Predicting Lotto Numbers

## Hypothesis:

Lottery Numbers that are drawn are determined to be pseudo random; that is given the same set of factors repeated it is assumed that the same Numbers would be drawn.

## Approach:

Gather data associated with UK National Lottery draws and to analyse past lottery results and see if there are any patterns for the numbers that are drawn. Source data (lotto numbers) can be found at <https://www.national-lottery.co.uk/player/lotto/results/results.ftl>. The following factors:

* Previous Results
* Draw Machine
* Draw Date
* Ball Set

The first phase is to identify and retrieve (**Capture**) the data that is thought to affect the lottery results. This may be from File, RESTful interface, Web Service, etc…

The next phase is to **Store** the data. Typically this will be in a Database. The backend technology that is currently being used is MongoDB, a NoSQL type of Database. Talend Big Data is used to create a job that takes data from the file and pushes it into Mongo. (Though the input can be pretty much any technology)

Final phase is to **Analyse** or Data Mine the data to discover hidden patterns that exists. In essence providing insight and prediction capabilities.

## Software Used

* MongoDB (<http://www.mongodb.org/downloads>)
* Talend Big Data (<http://www.talend.com/download/big-data>)
* Talend Open Studio (<http://www.talend.com/download/data-integration>)
* GitHub (<http://github.com>) and (<http://github.com/riotino/>)

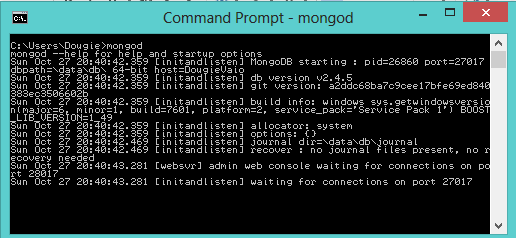
## Directory Setup

* bin – contains the Java executable for loading CSV data into Mongo
* Data – contains the simple CSV file
* Docs – contains associated documentation on the Project
* Software – contains relevant software
* Source – contains the TALEND source code

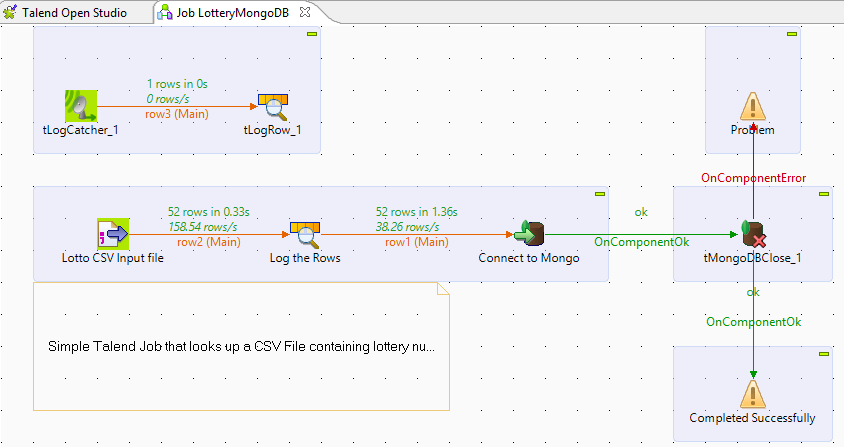
## Setup

Steps for getting set up and running:

1. Download and install the required Software identified above.
2. Setup required PATH variables for the Software
3. Connect and pull the repository <http://github.com/riotino/Lottery>. This will pull down this document, Talend Source Code and Java Executables
4. Assuming MongoDB has been installed start up a running instance of *mongod*.



1. Open up Talend Big Data and Import the Talend Project (source/INITIALDBPROJECT). Change the Context of the Talend job. Running the project should load the CSV file and push it into the Mongo Database.



1. Open up a *mongo* shell command line instance. This should have loaded the last 6 months of data from the downloadResults.csv file into Mongo.

## Factors

Figure : Some of the factors associated with drawing Lotto Numbers