

# Text-mining Airbnb reviews:

## Exploring Indoor Environmental Quality

By: Hannah Villeneuve | Supervisor: Prof. Liam O'Brien

# Presentation Outline

- Objectives
- About Airbnb
- Indoor Environmental Quality (IEQ)
- Methodology
- Findings

# Objectives

- To explore multi-aspect IEQ issues
- To explore use of unsolicited reviews to evaluate homes
- To inform hosts and those who design, operate, and maintain buildings of the most critical and complained about problems

# Airbnb: A peer-to-peer shared accommodation arrangement service

- Fast Facts:
  - Founded in 2008 ([Airbnb, 2019](#)).
  - 99800 listings worldwide
  - Listings range from a room in someone's house to castles
  - Guests and hosts can review; reviews posted simultaneously ([Bridges and Vasquez, 2016](#)).
  - Reviews are public



Alex

Overall our groups experience was great. We liked the place being just off the busy town but being able to use scooters provided by Nora's. When you arrive you will feel the actual place is better than the pictures. Nice and helpful staff and great pool. Will stay again if ever in Bohol.



 [New VIP Studio 6 with pool & FREE RENTAL BIKE](#)



Armarie

Two thumbs up! Generally, remarkable and great service and quality of rooms. Kitchen and free motorbike was steal! Hope to come back to Nora's Place and make it a yearly ritual.



 [Superior Studio 7 & FREE MOTORBIKE & swimming pool](#)

