Predicting World Revenue of Movies Through Statistical Analysis

Bilal Rehman

Summary of Goals

The goal of this report is to create a regression model that can help predict the world revenue of a movie based on information available shortly after its release. We will be using a dataset with various information for 40 movies for this report.

Summary of Statistics

Within the data set there are 16 variables, we can disregard the Title variable because it acts as like an id variable and it will not provide useful information. We can also disregard some of the redundant variables, like any variable that has to do with revenue, besides WorldRevenue, and we can also discount LOpening and Log.WorldRevenue which are log transformations of their respective variable. Not including these redundant variables can help us avoid multicollinearity since those variables are directly related to at least one other variable. Of the remaining 8 variables, 4 are quantitative, 3 are qualitative, and the last one is an indicator value which indicates whether or not the movie is a sequel.

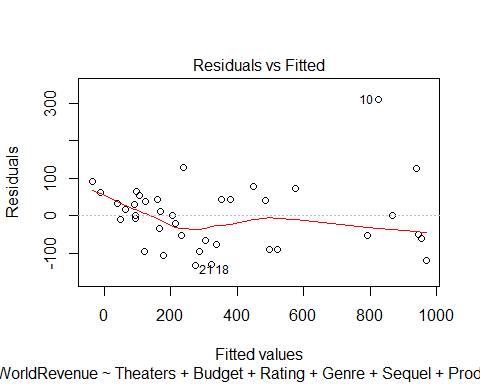
Regression

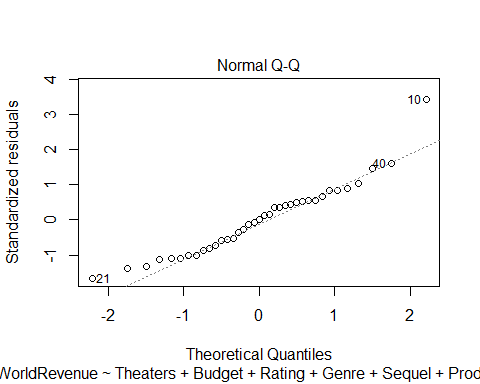
Regressing the data set by setting WorldRevenue as the dependent variable and every remaining variable as the dependent variable results in the following regression:

## Call:  
## lm(formula = WorldRevenue ~ Theaters + Budget + Rating + Genre +   
## Sequel + Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -131.65 -62.92 0.00 43.61 309.73   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 123.88735 285.10111 0.435 0.667780   
## Theaters 0.15808 0.09465 1.670 0.107865   
## Budget -0.71222 0.62473 -1.140 0.265518   
## RatingPG -443.69267 138.90056 -3.194 0.003894 \*\*   
## RatingPG-13 -378.46236 129.87987 -2.914 0.007607 \*\*   
## RatingR -295.80875 143.75614 -2.058 0.050637 .   
## GenreAdventure 165.43029 67.06526 2.467 0.021162 \*   
## GenreComedy 42.63922 76.78171 0.555 0.583808   
## GenreDrama 120.25767 92.74989 1.297 0.207104   
## GenreMusical 163.25128 141.89775 1.150 0.261279   
## GenreRomantic Comedy 252.60721 130.70568 1.933 0.065165 .   
## GenreThriller/Suspense 99.71452 94.85779 1.051 0.303637   
## Sequel 56.42361 80.87360 0.698 0.492084   
## Production.MethodDigital Animation -215.68263 99.06117 -2.177 0.039523 \*   
## Production.MethodLive Action -204.29728 69.80584 -2.927 0.007382 \*\*   
## Opening 4.64806 1.12636 4.127 0.000382 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 109.5 on 24 degrees of freedom  
## Multiple R-squared: 0.9233, Adjusted R-squared: 0.8754   
## F-statistic: 19.27 on 15 and 24 DF, p-value: 7.903e-10

Looking at the R-squared value (.9233) we can conclude that this model is very good because it explains the 92.33% of the variance of WorldRevenue.

