



Statistics
Canada

Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca



Crowdsourcing with Statistics Canada:

a pilot project to explore new frontiers in data
collection, mapping, and open data

Presented by
Alessandro Alasia

alessandro.alasia@canada.ca

Data Exploration and Integration Lab (DEIL)
Centre for Special Business Projects

Institute for Data Science, Carleton University
January 18, 2017

Evolution of data collection methods



Why this pilot project?

- We know that crowdsourcing exists, is out there, and works
- What is needed is an in-depth understanding of how a national statistical office can mobilize contributors and existing technologies for the purpose of official statistics

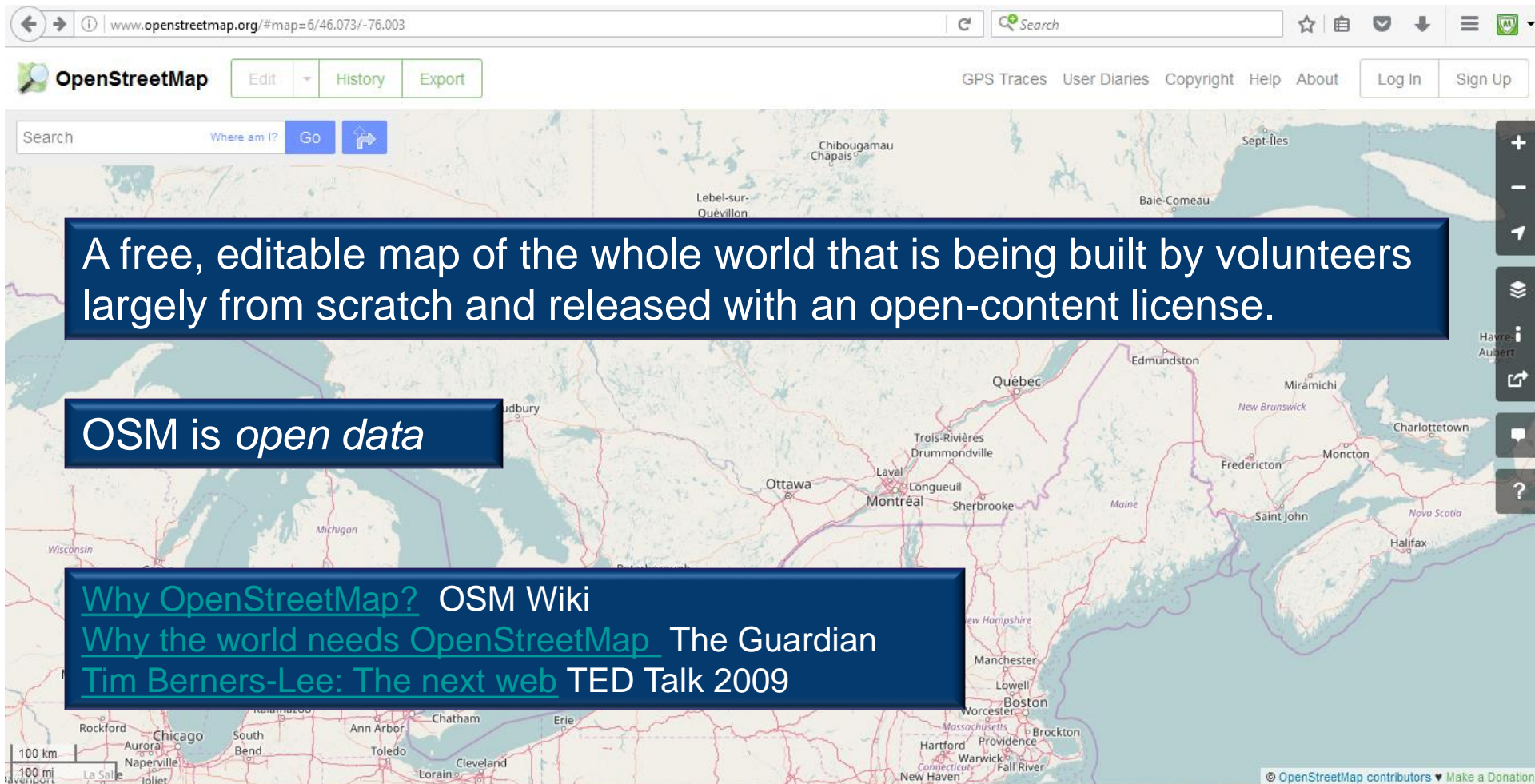
Crowdsourced data: a working definition for this project

Information of **statistical value** that is **volunteered** on an **open data platform**, which is freely **accessible, editable, and reusable** by anybody

What data are we crowdsourcing?

