Finding the Correlation between People Studying and Staying in Canada and Economic and Health Factors of their Source Countries using Semantic Web and Open Government Data

Alireza Manashty[1] and Christopher Baker[2]

[1]PhD Student, Department of Computer Science & Applied Statistics, University of New Brunswick, Saint John, NB, Canada and [2]Professor, Department of Computer Science & Applied Statistics, University of New Brunswick, Saint John, NB, Canada
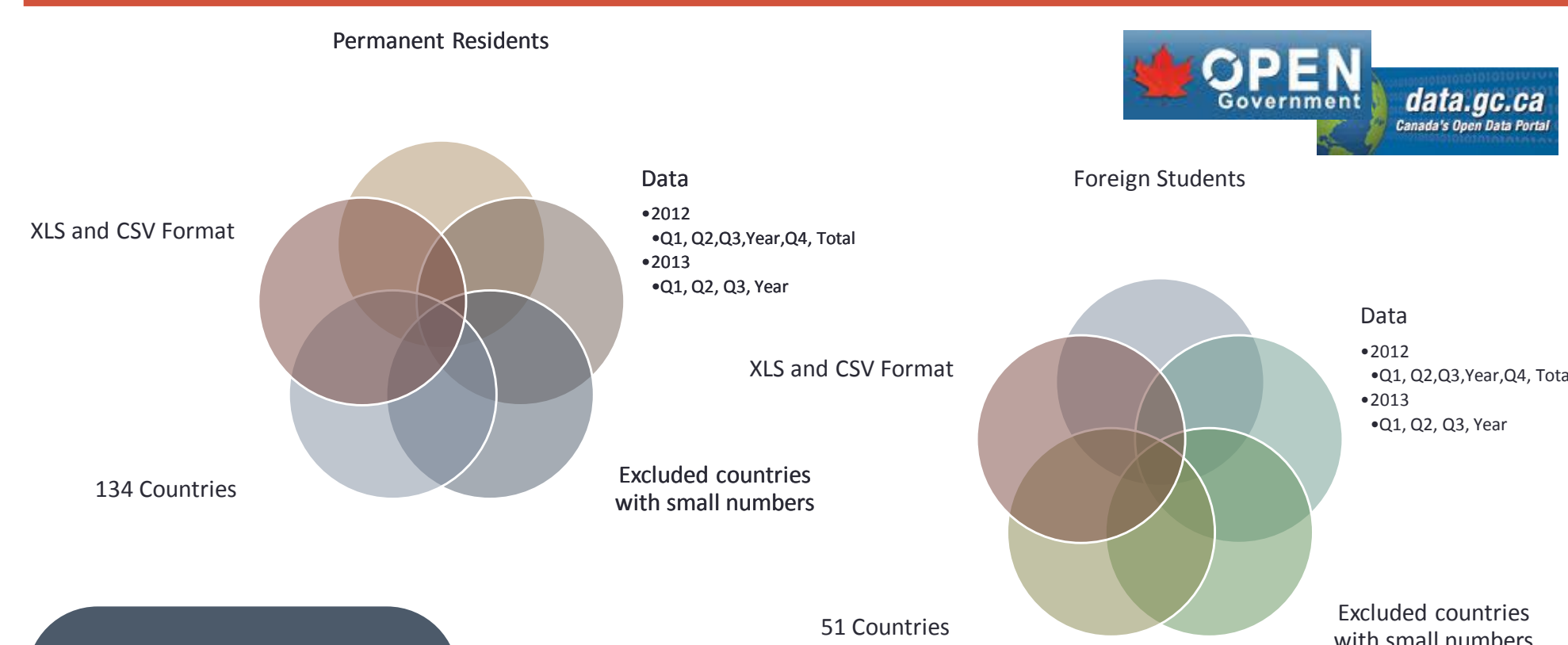
ABSTRACT

