

# Modeling and Visualization of the COVID-19 Outbreak in Canada

September 10, 2020

Ngan Lyle MD, MSDS





# Background

- Started as a Capstone project at UBC Okanagan
  - Master of Data Science students
  - Mentored by Statistics Canada
  - Objectives for the Capstone project
    - Collect and organize data about the COVID-19 outbreak in Ontario
    - Show that open data can be used to produce meaningful analyses
- Current work
  - Refine and expand the statistical analysis
  - Replicate the analysis for other provinces



# Aims

- Model the spread of COVID-19 among Long Term Care (LTC) homes
  - LTC homes analysis
- Model the spread of COVID-19 among health regions
  - Health regions analysis

---

# LTC Homes Analysis for Ontario - Updated

# LTC Homes - Data

- LTC homes Covid-19 data
  - Government of Ontario Data Catalogue
    - <https://data.ontario.ca/dataset/long-term-care-home-covid-19-data>
- LTC homes characteristics
  - Reports on Long-Term Care Homes
    - Ontario Ministry of Health and Long Term Care
    - <http://publicreporting.ltchomes.net/en-ca/Default.aspx>
  - Health Quality Ontario
    - Ontario Health
    - <https://www.hqontario.ca/>

The screenshot shows the Ontario Data Catalogue interface. At the top is the Ontario logo and the text "Data Catalogue". Below this is a navigation bar with links for "Datasets", "Organizations", "Groups", "About", and "Help". A search bar is on the right. A yellow banner below the navigation bar says "Tell us what you think about our data and how you're using it. [Take our survey](#)". The breadcrumb trail shows the path: Home / Organizations / Treasury Board Secretariat / Long-Term Care Home ... The main content area has tabs for "Dataset", "Groups", and "Activity Stream". The "Dataset" tab is selected, showing the title "Long-Term Care Home COVID-19 Data" and a description: "This dataset compiles daily snapshots of publicly reported data on 2019 Novel Coronavirus testing in Ontario in Long-Term (LTC) homes." There is a "Keep updated" section with a "Subscribe" button.

The screenshot shows the "Reports on Long-Term Care Homes" page. It has a dark blue header with the title and two buttons: "HOME" and "SEARCH AGAIN". Below the header is a "Home Report" section. It states: "The LTC home listed on this screen is the result of your search. To view details on this home: Click on the corresponding tabs to view the Home Profile or Inspections for a LTC home." There is a house icon and the name "AFTON PARK PLACE LONG TERM CARE COMMUNITY" followed by its address, phone, and fax numbers, and a link to visit the LTC home's website. Below this is a tabbed interface with "HOME PROFILE" and "INSPECTION(S)". The "INSPECTION(S)" tab is selected, showing a table of inspections for the year 2020. The table has columns for "Inspection Type", "Inspection Report Date", and "Document". There are two rows of data. To the right of the table is a section titled "How to interpret these reports" and "Terms and Definitions".

Inspection Type	Inspection Report Date	Document
Complaints Inspection	Jun 12, 2020	<a href="#">Complaints Inspection Jun 12, 2020 - PDF (144 KB)</a>
Complaints	Jan 17, 2020	<a href="#">Complaints Inspection</a>

The screenshot shows the Health Quality Ontario website. It has a teal header with the logo and the tagline "Let's make our health system healthier". Below the header are three columns: "What is Health Quality", "System Performance", and "Evidence to Improve Care". At the bottom of the header is a dark blue bar with the text "COVID-19: Get the latest updates".

The screenshot shows the "SYSTEM PERFORMANCE" page on the Health Quality Ontario website. It has a breadcrumb trail: Home > System Performance > Long-Term Care Home Performance. The title is "Long-Term Care Home Performance in Ontario". The text below the title says: "These indicators provide data on wait times for admission to long-term care homes in Ontario."

