

# **Springboard Data Science Intensive Capstone Project**

Patrick Long  
May, 2017

# **The Problem to be Addressed**

**Patient hospital experience plays a central role in patient well-being yet is an all too often neglected facet of our healthcare system.**

**While some hospital systems are highly attentive to the first-hand experiences of patients, it is difficult as patients – the healthcare consumer – to identify hospitals that excel at this dimension of patient care.**

**Moreover, many hospitals lack strong metrics to predict the quality of patient experience within their organization, which stifles clinical care.**

# **The Proposed Solution:**

**Hospital performance metrics are publically available through Medicare's Hospital Compare Database.**

**The goal of this study will be twofold:**

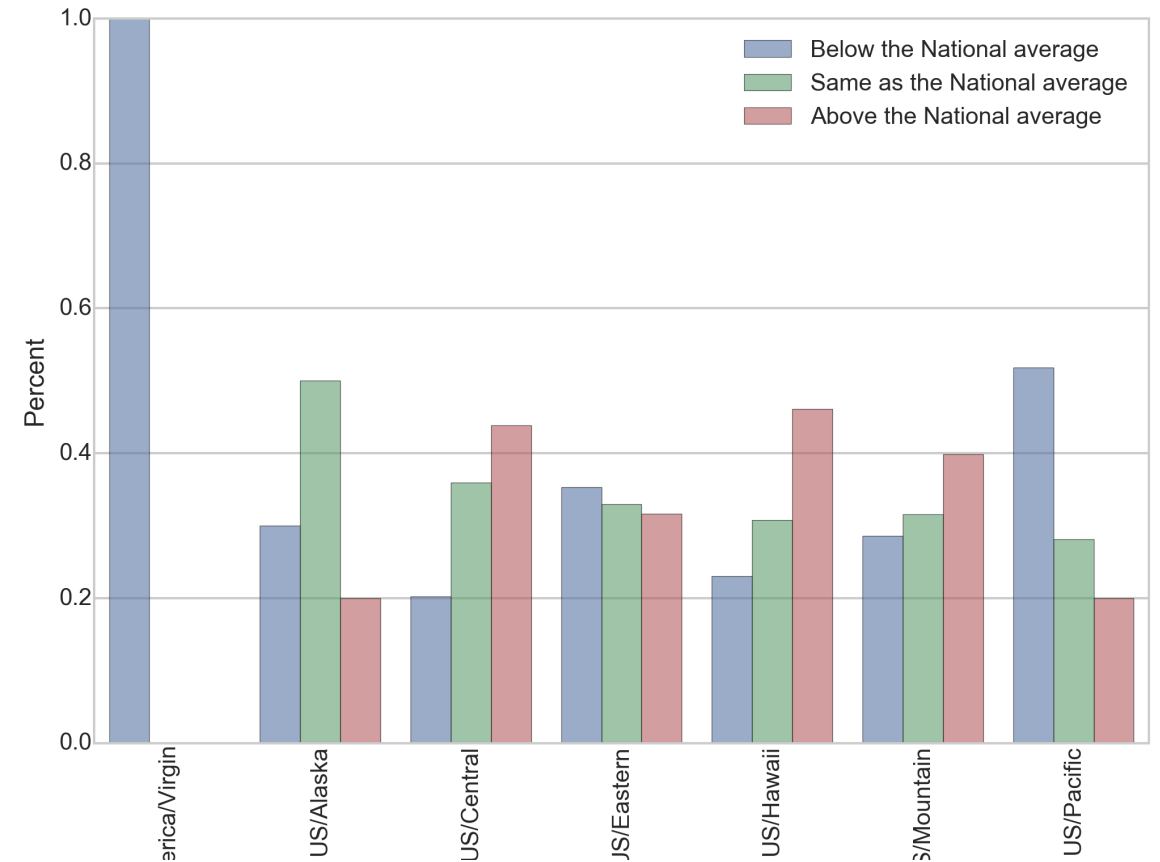
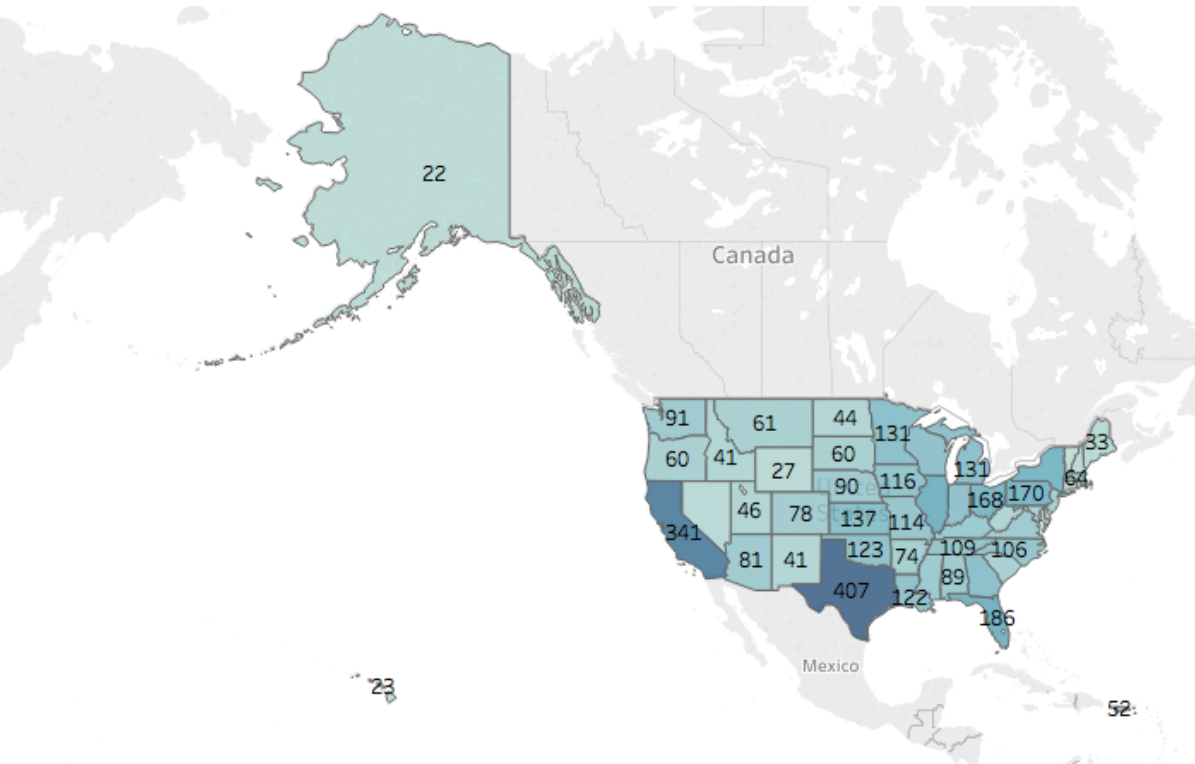
- 1. Develop a model to accurately predict patient hospital experience using Medicare Hospital Compare data.**
- 2. Identify a subset of key hospital attributes that provide the greatest value toward predicting patient hospital experience.**