# Indoor Environmental Quality



## Acoustic Comfort

- Noise levels



## Indoor Air Quality (IAQ)

- Air pollutants, odours



## Thermal Comfort

- Heating/cooling
- Humidity, ventilation levels

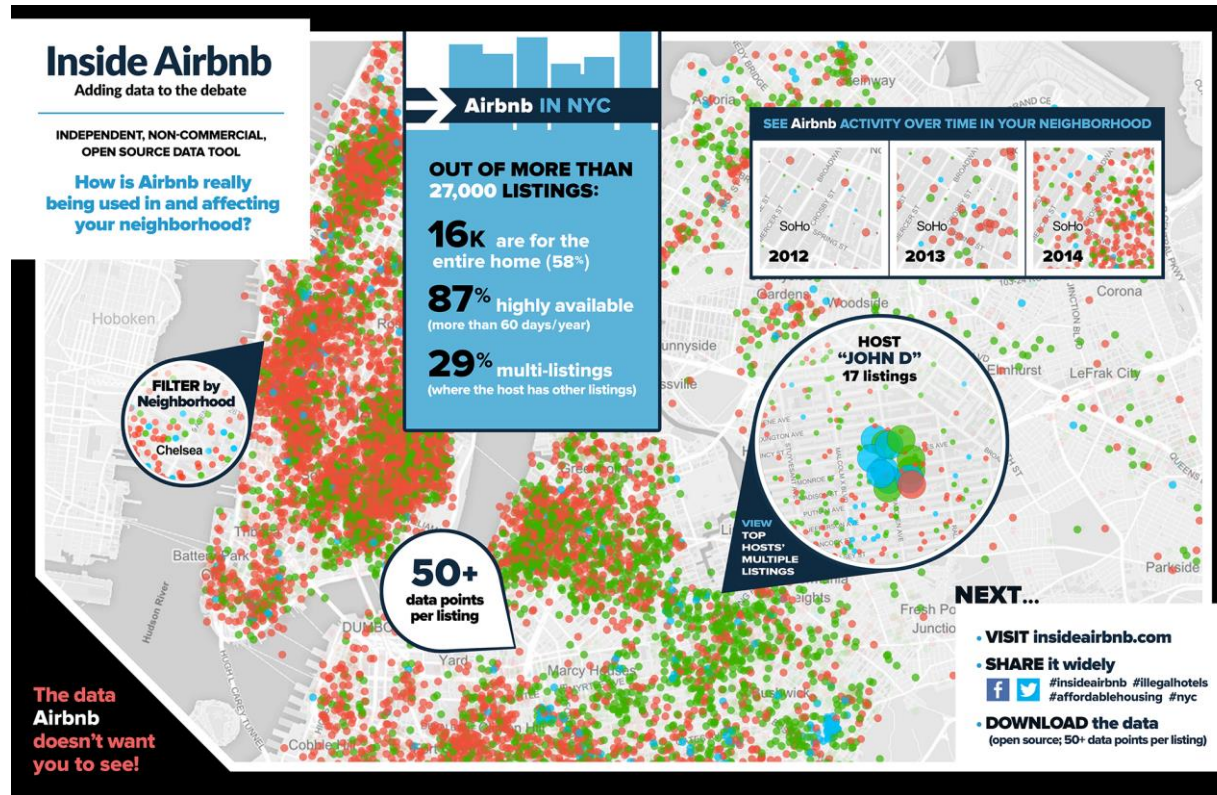


## Visual Comfort

- Lighting

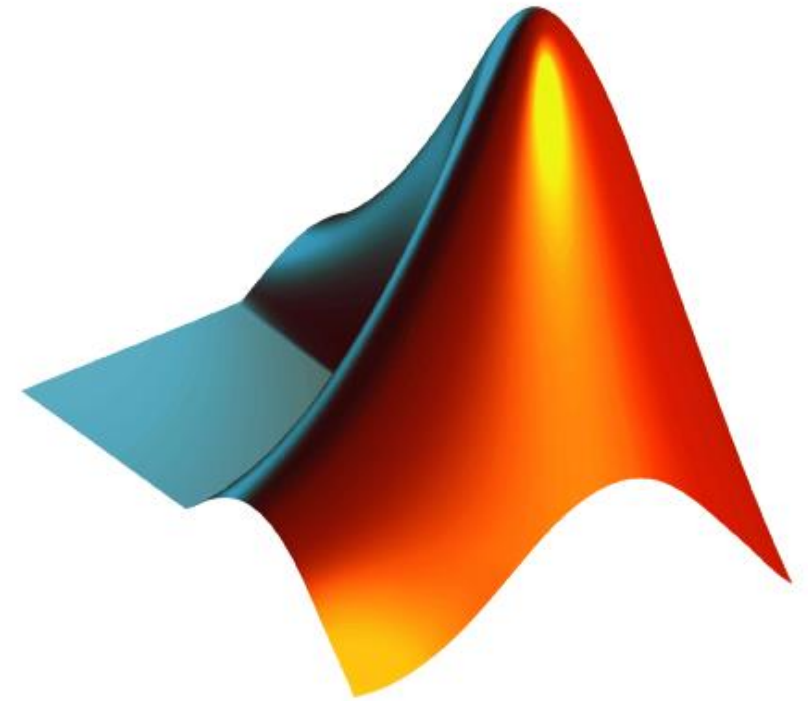
# Methodology

- Data came from Inside Airbnb → 6 cities, 10 years, 1.35 million reviews



# Methodology: Processing the Data

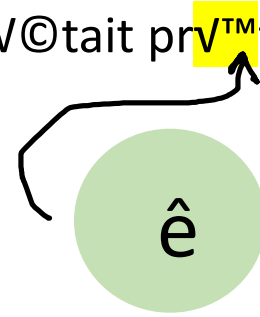
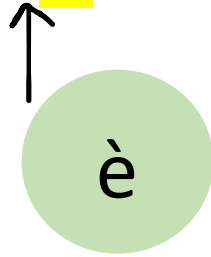
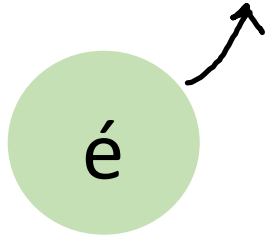
- **Find English comments**
- **Calculate sentiment scores for each comment**





## Removing non-English Comments:

“La chambre louv<sup>©</sup>e v<sup>©</sup>tait fid<sup>®</sup>le aux photos et aux conditions v<sup>©</sup>voquv<sup>©</sup>es. Tout v<sup>©</sup>tait pr<sup>v<sup>TM</sup></sup>t v<sup>†</sup>...”



