

**Data Mining Team Project**  
Nursing Home Performance Drivers – Care Quality  
Prepared for Professor & Peers  
12.8.09

Erica Eisenhart  
Alon Gotesman  
Audra Johnson  
Ben Meadema  
Shalin Saini

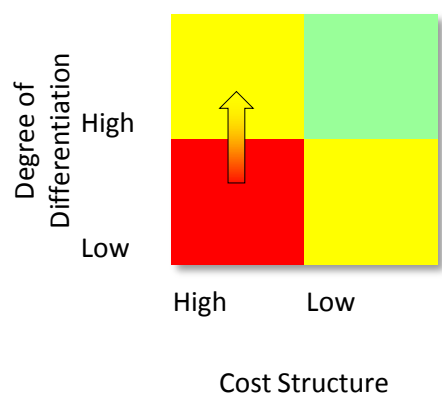
# Problem

Explanation

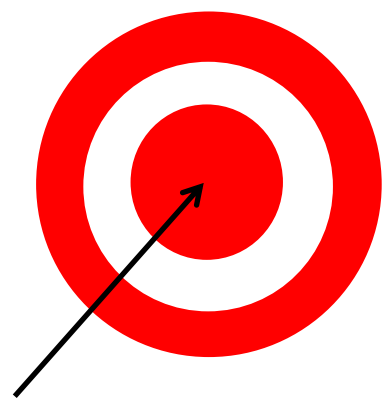
What organizational factors impact the quality of care in nursing homes; specifically which factors are associated with high quality care nursing homes?

Benefits obtained by analyzing this question:

Institution Differentiation



Consumer Decision Process



Institution Service Improvements

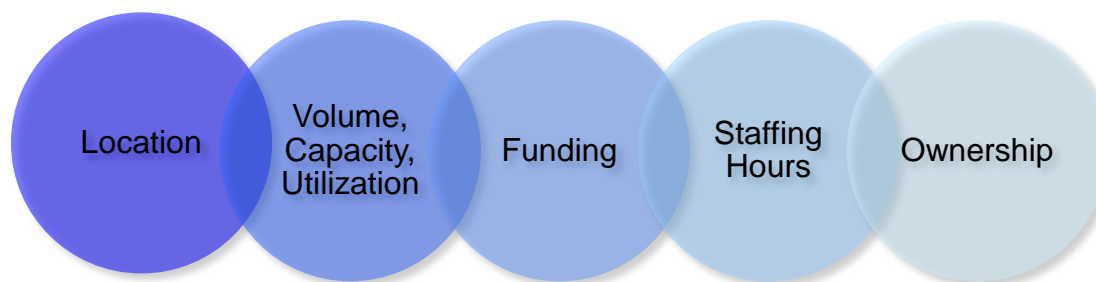


# Acquiring Data

Source & Preparation

Data Source: Centers for Medicare & Medicaid (CMS)

<http://www.medicare.gov/NHCompare/static/tabSI.asp?language=English&activeTab=3&subTab=16&version=default>



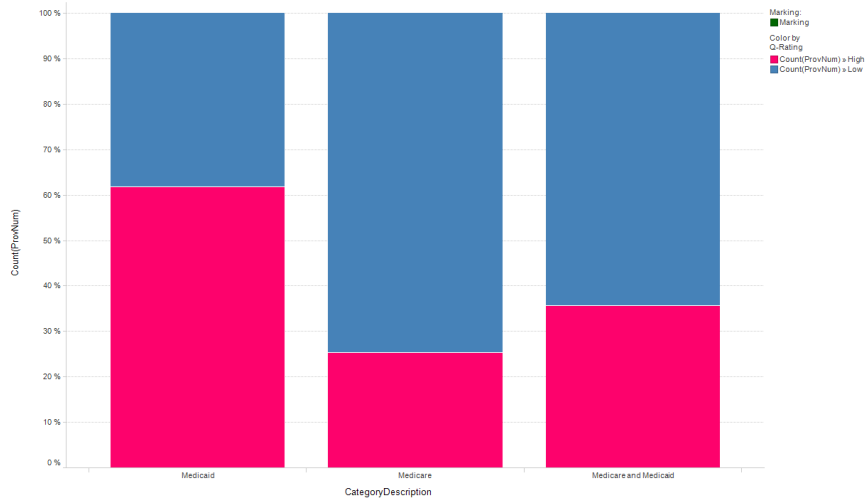
Preparation Process:



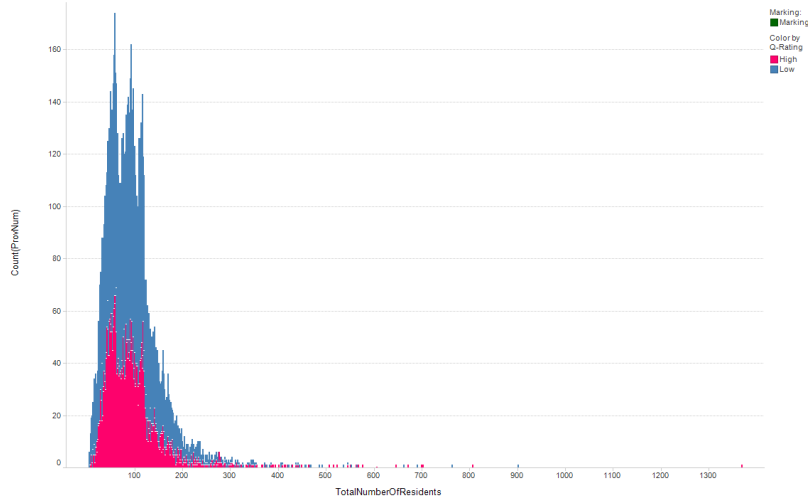
# Data Set Visualizations

## Original Data

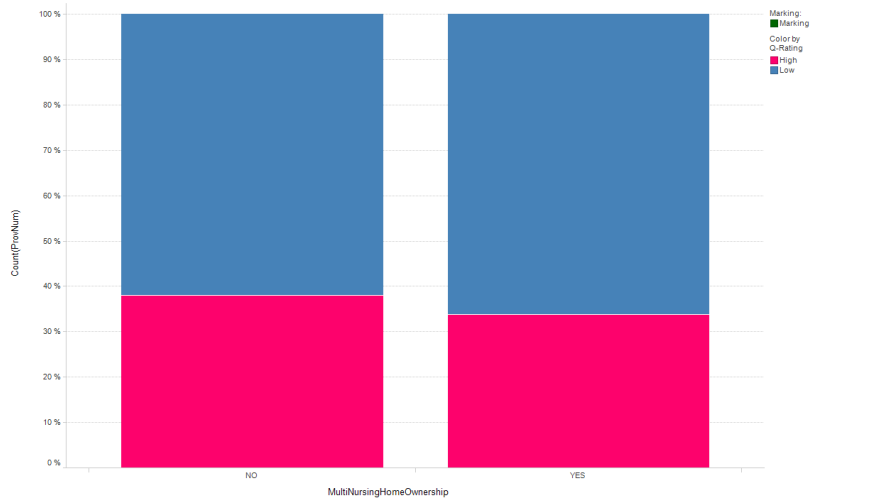
Bar Chart - Category Type



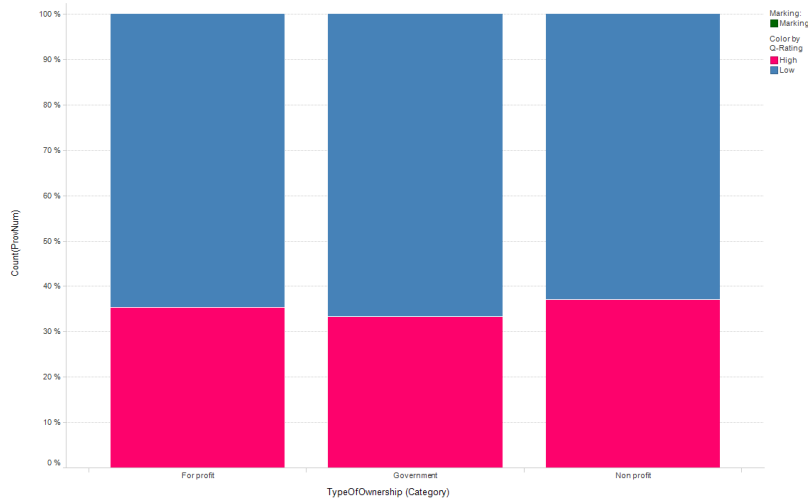
Bar Chart - Total Number of Residents



Bar Chart - MultiNursing Home Ownership



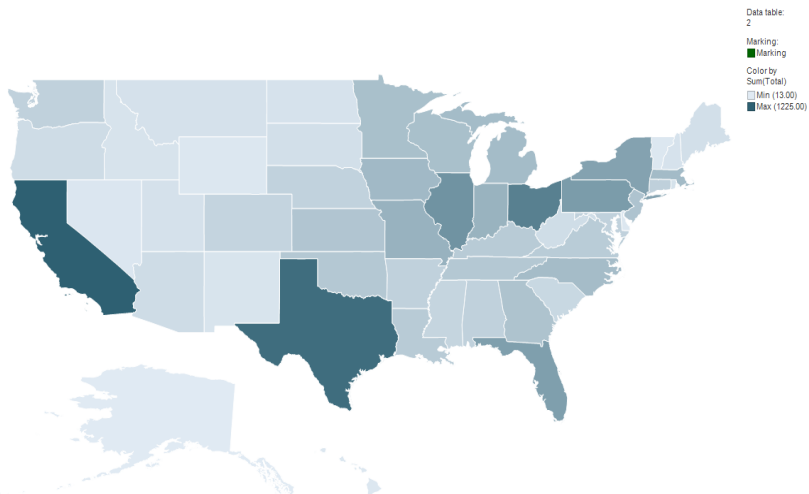
Bar Chart - Entity Type



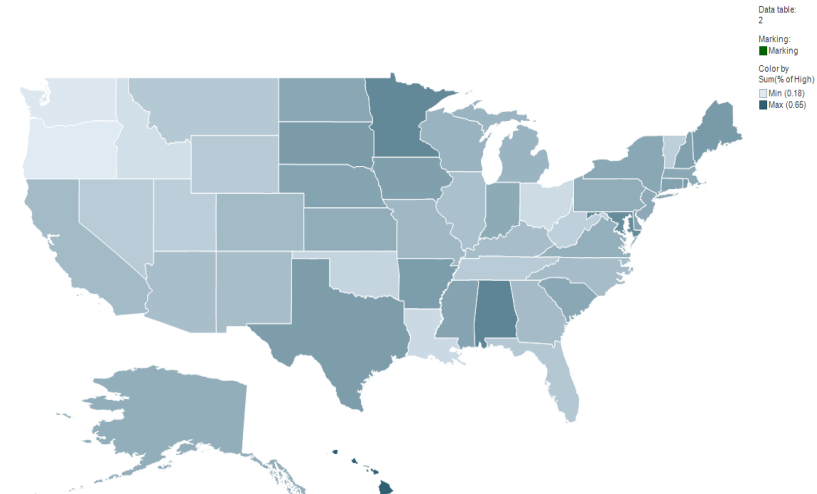
# Data Set Visualizations

## External Data

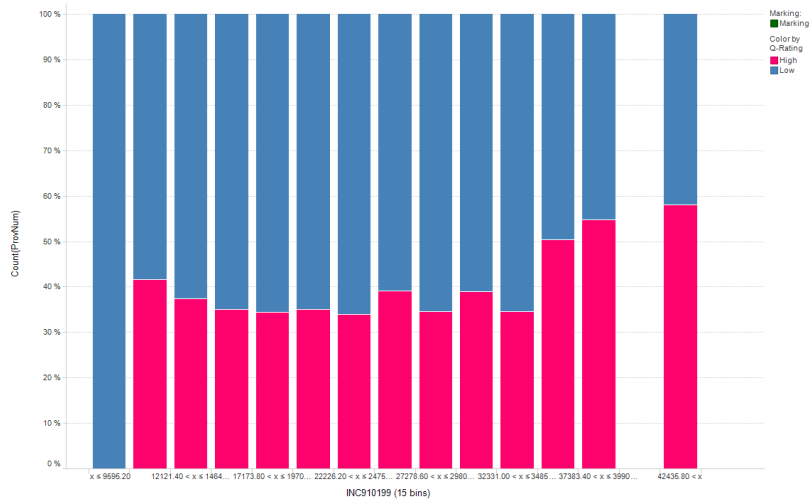
Map Chart - Provider Count



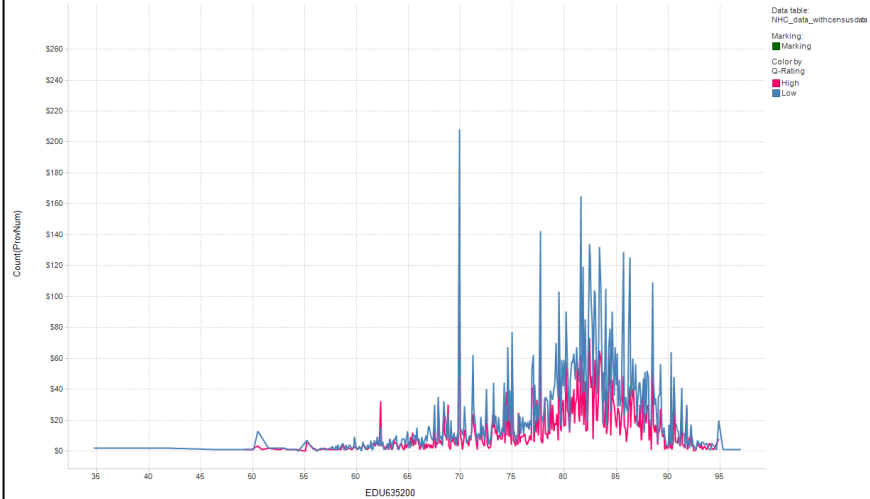
Map Chart - % of High Quality



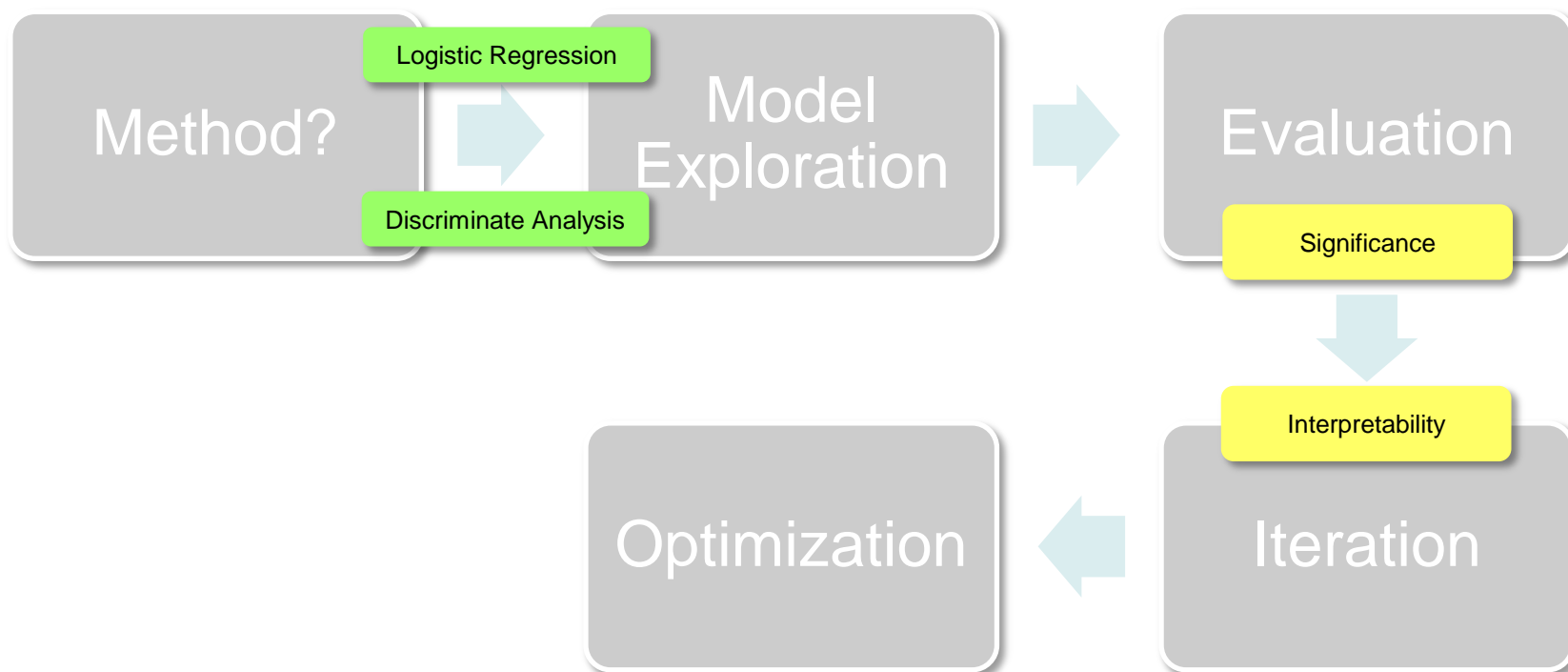
Bar Chart - County Income Level



Line Chart - % of 25+ as High School Graduates



