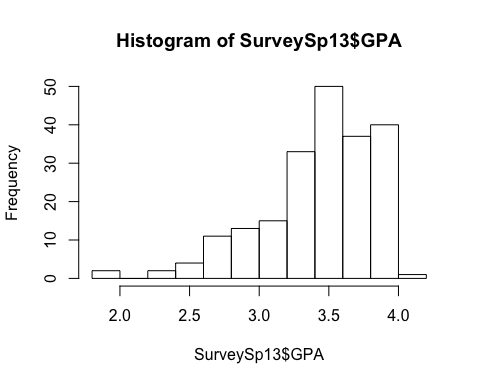
Lab10.R

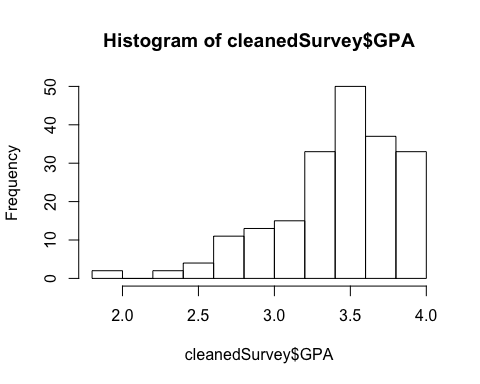
riserate

Thu Apr 13 16:00:45 2017

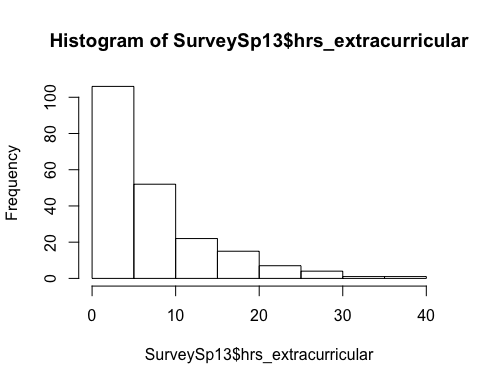
# Lab 10  
  
#Import the data into RStudio  
SurveySp13 <- read.csv("SurveySp13.csv", header = TRUE)  
  
#1(a)   
# Use a histogram to see the distrbution.  
# The GPA is left skewed   
hist(SurveySp13$GPA)



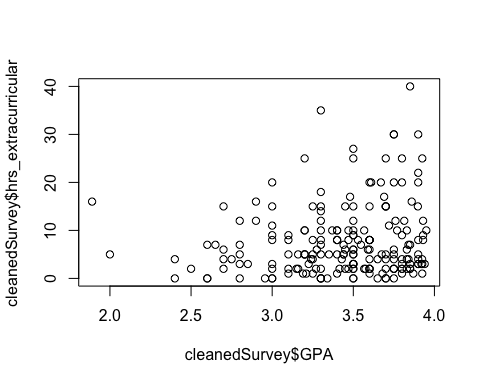
#1(b)  
#Creating a clean set due to some GPA's being higher than 4.0  
cleanedSurvey <-subset(SurveySp13,SurveySp13$GPA<4.0)   
hist(cleanedSurvey$GPA)



#Data contains 200 observations.  
  
#2(a)  
#Plotting histogram for hrs extracurriculars   
hist(SurveySp13$hrs\_extracurricular)



#Data is right skewed.  
  
#2(b)  
#Plotting a scatter plot to show the relationship between hrs extracurricular and GPA   
plot(cleanedSurvey$GPA,cleanedSurvey$hrs\_extracurricular)



#The scatter plot showed that there may be a slight positive correlation. Higher GPAs have more extracurricular involvement.  
  
#2(c)  
#Finding correation between hrs extracurricular and GPA   
cor(cleanedSurvey$GPA,cleanedSurvey$hrs\_extracurricular)

## [1] 0.1459643

cor.test(cleanedSurvey$GPA,cleanedSurvey$hrs\_extracurricular)

##   
## Pearson's product-moment correlation  
##   
## data: cleanedSurvey$GPA and cleanedSurvey$hrs\_extracurricular  
## t = 2.0761, df = 198, p-value = 0.03917  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.007372618 0.279054151  
## sample estimates:  
## cor   
## 0.1459643

#The correlation is 0.1459643. This is a weak positive correlation.   
#The 95% confidence interval is (0.007372618, 0.279054151) and does not include   
#zero, which means the null hypothesis (p=0) is rejected.  
  
