Prediction of California Hospital Quality Ratings

Maria P. Frushicheva with mentor Rachel Wilkerson

August 16, 2016

Questions to an audience

- Do you want to know what is the best healthcare available for you?
- Which hospital has the best overall quality ratings?
- Which hospital is recomended based on particular medical condition or procedure?

Introduction

Importance:

- Using hospital quality ratings, patients are able to make a better decision in what hospital they want to be treated and where the best care is available in state of California, based on overall hospital performance or based on particular medical condition or procedure.

Question:

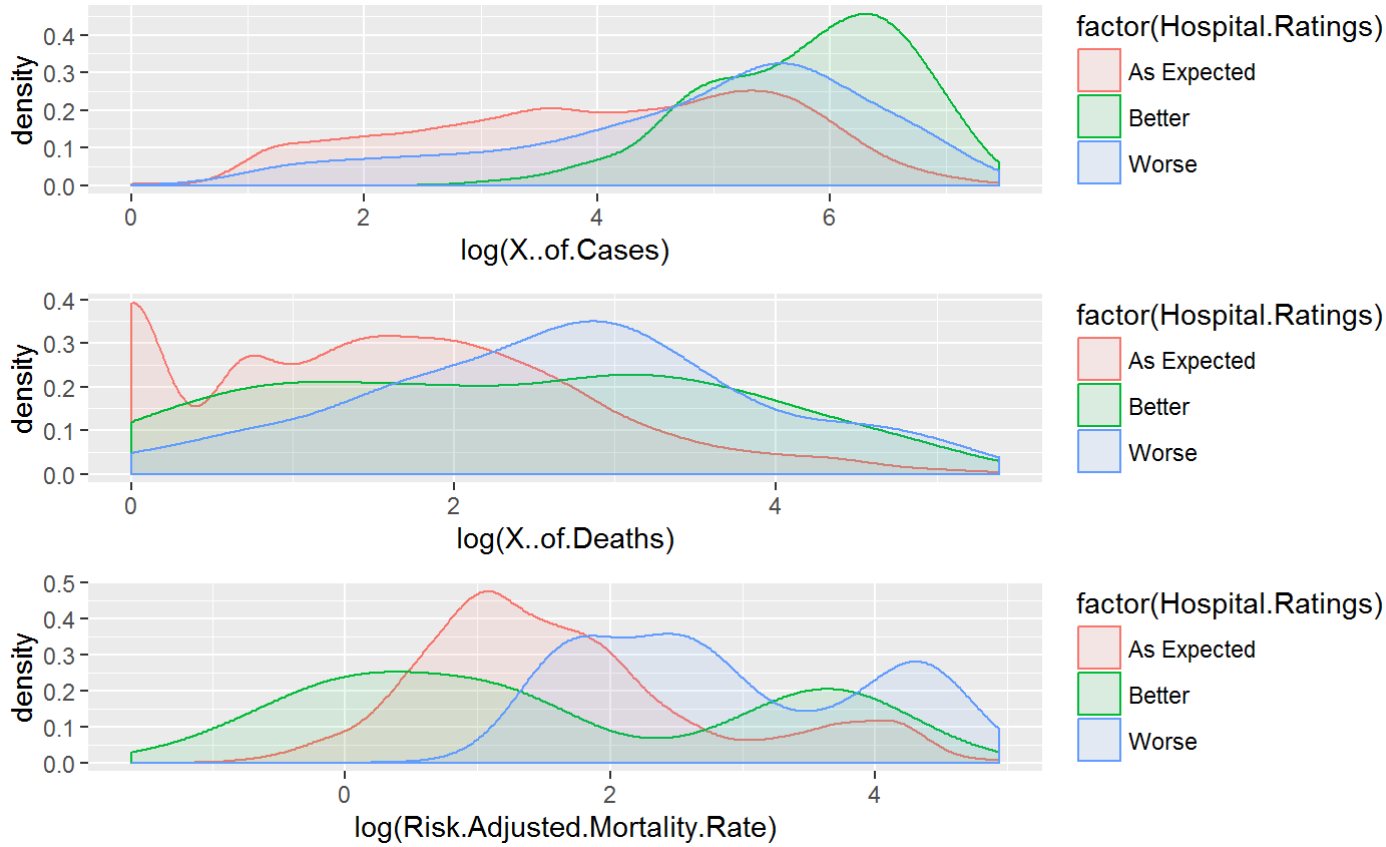
- Can we predict future hospital quality ratings based on risk adjusted mortality rates, number of deaths, number of cases, medical procedures performed and medical conditions treated in 2012-2013?

Description of Data Set

Dataset: is available online from California Hospital Inpatient Mortality Rates and Quality Ratings, 2012-2013