- Georeferenced attributes of buildings (e.g., building footprint, address, name of business in the building, etc.)
- Open data platform: OpenStreetMap (OSM)
- Selected municipalities. For the pilot project: Ottawa/Gatineau
- Two years pilot project

www.openstreetmap.org



Project implementation

May-August, 2016

- Background research /development (communication, outreach, IT)

September 15, 2016

- Launch of web page; start of communication campaign

October 17, 2016

- Deployment of customized OSM editor, social media campaign

**October 2016 to
February 2017**

- Outreach and monitoring

March 2017

- Preliminary assessment of results; possible revisions of work plan

Year II

- Experiment with a mobile application

Governance and collaborations



Collaborative agreement:
Open North, leading
open data non-profit

Educational institutions

Municipalities

Statistics Canada

- Working group (5 divisions)
- Steering committee

OSM community

- Local OSM Meet-Up group
- Canadian community @Talk-ca
- OSM Board of Directors
- OSM Institutions WG



**A customized version of an open
source OSM editor (iD-Editor)
to volunteer data to an open platform**

(Link to source code upcoming)



<http://www.statcan.gc.ca/eng/crowdsourcing>



Crowdsourcing





<https://crowdid.osmcanada.com>

File Edit View Favorites Tools Help

Crowdsourcing with Statistics Canada About Terms and conditions Français

+ Zoom in to Edit

Gatineau Ottawa

live 2.0.0



File Edit View Favorites Tools Help

Crowdsourcing with Statistics Canada

About Terms and conditions Français

Building ⓘ

▼ All fields

Name	⊗
Esplanade Laurier	+
Building	⊗
yes	▼
Levels	⊗
22	▲ ▼
Address	⊗

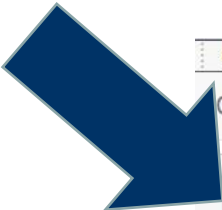
Add field:

▶ All tags (7)

[View on openstreetmap.org](#)

Point Area

live 107 hidden features 2.0.0



Preliminary results

**It is open data. You can monitor the project!
We have developed an open source
dashboard (beta version)**

(Link to source code upcoming)



Ottawa

Crowdsourcing - December 2016 Cities ▾ data Maps ▾

[Source Code](#)