# Modeling Process



Input Variables	Coefficient	Standard Error	P-value	Odds
Constant	4.35417128	0.35214877	0	*
Region 1 (C)	0.19545503	0.10281167	0.05728921	1.21586406
Region 2 (C)	0.01369609	0.10827228	0.89933872	1.01379037
Region 3 (C)	-0.14717098	0.09806134	0.13340592	0.86314636
Region 4 (C)	-0.1839278	0.09433191	0.05120067	0.83199584
Region 5 (C)	0.46744183	0.10930981	0.000019	1.59590638
Region 6 (C)	-0.23178081	0.10084127	0.02153495	0.79311997
Region 7 (C)	-0.21132194	0.10131818	0.03700347	0.80951339
Region 8 (C)	-0.11105476	0.12993658	0.39272588	0.89488977
Region 9 (C)	-0.52643698	0.09689617	0.00000006	0.59070593
Accepts Medicare (C)	-1.17600572	0.18138197	0	0.30850855
Accepts Medicare and Medicaid (C)	-1.07946002	0.13878389	0	0.33977896
Government Ownership (C)	0.31905618	0.10511177	0.00240217	1.37582862
Non-profit Ownership (C)	0.17079225	0.05396795	0.00155244	1.18624425
Located within a hospital (C)	-0.46381971	0.10523517	0.00001046	0.62887692
Multi-nursing home ownership (C)	-0.10479257	0.04487701	0.01953788	0.90051126
Licensed staff hours per resident per day (N)	-0.44295809	0.04342123	0	0.64213413
Percent of persons over age 25 with high school diploma (N)	-0.03787193	0.00446186	0	0.96283627
Percent of persons over age 25 with college degree (N)	0.0191048	0.00493862	0.00010953	1.01928842
Per capita income (N)	0.02599033	0.009237	0.0048972	1.02633107

# Final Model

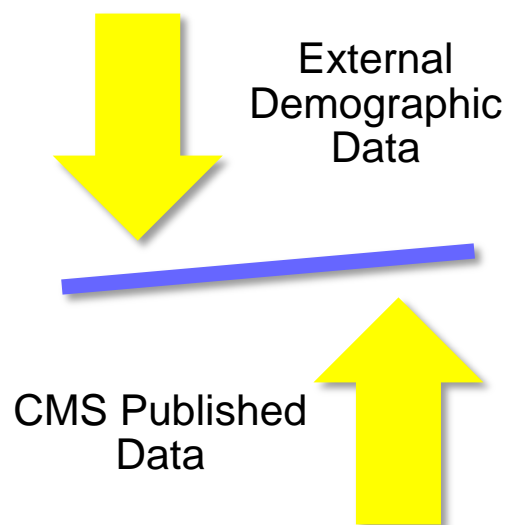