# Overview of Medicare Data used for this Study



- **General Hospital Information**
  - Location, ownership model, the inclusion of various clinical services (e.g. electronic health records, emergency care, medical imaging).
- **Hospital readmission metrics**
  - Number of readmissions, discharges, level of readmissions relative to hospitals treating similar patient populations.
- **Medicare Spending per Beneficiary/Patient (MSPB)**
  - Level of Medicare reimbursement per patient relative to the national average for Medicare reimbursement.
- **Patient experience** ratings used in this analysis are derived from Medicare's Hospital Consumer Assessment of Healthcare Providers and Systems Survey (CAHPS®).
- Data from a total of 2,208 hospitals were used in this study.

# Patient Experience Varies by Geography

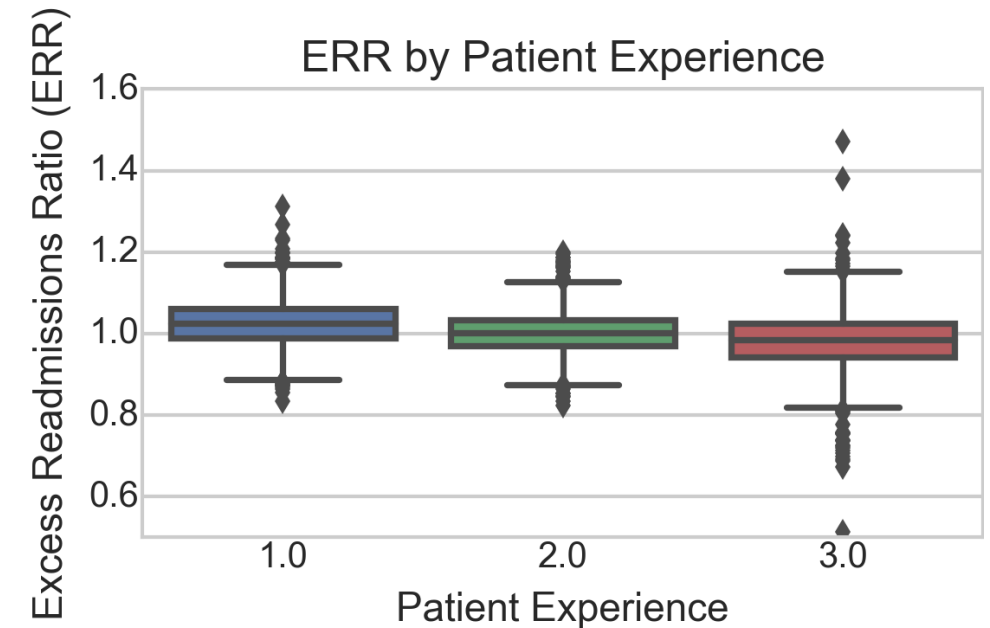
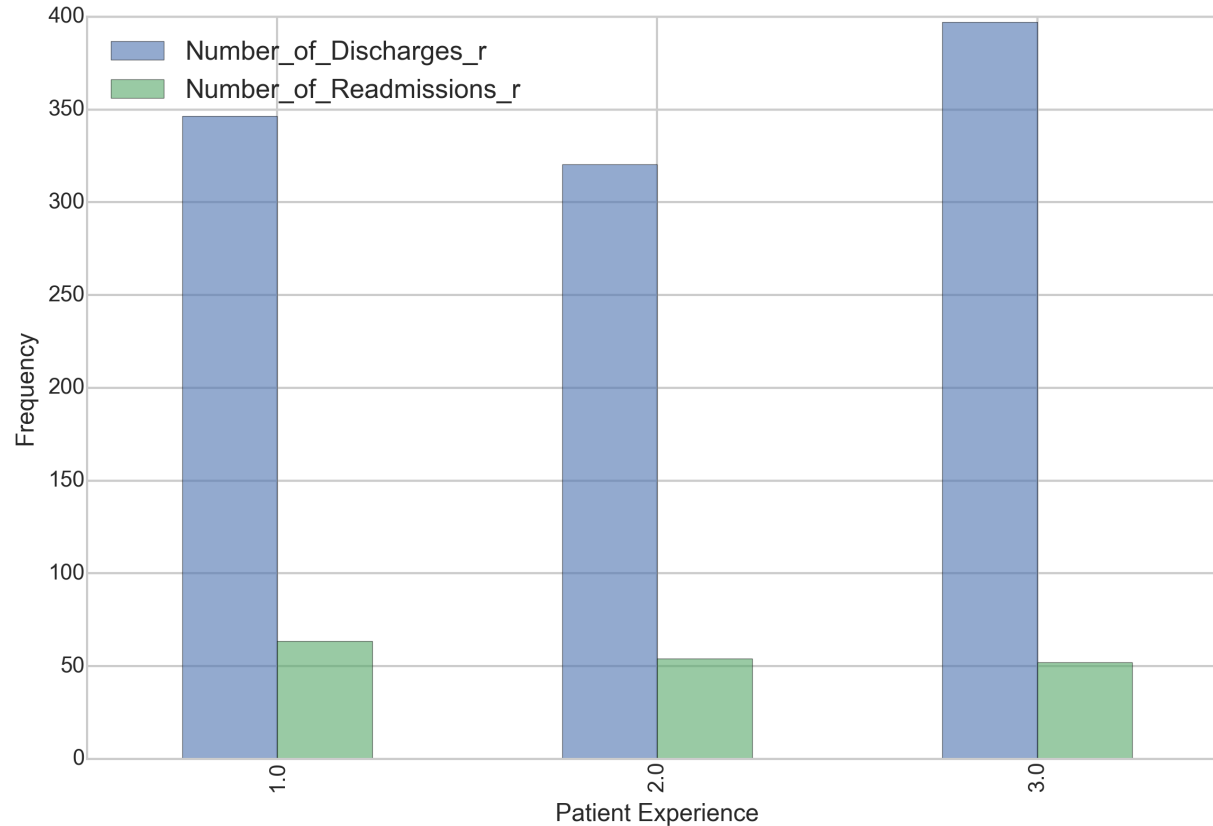