29,758

Buildings



184

Users

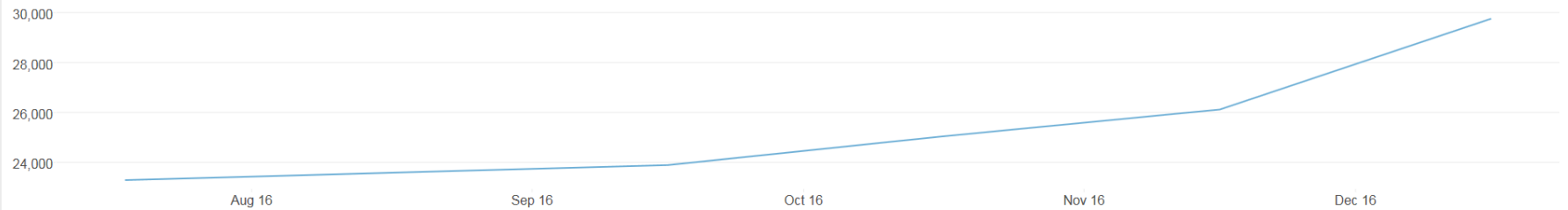


2.39

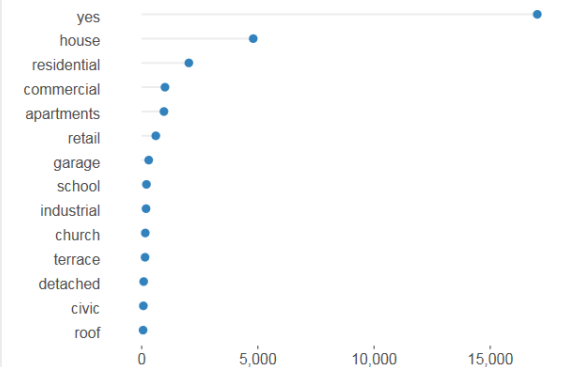
Average number of tags



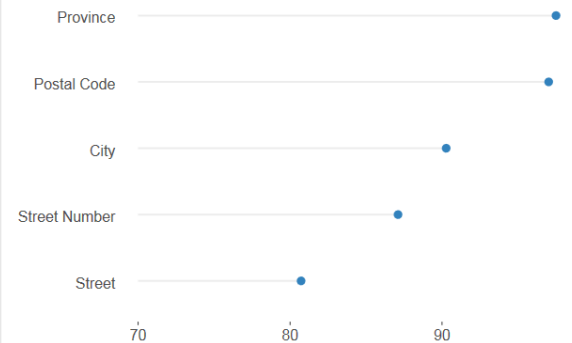
Buildings mapped



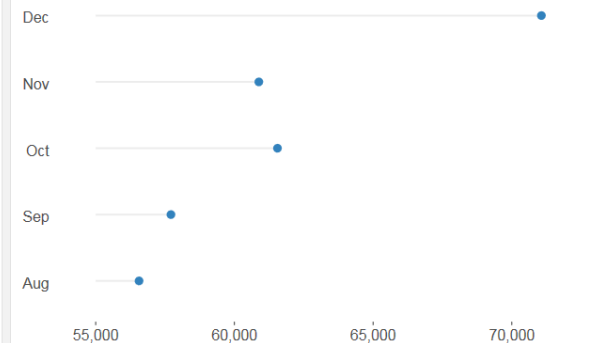
Buildings by type



Missing address fields in percentage






Number of tags

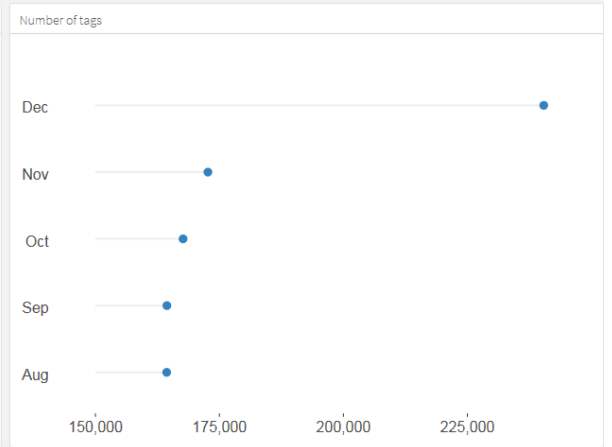
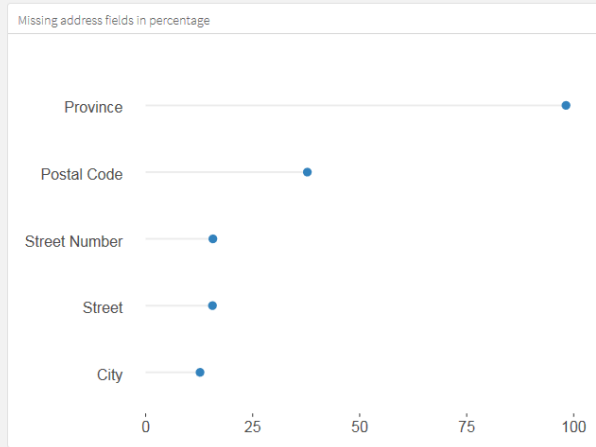
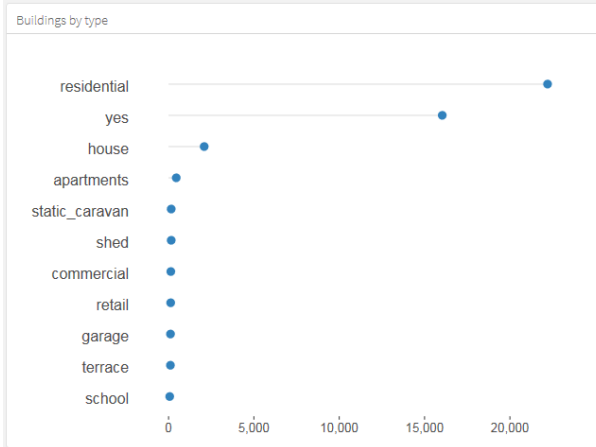


