DATA 608 – Final Proposal

Vijaya Cherukuri

**Data Source**

The UCI Machine Learning Repository: Diabetes 130-US Hospitals for years 1999-2008 of clinical care at 130 US hospitals and integrated delivery networks. It includes over 50 features representing patient and hospital outcomes. Data Set contains 100000 instances of data, with 55 attributes. The data contains attributes such as patient’s race, gender, age, time in hospital, number of lab tests performed, number of in-patient, outpatient, and emergency visits for prior year, etc.

<http://archive.ics.uci.edu/ml/datasets/diabetes+130-us+hospitals+for+years+1999-2008>

**Background**

Earlier I got a change to work on ADT project (Admission, Discharge and Transfers). Health Insurance companies and

Hospitals are penalized when patients are discharged and readmitted within 30 days.  Readmissions are associated with negative patient and financial outcomes. A readmission can be defined in multiple ways, including:

* Patients who are readmitted to the same hospital, or another applicable acute care hospital with same reason.
* Patients who are readmitted to the same hospital, or another applicable acute care hospital with different reason.

Health Insurance companies uses this data to categorize how can they avoid the readmissions by analyzing the data. Both Health Insurance companies and Hospitals were very keen on analyzing the readmissions data. Here using UCI dataset, we will attempt to create an interactive visualization that could help a hospital (business) to better understand their patient population, and what factors might determine if a patient will be readmitted.

**Tools**

We will create an app using Plotly Dash to create interactive visualizations of our dataset. The visualizations will include various forms including:

* Patient demographics (age, sex, etc.)
* Medication history and how medications impact a patient being readmitted to the hospital