In this research, semantic web and open government data is exploited to find the correlation between the number of people studying and staying in Canada and the economic and health factors of the countries they are from. The recently released Canadian government open data on number of foreign students and permanent residents by source country is used as a trusted statistical source of data for this study. Using semantic web for integration of data from different sources, the government data is matched with the CIA world factbook data published as SPARQL endpoint to find the correlation between 9 normalized factors of each source country and the percentage of people from each country who come to Canada as foreign students or stay as permanent residents. Previous studies on immigration data focus mostly on the destination country and the new economic and social behavior of the countries and the immigrants, whereas studying the factors on the source countries of immigrants reveals contributing factors that can be used by policy makers and source countries governments to control or predict the flow of migrants. Results indicate that from the factors studied, poverty and deathrate have the highest negative migration effect on foreign students and permanent residents, respectively. Also, life expectancy in the source countries has a positive correlation with both groups, while higher infant mortality rate means less people studying or staying in Canada. Visualizing the data on the globe map perfectly shows this correlations specially on certain continents like Africa and Europe with contrasting infant mortality rate and life expectancy. Finally, the open government data is published as RDF to enable further studies access and integrate the data easier.

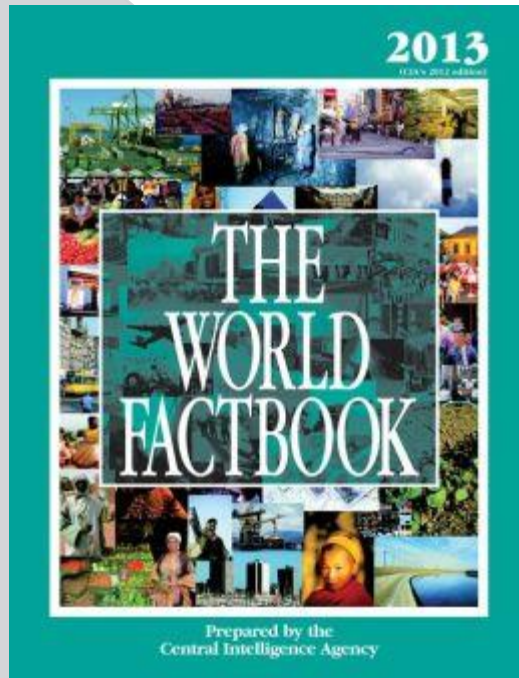Keywords: Semantic web, government open data, immigration

MOTIVATION

- Current Studies [1-8]
  - Focus of study on destination country economy (Canada, USA, Australia, etc.)
  - Considering immigrants characteristics, e.g. skills and investments
- Proposed Study
  - Focus on the statistical factors of the source countries
  - What is the causality from the source?
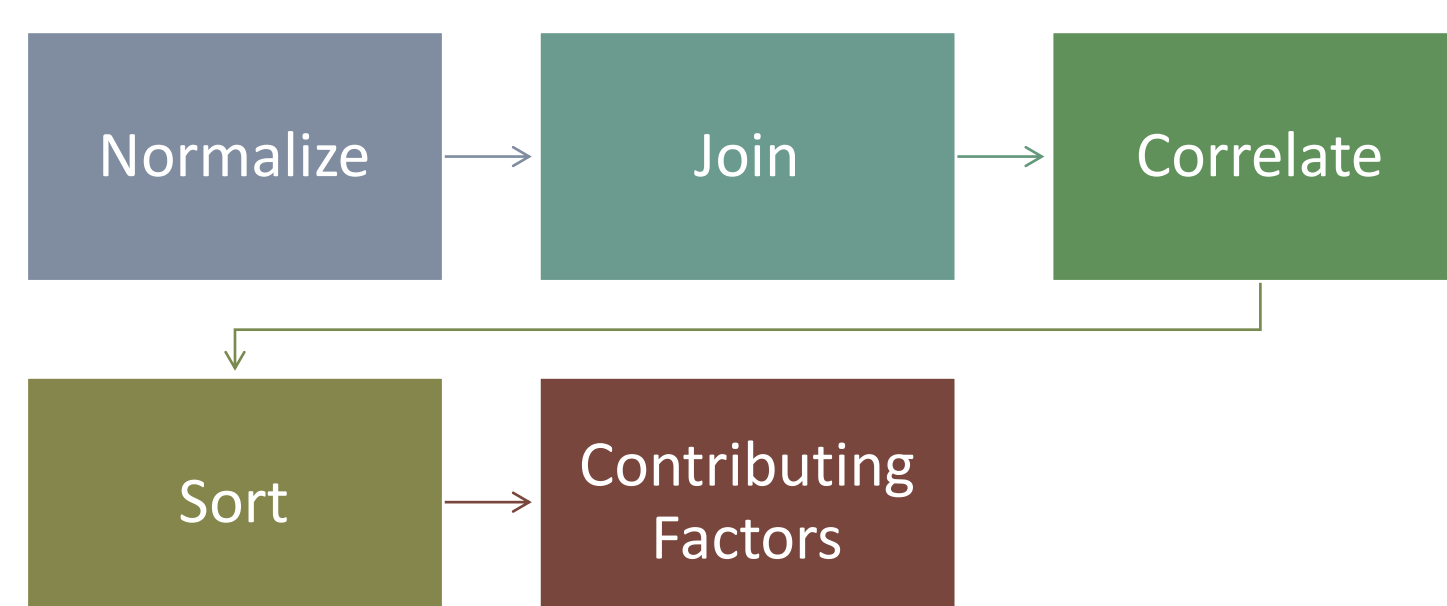
DATA

