

# The Big Data Institute and Open Data

Rob Warren<sup>1</sup>

<sup>1</sup>Big Data Institute  
Dalhousie University

and

School of Mathematics and Statistics  
Carleton University  
(Adjunct Professor)

January 30th, 2014

CODE EVENT

## First ★★★★★ data set on the Canada Open Data Portal



Government  
of Canada

Gouvernement  
du Canada

Jobs

Immigration

Travel

Business

Benefits

Health

Taxes

More services

[Home](#) [All Services](#) [Open Government](#) [Open Data](#) [Search Open Data](#)

[First World War Medical Case Files as a Linked Open Data \(LOD\) Trial](#)

### First World War Medical Case Files as a Linked Open Data (LOD) Trial

A trial set of 1,000 scanned Canadian Expeditionary Force (CEF), First World War personnel files, has been used to create a Linked Open Data (LOD) set in Resources Description Format (RDF). The medical case sheet information from up to 3,000 pages from these personnel files was extracted using both human transcription and computer-driven quality control. The transcribed data generated has value for researchers in handwriting recognition, and archival and medical institutions. A collaboration between Library and Archives Canada (LAC) and The Muninn Project.

Have your say



[Rate this dataset](#)

[0 Comment\(s\)](#)

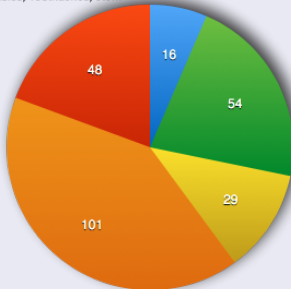
About this Dataset

[RSS Feed](#)

**Publisher**

## First ★★★★★ data set on the Canada Open Data Portal

- Veneral Diseases
- Shell Shock
- Measles, Toothaches, etc...
- Gun Shot and Shell Wounds
- Influenza and Respiratory Diseases

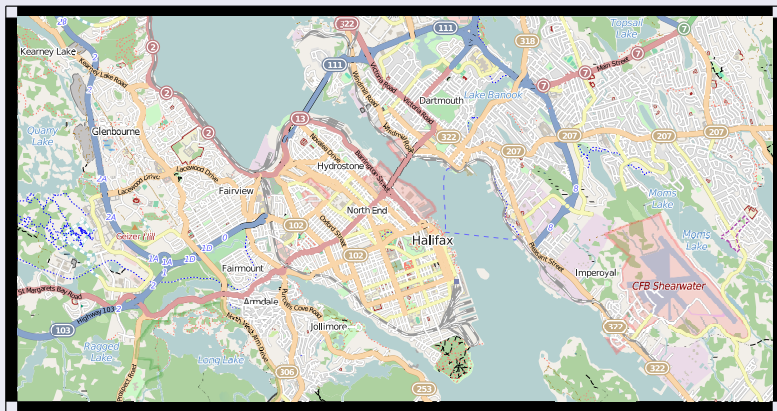


Coffee - the driving force behind our work!