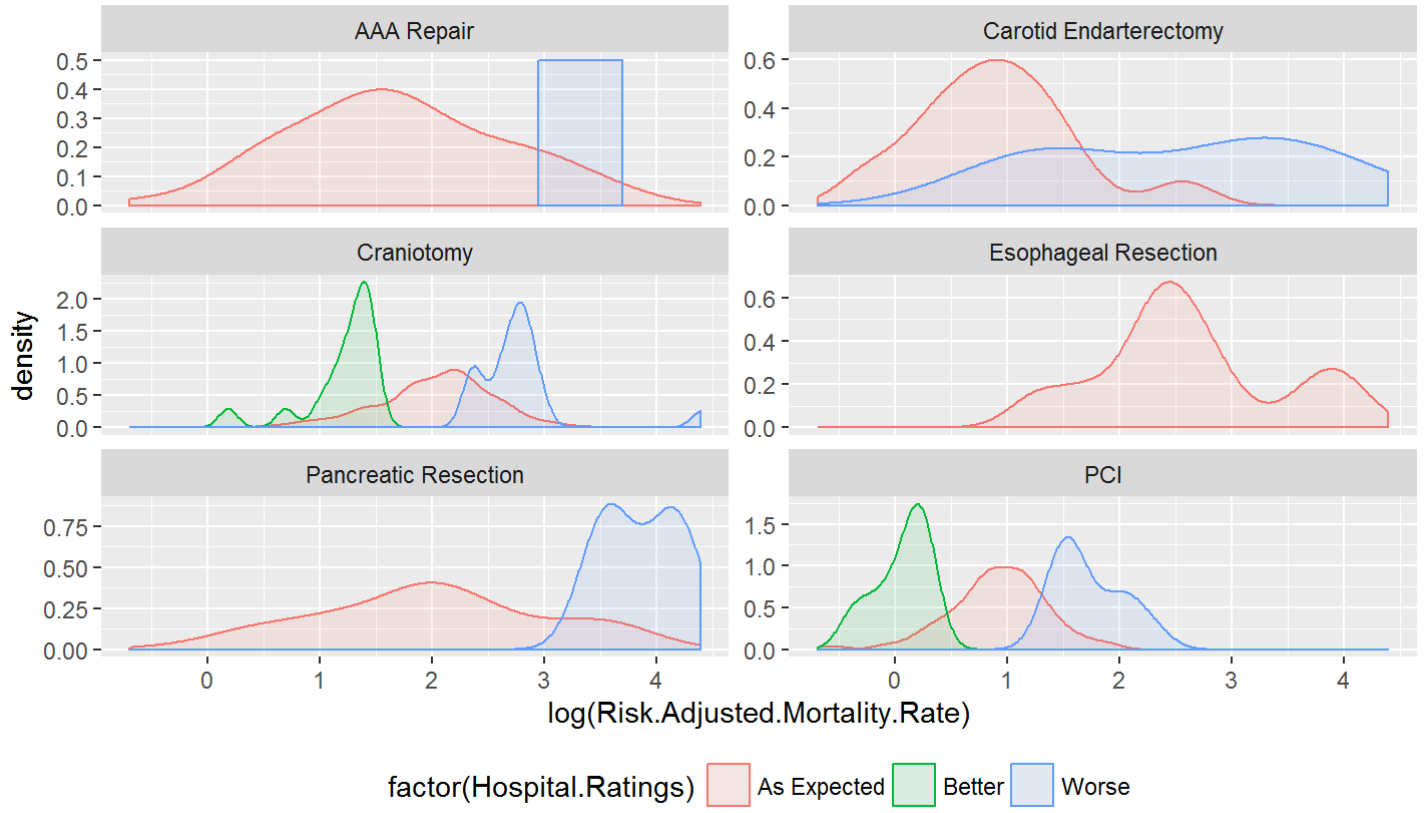
- Risk-adjusted mortality rates
- Number of deaths and number of cases
- 8 medical conditions treated and 6 medical procedures performed
- Year: 2012 and 2013
- County: 55 counties
- Hospital: 341 hospitals
- Longitude and latitude of hospitals
- Hospital Ratings: As Expected, Better and Worse

Density Plots for Number of Cases and Deaths, and Risk Adjusted Mortality Rate by Hospital Ratings



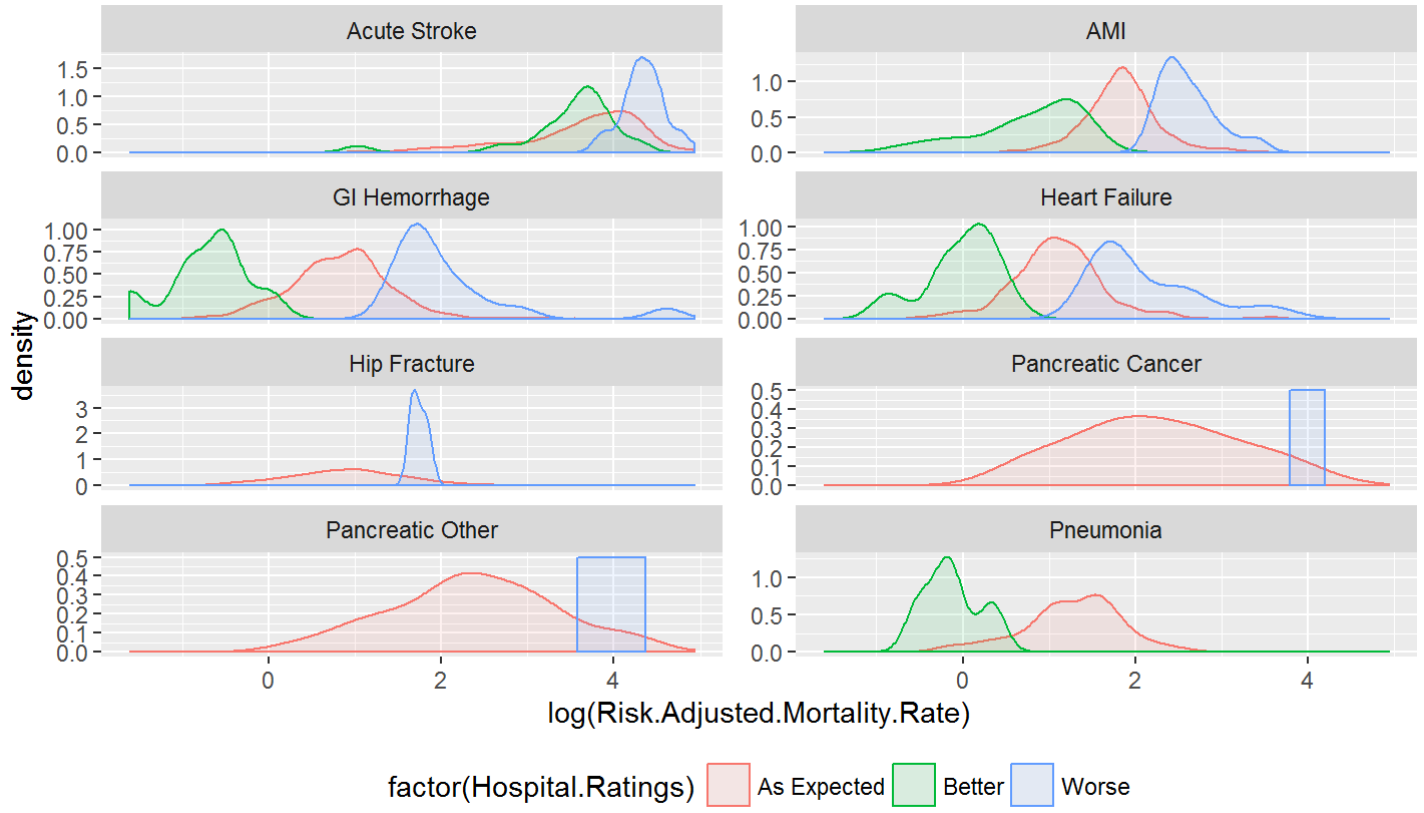
There is a possible **association** between the risk adjusted mortality rate and hospital ratings.