## Calculating Sentiments:

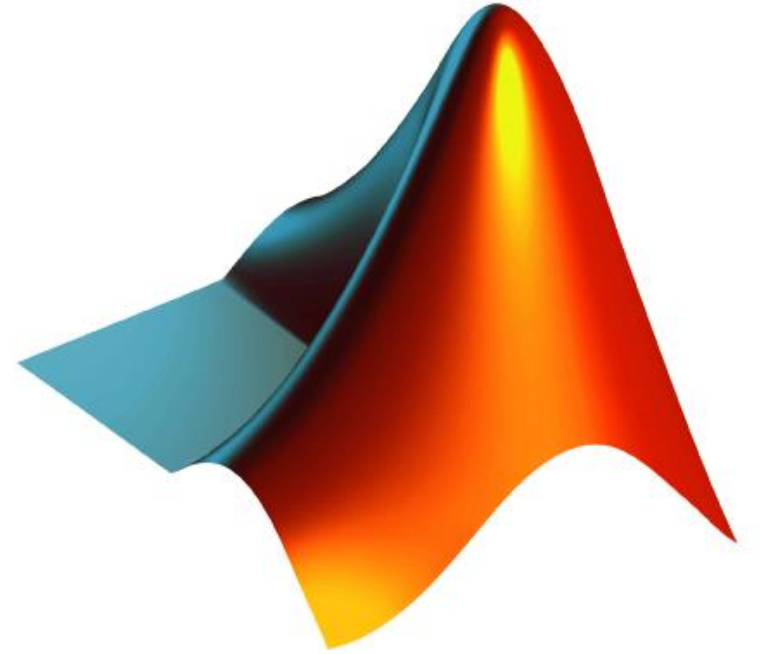
“Our trip was too short in Ottawa, we really wished we could have stayed longer because its a beautiful city. Donna was lovely, we didn't have much time to get to know her unfortunately.”

$$\text{sentiment score} = \log \left( \frac{\text{Positive} + 0.5}{\text{Negative} + 0.5} \right) = \log \left( \frac{2 + 0.5}{1 + 0.5} \right) = 0.2218$$



# Methodology: Processing the Data

- Find English comments
- Calculate sentiment scores for each comment
- **Search for key IEQ terms (ex. Noise, odour, warm, dim)**

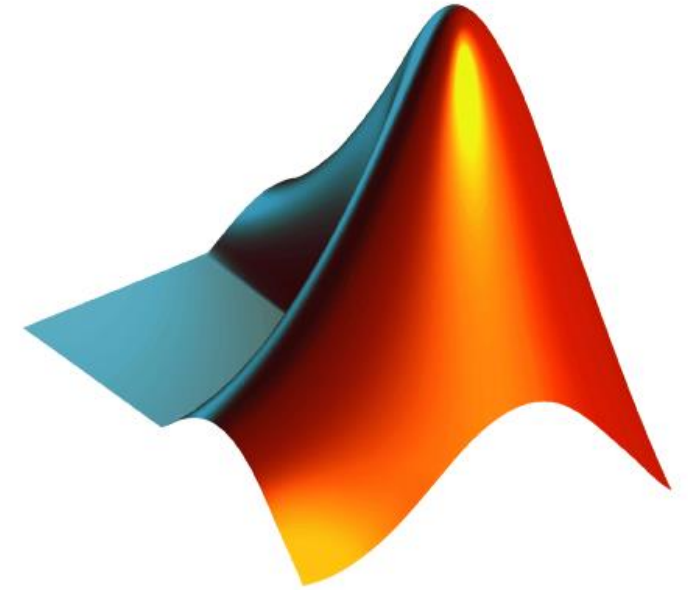


# List of IEQ Search Terms

Acoustic	IAQ	Thermal	Visual
Noise/Noisy	Odour/odor	Cold	Glare
Loud	Smell	Draughty/Drafty	Blinding
Blaring	Musty	Humid	Dim
Sound	Dusty	Chilly	Unlit
Racket	Allergy	Freezing	
Ear-splitting	Asthma	Scorching	
Heard/Hear	Stink	Sweltering	
Squeak	Stench	Hot	
Creak	Reek	Boiling	
Vibrations	Stale	Warm	
	Dry	Frozen	
	Steamy	Frigid	
	Damp	Shiver	

