

R Notebook

[Code ▾](#)[Hide](#)

```
data <- read.csv("modified_df6.csv")
data <- subset(data,select=-c(screen_KOR,revenue_KOR,vote_count_KOR))
movie <- na.omit(data)
```

[Hide](#)

```
movie$audience_KOR <- ifelse(movie$audience_KOR==0,mean(movie$audience_KOR),movie$audience_KOR)
movie$production_year <- as.factor(movie$production_year)
movie$budget <- ifelse(movie$budget==0,mean(movie$budget),movie$budget)
movie$revenue_US <- ifelse(movie$revenue_US==0,mean(movie$revenue_US),movie$revenue_US)
movie$vote_average_US <- ifelse(movie$vote_average_US==0,mean(movie$vote_average_US),movie$vote_average_US)
movie$vote_count_US <- ifelse(movie$vote_count_US==0,mean(movie$vote_count_US),movie$vote_count_US)
movie$vote_average_KOR <- ifelse(movie$vote_average_KOR==0,mean(movie$vote_average_KOR),movie$vote_average_KOR)
movie <- transform(movie,
  production_year_2003 = ifelse(production_year=="2003", 1, 0),
  production_year_2004 = ifelse(production_year=="2004", 1, 0),
  production_year_2005 = ifelse(production_year=="2005", 1, 0),
  production_year_2006 = ifelse(production_year=="2006", 1, 0),
  production_year_2007 = ifelse(production_year=="2007", 1, 0),
  production_year_2008 = ifelse(production_year=="2008", 1, 0),
  production_year_2009 = ifelse(production_year=="2009", 1, 0),
  production_year_2010 = ifelse(production_year=="2010", 1, 0),
  production_year_2011 = ifelse(production_year=="2011", 1, 0),
  production_year_2012 = ifelse(production_year=="2012", 1, 0),
  production_year_2013 = ifelse(production_year=="2013", 1, 0),
  production_year_2014 = ifelse(production_year=="2014", 1, 0),
  production_year_2015 = ifelse(production_year=="2015", 1, 0),
  production_year_2016 = ifelse(production_year=="2016", 1, 0),
  production_year_2017 = ifelse(production_year=="2017", 1, 0),
  production_year_2018 = ifelse(production_year=="2018", 1, 0),
  production_year_2019 = ifelse(production_year=="2019", 1, 0))
```

[Hide](#)

```
str(movie)
```

```

'data.frame':  1029 obs. of  34 variables:
 $ audience_KOR      : num  4966571 6396528 4935660 1011163 3002669 ...
 $ production_year   : Factor w/ 18 levels "2002","2003",...: 6 11 6 9 8 15 5 7 11 12 ...
 $ release_diff.US.KOR.: int   0 0 0 2 0 0 0 0 0 0 ...
 $ budget            : num   3.00e+08 2.50e+08 2.58e+08 2.60e+08 2.50e+08 2.50e+08 2.70e+08
2.00e+08 2.55e+08 2.25e+08 ...
 $ popularity_US     : num   139.1 112.3 115.7 48.7 98.9 ...
 $ revenue_US        : num   9.61e+08 1.08e+09 8.91e+08 5.92e+08 9.34e+08 ...
 $ runtime           : int   169 165 139 100 153 151 154 106 149 143 ...
 $ vote_average_US   : num    6.9 7.6 5.9 7.4 7.4 5.7 5.4 6.1 5.9 6.5 ...
 $ vote_count_US     : num   4500 9106 3576 3330 5293 ...
 $ vote_average_KOR  : num    8.21 9.02 7.44 9.34 6.95 6.86 6.9 7.09 7.32 7.92 ...
 $ vote_diff.US.KOR. : num   -1.31 -1.42 -1.54 -1.94 0.45 -1.16 -1.5 -0.99 -1.42 -1.42 ...
 $ nomi_DIR          : int    3 22 0 0 7 2 12 5 0 2 ...
 $ awards_DIR        : int    3 8 0 0 4 1 5 2 0 1 ...
 $ nomi_Actor1       : int   33 28 0 4 19 24 3 10 0 2 ...
 $ awards_Actor1     : int   28 8 0 1 5 10 2 0 0 1 ...
 $ nomi_Actor2       : int   13 15 0 0 22 2 23 1 0 32 ...
 $ awards_Actor2     : int    7 1 0 0 8 1 6 0 0 4 ...
 $ production_year_2003: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2004: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2005: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2006: num    0 0 0 0 0 0 1 0 0 0 ...
 $ production_year_2007: num    1 0 1 0 0 0 0 0 0 0 ...
 $ production_year_2008: num    0 0 0 0 0 0 0 1 0 0 ...
 $ production_year_2009: num    0 0 0 0 1 0 0 0 0 0 ...
 $ production_year_2010: num    0 0 0 1 0 0 0 0 0 0 ...
 $ production_year_2011: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2012: num    0 1 0 0 0 0 0 0 1 0 ...
 $ production_year_2013: num    0 0 0 0 0 0 0 0 0 1 ...
 $ production_year_2014: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2015: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2016: num    0 0 0 0 0 1 0 0 0 0 ...
 $ production_year_2017: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2018: num    0 0 0 0 0 0 0 0 0 0 ...
 $ production_year_2019: num    0 0 0 0 0 0 0 0 0 0 ...

```

Hide

```
summary(movie)
```

audience_KOR	production_year	release_diff.US.KOR.	budget
Min. : 1	2015 : 94	Min. : 0.000	Min. : 1
1st Qu.: 40573	2014 : 91	1st Qu.: 0.000	1st Qu.: 27000000
Median : 202667	2008 : 87	Median : 1.000	Median : 57000000
Mean : 659323	2013 : 87	Mean : 6.077	Mean : 69726741
3rd Qu.: 695566	2011 : 84	3rd Qu.: 4.000	3rd Qu.: 99000000
Max. : 10296101	2010 : 82	Max. : 180.000	Max. : 380000000

(Other):504

popularity_US	revenue_US	runtime	vote_average_US
Min. : 0.0181	Min. : 1.487e+04	Min. : 73.0	Min. : 2.800
1st Qu.: 19.5748	1st Qu.: 6.463e+07	1st Qu.: 97.0	1st Qu.: 5.700
Median : 34.5769	Median : 1.492e+08	Median : 108.0	Median : 6.300
Mean : 45.6601	Mean : 2.215e+08	Mean : 111.2	Mean : 6.246
3rd Qu.: 53.1619	3rd Qu.: 2.753e+08	3rd Qu.: 122.0	3rd Qu.: 6.700
Max. : 875.5813	Max. : 1.520e+09	Max. : 187.0	Max. : 8.200

vote_count_US	vote_average_KOR	vote_diff.US.KOR.	nomi_DIR
Min. : 2	Min. : 1.250	Min. : -723.300	Min. : 0.00
1st Qu.: 440	1st Qu.: 6.900	1st Qu.: -1.820	1st Qu.: 0.00
Median : 1045	Median : 7.620	Median : -1.330	Median : 0.00
Mean : 1664	Mean : 8.171	Mean : -1.937	Mean : 3.09
3rd Qu.: 2225	3rd Qu.: 8.350	3rd Qu.: -0.800	3rd Qu.: 2.00
Max. : 13752	Max. : 730.000	Max. : 5.150	Max. : 67.00

awards_DIR	nomi_Actor1	awards_Actor1	nomi_Actor2	awards_Actor2
Min. : 0.00	Min. : 0.000	Min. : 0.000	Min. : 0.000	Min. : 0.000
1st Qu.: 0.00	1st Qu.: 0.000	1st Qu.: 0.000	1st Qu.: 0.000	1st Qu.: 0.000
Median : 0.00	Median : 0.000	Median : 0.000	Median : 0.000	Median : 0.000
Mean : 1.31	Mean : 6.369	Mean : 2.129	Mean : 4.065	Mean : 1.289
3rd Qu.: 0.00	3rd Qu.: 7.000	3rd Qu.: 1.000	3rd Qu.: 4.000	3rd Qu.: 0.000
Max. : 34.00	Max. : 89.000	Max. : 35.000	Max. : 89.000	Max. : 35.000

production_year_2003	production_year_2004	production_year_2005	production_year_2006
Min. : 0.000000	Min. : 0.000000	Min. : 0.000000	Min. : 0.000000
1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000
Median : 0.000000	Median : 0.000000	Median : 0.000000	Median : 0.000000
Mean : 0.003887	Mean : 0.07094	Mean : 0.07386	Mean : 0.0758
3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000
Max. : 1.000000	Max. : 1.000000	Max. : 1.000000	Max. : 1.000000

production_year_2007	production_year_2008	production_year_2009	production_year_2010
Min. : 0.000000	Min. : 0.000000	Min. : 0.000000	Min. : 0.000000
1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000
Median : 0.000000	Median : 0.000000	Median : 0.000000	Median : 0.000000
Mean : 0.07191	Mean : 0.08455	Mean : 0.07677	Mean : 0.07969
3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000
Max. : 1.000000	Max. : 1.000000	Max. : 1.000000	Max. : 1.000000

production_year_2011	production_year_2012	production_year_2013	production_year_2014
Min. : 0.000000	Min. : 0.000000	Min. : 0.000000	Min. : 0.000000
1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000	1st Qu.: 0.000000
Median : 0.000000	Median : 0.000000	Median : 0.000000	Median : 0.000000
Mean : 0.08163	Mean : 0.06706	Mean : 0.08455	Mean : 0.08844
3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000	3rd Qu.: 0.000000
Max. : 1.000000	Max. : 1.000000	Max. : 1.000000	Max. : 1.000000

production_year_2015	production_year_2016	production_year_2017	production_year_2018
Min. :0.00000	Min. :0.00000	Min. :0.0000000	Min. :0.000000
1st Qu.:0.00000	1st Qu.:0.00000	1st Qu.:0.0000000	1st Qu.:0.000000
Median :0.00000	Median :0.00000	Median :0.0000000	Median :0.000000
Mean :0.09135	Mean :0.03693	Mean :0.0009718	Mean :0.007775
3rd Qu.:0.00000	3rd Qu.:0.00000	3rd Qu.:0.0000000	3rd Qu.:0.000000
Max. :1.00000	Max. :1.00000	Max. :1.0000000	Max. :1.000000

```
production_year_2019
Min. :0.000000
1st Qu.:0.000000
Median :0.000000
Mean :0.002916
3rd Qu.:0.000000
Max. :1.000000
```

Hide

```
par(mfrow=c(2,5))
hist(movie$audience_KOR)
hist(movie$release_diff.US.KOR.)
```

Hide

```
hist(movie$budget)
hist(movie$popularity_US)
```

Hide

```
hist(movie$revenue_US)
hist(movie$runtime)
```

Hide

```
hist(movie$vote_average_US)
hist(movie$vote_count_US)
```

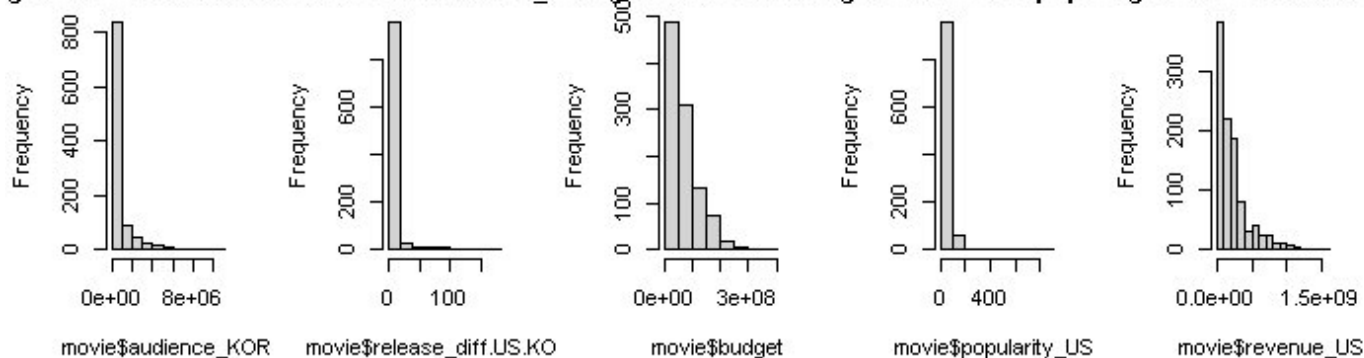
Hide

```
hist(movie$vote_average_KOR)
hist(movie$vote_diff.US.KOR.)
```

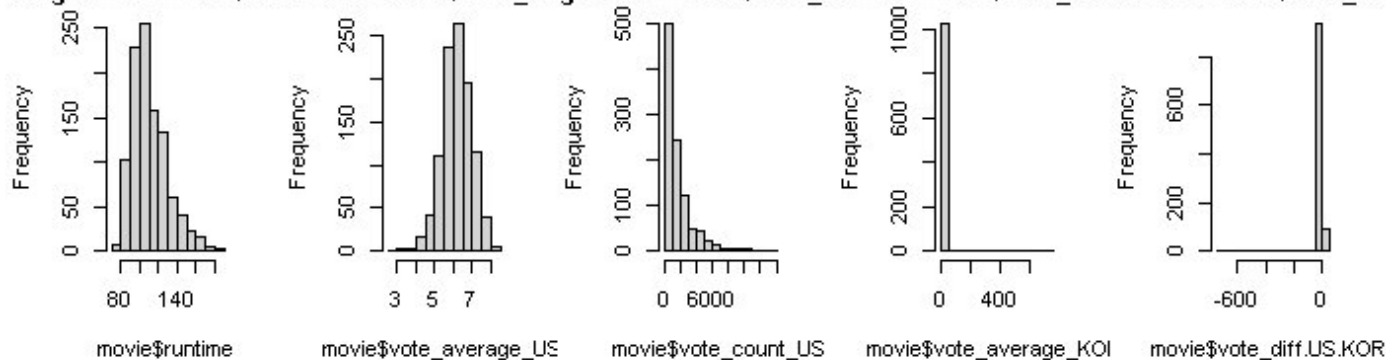
Hide

```
par(mfrow=c(1,1))
par(mfrow=c(2,3))
```

gram of movie\$audienm of movie\$release_ distogram of movie\$bugram of movie\$popultogram of movie\$rever



istogram of movie\$rurram of movie\$vote_avgram of movie\$vote_cam of movie\$vote_averam of movie\$vote_dif



