



FACTORS IMPACTING POST-ACUTE RE-ADMISSIONS

PREPARED BY HANNAH MCQUADE, MINH HA, AND LINDA BI



AGENDA

- Motivation
- Methodology
 - Data Source
 - Variable Selection
 - Model
- Conclusion
- Moving Forward

MOTIVATION

- 1/3 of healthcare spending is wasted
- One way to reduce waste is to decrease hospital readmissions
- Nursing homes are a major source of hospital admissions because they serve an elderly population who often have a complex mix of chronic and acute illnesses.
- Increasingly, policy and insurance companies are financially penalizing readmissions
- How do we decrease readmissions from nursing homes?

DATASET

■ CMS Nursing Home Compare

MSR_521_SCORE_ADJUSTED	CYCLE_2_NUMBER_OF_STANDARD_HEALTH_DEFICIENCIES	CNA_HOURS
MSR_522_SCORE_ADJUSTED	CYCLE_2_NUMBER_OF_COMPLAINT_HEALTH_DEFICIENCIES	LPN_HOURS
MSR_523_SCORE_ADJUSTED	CYCLE_2_HEALTH_DEFICIENCY_SCORE	RN_HOURS
PROVIDER_STATE	CYCLE_2_NUMBER_OF_HEALTH_REVISITS	TOTAL_NURSE_HOURS
OWNERSHIP_TYPE	CYCLE_2_HEALTH_REVISIT_SCORE	HEALTH_INSPECTION_RATING
NUMBER_OF_CERTIFIED_BEDS	CYCLE_2_TOTAL_HEALTH_SCORE	STAFFING_RATING_y
NUMBER_OF_RESIDENTS_IN_CERTIFIED_BEDS	CYCLE_3_TOTAL_NUMBER_OF_HEALTH_DEFICIENCIES	RN_STAFFING_RATING_y
TOTAL_WEIGHTED_HEALTH_SURVEY_SCORE	CYCLE_3_NUMBER_OF_STANDARD_HEALTH_DEFICIENCIES	
NUMBER_OF_FACILITY_REPORTED_INCIDENTS	CYCLE_3_NUMBER_OF_COMPLAINT_HEALTH_DEFICIENCIES	
NUMBER_OF_SUBSTANTIATED_COMPLAINTS	CYCLE_3_HEALTH_DEFICIENCY_SCORE	
PROVIDER_RESIDES_IN_HOSPITAL	CYCLE_3_NUMBER_OF_HEALTH_REVISITS	
CYCLE_1_TOTAL_NUMBER_OF_HEALTH_DEFICIENCIES	CYCLE_3_HEALTH_REVISIT_SCORE	
CYCLE_1_NUMBER_OF_STANDARD_HEALTH_DEFICIENCIES	CYCLE_3_TOTAL_HEALTH_SCORE	
CYCLE_1_NUMBER_OF_COMPLAINT_HEALTH_DEFICIENCIES	TOTAL_HEALTH_DEFICIENCY_SCORE	
CYCLE_1_HEALTH_DEFICIENCY_SCORE	TOTAL_FINE_AMT	
CYCLE_1_NUMBER_OF_HEALTH_REVISITS	NUM_PAYMENT_DENIALS	
CYCLE_1_HEALTH_REVISIT_SCORE	NUM_OF_PENALTIES	
CYCLE_1_TOTAL_HEALTH_SCORE	STAFFING_RATING_x	
CYCLE_2_TOTAL_NUMBER_OF_HEALTH_DEFICIENCIES	RN_STAFFING_RATING_x	