# Methodology: Processing the Data

- Find English comments
- Calculate sentiment scores for each comment
- Search for key IEQ terms (ex. Noise, odour, warm, dim)
- **Exclude comments with irrelevant instances of IEQ terms**
  - Ex. 'warm your soul', 'dim sum'

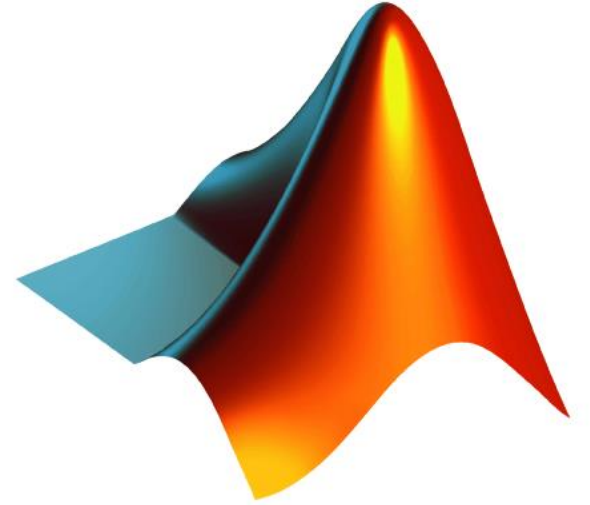


# Problem of Weather-Related Thermal Terms

- Indicates comfort:
  - "Air conditioning was very welcome during the very **hot weekend** we were there."
  - "The house stayed **very warm** despite crazy **cold weather** outside."
- Indicates discomfort:
  - "The only thing that was a bit of a surprise was that the a/c was a window unit. The apartment was **quite hot** when we arrived and it made sleeping a bit tricky the first night. It was a very **hot weekend** though. The unit **did eventually cool** the apartment."
  - "It was really **cold weather** outside during my stay. The apartment itself is **warm**, but you can feel **air blowing a bit from the windows**."

# Methodology: Processing the Data

- Find English comments
- Calculate sentiment scores for each comment
- Search for key IEQ terms (ex. Noise, odour, warm, dim)
- Exclude comments with irrelevant instances of IEQ terms
  - Ex. 'warm your soul', 'dim sum'
- **Sort the thermal complaints by season**
- **Look for comments with terms from multiple categories of IEQ**
- **Find causes:**
  1. Find top words in comments with IEQ terms
  2. Find top words in general comments
  3. Compare frequencies of top words that appear in IEQ comments and general comments



## Processing the Data: Finding Causes

Acoustic Top Words		IAQ Top Words	
word	difference	word	difference
hear	0.2906	dry	0.16071
night	0.18057	bit	0.1082
street	0.13559	building	0.10417
upstairs	0.12874	apartment	0.08978
bit	0.11735	bathroom	0.076321
sound	0.093924	house	0.07294
people	0.088332	old	0.0677
heard	0.087561	basement	0.065058
sleep	0.073326	towels	0.065004
traffic	0.06612	room	0.06284
door	0.065353	cigarette	0.057537
unit	0.059087	floor	0.056643
neighbours	0.056692	kitchen	0.05446
floor	0.053848	bed	0.052137
quite	0.052455	unit	0.051781
busy	0.050798	air	0.049985
little	0.050538	quite	0.042256
building	0.047315	overall	0.039535
basement	0.046377	window	0.039097
light	0.046159	itself	0.037628
bedroom	0.045651	night	0.037054
outside	0.043972	hair	0.036477
morning	0.041805	door	0.036042
music	0.04012	windows	0.035434
walls	0.037172	wash	0.035276
though	0.036923	inside	0.03494
guests	0.03626	little	0.034245
main	0.035285	sheets	0.034027

