

9-2016

# Spatio-Temporal Data in Tourism Research

Dmitry Shkolnik

*Destination Canada*, shkolnikd@gmail.com

Follow this and additional works at: [https://scholarworks.umass.edu/ttracanada\\_2016\\_conference](https://scholarworks.umass.edu/ttracanada_2016_conference)



Part of the [Tourism and Travel Commons](#)

---

Shkolnik, Dmitry, "Spatio-Temporal Data in Tourism Research" (2016). *TTRA Canada 2016 Conference*. 37.

Retrieved from [https://scholarworks.umass.edu/ttracanada\\_2016\\_conference/37](https://scholarworks.umass.edu/ttracanada_2016_conference/37)

This Presentation is brought to you for free and open access by the TTRA Canada at ScholarWorks@UMass Amherst. It has been accepted for inclusion in TTRA Canada 2016 Conference by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).

# Spatiotemporal Data in Tourism Research

**Applications from Flickr Photo Social Data**

TTRA Canada, “2.1. Understanding Visitors and Travel Behaviour”

Dmitry Shkolnik

Edmonton, AB, September 29, 2016




Canada

# Spatiotemporal

---

spa·ti·o·tem·po·ral

/ˌspāSHēōˈtempərəl/ 

*adjective*

PHYSICS

PHILOSOPHY

adjective: **spatio-temporal**

belonging to both space and time or to space-time.

Spatiotemporal data contains geographic and time information.

Ideal for analyzing activity, movement, and distribution.

Are we as tourism researchers taking enough advantage of this kind of data?

1

**Directions: How much  
do we know about  
visitor movement?**

# Not that much...

Q24. Which of the following Canadian destinations did you visit on this trip? (Select all that apply)

<Show the full answer list (codes 1-30, 99) for all country EXCEPT the US and Canada>

<If country=US or Canada and q27b=code 37, show only codes 28-30, and 99>

<If country=US or Canada and q27b=code 38, show only codes 1-4, and 99. Plus show the "British Columbia" header>

<If country=US or Canada and q27b=code 39, show only codes 5-9, and 99. Plus show the "Alberta" header>

<If country=US or Canada and q27b=code 42, show only codes 12-15, and 99. Plus show the "Ontario" header>

<If country=US or Canada and q27b=code 43, show only codes 16-18, and 99. Plus show the "Quebec" header>

<If country=US or Canada and q27b=code 44, show only codes 19-27, and 99. Plus show the following headers above the appropriate code groups per the answer list below, "New Brunswick," "Nova Scotia," "Prince Edward Island," and "Newfoundland & Labrador">

<If country= Canada and q27b=code 128, show only codes 19-21, and 99. Plus show the header, "New Brunswick,">

<If country= Canada and q27b=code 129, show only codes 22-23, and 99. Plus show the header, "Nova Scotia,">

<If country= Canada and q27b=code 130, show only codes 24-25, and 99. Plus show the header, "Prince Edward Island,">

<If country= Canada and q27b=code 131, show only codes 26-27, and 99. Plus show the header, "Newfoundland & Labrador">

## British Columbia

- |                        |                          |   |
|------------------------|--------------------------|---|
| Vancouver              | <input type="checkbox"/> | 1 |
| Victoria               | <input type="checkbox"/> | 2 |
| Whistler               | <input type="checkbox"/> | 3 |
| Other British Columbia | <input type="checkbox"/> | 4 |

<PN: Add a row space here>

## Alberta

- |               |                          |   |
|---------------|--------------------------|---|
| Calgary       | <input type="checkbox"/> | 5 |
| Edmonton      | <input type="checkbox"/> | 6 |
| Banff         | <input type="checkbox"/> | 7 |
| Jasper        | <input type="checkbox"/> | 8 |
| Other Alberta | <input type="checkbox"/> | 9 |

<PN: Add a row space here>

- |                                 |                          |    |
|---------------------------------|--------------------------|----|
| Saskatchewan (including Regina) | <input type="checkbox"/> | 10 |
|---------------------------------|--------------------------|----|

<PN: Add a row space here>

- |                               |                          |    |
|-------------------------------|--------------------------|----|
| Manitoba (including Winnipeg) | <input type="checkbox"/> | 11 |
|-------------------------------|--------------------------|----|

<PN: Add a row space here>

## Ontario

- |               |                          |    |
|---------------|--------------------------|----|
| Niagara Falls | <input type="checkbox"/> | 12 |
| Toronto       | <input type="checkbox"/> | 13 |

Other types of data can complement survey data.

Data that shows real, on-the-ground visitor behaviour can mitigate respondent recall biases and help assuage sampling concerns.

This is also relevant in sparsely populated areas like Canada's North where survey sample sizes are more likely to run into reliability issues.

2

**Origins: Tourist  
movement and activity  
from photo metadata**

# Digital footprints and photo metadata

---

**Digital footprints:** *the trail of meta-data emitted by users of social networks and other user-generated content*

Enormous quantities of posts, tweets, photos generated daily

The meta-data (user information, location, time, tags, etc.) of this content can be as valuable as the content itself

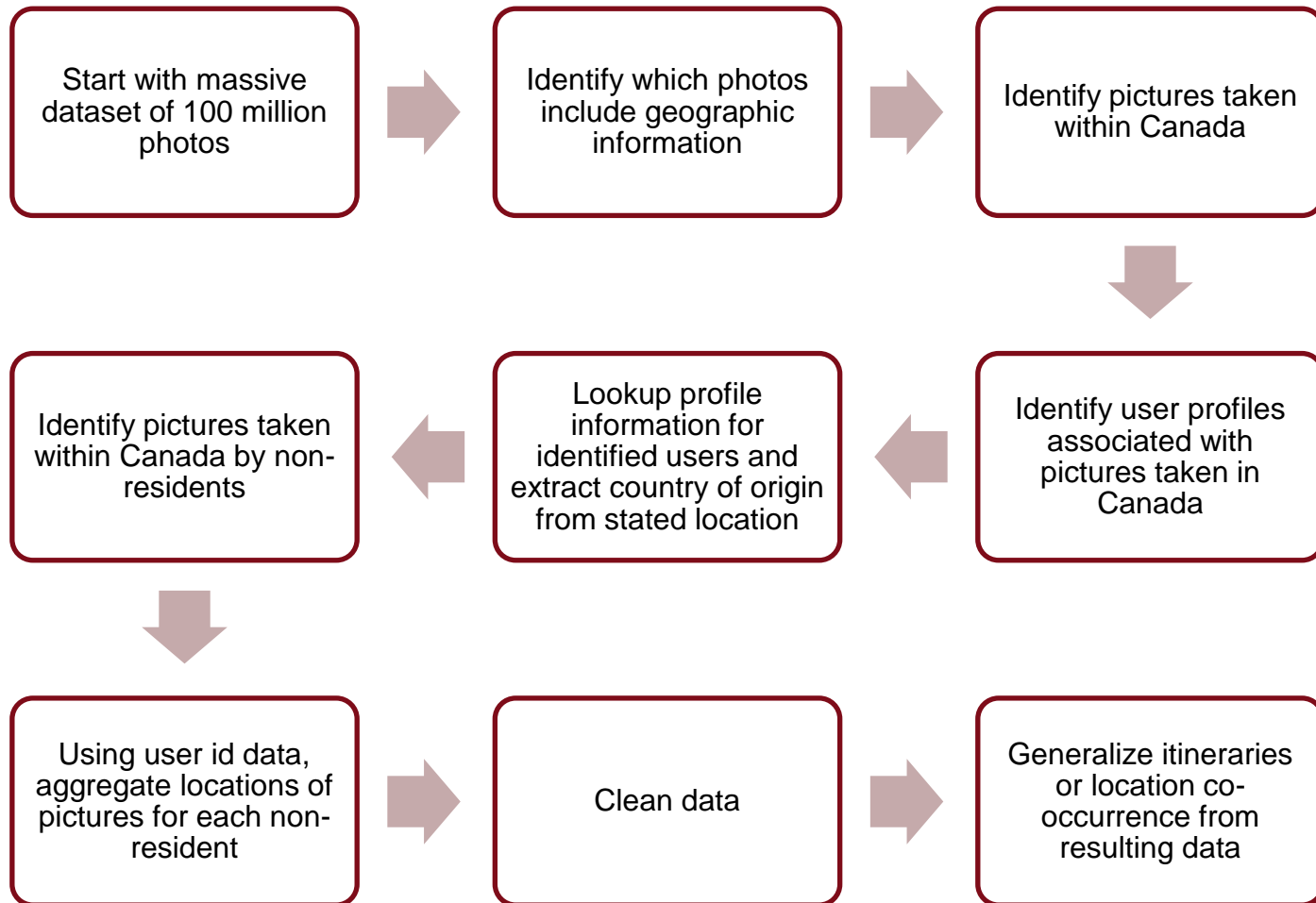
Flickr is a large photo sharing and hosting platform owned by Yahoo

- hosting over 10bn images from over 112m users across 63 countries
- 1m new photos uploaded daily

In 2015, Flickr released a Creative Commons 100M dataset including meta-data for 100m photos taken over 2005-2015