■ Engineered features: occupancy

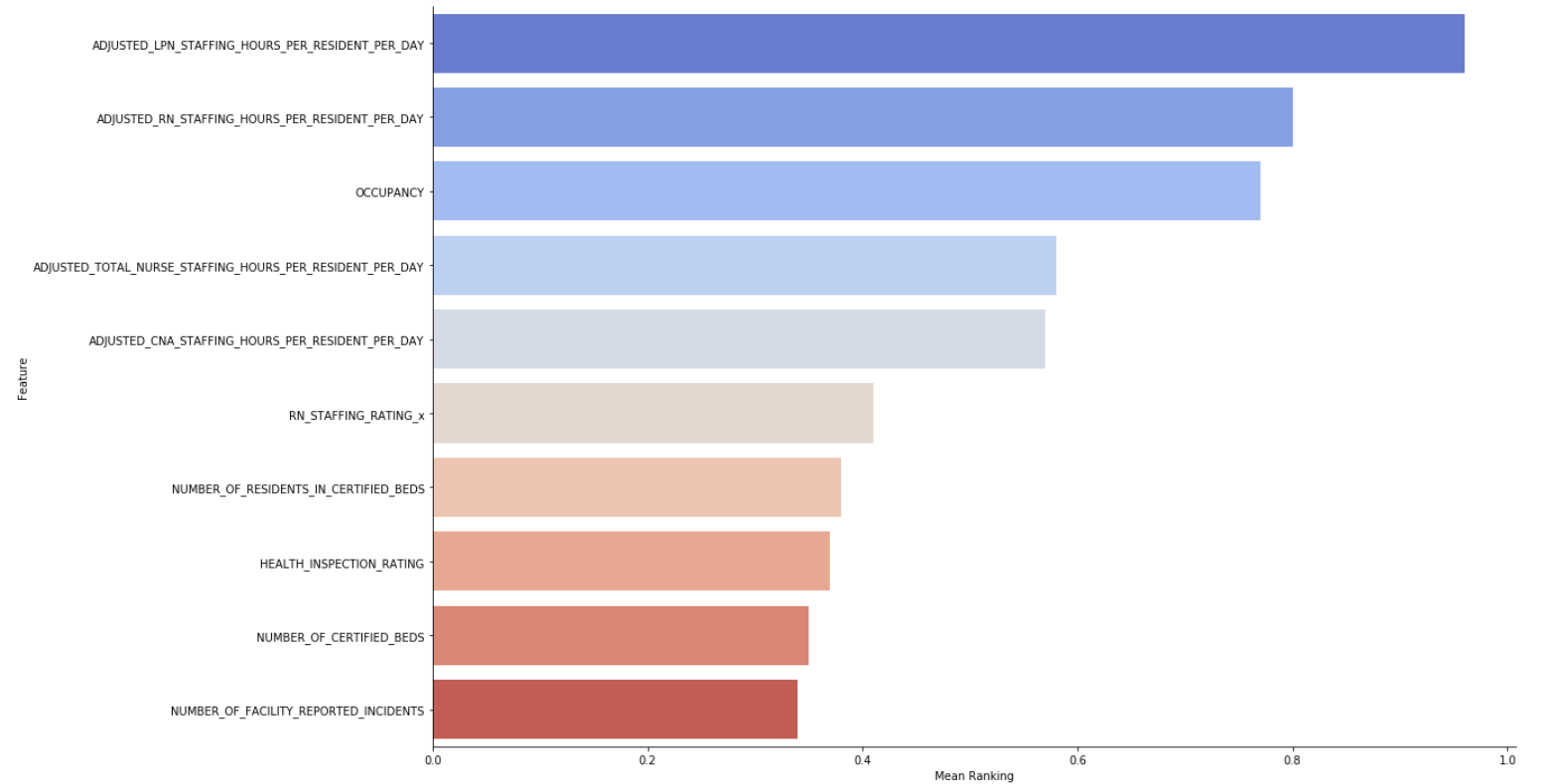
DATA EXTRACTION AND TRANSFORMATION

- Did not impute
- Transformations/pivots
- Convert response variables from percentage score to logits

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

NURSING HOME READMISSION ANALYSIS

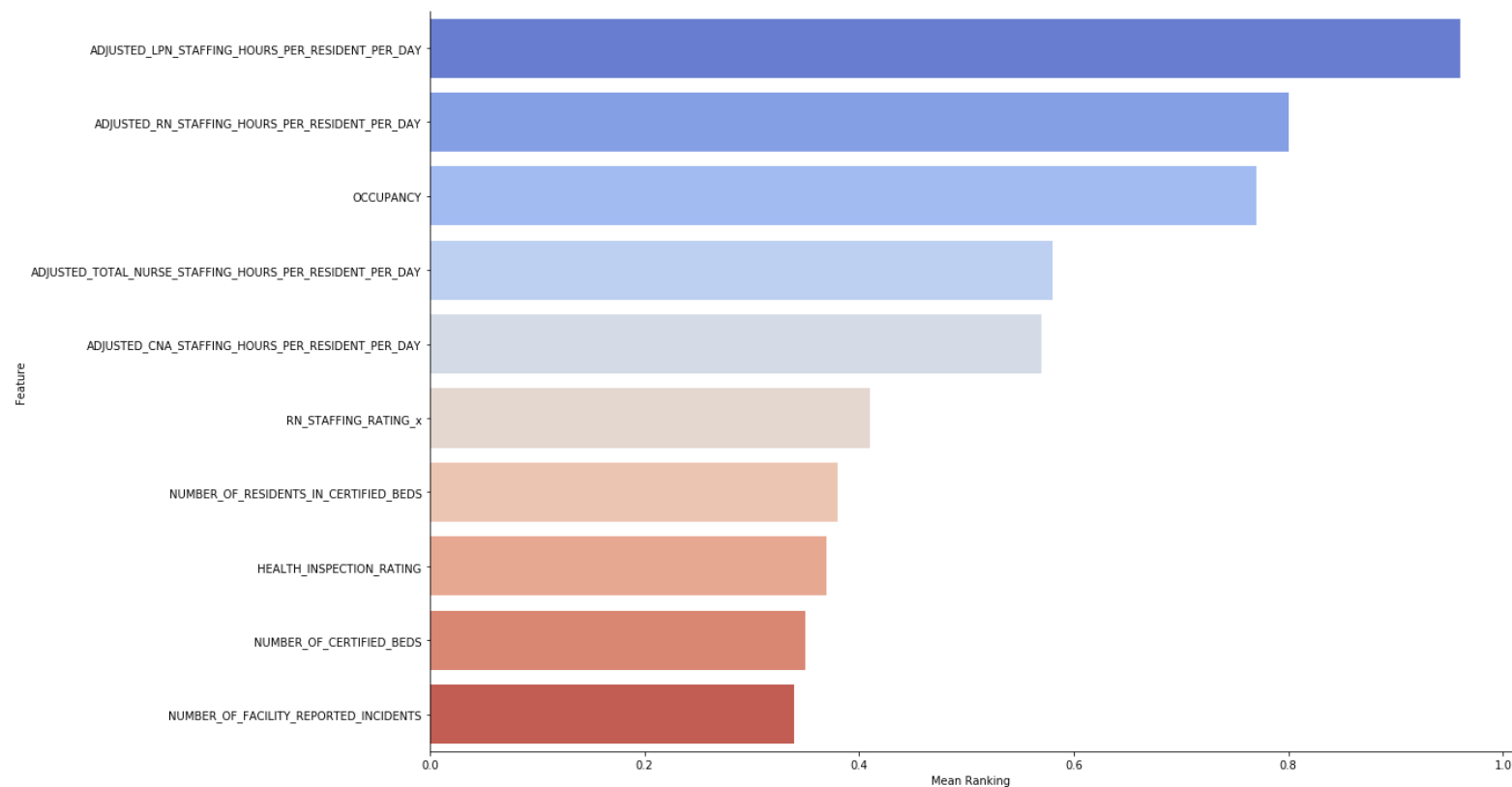
- Variable Selection Methods
 - Random Lasso
 - Recursive Feature Elimination
 - Random Forest Feature Ranking
- Mean Ranking Score



