

Open Projects at DEIL *exploring the potential of open data and open approaches*

Data Exploration and Integration Lab (DEIL)
Statistics Canada

Open First Day
Ottawa, September 28, 2018



100

STATISTICS CANADA
ONE HUNDRED YEARS AND COUNTING

STATISTIQUE CANADA
CENT ANS BIEN COMPTÉS



Statistics
Canada

Statistique
Canada

Canada

Context

- Statistics Canada has a history of producing open data but the environment is evolving
- NSOs are increasingly becoming consumers and stakeholders of open microdata; there are new producers of open microdata
- Crowdsourcing pilot project was a first step in exploring use of open data, open tools, open code
- And now we want more!

Take home messages

- Great potential for integration and use of open microdata from (municipal, provincial, federal) government sources
- Great potential for conducting this work in open spaces as an “open project”



1

Crowdsourcing pilot

2

Open microdata from government sources

3

***Open Projects* and collaborative data creation**



1

Crowdsourcing pilot

2

Open microdata from government sources

3

***Open Projects* and collaborative data creation**

Crowdsourcing Pilot: Overview

- Collaborated with OSM and local municipalities
- Built open-source editor and mobile app
- Virtually complete map of pilot area (Ottawa-Gatineau) achieved

MAPPING YOUR COMMUNITY THROUGH CROWDSOURCING

Statistics Canada is launching a new pilot project, and we need your help!

Crowdsourcing data collection was suggested by a Statistics Canada employee as a way to improve how we collect information. We are now the first to collaborate with us on this project by mapping buildings in Canada. Participating in this data-gathering project will improve your community's knowledge and understanding of buildings and their uses and help to inform future community and regional planning and policy.



Statistics
Canada

Statistique
Canada

www.statcan.gc.ca

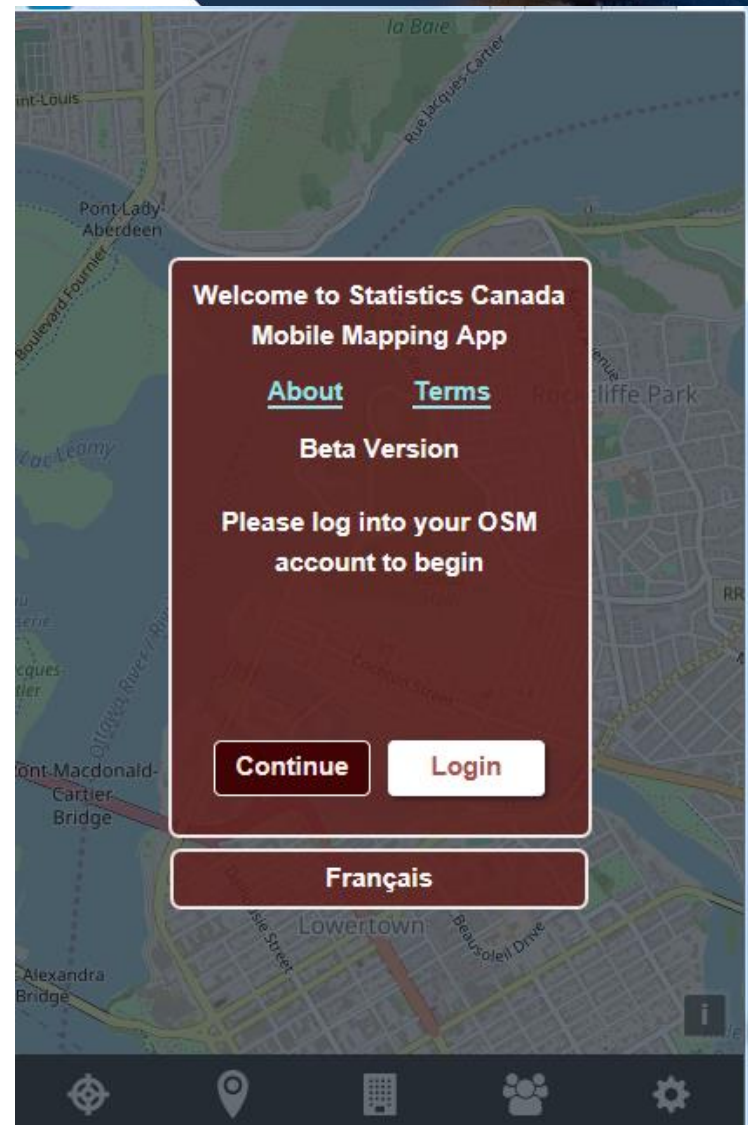
Canada

Crowdsourcing Pilot: Mobile App

- Created app to facilitate collecting data
- Used open source/code approach




[CODAP on GitHub](#)



Statistics
Canada Statistique
Canada

www.statcan.gc.ca

Canada



**Also, how much does a
gram of weed cost in
Ontario?**

Crowdsourcing cannabis prices

- \$7.37 per gram in ON
- Over 20k submissions
- Microdata available for download

➔ [StatsCannabis - Demo](#)

 StatsCannabis FRANÇAIS

Statistics Canada needs your help
collecting cannabis prices

Tell us about your last purchase of dried cannabis –
submissions are anonymous!

[Your data is protected](#)

I PAID

\$ Price

FOR

Quantity

OF

Quality

IN

City

TO

Primary Reason

 Submit

Cannabis prices in Canada

(Cannabis for medical and non-medical purposes, per gram)



Region	Price per gram
B.C.	\$6.88
Prairies	\$7.14
Ontario	\$7.37
Quebec	\$5.83
Atlantic	\$7.05
Territories	\$10.81
Canada	\$6.82

Source: Crowdsourced data, January 25 – September 7, 2018

[Show data table](#) [Download submitted data](#)

[Contact us](#) | [Terms and conditions](#) | [Privacy](#)



1

Crowdsourcing pilot

2

Open microdata from government sources

3

Open Projects and collaborative data creation

Experimental Open Database of Buildings “The ODB”