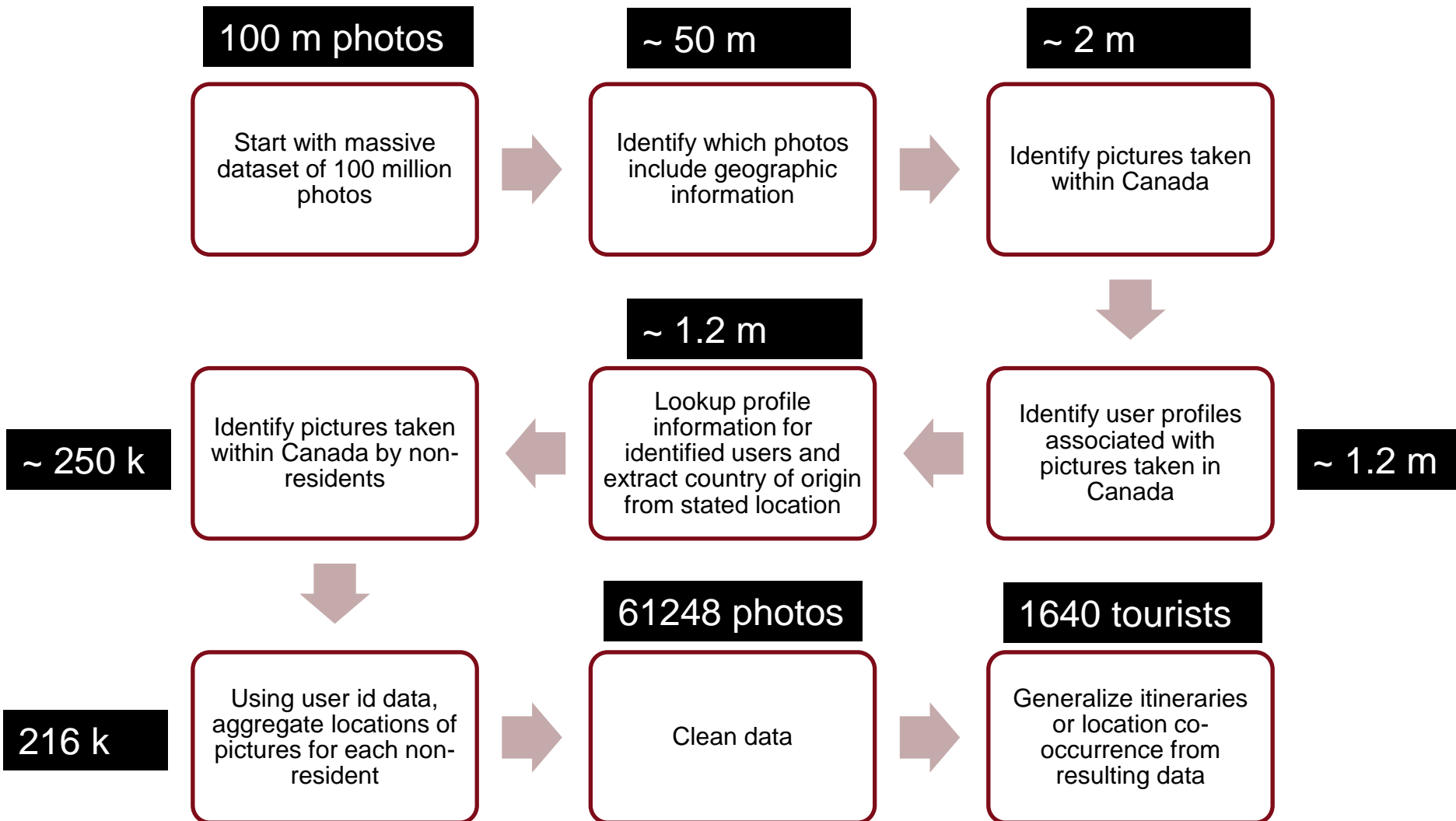
Yahoo Labs: <https://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67>

