

Big data Report

Healthcare Analytics

Ruchika Panjwani

San Emirza

Srijanani Chandrasekaran



imagination at work

Proposal

- The National Health Care Anti-Fraud Association estimates that outright fraud costs the health care industry an estimated \$70 billion to \$230 billion each year. In the EU the figures are similar – 5.6 percent of national health care budgets or \$30 billion to \$100 billion each year (SAS, 2012).
- Recently, CMS has released the claims processing data and Medicare expenditure to the public domain to encourage data transparency.
- This data could be fed into a variety of analytical tech stack to obtain a lot of useful insights and that would be helpful in preventing healthcare fraud and abuses.

Process

- Analyzing the tech stack
 - Big data
 - Apache Hadoop, Pig, Hive
 - Apache Mahout
 - Visualization
 - Mongo DB
 - Tableau
 - SocialLite
- Recommend a suitable tech stack for a preventive model and the possible capabilities

Status

- Downloaded the data from public domain
- Understanding the healthcare process in the US
- Studying the tech stack and its functionalities to analyze what kind of insights could be obtained from the data

Next Steps

- Research on existing preventive models
- Gather information on frauds that are taking place
- Try to analyze the data for related insights