Even though this is a fairly accurate model, there are a number of assumptions that need to be met to ensure that this is a valid regression. Two of the assumptions being that the variance of the errors is equal, and the distribution of the errors is normal. We can test for the presence of equal variance by looking at the residual vs fitted plot. If the equal variance assumption is valid the points on the plot should be random and not follow any trend. The residual vs fitted plot is below.

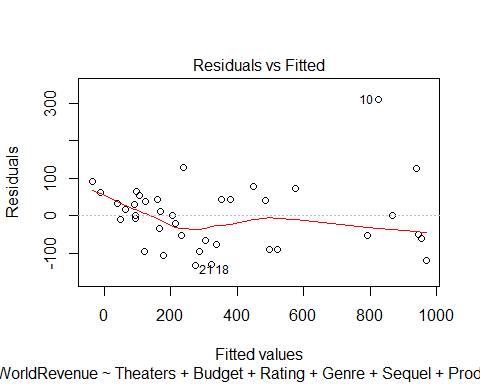
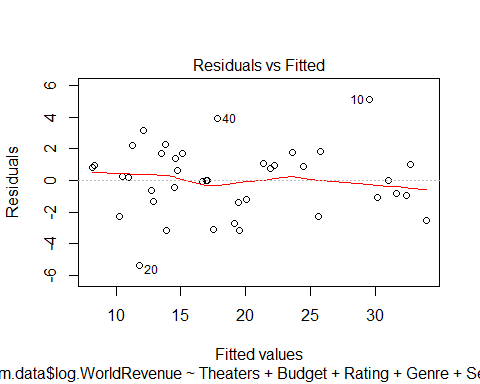
We can see a there is a trend amongst the points where they taper outwards. This means that the model fails the equal variance assumption. To test the normality of the errors, we can look at a QQ plot. If normality occurs, then the points should be roughly linear.



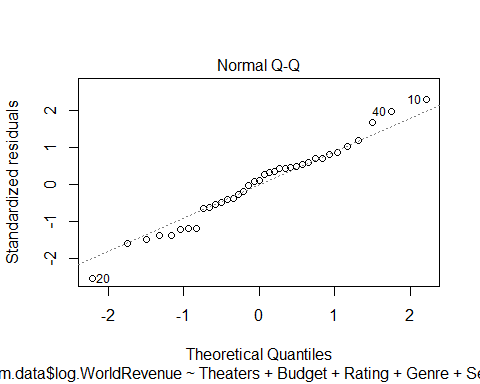
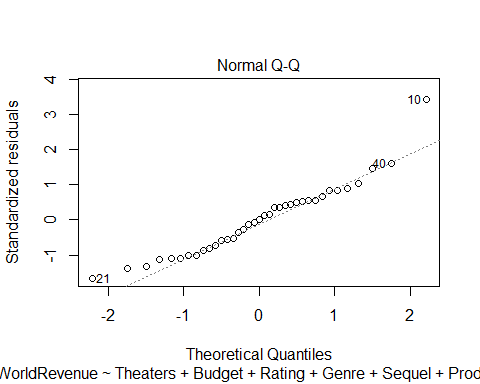
It’s hard to tell if the model is normal because there are points that don’t follow the linear trend towards the ends so it necessary to further test for normality. We can use the Shapiro-Wilk test, which is a hypothesis test with the null hypothesis being the distribution is normal and the alternative hypothesis being it is not normal, so if there is a small p-value we would reject the null. After performing the Shapiro-Wilk test we get a value of 0.01271. From the Shapiro-Wilk test we can conclude that normality assumption is not met. To get around our model’s failure of the two assumptions, we can perform an Box-Cox transform on the dependent variable by finding the lambda value (.375) and using the following formula: (Y^(lm)-1)/lm where Y is our dependent variable and lm is our lambda value. After the transformation, we use the transformed dependent variable in a new model.