#3(a)  
#Estimating regression model   
model1<-lm(cleanedSurvey$hrs\_extracurricular~cleanedSurvey$GPA)   
summary(model1)

##   
## Call:  
## lm(formula = cleanedSurvey$hrs\_extracurricular ~ cleanedSurvey$GPA)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.575 -5.459 -1.999 2.482 31.149   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.792 4.601 -0.389 0.6974   
## cleanedSurvey$GPA 2.764 1.331 2.076 0.0392 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7.303 on 198 degrees of freedom  
## Multiple R-squared: 0.02131, Adjusted R-squared: 0.01636   
## F-statistic: 4.31 on 1 and 198 DF, p-value: 0.03917

#viewing regressing results ŷ = -1.792 + 2.764GPA  
  
#3(b)  
#Get confidence interval for intercept and slope   
confint(model1)

## 2.5 % 97.5 %  
## (Intercept) -10.8649885 7.281313  
## cleanedSurvey$GPA 0.1386326 5.390075

#The confidence interval for the intercept is (-10.8649885, 7.281313)   
#The intercept is not significantly different than zero since the p-value > 0.05 also since the confidence interval includes 0  
  
#3(c)  
#Get confidence interval for intercept and slope.   
confint(model1)

## 2.5 % 97.5 %  
## (Intercept) -10.8649885 7.281313  
## cleanedSurvey$GPA 0.1386326 5.390075

#The confidence interval for the slope is (0.1386326, 5.390075).   
#Since it does not include 0 we can conclude the slope is significantly different (greater than) 0  
  
#3(d)  
#The R 2 of the model is 0.02131.   
#This is quite low and so the association between hrs extrscurricula and GPA is rickety.  
  
#3(e)  
#The residual standard error is 7.303 on 198 degrees of freedom. It is moderately high.  
  
