**INTRODUCTION**

"The FDA Adverse Event Reporting System (FAERS) is a database that contains information on adverse event and medication error reports submitted to FDA."

The FDA produces FDA Adverse Event Reporting System (FAERS) quarterly data. The FAERS include the following:

* demographic and administrative information and the initial report image ID number (if available);
* drug information from the case reports;
* reaction information from the reports;
* patient outcome information from the reports;
* information on the source of the reports;
* a "README" file containing a description of the files.

**Part 1: Parse files & Handling Missing data values**

**Link to Download data files :**

<http://www.nber.org/data/fda-adverse-event-reporting-system-faers-data.html>

The goal of this exercise is to extract “all” csv files present on the faers website save them using script.

• Handle missing data

• Compute summary metrics

• Check for any observable anomalies

## **Part 2: Machine Learning: Prediction & Classification**

A prediction script in a Jupiter notebook in R/Python that builds a Regression model for durg information from the case report / reaction information & patient outcome from the reports using downloaded dataset. Following steps would be covered:

* Variable selection
* Compute MAE, RMS, MAPE for training and testing datasets using Regression algorithms
* Deploy the best algorithm/algorithms on Azure ML studio
* Create API(s)
* Deploy the model using web based front end

## **Part 3: Summarization**

* Summarize the key insights related to different variables
* Present the results using python notebook and Tableau/Power BI