Density Plots for Risk Adjusted Mortality Rate by Procedures Performed and Hospital Ratings



There is **association** between the risk adjusted mortality rate and hospital ratings.

Density Plots for Risk Adjusted Mortality Rate by Conditions Treated and Hospital Ratings



There is **association** between the risk adjusted mortality rate and hospital ratings.

Associations between medical procedures with number of cases and deaths, and risk adjusted mortality rate

```
## # A tibble: 6 x 4
##      Procedure.Condition all_cases all_deaths all_mortality_rate
##                   <fctr>     <int>      <int>              <dbl>
## 1             AAA Repair      4927         59              508.5
## 2 Carotid Endarterectomy     12478         60              290.2
## 3             Craniotomy     30164       2159             2354.6
## 4   Esophageal Resection       619         28              436.9
## 5   Pancreatic Resection      3356         93             1002.8
## 6                    PCI     78660       2028              793.6
```

The most severe outcomes are for PCI, Craniotomy and Pancreatic Resection procedures.

Associations between medical conditions with number of cases and deaths, and risk adjusted mortality rate

```
## # A tibble: 8 x 4
##   Procedure.Condition all_cases all_deaths all_mortality_rate
##                 <fctr>     <int>      <int>              <dbl>
## 1        Acute Stroke    217956      20461            26582.9
## 2                 AMI     93594       5731             3863.4
## 3       GI Hemorrhage     94804       2099             1597.8
## 4       Heart Failure    155066       4778             2200.0
## 5        Hip Fracture     32245        744              945.8
## 6   Pancreatic Cancer      1787         43              632.0
## 7    Pancreatic Other      1425         41              590.0
## 8           Pneumonia     20630       1019              552.5
```

The most severe outcomes are for Acute Stroke, Heart Failure, AMI and GI Hemorrhage conditions.

Hospital Ratings with medical procedures and conditions

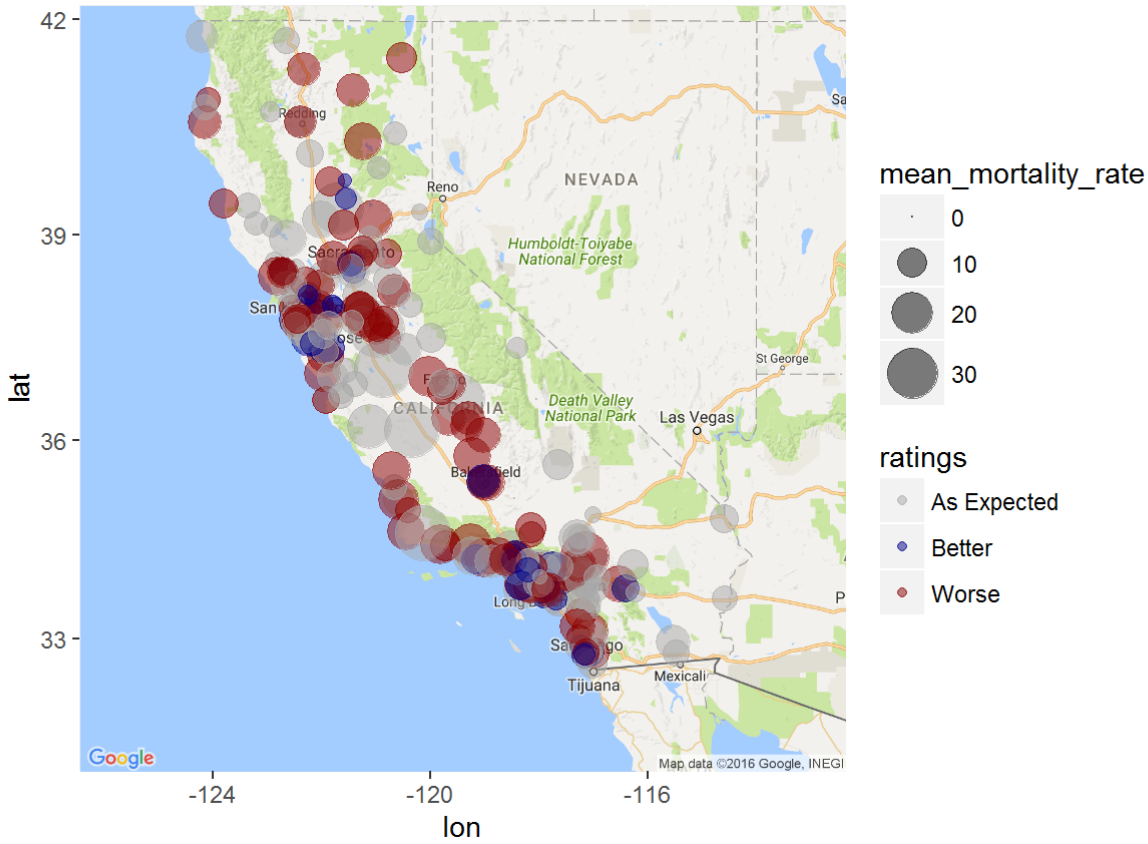
```
##
##                          As Expected      Better       Worse
##   AAA Repair              5.80218873  0.00000000  0.04129672
##   Acute Stroke           10.81973983  1.01176956  0.90852777
##   AMI                    11.06752013  0.47491224  0.64009911
##   Carotid Endarterectomy  8.19739831  0.00000000  0.14453851
##   Craniotomy              5.40986992  0.37167045  0.37167045
##   Esophageal Resection    1.54862688  0.00000000  0.00000000
##   GI Hemorrhage          12.28577328  0.20648358  0.35102209
##   Heart Failure          11.48048730  0.47491224  0.76398926
##   Hip Fracture            8.71360727  0.00000000  0.08259343
##   Pancreatic Cancer       2.89077018  0.00000000  0.04129672
##   Pancreatic Other        2.64298988  0.00000000  0.04129672
##   Pancreatic Resection    3.84059467  0.00000000  0.08259343
##   PCI                     5.78154037  0.12389015  0.26842866
##   Pneumonia               3.03530869  0.08259343  0.00000000
```