# What we did – the methodology

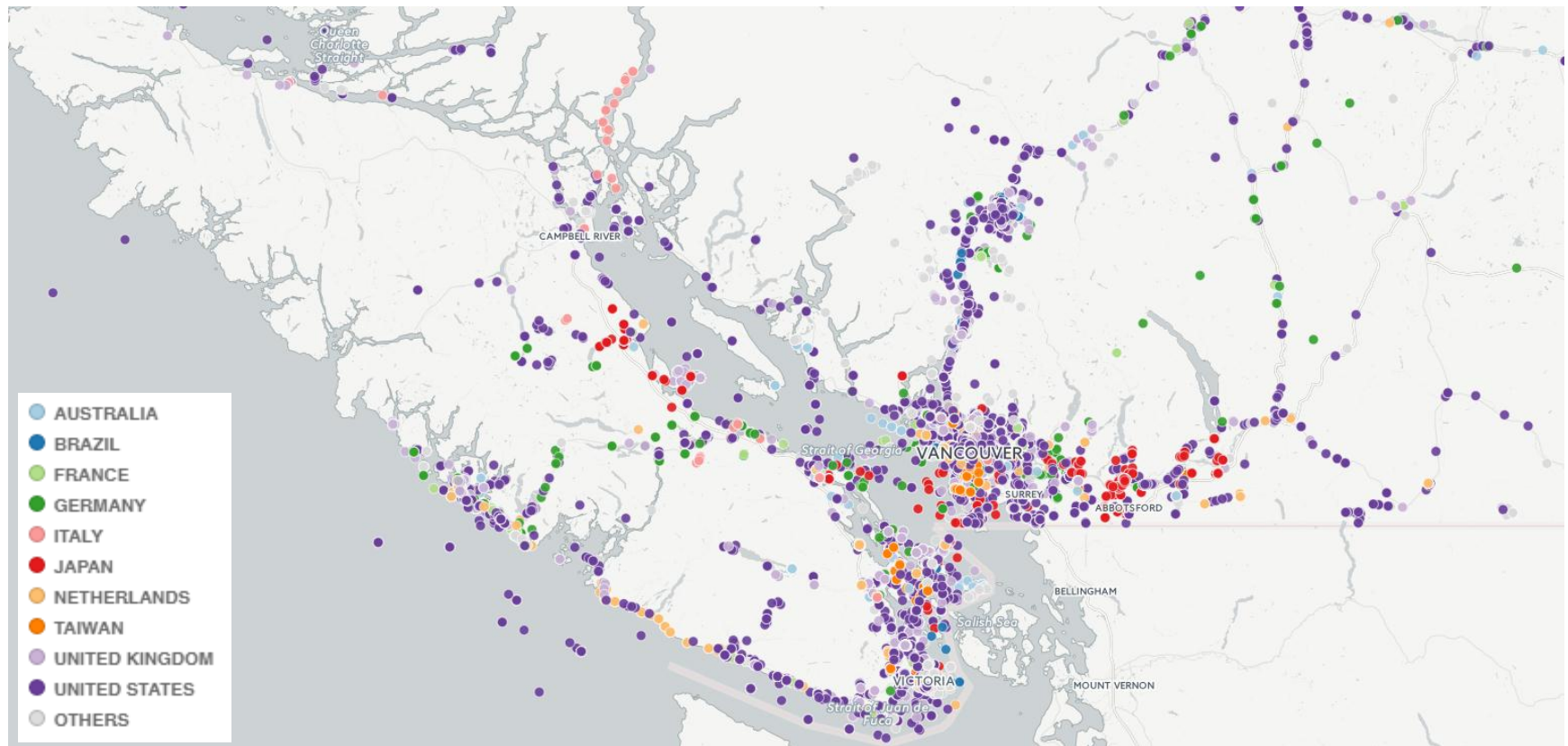


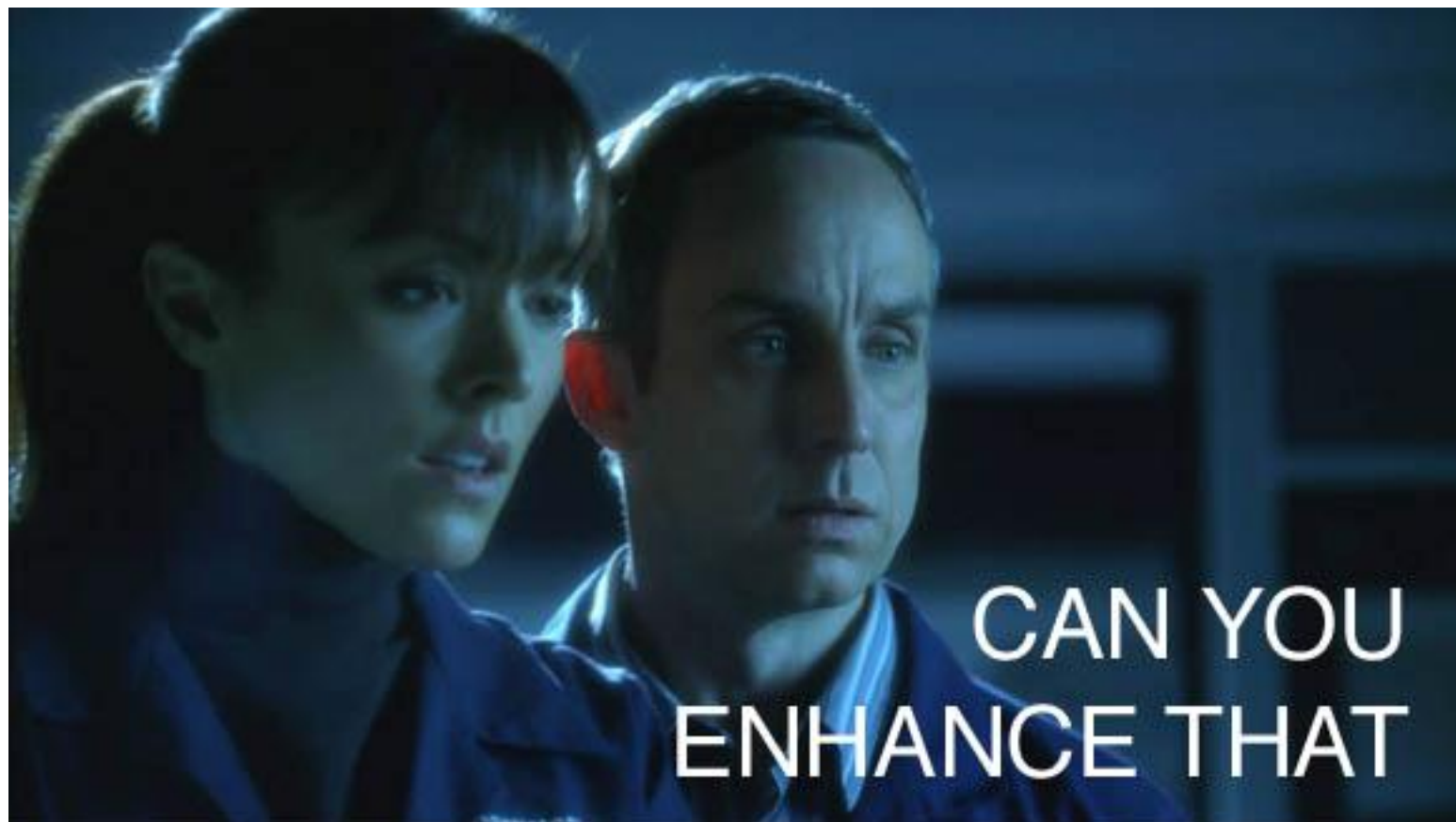


# Big data quickly turns into small data...

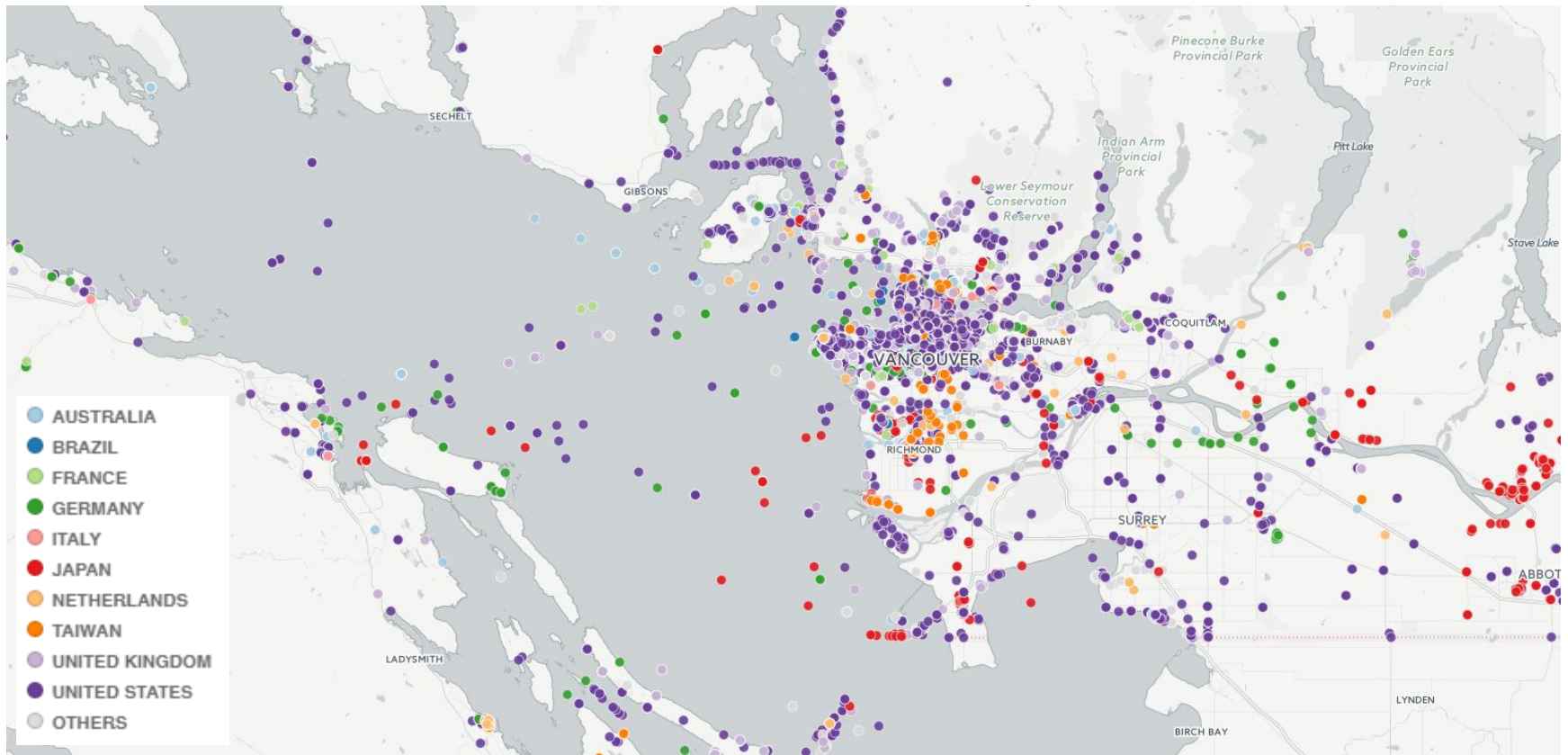


# Distribution of tourist activity from photos



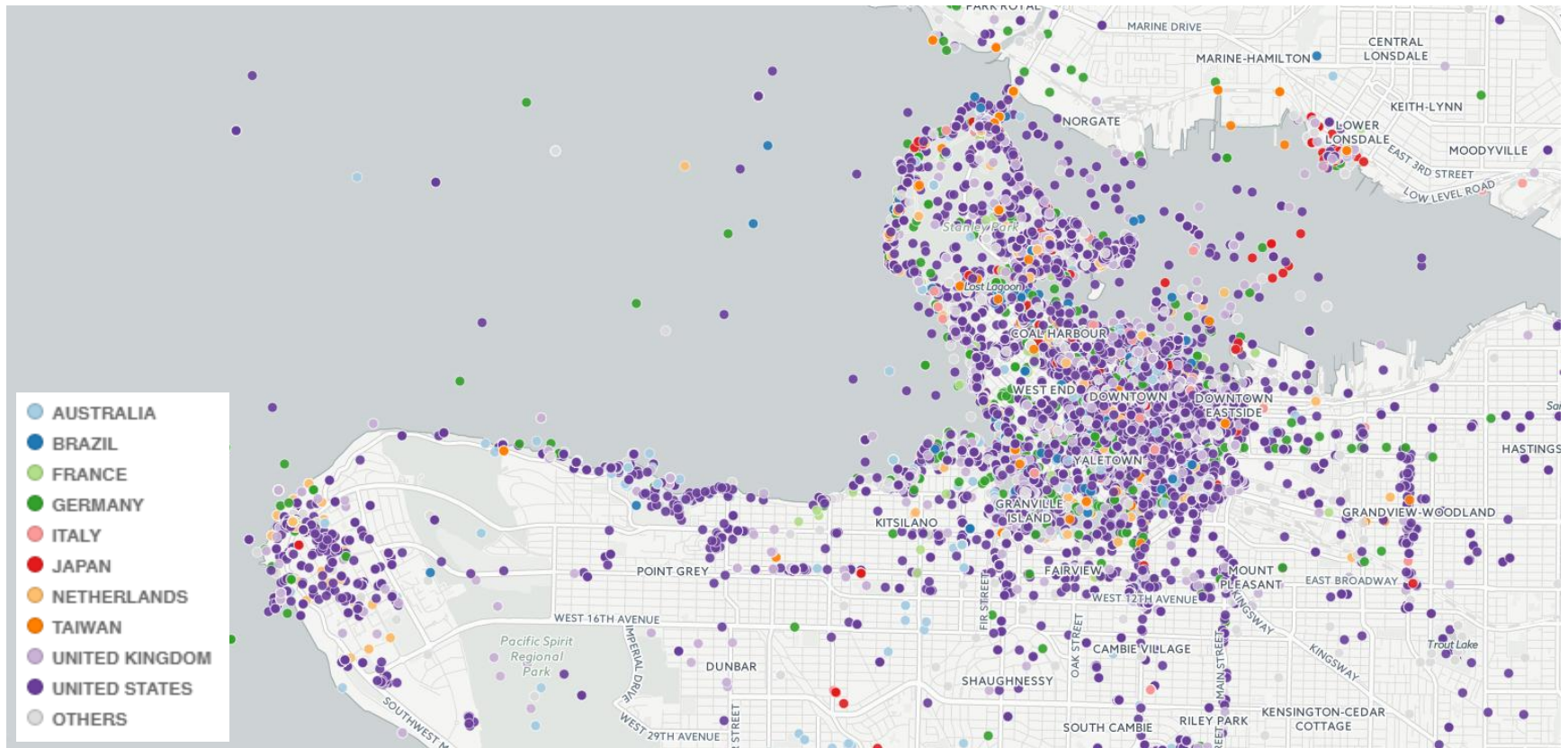


# Distribution of tourist activity from photos

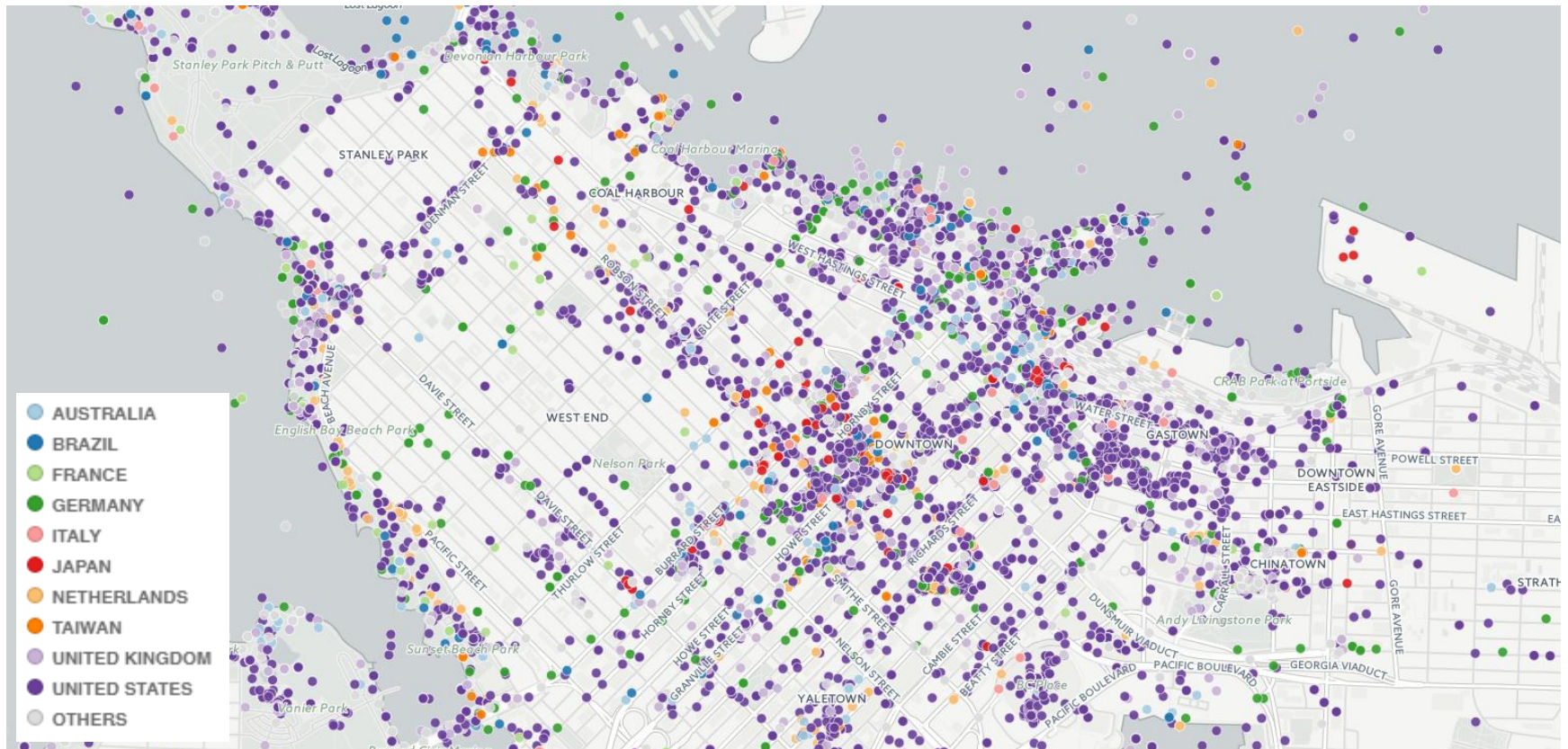