## Call:  
## lm(formula = lm.data$log.WorldRevenue ~ Theaters + Budget + Rating +   
## Genre + Sequel + Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.3386 -1.2485 0.0988 1.1687 5.0950   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.315026 6.963914 -0.045 0.96429   
## Theaters 0.007605 0.002312 3.289 0.00309 \*\*  
## Budget -0.007645 0.015260 -0.501 0.62094   
## RatingPG -11.707293 3.392802 -3.451 0.00208 \*\*  
## RatingPG-13 -9.200069 3.172461 -2.900 0.00786 \*\*  
## RatingR -4.555374 3.511405 -1.297 0.20686   
## GenreAdventure 4.395198 1.638144 2.683 0.01300 \*   
## GenreComedy 2.389334 1.875479 1.274 0.21487   
## GenreDrama 4.272066 2.265520 1.886 0.07150 .   
## GenreMusical 8.067067 3.466011 2.327 0.02870 \*   
## GenreRomantic Comedy 10.523387 3.192632 3.296 0.00304 \*\*  
## GenreThriller/Suspense 2.982563 2.317008 1.287 0.21028   
## Sequel 1.946630 1.975428 0.985 0.33425   
## Production.MethodDigital Animation -3.049256 2.419680 -1.260 0.21972   
## Production.MethodLive Action -4.159781 1.705086 -2.440 0.02247 \*   
## Opening 0.072087 0.027513 2.620 0.01500 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.674 on 24 degrees of freedom  
## Multiple R-squared: 0.926, Adjusted R-squared: 0.8797   
## F-statistic: 20.01 on 15 and 24 DF, p-value: 5.288e-10

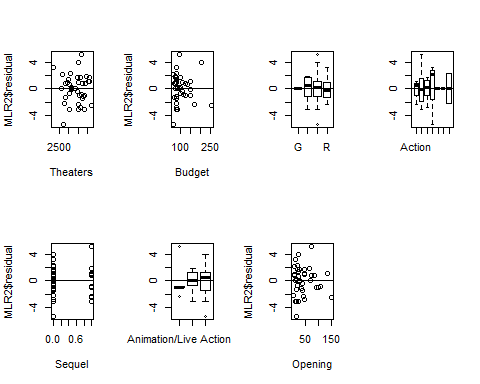
We can see that the transformation improved our R-squared value from .923 to .926 which means that the new model explains .3% more of the variance of WorldRevenue. To see if our transformation succeeded in meeting the assumptions, we look at the plots again.



We can see the effect the transformation had on the residual vs fitted graph. After the transformation there is no trend amongst the points which means that the equal variance assumption is met.



Looking at the QQ plot for the new model we can see the fitted line is close to having a slope of 1 which, along with the large p-value of the Shapiro-Wilk test (0.8845), means the normality assumption is met. Now that we have confirmed that these 2 assumptions have been met, we can test for the linearity assumption. To do this we plot each independent variable with the residuals. Ideally, we do not want to see any pattern.



We can see that the see that there is no pattern in any of the graphs so we can conclude that the linearity assumption is met. There is one more assumption to check for, independence of errors, but it does not apply here since this is not a time-series model.

Now that we have confirmed the validity of our model by checking each assumption, we can create the best possible model by removing unnecessary variables from our model. To do this we can run a backward elimination code that starts off with every variable in the model and removes variables one-by-one until the best possible model was made (had the lowest AIC). The variables Budget and Sequel were removed. The best possible model is as follows:

## Call:  
## lm(formula = lm.data$log.WorldRevenue ~ Theaters + Rating + Genre +   
## Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.4155 -1.3187 -0.0335 1.3217 5.6127   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.486966 6.682237 -0.223 0.825644   
## Theaters 0.008066 0.002111 3.821 0.000745 \*\*\*  
## RatingPG -11.073662 3.304296 -3.351 0.002470 \*\*   
## RatingPG-13 -9.169193 3.126566 -2.933 0.006926 \*\*   
## RatingR -4.551884 3.409386 -1.335 0.193415   
## GenreAdventure 3.963966 1.565538 2.532 0.017722 \*   
## GenreComedy 1.492637 1.590132 0.939 0.356530   
## GenreDrama 4.376956 2.228776 1.964 0.060325 .   
## GenreMusical 7.137342 3.296739 2.165 0.039751 \*   
## GenreRomantic Comedy 10.472023 3.149332 3.325 0.002637 \*\*   
## GenreThriller/Suspense 2.812569 2.273707 1.237 0.227145   
## Production.MethodDigital Animation -4.276815 1.956294 -2.186 0.037998 \*   
## Production.MethodLive Action -4.921447 1.553263 -3.168 0.003895 \*\*   
## Opening 0.080017 0.024617 3.250 0.003178 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.64 on 26 degrees of freedom  
## Multiple R-squared: 0.9218, Adjusted R-squared: 0.8827   
## F-statistic: 23.59 on 13 and 26 DF, p-value: 4.158e-11

The best model for calculating the world revenue of a movie is based on the rating, genre, the production method, the number of theaters it played in, and the opening weekend revenue. The rating with the most negative impact on the world revenue is PG, this means that, according to our model having a rating of PG is bad for the world revenue. Our model suggests that romantic comedies are most profitable genre. A movie with a hybrid production method of live action and digital animation has a better impact on revenue. Both the number of theaters the movie’s revenue on the opening weekend both have positive impacts on world revenue. If we wanted to create a movie with the highest world revenue based on our model, it would be a G rated romantic comedy with hybrid production that opened in as many theaters as possible.

Appendix

Project-Code.R

2019-12-10

lm.data<-read.delim("boxoffice.txt")  
attach(lm.data)  
summary(lm.data)

## Title USRelease Genre Rating   
## Beowulf : 1 5/27/2005 : 2 Action :10 G : 1   
## Cat in the Hat, The: 1 10/28/2005: 1 Adventure :15 PG :10   
## Catwoman : 1 11/14/2003: 1 Comedy : 8 PG-13:23   
## Click : 1 11/16/2007: 1 Drama : 3 R : 6   
## Cold Mountain : 1 11/18/2005: 1 Musical : 1   
## Core, The : 1 11/21/2003: 1 Romantic Comedy : 1   
## (Other) :34 (Other) :33 Thriller/Suspense: 2   
## Rating1 Production.Method Sequel Budget   
## Min. :0.000 Animation/Live Action: 5 Min. :0.0 Min. : 65.0   
## 1st Qu.:0.000 Digital Animation : 7 1st Qu.:0.0 1st Qu.: 80.0   
## Median :1.000 Live Action :28 Median :0.0 Median : 85.0   
## Mean :0.725 Mean :0.3 Mean :104.0   
## 3rd Qu.:1.000 3rd Qu.:1.0 3rd Qu.:117.5   
## Max. :1.000 Max. :1.0 Max. :258.0   
##   
## Opening Theaters USRevenue IntRevenue   
## Min. : 5.951 Min. :2163 Min. : 2.071 Min. : 20.90   
## 1st Qu.: 17.374 1st Qu.:3020 1st Qu.: 62.329 1st Qu.: 42.03   
## Median : 27.496 Median :3490 Median : 98.325 Median :108.80   
## Mean : 42.956 Mean :3380 Mean :136.890 Mean :205.16   
## 3rd Qu.: 53.365 3rd Qu.:3743 3rd Qu.:185.652 3rd Qu.:305.40   
## Max. :151.117 Max. :4252 Max. :423.316 Max. :756.00   
##   
## WorldRevenue Profit LOpening   
## Min. : 26.67 Min. :0.000 Min. :1.784   
## 1st Qu.: 130.76 1st Qu.:0.000 1st Qu.:2.855   
## Median : 198.31 Median :0.000 Median :3.314   
## Mean : 342.05 Mean :0.475 Mean :3.455   
## 3rd Qu.: 456.51 3rd Qu.:1.000 3rd Qu.:3.966   
## Max. :1133.03 Max. :1.000 Max. :5.018   
##

