**Project Plan – Forecasting Financial Foreign Aid**

Benjamin Arancibia

March 6th, 2015

IS661 – Independent Study

***Introduction***

Planning for financial foreign aid amounts are an important part of successfully planning and improving the socioeconomic status of citizens with a developing country. Three avenues were proposed in the project proposal and this document will outline a more definite project work plan for this exercise. As a reminder, the avenue chosen was to forecast financial foreign aid amounts using several country datasets.

The outputs of this document will be an exploratory data analysis and list of deliverables for the future of the project. All relevant data can be found on this Github repo: <https://github.com/bcarancibia/CUNY_IS_661>

***Exploratory Data Analysis***

Currently, there are nine datasets that can be used for the exploratory data analysis, with four being prepared. The ten datasets currently available for this exploratory data analysis are:

* Chad
* Cote d’Ivoire
* Haiti
* Kosovo
* Madagascar
* Malawi
* Moldova
* Timor
* Uganda
* Honduras

The three datasets that are being prepared and will have access to this week are:

* Nepal
* Burkina Faso
* Senegal

Making the grand total of available datasets at thirteen.

*Fields*

In terms of data each dataset has different number of fields (see attached r script) with the lowest of 14 fields and the largest of 111 fields. There are commitments and disbursements for each dataset (cash flows) and donor agencies linked to each funding flow. For projects that have the larger number of fields, there are project descriptions though not necessarily up to date, as well as, four date fields. The date fields are: Actual Completion, Actual Start, Proposed Completion, and Proposed Start Date.

The data completeness of the datasets varies from country to country. Moldova has the highest amount of data completeness (missing only 1% of fields) and Haiti is the highest missing around 20% of fields. Though this number is high for Haiti it makes sense, there are 2211 projects and 41 fields of data, while Moldova only has 377 projects and 17 fields of data.

The financial data for each country is relatively similar, but there are some differences that will require data cleaning. Almost all the countries record financial disbursements by year, but some countries i.e. Malawi record it by quarter. The other difference is that some countries have more historical data i.e. started recording aid in 2007, while others are more recent implementations and don’t have that level of historical data. Finally, the last aspect of the financial data is that two countries, Uganda and Haiti, have recorded their financial data with blanks. I need to follow up with the countries to determine if the blanks equate to 0 disbursements for that year.

*Data Clean Up*

There needs to be extensive data clean up in the data. There are fields in every country with multiple data entries in the field. As an example Timor has these donors listed today in one field:

- Asian Development Bank

- Government of Australia

- Ministry of Public Works

There needs to be a large data clean up to make it file, which includes merging all the dataset countries into one dataset. The final dataset will follow a similar data structure as found in the IATI dataset. This allows for easier data transformations.

*Workplan*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Activities | March 2015 | First Half of April 2015 | Second Half of April 2015 | First Half of May 2015 | Second Half of May 2015 |
|  | Data Clean Up |  |  |  |  |  |
| Forecasting Algorithm Research |  |  |  |  |  |
| Implementing Algorithm |  |  |  |  |  |
| Improving algorithm |  |  |  |  |  |
| Report Writing |  |  |  |  |  |

|  |  |  |
| --- | --- | --- |
| **Month/Phase** | **Activities** | **Expected Deliverables** |
| **March 2015** | Data Cleanup  Introduction writing of report  Forecast algorithm research | Cleaned data set  Introduction report done |
| **First Half of April 2015** | Implement first version of algorithm | Forecasting algorithm  Methodology written for report |
| **Second Half of April 2015** | Improving algorithm | Improved forecasting |
| **First Half of May 2015** | Improving algorithm | Improving algorithm  Writing report |
| **Second Half of Mary 2015** | Report writing | Report writing |
| **Final Report** | Final report due | Final Report |