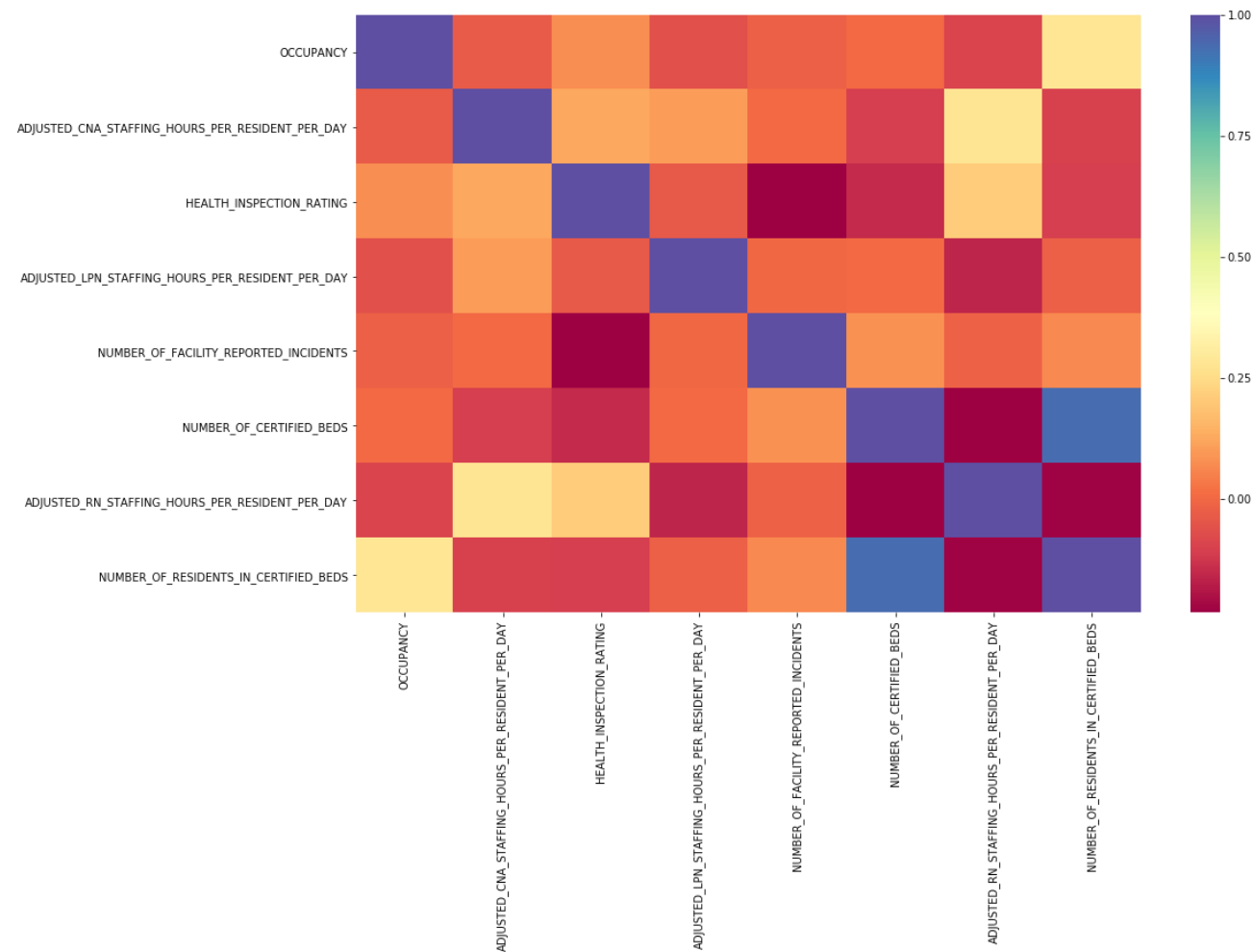


VARIABLE SELECTION

- Variable Selection Methods
 - Random Lasso
 - Recursive Feature Elimination
 - Random Forest feature Ranking
- Mean Ranking Score



VARIABLE SELECTION



MODELING – INITIAL LINEAR REGRESSION

```
call:
lm(formula = MSR_521_SCORE_ADJUSTED_LOGITS ~ NUMBER_OF_SUBSTANTIATED_COMPLAINTS +
    TOTAL_HEALTH_DEFICIENCY_SCORE + NUMBER_OF_CERTIFIED_BEDS +
    OCCUPANCY + ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY +
    ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY + ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY,
    data = train)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-1.58908 -0.13954  0.03144  0.18147  0.56612
```

```
Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                   -5.8298252   0.1563230  -37.293   < 2e-16 ***
NUMBER_OF_SUBSTANTIATED_COMPLAINTS    0.0035770   0.0017776    2.012   0.04502  *
TOTAL_HEALTH_DEFICIENCY_SCORE        -0.0003889   0.0001437   -2.706   0.00718  **
NUMBER_OF_CERTIFIED_BEDS              0.0005441   0.0003397    1.602   0.11019
OCCUPANCY                        -0.3601141   0.1389513   -2.592   0.00999  **
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY  0.0991845   0.0370222    2.679   0.00776  **
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY -0.0547722   0.0249581   -2.195   0.02891  *
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY -0.1212386   0.0704609   -1.721   0.08627  .
---
```