List of Factors
- Unemployment Rate
- PopulationBelowPovertyLine
- DeathRate
- InfantMortalityRate
- InflationRate
- LifeExpectancyBirth
- MedianAge
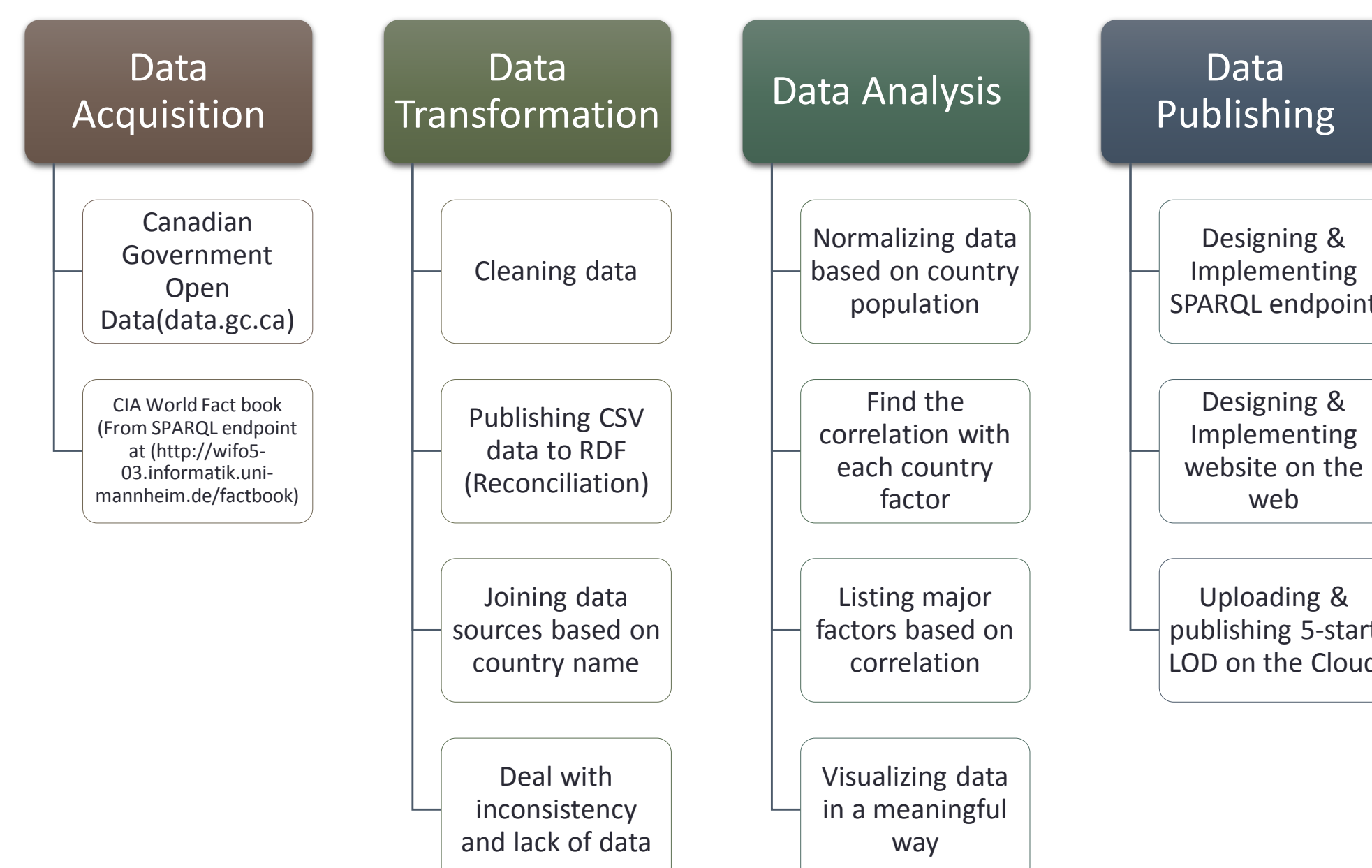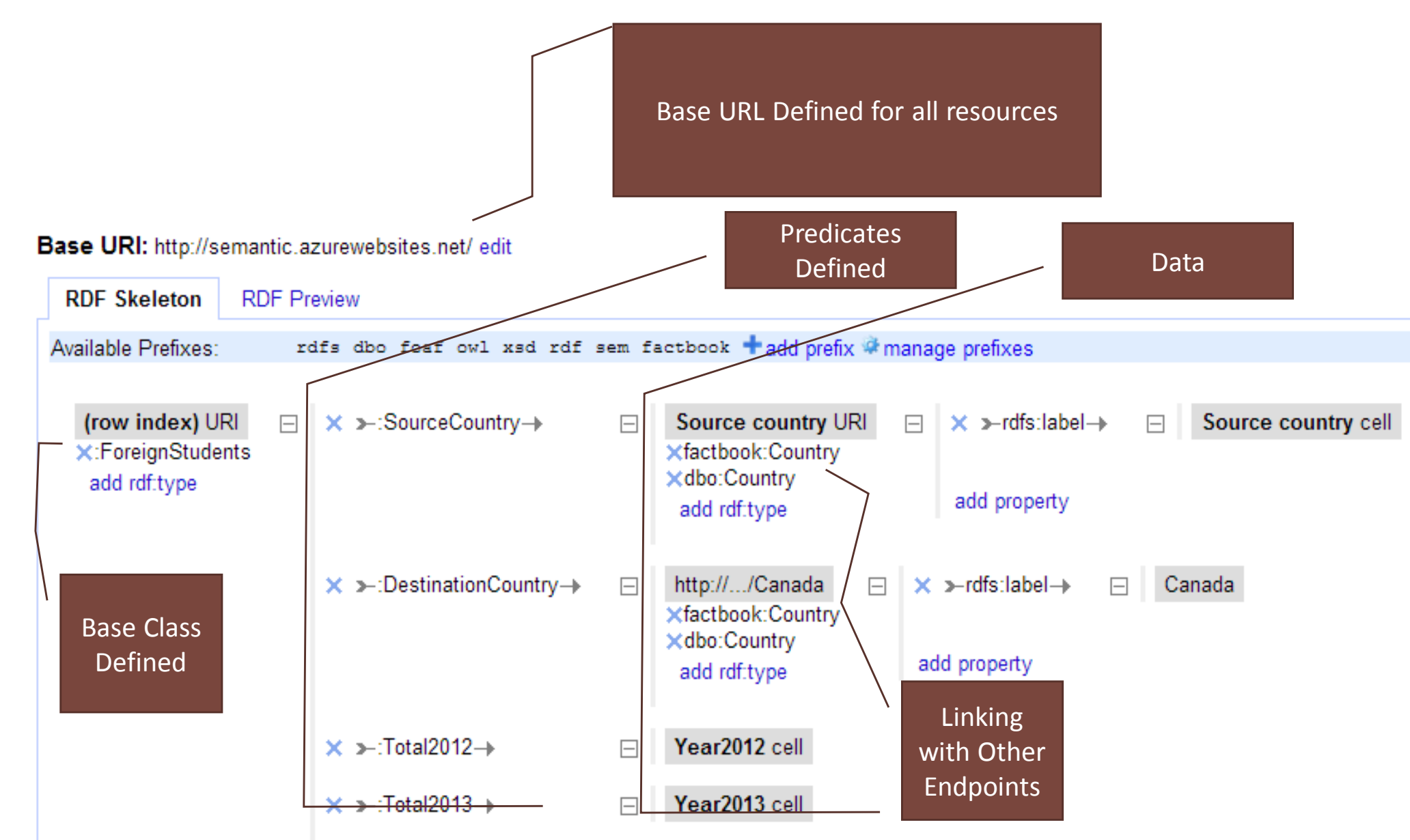- PopulationGrowthRate
- FertilityRate