# Distribution of tourist activity from photos



# Distribution of tourist activity from photos

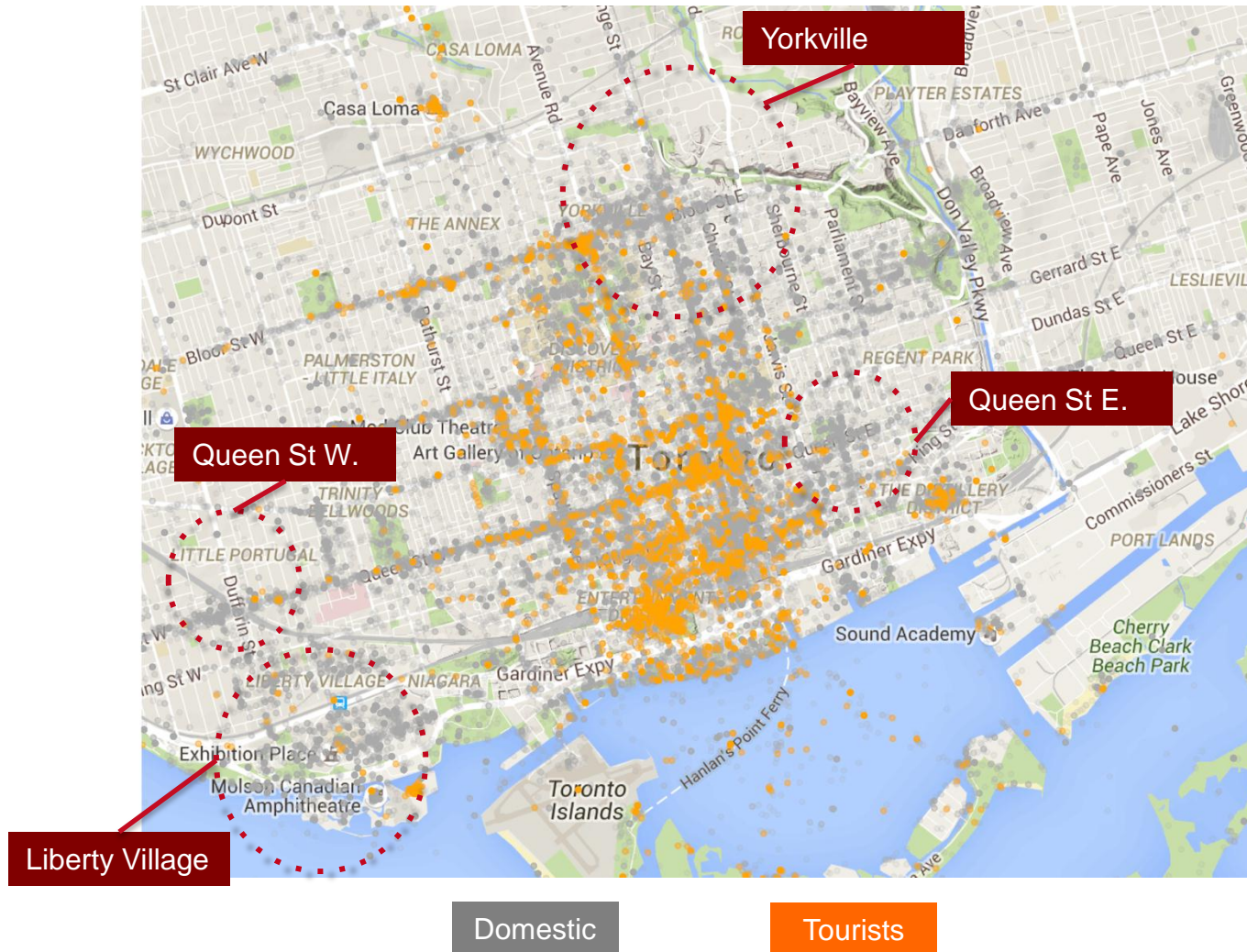


# Distribution of tourist activity from photos



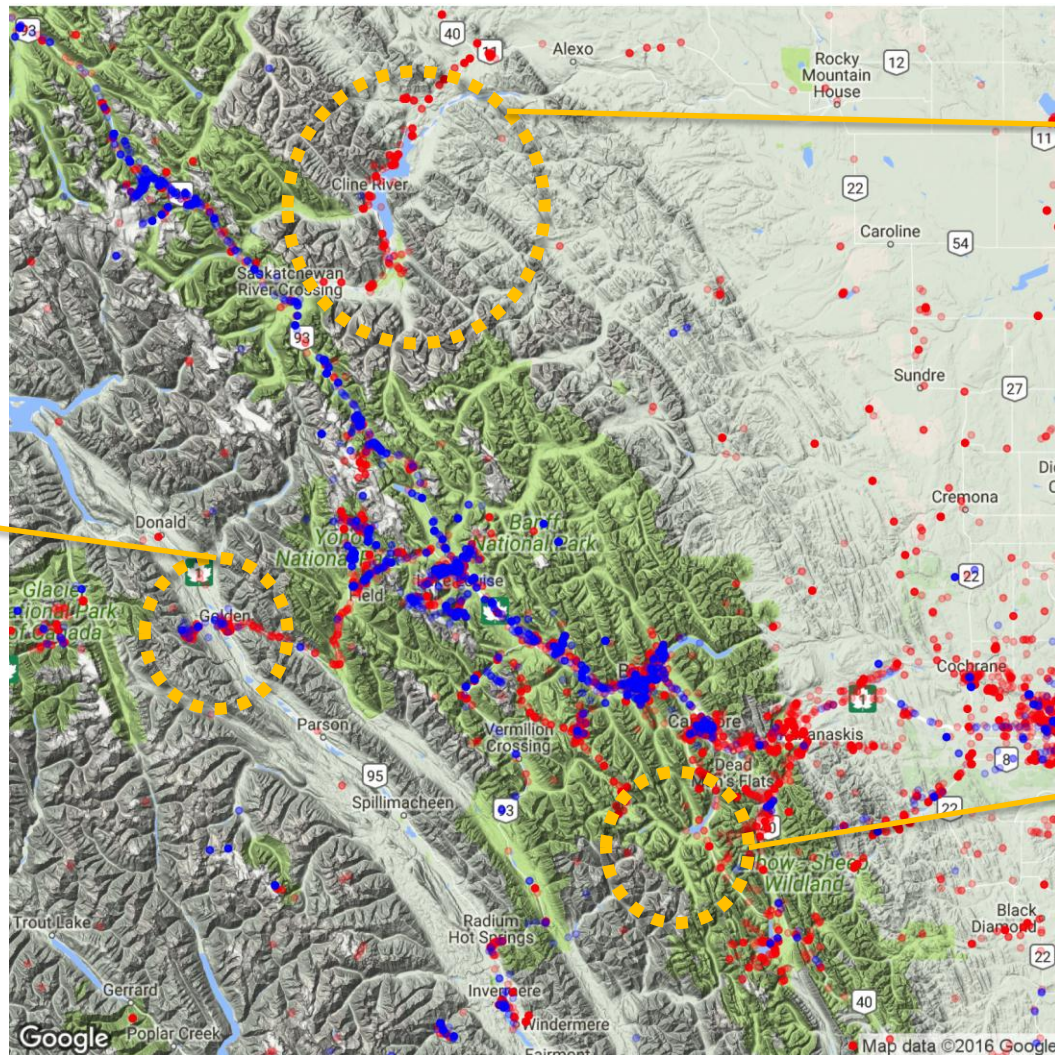


# Where do domestic residents go? - Toronto





# Where do domestic residents go? - Rockies



Golden, BC

Cline River



Mt. Assiniboine

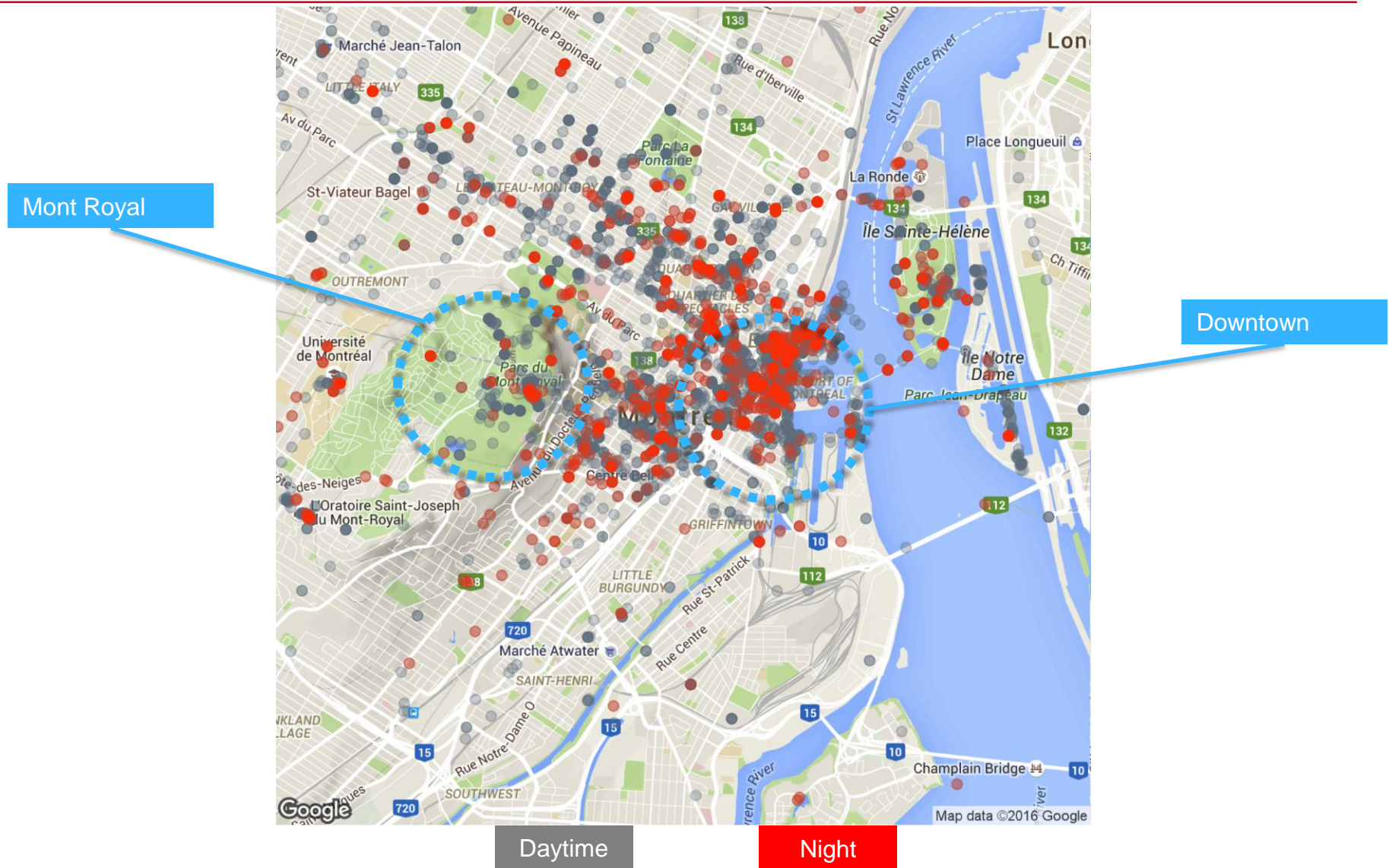


Domestic

