

Big Data Recipes

Rebecca Oakes

September 2016

**Dissertation submitted in partial fulfilment for the degree of
Master of Science in Big Data**

**Computing Science and Mathematics
University of Stirling**

Abstract

The focus of the project “Big Data Recipes” is dealing with the observed lack of Big Data coding assistance presently existing on the World Wide Web.

The problem which this project was designed to tackle was one half of a larger project, a website which is designed to assist its users with Big Data coding issues. The aspect of this website that I am tackling was the creation of the code examples to begin population of this website. While sites like this do exist, they tend to be more of a general purpose site or tutorial sites for specific programming languages. As such, a site geared towards Big Data specifically would be of help to other users of Big Data.

I set out to achieve a large number of examples of a number of different programming languages, with each example being of a Big Data orientation. Each example contained information beyond the code example such as the programming language being used, an explanation of the code example and any packages that needed to be installed for the code example to work.

The Project plan I devised consisted of 3 distinct steps

- 1) Research and selection of appropriate languages and tools
- 2) Collaboration and consideration of the web site requirements
- 3) Execution of suitable examples

I started this project with ideas about various languages I could use and through research and discussion with my supervisor, decided on what languages I would use for the project. From there, study and research was done into each language so that I could decide which examples I was going to use in the project. Also I was in discussion with the fellow MSc student doing the other side of the project. In particular it was necessary for collaboration on how the examples were going to be structured. This was an important aspect for presenting the information to the users who would be using this site. Once I finished with research, I made my decisions on what examples would be executed for the website. Once I finished deliberating, I spent the rest of the time creating and formatting the examples for the websites.

I consider that the project has been successful achieving the aims set out at the beginning and I have been able to create a wide array of examples in 5 different programming languages. While I had hoped to use more languages, I completed using the minimum. The examples are not evenly spread across the languages but they were never intended to be. I was also able to create far more examples than I expected to in the time the project ran, a pleasing outcome showing the value of the approach selected.