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

```
Residual standard error: 0.2696 on 323 degrees of freedom
Multiple R-squared:  0.1349,    Adjusted R-squared:  0.1161
F-statistic: 7.193 on 7 and 323 DF,  p-value: 5.22e-08
```

MODELING – LINEAR REGRESSION AFTER REMOVING HIGH LEVERAGE POINTS

```
call:
lm(formula = MSR_521_SCORE_ADJUSTED_LOGITS ~ NUMBER_OF_SUBSTANTIATED_COMPLAINTS +
    TOTAL_HEALTH_DEFICIENCY_SCORE + NUMBER_OF_CERTIFIED_BEDS +
    OCCUPANCY + ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY +
    ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY + ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY,
    data = train[lev <= 3 * 0.02416918, ])
```

Residuals:

Min	1Q	Median	3Q	Max
-1.56319	-0.13585	0.03082	0.17143	0.56312

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-5.9047283	0.1620230	-36.444	< 2e-16	***
NUMBER_OF_SUBSTANTIATED_COMPLAINTS	0.0033301	0.0018416	1.808	0.071526	.
TOTAL_HEALTH_DEFICIENCY_SCORE	-0.0003732	0.0001522	-2.452	0.014765	*
NUMBER_OF_CERTIFIED_BEDS	0.0005286	0.0003455	1.530	0.127036	
OCCUPANCY	-0.3245172	0.1399211	-2.319	0.021017	*
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.1494581	0.0410130	3.644	0.000314	***
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.0670363	0.0278247	-2.409	0.016557	*
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.0807446	0.0759394	-1.063	0.288468	

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

Residual standard error: 0.2686 on 316 degrees of freedom

Multiple R-squared: 0.1482, Adjusted R-squared: 0.1294

F-statistic: 7.857 on 7 and 316 DF, p-value: 8.789e-09

MODELING – BETA REGRESSION

- Used for percentage or ratio predictors where range is between 0-1

Coefficients (mean model with logit link):

	Estimate	Std. Error	z value
(Intercept)	-1.0208484	0.1542809	-6.617
NUMBER_OF_SUBSTANTIATED_COMPLAINTS	0.0043279	0.0016622	2.604
TOTAL_HEALTH_DEFICIENCY_SCORE	-0.0003437	0.0001462	-2.350
NUMBER_OF_CERTIFIED_BEDS	0.0004433	0.0003422	1.295
OCCUPANCY	-0.2897503	0.1322944	-2.190
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.0632081	0.0241721	-2.615
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.1141800	0.0362196	3.152
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.1212277	0.0715762	-1.694

Pr(>|z|)

(Intercept)	3.67e-11 ***
NUMBER_OF_SUBSTANTIATED_COMPLAINTS	0.00922 **
TOTAL_HEALTH_DEFICIENCY_SCORE	0.01875 *
NUMBER_OF_CERTIFIED_BEDS	0.19525
OCCUPANCY	0.02851 *
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.00893 **
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.00162 **
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.09032 .

Phi coefficients (precision model with identity link):

	Estimate	Std. Error	z value	Pr(> z)
(phi)	65.450	4.528	14.45	<2e-16 ***

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

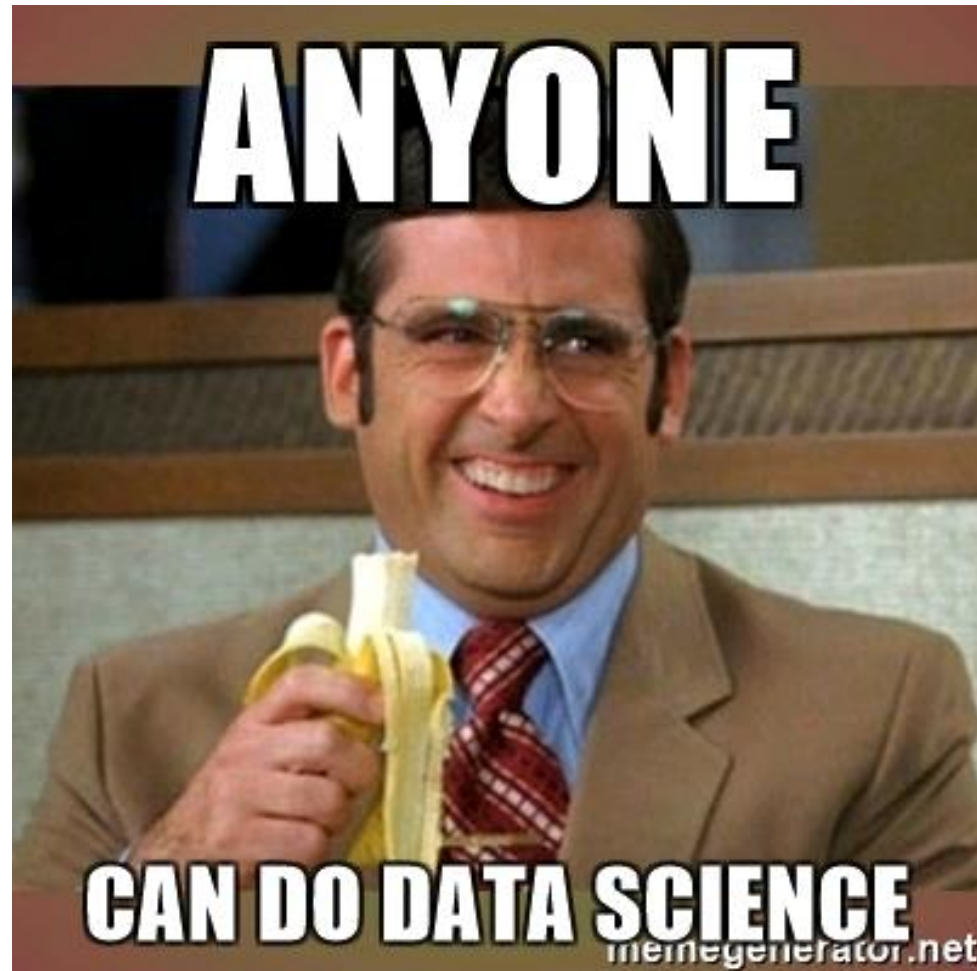
Type of estimator: ML (maximum likelihood)