## Take Home

- Hospitals in this data set are concentrated in large populous states like California and Texas.
- Regions with highest ratings for patient experience are in the central, mountain and US/Hawaii time zones.
- Regions with low patient experience ratings are in the Virgin Islands and the US/Pacific time zone.

# Patient Experience Varies by Readmissions Metrics

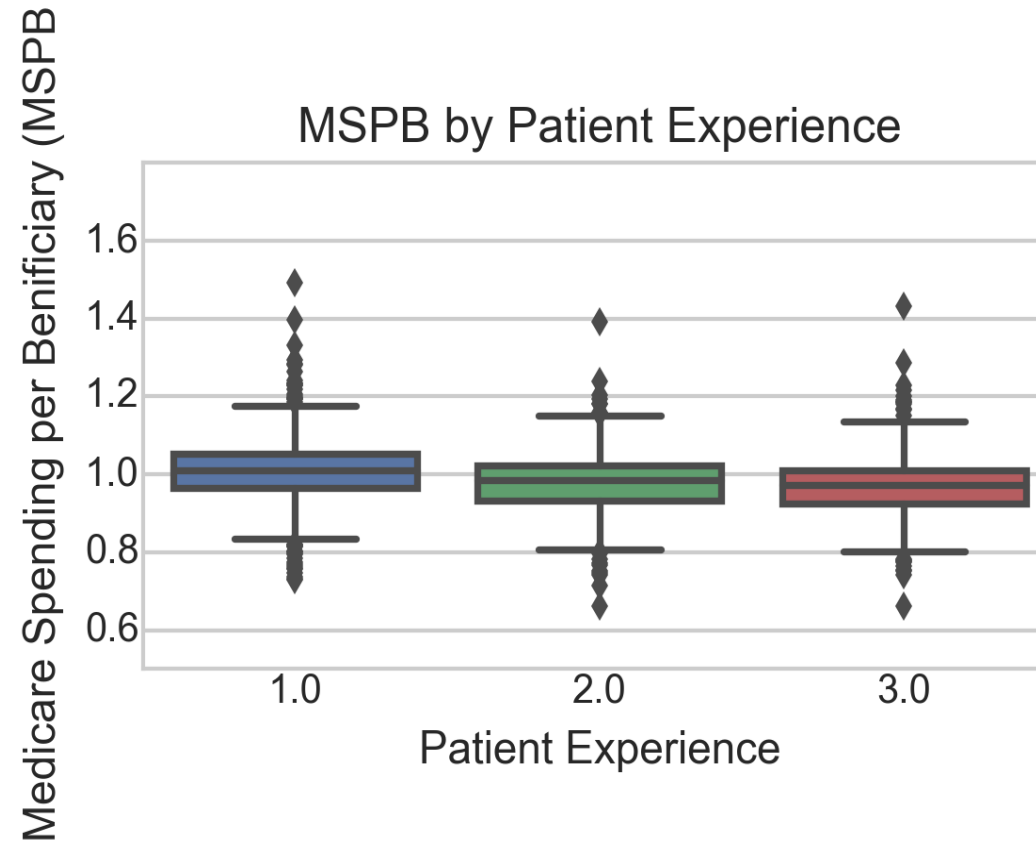
## Discharge and Readmission Number by Patient Experience



## Take Home

- Hospitals with '*above the national average*' rating in patient experience have the highest number of discharges yet the lowest number of readmissions. This trend is reflected in the hospital ERR, as well.
- MSPB inversely correlates with ratings of patient experience. This suggests that highly rated hospitals receive less Medicare reimbursement as measured by MSPB relative to the national average for similar hospitals.