MLR<-lm(WorldRevenue~Theaters+Budget+Rating+Genre+Sequel+Production.Method+Opening)  
MLR

##   
## Call:  
## lm(formula = WorldRevenue ~ Theaters + Budget + Rating + Genre +   
## Sequel + Production.Method + Opening)  
##   
## Coefficients:  
## (Intercept) Theaters   
## 123.8874 0.1581   
## Budget RatingPG   
## -0.7122 -443.6927   
## RatingPG-13 RatingR   
## -378.4624 -295.8087   
## GenreAdventure GenreComedy   
## 165.4303 42.6392   
## GenreDrama GenreMusical   
## 120.2577 163.2513   
## GenreRomantic Comedy GenreThriller/Suspense   
## 252.6072 99.7145   
## Sequel Production.MethodDigital Animation   
## 56.4236 -215.6826   
## Production.MethodLive Action Opening   
## -204.2973 4.6481

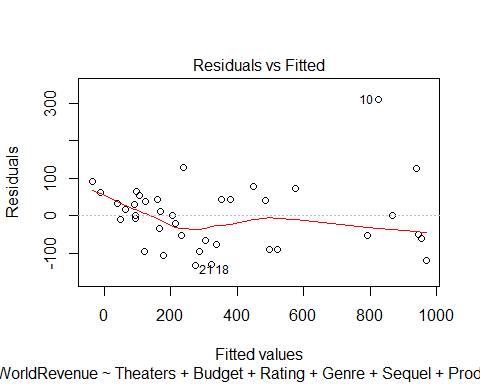
anova(MLR)

## Analysis of Variance Table  
##   
## Response: WorldRevenue  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Theaters 1 1596584 1596584 133.1997 2.789e-11 \*\*\*  
## Budget 1 145718 145718 12.1570 0.0019040 \*\*   
## Rating 3 560622 186874 15.5905 7.725e-06 \*\*\*  
## Genre 6 524033 87339 7.2865 0.0001620 \*\*\*  
## Sequel 1 306098 306098 25.5371 3.631e-05 \*\*\*  
## Production.Method 2 128133 64067 5.3449 0.0120255 \*   
## Opening 1 204116 204116 17.0290 0.0003824 \*\*\*  
## Residuals 24 287673 11986   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(MLR)

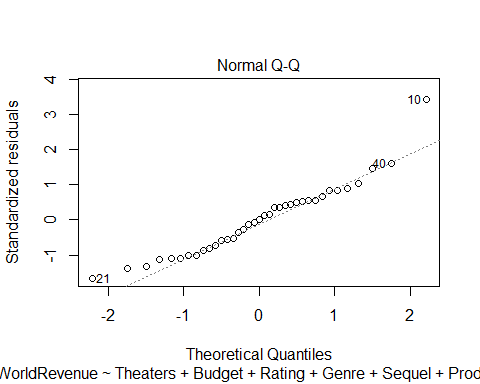
##   
## Call:  
## lm(formula = WorldRevenue ~ Theaters + Budget + Rating + Genre +   
## Sequel + Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -131.65 -62.92 0.00 43.61 309.73   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 123.88735 285.10111 0.435 0.667780   
## Theaters 0.15808 0.09465 1.670 0.107865   
## Budget -0.71222 0.62473 -1.140 0.265518   
## RatingPG -443.69267 138.90056 -3.194 0.003894 \*\*   
## RatingPG-13 -378.46236 129.87987 -2.914 0.007607 \*\*   
## RatingR -295.80875 143.75614 -2.058 0.050637 .   
## GenreAdventure 165.43029 67.06526 2.467 0.021162 \*   
## GenreComedy 42.63922 76.78171 0.555 0.583808   
## GenreDrama 120.25767 92.74989 1.297 0.207104   
## GenreMusical 163.25128 141.89775 1.150 0.261279   
## GenreRomantic Comedy 252.60721 130.70568 1.933 0.065165 .   
## GenreThriller/Suspense 99.71452 94.85779 1.051 0.303637   
## Sequel 56.42361 80.87360 0.698 0.492084   
## Production.MethodDigital Animation -215.68263 99.06117 -2.177 0.039523 \*   
## Production.MethodLive Action -204.29728 69.80584 -2.927 0.007382 \*\*   
## Opening 4.64806 1.12636 4.127 0.000382 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 109.5 on 24 degrees of freedom  
## Multiple R-squared: 0.9233, Adjusted R-squared: 0.8754   
## F-statistic: 19.27 on 15 and 24 DF, p-value: 7.903e-10

