Midterm: Working With Large Data Sets and Machine Learning

Team 4: Emily Strong and Raksha Kaverappa

Code Repository: <https://github.com/erstrong/INFO-7390-ADS-Fall-17-TeamNo.4/tree/master/Mid-Term>

**Part 1: Data Ingestion, EDA, Wrangling**

AWS Bucket:

EDA Notebook: https://github.com/erstrong/INFO-7390-ADS-Fall-17-TeamNo.4/blob/master/Mid-Term/Part%201%20Exploratory%20Data%20Analysis.ipynb

**Data Cleansing Design and Implementation:**

**Analysis:**

**Part 2: Build A Prediction Model**

Model Notebook:

**Multiple Linear Regression Design:**

**Random Forests Design:**

**Neural Networks Design:**

**Analysis:**

**Part 3: Model Deployment**

Azure Webservice:

API Notebook:

**Design and Implementation:**

**Analysis:**

**Part 4: Geospatial Search**

Azure Webservice:

API Notebook:

**Design and Implementation:**

**Analysis:**