Gatineau

Crowdsourcing - December 2016 Cities ▾ data Maps ▾

[Source Code](#)

41,732 Buildings		103 Users		5.76 Average number of tags	
----------------------------	---	---------------------	---	---------------------------------------	---



Ottawa - Gatineau

Crowdsourcing - December 2016 Cities ▾ data Maps ▾

[Source Code](#)

71,489

Buildings



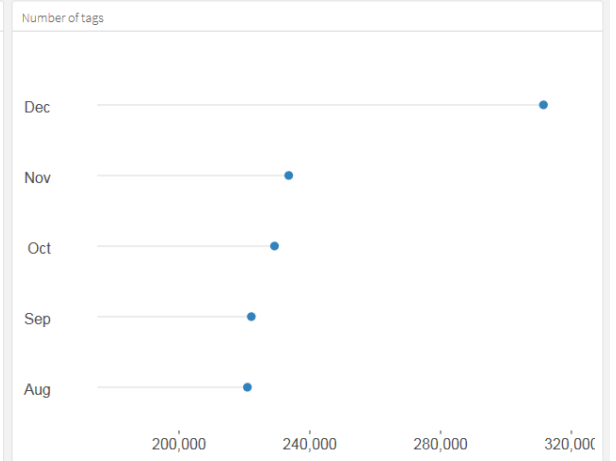
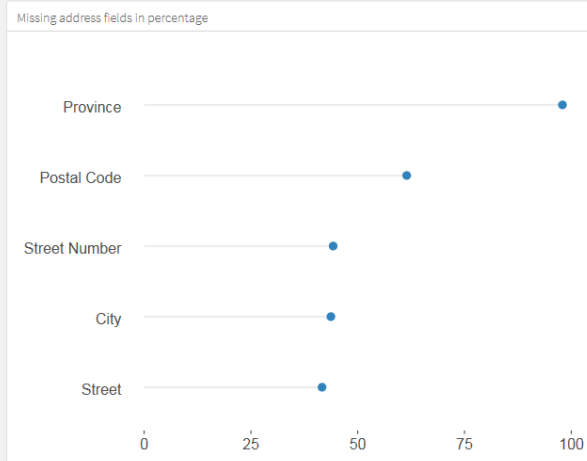
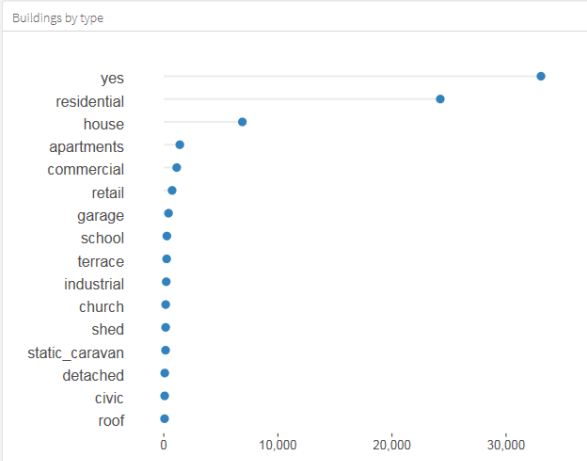
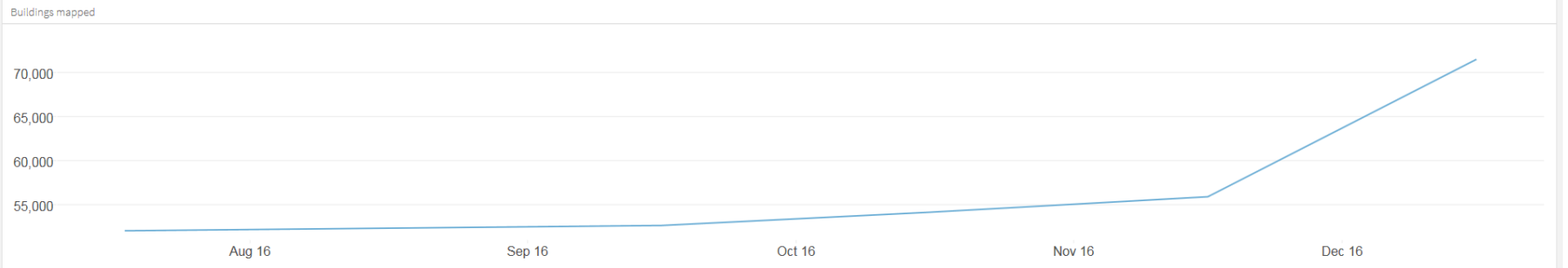
247