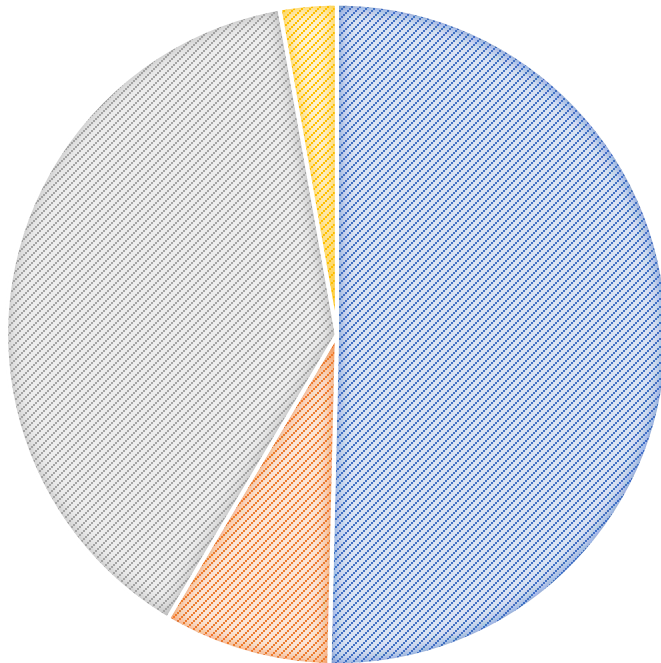
Thermal Top Words		Visual Top Words	
word	difference	word	difference
room	0.11198	room	0.23472
air	0.081463	night	0.19689
night	0.073568	light	0.17455
weather	0.067639	bedroom	0.17374
bit	0.062079	bit	0.16687
day	0.059021	bathroom	0.13685
summer	0.053729	sleep	0.13224
winter	0.049426	street	0.10985
bed	0.048457	lights	0.10865
shower	0.045736	basement	0.097822
heat	0.044225	small	0.087154
little	0.042343	curtains	0.086787
quite	0.040185	little	0.082858
conditioning	0.039785	door	0.070447
outside	0.039752	sleeping	0.06904
temperature	0.039696	windows	0.067555
basement	0.039469	find	0.064365
fan	0.039375	found	0.058637
unit	0.035937	window	0.054443
heater	0.03143	itâ	0.047712
bedroom	0.029751	floor	0.043003
heating	0.028394	day	0.041042
days	0.027008	living	0.039745
even	0.02674	outside	0.039246
floor	0.025907	arrived	0.038477
sleep	0.025811	mask	0.036451
got	0.02495	around	0.035696
get	0.023225	lighting	0.035353
though	0.02297	though	0.035276

# Findings: Prevalence of IEQ Complaints

- IEQ words per comment: 0.0662 in Toronto to 0.0865 in Vancouver

## DISTRIBUTION OF INSTANCES OF IEQ COMPLAINTS

■ Acoustic ■ IAQ ■ Thermal ■ Visual





# Findings: Co-occurrences of IEQ Terms

	Acoustic	IAQ	Thermal	Visual
Acoustic	45200	2380	5316	482
IAQ		7342	1923	205
Thermal			34704	389
Visual				2493

Acoustic, Visual, and Thermal	74
Acoustic, Thermal, and IAQ	378
Acoustic, Visual, and IAQ	24
Thermal, Visual, and IAQ	29