Log-likelihood: 662.6 on 9 Df

Pseudo R-squared: 0.1232

Number of iterations: 19 (BFGS) + 2 (Fisher scoring)

CONCLUSION



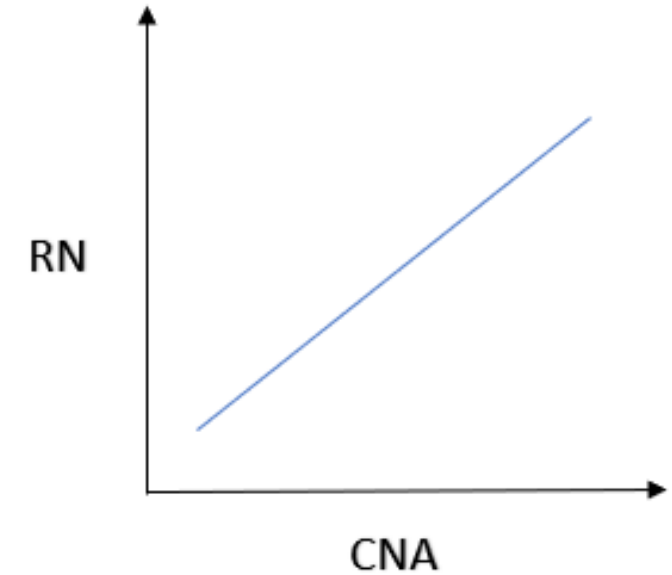
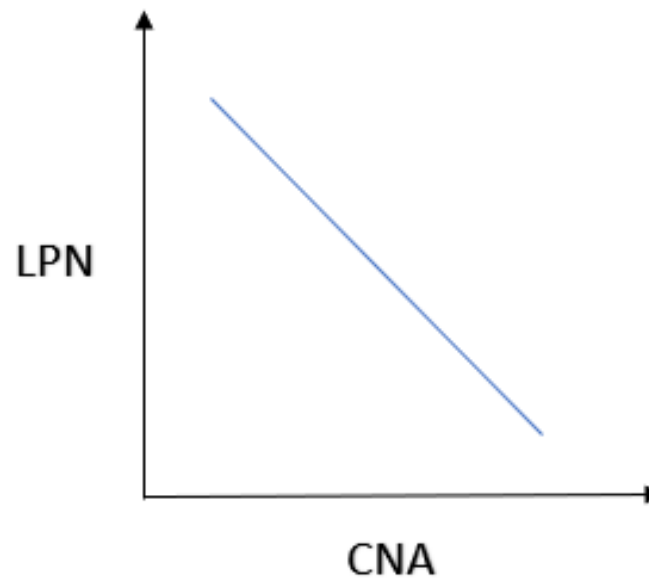
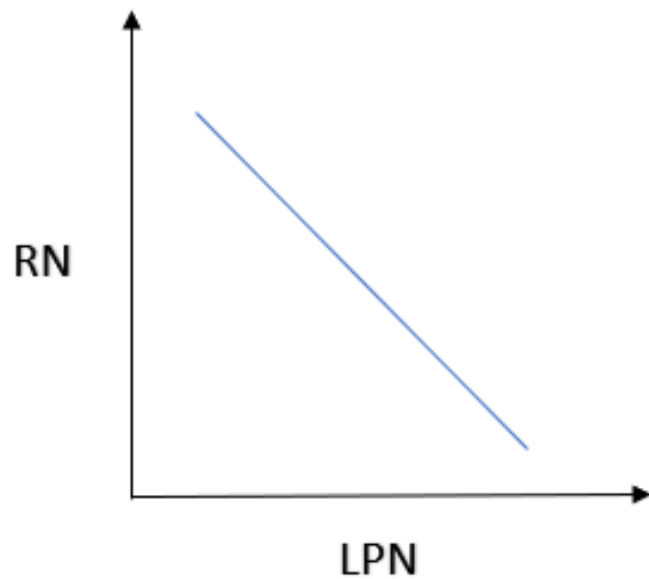
CONCLUSION

- Variable selection explained ~13% of the variation
- MSE similar between training and test, model does not over or under fit
 - Training MSE: 0.0703
 - Test MSE: 0.06669
- Using the beta regression, inferences:

Variable	Variable action	Response action
Number of complaints	increase	increase
Health Deficiency Score	increase	decrease
Occupancy	increase	decrease
CNA staff hours/resident/day	increase	decrease
LPN staff hours/resident/day	increase	increase
RN staff hours/resident/day	increase	decrease

CONCLUSION

- LPN/RN/CNA inferences
 - Increased LPN hours per resident increase the re-admittance rate
 - Increased LPN hours per resident are negatively correlated with RN and CNA hours



MOVING FORWARD

- Several variables explain ~13% of variation in readmissions from nursing homes back to hospitals
- Other potential factors:
 - Admission accepted 24/7/365
 - Participation in first fill prescription program
 - EHR access
 - Mandatory case review
 - Communication with PCP
 - Implement care pathways
 - Transition planning
 - Phone calls made to patients who transition home
 - Scheduling medical appointments prior to transition home
- Interaction term with RN/LPN/CPN hours



We would now be happy to avoid your questions

YOU ARE A DATA SCIENTIST

YOU ARE A DATA SCIENTIST

YOU ARE A DATA SCIENTIST

YOU ALL ARE DATA SCIENTISTS!