Users



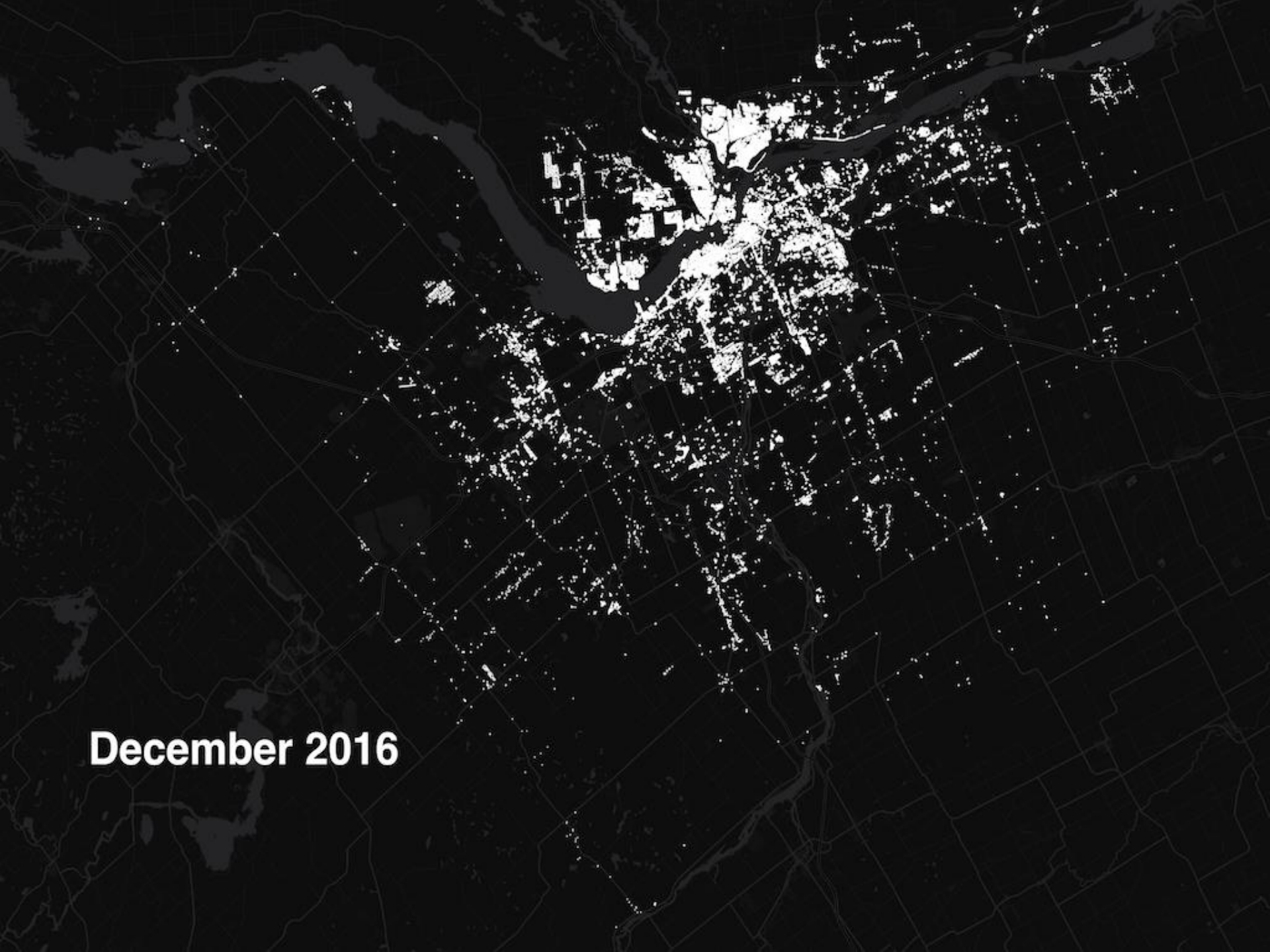
4.36

Average number of tags



An aerial night satellite image of a city, likely New York City, showing a dense grid of streets and illuminated areas. The image is in grayscale, with the city lights appearing as bright white and yellow spots against a dark background. A prominent river or waterway is visible in the upper left quadrant. The overall scene is a high-contrast, black and white representation of an urban landscape at night.

