Video Games Sales with Ratings
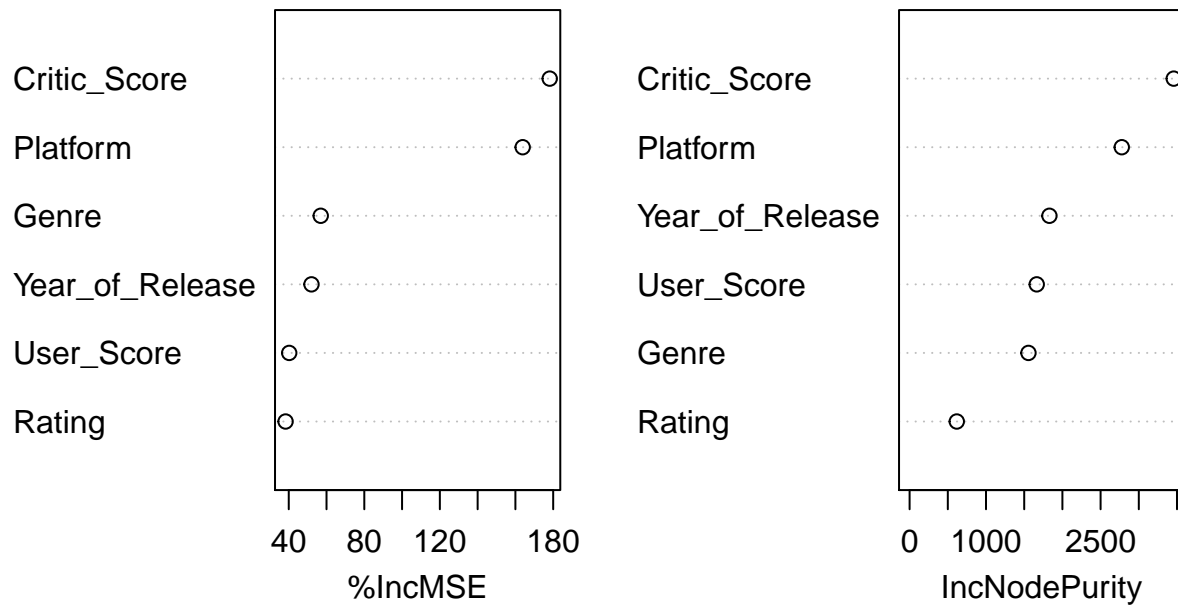
Math 4322 Final Project, Group 24

Jay Nguyen, Alonso Munoz, Andrew-Son Le

30 April 2021

```r
library(randomForest) # testing stuff
load("~/Desktop/paper.RData")
```

```r
varImpPlot(rf_full)
```

