

ORIE 4741 Project Proposal

Matthew Dalton and Katelyn Glassman

September 22, 2016

1 Our Problem

Our project's goal is to determine which hospitals provide the best care for different medical conditions. To do this we will examine which hospitals have a low ratio of severity of illness to length of stay as well as the costs for specific procedures. Going deeper we will also examine how demographic patterns and average costs in hospitals are correlated recovery rates. To specify this examination, we plan to focus in on particular condition and their varying recovery rates across different hospitals. Specifically, we plan to focus on sudden onset conditions such as stroke, aneurysm, and heart attack rather than chronic conditions. By choosing these conditions it will be easier to measure the efficacy of each hospital, as longer term diseases require more nuanced care and oftentimes patients are in and out of treatment making their recovery harder to track.

2 Example Dataset

One of the datasets we will be using is the SPARCS Hospital Inpatient Discharges for New York State, from which we will focus specifically on hospitals in New York City. This dataset provides basic information about all patients who have been discharged from a New York hospital in the year 2012, including their age, gender, and race. This dataset also provides information about each patient's length of stay, reason for admission, severity of illness, method of payment, and cost. Specifically, the dataset quantifies severity of illness on a scale from 1 to 4, with 1 being minor and 4 being extreme, as well as an ordinal field for risk of mortality, which will help our study. Furthermore, we plan to use this data to help manufacture other features for our analysis such as the ratio of severity of illness to length of stay for certain illnesses and how many patients each hospital treats in a year to help us determine which hospital is the "best". By examining this dataset as well as other we deem relevant, we will hopefully be able to find a solution to our problem.

Link to SPARCS data: <https://health.data.ny.gov/Health/Hospital-Inpatient-Discharges-SPARCS-De-Identified/u4ud-w55t>

3 Problem Significance

The significance of our proposed project is that it will allow us to develop a hospital recommendation system for the NYC community. This system will add tremendous value to community members by giving optimal hospital recommendations based on a prospective patients personal information. While this task will not be a small undertaking, given the relevance of the SPARCS dataset, we are confident that we will be able to develop a model that can classify the effectiveness of each hospital.