August 2016



December 2016

It is open data! You can monitor progress using OSM online tools

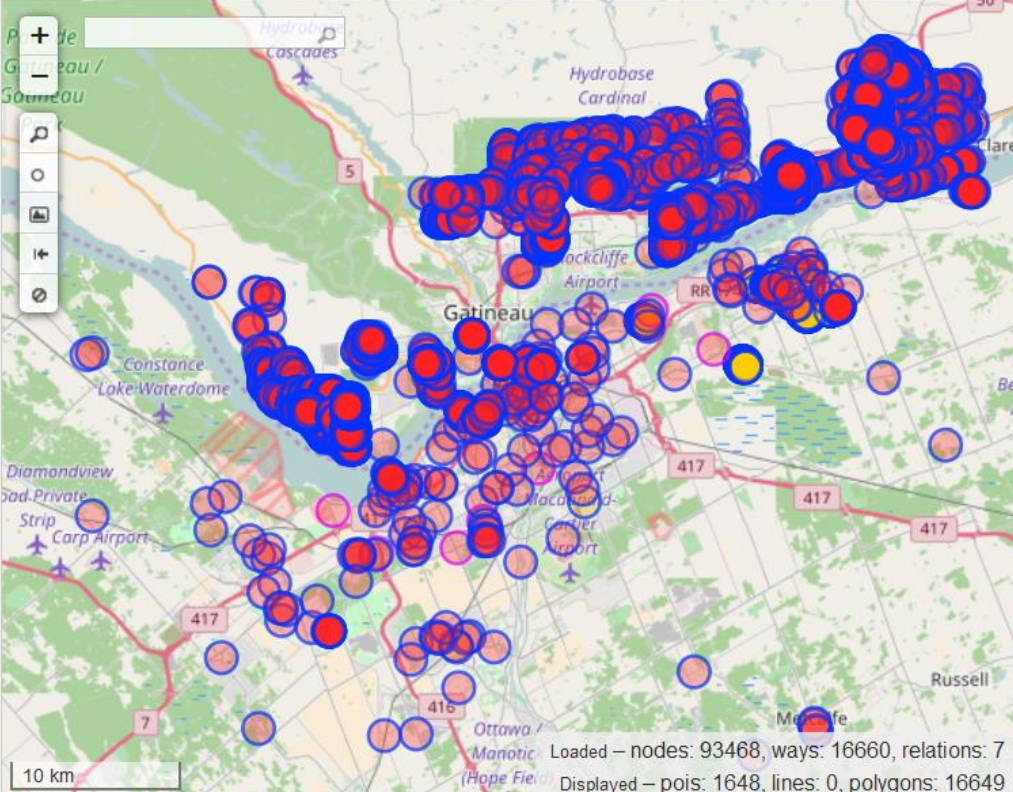
First 17 days of January 2017

Run Share Export Wizard Save Load Settings Help
overpass turbo
Flattr this!

Map Data

```

1  /*
2  This has been generated by the overpass-turbo wizard.
3  The original search was:
4  "building and newer:"17 days""
5  */
6  [out:json][timeout:25];
7  // gather results
8  (
9    // query part for: "building and newer:"17 days""
10   node["building"](newer:"{{date:17 days}}")
11   {{{bbox}}};
12   way["building"](newer:"{{date:17 days}}")
13   {{{bbox}}};
14   relation["building"](newer:"{{date:17 days}}")
15   {{{bbox}}};
16 )
17 // print results
18 out body;
19 >;
20 out skel qt;
          
```



Loaded – nodes: 93468, ways: 16660, relations: 7
 Displayed – pois: 1648, lines: 0, polygons: 16649