100

11



Statistics
Canada

Statistique
Canada

www.statcan.gc.ca

Canada

The ODB: *Details*

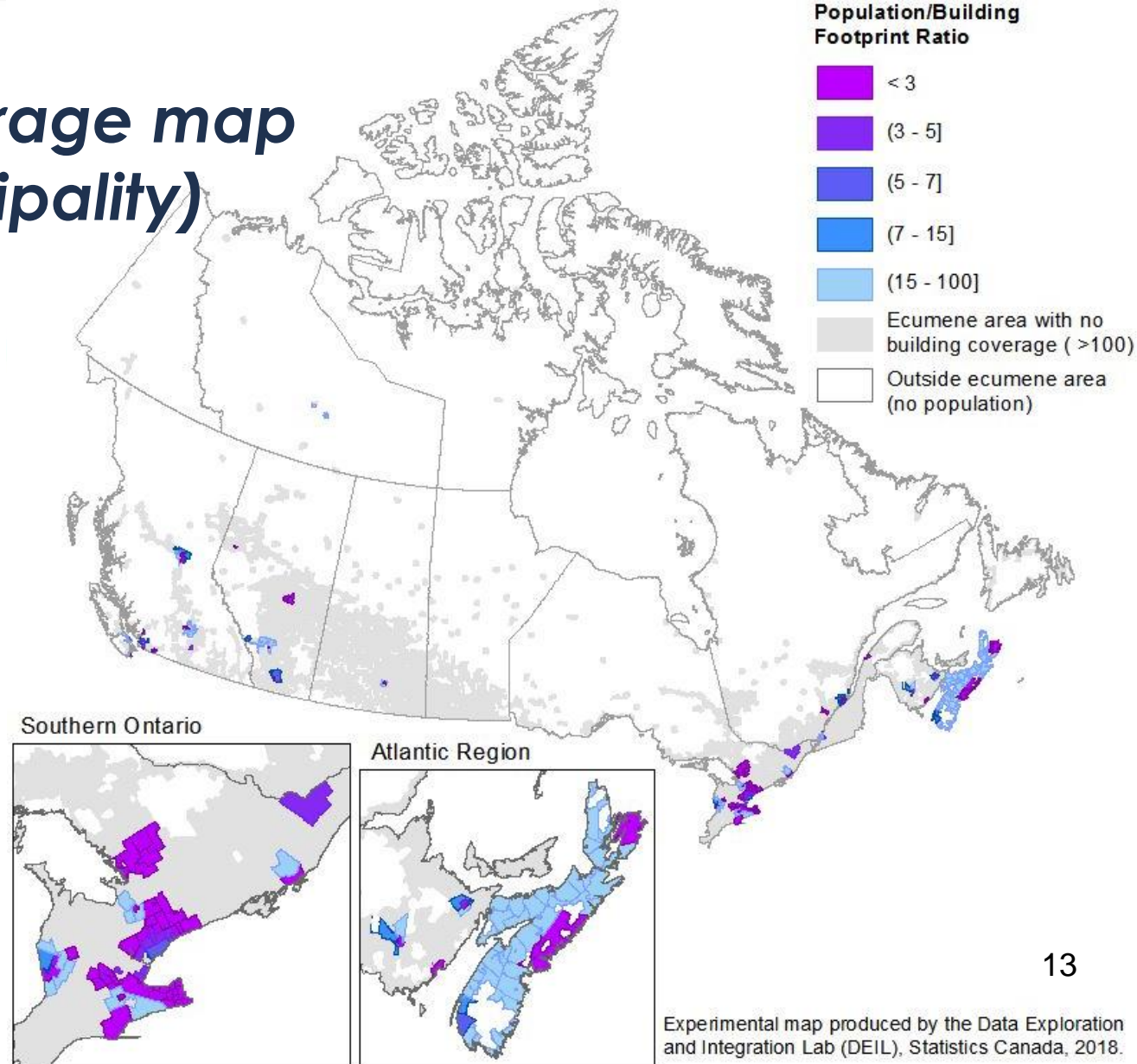
- 62 governmental datasets
- 4.3 million footprints (and counting)
- 100% PURE CANADIAN OPEN MICRODATA!

Open Data / Open License
**means everyone can access
and use the information**



The ODB: coverage map by CSD (municipality)

- 8 Provinces / Territories
- 40 CMA/CAs
- Urban and rural areas



13



Statistics
Canada

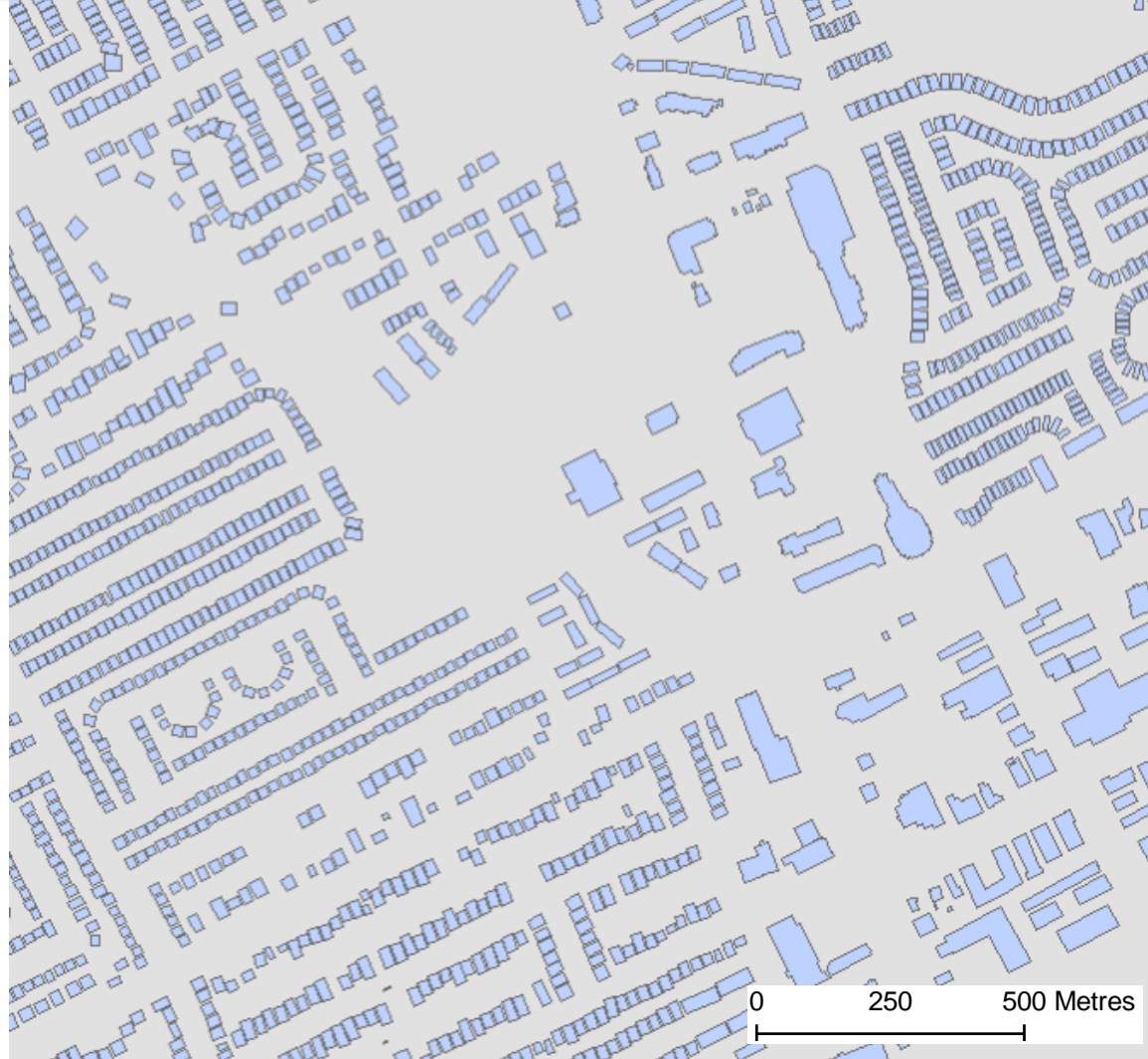
Statistique
Canada

www.statcan.gc.ca

Canada

The ODB: *example of the data*

- Ex: Footprints for Richmond Hill, Toronto
- Quality seems high, buildings are tightly knit
- Working to make ODB available for download



FID	Shape *	Longitude	Latitude	CSDUID	CSDNAME	Data_prov	Build_ID	Shape_Leng	Shape_Area
0	Polygon ZM	-79.726934	44.282285	3543021	Essa	Barrie	35430210000000	33.565106	63.971128
1	Polygon ZM	-79.699185	44.281962	3543017	Innisfil	Barrie	35430170000000	72.635583	269.749831
2	Polygon ZM	-79.726066	44.282264	3543021	Essa	Barrie	35430210000000	73.317978	278.524173