#4(a)  
#Checking model assumptions using residual plots   
model1$residuals

## 1 2 3 4 5 6   
## 10.66946972 -7.33053028 -7.44110444 -6.71270723 15.94174853 0.66946972   
## 7 8 9 10 11 12   
## -7.90621200 -1.07207323 -2.05409489 11.84016355 1.26312979 -6.71270723   
## 13 14 15 16 17 18   
## 21.42551047 -4.43627184 -4.88340106 -4.71270723 13.49877589 -5.71270723   
## 19 20 21 22 23 24   
## -2.19231259 27.66946972 0.32808206 7.80768741 -5.88340106 -1.67191794   
## 25 26 27 28 29 30   
## -6.50122411 -6.04442970 -1.88340106 -5.39548255 0.39303433 9.32808206   
## 31 32 33 34 35 36   
## 6.66946972 -1.26142143 2.11659894 -8.43627184 -1.74518336 16.56372816   
## 37 38 39 40 41 42   
## -6.05825147 -4.88340106 12.56720871 -1.82328138 9.17188602 -4.15983645   
## 43 44 45 46 47 48   
## -3.43627184 11.64665878 -6.09971677 -6.37959254 -5.24759966 -2.60696567   
## 49 50 51 52 53 54   
## -6.57448953 -6.19231259 -3.57448953 -0.15983645 -1.01332569 4.66946972   
## 55 56 57 58 59 60   
## -3.19231259 -5.80047044 -6.08173843 5.77521128 -3.67191794 -2.15983645   
## 61 62 63 64 65 66   
## -6.15983645 2.22234050 0.39303433 -3.16466905 -0.15983645 17.11659894   
## 67 68 69 70 71 72   
## -6.57448953 1.60451745 -5.39548255 -4.12320374 -2.94835333 -5.98914262   
## 73 74 75 76 77 78   
## -6.15983645 -2.57448953 -6.06861277 2.94590511 21.42551047 -6.85092492   
## 79 80 81 82 83 84   
## -4.82328138 13.01085738 -5.85092492 -6.60696567 0.03643268 6.56372816   
## 85 86 87 88 89 90   
## 2.50844108 -7.33053028 19.11659894 1.37022339 -0.98914262 -6.15983645   
## 91 92 93 94 95 96   
## -2.71753983 6.84016355 -4.50122411 -8.29805414 -4.88340106 -4.88340106   
## 97 98 99 100 101 102   
## 11.42551047 1.49877589 -6.85092492 -0.71753983 -4.77765950 7.11659894   
## 103 104 105 106 107 108   
## -2.88340106 2.49877589 -2.88340106 -4.57448953 3.24859182 1.11659894   
## 109 110 111 112 113 114   
## 1.11659894 6.01638609 -7.88340106 -5.64843098 2.66946972 -1.81013564   
## 115 116 117 118 119 120   
## -6.79563785 -3.50122411 6.56372816 17.94590511 0.80768741 1.05164667   
## 121 122 123 124 125 126   
## -2.71753983 11.01085738 2.39303433 -0.15983645 9.77521128 -3.36716299   
## 127 128 129 130 131 132   
## -6.50122411 2.19952956 -4.91587720 -7.88340106 7.12143154 -2.46874797   
## 133 134 135 136 137 138   
## -2.05409489 4.49877589 -3.11904716 11.28729277 -5.33053028 -7.38098476   
## 139 140 141 142 143 144   
## -4.79563785 -8.43627184 1.20436215 0.28729277 2.94590511 6.05164667   
## 145 146 147 148 149 150   
## -3.98914262 -3.68713193 -0.33053028 4.11659894 16.28729277 -6.57448953   
## 151 152 153 154 155 156   
## 8.59137170 -5.94835333 -6.02645135 -1.85092492 -6.43627184 0.19952956   
## 157 158 159 160 161 162   
## -2.05409489 -1.88340106 6.56372816 0.87263969 -2.33053028 0.39303433   
## 163 164 165 166 167 168   
## -4.29805414 1.46629975 2.92792677 -1.77765950 -5.88340106 -1.94352073   
## 169 170 171 172 173 174   
## 11.81252001 -4.94352073 21.01085738 -4.84261177 7.66946972 1.22234050   
## 175 176 177 178 179 180   
## 8.49877589 -8.57448953 -0.84261177 -0.15983645 -8.05825147 -4.98914262   
## 181 182 183 184 185 186   
## 1.93138723 -5.05548711 -5.74518336 -2.13219291 7.25481664 31.14907508   
## 187 188 189 190 191 192   
## -7.71270723 -5.77765950 -2.79563785 -0.07207323 -0.94835333 -7.15983645   
## 193 194 195 196 197 198   
## 2.47596495 -5.98914262 -6.09072760 -3.08657103 -7.33053028 2.39303433   
## 199 200   
## 3.39786693 -0.15983645

#regular residuals   
resid(model1)

