## IEOR 4727 Project: Missing Historical Data and Insert it Back

Group Members: Guoan Wang, Ellen Sun

1. I. Code Compilation
   1. OS, library, and other set ups
      1. Operation System: Mac
      2. ElasticSearch: Version 2.4.6.
         1. Download: <https://www.elastic.co/downloads/past-releases/elasticsearch-2-4-6> (For Mac, download .tar)
         2. Decompress the file.
         3. Open terminal. cd <your path>/elasticsearch-2.4.6
         4. In terminal, type bin/elasticsearch (Now the ElasticSearch starts running)
         5. Download postman. <https://www.getpostman.com/>
      3. RInside Library
         1. 1. Install R on your machine, and install Rcpp package and RInside package in R. Ex: install.packages("RInside")
         2. Follow <https://www.youtube.com/watch?v=sjiSaaNA8BY> to link RInside package in XCode. One should be able to compile the following code if the library is linked successfully:

*#include <RInside.h> // for the embedded R via RInside  
 int main(int argc, char \*argv[]) {  
 RInside R(argc, argv); // create an embedded R instance  
 R["txt"] = "Hello, world!\n"; // assign a char\* (string) to 'txt'  
 R.parseEvalQ("cat(txt)"); // eval the init string, ignoring any returns  
 exit(0);  
 }*

* 1. How to run our code
     1. Install ElasticSearch, let it start running.
     2. Install RInside and RCpp. Link.
     3. Put all the RData files in the same directory.
     4. Open main.cpp, in the main function, change RData file directory to the directory in Step3
     5. In main.cpp, look at row 88 to 93. If you want to test our code:
        1. If you want to test insert back missing data:
           1. In main.cpp, comment "UploadTestData(rhash);"
           2. Uncomment "TestCaseMissData(rhash);" and "InsertMissingData(rhash);"
           3. Run the code. This will create a random test case, and insert the missing data back.
        2. If you want to test correct inaccurate data:
           1. In main.cpp, comment "UploadTestData(rhash);"
           2. Uncomment "TestCaseFalseData(rhash);" and "InsertMissingData(rhash);"
           3. Run the code. This will create a random test case, and correct some intentionally modified datafields.
        3. If you want to upload all data from RData file to ElasticSearch:
           1. In main.cpp, uncomment "UploadTestData(rhash);"
           2. Comment "CreateTestCase(rhash);" and "InsertMissingData(rhash);".
           3. Run the code. This will read all RData files in the directory and upload all data onto ElasticSearch.
        4. If you already have some data on ElasticSearch and a corresponding RData file, and you want to insert missing data back:
           1. In main.cpp, comment "UploadTestData(rhash);" and "CreateTestCase(rhash);”
           2. Uncomment "InsertMissingData(rhash);".
           3. Run the code. This will check missing data and insert missing data back.
        5. If you only want to test the functionality of certain classes
           1. See TestCase.cpp for simple test cases of each individual class