Hide

```
hist(movie$nomi_DIR )
hist(movie$awards_DIR)
```

Hide

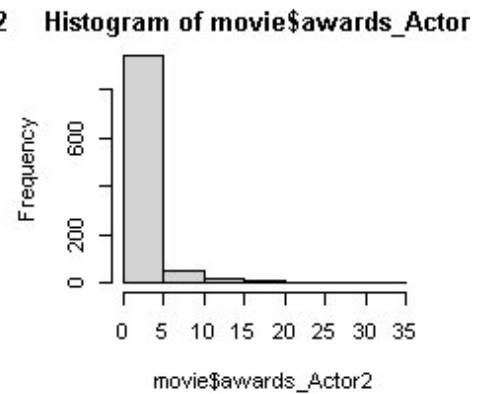
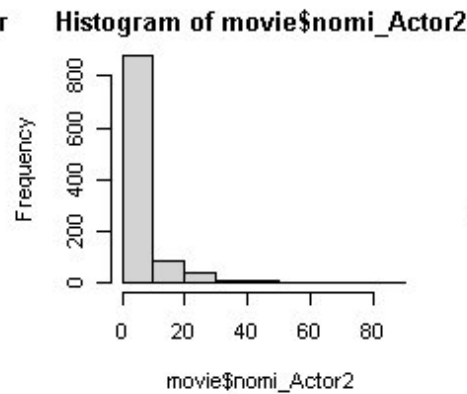
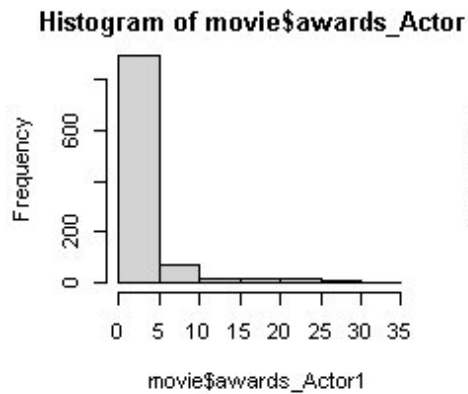
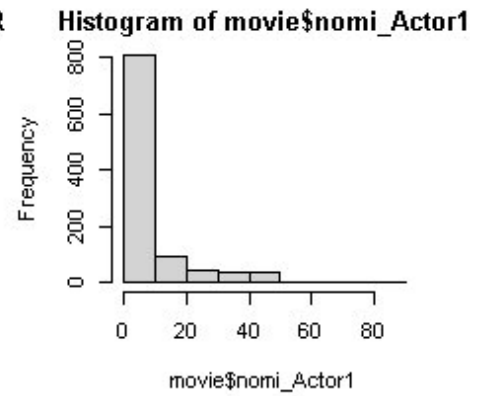
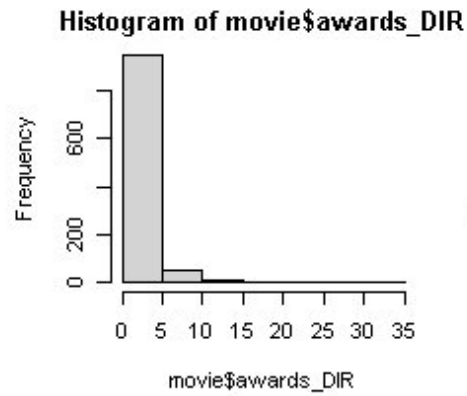
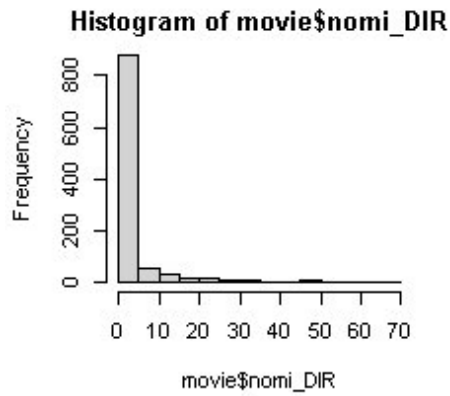
```
hist(movie$nomi_Actor1)
hist(movie$awards_Actor1)
```

Hide

```
hist(movie$nomi_Actor2)
hist(movie$awards_Actor2)
```

Hide

```
par(mfrow=c(1,1))
```



Hide

```
summary(lm(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005
+production_year_2006+production_year_2007+production_year_2008+production_year_2009+producti
on_year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2
014+production_year_2015+production_year_2016+production_year_2017+production_year_2018+produ
ction_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_a
verage_US+vote_count_US+vote_diff.US.KOR.+nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_
Actor2+awards_Actor2,data = movie))
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +  
  awards_Actor2, data = movie)
```

Residuals:

Min	1Q	Median	3Q	Max
-10.5345	-0.6914	0.5181	1.4037	12.2840

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-4.390e+00	2.893e+00	-1.517	0.129508
production_year_2003	-5.558e-01	2.723e+00	-0.204	0.838272
production_year_2004	-1.222e+00	2.454e+00	-0.498	0.618746
production_year_2005	-8.654e-01	2.453e+00	-0.353	0.724307
production_year_2006	-1.125e+00	2.453e+00	-0.459	0.646541
production_year_2007	-1.050e+00	2.454e+00	-0.428	0.668954
production_year_2008	-8.534e-01	2.451e+00	-0.348	0.727759
production_year_2009	-1.206e+00	2.453e+00	-0.492	0.623142
production_year_2010	-1.205e+00	2.452e+00	-0.491	0.623249
production_year_2011	-1.366e+00	2.451e+00	-0.557	0.577352
production_year_2012	-1.711e+00	2.455e+00	-0.697	0.486022
production_year_2013	-1.779e+00	2.451e+00	-0.726	0.468215
production_year_2014	-2.054e+00	2.451e+00	-0.838	0.402169
production_year_2015	-2.359e+00	2.450e+00	-0.963	0.335780
production_year_2016	-1.443e+00	2.468e+00	-0.585	0.558948
production_year_2017	-1.014e+01	3.447e+00	-2.941	0.003348 **
production_year_2018	2.442e-01	2.596e+00	0.094	0.925080
production_year_2019	3.152e-01	2.841e+00	0.111	0.911666
release_diff.US.KOR.	-6.540e-02	4.918e-03	-13.297	< 2e-16 ***
log(budget)	5.375e-01	6.917e-02	7.771	1.93e-14 ***
popularity_US	3.816e-03	1.871e-03	2.039	0.041715 *
log(revenue_US)	2.739e-01	6.657e-02	4.115	4.20e-05 ***
runtime	-1.364e-03	4.982e-03	-0.274	0.784257
vote_average_US	4.334e-01	1.221e-01	3.549	0.000405 ***
vote_count_US	2.293e-04	7.089e-05	3.235	0.001258 **
vote_diff.US.KOR.	-2.051e-03	3.394e-03	-0.604	0.545769
nomi_DIR	-1.146e-02	2.304e-02	-0.497	0.619018
awards_DIR	-9.606e-03	4.509e-02	-0.213	0.831359
nomi_Actor1	9.264e-05	1.333e-02	0.007	0.994457
awards_Actor1	3.370e-03	2.877e-02	0.117	0.906782
nomi_Actor2	-8.765e-03	1.689e-02	-0.519	0.603965
awards_Actor2	4.397e-03	3.784e-02	0.116	0.907502

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.434 on 997 degrees of freedom

Multiple R-squared: 0.4271, Adjusted R-squared: 0.4092

F-statistic: 23.97 on 31 and 997 DF, p-value: < 2.2e-16

Hide

```
a <- lm(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005+production_year_2006+production_year_2007+production_year_2008+production_year_2009+production_year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2014+production_year_2015+production_year_2016+production_year_2017+production_year_2018+production_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_average_US+vote_count_US+vote_diff.US.KOR.+nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_Actor2+awards_Actor2,data = movie)
par(mfrow=c(2,2))
plot(a,1)
plot(a,2)
```

not plotting observations with leverage one:
892, 971

Hide

```
plot(a,3)
```

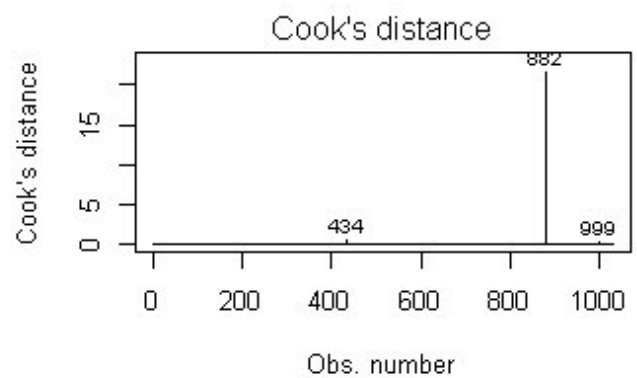
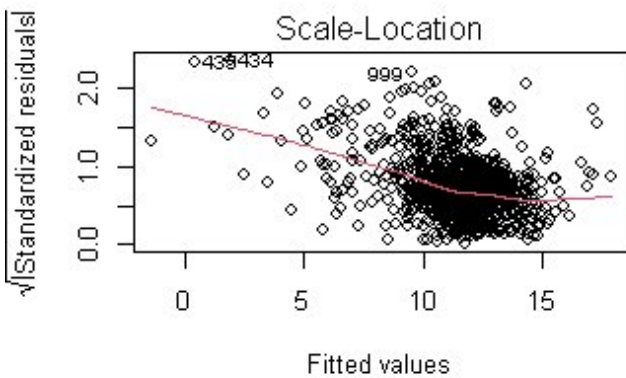
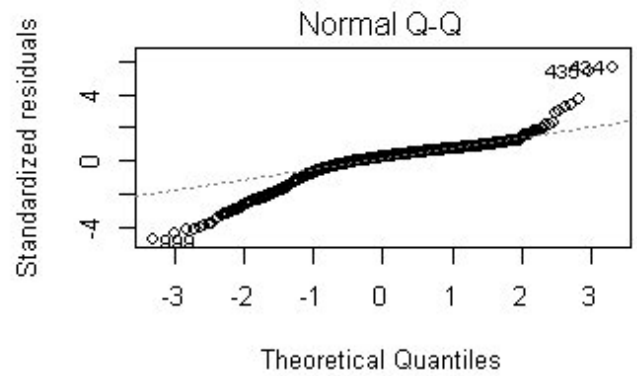
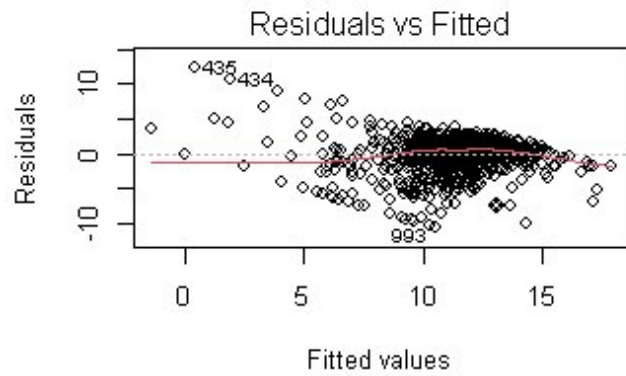
not plotting observations with leverage one:
892, 971

Hide

```
plot(a,4)
```

Hide

```
par(mfrow=c(1,1))
```

Hide

```
library(car)
outlierTest(a)
```

	rstudent	unadjusted p-value	Bonferroni p
434	5.683856	1.7289e-08	1.7756e-05
435	5.514242	4.4632e-08	4.5838e-05
999	-4.887743	1.1882e-06	1.2203e-03
993	-4.401161	1.1929e-05	1.2251e-02
982	-4.273127	2.1124e-05	2.1694e-02
371	-4.219596	2.6707e-05	2.7428e-02

Hide

```
movie.2 <- movie[-c(434,435,999,993,982,371),]
summary(lm(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005
+production_year_2006+production_year_2007+production_year_2008+production_year_2009+producti
on_year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2
014+production_year_2015+production_year_2016+production_year_2017+production_year_2018+produ
ction_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_a
verage_US+vote_count_US+vote_diff.US.KOR.+nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_
Actor2+awards_Actor2,data = movie.2))
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +  
  awards_Actor2, data = movie.2)
```

Residuals:

Min	1Q	Median	3Q	Max
-9.5893	-0.6712	0.4663	1.2937	10.1125

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.346e+00	2.715e+00	-1.232	0.218131
production_year_2003	-5.006e-01	2.551e+00	-0.196	0.844455
production_year_2004	-1.274e+00	2.300e+00	-0.554	0.579606
production_year_2005	-7.139e-01	2.298e+00	-0.311	0.756164
production_year_2006	-9.517e-01	2.298e+00	-0.414	0.678903
production_year_2007	-8.914e-01	2.300e+00	-0.388	0.698373
production_year_2008	-6.642e-01	2.297e+00	-0.289	0.772486
production_year_2009	-9.697e-01	2.298e+00	-0.422	0.673142
production_year_2010	-1.040e+00	2.298e+00	-0.453	0.650741
production_year_2011	-1.287e+00	2.296e+00	-0.560	0.575291
production_year_2012	-1.621e+00	2.300e+00	-0.705	0.481116
production_year_2013	-1.558e+00	2.297e+00	-0.678	0.497816
production_year_2014	-1.923e+00	2.297e+00	-0.837	0.402749
production_year_2015	-2.119e+00	2.296e+00	-0.923	0.356119
production_year_2016	-1.328e+00	2.313e+00	-0.574	0.565869
production_year_2017	-1.003e+01	3.230e+00	-3.105	0.001956 **
production_year_2018	1.044e+00	2.434e+00	0.429	0.668058
production_year_2019	-6.675e-01	3.234e+00	-0.206	0.836503
release_diff.US.KOR.	-8.342e-02	5.112e-03	-16.318	< 2e-16 ***
log(budget)	5.332e-01	6.490e-02	8.215	6.56e-16 ***
popularity_US	3.118e-03	1.756e-03	1.776	0.076089 .
log(revenue_US)	2.495e-01	6.258e-02	3.987	7.19e-05 ***
runtime	-2.760e-03	4.678e-03	-0.590	0.555286
vote_average_US	3.745e-01	1.148e-01	3.263	0.001140 **
vote_count_US	2.430e-04	6.664e-05	3.647	0.000279 ***
vote_diff.US.KOR.	-1.851e-03	3.180e-03	-0.582	0.560573
nomi_DIR	-7.604e-03	2.160e-02	-0.352	0.724830
awards_DIR	-2.507e-02	4.230e-02	-0.593	0.553533
nomi_Actor1	-7.166e-04	1.249e-02	-0.057	0.954271
awards_Actor1	5.635e-03	2.696e-02	0.209	0.834492
nomi_Actor2	-6.560e-03	1.584e-02	-0.414	0.678855
awards_Actor2	-7.429e-03	3.551e-02	-0.209	0.834322

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.281 on 991 degrees of freedom