1

Crowdsourcing pilot

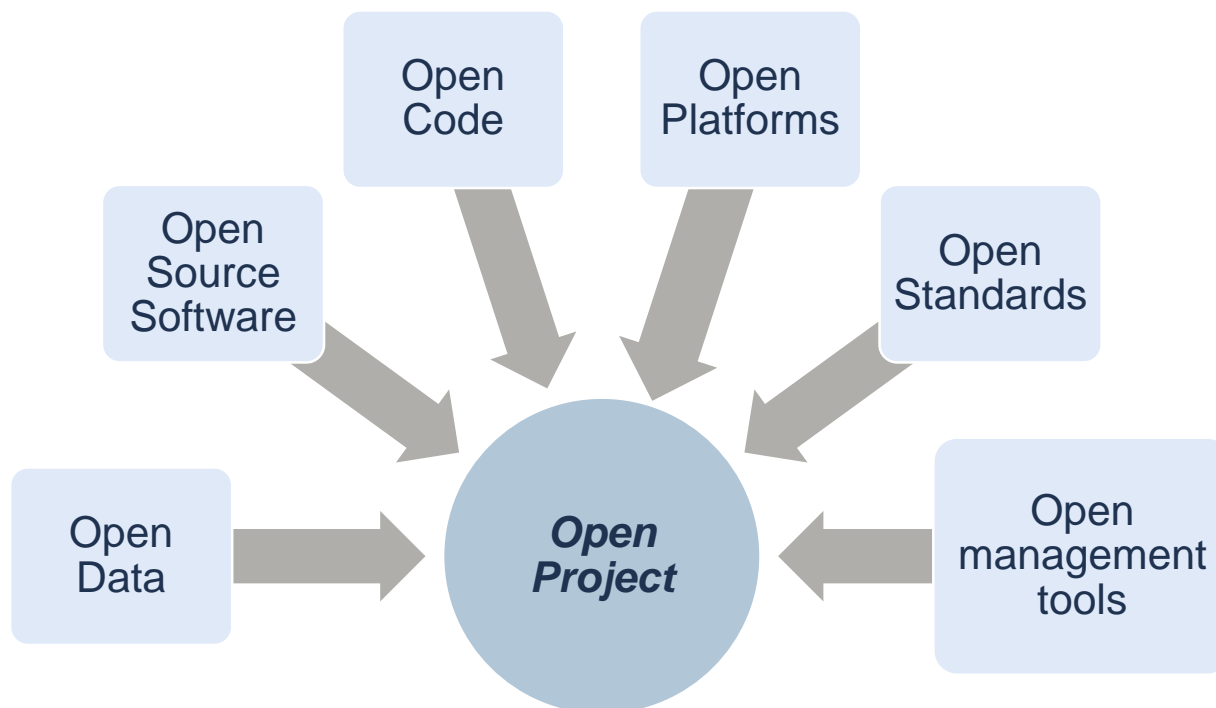
2

Open microdata from government sources

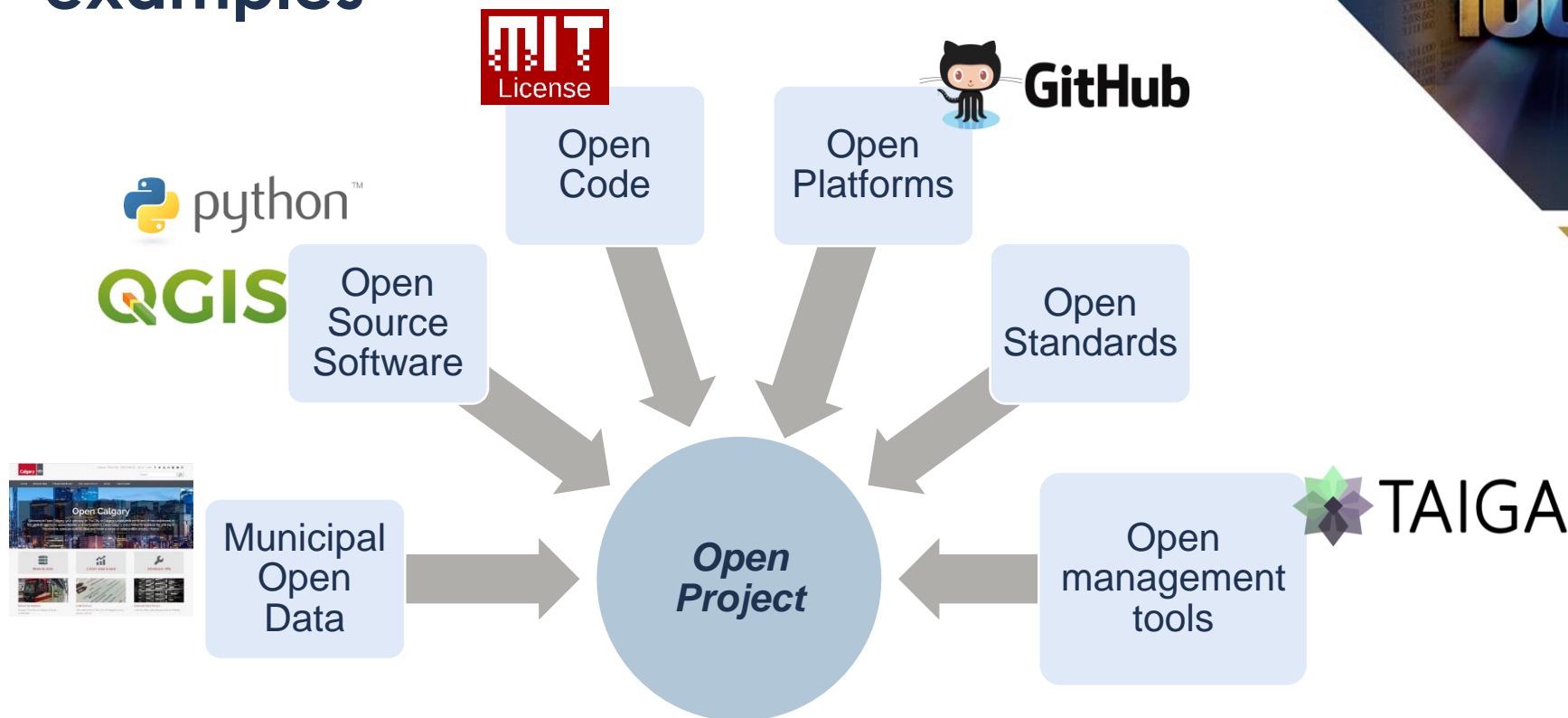
3

Open Projects and collaborative data creation

The *Open Project* concept



The Open Project concept: examples



Data development, an example: Open Businesses Repository “The OBR”

18

Collaborative data creation

- Data are becoming a basic infrastructure
- Partner across entities (private, public, community) for data creation
- Ex: OpenAddresses
- *Open Project* approach is an enabler
- Ex: Collaboration with Microsoft



19



Statistics
Canada Statistique
Canada

www.statcan.gc.ca

Canada

The ORB: *Details*

- Exploratory and experimental open project
(Note: work in progress)
- OBR is an open software solution \neq database
- Government/authoritative sources of open microdata on businesses
- OBR script (on GitHub) compiles, processes, integrates