## Equal variance meet due to flattness of line in first plot  
plot(MLR, 1)



## Normality not meet due to lack of fully diagonal line in second plot and since the p-value is small in the shapiro-wilk test  
plot(MLR, 2)

## Warning: not plotting observations with leverage one:  
## 31, 33, 38



shapiro.test(MLR$residual)

##   
## Shapiro-Wilk normality test  
##   
## data: MLR$residual  
## W = 0.92683, p-value = 0.01271

## Since equal variance is not met we must transform y  
library(MASS)

## Warning: package 'MASS' was built under R version 3.5.3

trans <- boxcox(WorldRevenue~Theaters+Budget+Rating+Genre+Sequel+Production.Method+Opening, data=lm.data, plotit=F, lambda = seq(-3, 3, by=0.125))  
maxyentry <- which.max(trans$y)  
trans$x[maxyentry]

## [1] 0.375

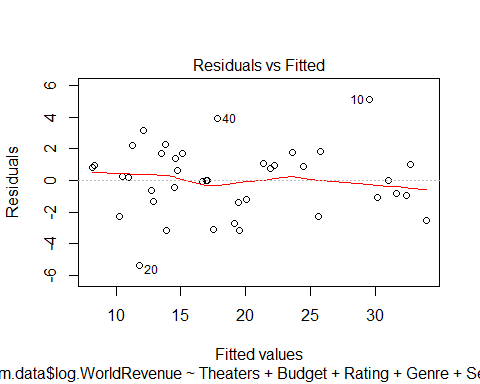
##Since lamda is .375 we transform cost with the following formula: (Y^(lm)-1)/lm where lm is .25  
lm.data$log.WorldRevenue <- ((WorldRevenue^.375)-1)/.375  
  
  
MLR2<-lm(lm.data$log.WorldRevenue~Theaters+Budget+Rating+Genre+Sequel+Production.Method+Opening)  
anova(MLR2)

## Analysis of Variance Table  
##   
## Response: lm.data$log.WorldRevenue  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Theaters 1 1190.49 1190.49 166.4673 2.747e-12 \*\*\*  
## Budget 1 87.64 87.64 12.2542 0.0018397 \*\*   
## Rating 3 319.71 106.57 14.9019 1.095e-05 \*\*\*  
## Genre 6 342.07 57.01 7.9720 8.433e-05 \*\*\*  
## Sequel 1 105.65 105.65 14.7734 0.0007815 \*\*\*  
## Production.Method 2 52.39 26.19 3.6627 0.0409043 \*   
## Opening 1 49.10 49.10 6.8652 0.0150039 \*   
## Residuals 24 171.64 7.15   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

summary(MLR2)