## 1 2 3 4 5 6   
## 10.66946972 -7.33053028 -7.44110444 -6.71270723 15.94174853 0.66946972   
## 7 8 9 10 11 12   
## -7.90621200 -1.07207323 -2.05409489 11.84016355 1.26312979 -6.71270723   
## 13 14 15 16 17 18   
## 21.42551047 -4.43627184 -4.88340106 -4.71270723 13.49877589 -5.71270723   
## 19 20 21 22 23 24   
## -2.19231259 27.66946972 0.32808206 7.80768741 -5.88340106 -1.67191794   
## 25 26 27 28 29 30   
## -6.50122411 -6.04442970 -1.88340106 -5.39548255 0.39303433 9.32808206   
## 31 32 33 34 35 36   
## 6.66946972 -1.26142143 2.11659894 -8.43627184 -1.74518336 16.56372816   
## 37 38 39 40 41 42   
## -6.05825147 -4.88340106 12.56720871 -1.82328138 9.17188602 -4.15983645   
## 43 44 45 46 47 48   
## -3.43627184 11.64665878 -6.09971677 -6.37959254 -5.24759966 -2.60696567   
## 49 50 51 52 53 54   
## -6.57448953 -6.19231259 -3.57448953 -0.15983645 -1.01332569 4.66946972   
## 55 56 57 58 59 60   
## -3.19231259 -5.80047044 -6.08173843 5.77521128 -3.67191794 -2.15983645   
## 61 62 63 64 65 66   
## -6.15983645 2.22234050 0.39303433 -3.16466905 -0.15983645 17.11659894   
## 67 68 69 70 71 72   
## -6.57448953 1.60451745 -5.39548255 -4.12320374 -2.94835333 -5.98914262   
## 73 74 75 76 77 78   
## -6.15983645 -2.57448953 -6.06861277 2.94590511 21.42551047 -6.85092492   
## 79 80 81 82 83 84   
## -4.82328138 13.01085738 -5.85092492 -6.60696567 0.03643268 6.56372816   
## 85 86 87 88 89 90   
## 2.50844108 -7.33053028 19.11659894 1.37022339 -0.98914262 -6.15983645   
## 91 92 93 94 95 96   
## -2.71753983 6.84016355 -4.50122411 -8.29805414 -4.88340106 -4.88340106   
## 97 98 99 100 101 102   
## 11.42551047 1.49877589 -6.85092492 -0.71753983 -4.77765950 7.11659894   
## 103 104 105 106 107 108   
## -2.88340106 2.49877589 -2.88340106 -4.57448953 3.24859182 1.11659894   
## 109 110 111 112 113 114   
## 1.11659894 6.01638609 -7.88340106 -5.64843098 2.66946972 -1.81013564   
## 115 116 117 118 119 120   
## -6.79563785 -3.50122411 6.56372816 17.94590511 0.80768741 1.05164667   
## 121 122 123 124 125 126   
## -2.71753983 11.01085738 2.39303433 -0.15983645 9.77521128 -3.36716299   
## 127 128 129 130 131 132   
## -6.50122411 2.19952956 -4.91587720 -7.88340106 7.12143154 -2.46874797   
## 133 134 135 136 137 138   
## -2.05409489 4.49877589 -3.11904716 11.28729277 -5.33053028 -7.38098476   
## 139 140 141 142 143 144   
## -4.79563785 -8.43627184 1.20436215 0.28729277 2.94590511 6.05164667   
## 145 146 147 148 149 150   
## -3.98914262 -3.68713193 -0.33053028 4.11659894 16.28729277 -6.57448953   
## 151 152 153 154 155 156   
## 8.59137170 -5.94835333 -6.02645135 -1.85092492 -6.43627184 0.19952956   
## 157 158 159 160 161 162   
## -2.05409489 -1.88340106 6.56372816 0.87263969 -2.33053028 0.39303433   
## 163 164 165 166 167 168   
## -4.29805414 1.46629975 2.92792677 -1.77765950 -5.88340106 -1.94352073   
## 169 170 171 172 173 174   
## 11.81252001 -4.94352073 21.01085738 -4.84261177 7.66946972 1.22234050   
## 175 176 177 178 179 180   
## 8.49877589 -8.57448953 -0.84261177 -0.15983645 -8.05825147 -4.98914262   
## 181 182 183 184 185 186   
## 1.93138723 -5.05548711 -5.74518336 -2.13219291 7.25481664 31.14907508   
## 187 188 189 190 191 192   
## -7.71270723 -5.77765950 -2.79563785 -0.07207323 -0.94835333 -7.15983645   
## 193 194 195 196 197 198   
## 2.47596495 -5.98914262 -6.09072760 -3.08657103 -7.33053028 2.39303433   
## 199 200   
## 3.39786693 -0.15983645