Multiple R-squared: 0.4774, Adjusted R-squared: 0.4611

F-statistic: 29.21 on 31 and 991 DF, p-value: < 2.2e-16

Hide

```
b <- lm(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005+production_year_2006+production_year_2007+production_year_2008+production_year_2009+production_year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2014+production_year_2015+production_year_2016+production_year_2017+production_year_2018+production_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_average_US+vote_count_US+vote_diff.US.KOR.++nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_Actor2+awards_Actor2,data = movie.2)
vif(b)
```

production_year_2003	production_year_2004	production_year_2005	production_year_2006
4.983830	68.051146	71.436222	73.154797
production_year_2007	production_year_2008	production_year_2009	production_year_2010
69.776107	80.698948	73.144699	76.527268
production_year_2011	production_year_2012	production_year_2013	production_year_2014
78.156117	65.431372	79.869425	84.068980
production_year_2015	production_year_2016	production_year_2017	production_year_2018
85.630373	37.618236	2.003498	9.039035
production_year_2019	release_diff.US.KOR.	log(budget)	popularity_US
2.008066	1.200650	1.440593	2.014181
log(revenue_US)	runtime	vote_average_US	vote_count_US
1.686639	1.512951	1.559324	3.069315
vote_diff.US.KOR.	nomi_DIR	awards_DIR	nomi_Actor1
1.014426	6.666450	6.158828	4.350778
awards_Actor1	nomi_Actor2	awards_Actor2	
3.568505	3.776937	3.223682	

Hide

```
library(tidyverse)
library(caret)
train_control <- trainControl(method = "cv",
                              number = 10)
model <- train(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005+production_year_2006+production_year_2007+production_year_2008+production_year_2009+production_year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2014+production_year_2015+production_year_2016+production_year_2017+production_year_2018+production_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_average_US+vote_count_US+vote_diff.US.KOR.+nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_Actor2+awards_Actor2,data = movie.2,
               method = "lm",trControl = train_control)
```

non-uniform 'Rounding' sampler usedprediction from a rank-deficient fit may be misleadingprediction from a rank-deficient fit may be misleading

Hide

```
print(model)
```

Linear Regression

1023 samples
31 predictor

No pre-processing

Resampling: Cross-Validated (10 fold)

Summary of sample sizes: 921, 922, 921, 921, 920, 920, ...

Resampling results:

RMSE	Rsquared	MAE
2.555372	0.4118231	1.679013

Tuning parameter 'intercept' was held constant at a value of TRUE

Hide

```
library(forecast)
n = nrow(movie.2)
i = 1:n
train_list <- sample(i,n*0.6)
valid_list <- setdiff(i,train_list)
train_df <- movie.2[train_list,]
valid_df <- movie.2[valid_list,]
```

Hide

```
d <- lm(log(audience_KOR) ~ production_year_2003+production_year_2004+production_year_2005+pr
oduction_year_2006+production_year_2007+production_year_2008+production_year_2009+production_
year_2010+production_year_2011+production_year_2012+production_year_2013+production_year_2014
+production_year_2015+production_year_2016+production_year_2017+production_year_2018+producti
on_year_2019+release_diff.US.KOR.+log(budget)+popularity_US+log(revenue_US)+runtime+vote_aver
age_US+vote_count_US+vote_diff.US.KOR.+nomi_DIR+awards_DIR+nomi_Actor1+awards_Actor1+nomi_Act
or2+awards_Actor2,data = movie.2)
```

Hide

```
movie.lm.step1 <- step(d,direction = "forward")
```

Start: AIC=1718.59

```
log(audience_KOR) ~ production_year_2003 + production_year_2004 +
  production_year_2005 + production_year_2006 + production_year_2007 +
  production_year_2008 + production_year_2009 + production_year_2010 +
  production_year_2011 + production_year_2012 + production_year_2013 +
  production_year_2014 + production_year_2015 + production_year_2016 +
  production_year_2017 + production_year_2018 + production_year_2019 +
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +
  awards_Actor2
```

Hide

```
summary(movie.lm.step1)
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +  
  awards_Actor2, data = movie.2)
```

Residuals:

Min	1Q	Median	3Q	Max
-9.5893	-0.6712	0.4663	1.2937	10.1125

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.346e+00	2.715e+00	-1.232	0.218131
production_year_2003	-5.006e-01	2.551e+00	-0.196	0.844455
production_year_2004	-1.274e+00	2.300e+00	-0.554	0.579606
production_year_2005	-7.139e-01	2.298e+00	-0.311	0.756164
production_year_2006	-9.517e-01	2.298e+00	-0.414	0.678903
production_year_2007	-8.914e-01	2.300e+00	-0.388	0.698373
production_year_2008	-6.642e-01	2.297e+00	-0.289	0.772486
production_year_2009	-9.697e-01	2.298e+00	-0.422	0.673142
production_year_2010	-1.040e+00	2.298e+00	-0.453	0.650741
production_year_2011	-1.287e+00	2.296e+00	-0.560	0.575291
production_year_2012	-1.621e+00	2.300e+00	-0.705	0.481116
production_year_2013	-1.558e+00	2.297e+00	-0.678	0.497816
production_year_2014	-1.923e+00	2.297e+00	-0.837	0.402749
production_year_2015	-2.119e+00	2.296e+00	-0.923	0.356119
production_year_2016	-1.328e+00	2.313e+00	-0.574	0.565869
production_year_2017	-1.003e+01	3.230e+00	-3.105	0.001956 **
production_year_2018	1.044e+00	2.434e+00	0.429	0.668058
production_year_2019	-6.675e-01	3.234e+00	-0.206	0.836503
release_diff.US.KOR.	-8.342e-02	5.112e-03	-16.318	< 2e-16 ***
log(budget)	5.332e-01	6.490e-02	8.215	6.56e-16 ***
popularity_US	3.118e-03	1.756e-03	1.776	0.076089 .
log(revenue_US)	2.495e-01	6.258e-02	3.987	7.19e-05 ***
runtime	-2.760e-03	4.678e-03	-0.590	0.555286
vote_average_US	3.745e-01	1.148e-01	3.263	0.001140 **
vote_count_US	2.430e-04	6.664e-05	3.647	0.000279 ***
vote_diff.US.KOR.	-1.851e-03	3.180e-03	-0.582	0.560573
nomi_DIR	-7.604e-03	2.160e-02	-0.352	0.724830
awards_DIR	-2.507e-02	4.230e-02	-0.593	0.553533
nomi_Actor1	-7.166e-04	1.249e-02	-0.057	0.954271
awards_Actor1	5.635e-03	2.696e-02	0.209	0.834492
nomi_Actor2	-6.560e-03	1.584e-02	-0.414	0.678855
awards_Actor2	-7.429e-03	3.551e-02	-0.209	0.834322

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.281 on 991 degrees of freedom

Multiple R-squared: 0.4774, Adjusted R-squared: 0.4611

F-statistic: 29.21 on 31 and 991 DF, p-value: $< 2.2e-16$

Hide

```
movie.lm.step_pred1 <- predict(movie.lm.step1,valid_df)
movie.lm.step2 <- step(d,direction = "backward")
```

Start: AIC=1718.59

```
log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +  
  awards_Actor2
```

	Df	Sum of Sq	RSS	AIC
- nomi_Actor1	1	0.02	5155.9	1716.6
- production_year_2003	1	0.20	5156.0	1716.6
- production_year_2019	1	0.22	5156.1	1716.6
- awards_Actor1	1	0.23	5156.1	1716.6
- awards_Actor2	1	0.23	5156.1	1716.6
- production_year_2008	1	0.44	5156.3	1716.7
- production_year_2005	1	0.50	5156.3	1716.7
- nomi_DIR	1	0.65	5156.5	1716.7
- production_year_2007	1	0.78	5156.6	1716.8
- production_year_2006	1	0.89	5156.7	1716.8
- nomi_Actor2	1	0.89	5156.7	1716.8
- production_year_2009	1	0.93	5156.8	1716.8
- production_year_2018	1	0.96	5156.8	1716.8
- production_year_2010	1	1.07	5156.9	1716.8
- production_year_2004	1	1.60	5157.4	1716.9
- production_year_2011	1	1.63	5157.5	1716.9
- production_year_2016	1	1.72	5157.6	1716.9
- vote_diff.US.KOR.	1	1.76	5157.6	1716.9
- runtime	1	1.81	5157.7	1717.0
- awards_DIR	1	1.83	5157.7	1717.0
- production_year_2013	1	2.39	5158.2	1717.1
- production_year_2012	1	2.58	5158.4	1717.1
- production_year_2014	1	3.65	5159.5	1717.3
- production_year_2015	1	4.43	5160.3	1717.5
<none>			5155.8	1718.6
- popularity_US	1	16.40	5172.2	1719.8
- production_year_2017	1	50.16	5206.0	1726.5
- vote_average_US	1	55.39	5211.2	1727.5
- vote_count_US	1	69.21	5225.1	1730.2
- log(revenue_US)	1	82.69	5238.5	1732.9
- log(budget)	1	351.15	5507.0	1784.0
- release_diff.US.KOR.	1	1385.27	6541.1	1960.0

Step: AIC=1716.59

```
log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + awards_Actor1 + nomi_Actor2 + awards_Actor2
```


	Df	Sum of Sq	RSS	AIC
- production_year_2003	1	0.20	5156.1	1714.6
- awards_Actor2	1	0.21	5156.1	1714.6
- production_year_2019	1	0.22	5156.1	1714.6
- awards_Actor1	1	0.34	5156.2	1714.7
- production_year_2008	1	0.44	5156.3	1714.7
- production_year_2005	1	0.50	5156.4	1714.7
- nomi_DIR	1	0.70	5156.6	1714.7
- production_year_2007	1	0.78	5156.6	1714.8
- production_year_2006	1	0.90	5156.8	1714.8
- production_year_2009	1	0.93	5156.8	1714.8
- production_year_2018	1	0.96	5156.8	1714.8
- production_year_2010	1	1.07	5156.9	1714.8
- nomi_Actor2	1	1.18	5157.0	1714.8
- production_year_2004	1	1.60	5157.5	1714.9
- production_year_2011	1	1.64	5157.5	1714.9
- production_year_2016	1	1.72	5157.6	1714.9
- vote_diff.US.KOR.	1	1.77	5157.6	1714.9
- runtime	1	1.80	5157.7	1715.0
- awards_DIR	1	1.81	5157.7	1715.0
- production_year_2013	1	2.39	5158.3	1715.1
- production_year_2012	1	2.58	5158.4	1715.1
- production_year_2014	1	3.65	5159.5	1715.3
- production_year_2015	1	4.44	5160.3	1715.5
<none>			5155.9	1716.6
- popularity_US	1	16.55	5172.4	1717.9
- production_year_2017	1	50.17	5206.0	1724.5
- vote_average_US	1	55.40	5211.3	1725.5
- vote_count_US	1	69.37	5225.2	1728.3
- log(revenue_US)	1	82.73	5238.6	1730.9
- log(budget)	1	351.42	5507.3	1782.0
- release_diff.US.KOR.	1	1387.09	6542.9	1958.3

Step: AIC=1714.63

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + production_year_2019 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + vote_diff.US.KOR. + nomi_DIR +
 awards_DIR + awards_Actor1 + nomi_Actor2 + awards_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2019	1	0.06	5156.1	1712.7
- awards_Actor2	1	0.21	5156.3	1712.7
- production_year_2008	1	0.33	5156.4	1712.7
- awards_Actor1	1	0.33	5156.4	1712.7
- production_year_2005	1	0.46	5156.5	1712.7
- nomi_DIR	1	0.70	5156.8	1712.8
- production_year_2007	1	1.12	5157.2	1712.9
- nomi_Actor2	1	1.19	5157.3	1712.9
- production_year_2006	1	1.42	5157.5	1712.9
- production_year_2009	1	1.50	5157.6	1712.9

- vote_diff.US.KOR.	1	1.77	5157.8	1713.0
- runtime	1	1.79	5157.8	1713.0
- awards_DIR	1	1.81	5157.9	1713.0
- production_year_2010	1	1.91	5158.0	1713.0
- production_year_2004	1	3.53	5159.6	1713.3
- production_year_2011	1	3.68	5159.7	1713.4
- production_year_2016	1	3.78	5159.8	1713.4
- production_year_2018	1	6.17	5162.2	1713.9
- production_year_2013	1	6.27	5162.3	1713.9
- production_year_2012	1	6.86	5162.9	1714.0
<none>			5156.1	1714.6
- production_year_2014	1	10.85	5166.9	1714.8
- production_year_2015	1	13.91	5170.0	1715.4
- popularity_US	1	16.54	5172.6	1715.9
- vote_average_US	1	55.38	5211.4	1723.6
- vote_count_US	1	69.32	5225.4	1726.3
- production_year_2017	1	76.96	5233.0	1727.8
- log(revenue_US)	1	82.68	5238.7	1728.9
- log(budget)	1	351.54	5507.6	1780.1
- release_diff.US.KOR.	1	1387.39	6543.4	1956.4

Step: AIC=1712.65

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
 awards_Actor1 + nomi_Actor2 + awards_Actor2

	Df	Sum of Sq	RSS	AIC
- awards_Actor2	1	0.22	5156.3	1710.7
- production_year_2008	1	0.27	5156.4	1710.7
- awards_Actor1	1	0.34	5156.5	1710.7
- production_year_2005	1	0.40	5156.5	1710.7
- nomi_DIR	1	0.70	5156.8	1710.8
- production_year_2007	1	1.09	5157.2	1710.9
- nomi_Actor2	1	1.17	5157.3	1710.9
- production_year_2006	1	1.42	5157.5	1710.9
- production_year_2009	1	1.52	5157.6	1711.0
- vote_diff.US.KOR.	1	1.76	5157.9	1711.0
- runtime	1	1.80	5157.9	1711.0
- awards_DIR	1	1.83	5157.9	1711.0
- production_year_2010	1	1.96	5158.1	1711.0
- production_year_2004	1	3.77	5159.9	1711.4
- production_year_2011	1	3.95	5160.1	1711.4
- production_year_2016	1	4.02	5160.1	1711.4
- production_year_2013	1	6.89	5163.0	1712.0
- production_year_2018	1	7.30	5163.4	1712.1
- production_year_2012	1	7.55	5163.7	1712.1
<none>			5156.1	1712.7
- production_year_2014	1	12.17	5168.3	1713.1
- production_year_2015	1	15.70	5171.8	1713.8
- popularity_US	1	16.55	5172.7	1713.9