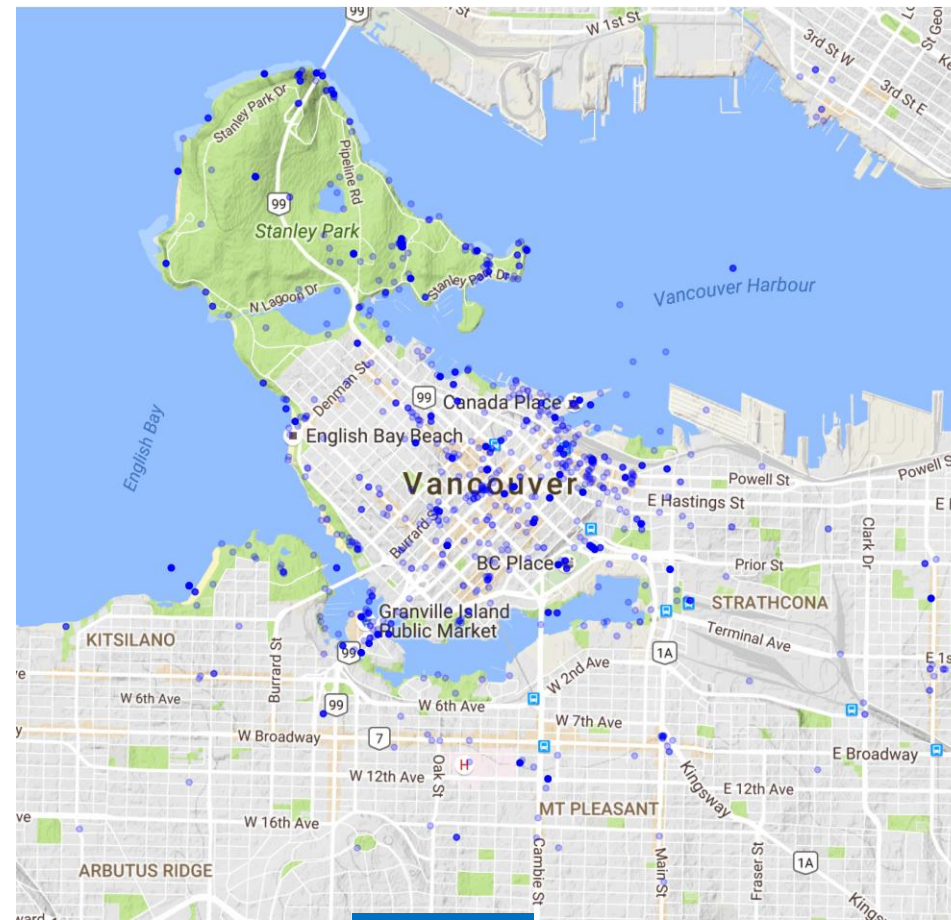
Tourists



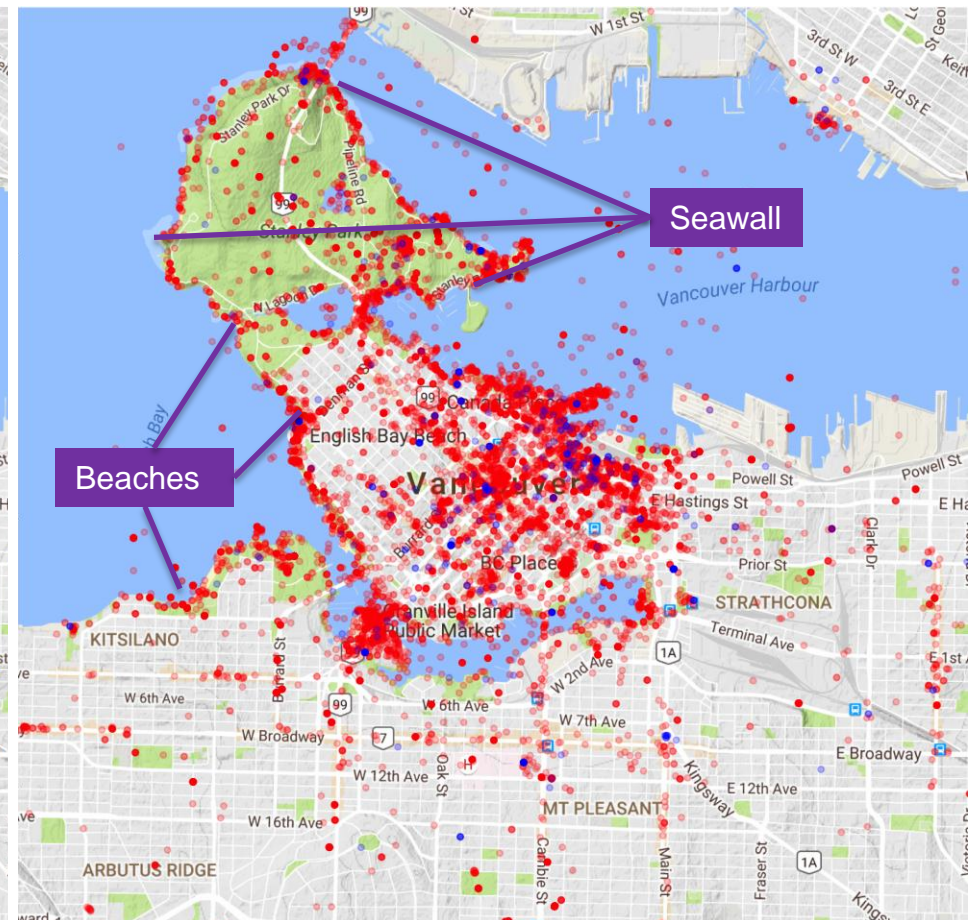
# Do day and night differ? - Montreal



# Rain or shine? - Vancouver



Rain



No Rain



# Connecting the dots





# Connecting the dots

User: 32781193  
Origin: UK  
Duration: 14 days  
Visited in July 2009

## Itinerary:

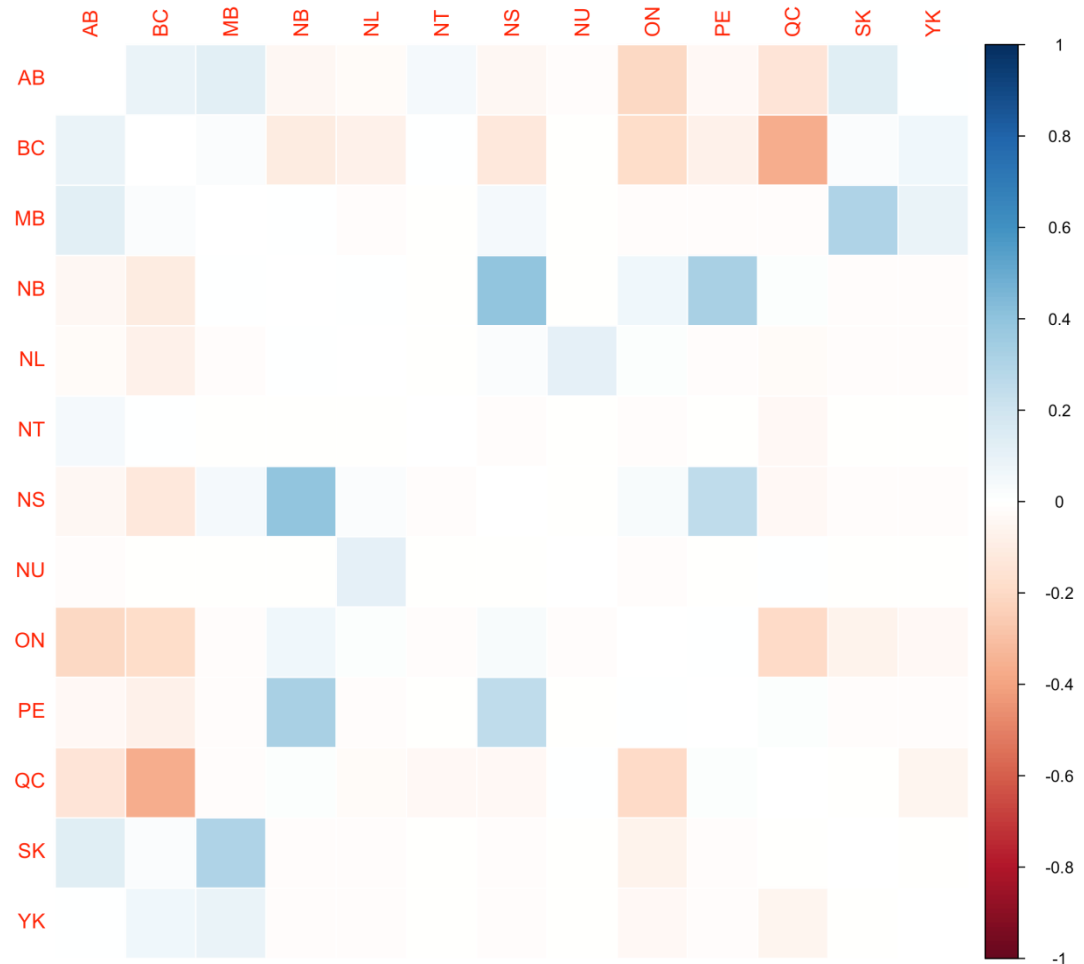
Vancouver, Victoria,  
Nanaimo, Tofino,  
Whistler, Kamloops,  
