           11251.0      819.3  13.733 <0.0000000000000002 ***
## catcher       -481230.5   218048.7  -2.207              0.0280 *
## hruns:catcher   12479.1     5097.9   2.448              0.0149 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1106000 on 349 degrees of freedom
## Multiple R-squared:  0.3878, Adjusted R-squared:  0.3825
## F-statistic: 73.69 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
data$non_catcher <- data$catcher - 1
Output_non_catcher <- lm(salary~hruns*non_catcher,data)
summary(Output_non_catcher)
```

```
##
## Call:
## lm(formula = salary ~ hruns * non_catcher, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5001906  -639010  -346512   455946  4163420
##
## Coefficients:
##                   Estimate Std. Error t value   Pr(>|t|)
## (Intercept)         267780     202648   1.321     0.1872
## hruns                23730       5032   4.716 0.00000348 ***
## non_catcher        -481230     218049  -2.207     0.0280 *
## hruns:non_catcher    12479       5098   2.448     0.0149 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1106000 on 349 degrees of freedom
## Multiple R-squared:  0.3878, Adjusted R-squared:  0.3825
## F-statistic: 73.69 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

GSB 516: Group 3

Table 2A:

Question 2:

Table 2A:

Use the same data set to present a meaningful Spotlight analysis on moderated regression analyses using a continuous moderator.

```
data$cHruns <- data$hruns - mean(data$hruns)
data$cGames <- data$games - mean(data$games)

out <- lm(salary~ cHruns*cGames,data)
summary(out)
```

```
##
## Call:
## lm(formula = salary ~ cHruns * cGames, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3157323  -420852   -80015   376365  3773905
##
## Coefficients:
##                  Estimate   Std. Error t value             Pr(>|t|)
## (Intercept)   1605953.8775   60635.2549  26.485 < 0.0000000000000002 ***
## cHruns          15277.3226    1445.4579  10.569 < 0.0000000000000002 ***
## cGames            617.5374     154.1500   4.006 0.000075480537476104 ***
## cHruns:cGames      -8.5076       0.9917  -8.579 0.000000000000000318 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 986400 on 349 degrees of freedom
## Multiple R-squared:  0.513,  Adjusted R-squared:  0.5088
## F-statistic: 122.5 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 2B:

Table 2B:

Table 2B:

```
data$HIGames <- data$cGames - sd(data$games)
data$LOGames <- data$cGames + sd(data$games)
outHI <- lm(salary~cHruns*HIGames,data)
outLO <- lm(salary~cHruns*LOGames,data)

summary(outHI)
```

```
##
## Call:
## lm(formula = salary ~ cHruns * HIGames, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3157323  -420852   -80015   376365  3773905
##
## Coefficients:
##                    Estimate   Std. Error t value             Pr(>|t|)
## (Intercept)    1938619.7511  100163.7461  19.355 < 0.0000000000000002 ***
## cHruns           10694.3042    1181.3177   9.053 < 0.0000000000000002 ***
## HIGames            617.5374     154.1500   4.006 0.000075480537476103 ***
## cHruns:HIGames      -8.5076       0.9917  -8.579 0.000000000000000318 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 986400 on 349 degrees of freedom
## Multiple R-squared:  0.513,  Adjusted R-squared:  0.5088
## F-statistic: 122.5 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

```
summary(outLO)
```

```
##
## Call:
## lm(formula = salary ~ cHruns * LOGames, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3157323  -420852   -80015   376365  3773905
##
## Coefficients:
##                    Estimate   Std. Error t value             Pr(>|t|)
## (Intercept)    1273288.0040  105412.7239  12.079 < 0.0000000000000002 ***
## cHruns           19860.3409    1831.3830  10.844 < 0.0000000000000002 ***
## LOGames            617.5374     154.1500   4.006 0.000075480537476104 ***
## cHruns:LOGames      -8.5076       0.9917  -8.579 0.000000000000000318 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 986400 on 349 degrees of freedom
## Multiple R-squared:  0.513,  Adjusted R-squared:  0.5088
## F-statistic: 122.5 on 3 and 349 DF,  p-value: < 0.00000000000000022
```

Table 3:

Question 3:

Use the same data set to present a meaningful Floodlight analysis using a continuous moderator.

Table 3:

```
library(interactions)
m <- lm(salary ~ hruns*games,data)
johnson_neyman(m, pred = hruns, modx = games, alpha = 0.05,  title = "Johnson-Neyman Plot")
```

```
## JOHNSON-NEYMAN INTERVAL
##
## When games is OUTSIDE the interval [2153.46, 2831.55], the slope of hruns
## is p < .05.
##
## Note: The range of observed values of games is [7.00, 2729.00]
```