Better/worse ratings are for Craniotomy, PCI, Acute Stroke, AMI, Heart Failure and GI Hemorrhage.

Conclusions

- There is **association** between the risk adjusted mortality rate and hospital ratings.
  - Lower the risk adjusted mortality rate, better the hospital ratings.
  - Higher the risk adjusted mortality rate, worse the hospital ratings.
- **Procedures** with severe outcomes:
  - PCI, Craniotomy and Pancreatic Resection.
- **Conditions** with severe outcomes:
  - Acute Stroke, AMI, Heart Failure and GI Hemorrhage.

Mapping of overall hospital quality ratings and mean mortality rate over all conditions and procedures



172 hospitals with "As Expected", 69 with "Better" and 99 with "Worse" ratings.

Top 5 hospitals with the best quality ratings:

```
## # A tibble: 5 x 1
##                                   Hospital
##                                      <chr>
## 1 Kaiser Foundation Hospital  Redwood City
## 2       Kaiser Foundation Hospital  Sunset
## 3        Centinela Hospital Medical Center
## 4                   Scripps Green Hospital
## 5              Cedars Sinai Medical Center
```

Top 5 hospitals with the lowest mean mortality rate:

```
## # A tibble: 5 x 1
##                                  Hospital
##                                     <chr>
## 1             Corcoran District Hospital
## 2 Eastern Plumas Hospital  Portola Campus
## 3                   Glenn Medical Center
## 4    Good Samaritan Hospital  Bakersfield
## 5               Hoag Orthopedic Institute
```

Top 5 hospitals for treatment of Heart Failure condition:

```
## # A tibble: 5 x 1
##                            Hospital
##                               <chr>
## 1 Adventist Medical Center  Reedley
## 2          Anaheim General Hospital
## 3             Sherman Oaks Hospital
## 4        Barstow Community Hospital
## 5 Centinela Hospital Medical Center
```

Top 5 hospitals to perform the Pancreatic Resection:

```
## # A tibble: 5 x 1
##                                       Hospital
##                                          <chr>
## 1                 Alameda County Medical Center
## 2         Community Hospital Monterey Peninsula
## 3  Community Hospital of The Monterey Peninsula
## 4 Community Memorial Hospital  San Buenaventura
## 5                          Eden Medical Center
```