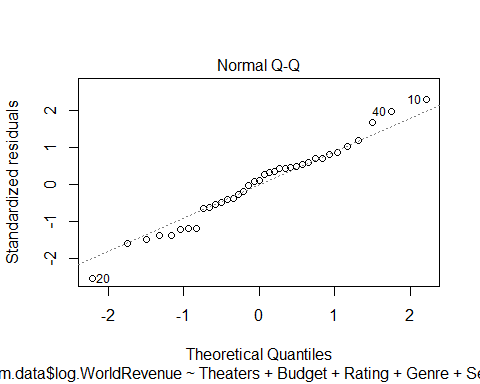
##   
## Call:  
## lm(formula = lm.data$log.WorldRevenue ~ Theaters + Budget + Rating +   
## Genre + Sequel + Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.3386 -1.2485 0.0988 1.1687 5.0950   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.315026 6.963914 -0.045 0.96429   
## Theaters 0.007605 0.002312 3.289 0.00309 \*\*  
## Budget -0.007645 0.015260 -0.501 0.62094   
## RatingPG -11.707293 3.392802 -3.451 0.00208 \*\*  
## RatingPG-13 -9.200069 3.172461 -2.900 0.00786 \*\*  
## RatingR -4.555374 3.511405 -1.297 0.20686   
## GenreAdventure 4.395198 1.638144 2.683 0.01300 \*   
## GenreComedy 2.389334 1.875479 1.274 0.21487   
## GenreDrama 4.272066 2.265520 1.886 0.07150 .   
## GenreMusical 8.067067 3.466011 2.327 0.02870 \*   
## GenreRomantic Comedy 10.523387 3.192632 3.296 0.00304 \*\*  
## GenreThriller/Suspense 2.982563 2.317008 1.287 0.21028   
## Sequel 1.946630 1.975428 0.985 0.33425   
## Production.MethodDigital Animation -3.049256 2.419680 -1.260 0.21972   
## Production.MethodLive Action -4.159781 1.705086 -2.440 0.02247 \*   
## Opening 0.072087 0.027513 2.620 0.01500 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.674 on 24 degrees of freedom  
## Multiple R-squared: 0.926, Adjusted R-squared: 0.8797   
## F-statistic: 20.01 on 15 and 24 DF, p-value: 5.288e-10

## Equal variance meet due to flattness of line in first plot  
plot(MLR2, 1)



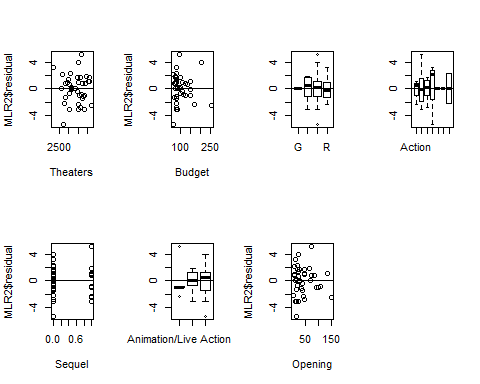
## Normality meet due to diagonal line in qq plot and the large p-value in the shapiro-wilk test  
plot(MLR2, 2)

## Warning: not plotting observations with leverage one:  
## 31, 33, 38



shapiro.test(MLR2$residual)

##   
## Shapiro-Wilk normality test  
##   
## data: MLR2$residual  
## W = 0.98568, p-value = 0.8845

## Linearity Check  
par(mfrow=c(2,4))  
plot(x = Theaters, y = MLR2$residual)  
abline(h=0)  
plot(x = Budget, y = MLR2$residual)  
abline(h=0)  
plot(x = Rating, y = MLR2$residual)  
abline(h=0)  
plot(x = Genre, y = MLR2$residual)  
abline(h=0)  
plot(x = Sequel, y = MLR2$residual)  
abline(h=0)  
plot(x = Production.Method, y = MLR2$residual)  
abline(h=0)  
plot(x = Opening, y = MLR2$residual)  
abline(h=0)  


