

Regression Model to Predict Number of US Theaters in Which Foreign Language Film will be Released - Rebecca Hyde

The best adjusted R-squared that I have been able to achieve so far is .113. The predictor variables unfortunately do not explain a large amount of the variance in the target variable. This regression model includes the following features: release month, release year, runtime, MPAA rating, country, distributor and uses 1079 observations. The residual plot looks random. I would like to work further on feature selection to see if the adj. R-squared can be improved.

Dep. Variable:	Y	R-squared:	0.151
Model:	OLS	Adj. R-squared:	0.113
Method:	Least Squares	F-statistic:	3.983
Date:	Tue, 04 Oct 2016	Prob (F-statistic):	1.80e-16
Time:	20:27:14	Log-Likelihood:	-7077.7
No. Observations:	1079	AIC:	1.425e+04
Df Residuals:	1032	BIC:	1.448e+04
Df Model:	46		
Covariance Type:	nonrobust		

Residual Plot