Predictions

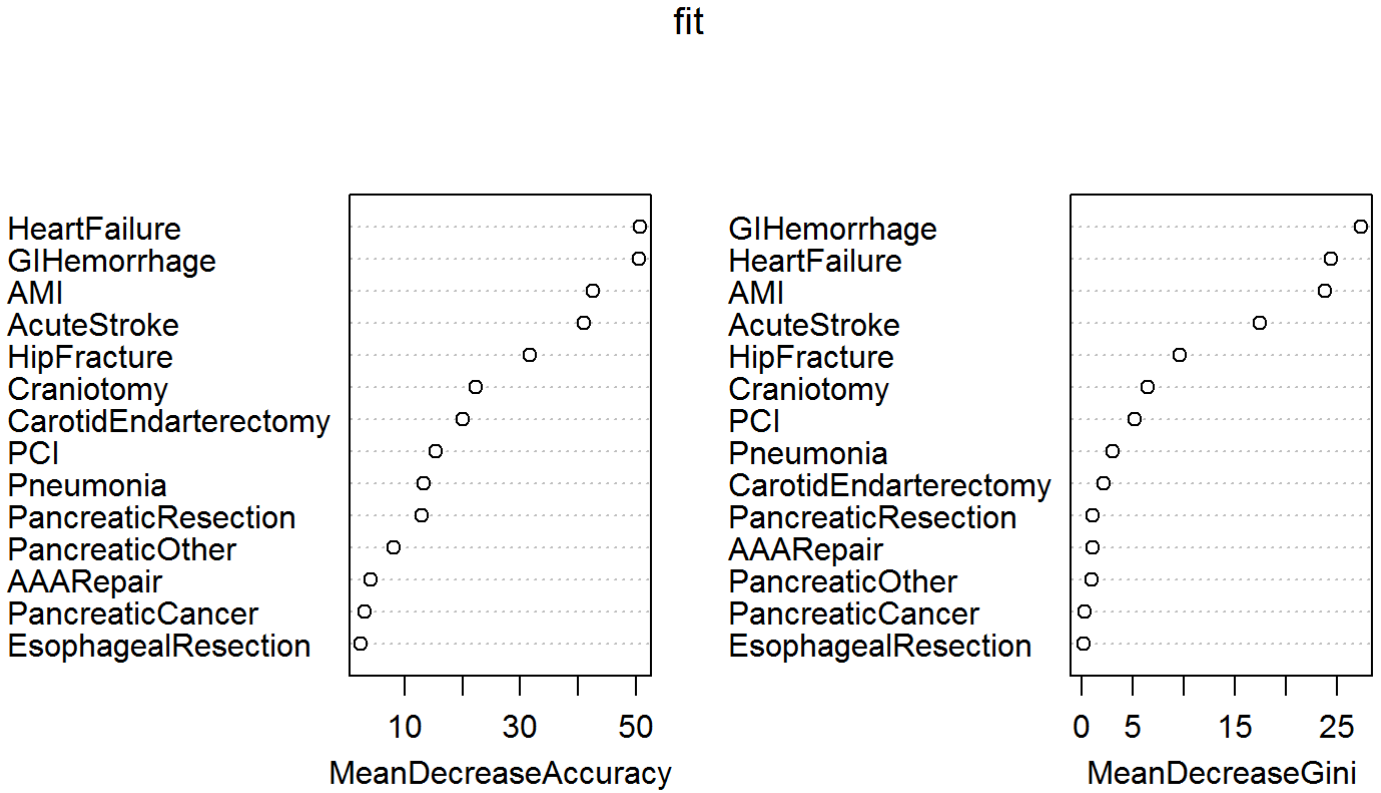
Approach:

- Predict hospital quality ratings using **random forests and classification decision trees**.
- Train the models and evaluate the model performances on 2012 training data.
- Test the model performances on 2013 test data.

Convert the dataset into **wide** format and split it into 2012 training and 2013 test sets:

```
train_wide <- df_wide[which(df_wide$Year==2012),]
test_wide_original <- df_wide[which(df_wide$Year==2013),]
test_wide <- subset(test_wide_original, select = -Hospital.Ratings)
```

Feature Enginering with Random Forests



The most important variables are Heart Failure, GI Hemorrhage, AMI and Acute Stroke, Hip Fracture **conditions**.

Random Forests Model Performance on training set using all variables

```
##             As Expected Better Worse class.error
## As Expected         297     17    12  0.08895706
## Better                8     37     2  0.21276596
## Worse                25      5    35  0.46153846
```

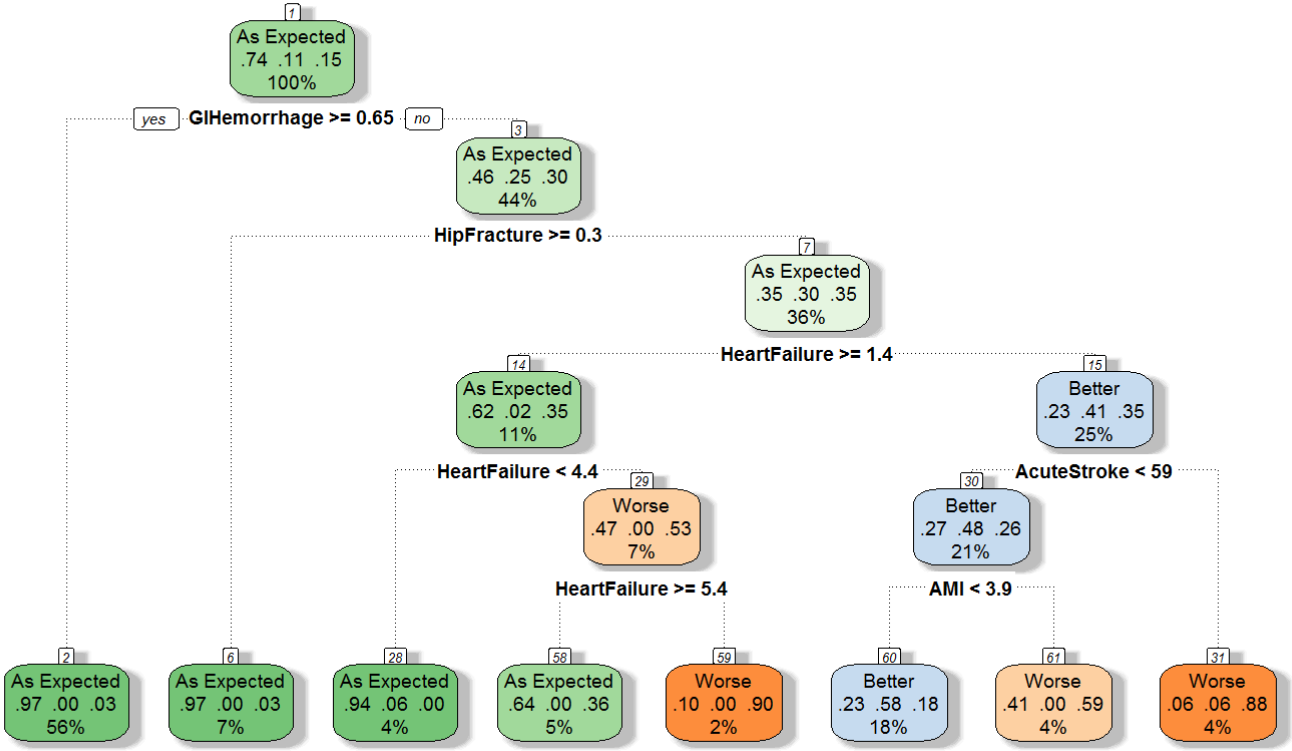
- Accuracy: 0.841 and Error Rate: 0.159

Predictions on test set using all variables

```
##              Predicted
## Actual        As Expected Better Worse
##   As Expected         291     20    13
##   Better                7     40     4
##   Worse                26      2    40
```