Wells Gray Park,  
Jasper, Rockies,  
Banff, Calgary



# Do tourists visit multiple provinces?

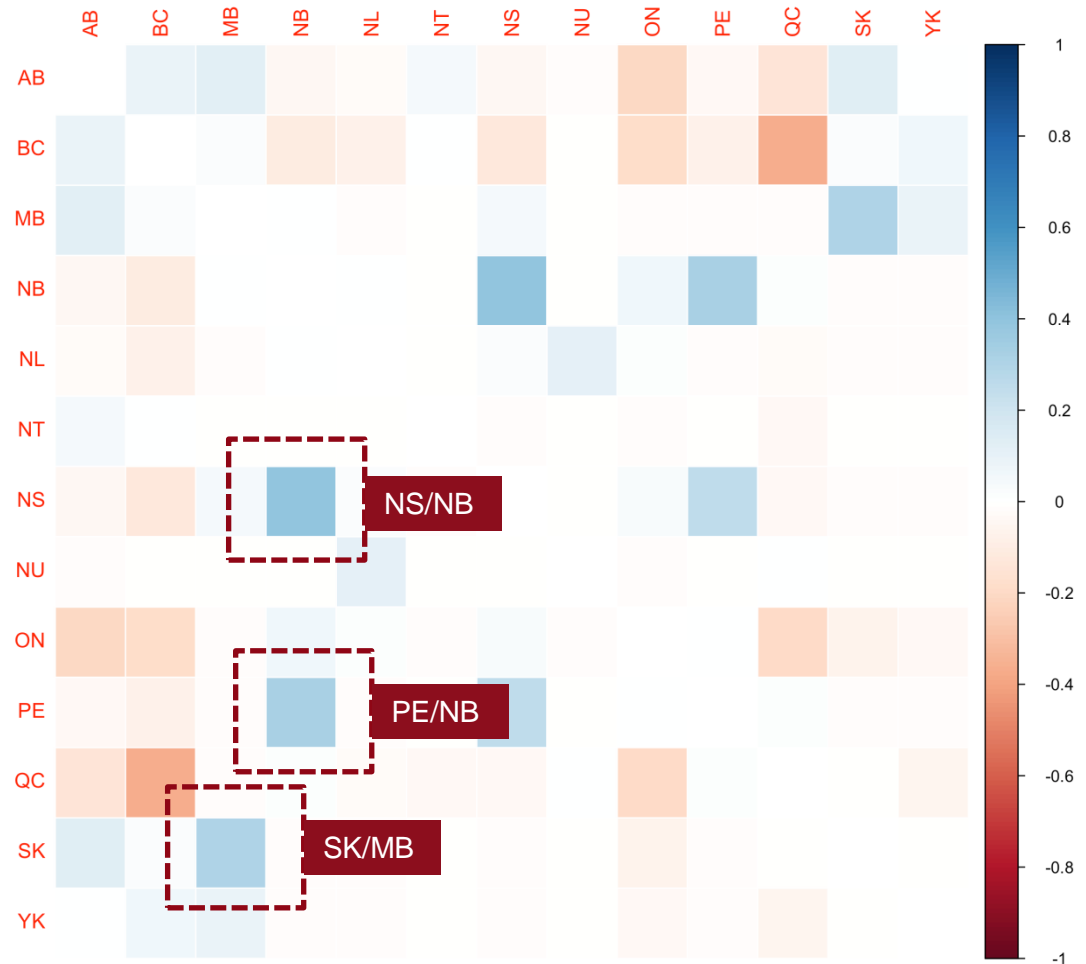
Despite relatively large co-occurrence between ON and BC or QC and BC, these provinces are actually negatively correlated with one another.

Most visitors to BC do not also visit ON or QC.



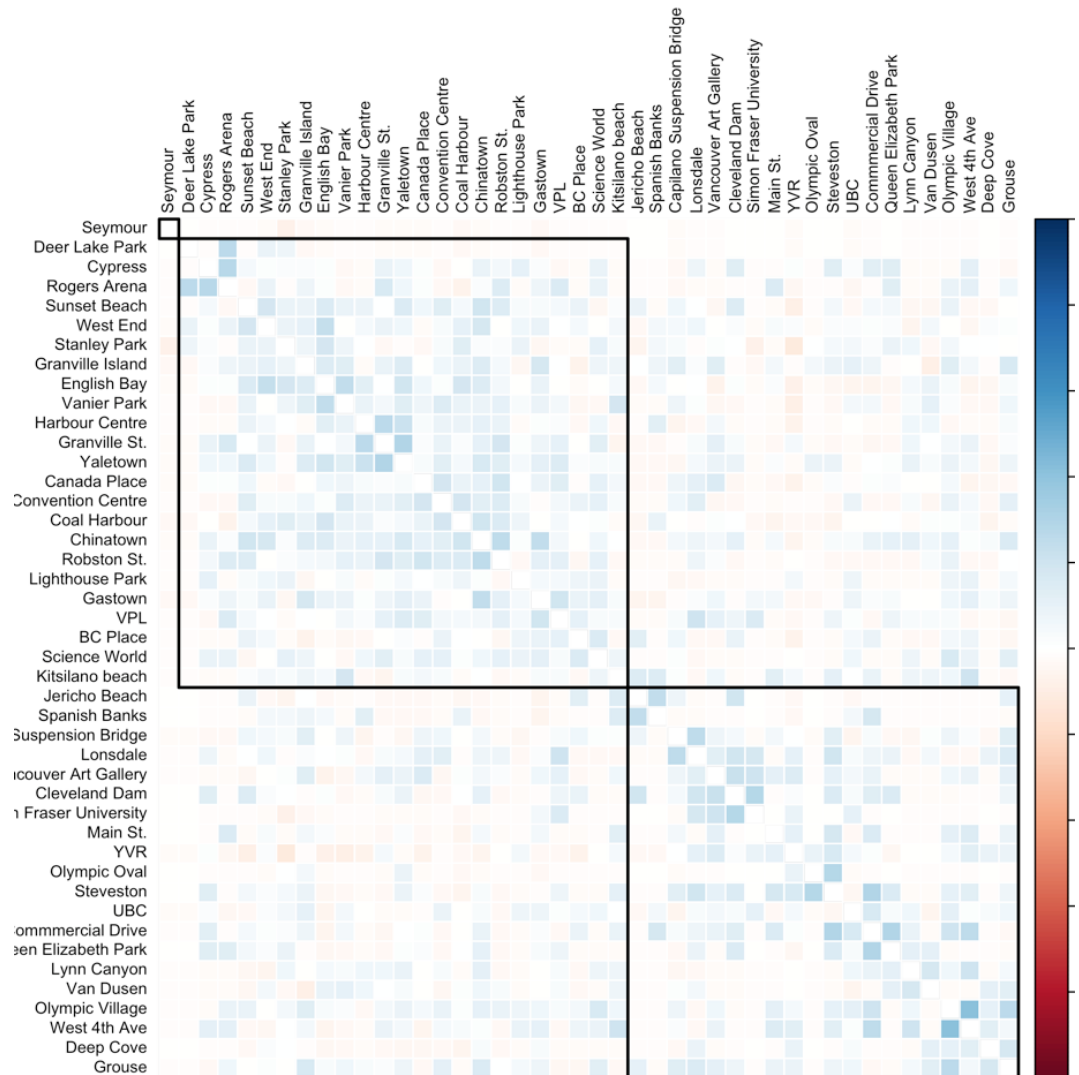
# Do tourists visit multiple provinces?