## Linearity met due to randomness of every plot   
  
  
## Backward selection  
step(MLR2, direction = "backward")

## Start: AIC=90.26  
## lm.data$log.WorldRevenue ~ Theaters + Budget + Rating + Genre +   
## Sequel + Production.Method + Opening  
##   
## Df Sum of Sq RSS AIC  
## - Budget 1 1.795 173.43 88.676  
## - Sequel 1 6.945 178.58 89.846  
## <none> 171.64 90.260  
## - Production.Method 2 44.049 215.69 95.398  
## - Opening 1 49.097 220.73 98.323  
## - Genre 6 116.883 288.52 99.035  
## - Theaters 1 77.382 249.02 103.146  
## - Rating 3 121.520 293.16 105.673  
##   
## Step: AIC=88.68  
## lm.data$log.WorldRevenue ~ Theaters + Rating + Genre + Sequel +   
## Production.Method + Opening  
##   
## Df Sum of Sq RSS AIC  
## - Sequel 1 7.813 181.25 88.439  
## <none> 173.43 88.676  
## - Production.Method 2 51.021 224.45 94.991  
## - Opening 1 47.306 220.74 96.324  
## - Genre 6 115.158 288.59 97.045  
## - Theaters 1 76.101 249.53 101.228  
## - Rating 3 120.030 293.46 103.715  
##   
## Step: AIC=88.44  
## lm.data$log.WorldRevenue ~ Theaters + Rating + Genre + Production.Method +   
## Opening  
##   
## Df Sum of Sq RSS AIC  
## <none> 181.25 88.439  
## - Genre 6 118.615 299.86 96.577  
## - Production.Method 2 70.269 251.51 97.545  
## - Opening 1 73.652 254.90 100.079  
## - Rating 3 116.555 297.80 102.302  
## - Theaters 1 101.769 283.01 104.265

##   
## Call:  
## lm(formula = lm.data$log.WorldRevenue ~ Theaters + Rating + Genre +   
## Production.Method + Opening)  
##   
## Coefficients:  
## (Intercept) Theaters   
## -1.486966 0.008066   
## RatingPG RatingPG-13   
## -11.073662 -9.169193   
## RatingR GenreAdventure   
## -4.551884 3.963966   
## GenreComedy GenreDrama   
## 1.492637 4.376956   
## GenreMusical GenreRomantic Comedy   
## 7.137342 10.472023   
## GenreThriller/Suspense Production.MethodDigital Animation   
## 2.812569 -4.276815   
## Production.MethodLive Action Opening   
## -4.921447 0.080017

final<-lm(formula = lm.data$log.WorldRevenue ~ Theaters + Rating + Genre +   
 Production.Method + Opening)  
summary(final)

##   
## Call:  
## lm(formula = lm.data$log.WorldRevenue ~ Theaters + Rating + Genre +   
## Production.Method + Opening)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.4155 -1.3187 -0.0335 1.3217 5.6127   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.486966 6.682237 -0.223 0.825644   
## Theaters 0.008066 0.002111 3.821 0.000745 \*\*\*  
## RatingPG -11.073662 3.304296 -3.351 0.002470 \*\*   
## RatingPG-13 -9.169193 3.126566 -2.933 0.006926 \*\*   
## RatingR -4.551884 3.409386 -1.335 0.193415   
## GenreAdventure 3.963966 1.565538 2.532 0.017722 \*   
## GenreComedy 1.492637 1.590132 0.939 0.356530   
## GenreDrama 4.376956 2.228776 1.964 0.060325 .   
## GenreMusical 7.137342 3.296739 2.165 0.039751 \*   
## GenreRomantic Comedy 10.472023 3.149332 3.325 0.002637 \*\*   
## GenreThriller/Suspense 2.812569 2.273707 1.237 0.227145   
## Production.MethodDigital Animation -4.276815 1.956294 -2.186 0.037998 \*   
## Production.MethodLive Action -4.921447 1.553263 -3.168 0.003895 \*\*   
## Opening 0.080017 0.024617 3.250 0.003178 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.64 on 26 degrees of freedom  
## Multiple R-squared: 0.9218, Adjusted R-squared: 0.8827   
## F-statistic: 23.59 on 13 and 26 DF, p-value: 4.158e-11