ANALYSIS

OBJECTIVES

- Rating of the statistical factors of source countries on immigration to Canada
- Data Required
- Data on number of immigrants of Canada by source country
- Data on statistical factors of the source countries
- Integrating different data sources
- Analyzing the correlation between data sources
- Visualizing the data
- Publishing any unlinked data as 5-start linked open data endpoint (Online on Cloud not even local)
- Creating a web site UI for data manipulation
- Publish the result as a paper or poster
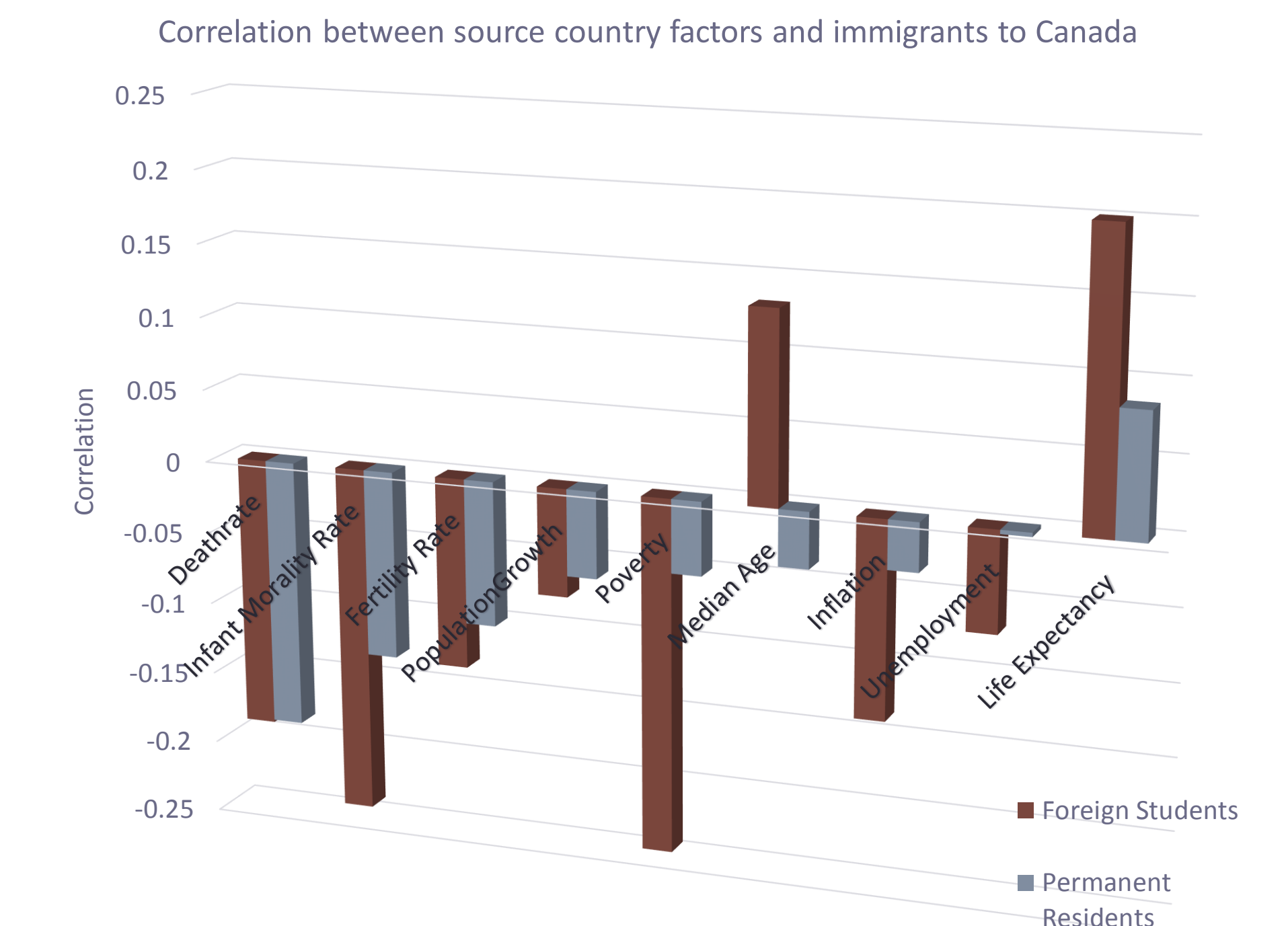
METHODS

