Visualizing Linguistic Diversity in Vancouver

CPSC 547 Project Presentation

Anika Sayara, Namratha Rao, Roger Yu-Hsiang Lo

10 December, 2020

Road Map

- Motivation and goal / Why?
- Data / What?
- Visualization / How?
- Demo

7000 +

Languages spoken globally today

79 th

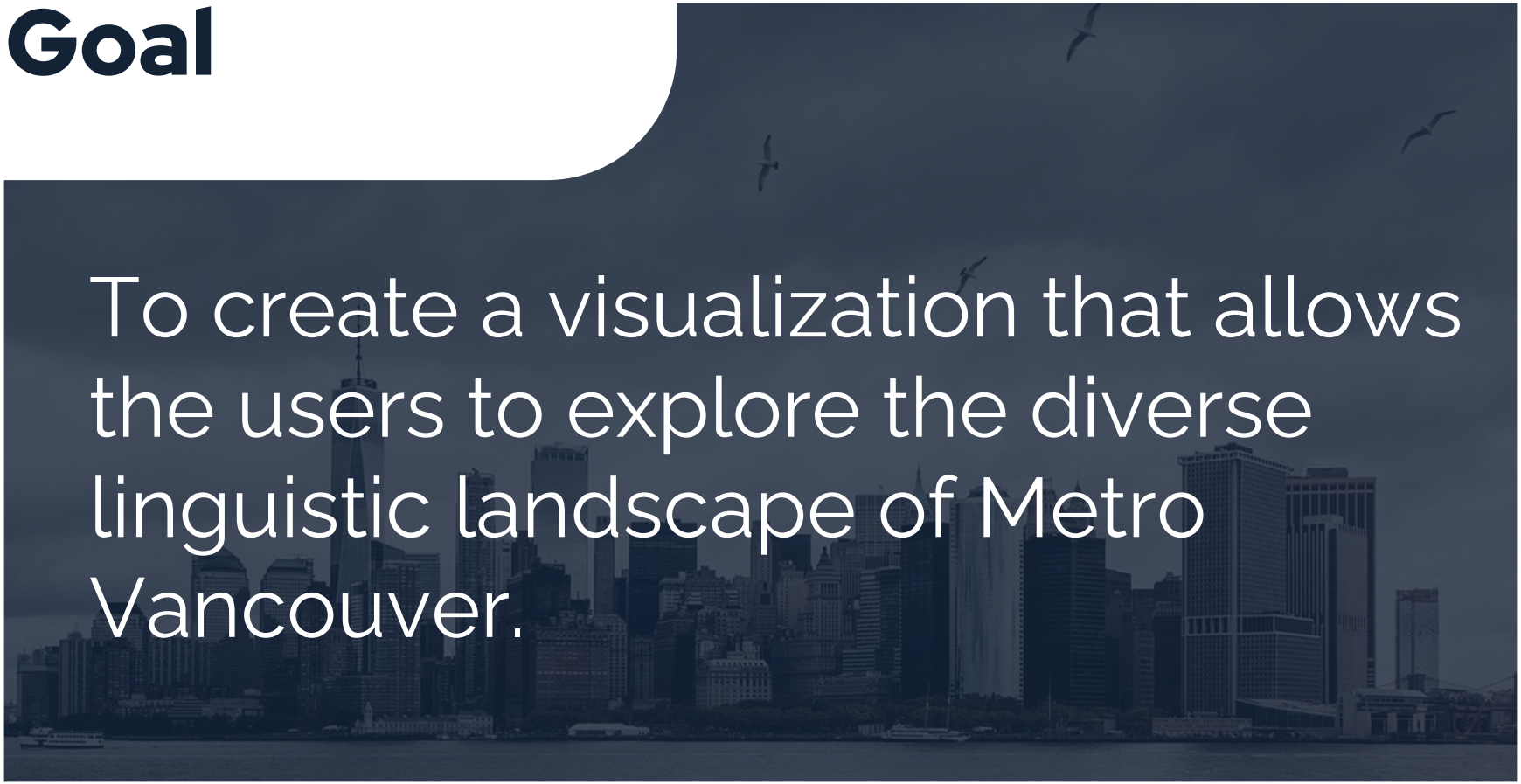
Canada's linguistic diversity rank

14.5%

Rise in non-English/non-French speakers in Canada (2016)

“Have you ever wondered to what extent your language is spoken in different neighborhoods in Metro Vancouver?

Goal



To create a visualization that allows the users to explore the diverse linguistic landscape of Metro Vancouver.

