Applied Data Science Capstone Project Overview

- **For the second week, the final deliverables of the project will be:**
  - A description of the problem and a discussion of the background. (**15 marks**)
  - A description of the data and how it will be used to solve the problem. (**15 marks)**

- **For the second week, the final deliverables of the project will be:**
  - A link to your Notebook on your Github repository, showing your code. (**15 marks**)
  - A full report consisting of all of the following components (**15 marks**):
    1) Introduction where you discuss the business problem and who would be interested in this project.
    2) Data where you describe the data that will be used to solve the problem and the source of the data.
    3) Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.
    4) Results section where you discuss the results.
    5) Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.
    6) Conclusion section where you conclude the report
  - Your choice of a presentation or blogpost. (10 marks)

Framing the problem: Background Information

A Changing Healthcare Landscape:

Healthcare delivery models are undergoing substantial changes given evolving reimbursement structures. As part of this process, patients are increasingly responsible for making financial decisions about their own health care as ever larger portions of healthcare costs are now the responsibility of the individual patient. Consequently, individuals seeking health services are acting more and more like consumers of other types of goods and services, particularly when it comes to drivers such as quality and convenience.

To meet the demands of savvy health consumers, in the past decade there has been a precipitous rise in so-called “doc-in-a-box” clinics and urgent-care centers which operate differently than either the traditional physician office or the acute-care hospital emergency room.

Layered atop this discussion are differences in medical care access between the current US system - a mesh of private/employer insurance and government pay (Medicare and Medicaid) – and the Canadian single-payer approach.

Problem Statement

Problem Statement:

This project will attempt to provide some information about the density of minute-clinic/urgent care between an American city and a physically proximate Canadian city – Seattle Washington and Vancouver, British Columbia. Both cities are in the “Pacific Northwest” region of North America., and have comparable overall population sizes (754,000 and 675,000 in city and 3.5 million and 2.5 million in the greater metro areas, respectively).

The Analysis:

Utilize the Foursquare API to create a visualization of ”urgent care” locations within the city limits of Seattle, WA and Vancouver BC and gain some possible insights into difference in care delivery from such locations between the two cities.

Data Discussion

Data Details:

- Using Foursquare, the data for this project will use a combination of three elements for the filter description:
  - "Urgent care"
  - And "Medical clinic"
  - And NOT "Hospital"

- Although the intent of this combination of filter keywords is to identify the types of healthcare delivery locations that are not stand-alone physician offices nor full-acute care hospital facilities, the result set could still be either too narrow or too broad given the great variation in care delivery. Nevertheless, this keyword combination should provide an adequate sample for the purposes of this data science project.
- The Folium package will then be used to provide a mapping which will provide a visual representation of the overall density of such health provider locations.