20

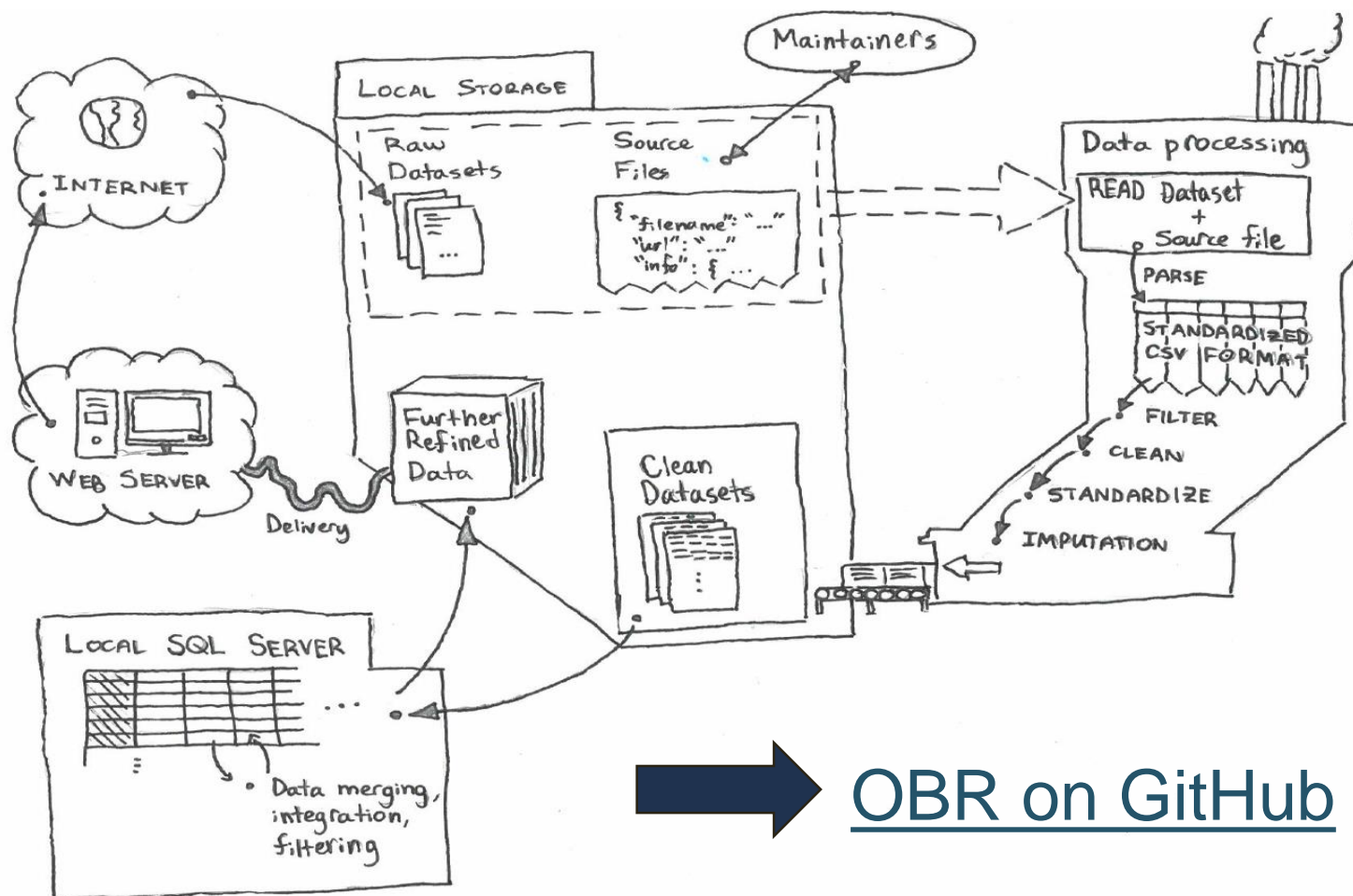


Statistics
Canada Statistique
Canada

www.statcan.gc.ca

Canada

The OBR



21



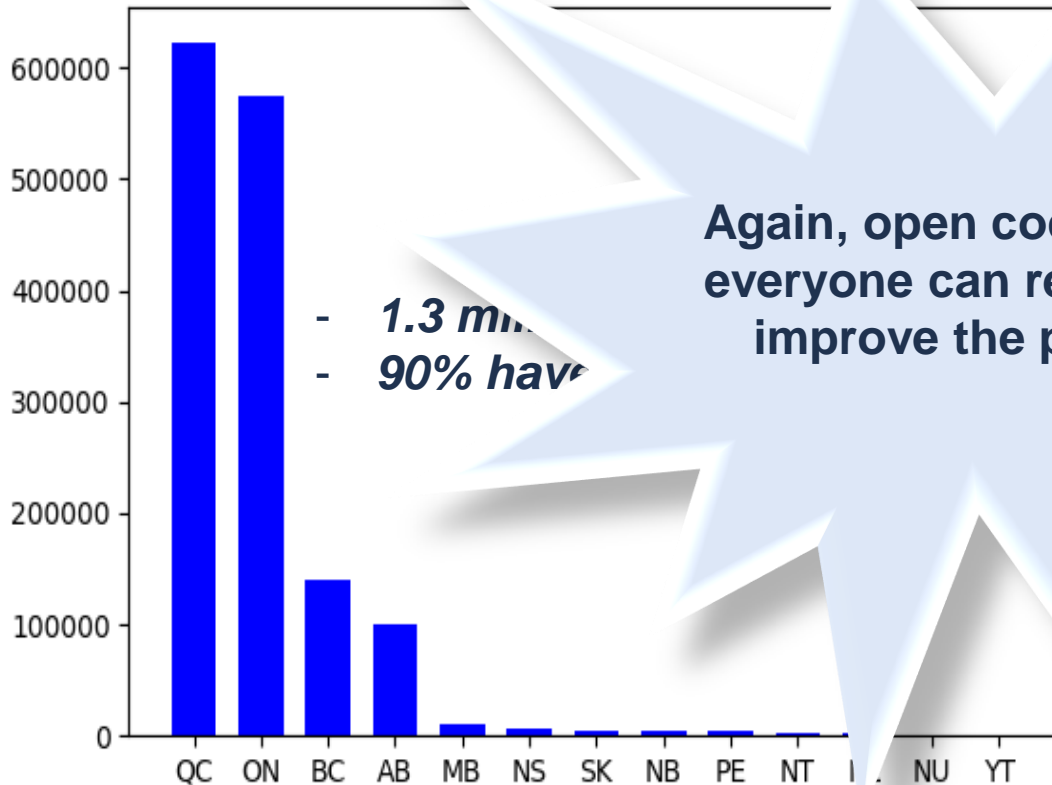
Statistics
Canada Statistique
Canada

www.statcan.gc.ca

Canada

The OBR: *preliminary results*

100



- 1.3 million
- 90% have

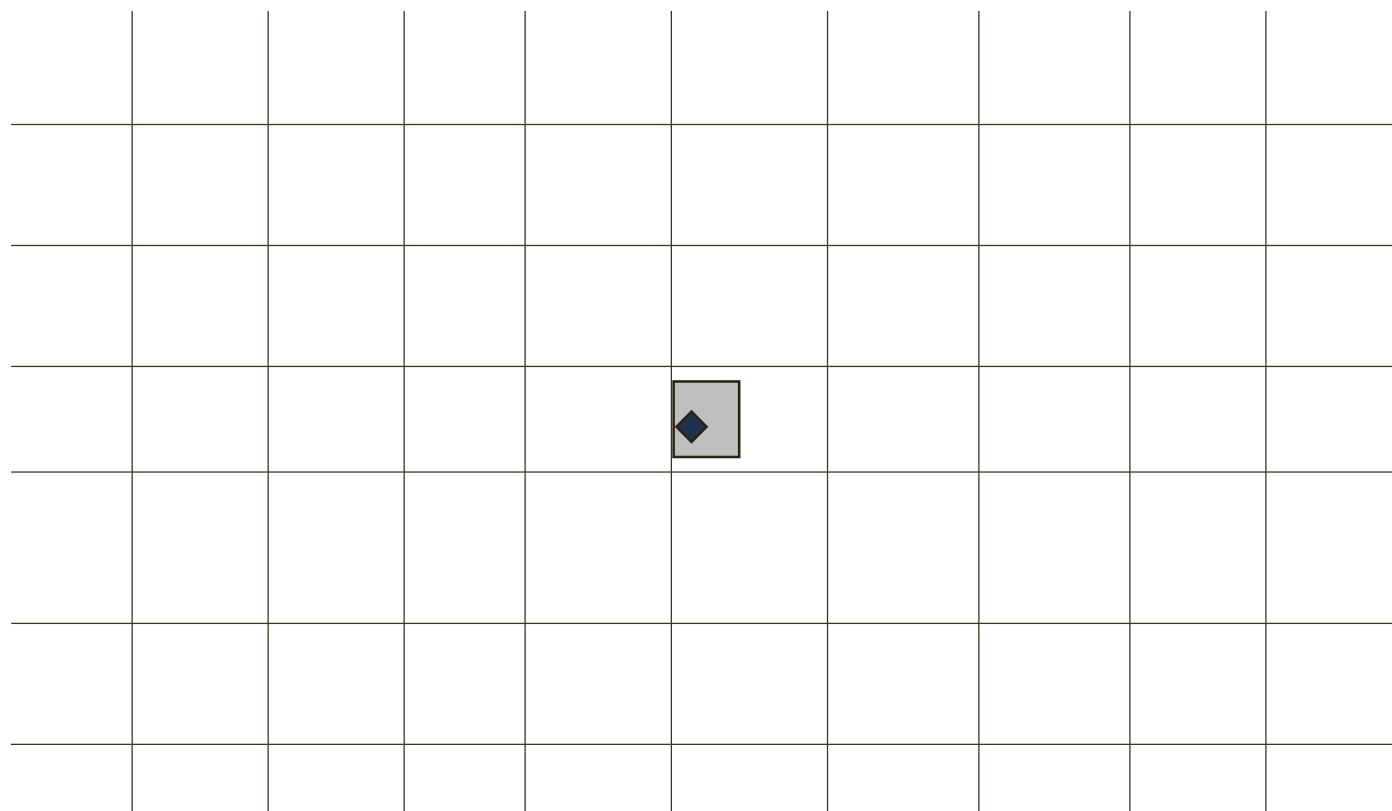
