William Sanders
Project Luther

I am a film producer trying to find the best ways to optimize the movie net promoter score (NPS) for the next film for Wayward Film Producer Company. With movie data from 1980 – 2005, I want to discover what the ideal optimization for budget, genre, rating, and any other criteria to get the highest movie NPS. Looking at the preliminary analysis, I was able to return to find this data:

```
listed = []
for i in xrange(20):
    X = nonull_combined_data[cols]
    y = nonull_combined_data['movie_nps']
    X_train, X_test, y_train, y_test = train_test_split(X,y,test_size
    linear = linearR(X_train, X_test, y_train, y_test)
    ridge = ridgeCVs(X_train, X_test, y_train, y_test)
    lineartrain = linear_train(X_train, X_test, y_train, y_test)
    listed.append((linear,ridge, lineartrain))
listed
```

```
[(0.35854221466659481, 0.4006317494905054, 0.35854221466659481),
 (0.24049028502388448, 0.30240481150338672, 0.24049028502388448),
 (0.20459133459482126, 0.2497091695117524, 0.20459133459482126),
 (0.36191173904187524, 0.40934138830372713, 0.36191173904187524),
 (0.23238562118915806, 0.30967388274514762, 0.23238562118915806),
 (-0.048004186814684102, 0.29087705654603602, -0.04800418681468410
2),
 (0.18698433788639057, 0.36845409492578335, 0.18698433788639057),
 (0.2854157250427356, 0.30846886235245241, 0.2854157250427356),
 (0.32924150271485908, 0.3805606704766602, 0.32924150271485908),
 (0.29823164176989769, 0.35547106693536301, 0.29823164176989769),
 (0.36291812387246902, 0.3930064595383933, 0.36291812387246902),
 (0.24993180829516604, 0.32008450859830551, 0.24993180829516604),
 (0.38326149100927764, 0.41068371127829462, 0.38326149100927764),
 (0.16728571295657244, 0.29510154126417032, 0.16728571295657244),
 (0.3269539673675016, 0.3768699392167234, 0.3269539673675016),
 (0.27722625407321166, 0.29795936897708808, 0.27722625407321166),
 (0.30428747975477222, 0.36039498926528168, 0.30428747975477222),
 (0.32135330391974626, 0.34148694809475072, 0.32135330391974626),
 (0.090445821805830384, 0.30080708337215722, 0.090445821805830384),
 (0.37491962771760662, 0.3933855298246261, 0.37491962771760662)]
```

With this preliminary analysis, we will try to find another model that optimizes the highest score. We should be able to make a choice of specific movie variables and optimize for highest NPS.