- vote_average_US	1	55.34	5211.5	1721.6
- vote_count_US	1	69.33	5225.5	1724.3
- production_year_2017	1	78.37	5234.5	1726.1
- log(revenue_US)	1	82.77	5238.9	1726.9
- log(budget)	1	351.62	5507.7	1778.1
- release_diff.US.KOR.	1	1387.38	6543.5	1954.4

Step: AIC=1710.69

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
 awards_Actor1 + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- awards_Actor1	1	0.26	5156.6	1708.7
- production_year_2008	1	0.27	5156.6	1708.7
- production_year_2005	1	0.40	5156.7	1708.8
- nomi_DIR	1	0.68	5157.0	1708.8
- production_year_2007	1	1.09	5157.4	1708.9
- production_year_2006	1	1.40	5157.7	1709.0
- production_year_2009	1	1.52	5157.9	1709.0
- vote_diff.US.KOR.	1	1.76	5158.1	1709.0
- runtime	1	1.82	5158.2	1709.0
- awards_DIR	1	1.82	5158.2	1709.0
- production_year_2010	1	1.95	5158.3	1709.1
- production_year_2004	1	3.75	5160.1	1709.4
- production_year_2011	1	3.93	5160.3	1709.5
- production_year_2016	1	4.01	5160.4	1709.5
- nomi_Actor2	1	4.54	5160.9	1709.6
- production_year_2013	1	6.82	5163.2	1710.0
- production_year_2018	1	7.35	5163.7	1710.2
- production_year_2012	1	7.52	5163.9	1710.2
<none>			5156.3	1710.7
- production_year_2014	1	12.14	5168.5	1711.1
- production_year_2015	1	15.65	5172.0	1711.8
- popularity_US	1	16.46	5172.8	1712.0
- vote_average_US	1	55.53	5211.9	1719.7
- vote_count_US	1	69.59	5225.9	1722.4
- production_year_2017	1	78.32	5234.7	1724.1
- log(revenue_US)	1	82.62	5239.0	1725.0
- log(budget)	1	352.06	5508.4	1776.3
- release_diff.US.KOR.	1	1393.18	6549.5	1953.4

Step: AIC=1708.74

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +

vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2008	1	0.26	5156.9	1706.8
- production_year_2005	1	0.39	5157.0	1706.8
- nomi_DIR	1	0.62	5157.2	1706.9
- production_year_2007	1	1.08	5157.7	1707.0
- production_year_2006	1	1.40	5158.0	1707.0
- production_year_2009	1	1.52	5158.1	1707.0
- vote_diff.US.KOR.	1	1.76	5158.4	1707.1
- awards_DIR	1	1.78	5158.4	1707.1
- runtime	1	1.81	5158.4	1707.1
- production_year_2010	1	1.92	5158.5	1707.1
- production_year_2004	1	3.70	5160.3	1707.5
- production_year_2011	1	3.94	5160.5	1707.5
- production_year_2016	1	4.03	5160.6	1707.5
- nomi_Actor2	1	4.44	5161.0	1707.6
- production_year_2013	1	6.85	5163.4	1708.1
- production_year_2018	1	7.37	5164.0	1708.2
- production_year_2012	1	7.53	5164.1	1708.2
<none>			5156.6	1708.7
- production_year_2014	1	12.20	5168.8	1709.2
- production_year_2015	1	15.64	5172.2	1709.8
- popularity_US	1	16.52	5173.1	1710.0
- vote_average_US	1	55.45	5212.0	1717.7
- vote_count_US	1	69.72	5226.3	1720.5
- production_year_2017	1	78.37	5235.0	1722.2
- log(revenue_US)	1	83.21	5239.8	1723.1
- log(budget)	1	351.84	5508.4	1774.3
- release_diff.US.KOR.	1	1398.35	6554.9	1952.2

Step: AIC=1706.79

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
production_year_2006 + production_year_2007 + production_year_2009 +
production_year_2010 + production_year_2011 + production_year_2012 +
production_year_2013 + production_year_2014 + production_year_2015 +
production_year_2016 + production_year_2017 + production_year_2018 +
release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +
nomi_DIR + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2005	1	0.18	5157.0	1704.8
- nomi_DIR	1	0.64	5157.5	1704.9
- awards_DIR	1	1.74	5158.6	1705.1
- vote_diff.US.KOR.	1	1.75	5158.6	1705.1
- runtime	1	1.83	5158.7	1705.2
- production_year_2007	1	2.35	5159.2	1705.3
- production_year_2006	1	3.82	5160.7	1705.5
- production_year_2009	1	4.37	5161.2	1705.7
- nomi_Actor2	1	4.49	5161.3	1705.7
- production_year_2010	1	6.45	5163.3	1706.1
<none>			5156.9	1706.8
- production_year_2016	1	12.32	5169.2	1707.2
- production_year_2004	1	15.27	5172.1	1707.8

- popularity_US	1	16.53	5173.4	1708.1
- production_year_2011	1	17.93	5174.8	1708.3
- production_year_2018	1	20.07	5176.9	1708.8
- production_year_2013	1	36.09	5192.9	1711.9
- production_year_2012	1	36.19	5193.0	1712.0
- vote_average_US	1	55.25	5212.1	1715.7
- vote_count_US	1	69.94	5226.8	1718.6
- production_year_2014	1	71.31	5228.2	1718.8
- log(revenue_US)	1	83.32	5240.2	1721.2
- production_year_2017	1	86.34	5243.2	1721.8
- production_year_2015	1	96.53	5253.4	1723.8
- log(budget)	1	351.99	5508.8	1772.3
- release_diff.US.KOR.	1	1400.66	6557.5	1950.6

Step: AIC=1704.83

```
log(audience_KOR) ~ production_year_2004 + production_year_2006 +
  production_year_2007 + production_year_2009 + production_year_2010 +
  production_year_2011 + production_year_2012 + production_year_2013 +
  production_year_2014 + production_year_2015 + production_year_2016 +
  production_year_2017 + production_year_2018 + release_diff.US.KOR. +
  log(budget) + popularity_US + log(revenue_US) + runtime +
  vote_average_US + vote_count_US + vote_diff.US.KOR. + nomi_DIR +
  awards_DIR + nomi_Actor2
```

	Df	Sum of Sq	RSS	AIC
- nomi_DIR	1	0.62	5157.7	1703.0
- vote_diff.US.KOR.	1	1.75	5158.8	1703.2
- awards_DIR	1	1.76	5158.8	1703.2
- runtime	1	1.99	5159.0	1703.2
- production_year_2007	1	2.26	5159.3	1703.3
- production_year_2006	1	3.91	5160.9	1703.6
- nomi_Actor2	1	4.47	5161.5	1703.7
- production_year_2009	1	4.55	5161.6	1703.7
- production_year_2010	1	6.94	5164.0	1704.2
<none>			5157.0	1704.8
- production_year_2016	1	12.94	5170.0	1705.4
- popularity_US	1	16.54	5173.6	1706.1
- production_year_2004	1	17.32	5174.4	1706.3
- production_year_2011	1	20.68	5177.7	1706.9
- production_year_2018	1	21.52	5178.6	1707.1
- production_year_2012	1	41.86	5198.9	1711.1
- production_year_2013	1	42.70	5199.7	1711.3
- vote_average_US	1	55.25	5212.3	1713.7
- vote_count_US	1	70.59	5227.6	1716.7
- log(revenue_US)	1	83.24	5240.3	1719.2
- production_year_2014	1	86.13	5243.2	1719.8
- production_year_2017	1	86.20	5243.2	1719.8
- production_year_2015	1	117.91	5274.9	1726.0
- log(budget)	1	351.86	5508.9	1770.3
- release_diff.US.KOR.	1	1400.53	6557.6	1948.6

Step: AIC=1702.95

```
log(audience_KOR) ~ production_year_2004 + production_year_2006 +
  production_year_2007 + production_year_2009 + production_year_2010 +
  production_year_2011 + production_year_2012 + production_year_2013 +
  production_year_2014 + production_year_2015 + production_year_2016 +
```

production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + vote_diff.US.KOR. + awards_DIR +
 nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- vote_diff.US.KOR.	1	1.77	5159.4	1701.3
- production_year_2007	1	2.23	5159.9	1701.4
- runtime	1	2.30	5160.0	1701.4
- production_year_2006	1	3.78	5161.4	1701.7
- production_year_2009	1	4.40	5162.1	1701.8
- nomi_Actor2	1	5.56	5163.2	1702.0
- production_year_2010	1	6.75	5164.4	1702.3
<none>			5157.7	1703.0
- production_year_2016	1	12.79	5170.4	1703.5
- popularity_US	1	16.47	5174.1	1704.2
- production_year_2004	1	16.97	5174.6	1704.3
- production_year_2011	1	20.61	5178.3	1705.0
- awards_DIR	1	21.40	5179.1	1705.2
- production_year_2018	1	21.77	5179.4	1705.3
- production_year_2012	1	41.62	5199.3	1709.2
- production_year_2013	1	42.77	5200.4	1709.4
- vote_average_US	1	55.77	5213.4	1712.0
- vote_count_US	1	70.57	5228.2	1714.8
- log(revenue_US)	1	83.18	5240.8	1717.3
- production_year_2014	1	85.84	5243.5	1717.8
- production_year_2017	1	86.09	5243.7	1717.9
- production_year_2015	1	118.84	5276.5	1724.3
- log(budget)	1	353.67	5511.3	1768.8
- release_diff.US.KOR.	1	1414.52	6572.2	1948.9

Step: AIC=1701.3

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2007 + production_year_2009 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2007	1	1.88	5161.3	1699.7
- runtime	1	2.20	5161.6	1699.7
- production_year_2006	1	3.77	5163.2	1700.0
- production_year_2009	1	4.41	5163.8	1700.2
- nomi_Actor2	1	5.61	5165.0	1700.4
- production_year_2010	1	6.75	5166.2	1700.6
<none>			5159.4	1701.3
- production_year_2016	1	12.75	5172.2	1701.8
- popularity_US	1	16.47	5175.9	1702.6
- production_year_2004	1	16.96	5176.4	1702.7
- production_year_2011	1	20.61	5180.0	1703.4
- awards_DIR	1	21.51	5180.9	1703.6
- production_year_2018	1	21.83	5181.3	1703.6
- production_year_2012	1	41.67	5201.1	1707.5
- production_year_2013	1	42.81	5202.2	1707.8

- vote_average_US	1	55.77	5215.2	1710.3
- vote_count_US	1	70.59	5230.0	1713.2
- log(revenue_US)	1	83.10	5242.5	1715.7
- production_year_2014	1	85.88	5245.3	1716.2
- production_year_2017	1	86.15	5245.6	1716.2
- production_year_2015	1	118.90	5278.3	1722.6
- log(budget)	1	352.67	5512.1	1766.9
- release_diff.US.KOR.	1	1416.88	6576.3	1947.5

Step: AIC=1699.67

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- runtime	1	2.14	5163.4	1698.1
- production_year_2006	1	2.53	5163.8	1698.2
- production_year_2009	1	3.11	5164.4	1698.3
- production_year_2010	1	5.20	5166.5	1698.7
- nomi_Actor2	1	5.32	5166.6	1698.7
<none>			5161.3	1699.7
- production_year_2016	1	11.22	5172.5	1699.9
- production_year_2004	1	15.09	5176.4	1700.7
- popularity_US	1	16.41	5177.7	1700.9
- production_year_2011	1	18.73	5180.0	1701.4
- awards_DIR	1	21.11	5182.4	1701.8
- production_year_2018	1	23.72	5185.0	1702.4
- production_year_2012	1	40.01	5201.3	1705.6
- production_year_2013	1	41.43	5202.7	1705.8
- vote_average_US	1	54.12	5215.4	1708.3
- vote_count_US	1	72.10	5233.4	1711.9
- log(revenue_US)	1	82.71	5244.0	1713.9
- production_year_2017	1	85.37	5246.7	1714.5
- production_year_2014	1	86.63	5247.9	1714.7
- production_year_2015	1	122.02	5283.3	1721.6
- log(budget)	1	351.92	5513.2	1765.2
- release_diff.US.KOR.	1	1415.61	6576.9	1945.6

Step: AIC=1698.1

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
 awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2006	1	2.39	5165.8	1696.6
- production_year_2009	1	3.15	5166.6	1696.7
- production_year_2010	1	4.68	5168.1	1697.0
- nomi_Actor2	1	5.04	5168.5	1697.1

<none>			5163.4	1698.1
- production_year_2016	1	11.21	5174.7	1698.3
- production_year_2004	1	15.36	5178.8	1699.1
- popularity_US	1	16.90	5180.3	1699.4
- production_year_2011	1	18.05	5181.5	1699.7
- production_year_2018	1	23.61	5187.1	1700.8
- awards_DIR	1	24.05	5187.5	1700.8
- production_year_2012	1	40.30	5203.7	1704.0
- production_year_2013	1	40.73	5204.2	1704.1
- vote_average_US	1	52.11	5215.6	1706.4
- vote_count_US	1	70.02	5233.5	1709.9
- log(revenue_US)	1	84.39	5247.8	1712.7
- production_year_2017	1	85.49	5248.9	1712.9
- production_year_2014	1	86.11	5249.6	1713.0
- production_year_2015	1	121.33	5284.8	1719.9
- log(budget)	1	352.28	5515.7	1763.6
- release_diff.US.KOR.	1	1416.23	6579.7	1944.0

Step: AIC=1696.57

log(audience_KOR) ~ production_year_2004 + production_year_2009 +
 production_year_2010 + production_year_2011 + production_year_2012 +
 production_year_2013 + production_year_2014 + production_year_2015 +
 production_year_2016 + production_year_2017 + production_year_2018 +
 release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
 vote_average_US + vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2009	1	2.08	5167.9	1695.0
- production_year_2010	1	3.38	5169.2	1695.2
- nomi_Actor2	1	5.18	5171.0	1695.6
- production_year_2016	1	9.77	5175.6	1696.5
<none>			5165.8	1696.6
- production_year_2004	1	13.38	5179.2	1697.2
- production_year_2011	1	15.93	5181.8	1697.7
- popularity_US	1	16.99	5182.8	1697.9
- awards_DIR	1	24.13	5190.0	1699.3
- production_year_2018	1	25.18	5191.0	1699.5
- production_year_2012	1	37.94	5203.8	1702.1
- production_year_2013	1	38.35	5204.2	1702.1
- vote_average_US	1	51.70	5217.5	1704.8
- vote_count_US	1	70.46	5236.3	1708.4
- log(revenue_US)	1	83.40	5249.2	1711.0
- production_year_2014	1	84.31	5250.1	1711.1
- production_year_2017	1	84.68	5250.5	1711.2
- production_year_2015	1	120.47	5286.3	1718.2
- log(budget)	1	352.56	5518.4	1762.1
- release_diff.US.KOR.	1	1417.98	6583.8	1942.7