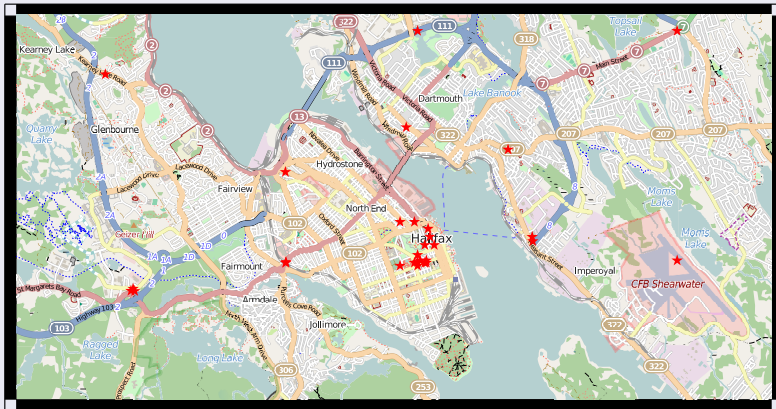
## Halifax Peninsula



a

<sup>a</sup>Open Street Map

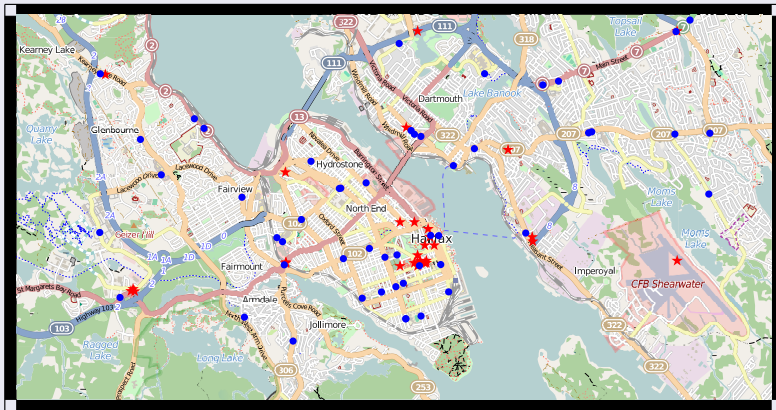
## Available real estate



a

<sup>a</sup>Local Realtors

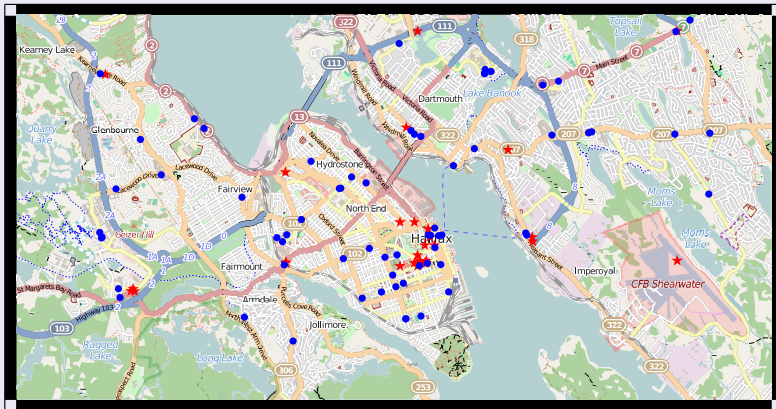
## Current competitors, chain #1



a

<sup>a</sup>Coffee Chain Website

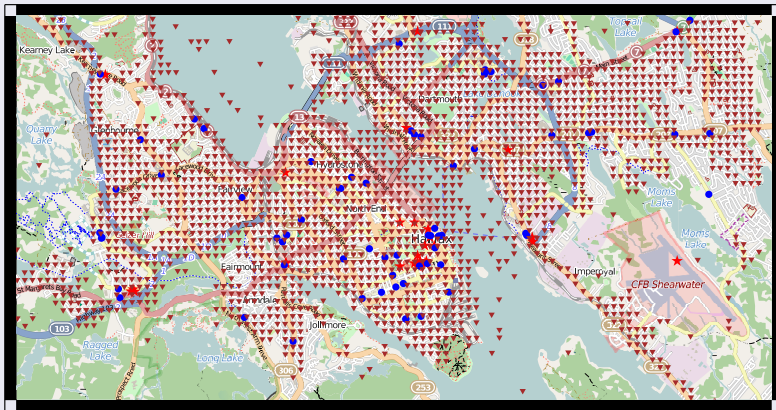
## Current competitors, chain #2



a

<sup>a</sup>Another Coffee Chain Website

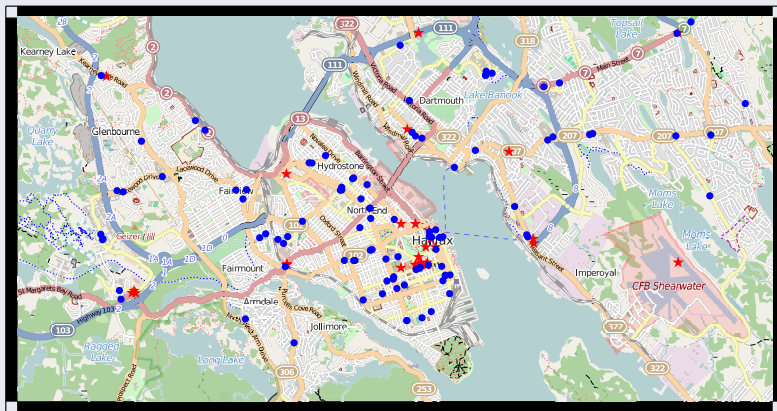
## Non chain cafe's and coffee houses; Wifi Signals?



