

Problem 5

```
library(MASS)
library(nnet)
library(ggplot2)

rosters <- read.csv('rosters.csv')
summary(rosters)
```

```
##           X           gender           height           homestate
## Min.      : 0.00   male:204   Min.      :62.00   CA      :70
## 1st Qu.: 50.75                1st Qu.:71.00   WA      :12
## Median :101.50                Median :74.00   TX      :11
## Mean    :101.63                Mean    :73.33   GA      : 9
## 3rd Qu.:152.25                3rd Qu.:76.00   AZ      : 8
## Max.    :206.00                Max.    :84.00   FL      : 8
##                                     (Other):86
##           name           sport           weight
## Alabi, Adrian      : 1   Baseball   :29   Min.    :125.0
## Alexander, Terrence: 1   Basketball:12   1st Qu.:176.5
## Alfieri, Joey       : 1   Football   :95   Median  :197.0
## Allen, Malcolm      : 1   Soccer      :25   Mean    :205.4
## Allen, Marcus       : 1   Tennis      :10   3rd Qu.:229.2
## Allen, Rosco        : 1   Wrestling   :33   Max.    :321.0
## (Other)             :198
```

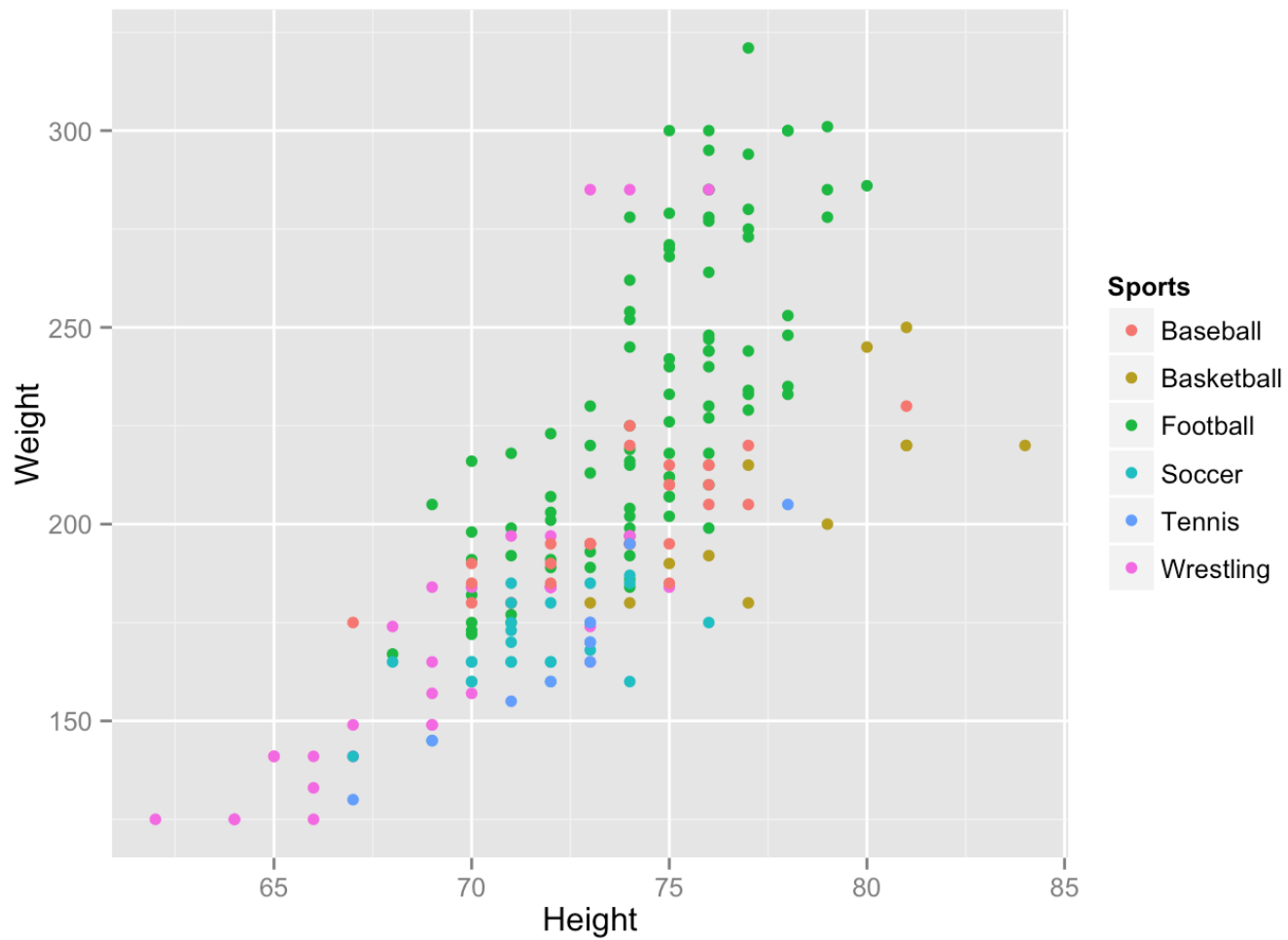
```
fit = multinom(sport ~ height + weight, rosters)
```

```
## # weights:  24 (15 variable)
## initial  value 365.518932
## iter   10 value 266.164885
## iter   20 value 207.817029
## iter   30 value 200.902009
## iter   40 value 200.198801
## iter   50 value 199.808004
## iter   60 value 199.680438
## iter   70 value 199.653822
## iter   80 value 199.647074
## iter   90 value 199.645164
## final   value 199.644727
## converged
```

```
summary(fit)
```

```
## Call:
## multinom(formula = sport ~ height + weight, data = rosters)
##
## Coefficients:
##           (Intercept)      height      weight
## Basketball  -71.754980   1.1940557 -0.09567987
## Football     16.366376  -0.3500011  0.05057577
## Soccer       -6.939543   0.3534022 -0.10321321
## Tennis      -30.636220   0.8123257 -0.16392309
## Wrestling    44.531678  -0.6709520  0.02033305
##
## Std. Errors:
##           (Intercept)      height      weight
## Basketball  0.140277235  0.06577436  0.02461660
## Football    4.953425402  0.08643443  0.01180643
## Soccer      0.083352009  0.04979394  0.01992204
## Tennis      0.004621523  0.07325690  0.03054070
## Wrestling   3.278747021  0.06415213  0.01411842
##
## Residual Deviance: 399.2895
## AIC: 429.2895
```

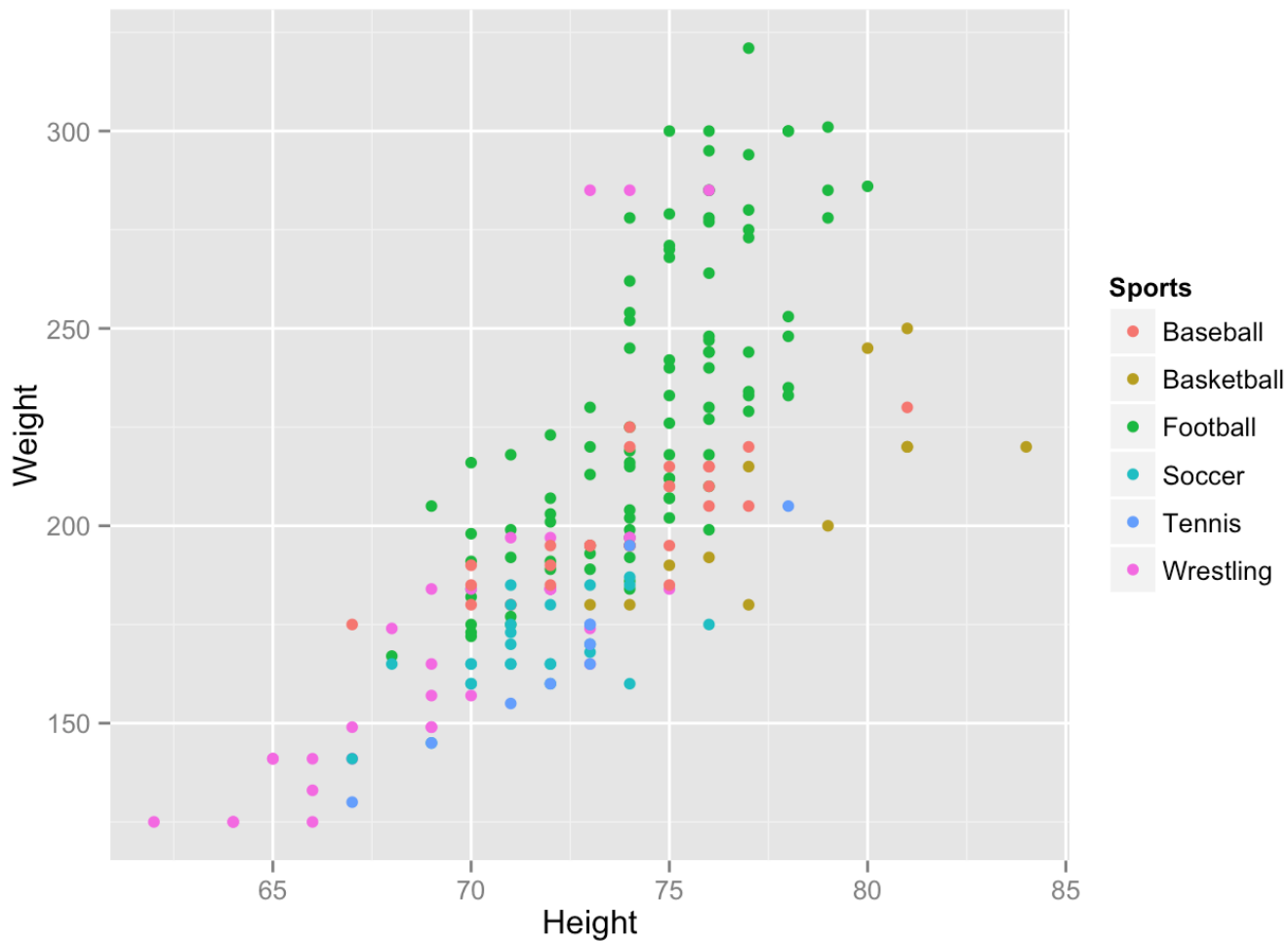
```
Height = rosters$height
Weight = rosters$weight
Sports = rosters$sport
qplot(Height, Weight, rosters, color=Sports)
```



```
pred = predict(fit, rosters)
pred
```

##	[1]	Wrestling	Football	Football	Football	Football
##	[6]	Football	Soccer	Football	Football	Football
##	[11]	Football	Wrestling	Football	Football	Football
##	[16]	Football	Football	Football	Football	Football
##	[21]	Football	Football	Football	Baseball	Football
##	[26]	Football	Football	Football	Football	Football
##	[31]	Football	Wrestling	Football	Football	Football
##	[36]	Wrestling	Football	Football	Soccer	Football
##	[41]	Football	Football	Football	Football	Football
##	[46]	Football	Football	Wrestling	Football	Football
##	[51]	Wrestling	Football	Football	Football	Football
##	[56]	Football	Football	Football	Football	Football
##	[61]	Football	Football	Football	Football	Football
##	[66]	Football	Football	Football	Football	Football
##	[71]	Football	Football	Football	Football	Football
##	[76]	Football	Football	Football	Football	Football
##	[81]	Football	Football	Wrestling	Football	Football
##	[86]	Wrestling	Football	Football	Football	Football
##	[91]	Football	Football	Football	Football	Football
##	[96]	Football	Soccer	Wrestling	Football	Soccer
##	[101]	Football	Football	Wrestling	Wrestling	Wrestling
##	[106]	Wrestling	Soccer	Football	Football	