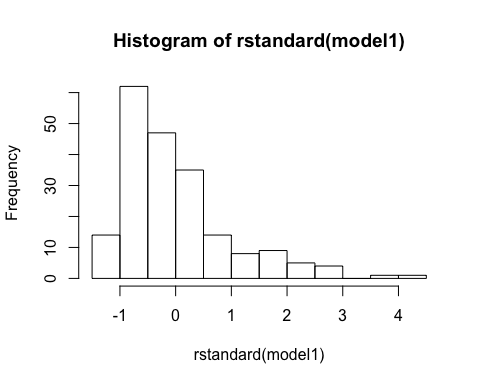
#regular residuals   
rstandard(model1)

## 1 2 3 4 5   
## 1.465072924 -1.006588116 -1.021616309 -0.923549719 2.197258725   
## 6 7 8 9 10   
## 0.091927901 -1.088781836 -0.147776650 -0.282230337 1.626092725   
## 11 12 13 14 15   
## 0.179672623 -0.923549719 2.946084187 -0.609705574 -0.670410958   
## 16 17 18 19 20   
## -0.648385114 1.858872972 -0.785967416 -0.301116170 3.799419461   
## 21 22 23 24 25   
## 0.045447447 1.072393120 -0.807694575 -0.231601819 -0.895262644   
## 26 27 28 29 30   
## -0.833038025 -0.258560109 -0.749407806 0.053954228 1.292169143   
## 31 32 33 34 35   
## 0.915814915 -0.173232616 0.290574357 -1.159451484 -0.239568537   
## 36 37 38 39 40   
## 2.276460451 -0.835011659 -0.670410958 1.798203704 -0.250981446   
## 41 42 43 44 45   
## 1.259102275 -0.571299523 -0.472269097 1.600271729 -0.840938796   
## 46 47 48 49 50   
## -0.879105890 -0.720678980 -0.357874131 -0.904015785 -0.850519887   
## 51 52 53 54 55   
## -0.491505073 -0.021951461 -0.139132648 0.641185910 -0.438467099   
## 56 57 58 59 60   
## -0.796268579 -0.835560107 0.796582055 -0.508651085 -0.296625492   
## 61 62 63 64 65   
## -0.845973554 0.305637394 0.053954228 -0.434697663 -0.021951461   
## 66 67 68 69 70   
## 2.349828605 -0.904015785 0.222860122 -0.749407806 -0.566418772   
## 71 72 73 74 75   
## -0.407472974 -0.825153734 -0.845973554 -0.354001503 -0.833307389   
## 76 77 78 79 80   
## 0.404764063 2.946084187 -0.943185570 -0.663942573 1.792570028   
## 81 82 83 84 85   
## -0.805512835 -0.906978609 0.005002002 0.902095677 0.344815131   
## 86 87 88 89 90   
## -1.006588116 2.624395838 0.188451843 -0.136279060 -0.845973554   
## 91 92 93 94 95   
## -0.373046515 0.939407647 -0.619849082 -1.139994753 -0.670410958   
## 96 97 98 99 100   
## -0.670410958 1.571048483 0.206391603 -0.943185570 -0.098499285   
## 101 102 103 104 105   
## -0.657069158 0.976992440 -0.395843725 0.344098384 -0.395843725   
## 106 107 108 109 110   
## -0.629008644 0.447027161 0.153290741 0.153290741 0.828881241   
## 111 112 113 114 115   
## -1.082261808 -0.775384523 0.366556906 -0.250446666 -0.935318661   
## 116 117 118 119 120   
## -0.482142301 0.902095677 2.465747262 0.110936617 0.145341330   
## 121 122 123 124 125   
## -0.373046515 1.517020158 0.328506467 -0.021951461 1.348307017   
## 126 127 128 129 130   
## -0.462672706 -0.895262644 0.301943832 -0.675728570 -1.082261808   
## 131 132 133 134 135   
## 0.980565796 -0.338933353 -0.282230337 0.619511945 -0.434536641   
## 136 137 138 139 140   
## 1.552931734 -0.731959112 -1.014242429 -0.660048353 -1.159451484   
## 141 142 143 144 145   
## 0.165762571 0.039526401 0.404764063 0.836359211 -0.549603864   
## 146 147 148 149 150   
## -0.506145771 -0.045386601 0.565141590 2.240843248 -0.904015785   
## 151 152 153 154 155   
## 1.180664539 -0.822083703 -0.828093813 -0.254821896 -0.884578529   
## 156 157 158 159 160   
## 0.027390730 -0.282230337 -0.258560109 0.902095677 0.120327398   
## 161 162 163 164 165   
## -0.320015605 0.053954228 -0.590470861 0.203381260 0.403591093   
## 166 167 168 169 170   
## -0.244480636 -0.807694575 -0.267128300 1.622389202 -0.679464986   
## 171 172 173 174 175   
## 2.894769506 -0.676953900 1.053129417 0.168107887 1.170339068   
## 176 177 178 179 180   
## -1.179022925 -0.117789605 -0.021951461 -1.110672603 -0.687378799   
## 181 182 183 184 185   
## 0.265207109 -0.696788656 -0.788665076 -0.292813213 0.995898678   
## 186 187 188 189 190   
## 4.288378351 -1.061132022 -0.794598666 -0.384778045 -0.009934714   
## 191 192 193 194 195   
## -0.131065822 -0.983310569 0.339907436 -0.825153734 -0.836374792   
## 196 197 198 199 200   
## -0.426136239 -1.006588116 0.328506467 0.467269169 -0.021951461