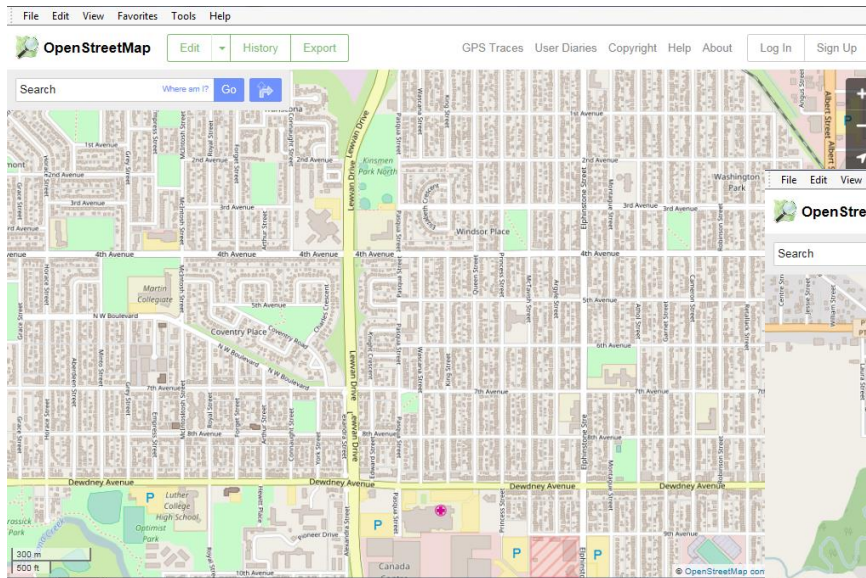
Lessons Learned

**it is still too early to make an
assessment of data quality issues,
but we have learned something!**

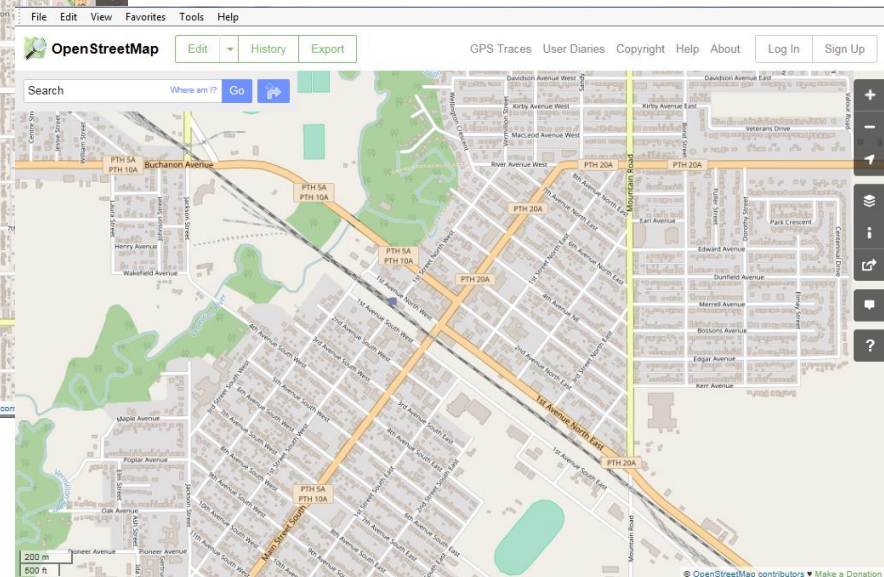
We are not alone....

- Over 1.7 mln buildings currently mapped on OSM across Canada

Rural: Dauphin (MB)
 about 4,400 buildings
 about 8,300 people




Urban: Regina (SK)
 about 95,000 buildings
 about 210,000 people





A new data ecosystem: open data platform



We are a new species in
this ecosystem

We need to learn about
and adapt to it



Expand collaborations and partnerships

Local(s): communities,
organizations, institutions

Global: communities,
organizations, institutions

OSM is a global open data platform

....starting by connecting at the local level

- Municipalities (and municipal open data)
- Local OSM groups and expert mappers
- Local for-profit and non-profit organizations
- Interest groups
- Educational institutions (high schools, colleges, universities)

Improve OSM data usability

- OSM data are great, but how easy is it to access, analyze and process these data?
- We are working on improving the processing of OSM data and enriching the tools to analyze them
- (OSM) data literacy is also a challenge