Step: AIC=1694.98

log(audience_KOR) ~ production_year_2004 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + vote_average_US +
 vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2010	1	2.51	5170.4	1693.5
- nomi_Actor2	1	5.49	5173.4	1694.1
- production_year_2016	1	8.69	5176.6	1694.7
<none>			5167.9	1695.0
- production_year_2004	1	11.86	5179.8	1695.3
- production_year_2011	1	14.25	5182.2	1695.8
- popularity_US	1	17.23	5185.1	1696.4
- awards_DIR	1	24.33	5192.2	1697.8
- production_year_2018	1	25.96	5193.9	1698.1
- production_year_2012	1	35.94	5203.8	1700.1
- production_year_2013	1	36.29	5204.2	1700.1
- vote_average_US	1	52.09	5220.0	1703.2
- vote_count_US	1	69.71	5237.6	1706.7
- production_year_2014	1	82.42	5250.3	1709.2
- log(revenue_US)	1	83.85	5251.8	1709.5
- production_year_2017	1	84.12	5252.0	1709.5
- production_year_2015	1	119.00	5286.9	1716.3
- log(budget)	1	352.38	5520.3	1760.5
- release_diff.US.KOR.	1	1417.10	6585.0	1940.9

Step: AIC=1693.48

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
 awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- nomi_Actor2	1	5.44	5175.9	1692.6
- production_year_2016	1	7.69	5178.1	1693.0
<none>			5170.4	1693.5
- production_year_2004	1	10.48	5180.9	1693.5
- production_year_2011	1	12.62	5183.0	1694.0
- popularity_US	1	17.55	5188.0	1695.0
- awards_DIR	1	23.90	5194.3	1696.2
- production_year_2018	1	27.00	5197.4	1696.8
- production_year_2012	1	33.78	5204.2	1698.1
- production_year_2013	1	34.04	5204.5	1698.2
- vote_average_US	1	52.34	5222.8	1701.8
- vote_count_US	1	68.30	5238.7	1704.9
- production_year_2014	1	79.92	5250.3	1707.2
- production_year_2017	1	83.57	5254.0	1707.9
- log(revenue_US)	1	84.49	5254.9	1708.1
- production_year_2015	1	116.60	5287.0	1714.3
- log(budget)	1	352.69	5523.1	1759.0
- release_diff.US.KOR.	1	1420.20	6590.6	1939.8

Step: AIC=1692.56

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
 awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2016	1	7.00	5182.9	1691.9
<none>			5175.9	1692.6
- production_year_2004	1	10.55	5186.4	1692.6
- production_year_2011	1	12.52	5188.4	1693.0
- popularity_US	1	17.21	5193.1	1694.0
- production_year_2018	1	27.77	5203.6	1696.0
- production_year_2012	1	32.35	5208.2	1696.9
- awards_DIR	1	32.98	5208.8	1697.0
- production_year_2013	1	34.43	5210.3	1697.3
- vote_average_US	1	51.73	5227.6	1700.7
- vote_count_US	1	67.93	5243.8	1703.9
- production_year_2014	1	77.12	5253.0	1705.7
- production_year_2017	1	83.13	5259.0	1706.9
- log(revenue_US)	1	83.99	5259.9	1707.0
- production_year_2015	1	112.86	5288.7	1712.6
- log(budget)	1	350.24	5526.1	1757.5
- release_diff.US.KOR.	1	1416.31	6592.2	1938.0

Step: AIC=1691.94

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2017 + production_year_2018 +
 release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
 vote_average_US + vote_count_US + awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2004	1	9.04	5191.9	1691.7
<none>			5182.9	1691.9
- production_year_2011	1	10.67	5193.5	1692.0
- popularity_US	1	16.21	5199.1	1693.1
- production_year_2018	1	28.57	5211.4	1695.6
- production_year_2012	1	29.41	5212.3	1695.7
- production_year_2013	1	31.34	5214.2	1696.1
- awards_DIR	1	33.16	5216.0	1696.5
- vote_average_US	1	58.51	5241.4	1701.4
- vote_count_US	1	65.36	5248.2	1702.8
- production_year_2014	1	72.33	5255.2	1704.1
- production_year_2017	1	82.09	5265.0	1706.0
- log(revenue_US)	1	83.18	5266.0	1706.2
- production_year_2015	1	107.49	5290.4	1710.9
- log(budget)	1	352.49	5535.3	1757.2
- release_diff.US.KOR.	1	1413.77	6596.6	1936.7

Step: AIC=1691.72

log(audience_KOR) ~ production_year_2011 + production_year_2012 +
 production_year_2013 + production_year_2014 + production_year_2015 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + vote_average_US +
 vote_count_US + awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2011	1	8.47	5200.4	1691.4
<none>			5191.9	1691.7
- popularity_US	1	15.98	5207.9	1692.9

```

- production_year_2012 1      26.35 5218.3 1694.9
- production_year_2013 1      27.85 5219.8 1695.2
- production_year_2018 1      29.93 5221.8 1695.6
- awards_DIR           1      34.32 5226.2 1696.5
- vote_average_US      1      58.34 5250.2 1701.2
- production_year_2014 1      67.07 5259.0 1702.8
- vote_count_US        1      68.92 5260.8 1703.2
- log(revenue_US)      1      81.10 5273.0 1705.6
- production_year_2017 1      81.35 5273.3 1705.6
- production_year_2015 1     101.39 5293.3 1709.5
- log(budget)          1     347.93 5539.8 1756.1
- release_diff.US.KOR. 1    1407.73 6599.6 1935.2

```

Step: AIC=1691.39

```

log(audience_KOR) ~ production_year_2012 + production_year_2013 +
  production_year_2014 + production_year_2015 + production_year_2017 +
  production_year_2018 + release_diff.US.KOR. + log(budget) +
  popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
  awards_DIR

```

	Df	Sum of Sq	RSS	AIC
<none>			5200.4	1691.4
- popularity_US	1	16.40	5216.8	1692.6
- production_year_2012	1	23.04	5223.4	1693.9
- production_year_2013	1	24.31	5224.7	1694.2
- production_year_2018	1	30.53	5230.9	1695.4
- awards_DIR	1	33.92	5234.3	1696.0
- vote_average_US	1	59.91	5260.3	1701.1
- production_year_2014	1	62.15	5262.5	1701.5
- vote_count_US	1	66.18	5266.6	1702.3
- production_year_2017	1	80.61	5281.0	1705.1
- log(revenue_US)	1	85.01	5285.4	1706.0
- production_year_2015	1	95.56	5295.9	1708.0
- log(budget)	1	349.59	5550.0	1756.0
- release_diff.US.KOR.	1	1399.36	6599.7	1933.2

Hide

```
summary(movie.lm.step2)
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2017 +  
  production_year_2018 + release_diff.US.KOR. + log(budget) +  
  popularity_US + log(revenue_US) + vote_average_US + vote_count_US +  
  awards_DIR, data = movie.2)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-9.7652	-0.6878	0.4658	1.2842	10.0785

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-4.443e+00	1.504e+00	-2.955	0.003203	**
production_year_2012	-6.193e-01	2.929e-01	-2.114	0.034743	*
production_year_2013	-5.727e-01	2.637e-01	-2.172	0.030101	*
production_year_2014	-9.143e-01	2.633e-01	-3.472	0.000538	***
production_year_2015	-1.113e+00	2.586e-01	-4.306	1.83e-05	***
production_year_2017	-9.031e+00	2.284e+00	-3.955	8.19e-05	***
production_year_2018	2.027e+00	8.327e-01	2.434	0.015109	*
release_diff.US.KOR.	-8.242e-02	5.002e-03	-16.478	< 2e-16	***
log(budget)	5.215e-01	6.332e-02	8.236	5.49e-16	***
popularity_US	3.100e-03	1.738e-03	1.784	0.074744	.
log(revenue_US)	2.511e-01	6.182e-02	4.061	5.26e-05	***
vote_average_US	3.696e-01	1.084e-01	3.409	0.000677	***
vote_count_US	2.293e-04	6.399e-05	3.583	0.000355	***
awards_DIR	-4.477e-02	1.745e-02	-2.565	0.010450	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.27 on 1009 degrees of freedom

Multiple R-squared: 0.4729, Adjusted R-squared: 0.4661

F-statistic: 69.64 on 13 and 1009 DF, p-value: < 2.2e-16

Hide

```
movie.lm.step_pred2 <- predict(movie.lm.step2,valid_df)  
movie.lm.step3 <- step(d,direction = "both")
```

Start: AIC=1718.59

```
log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + nomi_Actor1 + awards_Actor1 + nomi_Actor2 +  
  awards_Actor2
```

	Df	Sum of Sq	RSS	AIC
- nomi_Actor1	1	0.02	5155.9	1716.6
- production_year_2003	1	0.20	5156.0	1716.6
- production_year_2019	1	0.22	5156.1	1716.6
- awards_Actor1	1	0.23	5156.1	1716.6
- awards_Actor2	1	0.23	5156.1	1716.6
- production_year_2008	1	0.44	5156.3	1716.7
- production_year_2005	1	0.50	5156.3	1716.7
- nomi_DIR	1	0.65	5156.5	1716.7
- production_year_2007	1	0.78	5156.6	1716.8
- production_year_2006	1	0.89	5156.7	1716.8
- nomi_Actor2	1	0.89	5156.7	1716.8
- production_year_2009	1	0.93	5156.8	1716.8
- production_year_2018	1	0.96	5156.8	1716.8
- production_year_2010	1	1.07	5156.9	1716.8
- production_year_2004	1	1.60	5157.4	1716.9
- production_year_2011	1	1.63	5157.5	1716.9
- production_year_2016	1	1.72	5157.6	1716.9
- vote_diff.US.KOR.	1	1.76	5157.6	1716.9
- runtime	1	1.81	5157.7	1717.0
- awards_DIR	1	1.83	5157.7	1717.0
- production_year_2013	1	2.39	5158.2	1717.1
- production_year_2012	1	2.58	5158.4	1717.1
- production_year_2014	1	3.65	5159.5	1717.3
- production_year_2015	1	4.43	5160.3	1717.5
<none>			5155.8	1718.6
- popularity_US	1	16.40	5172.2	1719.8
- production_year_2017	1	50.16	5206.0	1726.5
- vote_average_US	1	55.39	5211.2	1727.5
- vote_count_US	1	69.21	5225.1	1730.2
- log(revenue_US)	1	82.69	5238.5	1732.9
- log(budget)	1	351.15	5507.0	1784.0
- release_diff.US.KOR.	1	1385.27	6541.1	1960.0

Step: AIC=1716.59