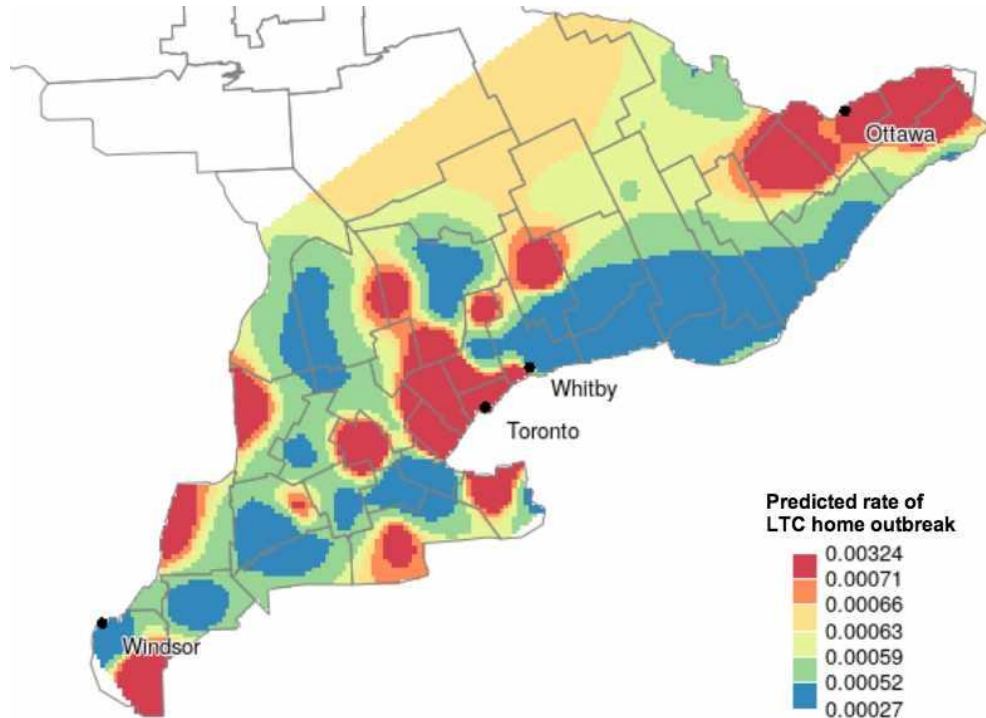
# LTC Homes - Logistic Regression

- 615/625 homes without missing data
- Binary outcome - outbreak status
- Predictors (18)
  - Numeric (12)
    - Number of beds
    - Quality (5)
    - Inspections (6)
  - Categorical (2)
    - Operator type
    - Region
  - Binary (4)
    - Residents council
    - Family council
    - Accreditation
    - Short stay

Variable	Estimate	Standard Error	P-value
Intercept	- 3.57	0.59	$1.28 * 10^{-9}$
Number of beds	+ 0.36	0.04	$<2 * 10^{-16} ***$
Total complaints/Number of beds	+ 2.90	0.96	0.0024 **
Total number non-complaints in the last 5 years	- 0.56	0.25	0.026 *
Municipal home type	- 0.61	0.26	0.020 *
Non-profit home type	+ 0.22	0.22	0.31
For-profit home type (ref)	(ref)	(ref)	(ref)

# LTC Homes - Geostatistical Model

- Binomial model with individual effect
- Focus on LTC homes in Southern Ontario



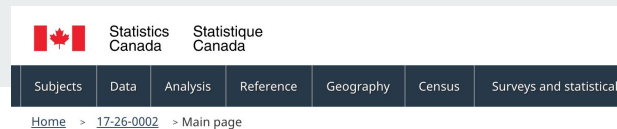
---

# Health Regions Analysis for Ontario - Updated



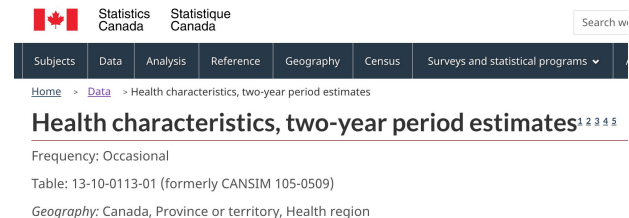
# Health Regions - Data

- Statistics Canada Data
  - Proximity data (10)
  - Health data (19)
    - Canadian Community Health Survey
  - **Census Profile data (26)**
    - **2016**
- Government of Ontario Data Catalogue
  - Confirmed positive cases of COVID-19 in Ontario
    - Cases
    - Fatalities