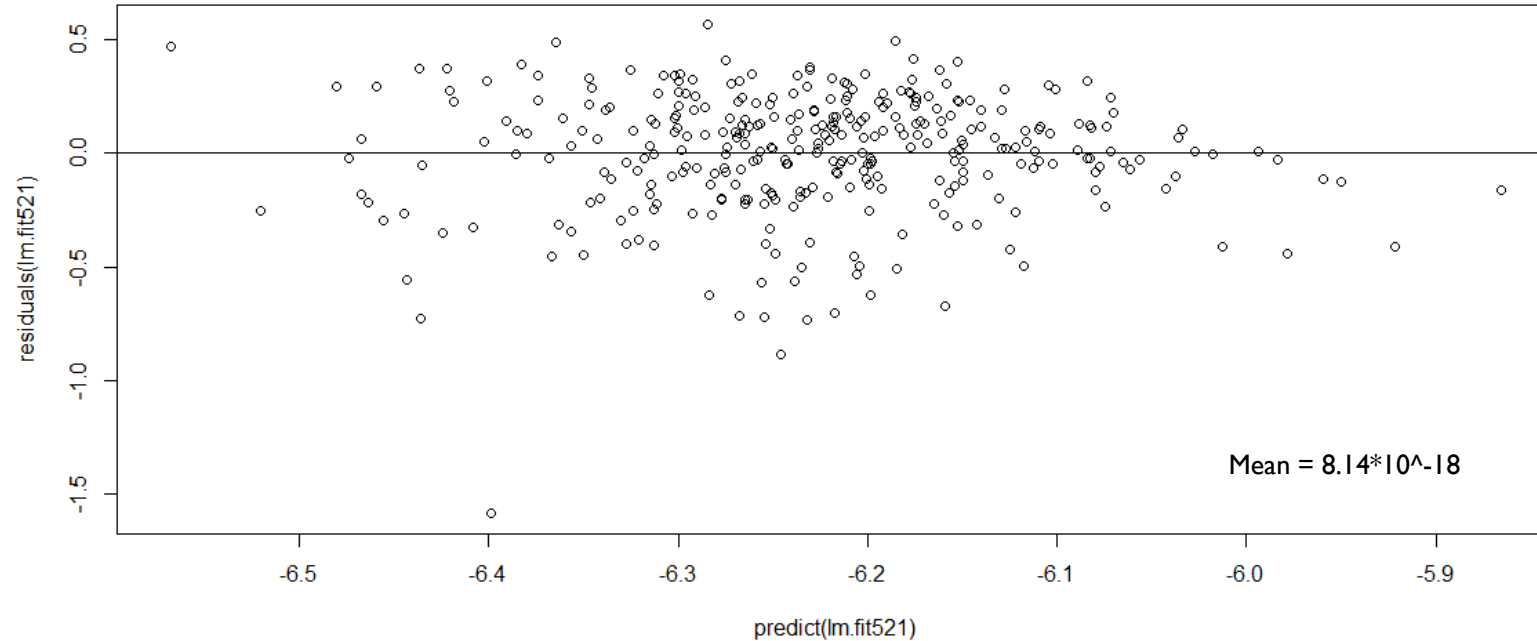
APPENDIX

- VIF for Linear Model

```
> vif(lm.fit521)
      NUMBER_OF_SUBSTANTIATED_COMPLAINTS
                        2.503522
      TOTAL_HEALTH_DEFICIENCY_SCORE
                        2.297412
      NUMBER_OF_CERTIFIED_BEDS
                        1.187816
      OCCUPANCY
                        1.141149
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY
                        1.210729
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY
                        1.201973
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY
                        1.509753
```