Provinces with blue squares are positively correlated.



# What areas of a city are jointly visited?

Defining a large number of attractions and areas, we see if natural clusters emerge among destinations.



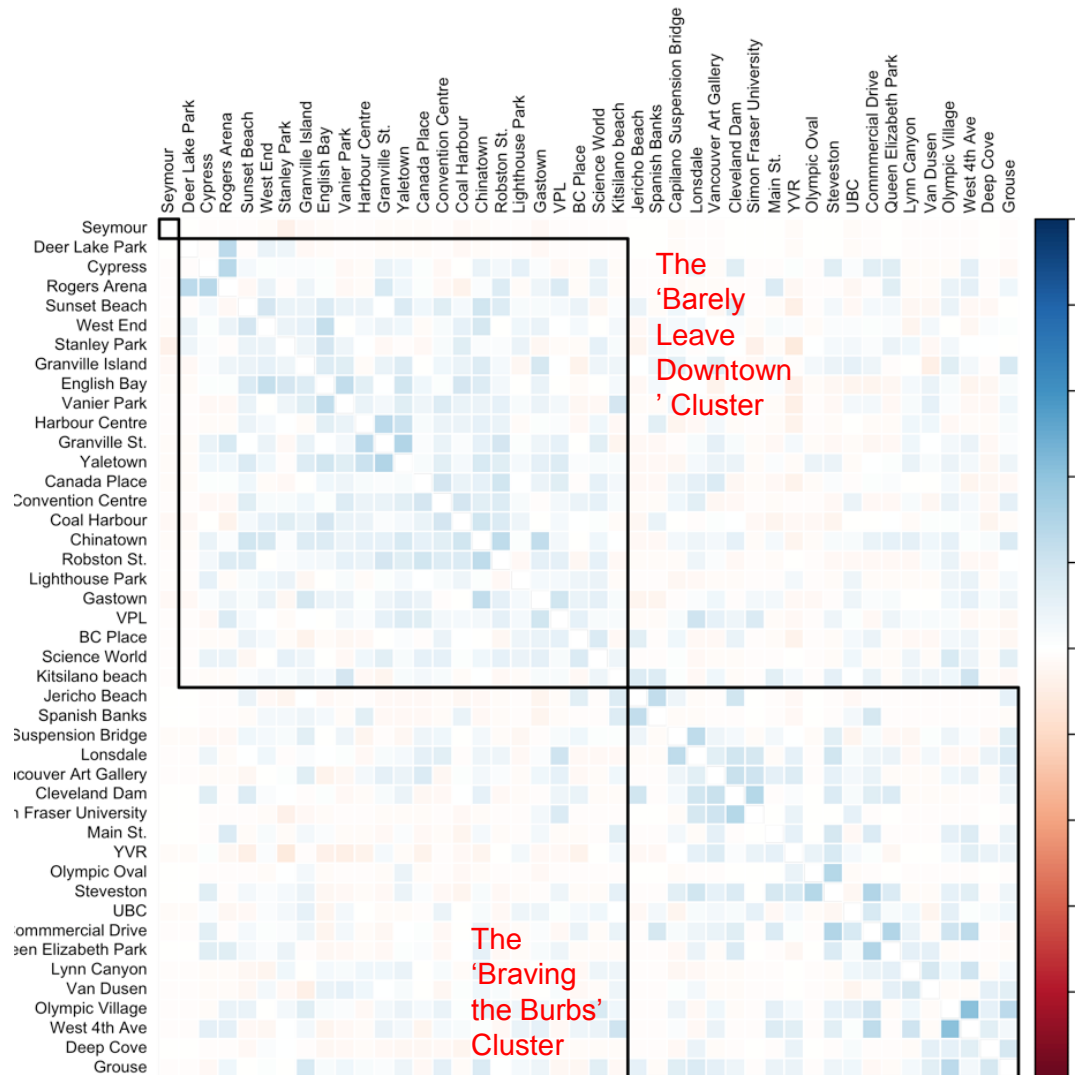


# What areas of a city are jointly visited?

Using a hierarchical clustering algorithm, we can sort our table into a pair of clusters.

Attractions further away from downtown are more closely correlated with each other.

One group of tourists will barely leave downtown, while the other is willing to explore a little bit more.

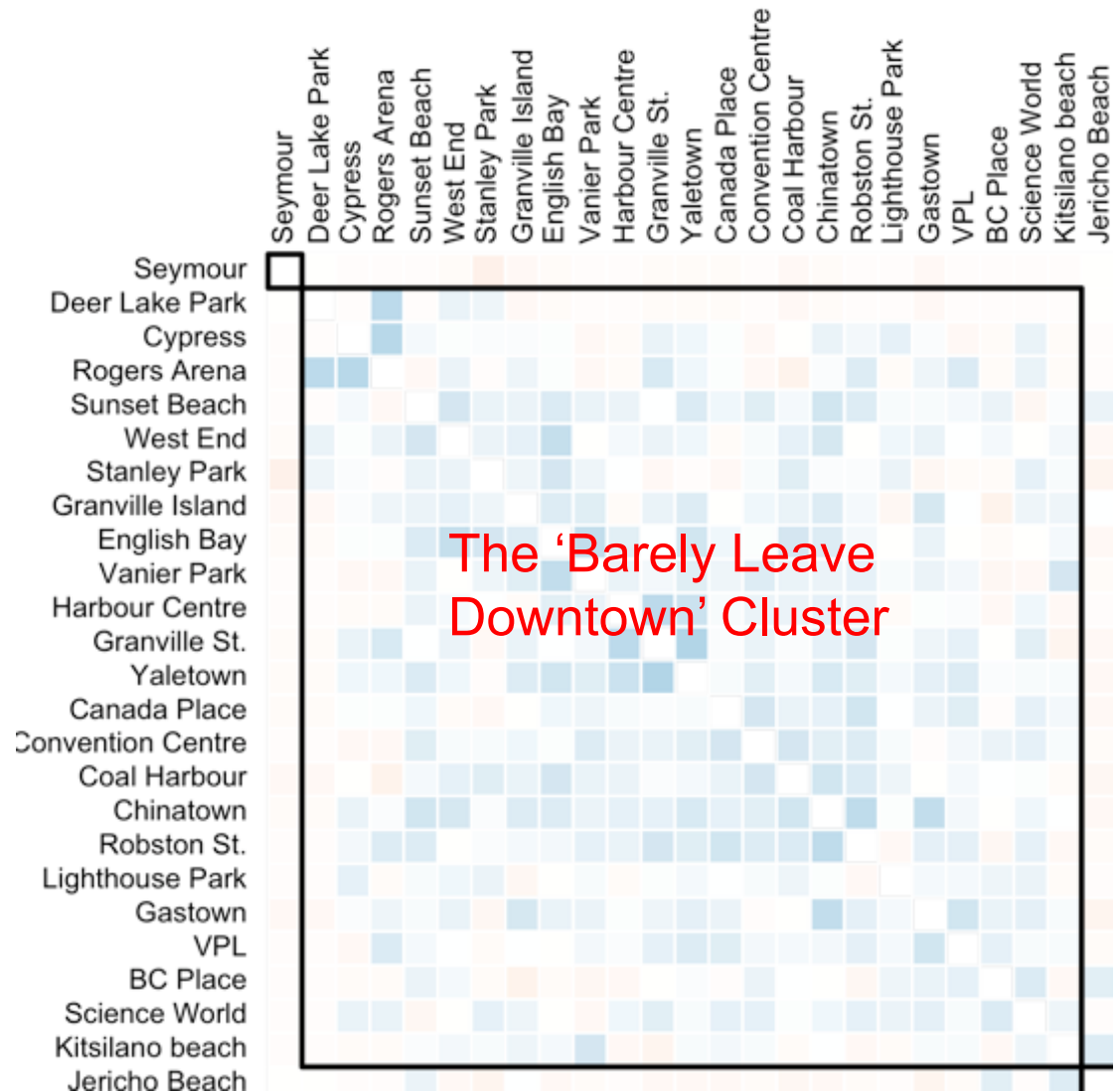


# What areas of a city are jointly visited?

Using a hierarchical clustering algorithm, we can sort our table into a pair of clusters.

Attractions further away from downtown are more closely correlated with each other.

One group of tourists will barely leave downtown, while the other is willing to explore a little bit more.



# Outcomes

---

That we used photo data is irrelevant.