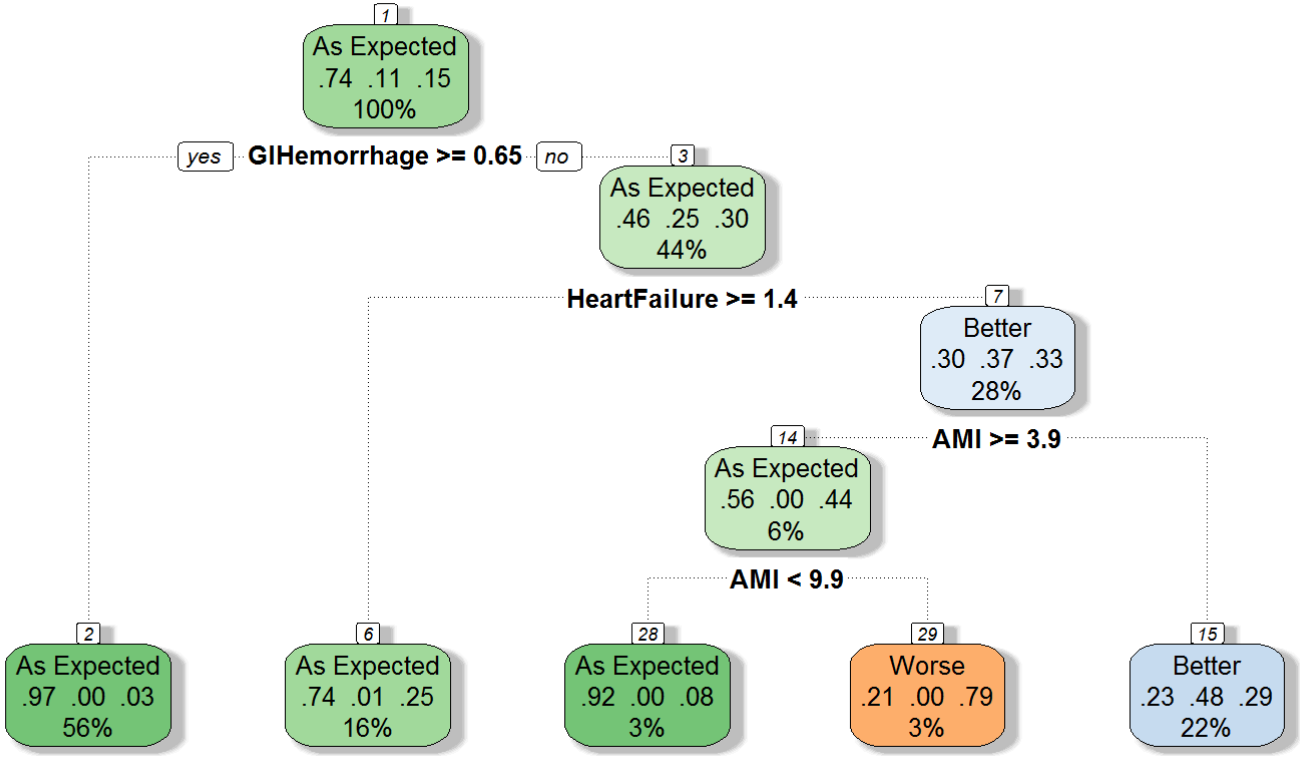
- Accuracy: 0.8375 and Error Rate: 0.1625

Hospital Ratings Prediction Using Classification Decision Trees (CART) with all variables



Rattle 2016-Aug-16 14:07:28 maria

Hospital Ratings Prediction with CART using AMI, GI Hemorrhage and Heart Failure variables



Predictions on test dataset: Comparison of two models

Model with all variables has accuracy 0.8126 and error rate 0.1874.

```
##              Predicted
## Actual        As Expected Better Worse
##   As Expected         287     24    13
##   Better                4     44     3
##   Worse                26     13    29
```

Model with three variables has accuracy 0.7833 and error rate 0.2167.

```
##              Predicted
## Actual        As Expected Better Worse
##   As Expected         292     27     5
##   Better                5     46     0
##   Worse                29     30     9
```