rf_full


```r
summary(lm_full)
```

```
##
## Call:
## lm(formula = log(Global_Sales) ~ Platform + Year_of_Release +
##       Genre + Critic_Score + User_Score + Rating, data = data_global)
##
## Residuals:
```

```
##      Min      1Q  Median      3Q     Max
## -4.6119 -0.7323  0.0002  0.7478  5.1099
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -2.229366   0.489715  -4.552 5.40e-06 ***
## PlatformDC         -1.339017   0.335791  -3.988 6.74e-05 ***
## PlatformDS         -0.010175   0.114659  -0.089 0.929292
## PlatformGBA        -0.338122   0.135736  -2.491 0.012761 *
## PlatformGC         -0.390567   0.128710  -3.034 0.002419 **
## PlatformPC         -1.727686   0.107244 -16.110  < 2e-16 ***
## PlatformPS          0.024085   0.180352   0.134 0.893766
## PlatformPS2         0.214312   0.116090   1.846 0.064924 .
## PlatformPS3         0.157998   0.104710   1.509 0.131371
## PlatformPS4         0.174147   0.128087   1.360 0.174002
## PlatformPSP        -0.204708   0.117623  -1.740 0.081840 .
## PlatformPSV        -0.665467   0.138827  -4.793 1.67e-06 ***
## PlatformWii         0.412782   0.113069   3.651 0.000264 ***
## PlatformWiiU       -0.155960   0.151619  -1.029 0.303691
## PlatformX360        0.137064   0.105059   1.305 0.192058
## PlatformXB         -0.592264   0.123708  -4.788 1.72e-06 ***
## PlatformXOne        0.060666   0.137367   0.442 0.658769
## Year_of_Release1997 -0.726793   0.554212  -1.311 0.189767
## Year_of_Release1998 -1.077743   0.510138  -2.113 0.034667 *
## Year_of_Release1999 -0.949456   0.503195  -1.887 0.059223 .
## Year_of_Release2000 -1.236553   0.472280  -2.618 0.008858 **
## Year_of_Release2001 -1.223498   0.469827  -2.604 0.009231 **
## Year_of_Release2002 -1.409091   0.471339  -2.990 0.002804 **
## Year_of_Release2003 -1.515330   0.471162  -3.216 0.001305 **
## Year_of_Release2004 -1.341904   0.471065  -2.849 0.004404 **
## Year_of_Release2005 -1.528017   0.470406  -3.248 0.001167 **
## Year_of_Release2006 -1.629066   0.470786  -3.460 0.000543 ***
## Year_of_Release2007 -1.435650   0.471085  -3.048 0.002316 **
## Year_of_Release2008 -1.281816   0.471345  -2.719 0.006555 **
## Year_of_Release2009 -1.388055   0.471371  -2.945 0.003244 **
## Year_of_Release2010 -1.097769   0.472211  -2.325 0.020115 *
## Year_of_Release2011 -1.165833   0.472206  -2.469 0.013577 *
## Year_of_Release2012 -1.281404   0.473680  -2.705 0.006843 **
## Year_of_Release2013 -1.172757   0.474853  -2.470 0.013546 *
## Year_of_Release2014 -1.470740   0.476074  -3.089 0.002014 **
## Year_of_Release2015 -1.636432   0.478963  -3.417 0.000638 ***
## Year_of_Release2016 -2.348183   0.479449  -4.898 9.92e-07 ***
## GenreAdventure     -0.557443   0.076865  -7.252 4.56e-13 ***
## GenreFighting      -0.045047   0.065896  -0.684 0.494246
## GenreMisc           0.096740   0.066005   1.466 0.142791
## GenrePlatform       0.007374   0.066434   0.111 0.911622
## GenrePuzzle        -0.648282   0.111817  -5.798 7.03e-09 ***
## GenreRacing        -0.128074   0.058928  -2.173 0.029784 *
## GenreRole-Playing  -0.157703   0.051807  -3.044 0.002343 **
## GenreShooter        0.007235   0.048960   0.148 0.882520
## GenreSimulation     0.109498   0.073574   1.488 0.136728
## GenreSports        -0.203707   0.055577  -3.665 0.000249 ***
## GenreStrategy      -0.576644   0.077573  -7.434 1.19e-13 ***
## Critic_Score        0.049864   0.001311  38.046  < 2e-16 ***
```

```
## User_Score             -0.104630   0.012763  -8.198 2.91e-16 ***
## RatingE10+             -0.203597   0.049220  -4.136 3.57e-05 ***
## RatingM                0.044288   0.053417   0.829 0.407071
## RatingT               -0.200275   0.043182  -4.638 3.59e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.118 on 6766 degrees of freedom
## Multiple R-squared:  0.3653, Adjusted R-squared:  0.3604
## F-statistic: 74.87 on 52 and 6766 DF,  p-value: < 2.2e-16
```

Introduction

fasdfas

Just temporary

A multiple regression model follows $Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + ... + \beta_p * X_p + \varepsilon$ where X_p represents our predictors and β_p represents the coefficients related to the predictors in the linear model. The goal of the multiple linear regression model is to estimate each β_p coefficient.

m ~$\sqrt{p}$

Global_Sales $= \beta_0 +$ Platform $* \beta_1 +$ Year_of_Release $* \beta_2 +$ Genre $* \beta_3 +$ Critic_Score $* \beta4 +$ User_Score $* \beta_5 +$ Rating $* \beta_6 + \varepsilon$.

Linear Model

asdfasd

Random Forest Model

asdfasd

Conclusion

asdfasdf