# Patient Experience Varies by Medicare Spending per Patient

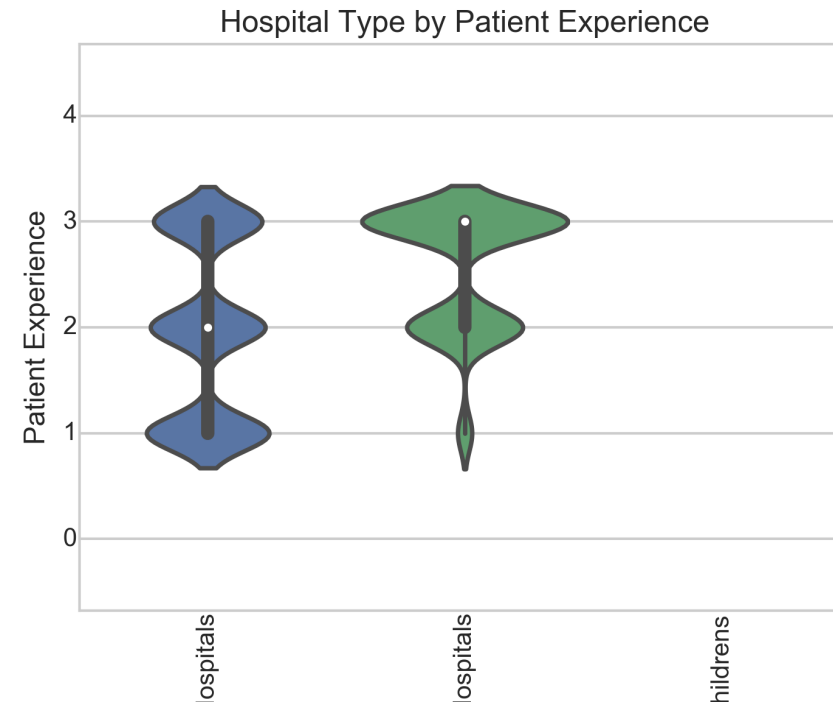
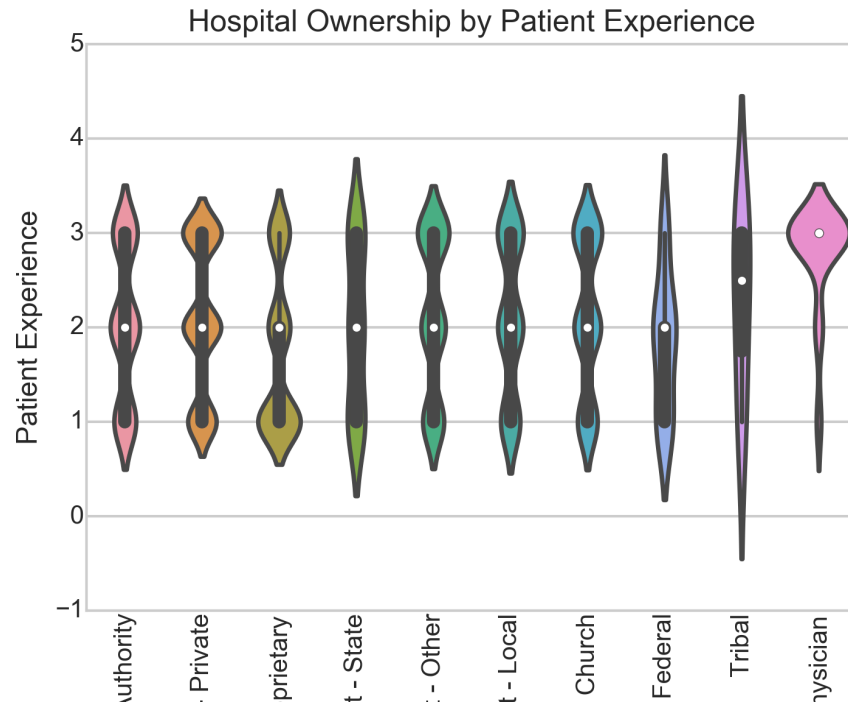


## Take Home

- Hospitals with “*Above the national average*” rating in patient experience have the highest number of discharges yet the lowest number of readmissions. This trend is reflected in the hospital ERR, as well.
- MSPB inversely correlates with ratings of patient experience. This suggests that highly rated hospitals receive less Medicare reimbursement as measured by MSPB relative to the national average for similar hospitals.



