Coding for Sustainability: Optimizing for Food Self-Reliance in the South West BC

Keystone Project

by Rachel Surtshin

Question: *How Can Big Data Inform Sustainability?*

Quest University Canada

April, 2018

**Introduction**

In the age of information we find ourselves with abundance data and an abundance of problems to solve. My question, *how can big data inform sustainability,* asks how we might create useful tools for meaningful interpretation of data to ameliorate 21st century dilemas, particularly those that threaten climate stability, resource availability, and environmental (opposite of pollution). I am interested in the ways that we can design a more sustainable future. Sustainability has a myriad of definitions. The definition widely accepted by WHO is defined as DEFINE.

In designing my keystone I was looked for a project that would allow me to practice computer and data science skills for a sustainability project. I wanted my finished product to be something needed and useful. My project was a computer and data management effort to optimize the local food system for increased food self-reliance.

Does food self-reliance = sustainability? Why or why not?

The Project

In 2014 the team at the Institute for Sustainable Food Systems at Kwantlen Polytechnic built a mathematical model of the food system in South West British Columbia, called the NAME. The NAME used national and provincial data sets to model and optimize for food-self reliance SOURCE. Food self-reliance is defined as DEFINITION.

The Rationale

The model was built to serve as a tool to inform policy decision as well as a quantitative basis for advocating for the economic development of resilient local food systems. The model was developed into a report called NAME, that was WHAT

The Problem

The model was built using Microsoft Excel and Open Solver. While Microsoft Excel is accessible to a wide variety of users the sheer depth of the model made a conglomeration of Workbooks that were complex and unwieldy. Modifying, comparing, and updating data sets is tedious, if not impossible. A modification of a single cell can break the model

My Keystone

For my Keystone I undertook a translation of the model from Microsoft Excel and Open Solver into a computer program written in the Python programming language. My goal was to transform the model into something more functional and user friendly. The first step was to collect all the data and write a Python code that cleaned and organized the data before executing the functions of the model. Download instructions for the data sets were created so that data sets could be retrieved in a reproducible manner. The second step was to streamline, simplify, and debug all the code, verifying the accuracy of the program. The third step was to add automated data acquisition features and build a graphical user interface (GUI) with integrated options for selecting different data sets.

The Process

My program was written in the Python programming language using Spyder Integrated Development Environment on the Anaconda Platform. My code made use of Numpy, Scipy, and Pandas libraries as well as the fuzzywuzzy library for fuzzy string matching. Much of my time was spent becoming familiar with the Python language, getting to know the data sets, and understanding some of the idiosyncrasies of the model.

How The Project Connects to My Question