```
log(audience_KOR) ~ production_year_2003 + production_year_2004 +  
  production_year_2005 + production_year_2006 + production_year_2007 +  
  production_year_2008 + production_year_2009 + production_year_2010 +  
  production_year_2011 + production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2016 +  
  production_year_2017 + production_year_2018 + production_year_2019 +  
  release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +  
  runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +  
  nomi_DIR + awards_DIR + awards_Actor1 + nomi_Actor2 + awards_Actor2
```

	Df	Sum of Sq	RSS	AIC
- production_year_2003	1	0.20	5156.1	1714.6
- awards_Actor2	1	0.21	5156.1	1714.6
- production_year_2019	1	0.22	5156.1	1714.6
- awards_Actor1	1	0.34	5156.2	1714.7
- production_year_2008	1	0.44	5156.3	1714.7
- production_year_2005	1	0.50	5156.4	1714.7
- nomi_DIR	1	0.70	5156.6	1714.7
- production_year_2007	1	0.78	5156.6	1714.8
- production_year_2006	1	0.90	5156.8	1714.8
- production_year_2009	1	0.93	5156.8	1714.8
- production_year_2018	1	0.96	5156.8	1714.8
- production_year_2010	1	1.07	5156.9	1714.8
- nomi_Actor2	1	1.18	5157.0	1714.8
- production_year_2004	1	1.60	5157.5	1714.9
- production_year_2011	1	1.64	5157.5	1714.9
- production_year_2016	1	1.72	5157.6	1714.9
- vote_diff.US.KOR.	1	1.77	5157.6	1714.9
- runtime	1	1.80	5157.7	1715.0
- awards_DIR	1	1.81	5157.7	1715.0
- production_year_2013	1	2.39	5158.3	1715.1
- production_year_2012	1	2.58	5158.4	1715.1
- production_year_2014	1	3.65	5159.5	1715.3
- production_year_2015	1	4.44	5160.3	1715.5
<none>			5155.9	1716.6
- popularity_US	1	16.55	5172.4	1717.9
+ nomi_Actor1	1	0.02	5155.8	1718.6
- production_year_2017	1	50.17	5206.0	1724.5
- vote_average_US	1	55.40	5211.3	1725.5
- vote_count_US	1	69.37	5225.2	1728.3
- log(revenue_US)	1	82.73	5238.6	1730.9
- log(budget)	1	351.42	5507.3	1782.0
- release_diff.US.KOR.	1	1387.09	6542.9	1958.3

Step: AIC=1714.63

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + production_year_2019 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + vote_diff.US.KOR. + nomi_DIR +
 awards_DIR + awards_Actor1 + nomi_Actor2 + awards_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2019	1	0.06	5156.1	1712.7
- awards_Actor2	1	0.21	5156.3	1712.7
- production_year_2008	1	0.33	5156.4	1712.7
- awards_Actor1	1	0.33	5156.4	1712.7
- production_year_2005	1	0.46	5156.5	1712.7
- nomi_DIR	1	0.70	5156.8	1712.8
- production_year_2007	1	1.12	5157.2	1712.9
- nomi_Actor2	1	1.19	5157.3	1712.9
- production_year_2006	1	1.42	5157.5	1712.9

- production_year_2009	1	1.50	5157.6	1712.9
- vote_diff.US.KOR.	1	1.77	5157.8	1713.0
- runtime	1	1.79	5157.8	1713.0
- awards_DIR	1	1.81	5157.9	1713.0
- production_year_2010	1	1.91	5158.0	1713.0
- production_year_2004	1	3.53	5159.6	1713.3
- production_year_2011	1	3.68	5159.7	1713.4
- production_year_2016	1	3.78	5159.8	1713.4
- production_year_2018	1	6.17	5162.2	1713.9
- production_year_2013	1	6.27	5162.3	1713.9
- production_year_2012	1	6.86	5162.9	1714.0
<none>			5156.1	1714.6
- production_year_2014	1	10.85	5166.9	1714.8
- production_year_2015	1	13.91	5170.0	1715.4
- popularity_US	1	16.54	5172.6	1715.9
+ production_year_2003	1	0.20	5155.9	1716.6
+ nomi_Actor1	1	0.02	5156.0	1716.6
- vote_average_US	1	55.38	5211.4	1723.6
- vote_count_US	1	69.32	5225.4	1726.3
- production_year_2017	1	76.96	5233.0	1727.8
- log(revenue_US)	1	82.68	5238.7	1728.9
- log(budget)	1	351.54	5507.6	1780.1
- release_diff.US.KOR.	1	1387.39	6543.4	1956.4

Step: AIC=1712.65

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
 awards_Actor1 + nomi_Actor2 + awards_Actor2

	Df	Sum of Sq	RSS	AIC
- awards_Actor2	1	0.22	5156.3	1710.7
- production_year_2008	1	0.27	5156.4	1710.7
- awards_Actor1	1	0.34	5156.5	1710.7
- production_year_2005	1	0.40	5156.5	1710.7
- nomi_DIR	1	0.70	5156.8	1710.8
- production_year_2007	1	1.09	5157.2	1710.9
- nomi_Actor2	1	1.17	5157.3	1710.9
- production_year_2006	1	1.42	5157.5	1710.9
- production_year_2009	1	1.52	5157.6	1711.0
- vote_diff.US.KOR.	1	1.76	5157.9	1711.0
- runtime	1	1.80	5157.9	1711.0
- awards_DIR	1	1.83	5157.9	1711.0
- production_year_2010	1	1.96	5158.1	1711.0
- production_year_2004	1	3.77	5159.9	1711.4
- production_year_2011	1	3.95	5160.1	1711.4
- production_year_2016	1	4.02	5160.1	1711.4
- production_year_2013	1	6.89	5163.0	1712.0
- production_year_2018	1	7.30	5163.4	1712.1
- production_year_2012	1	7.55	5163.7	1712.1
<none>			5156.1	1712.7

- production_year_2014	1	12.17	5168.3	1713.1
- production_year_2015	1	15.70	5171.8	1713.8
- popularity_US	1	16.55	5172.7	1713.9
+ production_year_2019	1	0.06	5156.1	1714.6
+ production_year_2003	1	0.04	5156.1	1714.6
+ nomi_Actor1	1	0.02	5156.1	1714.6
- vote_average_US	1	55.34	5211.5	1721.6
- vote_count_US	1	69.33	5225.5	1724.3
- production_year_2017	1	78.37	5234.5	1726.1
- log(revenue_US)	1	82.77	5238.9	1726.9
- log(budget)	1	351.62	5507.7	1778.1
- release_diff.US.KOR.	1	1387.38	6543.5	1954.4

Step: AIC=1710.69

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
 production_year_2006 + production_year_2007 + production_year_2008 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
 awards_Actor1 + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- awards_Actor1	1	0.26	5156.6	1708.7
- production_year_2008	1	0.27	5156.6	1708.7
- production_year_2005	1	0.40	5156.7	1708.8
- nomi_DIR	1	0.68	5157.0	1708.8
- production_year_2007	1	1.09	5157.4	1708.9
- production_year_2006	1	1.40	5157.7	1709.0
- production_year_2009	1	1.52	5157.9	1709.0
- vote_diff.US.KOR.	1	1.76	5158.1	1709.0
- runtime	1	1.82	5158.2	1709.0
- awards_DIR	1	1.82	5158.2	1709.0
- production_year_2010	1	1.95	5158.3	1709.1
- production_year_2004	1	3.75	5160.1	1709.4
- production_year_2011	1	3.93	5160.3	1709.5
- production_year_2016	1	4.01	5160.4	1709.5
- nomi_Actor2	1	4.54	5160.9	1709.6
- production_year_2013	1	6.82	5163.2	1710.0
- production_year_2018	1	7.35	5163.7	1710.2
- production_year_2012	1	7.52	5163.9	1710.2
<none>			5156.3	1710.7
- production_year_2014	1	12.14	5168.5	1711.1
- production_year_2015	1	15.65	5172.0	1711.8
- popularity_US	1	16.46	5172.8	1712.0
+ awards_Actor2	1	0.22	5156.1	1712.7
+ production_year_2019	1	0.07	5156.3	1712.7
+ production_year_2003	1	0.03	5156.3	1712.7
+ nomi_Actor1	1	0.00	5156.3	1712.7
- vote_average_US	1	55.53	5211.9	1719.7
- vote_count_US	1	69.59	5225.9	1722.4
- production_year_2017	1	78.32	5234.7	1724.1
- log(revenue_US)	1	82.62	5239.0	1725.0
- log(budget)	1	352.06	5508.4	1776.3

- release_diff.US.KOR. 1 1393.18 6549.5 1953.4

Step: AIC=1708.74

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
production_year_2006 + production_year_2007 + production_year_2008 +
production_year_2009 + production_year_2010 + production_year_2011 +
production_year_2012 + production_year_2013 + production_year_2014 +
production_year_2015 + production_year_2016 + production_year_2017 +
production_year_2018 + release_diff.US.KOR. + log(budget) +
popularity_US + log(revenue_US) + runtime + vote_average_US +
vote_count_US + vote_diff.US.KOR. + nomi_DIR + awards_DIR +
nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2008	1	0.26	5156.9	1706.8
- production_year_2005	1	0.39	5157.0	1706.8
- nomi_DIR	1	0.62	5157.2	1706.9
- production_year_2007	1	1.08	5157.7	1707.0
- production_year_2006	1	1.40	5158.0	1707.0
- production_year_2009	1	1.52	5158.1	1707.0
- vote_diff.US.KOR.	1	1.76	5158.4	1707.1
- awards_DIR	1	1.78	5158.4	1707.1
- runtime	1	1.81	5158.4	1707.1
- production_year_2010	1	1.92	5158.5	1707.1
- production_year_2004	1	3.70	5160.3	1707.5
- production_year_2011	1	3.94	5160.5	1707.5
- production_year_2016	1	4.03	5160.6	1707.5
- nomi_Actor2	1	4.44	5161.0	1707.6
- production_year_2013	1	6.85	5163.4	1708.1
- production_year_2018	1	7.37	5164.0	1708.2
- production_year_2012	1	7.53	5164.1	1708.2
<none>			5156.6	1708.7
- production_year_2014	1	12.20	5168.8	1709.2
- production_year_2015	1	15.64	5172.2	1709.8
- popularity_US	1	16.52	5173.1	1710.0
+ awards_Actor1	1	0.26	5156.3	1710.7
+ awards_Actor2	1	0.14	5156.5	1710.7
+ nomi_Actor1	1	0.14	5156.5	1710.7
+ production_year_2019	1	0.08	5156.5	1710.7
+ production_year_2003	1	0.03	5156.6	1710.7
- vote_average_US	1	55.45	5212.0	1717.7
- vote_count_US	1	69.72	5226.3	1720.5
- production_year_2017	1	78.37	5235.0	1722.2
- log(revenue_US)	1	83.21	5239.8	1723.1
- log(budget)	1	351.84	5508.4	1774.3
- release_diff.US.KOR.	1	1398.35	6554.9	1952.2

Step: AIC=1706.79

log(audience_KOR) ~ production_year_2004 + production_year_2005 +
production_year_2006 + production_year_2007 + production_year_2009 +
production_year_2010 + production_year_2011 + production_year_2012 +
production_year_2013 + production_year_2014 + production_year_2015 +
production_year_2016 + production_year_2017 + production_year_2018 +
release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
runtime + vote_average_US + vote_count_US + vote_diff.US.KOR. +
nomi_DIR + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2005	1	0.18	5157.0	1704.8
- nomi_DIR	1	0.64	5157.5	1704.9
- awards_DIR	1	1.74	5158.6	1705.1
- vote_diff.US.KOR.	1	1.75	5158.6	1705.1
- runtime	1	1.83	5158.7	1705.2
- production_year_2007	1	2.35	5159.2	1705.3
- production_year_2006	1	3.82	5160.7	1705.5
- production_year_2009	1	4.37	5161.2	1705.7
- nomi_Actor2	1	4.49	5161.3	1705.7
- production_year_2010	1	6.45	5163.3	1706.1
<none>			5156.9	1706.8
- production_year_2016	1	12.32	5169.2	1707.2
- production_year_2004	1	15.27	5172.1	1707.8
- popularity_US	1	16.53	5173.4	1708.1
- production_year_2011	1	17.93	5174.8	1708.3
+ production_year_2008	1	0.26	5156.6	1708.7
+ awards_Actor1	1	0.25	5156.6	1708.7
- production_year_2018	1	20.07	5176.9	1708.8
+ awards_Actor2	1	0.14	5156.7	1708.8
+ nomi_Actor1	1	0.12	5156.7	1708.8
+ production_year_2003	1	0.10	5156.8	1708.8
+ production_year_2019	1	0.00	5156.9	1708.8
- production_year_2013	1	36.09	5192.9	1711.9
- production_year_2012	1	36.19	5193.0	1712.0
- vote_average_US	1	55.25	5212.1	1715.7
- vote_count_US	1	69.94	5226.8	1718.6
- production_year_2014	1	71.31	5228.2	1718.8
- log(revenue_US)	1	83.32	5240.2	1721.2
- production_year_2017	1	86.34	5243.2	1721.8
- production_year_2015	1	96.53	5253.4	1723.8
- log(budget)	1	351.99	5508.8	1772.3
- release_diff.US.KOR.	1	1400.66	6557.5	1950.6

Step: AIC=1704.83

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2007 + production_year_2009 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + vote_diff.US.KOR. + nomi_DIR +
 awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- nomi_DIR	1	0.62	5157.7	1703.0
- vote_diff.US.KOR.	1	1.75	5158.8	1703.2
- awards_DIR	1	1.76	5158.8	1703.2
- runtime	1	1.99	5159.0	1703.2
- production_year_2007	1	2.26	5159.3	1703.3
- production_year_2006	1	3.91	5160.9	1703.6
- nomi_Actor2	1	4.47	5161.5	1703.7
- production_year_2009	1	4.55	5161.6	1703.7
- production_year_2010	1	6.94	5164.0	1704.2
<none>			5157.0	1704.8

- production_year_2016	1	12.94	5170.0	1705.4
- popularity_US	1	16.54	5173.6	1706.1
- production_year_2004	1	17.32	5174.4	1706.3
+ awards_Actor1	1	0.26	5156.8	1706.8
+ production_year_2005	1	0.18	5156.9	1706.8
+ awards_Actor2	1	0.14	5156.9	1706.8
+ production_year_2003	1	0.14	5156.9	1706.8
+ nomi_Actor1	1	0.14	5156.9	1706.8
+ production_year_2008	1	0.04	5157.0	1706.8
+ production_year_2019	1	0.00	5157.0	1706.8
- production_year_2011	1	20.68	5177.7	1706.9
- production_year_2018	1	21.52	5178.6	1707.1
- production_year_2012	1	41.86	5198.9	1711.1
- production_year_2013	1	42.70	5199.7	1711.3
- vote_average_US	1	55.25	5212.3	1713.7
- vote_count_US	1	70.59	5227.6	1716.7
- log(revenue_US)	1	83.24	5240.3	1719.2
- production_year_2014	1	86.13	5243.2	1719.8
- production_year_2017	1	86.20	5243.2	1719.8
- production_year_2015	1	117.91	5274.9	1726.0
- log(budget)	1	351.86	5508.9	1770.3
- release_diff.US.KOR.	1	1400.53	6557.6	1948.6

Step: AIC=1702.95

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2007 + production_year_2009 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + vote_diff.US.KOR. + awards_DIR +
 nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- vote_diff.US.KOR.	1	1.77	5159.4	1701.3
- production_year_2007	1	2.23	5159.9	1701.4
- runtime	1	2.30	5160.0	1701.4
- production_year_2006	1	3.78	5161.4	1701.7
- production_year_2009	1	4.40	5162.1	1701.8
- nomi_Actor2	1	5.56	5163.2	1702.0
- production_year_2010	1	6.75	5164.4	1702.3
<none>			5157.7	1703.0
- production_year_2016	1	12.79	5170.4	1703.5
- popularity_US	1	16.47	5174.1	1704.2
- production_year_2004	1	16.97	5174.6	1704.3
+ nomi_DIR	1	0.62	5157.0	1704.8
+ awards_Actor1	1	0.20	5157.5	1704.9
+ production_year_2005	1	0.16	5157.5	1704.9
+ production_year_2003	1	0.14	5157.5	1704.9
+ awards_Actor2	1	0.14	5157.5	1704.9
+ nomi_Actor1	1	0.05	5157.6	1704.9
+ production_year_2008	1	0.03	5157.6	1704.9
+ production_year_2019	1	0.00	5157.7	1705.0
- production_year_2011	1	20.61	5178.3	1705.0
- awards_DIR	1	21.40	5179.1	1705.2
- production_year_2018	1	21.77	5179.4	1705.3

- production_year_2012	1	41.62	5199.3	1709.2
- production_year_2013	1	42.77	5200.4	1709.4
- vote_average_US	1	55.77	5213.4	1712.0
- vote_count_US	1	70.57	5228.2	1714.8
- log(revenue_US)	1	83.18	5240.8	1717.3
- production_year_2014	1	85.84	5243.5	1717.8
- production_year_2017	1	86.09	5243.7	1717.9
- production_year_2015	1	118.84	5276.5	1724.3
- log(budget)	1	353.67	5511.3	1768.8
- release_diff.US.KOR.	1	1414.52	6572.2	1948.9

Step: AIC=1701.3

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2007 + production_year_2009 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + runtime +
 vote_average_US + vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2007	1	1.88	5161.3	1699.7
- runtime	1	2.20	5161.6	1699.7
- production_year_2006	1	3.77	5163.2	1700.0
- production_year_2009	1	4.41	5163.8	1700.2
- nomi_Actor2	1	5.61	5165.0	1700.4
- production_year_2010	1	6.75	5166.2	1700.6
<none>			5159.4	1701.3
- production_year_2016	1	12.75	5172.2	1701.8
- popularity_US	1	16.47	5175.9	1702.6
- production_year_2004	1	16.96	5176.4	1702.7
+ vote_diff.US.KOR.	1	1.77	5157.7	1703.0
+ nomi_DIR	1	0.63	5158.8	1703.2
+ awards_Actor1	1	0.19	5159.2	1703.3
+ production_year_2005	1	0.16	5159.3	1703.3
+ production_year_2003	1	0.14	5159.3	1703.3
+ awards_Actor2	1	0.14	5159.3	1703.3
+ nomi_Actor1	1	0.05	5159.4	1703.3
+ production_year_2008	1	0.03	5159.4	1703.3
+ production_year_2019	1	0.00	5159.4	1703.3
- production_year_2011	1	20.61	5180.0	1703.4
- awards_DIR	1	21.51	5180.9	1703.6
- production_year_2018	1	21.83	5181.3	1703.6
- production_year_2012	1	41.67	5201.1	1707.5
- production_year_2013	1	42.81	5202.2	1707.8
- vote_average_US	1	55.77	5215.2	1710.3
- vote_count_US	1	70.59	5230.0	1713.2
- log(revenue_US)	1	83.10	5242.5	1715.7
- production_year_2014	1	85.88	5245.3	1716.2
- production_year_2017	1	86.15	5245.6	1716.2
- production_year_2015	1	118.90	5278.3	1722.6
- log(budget)	1	352.67	5512.1	1766.9
- release_diff.US.KOR.	1	1416.88	6576.3	1947.5

Step: AIC=1699.67

log(audience_KOR) ~ production_year_2004 + production_year_2006 +

production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + runtime + vote_average_US +
 vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- runtime	1	2.14	5163.4	1698.1
- production_year_2006	1	2.53	5163.8	1698.2
- production_year_2009	1	3.11	5164.4	1698.3
- production_year_2010	1	5.20	5166.5	1698.7
- nomi_Actor2	1	5.32	5166.6	1698.7
<none>			5161.3	1699.7
- production_year_2016	1	11.22	5172.5	1699.9
- production_year_2004	1	15.09	5176.4	1700.7
- popularity_US	1	16.41	5177.7	1700.9
+ production_year_2007	1	1.88	5159.4	1701.3
- production_year_2011	1	18.73	5180.0	1701.4
+ vote_diff.US.KOR.	1	1.42	5159.9	1701.4
+ production_year_2008	1	0.70	5160.6	1701.5
+ nomi_DIR	1	0.60	5160.7	1701.6
+ production_year_2003	1	0.24	5161.1	1701.6
+ awards_Actor1	1	0.22	5161.1	1701.6
+ awards_Actor2	1	0.12	5161.2	1701.7
+ production_year_2005	1	0.07	5161.2	1701.7
+ nomi_Actor1	1	0.07	5161.2	1701.7
+ production_year_2019	1	0.00	5161.3	1701.7
- awards_DIR	1	21.11	5182.4	1701.8
- production_year_2018	1	23.72	5185.0	1702.4
- production_year_2012	1	40.01	5201.3	1705.6
- production_year_2013	1	41.43	5202.7	1705.8
- vote_average_US	1	54.12	5215.4	1708.3
- vote_count_US	1	72.10	5233.4	1711.9
- log(revenue_US)	1	82.71	5244.0	1713.9
- production_year_2017	1	85.37	5246.7	1714.5
- production_year_2014	1	86.63	5247.9	1714.7
- production_year_2015	1	122.02	5283.3	1721.6
- log(budget)	1	351.92	5513.2	1765.2
- release_diff.US.KOR.	1	1415.61	6576.9	1945.6

Step: AIC=1698.1

log(audience_KOR) ~ production_year_2004 + production_year_2006 +
 production_year_2009 + production_year_2010 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2016 + production_year_2017 +
 production_year_2018 + release_diff.US.KOR. + log(budget) +
 popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
 awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2006	1	2.39	5165.8	1696.6
- production_year_2009	1	3.15	5166.6	1696.7
- production_year_2010	1	4.68	5168.1	1697.0
- nomi_Actor2	1	5.04	5168.5	1697.1
<none>			5163.4	1698.1

- production_year_2016	1	11.21	5174.7	1698.3
- production_year_2004	1	15.36	5178.8	1699.1
- popularity_US	1	16.90	5180.3	1699.4
- production_year_2011	1	18.05	5181.5	1699.7
+ runtime	1	2.14	5161.3	1699.7
+ production_year_2007	1	1.83	5161.6	1699.7
+ vote_diff.US.KOR.	1	1.33	5162.1	1699.8
+ production_year_2008	1	0.89	5162.5	1699.9
+ nomi_DIR	1	0.89	5162.6	1699.9
+ production_year_2003	1	0.29	5163.2	1700.0
+ awards_Actor1	1	0.21	5163.2	1700.1
+ awards_Actor2	1	0.14	5163.3	1700.1
+ nomi_Actor1	1	0.08	5163.4	1700.1
+ production_year_2005	1	0.01	5163.4	1700.1
+ production_year_2019	1	0.00	5163.4	1700.1
- production_year_2018	1	23.61	5187.1	1700.8
- awards_DIR	1	24.05	5187.5	1700.8
- production_year_2012	1	40.30	5203.7	1704.0
- production_year_2013	1	40.73	5204.2	1704.1
- vote_average_US	1	52.11	5215.6	1706.4
- vote_count_US	1	70.02	5233.5	1709.9
- log(revenue_US)	1	84.39	5247.8	1712.7
- production_year_2017	1	85.49	5248.9	1712.9
- production_year_2014	1	86.11	5249.6	1713.0
- production_year_2015	1	121.33	5284.8	1719.9
- log(budget)	1	352.28	5515.7	1763.6
- release_diff.US.KOR.	1	1416.23	6579.7	1944.0

Step: AIC=1696.57

log(audience_KOR) ~ production_year_2004 + production_year_2009 +
 production_year_2010 + production_year_2011 + production_year_2012 +
 production_year_2013 + production_year_2014 + production_year_2015 +
 production_year_2016 + production_year_2017 + production_year_2018 +
 release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
 vote_average_US + vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2009	1	2.08	5167.9	1695.0
- production_year_2010	1	3.38	5169.2	1695.2
- nomi_Actor2	1	5.18	5171.0	1695.6
- production_year_2016	1	9.77	5175.6	1696.5
<none>			5165.8	1696.6
- production_year_2004	1	13.38	5179.2	1697.2
- production_year_2011	1	15.93	5181.8	1697.7
- popularity_US	1	16.99	5182.8	1697.9
+ production_year_2006	1	2.39	5163.4	1698.1
+ production_year_2008	1	2.02	5163.8	1698.2
+ runtime	1	1.99	5163.8	1698.2
+ vote_diff.US.KOR.	1	1.44	5164.4	1698.3
+ nomi_DIR	1	0.76	5165.1	1698.4
+ production_year_2007	1	0.64	5165.2	1698.4
+ production_year_2003	1	0.41	5165.4	1698.5
+ production_year_2005	1	0.34	5165.5	1698.5
+ awards_Actor1	1	0.24	5165.6	1698.5
+ awards_Actor2	1	0.08	5165.7	1698.5
+ nomi_Actor1	1	0.07	5165.8	1698.6

+ production_year_2019	1	0.01	5165.8	1698.6
- awards_DIR	1	24.13	5190.0	1699.3
- production_year_2018	1	25.18	5191.0	1699.5
- production_year_2012	1	37.94	5203.8	1702.1
- production_year_2013	1	38.35	5204.2	1702.1
- vote_average_US	1	51.70	5217.5	1704.8
- vote_count_US	1	70.46	5236.3	1708.4
- log(revenue_US)	1	83.40	5249.2	1711.0
- production_year_2014	1	84.31	5250.1	1711.1
- production_year_2017	1	84.68	5250.5	1711.2
- production_year_2015	1	120.47	5286.3	1718.2
- log(budget)	1	352.56	5518.4	1762.1
- release_diff.US.KOR.	1	1417.98	6583.8	1942.7

Step: AIC=1694.98

log(audience_KOR) ~ production_year_2004 + production_year_2010 +
 production_year_2011 + production_year_2012 + production_year_2013 +
 production_year_2014 + production_year_2015 + production_year_2016 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + vote_average_US +
 vote_count_US + awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- production_year_2010	1	2.51	5170.4	1693.5
- nomi_Actor2	1	5.49	5173.4	1694.1
- production_year_2016	1	8.69	5176.6	1694.7
<none>			5167.9	1695.0
- production_year_2004	1	11.86	5179.8	1695.3
- production_year_2011	1	14.25	5182.2	1695.8
+ production_year_2008	1	3.07	5164.8	1696.4
- popularity_US	1	17.23	5185.1	1696.4
+ production_year_2009	1	2.08	5165.8	1696.6
+ runtime	1	2.06	5165.8	1696.6
+ vote_diff.US.KOR.	1	1.53	5166.4	1696.7
+ production_year_2006	1	1.32	5166.6	1696.7
+ production_year_2005	1	0.83	5167.1	1696.8
+ nomi_DIR	1	0.68	5167.2	1696.8
+ production_year_2003	1	0.50	5167.4	1696.9
+ awards_Actor1	1	0.29	5167.6	1696.9
+ production_year_2007	1	0.20	5167.7	1696.9
+ nomi_Actor1	1	0.14	5167.8	1697.0
+ awards_Actor2	1	0.11	5167.8	1697.0
+ production_year_2019	1	0.01	5167.9	1697.0
- awards_DIR	1	24.33	5192.2	1697.8
- production_year_2018	1	25.96	5193.9	1698.1
- production_year_2012	1	35.94	5203.8	1700.1
- production_year_2013	1	36.29	5204.2	1700.1
- vote_average_US	1	52.09	5220.0	1703.2
- vote_count_US	1	69.71	5237.6	1706.7
- production_year_2014	1	82.42	5250.3	1709.2
- log(revenue_US)	1	83.85	5251.8	1709.5
- production_year_2017	1	84.12	5252.0	1709.5
- production_year_2015	1	119.00	5286.9	1716.3
- log(budget)	1	352.38	5520.3	1760.5
- release_diff.US.KOR.	1	1417.10	6585.0	1940.9

Step: AIC=1693.48

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
production_year_2012 + production_year_2013 + production_year_2014 +
production_year_2015 + production_year_2016 + production_year_2017 +
production_year_2018 + release_diff.US.KOR. + log(budget) +
popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
awards_DIR + nomi_Actor2

	Df	Sum of Sq	RSS	AIC
- nomi_Actor2	1	5.44	5175.9	1692.6
- production_year_2016	1	7.69	5178.1	1693.0
<none>			5170.4	1693.5
- production_year_2004	1	10.48	5180.9	1693.5
- production_year_2011	1	12.62	5183.0	1694.0
+ production_year_2008	1	4.21	5166.2	1694.7
- popularity_US	1	17.55	5188.0	1695.0
+ production_year_2010	1	2.51	5167.9	1695.0
+ runtime	1	1.68	5168.7	1695.2
+ vote_diff.US.KOR.	1	1.62	5168.8	1695.2
+ production_year_2005	1	1.45	5169.0	1695.2
+ production_year_2009	1	1.20	5169.2	1695.2
+ production_year_2006	1	0.66	5169.8	1695.3
+ production_year_2003	1	0.60	5169.8	1695.4
+ nomi_DIR	1	0.58	5169.8	1695.4
+ awards_Actor1	1	0.21	5170.2	1695.4
+ nomi_Actor1	1	0.13	5170.3	1695.5
+ awards_Actor2	1	0.09	5170.3	1695.5
+ production_year_2019	1	0.02	5170.4	1695.5
+ production_year_2007	1	0.02	5170.4	1695.5
- awards_DIR	1	23.90	5194.3	1696.2
- production_year_2018	1	27.00	5197.4	1696.8
- production_year_2012	1	33.78	5204.2	1698.1
- production_year_2013	1	34.04	5204.5	1698.2
- vote_average_US	1	52.34	5222.8	1701.8
- vote_count_US	1	68.30	5238.7	1704.9
- production_year_2014	1	79.92	5250.3	1707.2
- production_year_2017	1	83.57	5254.0	1707.9
- log(revenue_US)	1	84.49	5254.9	1708.1
- production_year_2015	1	116.60	5287.0	1714.3
- log(budget)	1	352.69	5523.1	1759.0
- release_diff.US.KOR.	1	1420.20	6590.6	1939.8

Step: AIC=1692.56

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
production_year_2012 + production_year_2013 + production_year_2014 +
production_year_2015 + production_year_2016 + production_year_2017 +
production_year_2018 + release_diff.US.KOR. + log(budget) +
popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2016	1	7.00	5182.9	1691.9
<none>			5175.9	1692.6
- production_year_2004	1	10.55	5186.4	1692.6
- production_year_2011	1	12.52	5188.4	1693.0
+ nomi_Actor2	1	5.44	5170.4	1693.5

+ awards_Actor2	1	4.13	5171.7	1693.7
+ production_year_2008	1	3.94	5171.9	1693.8
- popularity_US	1	17.21	5193.1	1694.0
+ production_year_2010	1	2.46	5173.4	1694.1
+ vote_diff.US.KOR.	1	1.71	5174.2	1694.2
+ nomi_DIR	1	1.58	5174.3	1694.2
+ production_year_2005	1	1.54	5174.3	1694.2
+ production_year_2009	1	1.44	5174.4	1694.3
+ runtime	1	1.42	5174.4	1694.3
+ nomi_Actor1	1	0.92	5174.9	1694.4
+ production_year_2006	1	0.70	5175.2	1694.4
+ production_year_2003	1	0.66	5175.2	1694.4
+ awards_Actor1	1	0.34	5175.5	1694.5
+ production_year_2019	1	0.03	5175.8	1694.5
+ production_year_2007	1	0.00	5175.9	1694.6
- production_year_2018	1	27.77	5203.6	1696.0
- production_year_2012	1	32.35	5208.2	1696.9
- awards_DIR	1	32.98	5208.8	1697.0
- production_year_2013	1	34.43	5210.3	1697.3
- vote_average_US	1	51.73	5227.6	1700.7
- vote_count_US	1	67.93	5243.8	1703.9
- production_year_2014	1	77.12	5253.0	1705.7
- production_year_2017	1	83.13	5259.0	1706.9
- log(revenue_US)	1	83.99	5259.9	1707.0
- production_year_2015	1	112.86	5288.7	1712.6
- log(budget)	1	350.24	5526.1	1757.5
- release_diff.US.KOR.	1	1416.31	6592.2	1938.0

Step: AIC=1691.94

log(audience_KOR) ~ production_year_2004 + production_year_2011 +
 production_year_2012 + production_year_2013 + production_year_2014 +
 production_year_2015 + production_year_2017 + production_year_2018 +
 release_diff.US.KOR. + log(budget) + popularity_US + log(revenue_US) +
 vote_average_US + vote_count_US + awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2004	1	9.04	5191.9	1691.7
<none>			5182.9	1691.9
- production_year_2011	1	10.67	5193.5	1692.0
+ production_year_2016	1	7.00	5175.9	1692.6
+ production_year_2008	1	5.38	5177.5	1692.9
+ nomi_Actor2	1	4.75	5178.1	1693.0
- popularity_US	1	16.21	5199.1	1693.1
+ awards_Actor2	1	3.64	5179.2	1693.2
+ production_year_2005	1	2.37	5180.5	1693.5
+ vote_diff.US.KOR.	1	1.75	5181.1	1693.6
+ runtime	1	1.52	5181.3	1693.6
+ production_year_2010	1	1.51	5181.4	1693.6
+ nomi_DIR	1	1.45	5181.4	1693.7
+ production_year_2003	1	0.81	5182.1	1693.8
+ production_year_2009	1	0.76	5182.1	1693.8
+ nomi_Actor1	1	0.62	5182.2	1693.8
+ production_year_2006	1	0.28	5182.6	1693.9
+ awards_Actor1	1	0.17	5182.7	1693.9
+ production_year_2007	1	0.07	5182.8	1693.9
+ production_year_2019	1	0.04	5182.8	1693.9

- production_year_2018	1	28.57	5211.4	1695.6
- production_year_2012	1	29.41	5212.3	1695.7
- production_year_2013	1	31.34	5214.2	1696.1
- awards_DIR	1	33.16	5216.0	1696.5
- vote_average_US	1	58.51	5241.4	1701.4
- vote_count_US	1	65.36	5248.2	1702.8
- production_year_2014	1	72.33	5255.2	1704.1
- production_year_2017	1	82.09	5265.0	1706.0
- log(revenue_US)	1	83.18	5266.0	1706.2
- production_year_2015	1	107.49	5290.4	1710.9
- log(budget)	1	352.49	5535.3	1757.2
- release_diff.US.KOR.	1	1413.77	6596.6	1936.7

Step: AIC=1691.72

log(audience_KOR) ~ production_year_2011 + production_year_2012 +
 production_year_2013 + production_year_2014 + production_year_2015 +
 production_year_2017 + production_year_2018 + release_diff.US.KOR. +
 log(budget) + popularity_US + log(revenue_US) + vote_average_US +
 vote_count_US + awards_DIR

	Df	Sum of Sq	RSS	AIC
- production_year_2011	1	8.47	5200.4	1691.4
<none>			5191.9	1691.7
+ production_year_2004	1	9.04	5182.9	1691.9
+ production_year_2008	1	7.55	5184.4	1692.2
+ production_year_2016	1	5.49	5186.4	1692.6
+ nomi_Actor2	1	4.89	5187.0	1692.8
- popularity_US	1	15.98	5207.9	1692.9
+ production_year_2005	1	3.84	5188.1	1693.0
+ awards_Actor2	1	3.64	5188.3	1693.0
+ vote_diff.US.KOR.	1	1.85	5190.1	1693.4
+ runtime	1	1.83	5190.1	1693.4
+ nomi_DIR	1	1.21	5190.7	1693.5
+ production_year_2003	1	0.96	5190.9	1693.5
+ nomi_Actor1	1	0.94	5191.0	1693.5
+ production_year_2010	1	0.62	5191.3	1693.6
+ production_year_2007	1	0.48	5191.4	1693.6
+ awards_Actor1	1	0.33	5191.6	1693.7
+ production_year_2009	1	0.17	5191.7	1693.7
+ production_year_2019	1	0.06	5191.8	1693.7
+ production_year_2006	1	0.01	5191.9	1693.7
- production_year_2012	1	26.35	5218.3	1694.9
- production_year_2013	1	27.85	5219.8	1695.2
- production_year_2018	1	29.93	5221.8	1695.6
- awards_DIR	1	34.32	5226.2	1696.5
- vote_average_US	1	58.34	5250.2	1701.2
- production_year_2014	1	67.07	5259.0	1702.8
- vote_count_US	1	68.92	5260.8	1703.2
- log(revenue_US)	1	81.10	5273.0	1705.6
- production_year_2017	1	81.35	5273.3	1705.6
- production_year_2015	1	101.39	5293.3	1709.5
- log(budget)	1	347.93	5539.8	1756.1
- release_diff.US.KOR.	1	1407.73	6599.6	1935.2

Step: AIC=1691.39

log(audience_KOR) ~ production_year_2012 + production_year_2013 +