What?

What?

- GeoJSON format, 3450 tuples & 230 attributes.
- Each tuple represents a single dissemination area.

- ***Dissemination Area (DA):*** Smallest standard geographic area for which all census data are disseminated.

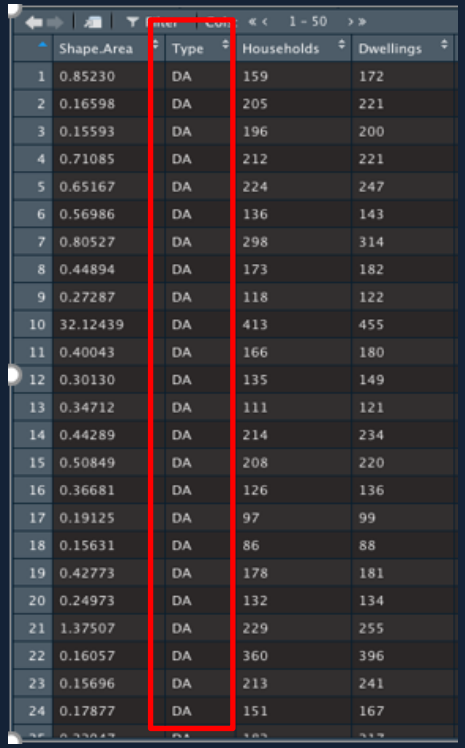
Data View

Dissemination Area

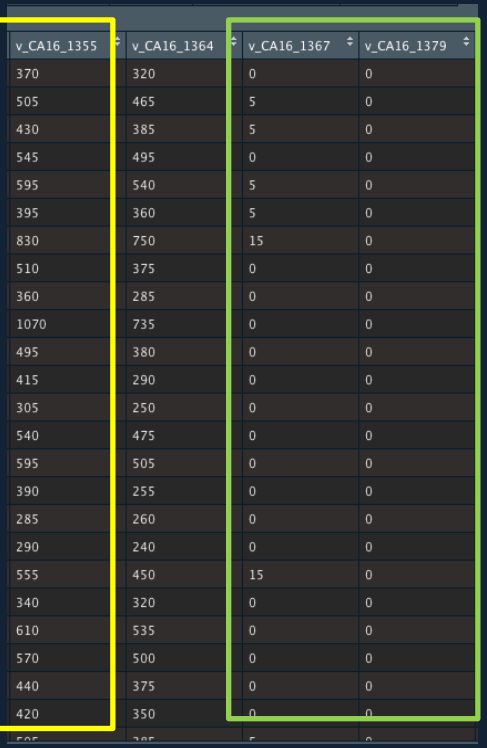
Total # of lang speakers

lang1, lang2..

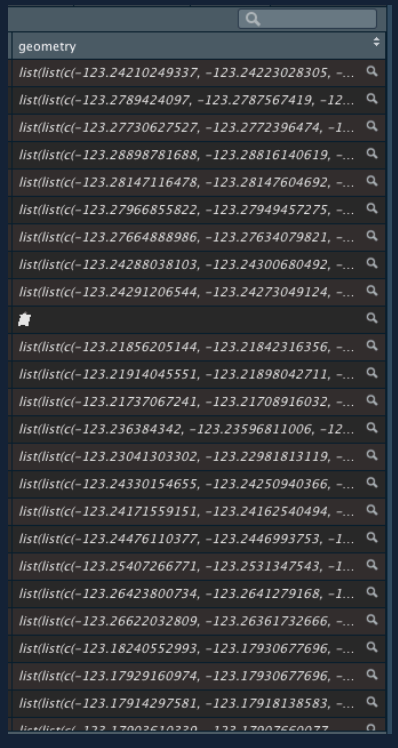
Shape coordinates



	Shape.Area	Type	Households	Dwellings
1	0.85230	DA	159	172
2	0.16598	DA	205	221
3	0.15593	DA	196	200
4	0.71085	DA	212	221
5	0.65167	DA	224	247
6	0.56986	DA	136	143
7	0.80527	DA	298	314
8	0.44894	DA	173	182
9	0.27287	DA	118	122
10	32.12439	DA	413	455
11	0.40043	DA	166	180
12	0.30130	DA	135	149
13	0.34712	DA	111	121
14	0.44289	DA	214	234
15	0.50849	DA	208	220
16	0.36681	DA	126	136
17	0.19125	DA	97	99
18	0.15631	DA	86	88
19	0.42773	DA	178	181
20	0.24973	DA	132	134
21	1.37507	DA	229	255
22	0.16057	DA	360	396
23	0.15696	DA	213	241
24	0.17877	DA	151	167