# Patient Experience Varies by Hospital Ownership and Type



## Take Home

- Physician run hospitals show the highest level of patient satisfaction, while proprietary hospital ownership is associated with poor patient satisfaction.
- For hospital type, critical access hospitals have more favorable patient experience ratings relative to acute care hospitals.
- None of the children's hospitals have patient experience ratings and will thus be not be used for machine learning predictions of patient experience.



## Key Questions

*Can patient hospital experience be predicted using Medicare Hospital Compare Data?*

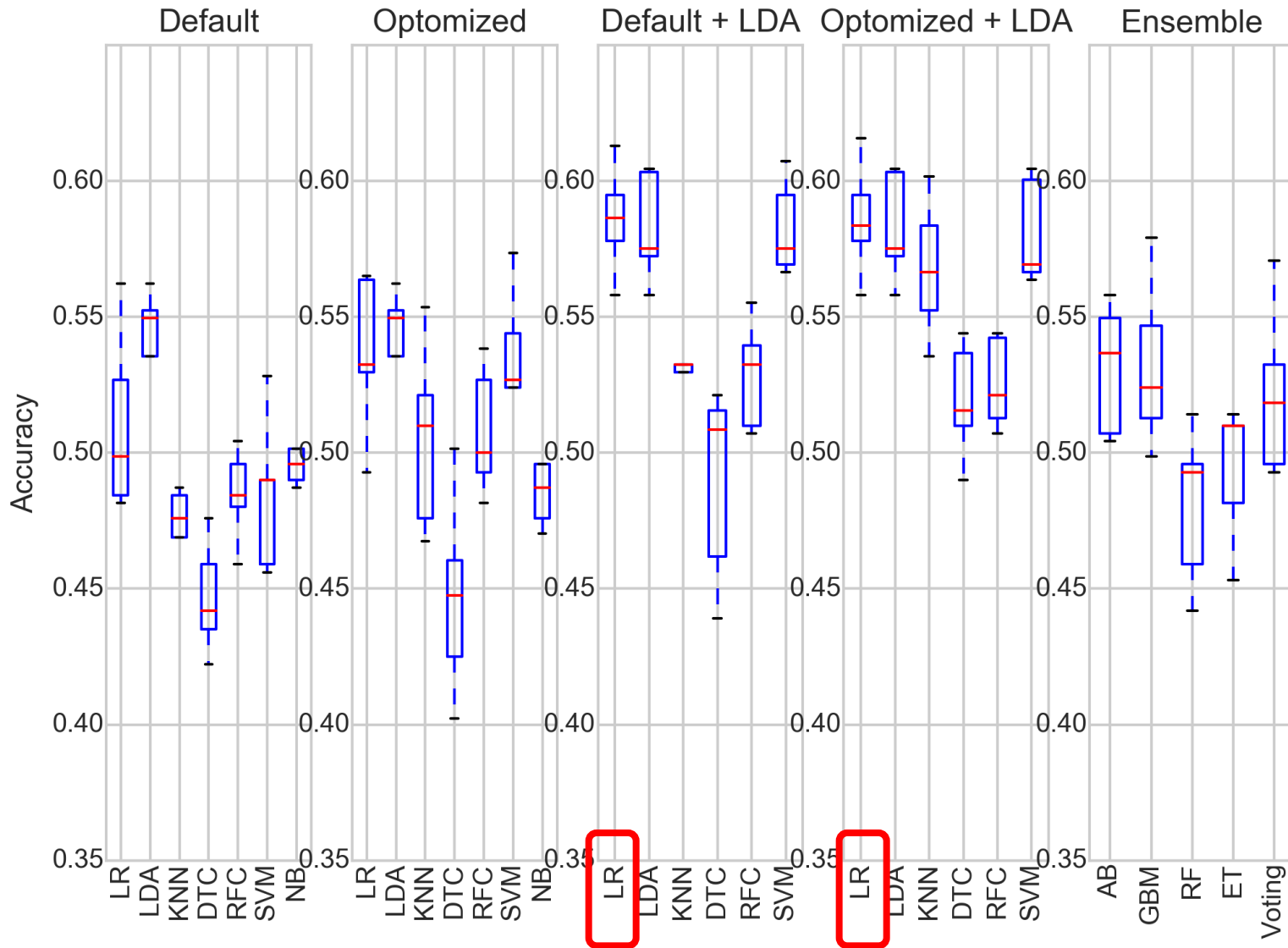
*What key hospital attributes are most effective at predicting patient experience?*

## Approach

*Test various machine learning models to see how accurately one can predict patient experience.*

*Identify the key hospital attributes responsible for machine learning predictive accuracy.*

# What Models are Most Accurately Predict Patient Hospital Experience?



## Model Optimization

**Logistic regression (LR)** was consistently among the most accurate algorithms at predicting patient hospital experience using Medicare Hospital Compare data during the model optimization state of this analysis.