Wrestling
##	[111]	Soccer	Football	Football	Wrestling	Soccer
##	[116]	Soccer	Football	Wrestling	Wrestling	Football
##	[121]	Soccer	Wrestling	Wrestling	Football	Wrestling
##	[126]	Football	Wrestling	Wrestling	Basketball	Soccer
##	[131]	Soccer	Basketball	Football	Basketball	Baseball
##	[136]	Basketball	Basketball	Football	Basketball	Basketball
##	[141]	Football	Wrestling	Football	Football	Wrestling
##	[146]	Football	Football	Football	Football	Wrestling
##	[151]	Football	Football	Football	Football	Football
##	[156]	Football	Football	Football	Soccer	Football
##	[161]	Football	Football	Baseball	Football	Basketball
##	[166]	Football	Basketball	Football	Football	Football
##	[171]	Wrestling	Wrestling	Soccer	Soccer	Soccer
##	[176]	Wrestling	Football	Football	Soccer	Soccer
##	[181]	Soccer	Football	Soccer	Soccer	Tennis
##	[186]	Soccer	Soccer	Wrestling	Wrestling	Tennis
##	[191]	Soccer	Wrestling	Baseball	Soccer	Basketball
##	[196]	Soccer	Soccer	Soccer	Football	Soccer
##	[201]	Soccer	Soccer	Soccer	Soccer	
##	Levels:	Baseball	Basketball	Football	Soccer	Tennis Wrestling

```
qplot(Height, Weight, pred, color=Sports)
```



```
table(rosters$sport, pred)
```

```
##          pred
##      Baseball Basketball  Football  Soccer  Tennis  Wrestling
## Baseball          1           2        22         1         0          3
## Basketball         1           7         2         2         0          0
## Football           1           0        84         2         0          8
## Soccer             1           0         4        12         2          6
## Tennis             0           1         1         8         0          0
## Wrestling          0           0        12         7         0         14
```

```
success_rate = mean(pred == rosters$sport)
```

```
## Our success rate is 0.5784314
```

```
## Our 0-1 loss error rate is 0.4215686
```