Conclusions

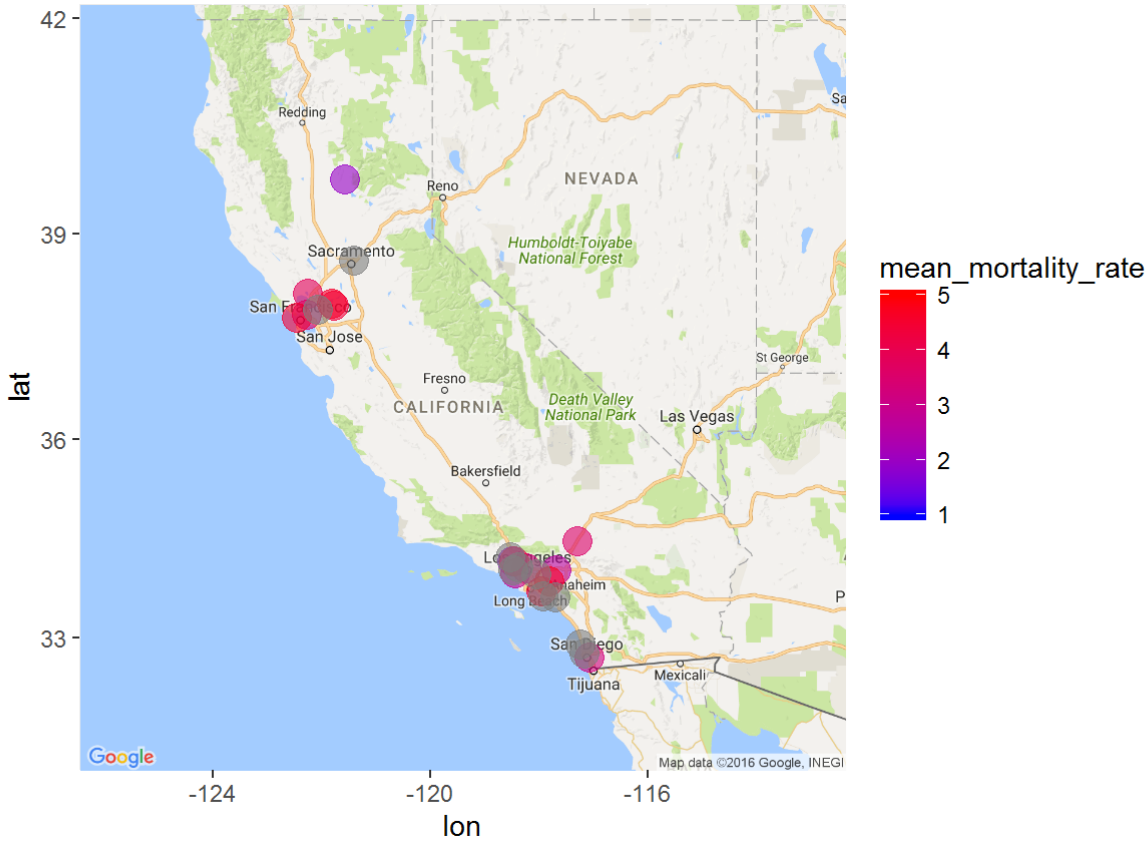
- **Accuracy on the test data set using**
  - Random Forests with all variables: 0.8375
  - CART with all variables: 0.8126
  - CART with three variables: 0.7833
- **Random forests** gives the best performance, however is not good enough to predict hospitals with the best care in future.
- **Random forests** predicts that classification of **hospital ratings** depends on **conditions and not procedures** with the most severe patient outcomes.

Future Work

- Predict hospital quality ratings using **multinomial logistic regression**.
  - Train the model and evaluate the model performance on 2012 training data.
  - Test the model performance on 2013 test data.
- Compare three models: random forests, classification decision trees and multinomial logistic regression.
  - Summarize which model gives the best performance on 2012 training data and on 2013 test data.
  - Choose the best model and test its performance on 2014 test data.
- Recommend which hospitals will have the best care in future using predicted hospital ratings.

Recommendations to Patients

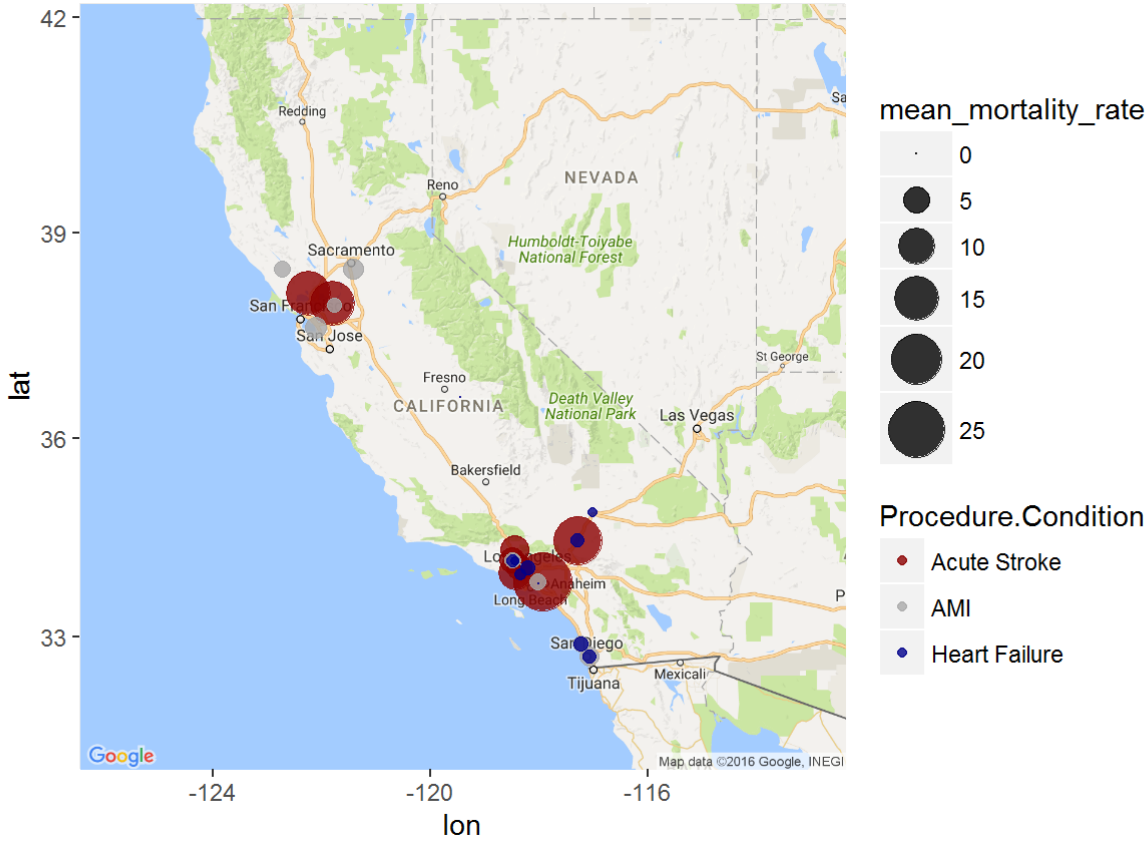
Top 25 hospitals with the best overall ratings and the lowest mean mortality rate in state of California