# How Does Logistic Regression Perform as a Predictive Model of Patient Hospital Experience?

## Confusion Matrix

	Predicted 0	Predicted 1	Predicted 2
Actual 0	127	33	11
Actual 1	49	57	41
Actual 2	32	42	50

## Classification Report

	precision	recall	f1-score	support
0	0.61	0.74	0.67	171
1	0.43	0.39	0.41	147
2	0.49	0.40	0.44	124
avg / total	0.52	0.53	0.52	442

## Model Validation

- This LR has about 52% **accuracy** in predicting patient hospital experience.
- This LR shows greatest **precision** when predicting poor patient experiences (61%).
- This LR shows greatest **recall** when predicting poor patient experiences (74%).

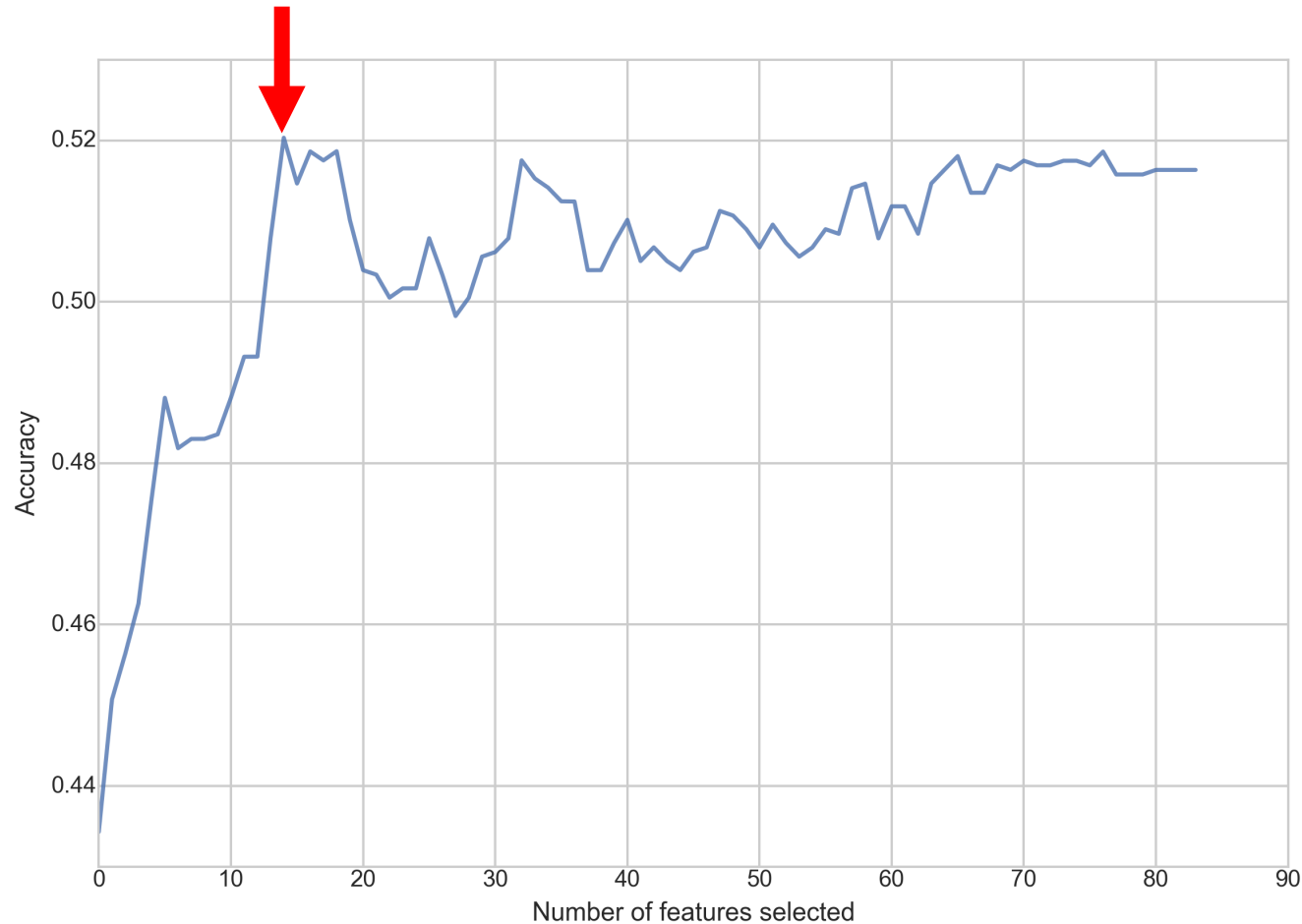
*accuracy= True Positives / All Instances*