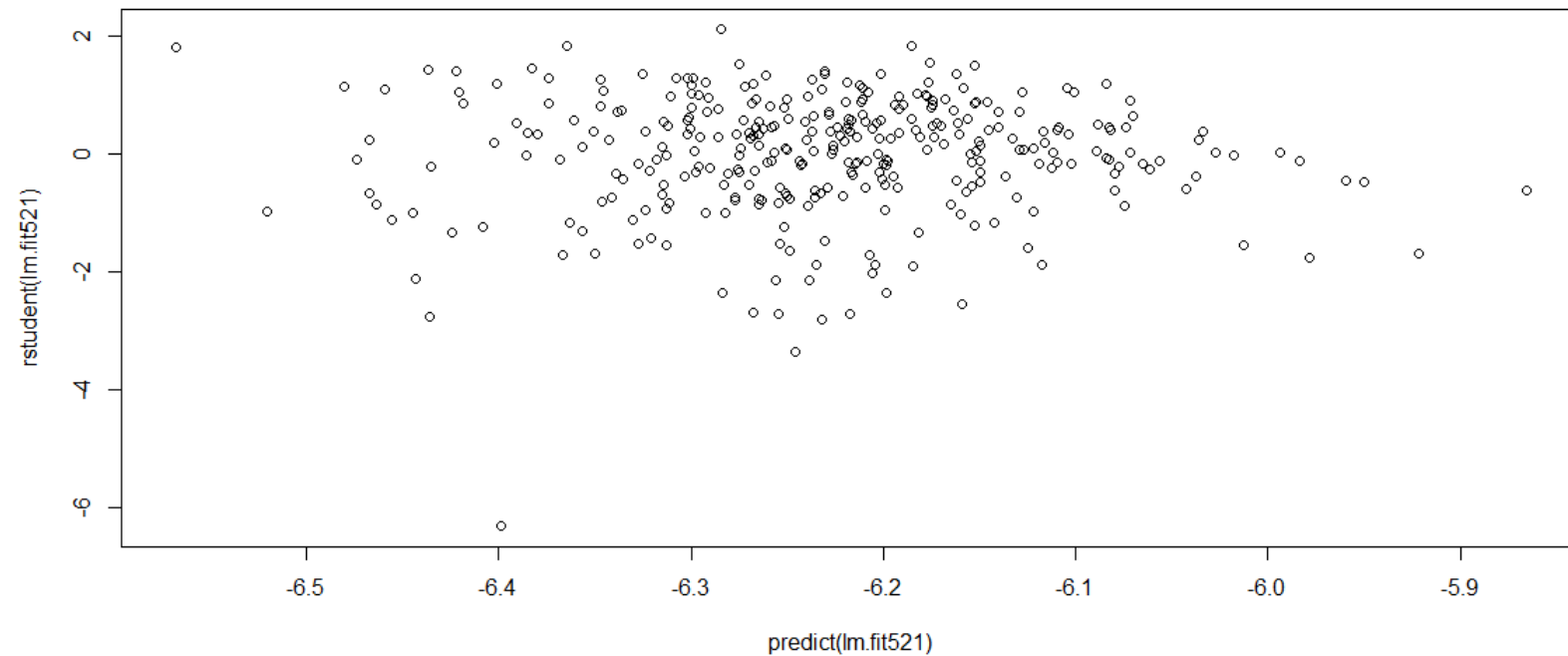
APPENDIX

- Residuals Plot for Linear Model
 - Does not appear to violate assumptions of linear model (mean of zero, constant variance, no correlation of error terms)