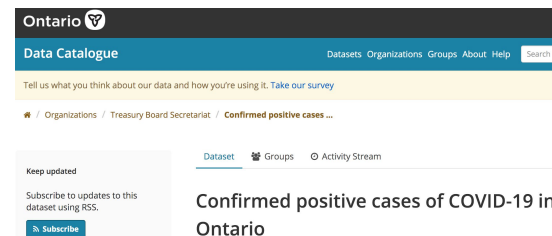


## Proximity Measures Database – Early release

Release date: April 17, 2020



## Download, Census Profile, 2016 Census



# Health Regions - Principal Components Analysis (PCA)

PC1 42.7%

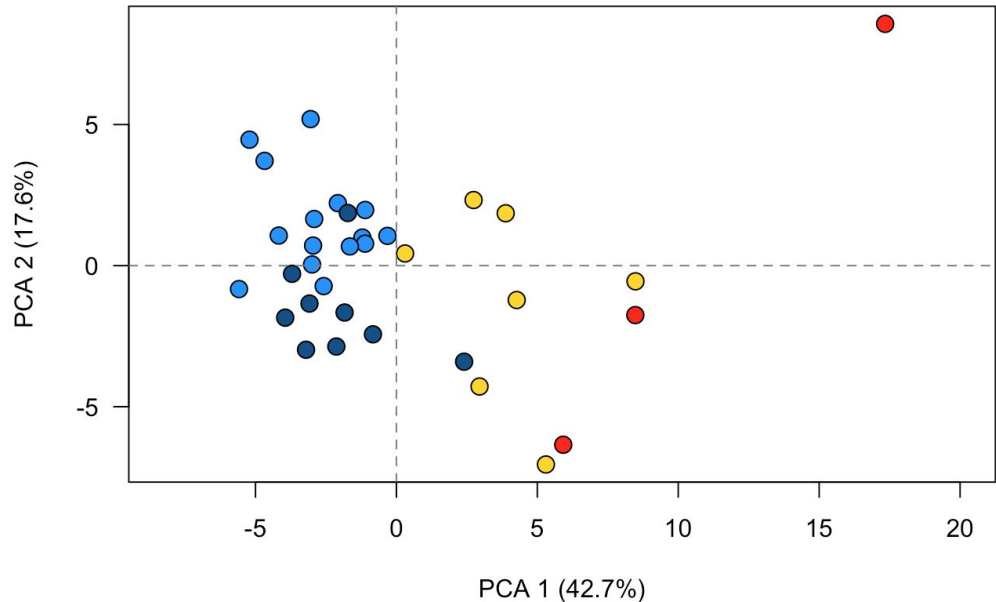


Variable type	Positive	Negative
Proximity	All 10 variables	None
Health	<ul style="list-style-type: none"> <li>• Life satisfaction</li> <li>• Perceived good health</li> <li>• Perceived good mental health</li> <li>• Perceived life stress</li> <li>• Physical activity &gt; 150 min per week</li> </ul> <p><b>Good health factors</b></p>	<ul style="list-style-type: none"> <li>• Medical problems</li> <li>• Mood disorder</li> <li>• Heavy drinking</li> <li>• Influenza immunization</li> <li>• Has a regular healthcare provider</li> <li>• Sense of belonging</li> </ul> <p><b>Poor health factors</b></p>
Census	<ul style="list-style-type: none"> <li>• Population density</li> <li>• Demographic                             <ul style="list-style-type: none"> <li>◦ Younger age</li> <li>◦ Single</li> <li>◦ Higher income</li> <li>◦ Higher level of education</li> <li>◦ Working</li> </ul> </li> <li>• Housing                             <ul style="list-style-type: none"> <li>◦ Renter</li> <li>◦ Mover</li> <li>◦ Non-suitable</li> <li>◦ Larger household size</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Older age</li> <li>• Married</li> <li>• Housing                             <ul style="list-style-type: none"> <li>◦ Detached</li> <li>◦ Owner</li> <li>◦ Non-mover</li> <li>◦ Suitable</li> </ul> </li> <li>• Driver</li> </ul>

# Health Regions - PCA

- Dots represent the health regions
- Similar health regions cluster together

More urban health regions;  
1. More connected  
2. Young, healthy, mobile population