v_CA16_1355	v_CA16_1364	v_CA16_1367	v_CA16_1379
370	320	0	0
505	465	5	0
430	385	5	0
545	495	0	0
595	540	5	0
395	360	5	0
830	750	15	0
510	375	0	0
360	285	0	0
1070	735	0	0
495	380	0	0
415	290	0	0
305	250	0	0
540	475	0	0
595	505	0	0
390	255	0	0
285	260	0	0
290	240	0	0
555	450	15	0
340	320	0	0
610	535	0	0
570	500	0	0
440	375	0	0
420	350	0	0



geometry
list(list(c(-123.24210249337, -123.24223028305, -...
list(list(c(-123.2789424097, -123.2787567419, -12...
list(list(c(-123.27730627527, -123.2772396474, -1...
list(list(c(-123.28898781688, -123.28816140619, -...
list(list(c(-123.28147116478, -123.28147604692, -...
list(list(c(-123.27966855822, -123.27949457275, -...
list(list(c(-123.27664888986, -123.27634079821, -...
list(list(c(-123.24288038103, -123.24300680492, -...
list(list(c(-123.24291206544, -123.24273049124, -...
list(list(c(-123.21856205144, -123.21842316356, -...
list(list(c(-123.21914045551, -123.21898042711, -...
list(list(c(-123.21737067241, -123.21708916032, -...
list(list(c(-123.236384342, -123.23596811006, -12...
list(list(c(-123.23041303302, -123.22981813119, -...
list(list(c(-123.24330154655, -123.24250940366, -...
list(list(c(-123.24171559151, -123.24162540494, -...
list(list(c(-123.24476110377, -123.2446993753, -1...
list(list(c(-123.25407266771, -123.2531347543, -1...
list(list(c(-123.26423800734, -123.2641279168, -1...
list(list(c(-123.26622032809, -123.26361732666, -...
list(list(c(-123.18240552993, -123.17930677696, -...
list(list(c(-123.17929160974, -123.17930677696, -...
list(list(c(-123.17914297581, -123.17918138583, -...

Categorical Attributes

Name	Description	Cardinality	Sample value
TYPE	Constant value field indicating DA	1	‘DA’
REGION NAME	The name of the neighborhood the DA is in	25	‘Richmond’
GEOMETRY	Polygon geometry of the DA	3450	POLYGON ((-123.28147116478 49.36803241352), ...)

Quantitative Attributes

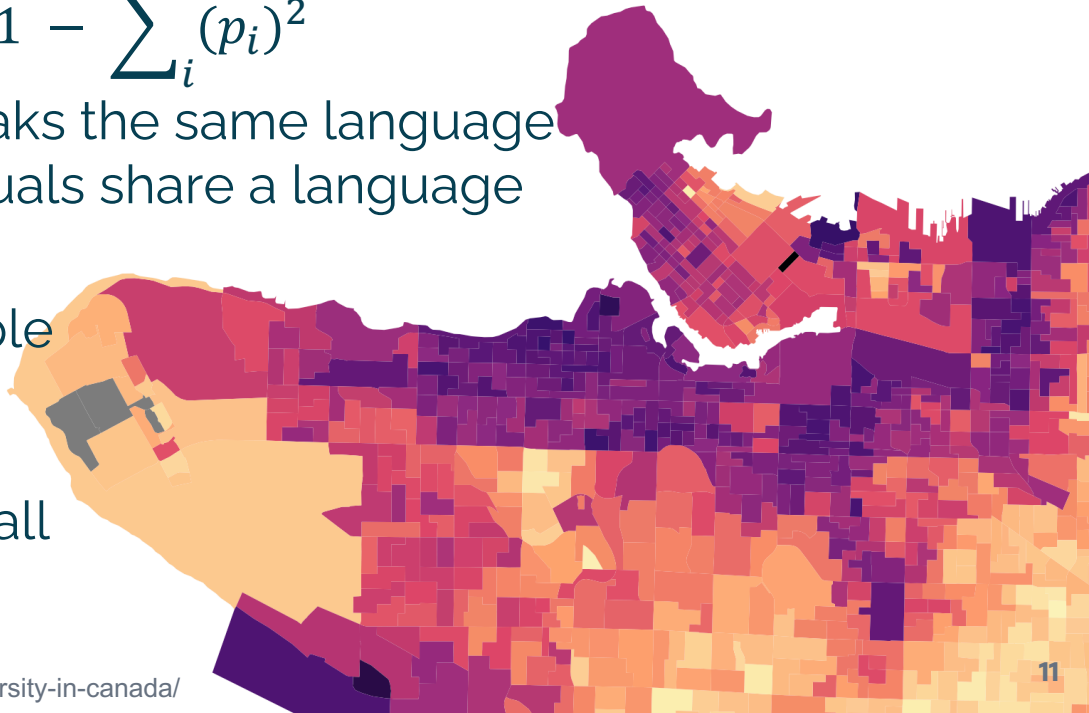
Name	Description	Min	Max	Median
HOUSEHOLDS	The number of families in the DA.	0	4923	213
DWELLINGS	The number of dwelling units in the DA	0	5631	219
POPULATION	The total number of people in the DA	0	8778	586
SHAPE AREA	Area of the DA in sq. km	0.002	846.8	0.1
V_CA16_1367	The number of speakers of language 1367 (French)	0	85	0