Again, open code means
everyone can re-use and
improve the product

Data integration and analysis, an example: Proximity measures

What can we do with the data?

Example - Proximity Measures

100



24



Statistics
Canada

Statistique
Canada

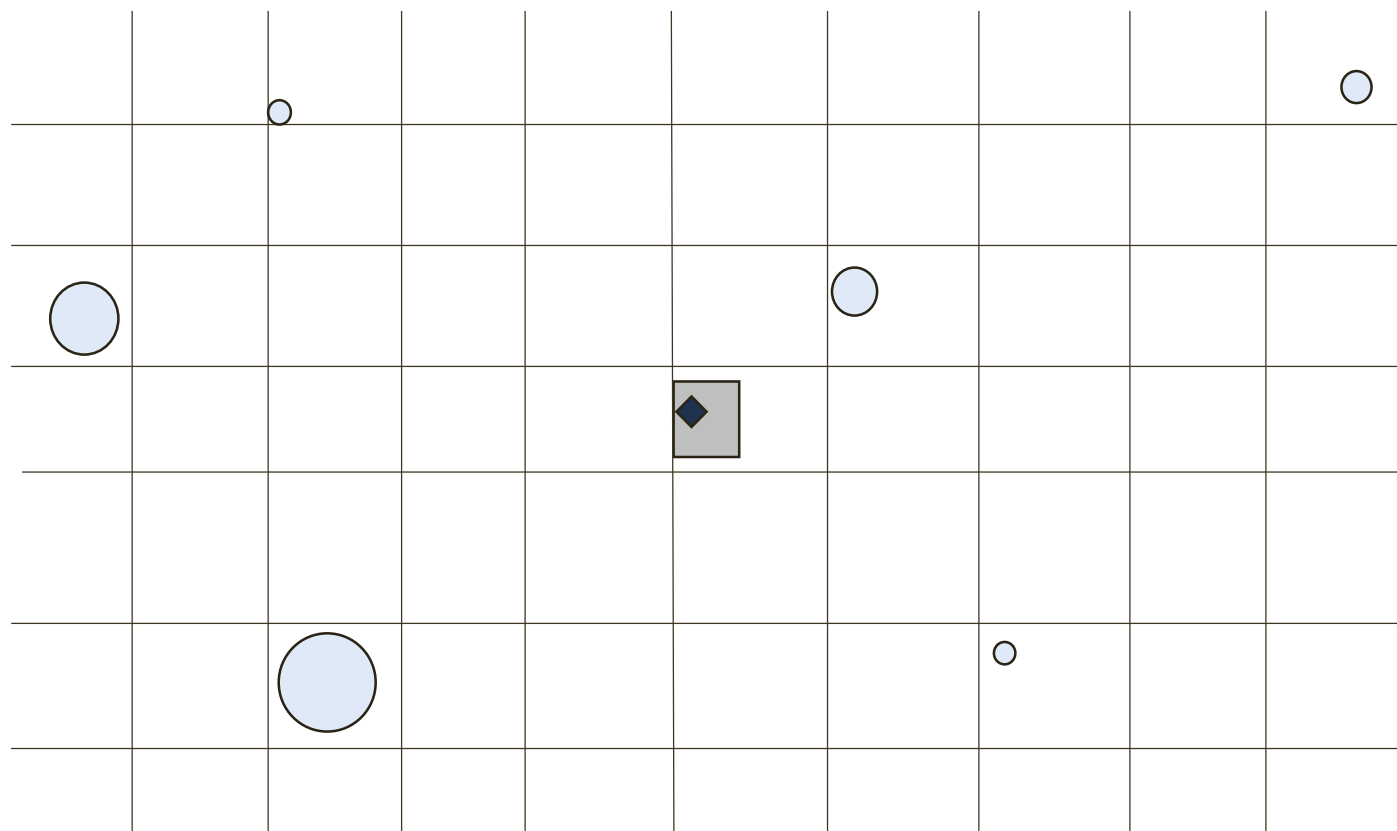
www.statcan.gc.ca

Canada

What can we do with the data?