## Logistic Regression

Residual df	9980
Residual Dev.	12815.02344
% Success in training data	60.87
# Iterations used	11
Multiple R-squared	0.04269204

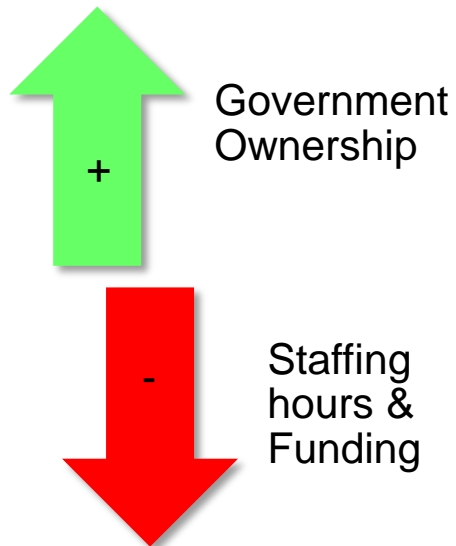
Classification Confusion Matrix				
Actual Class	Predicted Class			
	1	0		
1	5630	457		
0	3230	683		
Error Report				
Class	# Cases	# Errors	% Error	
1	6087	457	7.51	
0	3913	3230	82.55	
Overall	10000	3687	36.87	

# Model Interpretation

## Data Source Importance



## Non-Intuitive Signs



## Insignificant Variables

CNA Hours Per Res Per Day



# Recommendations

---

Data Segmentation

**Institution Type**

**Patient Type**

**Services Type**

Risk Profile

Demographics

Specialty Care

Traditional  
Care