```
production_year_2014 + production_year_2015 + production_year_2017 +
production_year_2018 + release_diff.US.KOR. + log(budget) +
popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
awards_DIR
```

	Df	Sum of Sq	RSS	AIC
<none>			5200.4	1691.4
+ production_year_2008	1	9.86	5190.5	1691.5
+ production_year_2011	1	8.47	5191.9	1691.7
+ production_year_2004	1	6.84	5193.5	1692.0
+ production_year_2005	1	5.43	5194.9	1692.3
+ nomi_Actor2	1	4.87	5195.5	1692.4
+ production_year_2016	1	4.18	5196.2	1692.6
- popularity_US	1	16.40	5216.8	1692.6
+ awards_Actor2	1	3.51	5196.9	1692.7
+ vote_diff.US.KOR.	1	1.98	5198.4	1693.0
+ runtime	1	1.45	5198.9	1693.1
+ nomi_DIR	1	1.32	5199.0	1693.1
+ production_year_2003	1	1.15	5199.2	1693.2
+ production_year_2007	1	1.13	5199.2	1693.2
+ nomi_Actor1	1	0.81	5199.6	1693.2
+ awards_Actor1	1	0.20	5200.2	1693.3
+ production_year_2010	1	0.13	5200.2	1693.4
+ production_year_2006	1	0.09	5200.3	1693.4
+ production_year_2019	1	0.08	5200.3	1693.4
+ production_year_2009	1	0.00	5200.4	1693.4
- production_year_2012	1	23.04	5223.4	1693.9
- production_year_2013	1	24.31	5224.7	1694.2
- production_year_2018	1	30.53	5230.9	1695.4
- awards_DIR	1	33.92	5234.3	1696.0
- vote_average_US	1	59.91	5260.3	1701.1
- production_year_2014	1	62.15	5262.5	1701.5
- vote_count_US	1	66.18	5266.6	1702.3
- production_year_2017	1	80.61	5281.0	1705.1
- log(revenue_US)	1	85.01	5285.4	1706.0
- production_year_2015	1	95.56	5295.9	1708.0
- log(budget)	1	349.59	5550.0	1756.0
- release_diff.US.KOR.	1	1399.36	6599.7	1933.2

Hide