#regular residuals   
predict(model1)

## 1 2 3 4 5 6 7 8   
## 7.330530 7.330530 7.441104 8.712707 9.058251 7.330530 8.906212 9.072073   
## 9 10 11 12 13 14 15 16   
## 7.054095 8.159836 3.736870 8.712707 8.574490 8.436272 7.883401 8.712707   
## 17 18 19 20 21 22 23 24   
## 6.501224 8.712707 7.192313 7.330530 5.671918 7.192313 7.883401 5.671918   
## 25 26 27 28 29 30 31 32   
## 6.501224 9.044430 7.883401 5.395483 7.606966 5.671918 7.330530 7.261421   
## 33 34 35 36 37 38 39 40   
## 7.883401 8.436272 7.745183 8.436272 9.058251 7.883401 3.432791 8.823281   
## 41 42 43 44 45 46 47 48   
## 7.828114 8.159836 8.436272 8.353341 9.099717 6.379593 7.247600 7.606966   
## 49 50 51 52 53 54 55 56   
## 8.574490 7.192313 8.574490 8.159836 8.013326 7.330530 7.192313 7.800470   
## 57 58 59 60 61 62 63 64   
## 7.081738 6.224789 5.671918 8.159836 8.159836 6.777660 7.606966 7.164669   
## 65 66 67 68 69 70 71 72   
## 8.159836 7.883401 8.574490 5.395483 5.395483 7.123204 5.948353 8.989143   
## 73 74 75 76 77 78 79 80   
## 8.159836 8.574490 8.068613 7.054095 8.574490 8.850925 8.823281 8.989143   
## 81 82 83 84 85 86 87 88   
## 8.850925 7.606966 7.963567 8.436272 8.491559 7.330530 7.883401 8.629777   
## 89 90 91 92 93 94 95 96   
## 8.989143 8.159836 7.717540 8.159836 6.501224 8.298054 7.883401 7.883401   
## 97 98 99 100 101 102 103 104   
## 8.574490 6.501224 8.850925 7.717540 6.777660 7.883401 7.883401 6.501224   
## 105 106 107 108 109 110 111 112   
## 7.883401 8.574490 8.751408 7.883401 7.883401 8.983614 7.883401 7.648431   
## 113 114 115 116 117 118 119 120   
## 7.330530 5.810136 8.795638 6.501224 8.436272 7.054095 7.192313 5.948353   
## 121 122 123 124 125 126 127 128   
## 7.717540 8.989143 7.606966 8.159836 6.224789 8.367163 6.501224 7.800470   
## 129 130 131 132 133 134 135 136   
## 6.915877 7.883401 8.878568 7.468748 7.054095 6.501224 5.119047 8.712707   
## 137 138 139 140 141 142 143 144   
## 7.330530 8.380985 8.795638 8.436272 8.795638 8.712707 7.054095 5.948353   
## 145 146 147 148 149 150 151 152   
## 8.989143 7.687132 7.330530 7.883401 8.712707 8.574490 8.408628 5.948353   
## 153 154 155 156 157 158 159 160   
## 7.026451 8.850925 8.436272 7.800470 7.054095 7.883401 8.436272 9.127360   
## 161 162 163 164 165 166 167 168   
## 7.330530 7.606966 8.298054 5.533700 9.072073 6.777660 7.883401 6.943521   
## 169 170 171 172 173 174 175 176   
## 8.187480 6.943521 8.989143 4.842612 7.330530 6.777660 6.501224 8.574490   
## 177 178 179 180 181 182 183 184   
## 4.842612 8.159836 9.058251 8.989143 8.068613 9.055487 7.745183 8.132193   
## 185 186 187 188 189 190 191 192   
## 7.745183 8.850925 8.712707 6.777660 8.795638 9.072073 5.948353 8.159836   
## 193 194 195 196 197 198 199 200   
## 7.524035 8.989143 8.090728 6.086571 7.330530 7.606966 8.602133 8.159836