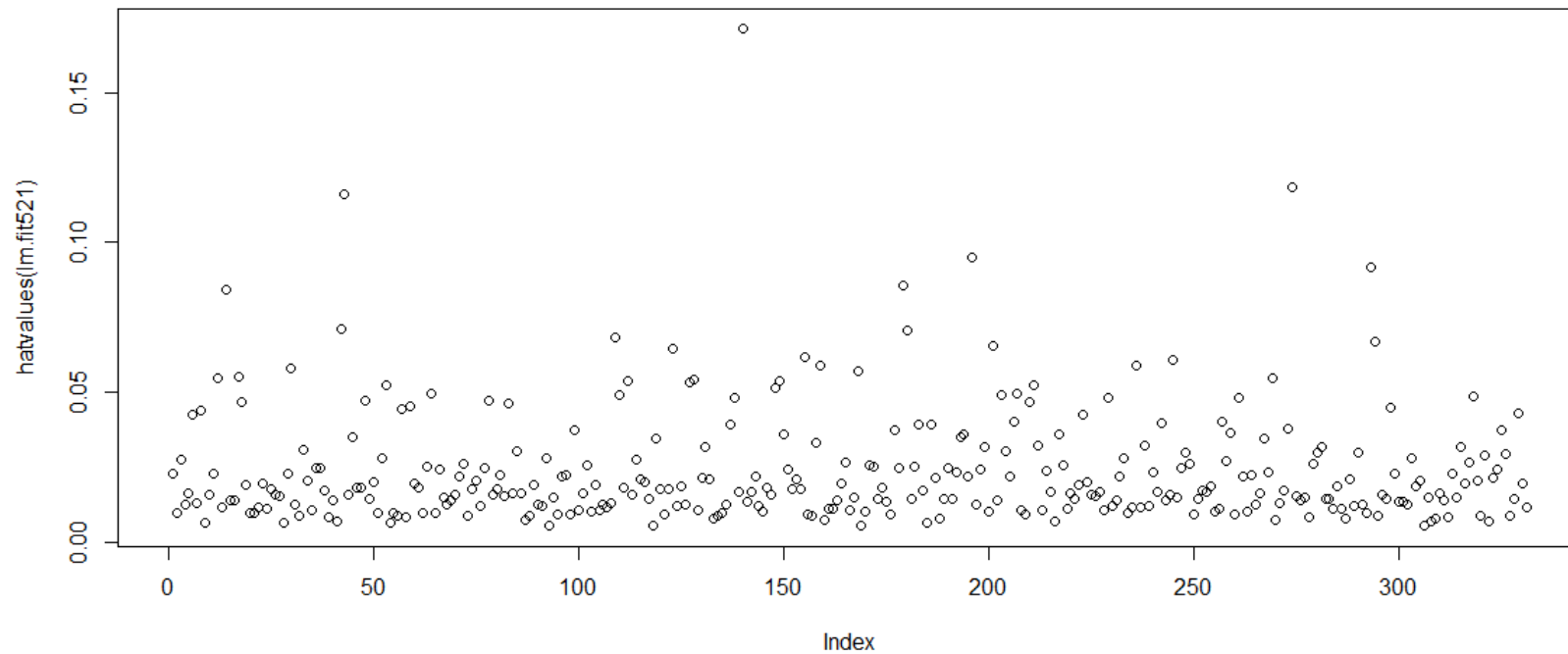
APPENDIX

- Studentized Residuals Plot for Linear Model
 - Does not appear to be outliers



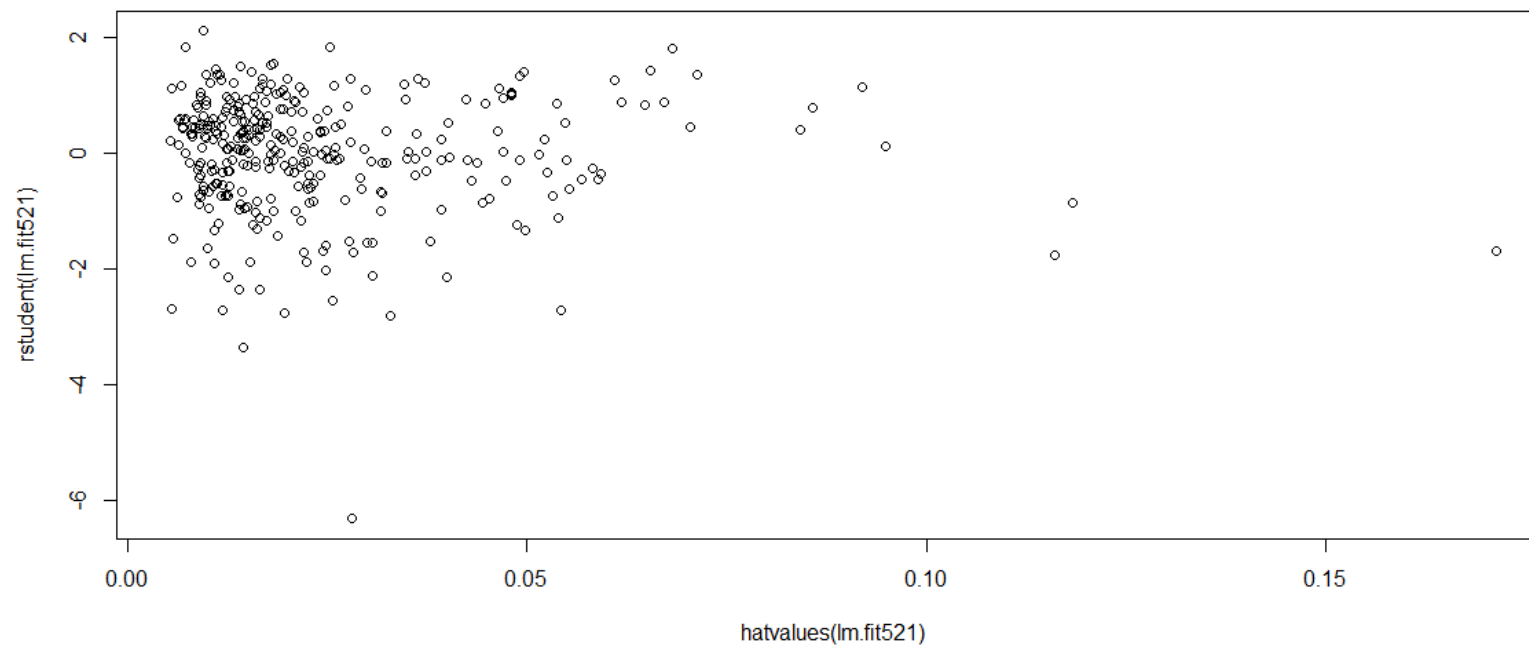
APPENDIX

- Leverage
 - 7 high leverage point that should be removed



APPENDIX

- Looking for high leverage points part 2
 - 7 high leverage point that should be removed



APPENDIX

■ Linear regression model results with interaction term

Residuals:
 Min 1Q Median 3Q Max
 -1.55025 -0.14275 0.03368 0.17323 0.58128

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-5.0126601	0.4270595	-11.738	<2e-16 ***
NUMBER_OF_SUBSTANTIATED_COMPLAINTS	0.0028609	0.0018018	1.588	0.1133
TOTAL_HEALTH_DEFICIENCY_SCORE	-0.0003709	0.0001443	-2.570	0.0106 *
NUMBER_OF_CERTIFIED_BEDS	0.0006946	0.0003431	2.024	0.0438 *
OCCUPANCY	-0.3598604	0.1392874	-2.584	0.0102 *
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.3511572	0.1478560	-2.375	0.0181 *
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.2402598	0.3406889	-0.705	0.4812
ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-1.0005100	0.4325095	-2.313	0.0213 *
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY:ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.1075709	0.1269589	0.847	0.3975
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY:ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.2935115	0.1396608	2.102	0.0364 *
ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY:ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	0.2808500	0.3482840	0.806	0.4206
ADJUSTED_CNA_STAFFING_HOURS_PER_RESIDENT_PER_DAY:ADJUSTED_LPN_STAFFING_HOURS_PER_RESIDENT_PER_DAY:ADJUSTED_RN_STAFFING_HOURS_PER_RESIDENT_PER_DAY	-0.0698205	0.1227321	-0.569	0.5698

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

Residual standard error: 0.2674 on 319 degrees of freedom
 Multiple R-squared: 0.1595, Adjusted R-squared: 0.1305
 F-statistic: 5.504 on 11 and 319 DF, p-value: 4.658e-08

	Net Effect of Increase by One on Logit
CNA	-0.0198953
LPN	0.0783406
RN	-0.495969