How?

- Which neighbourhood is most diverse linguistically?
  - **Measure diversity**: Language Diversity Index (Greenberg, 1956)

$$LDI = 1 - \sum_i (p_i)^2$$

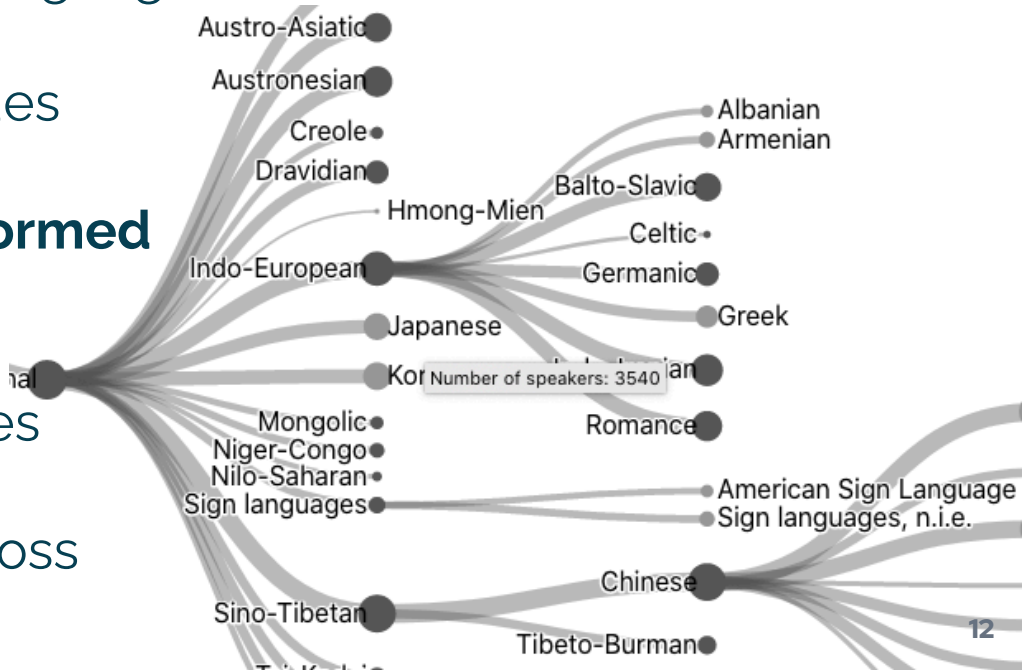
  - **LDI = 0**: everyone speaks the same language
  - **LDI = 1**: no two individuals share a language
- Choropleth map
  - **Familiar** to most people
  - LDI is **colour-coded**
  - Limitations
    - Some DAs are small



https://www.dshkol.com/2017/language-diversity-in-canada/

How?