*precision = True Positives / (True Positives  
+ False Positives)*

*recall= True Positives / (True Positives  
+ False Negatives)*

# What Hospital Attributes/Features Most Accurately Predict Patient Hospital Experience?

Optimal Feature Number



**15 of 84 features were identified as important to predicting patient hospital experience:**

- Number of Discharges
- Number of Readmissions
- MSPB
- Proprietary Hospital Ownership
- *“Below National Average”* Timeliness of Care
- Hospital Location in 1 of 10 states/districts:
  - AR, DC, LA, NE, NH, NM, NY, SC, SD, WI

## Summary of Key Findings

- Patient hospital experience can be predicted with approximately 52% accuracy using publically available Medicare Hospital Compare Data.
  - This represents a ~20% improvement over random guess (i.e. 33% likelihood of a correct guess since low-mid-high patient experience ratings are nearly evenly distributed).
- Fifteen hospital attributes were identified that account for the greatest predictive accuracy of patient hospital experience.
  - These include readmissions metrics, Medicare spending per patient, hospital ownership model, timeliness of care, and hospital geography.

## Key Recommendations

- When predicting and/or improving patient hospital experience, attention should focus on readmissions metrics, spending per patient, ownership model, timeliness of care, and hospital geography.
- While the Medicare Hospital Compare Data utilized in this study increased predictive accuracy  $\sim 20\%$  over a random guess, it is advisable to investigate additional hospital attributes that could provide greater predictive accuracy.
- Reliably detecting negative and positive patient experiences may be of the greater value to both the patient/healthcare consumer and healthcare provider than identifying neutral patient experiences. Using a binary classification system to assess and predict the quality of patient experience system is recommended.

# Appendix



# Key Terms

**Number of Discharges:** Number of patients discharged over 30-day period for heart attack, heart failure, and pneumonia.

**Number of Readmissions:** Number of patients readmitted over 30-day period for heart attack, heart failure, and pneumonia.

**Excess Readmission Ratio (ERR):** This ratio is created by dividing a hospital's number of "predicted" 30-day readmissions for heart attack, heart failure, and pneumonia by the number that would be "expected," based on an average hospital with similar patients. A ratio greater than 1 indicates excess readmissions.

**MSPB:** The Medicare hospital spending per patient (Medicare Spending per Beneficiary, MSPB) measure indicates whether Medicare spends more, less or about the same per Medicare patient treated in a specific hospital, compared to how much Medicare spends per patient nationally. This measure includes any Medicare Part A and Part B payments made for services provided to a patient during the 3 days prior to the hospital stay, during the stay, and during the 30 days after discharge from the hospital.

**EHR:** The use of electronic health records (EHR).

**Hospital Type:** Hospital Type i.e. critical access, acute care, children's'.

**Hospital Ownership:** Model of hospital ownership e.g. government run, proprietary, volunteer/non-profit.

**Emergency Services:** Meets emergency services criteria.

**Safety of Care:** Safety of care ratings for a hospital (below, same as, or above national average).

**Effectiveness of Care:** Effectiveness of care ratings for a hospital (below, same as, or above national average).

**Timeliness of Care:** Timeliness of care ratings for a hospital (below, same as, or above national average).

**Efficient use of Medical Imaging:** Medicare score of medical imaging use.

**Hospital Overall rating:** An aggregate score of hospital rating from low to high (1 - 5) derived from the above measures. Will be used for data exploration but not prediction of patient experience.