Example - Proximity Measures

100



25



Statistics
Canada

Statistique
Canada

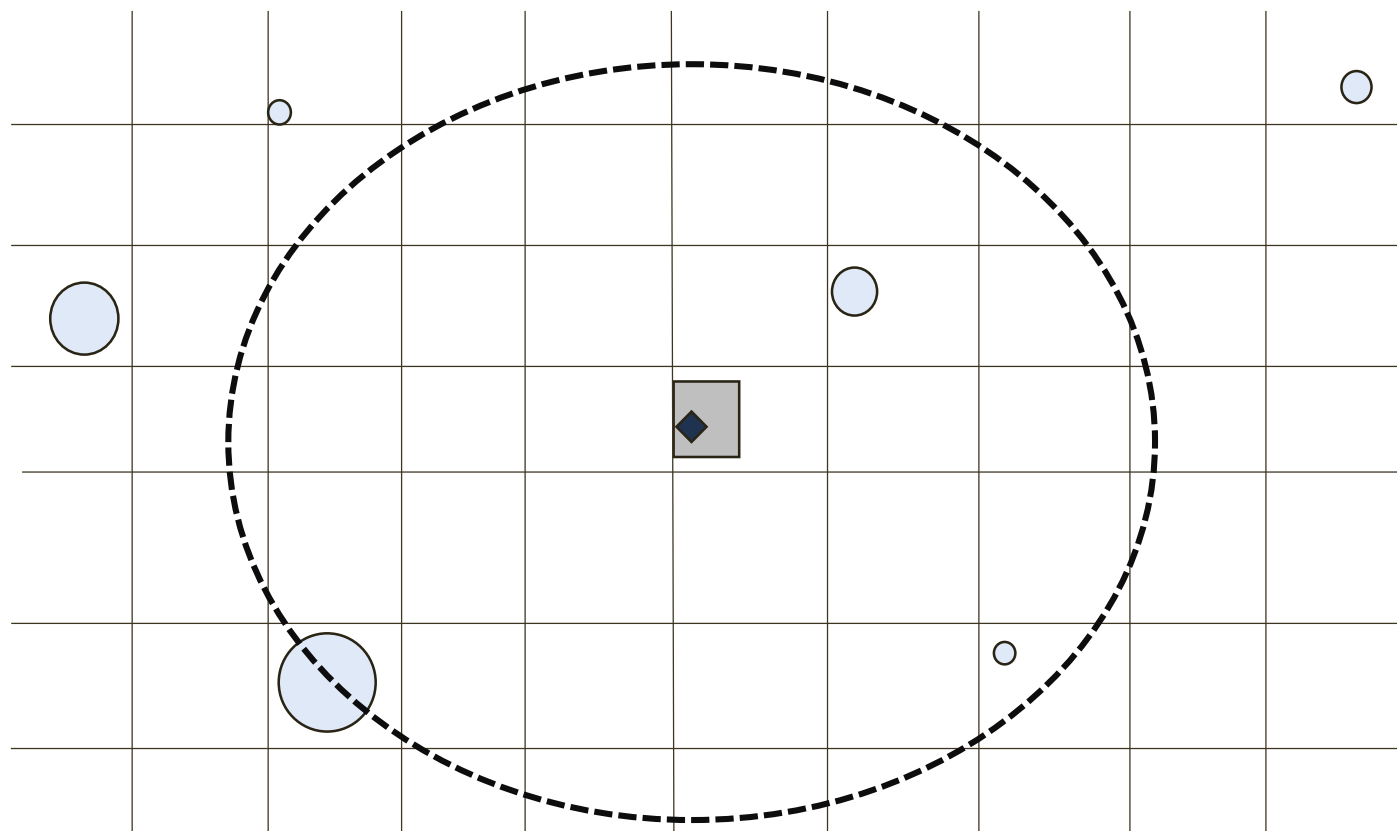
www.statcan.gc.ca

Canada

What can we do with the data?

Example - Proximity Measures

100



26



Statistics
Canada

Statistique
Canada

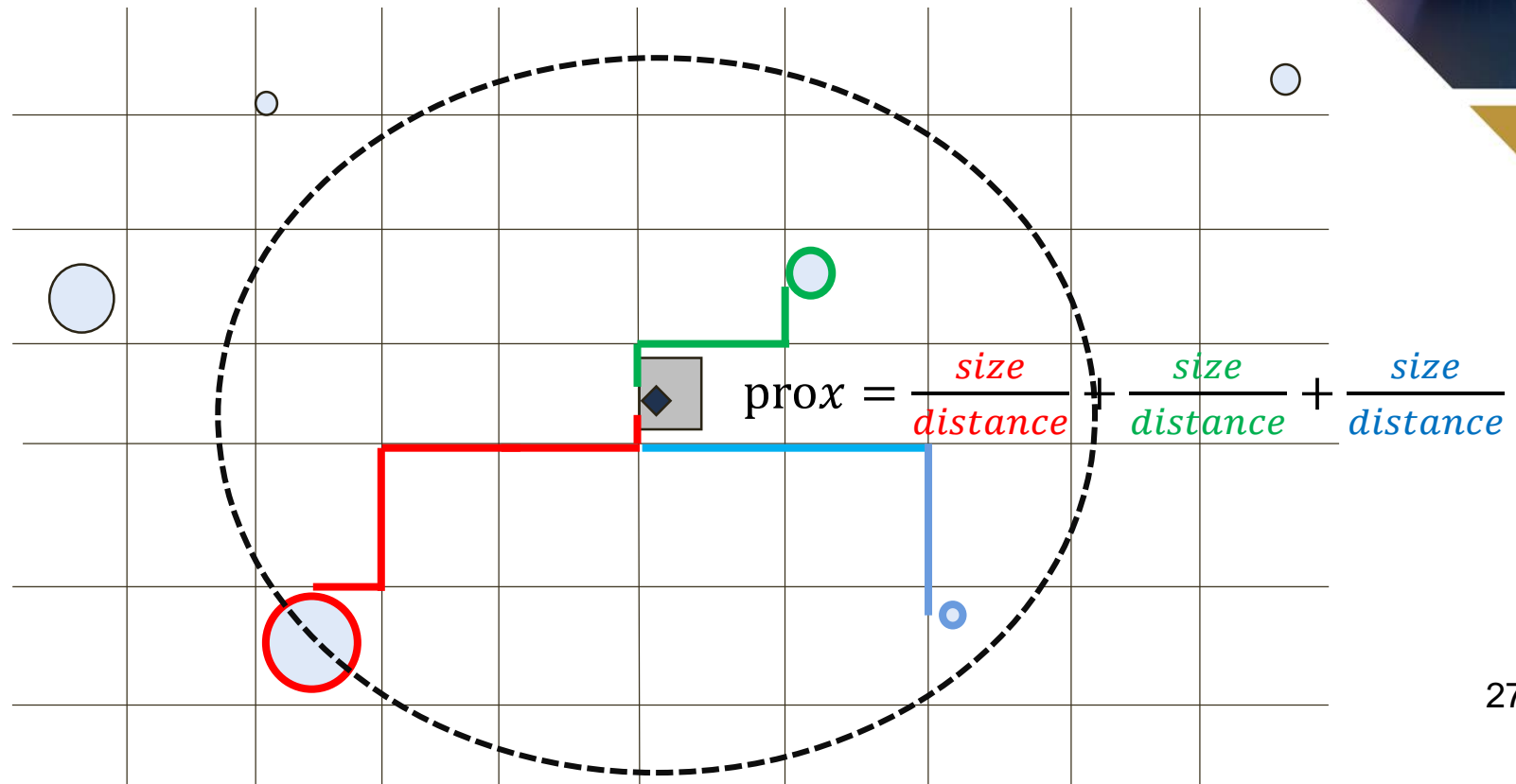
www.statcan.gc.ca

Canada

What can we do with the data?