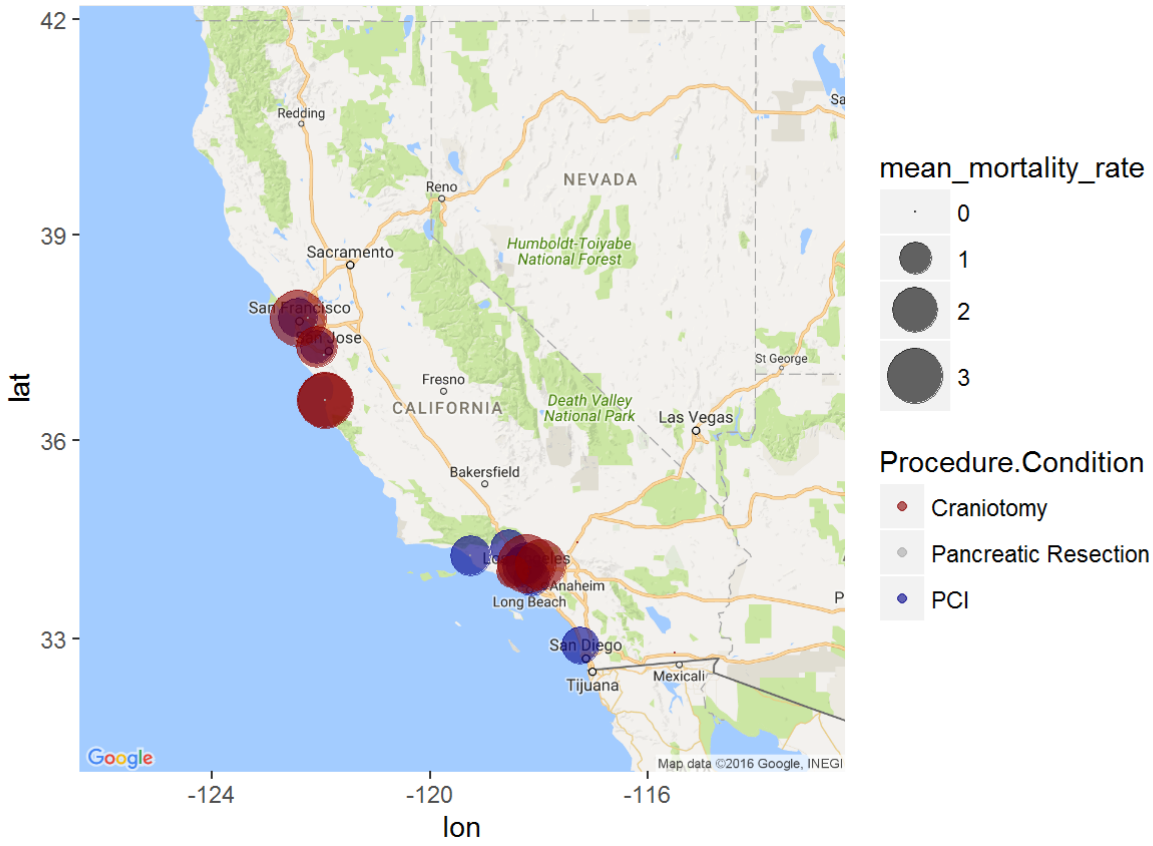
Top 10 hospitals with the best overall ratings and the lowest mean mortality rate in state of California

```
##  [1] "Encino Hospital Medical Center"
##  [2] "Feather River Hospital"
##  [3] "Chino Valley Medical Center"
##  [4] "Marina Del Rey Hospital"
##  [5] "Paradise Valley Hospital"
##  [6] "Pacific Alliance Medical Center, Inc."
##  [7] "Kaiser Foundation Hospital  Oakland Campus"
##  [8] "Desert Valley Hospital"
##  [9] "Sutter Solano Medical Center"
## [10] "Sherman Oaks Hospital"
```

Top hospitals with the best ratings and the lowest mean mortality rate for Acute Stroke, AMI and Heart Failure conditions



Top hospitals with the best ratings and the lowest mean mortality rate for PCI, Craniotomy and Pancreatic Resection procedures



There are 7 hospitals that have the best ratings and the lowest mortality rate for the most severe conditions:

```
## [1] "Encino Hospital Medical Center" "Anaheim General Hospital"
## [3] "Sherman Oaks Hospital"          "Encino Hospital Medical Center"
## [5] "Desert Valley Hospital"         "Paradise Valley Hospital"
## [7] "Scripps Green Hospital"
```

There are 7 hospitals that have the best ratings and the lowest mortality rate for the most severe procedures:

```
## [1] "El Camino Hospital"
## [2] "California Pacific Medical Center  Pacific Campus"
## [3] "Glendale Adventist Medical Center  Wilson Terrace"
## [4] "Community Hospital Monterey Peninsula"
## [5] "Community Hospital of The Monterey Peninsula"
## [6] "Community Memorial Hospital  San Buenaventura"
## [7] "Fresno Heart and Surgical Hospital"
```

Resources

Datasets are available online

- **2012-2013 dataset:** is available from California Hospital Inpatient Mortality Rates and Quality Ratings
- 2014 dataset

Files for Capstone Project are available online

- Rpubs
- github