# Health Regions - PC Regression

PCR on the proportion of cases

Principal Component	Estimate	Standard Error	P-value
Intercept	+ $1.43 \cdot 10^{-3}$	$1.33 \cdot 10^{-4}$	$2.84 \cdot 10^{-11}$ ***
PC1	+ $1.21 \cdot 10^{-4}$	$3.34 \cdot 10^{-5}$	$0.0012 \cdot 10^{-6}$ **
PC3	+ $1.83 \cdot 10^{-4}$	$7.19 \cdot 10^{-5}$	0.017*
PC5	+ $1.21 \cdot 10^{-4}$	$1.04 \cdot 10^{-4}$	0.015 *

Regions that are “connected” with a “young and healthy” population are at increased risk

---

# Replicate the Analysis for Other Provinces

# BC

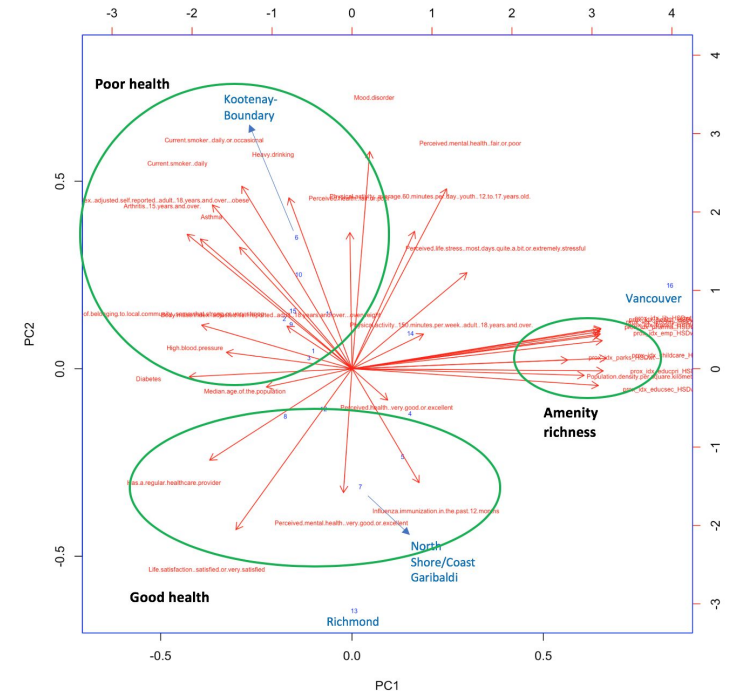
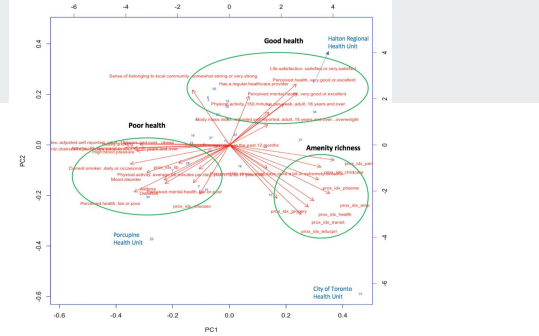
## • LTC Analysis

- LTC Homes data rich and complex
  - > 100 predictors but many missing values

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-5.14887	0.90350	-5.699	1.21e-08 ***
BEDS_TOTAL	0.37566	0.07486	5.018	5.21e-07 ***
FAMILY_COUNCILTrue	-1.00846	0.49201	-2.050	0.0404 *

## • Health Regions Analysis

- Regions are not as clearly separate based on health factors
- PC Regression showed a null result





## Future Work and Challenges

- Ongoing work with the Ontario and British Columbia data
- Quebec
  - LTC data are not accessible
  - No central registry of LTC homes
  - Only homes with an active outbreak are identified on government websites
- Alberta
- Manitoba





# Thanks!



OKANAGAN

- **UBCO team**
  - Prof. John Braun
    - KT Hobbs (BC/AB data)
    - Shahrukh Alvi (QC/MB data)
- **University of Toronto team**
  - Prof. Patrick Brown
    - Maggie Ma (Spatial Statistics)



# LTC Homes - Logistic Regression

Log odds of an outbreak =

- 3.57
- + 0.36 \* Number of beds
- + 2.90 \* Total complaints (normalized)
- 0.56 \* Total number non-complaints (5 years)
- 0.61 \* Municipal home type