Foster the diversity of motivations to volunteer on an open data platform

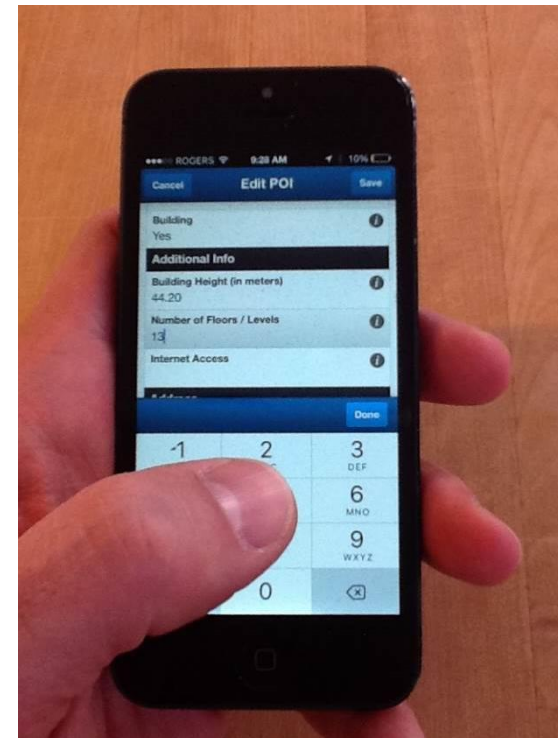
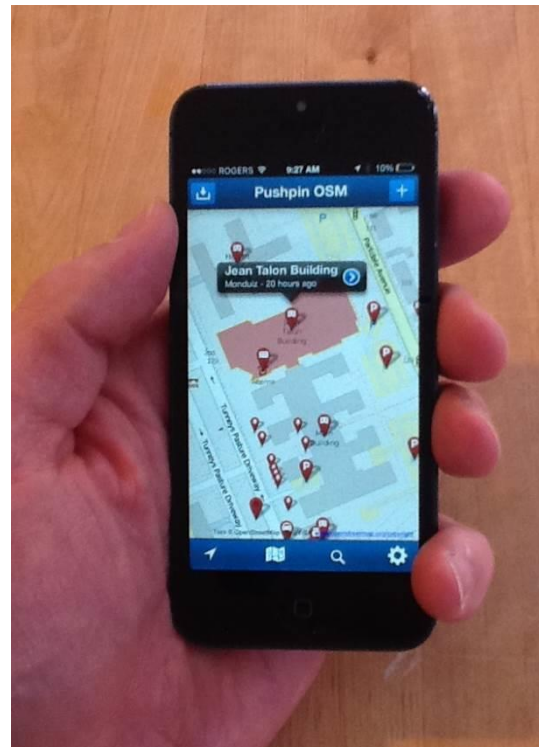
- There are a variety of motivations
 - Fun, curiosity
 - Civic engagement
 - Passion for maps and cartography, open data
 - Professional interest or development (learning)
- Ideally: collaborate with a core group of volunteer mappers as well as engaging the general public

Recognizing the potential of open data and open source

- The fundamental motivations:
 - Contributing to a public good (open data)
 - Learning and sharing (open source code)
- Explore the potential of source code development on collaborative platforms
- Explore the potential of light IT infrastructures that could be accessible by any organization at low cost

Next step: mobile version

- Engage users of mobile/portable devices, with the goal of “making data entry easy”



Can we work together?

- Engage communities
 - Community mapping and capacity building
 - High school projects: “map your neighborhood”
- Develop (open source) tools for organizations
 - [New York City and OpenStreetMap Collaborating Through Open Data](#). Code available through [GitHub](#)
 - Crowd-id Editor
 - Project dashboard and monitoring tool (beta version)
- Use the data and develop data applications
 - Automated applications for analysis, data processing and management
 - Community-based applications, or applications for businesses

Acknowledgments

- Over 200 OSM contributors
- [Openstreetmap – Ottawa Meetup](#) group and core group of expert contributors
- Inputs and ideas from many OSM developers including Mapbox
- City of Ottawa and Gatineau
- [Open North](#)
- Working Group and Steering Committee @Statcan

Questions?