This analysis can be adapted to any kind of data that:

- is large in scale
- has a time stamp and a location, and
- that can connected to cohort (e.g., country of origin, Millennials)

This can include:

- mobile phone data
- web-data and social media data
- transactional data (e.g. financial transactions)
- and on and on.

3

**Destinations: Deep dive  
at the regional level**

# Phase 2: Deep dive at the regional level

---

Cities, Regions and Provinces are making large investments based on data that is:

- Old (most recent is 2014)
- Incomplete (only provides information at the provincial level)
- Inaccurate (survey based data with small sample sizes)

Tourism Vancouver, a partner, has identified this approach could help solve:

- Economic value of cruise ships (increase / decrease #s)
- Economic value of conventions (increase infrastructure)
- Improved marketing, and itinerary building, including how to work with other destination marketing organizations

# Destination Canada's Mobile Phone Project

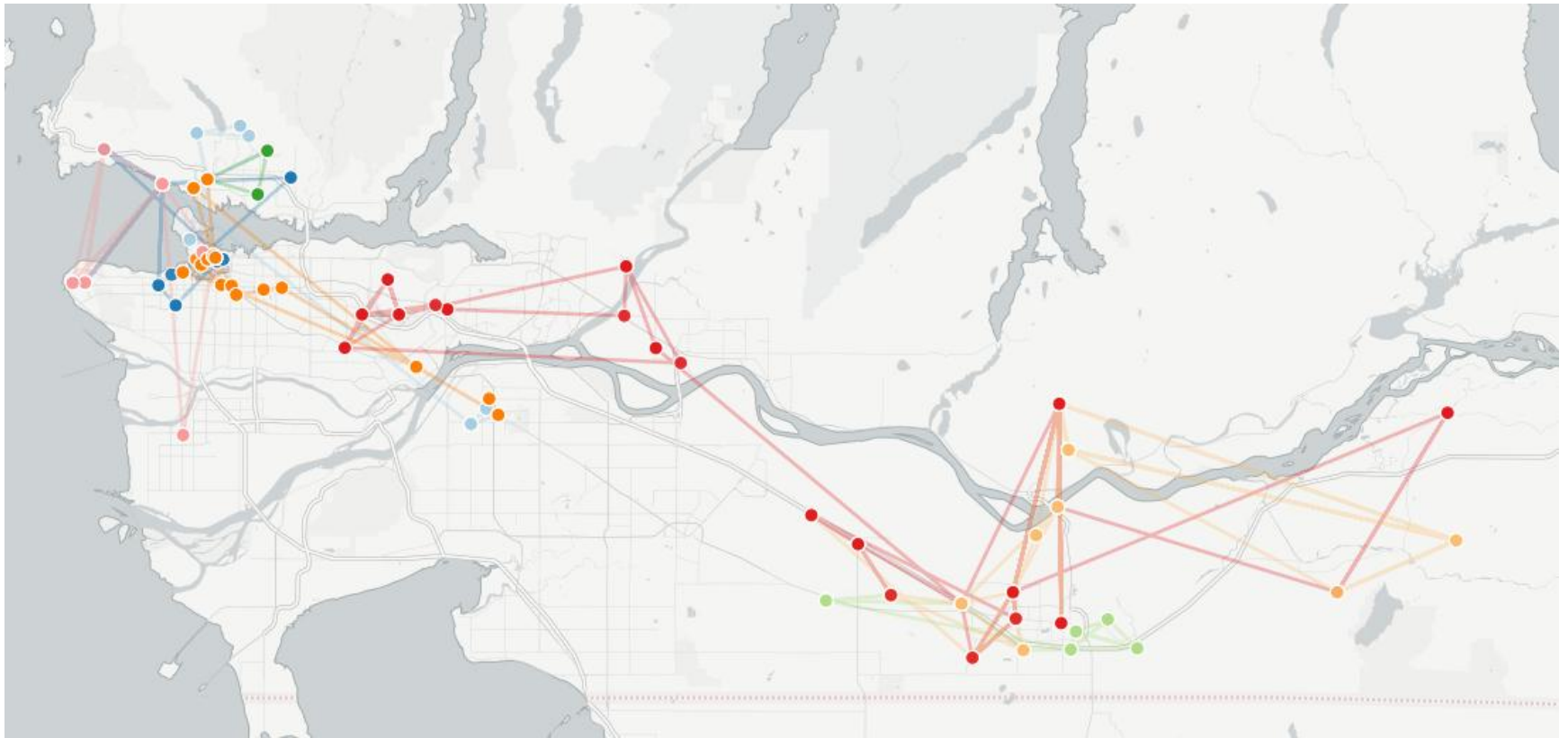
---

Using mobile phone signals within a region to understand tourist movements

## Objectives:

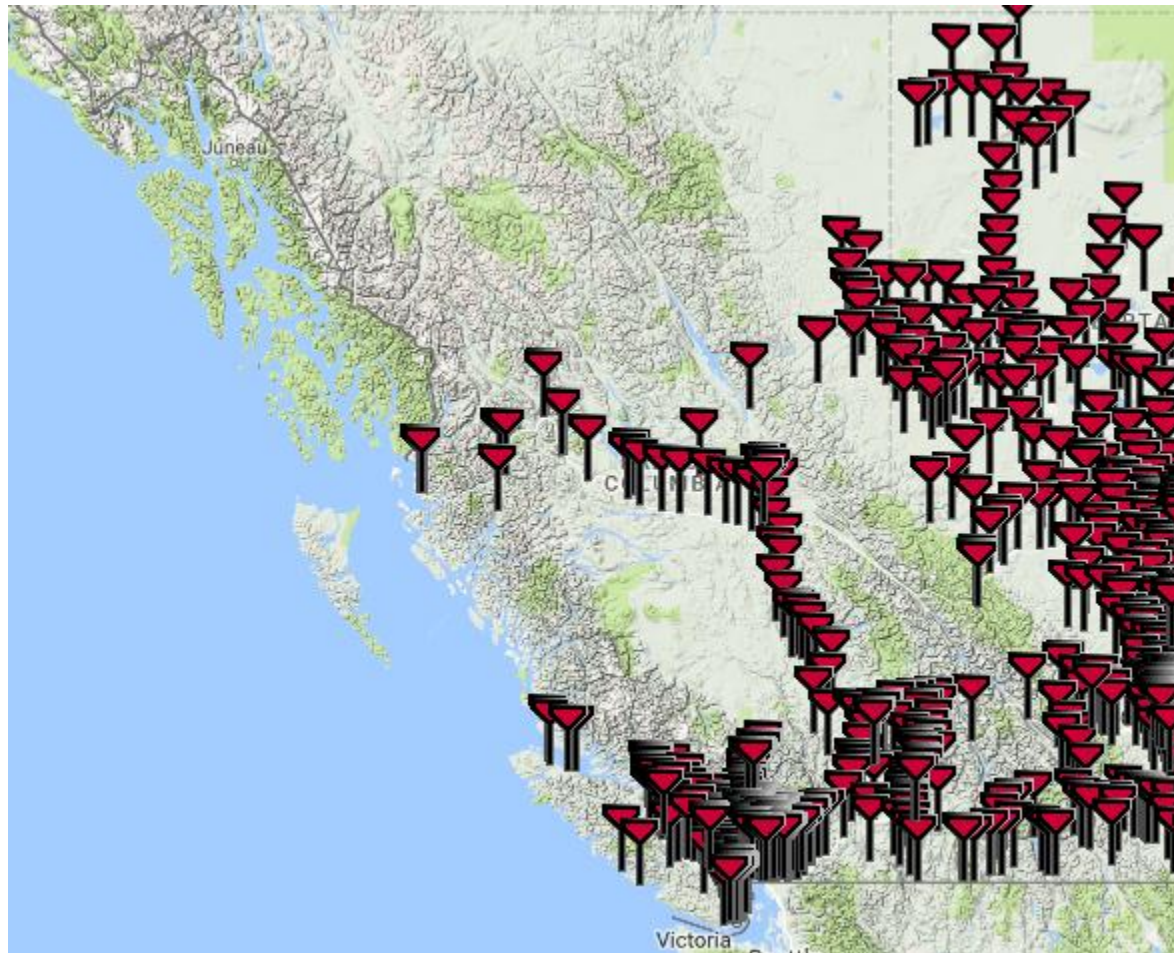
- Determine if and to what extent mobile phone data can provide:
  - accurate,
  - timely and
  - cost-effective information on tourist movements
- Develop processes to clean and process data
- Identify and address any legal and privacy issues of using this data
- Apply this information to real business needs of tourism sector
- Build analytical capacity within DC and partners
- Understand how to link passive, observational data to existing survey data

# Example mobile phone data (7 devices)





# Cell Phone Towers in BC (LTE, GSM, GSM)





# Expected Challenges

---

Working with massive volumes of data

Evaluating data in context of other data sources (ITS, TSRC, Global Tourism Watch)

Responsibility to maintain acceptable privacy standards

Managing realistic expectations of what we can do with the data

# Realistic Expectations

---

We expect to use this data to:

- Identify where tourists from different markets are visiting in Metro Vancouver and British Columbia
- Estimate duration of stay (in an area/in a city/in a region/in BC) for visitors
- Evaluate which areas and regions tend to be jointly visited

More difficult but realistic goals:

- Correlate with other data (social media, TripAdvisor) to generate some idea about tourist activities
- Analyze whether there exist temporal differences in tourist movement (e.g. by month, by day of week)
- Analyze whether external factors like weather affect movement

# Realistic Expectations

---

We will not have the following information:

- Personal traveller and travel party characteristics
- Spending data
- Fully-accurate attraction and activity data
- Trip-purpose

And that's ok.

# Questions?



Dmitry Shkolnik  
shkolnik.dmitry@destinationcanada.com  
tw: @dshkolnik /dshkolnik