a

<sup>a</sup>Wiggle.net

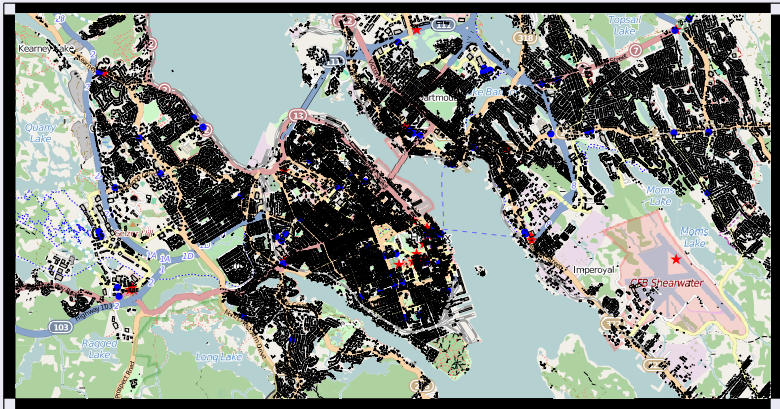
## Non chain cafe's and coffee houses



a

<sup>a</sup>Wiggle.net

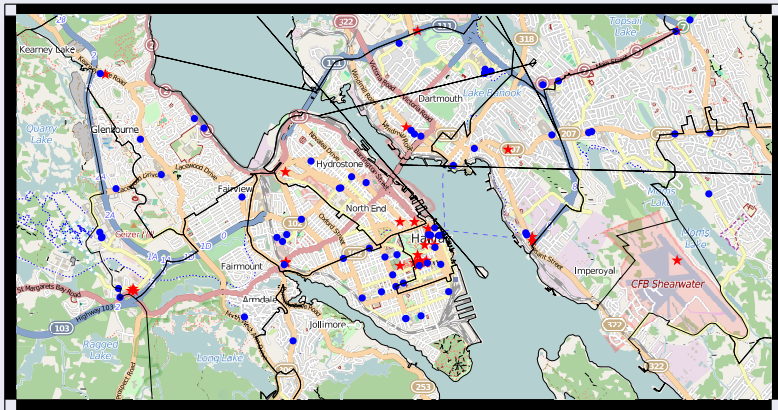
## Where are the customers?



a

<sup>a</sup>HRM Open Data Portal

## Where are the customers?

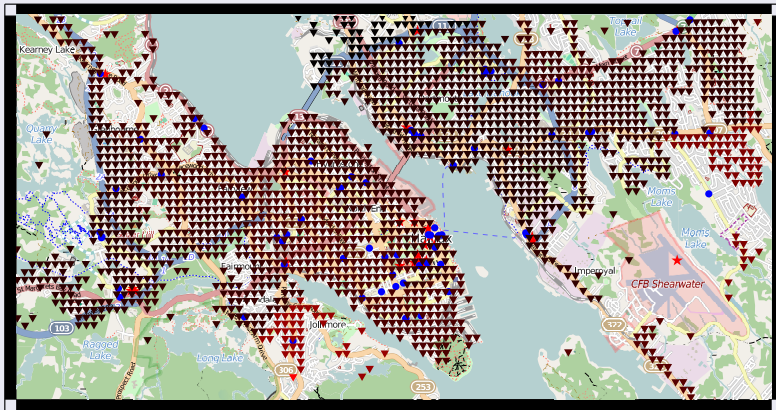


a

<sup>a</sup>Stats Can. / Open Data Portal



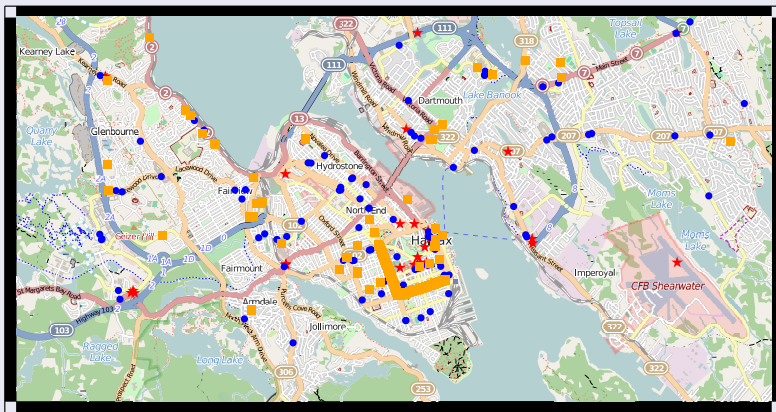
## Where are the customers?