#regular residuals   
hist(rstandard(model1))



qqnorm(rstandard(model1))   
#produce qq plot   
qqline(rstandard(model1))   
library(mosaic)

## Loading required package: dplyr

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

## Loading required package: lattice

## Loading required package: ggplot2

## Loading required package: mosaicData

## Loading required package: Matrix

##   
## The 'mosaic' package masks several functions from core packages in order to add additional features.   
## The original behavior of these functions should not be affected by this.

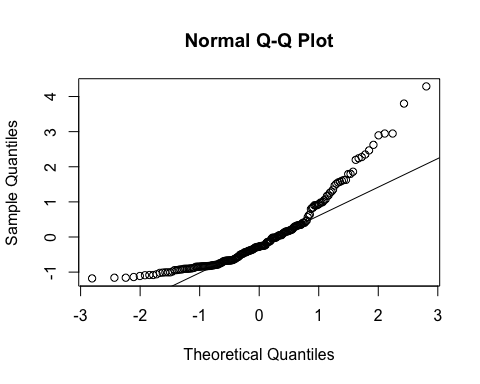
##   
## Attaching package: 'mosaic'

## The following object is masked from 'package:Matrix':  
##   
## mean

## The following objects are masked from 'package:dplyr':  
##   
## count, do, tally

## The following objects are masked from 'package:stats':  
##   
## binom.test, cor, cov, D, fivenum, IQR, median, prop.test,  
## quantile, sd, t.test, var

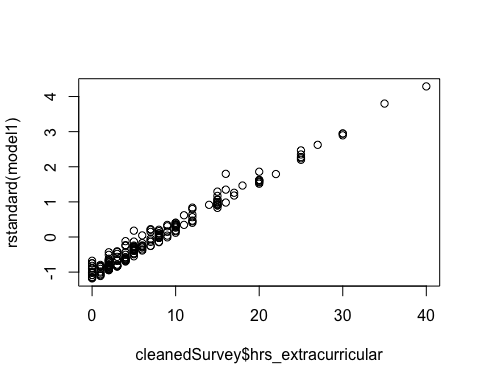
## The following objects are masked from 'package:base':  
##   
## max, mean, min, prod, range, sample, sum



favstats(rstandard(model1))

## min Q1 median Q3 max mean sd  
## -1.179023 -0.755902 -0.2746793 0.3409552 4.288378 0.0003007109 1.002529  
## n missing  
## 200 0

#It does not appear that the residuals are approximately normally distributed. They are right skewed.  
  
#4(b)  
#scatter plot of hrs extracurricular and standardized residuals   
plot(cleanedSurvey$hrs\_extracurricular,rstandard(model1))



#Is there any evidence of a non-linear trend in the residuals?   
# No  
#Is there any evidence of non-constant variance in the residuals?   
# No