```
summary(movie.lm.step3)
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2012 + production_year_2013 +  
  production_year_2014 + production_year_2015 + production_year_2017 +  
  production_year_2018 + release_diff.US.KOR. + log(budget) +  
  popularity_US + log(revenue_US) + vote_average_US + vote_count_US +  
  awards_DIR, data = movie.2)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-9.7652	-0.6878	0.4658	1.2842	10.0785

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-4.443e+00	1.504e+00	-2.955	0.003203	**
production_year_2012	-6.193e-01	2.929e-01	-2.114	0.034743	*
production_year_2013	-5.727e-01	2.637e-01	-2.172	0.030101	*
production_year_2014	-9.143e-01	2.633e-01	-3.472	0.000538	***
production_year_2015	-1.113e+00	2.586e-01	-4.306	1.83e-05	***
production_year_2017	-9.031e+00	2.284e+00	-3.955	8.19e-05	***
production_year_2018	2.027e+00	8.327e-01	2.434	0.015109	*
release_diff.US.KOR.	-8.242e-02	5.002e-03	-16.478	< 2e-16	***
log(budget)	5.215e-01	6.332e-02	8.236	5.49e-16	***
popularity_US	3.100e-03	1.738e-03	1.784	0.074744	.
log(revenue_US)	2.511e-01	6.182e-02	4.061	5.26e-05	***
vote_average_US	3.696e-01	1.084e-01	3.409	0.000677	***
vote_count_US	2.293e-04	6.399e-05	3.583	0.000355	***
awards_DIR	-4.477e-02	1.745e-02	-2.565	0.010450	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.27 on 1009 degrees of freedom
Multiple R-squared: 0.4729, Adjusted R-squared: 0.4661
F-statistic: 69.64 on 13 and 1009 DF, p-value: < 2.2e-16

Hide

```
movie.lm.step_pred3 <- predict(movie.lm.step3,valid_df)  
accuracy(movie.lm.step_pred1,valid_df$audience_KOR)
```

	ME	RMSE	MAE	MPE	MAPE
Test set	641355.7	1331129	641355.8	89.10523	105.0101

Hide

```
accuracy(movie.lm.step_pred2,valid_df$audience_KOR)
```

	ME	RMSE	MAE	MPE	MAPE
Test set	641355.7	1331129	641355.8	89.06951	105.0326

Hide

```
accuracy(movie.lm.step_pred3,valid_df$audience_KOR)
```

	ME	RMSE	MAE	MPE	MAPE
Test set	641355.7	1331129	641355.8	89.06951	105.0326

[Hide](#)

```
fit <- movie.lm.step3
summary(fit)
```

Call:

```
lm(formula = log(audience_KOR) ~ production_year_2012 + production_year_2013 +
  production_year_2014 + production_year_2015 + production_year_2017 +
  production_year_2018 + release_diff.US.KOR. + log(budget) +
  popularity_US + log(revenue_US) + vote_average_US + vote_count_US +
  awards_DIR, data = movie.2)
```

Residuals:

Min	1Q	Median	3Q	Max
-9.7652	-0.6878	0.4658	1.2842	10.0785

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-4.443e+00	1.504e+00	-2.955	0.003203	**
production_year_2012	-6.193e-01	2.929e-01	-2.114	0.034743	*
production_year_2013	-5.727e-01	2.637e-01	-2.172	0.030101	*
production_year_2014	-9.143e-01	2.633e-01	-3.472	0.000538	***
production_year_2015	-1.113e+00	2.586e-01	-4.306	1.83e-05	***
production_year_2017	-9.031e+00	2.284e+00	-3.955	8.19e-05	***
production_year_2018	2.027e+00	8.327e-01	2.434	0.015109	*
release_diff.US.KOR.	-8.242e-02	5.002e-03	-16.478	< 2e-16	***
log(budget)	5.215e-01	6.332e-02	8.236	5.49e-16	***
popularity_US	3.100e-03	1.738e-03	1.784	0.074744	.
log(revenue_US)	2.511e-01	6.182e-02	4.061	5.26e-05	***
vote_average_US	3.696e-01	1.084e-01	3.409	0.000677	***
vote_count_US	2.293e-04	6.399e-05	3.583	0.000355	***
awards_DIR	-4.477e-02	1.745e-02	-2.565	0.010450	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.27 on 1009 degrees of freedom

Multiple R-squared: 0.4729, Adjusted R-squared: 0.4661

F-statistic: 69.64 on 13 and 1009 DF, p-value: < 2.2e-16

[Hide](#)

```

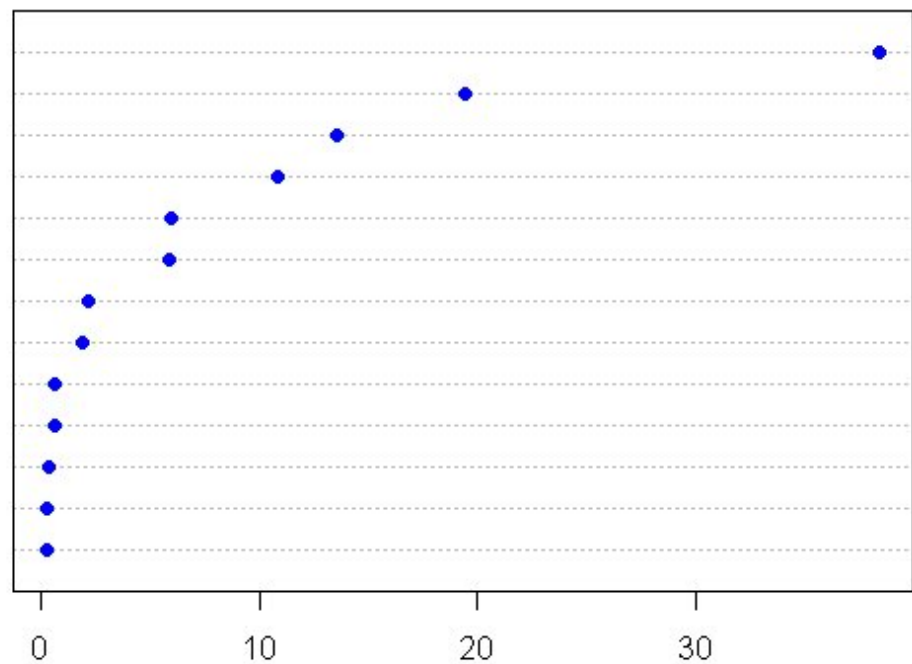
relweights <- function(fit,...){
  R <- cor(fit$model)
  nvar <- ncol(R)
  rxx <- R[2:nvar, 2:nvar]
  rxy <- R[2:nvar, 1]
  svd <- eigen(rxx)
  evec <- svd$vectors
  ev <- svd$values
  delta <- diag(sqrt(ev))
  lambda <- evec %*% delta %*% t(evec)
  lambdasq <- lambda ^ 2
  beta <- solve(lambda) %*% rxy
  rsquare <- colSums(beta ^ 2)
  rawwgt <- lambdasq %*% beta ^ 2
  import <- (rawwgt / rsquare) * 100
  import <- as.data.frame(import)
  row.names(import) <- names(fit$model[2:nvar])
  names(import) <- "Weights"
  import <- import[order(import),1, drop=FALSE]
  dotchart(import$Weights, labels=row.names(import),
            xlab="% of R-Square", pch=19,
            main="Relative Importance of Predictor Variables",
            sub=paste("Total R-Square=", round(rsquare, digits=3)),
            ...)
  return(import)
}
relweights(fit, col="blue")

```

	Weights
production_year_2013	0.2078753
production_year_2012	0.2177176
awards_DIR	0.2870244
production_year_2014	0.5559433
production_year_2018	0.6334703
production_year_2015	1.8690247
production_year_2017	2.1584233
popularity_US	5.8044345
vote_average_US	5.9312637
vote_count_US	10.8449404
log(revenue_US)	13.5855447
log(budget)	19.4402648
release_diff.US.KOR.	38.4640732

Relative Importance of Predictor Variables

release_diff.US.KOR.
log(budget)
log(revenue_US)
vote_count_US
vote_average_US
popularity_US
production_year_2017
production_year_2015
production_year_2018
production_year_2014
awards_DIR
production_year_2012
production_year_2013



% of R-Square
Total R-Square= 0.473

Hide

```
hist(fit$residuals, breaks = 20, xlab = "Residuals", main = "Distribution of errors")
```

Distribution of errors