PUBLISHING ON SEMANTIC WEB

RESULTS

- Life Expectancy has a major positive impact both on foreign students and permanent residents
- Infant Mortality Rate has a major negative impact on both
  - Africa has a few foreign students or permanent residents in Canada
- Poverty has the major negative impact on foreign students
- Deathrate has the major negative impact on permanent residents
- Correlation between major factors perfectly visible on the globe map
- The new point of view looks valid and important

RESULTS

Correlation between source country factors and immigrants to Canada

CONCLUSIONS

- Novel point of view and focus
- Data from trusted sources
- Nearly lossless data transformation
- Statistical analysis for factor correlations
- State of the art multidimensional visualization (Location, Three Major Factors, normalized data)
- Real 5-Star Linked Open Data provided
- Online cloud-based data publishing
- Customized code and web application

REFERENCES

- [1] L. V. Azarnert, "Is skilled immigration always good for growth in the receiving economy?," Economics Letters, vol. 108, pp. 116-118, 8// 2010.
- [2] A. Chassamboulli and T. Palivos, "The impact of immigration on the employment and wages of native workers," Journal of Macroeconomics, vol. 38, Part A, pp. 19-34, 12// 2013.
- [3] S. Fuller, "Up and on or down and out? Gender, immigration and the consequences of temporary employment in Canada," Research in Social Stratification and Mobility, vol. 29, pp. 155-180, 6// 2011.
- [4] L. González and F. Ortega, "How do very open economies adjust to large immigration flows? Evidence from Spanish regions," Labour Economics, vol. 18, pp. 57-70, 1// 2011.
- [5] L. E. Masselink and C. B. Jones, "Immigration policy and internationally educated nurses in the United States: A brief history," Nursing Outlook, vol. 62, pp. 39-45, 1// 2014.
- [6] S. Mocetti and C. Porello, "How does immigration affect native internal mobility? New evidence from Italy," Regional Science and Urban Economics, vol. 40, pp. 427-439, 11// 2010.
- [7] A. Saiz, "Immigration and housing rents in American cities," Journal of Urban Economics, vol. 61, pp. 345-371, 3// 2007.
- [8] B. Speciale, "Does immigration affect public education expenditures? Quasi-experimental evidence," Journal of Public Economics, vol. 96, pp. 773-783, 10// 2012.