Example - Proximity Measures

100



27



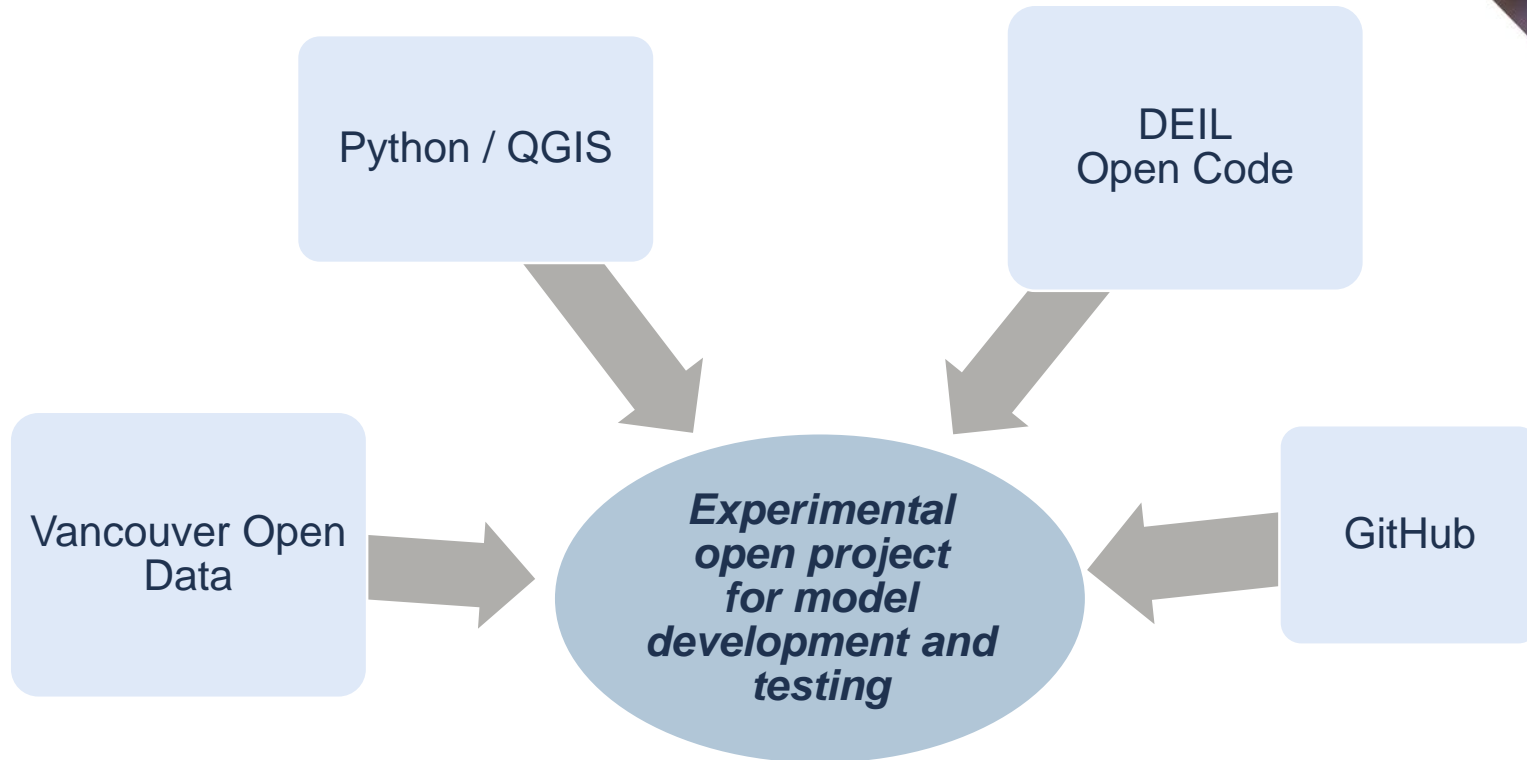
Statistics
Canada

Statistique
Canada

www.statcan.gc.ca

Canada

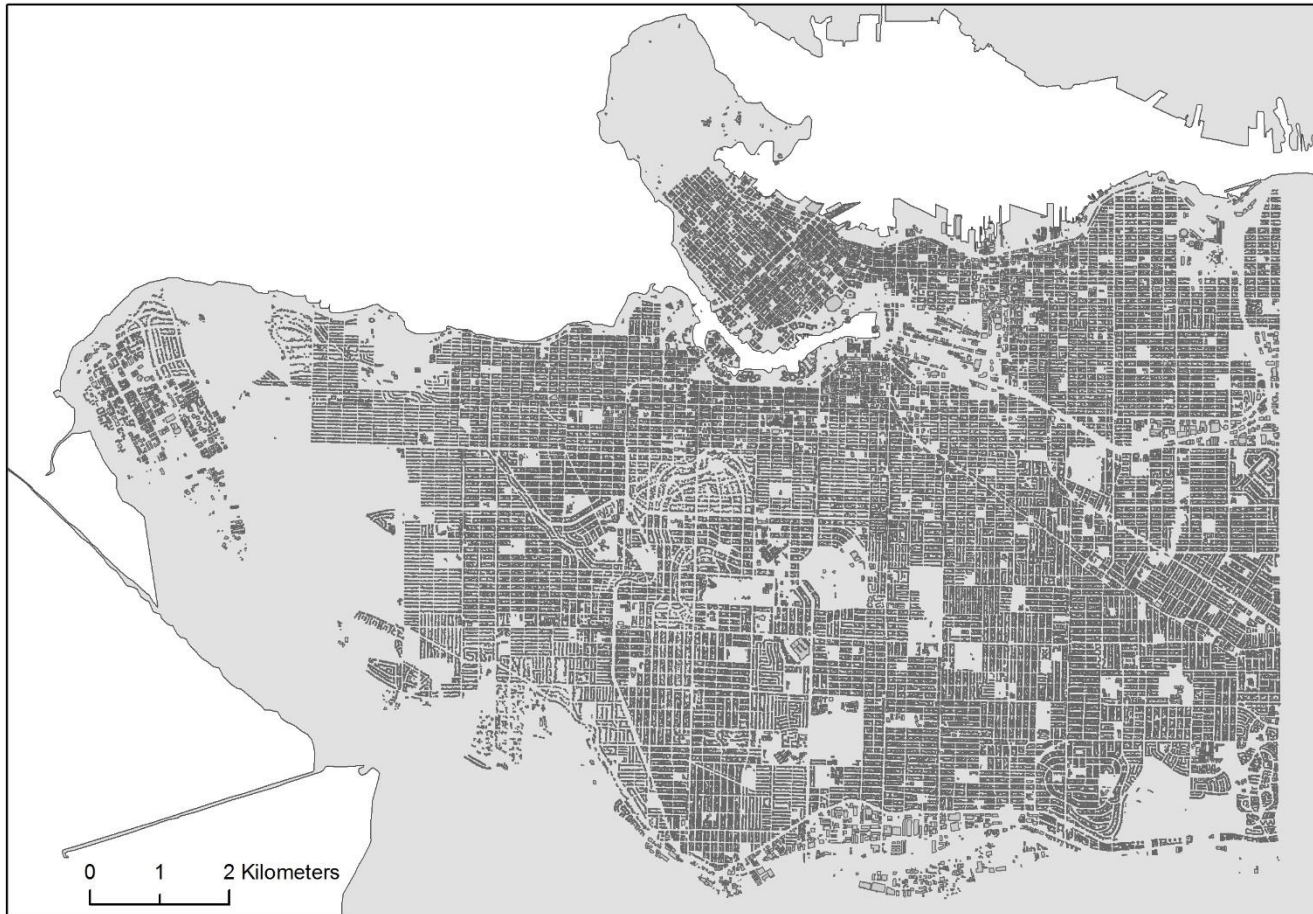
Is it possible with an open project approach?