a

<sup>a</sup>Stats Can. / Open Data Portal + Wiggle.net

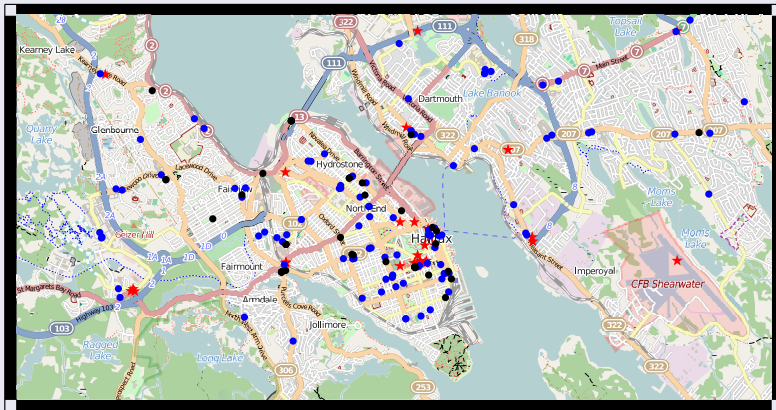
## Where are the customers?



a

<sup>a</sup>Data from KnowTime Halifax

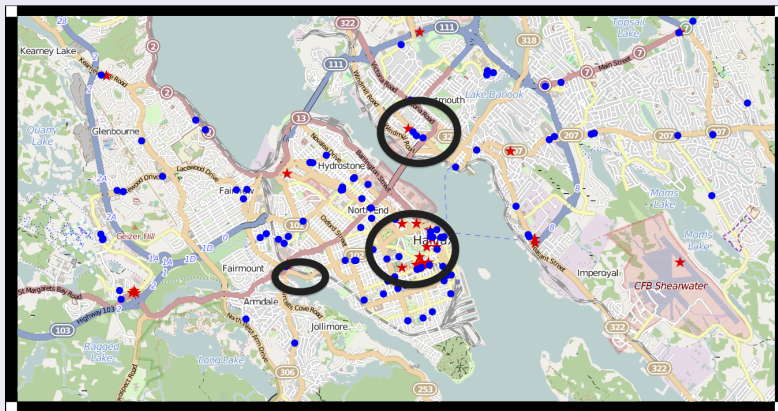
## Where have people failed before?



a

<sup>a</sup>Data from Nova Scotia Government

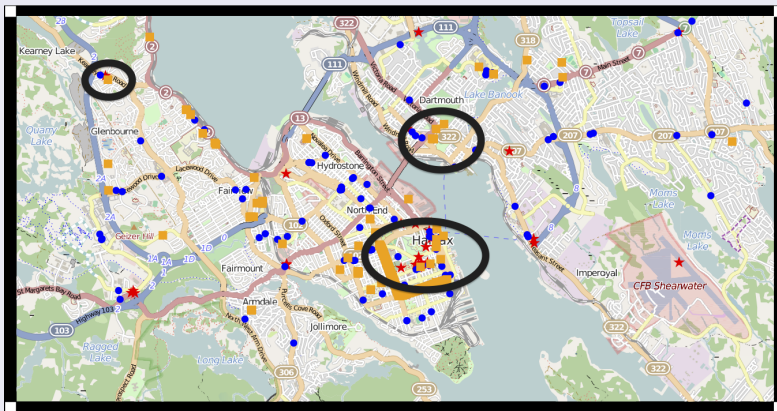
## Where to rent?



a

<sup>a</sup>Data from KnowTime Halifax

## Where to rent?



a

<sup>a</sup>Data from KnowTime Halifax

**Only one question remains!**

You have the data to do it, now what do you want to do?

Analytics