Northeastern University

Course: DA5020

Assignment: Module 4 - Data Import - A

Total Points: 100

Date Due: Posted on Blackboard

Learning Objectives

In this assignment, you will learn how to:
- read and write Excel files
- shape data

Tasks

Before diving into the programming problems, study the data file "2013 Geographic Coordinate Spreadsheet for U S Farmers Markets 8'3'1013" that is provided for the assignment:

1. (25 Points) Load the data file into a data frame.

2. (50 Points) The seasons are not standardized and would make analysis difficult. Create six levels of seasons: Summer, Fall, Winter, Spring, Year-Round, Half-Year and convert each provided season in the **Season1Date** column to one of the season levels. Come up with reasonable rules, for example, June to August would be Summer, while 5/1 to 10/30 would be Half-Year. You need to use string processing functions to parse the strings and shape the data into the categories. If there are missing end dates, ignore the entire row.

3. (25 Points) Write a retrieval function *acceptsWIC()* that allows a data scientist to find which markets accept WIC. The function should output a data frame containing only the markets that accept WIC. Use the function and output the data frame.