28

Example of Vancouver Buildings

100



29



Statistics
Canada

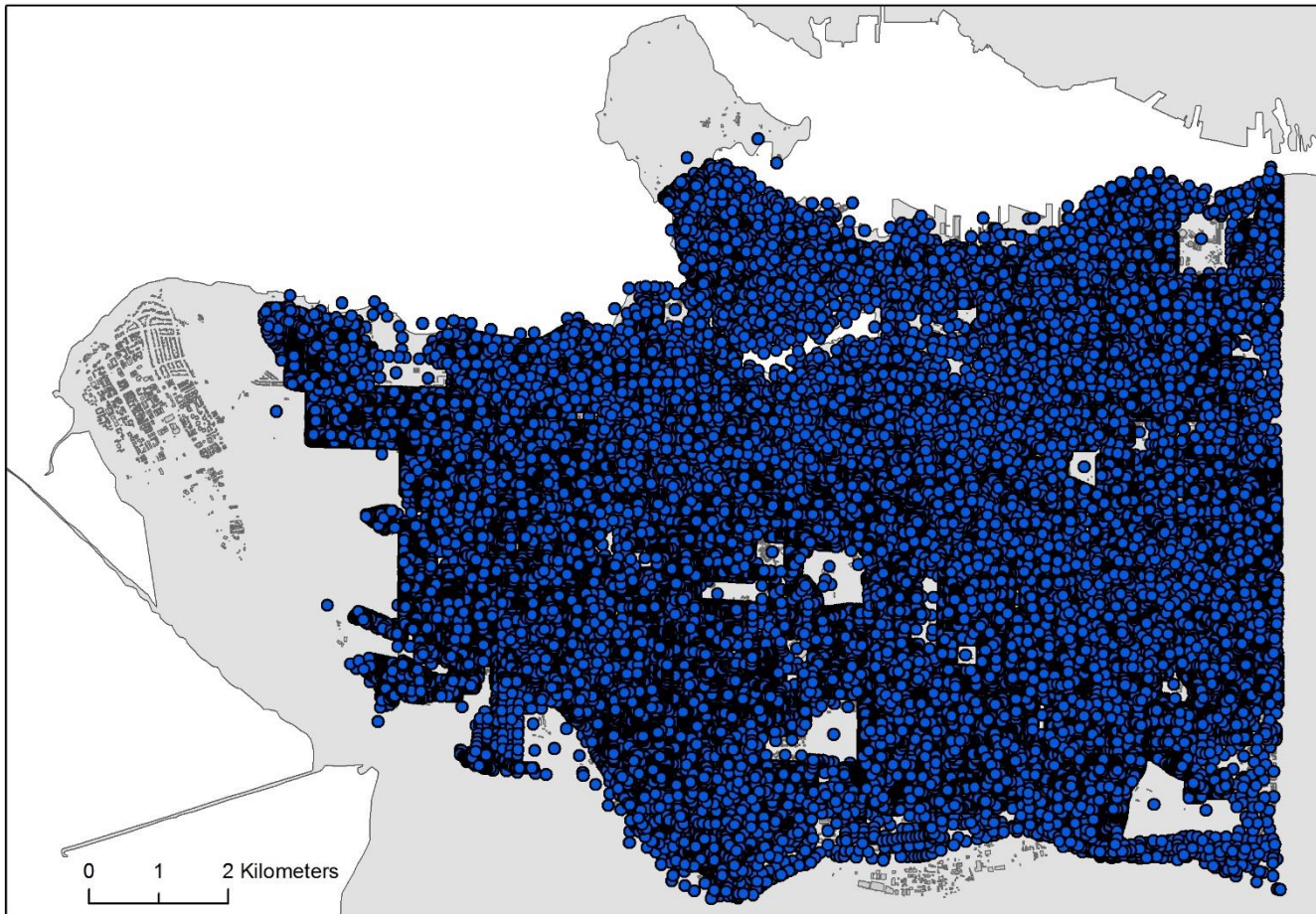
Statistique
Canada

www.statcan.gc.ca

Canada

Example of Vancouver Addresses

100



30



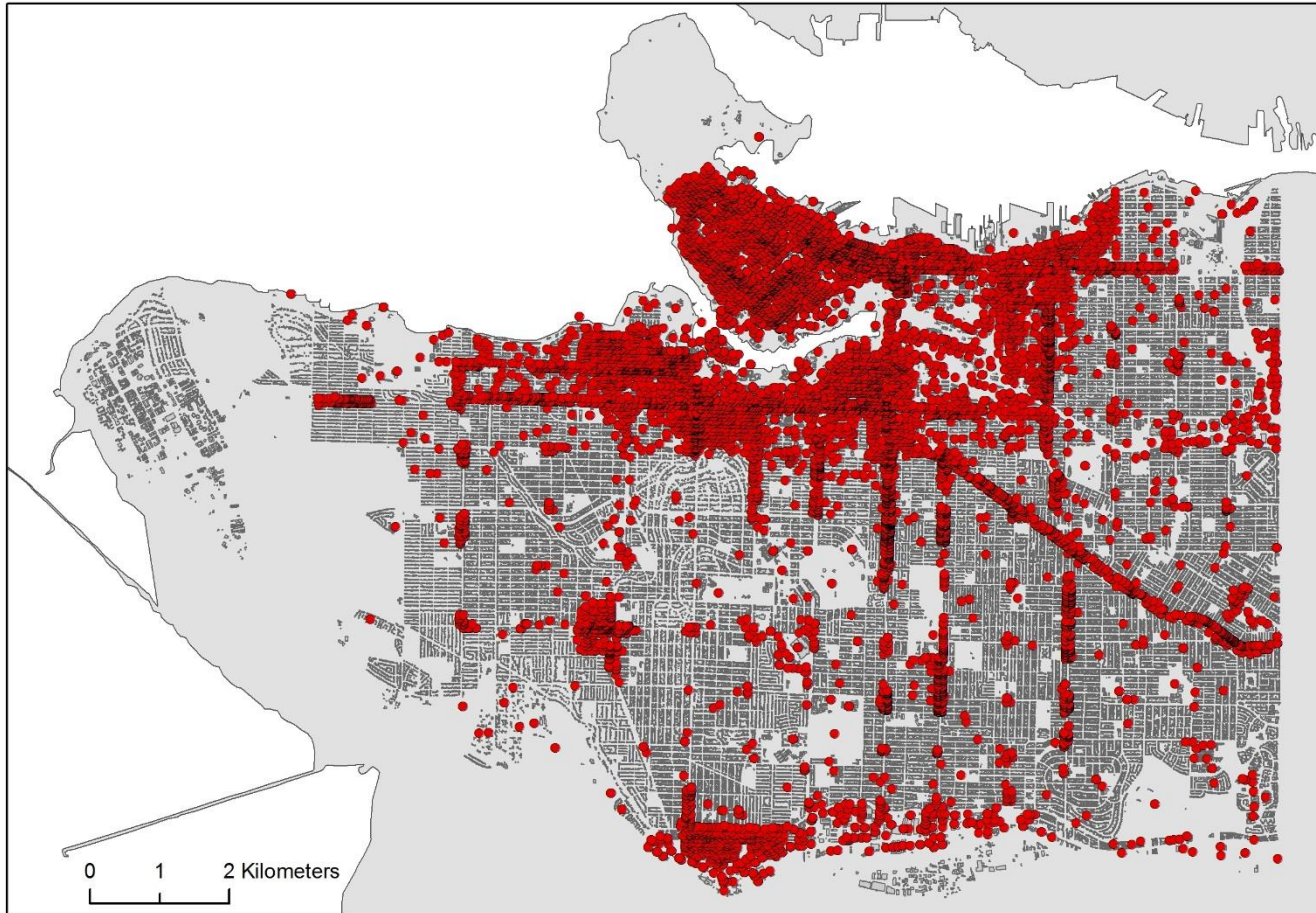
Statistics
Canada

Statistique
Canada

www.statcan.gc.ca

Canada

Example of Vancouver Businesses



Example of Vancouver Buildings + addresses



Example of Vancouver Buildings+addresses+businesses

100



Example of Vancouver Buildings+addresses+food businesses



Challenges

- Coverage of open microdata
- Governance and licensing
- Heterogeneity of formats, concepts, definitions

Take home messages

- Different options for use and integration of open data
- Great opportunities for integrating and using open microdata from local/prov./fed. government sources
 - Buildings, businesses, addresses and much more
- Great potential for conducting this work in open spaces as an “open project”
 - You can contribute, develop, re-use!

THANK YOU!

For more information,

alessandro.alasia@canada.ca

jean.lemoullec@canada.ca

haaris.jafri@canada.ca

MERCI!

Pour de plus amples renseignements,

alessandro.alasia@canada.ca

jean.lemoullec@canada.ca

haaris.jafri@canada.ca



#StatCan100



Statistics
Canada

Statistique
Canada

www.statcan.gc.ca

Canada

100