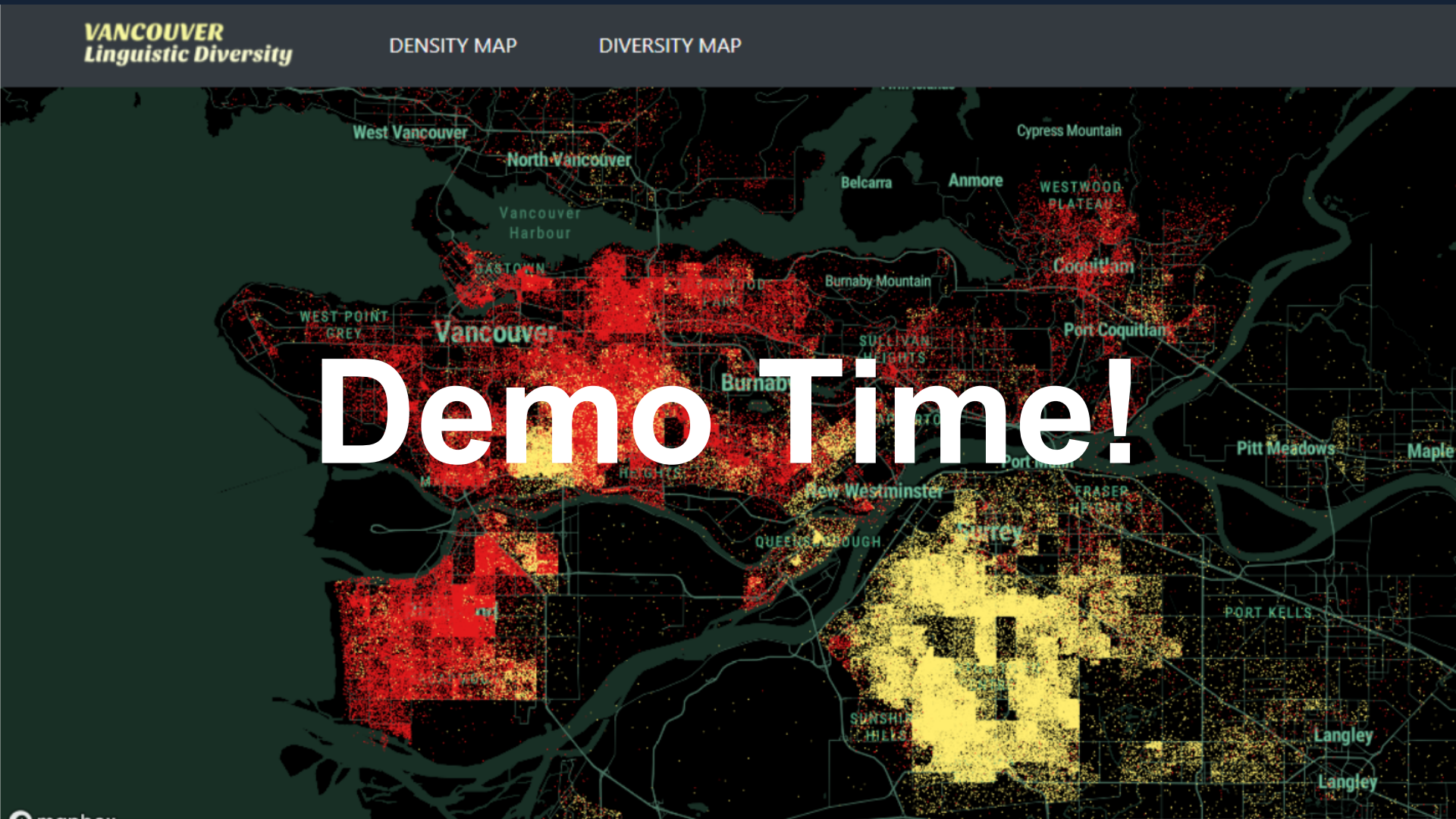
- **What languages** are spoken in a neighbourhood? **How many speakers** does each language have?
- Collapsible tree
  - **Aggregation** based on language relatedness
  - Number of speakers coded by the **size** of nodes and links
  - Numbers are **log-transformed** to fit the screen
  - Limitations:
    - Log scale compresses actual difference
    - Hard to compare across families



How?

- How different languages are **distributed** geographically?
- Dot density map
  - A dot represent a single speaker of a language
  - Languages coded by hues
  - **Clusters** are easily observed
  - Limitations
    - Only allows comparison among a handful of languages
    - Occlusion of dots
    - Computationally expensive

https://labs.mapbox.com/education/thematic-map-types/dot-density/


VANCOUVER
Linguistic Diversity
DENSITY MAP
DIVERSITY MAP
West Vancouver
North Vancouver
Cypress Mountain
Belcarra
Anmore
WESTWOOD PLATEAU
Vancouver Harbour
GASTOWN
Burnaby Mountain
Coquitlam
WEST POINT GREY
Vancouver
Port Coquitlam
SULLIVAN HEIGHTS
Burnaby
Demo Time!
Pitt Meadows
Maple
New Westminster
FRASER
Surrey
QUEENSBOROUGH
PORT KELLS
SUNSHINE HILLS
Langley
Langley