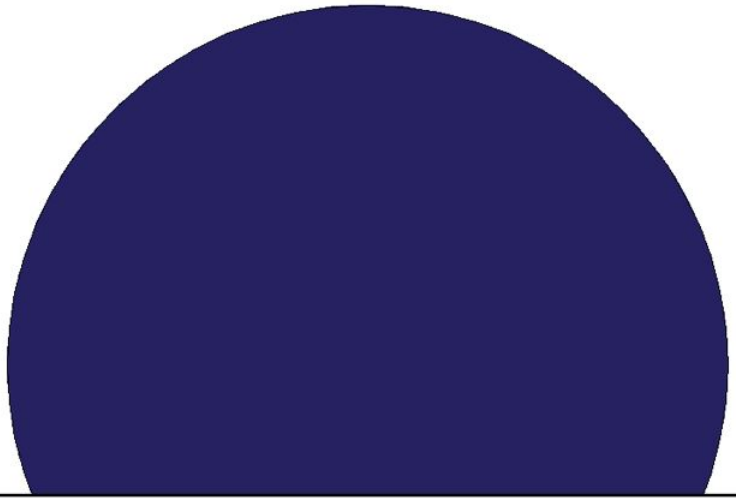

Business Classification Index,

Version: 1.0 R1

Released: Monday, October 16, 2017

Team Size: 3 Developers

Duration: 3-5 Days (full time)

Problem Statement:

For any benchmarking company, it is crucial to have accurate classification of businesses by sectors and subsectors. These are published for public consumption but are never available in a user friendly way or easy to Integrate to an application.

One such classification index is published by Thomson Reuters and an extract of various sectors and subsectors of such classification is provided in the attached spreadsheet.

This work sample requires you to write an application that can crawl through Thomson Reuters company ranking pages and map businesses back to the PermID as available in the attached spreadsheet.

For example: URL **https://www.reuters.com/sectors/industries/rankings?view=size&industryCode=179** maps to PermID **4294952822** or Hierarchical ID **58101020** present in row 1079 of the attached sheet.

You would be required to extract all Industries along with following information and map them back to the PermID or Hierarchical ID as in the spreadsheet.

1. Ticker
2. Name
3. Market Capitalization
4. TTM Sales $
5. Employees

Finally, for each company listed in the index, we need to extract information on executives involved with the business. This information is available in the People section of all Industries listed in the categorization table.

I.e. taking example of N**TT Docomo Inc** (first company in the link mentioned above), People information is available at https://www.reuters.com/finance/stocks/company-officers/9437.T. From this page, we need following information extracted

1. Name
2. Age
3. Since
4. Current Position
5. Description (BIOGRAPHIES)

The output format needs to be designed as part of the work sample, and should take into consideration usability of the data. It would also be ideal to have the URLs mapped back and present in the output in case a manual verification step is required to be carried out.

Evaluation:

Your submission will be evaluated on

1. **Solution Analysis:** This is would be how you have approached the problem, details considered before getting to write solution, any alternatives investigated etc. This will be subjective analysis and will be covered as part of the work sample discussion.

2. **Solution Implementation:** This is how the solution has been implemented, error handling, libraries used, efficiency of the solution, etc. (Off course,this will take into account the time available for the work sample, i.e. we expect reasonable quality solution and avoid expecting the ultimate solution.)

Important: To make evaluation accurate and less time consuming please take note of following.

1. **Code formatting:** Please indent the code (using tabs or space) and use camel case when defining variables and functions.
2. **Documentation:** Please document the code where necessary and document it just enough. Excessive documentation is worse than no documentation.
3. **Overall summary:** Please explain on how to build and execute the code. How and what inputs to be passed and where output can be seen. A general execution flow will also be good for the evaluator to understand the solution that is being provided. As mentioned above, this needs to be just enough to make evaluation effective and avoid excessive documentation.