Acoustic, Visual, Thermal, and IAQ	
------------------------------------	--

# Findings: Co-occurrences of IEQ Terms

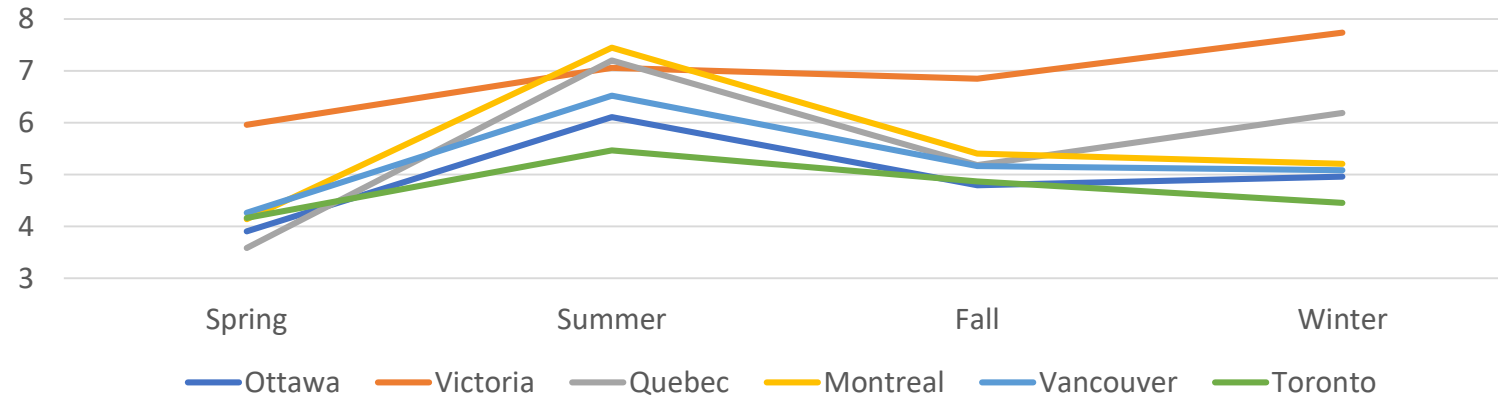
- **Thermal and IAQ:** "One of the 2 **ceiling fans did not work, and they were a bit dusty, so when we turned the one on because we had no ac** , dust fell all over the bed. 2 of the bedrooms-had no screens on the windows..."
- **Thermal and Acoustic:** "There wasn't an ac when I stayed in July 2017 and the fan didn't keep the room cool enough, so it felt like a **trade off between really loud or a little bit too hot when sleeping.**"

# Findings: Causes of Complaints

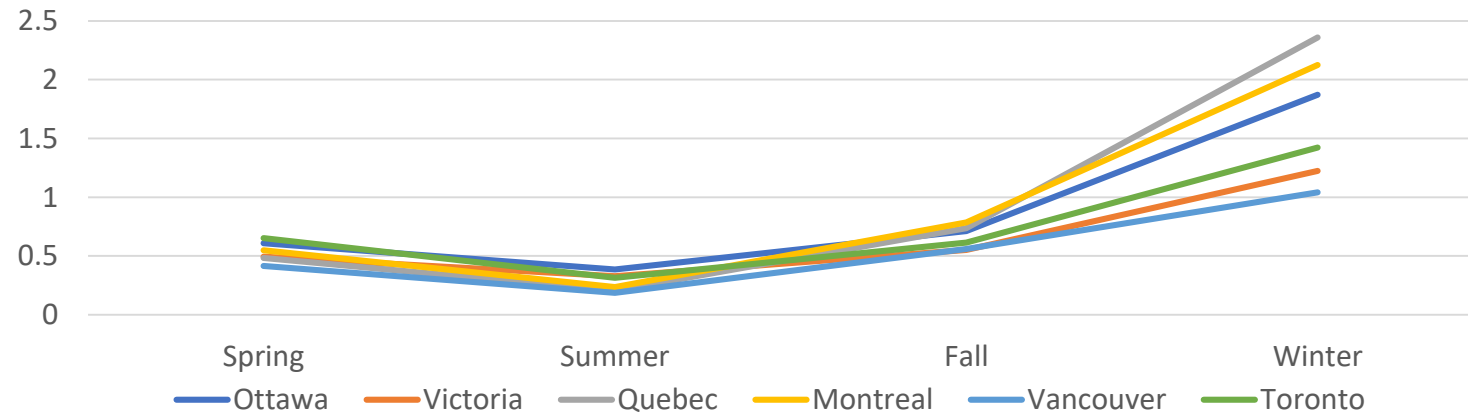
Acoustic	IAQ	Thermal	Visual
Streets	Bathroom	Air	Light
Upstairs	Cigarette	Weather	Street
People	Room	Summer	Curtains
Windows	Sheets	Conditioning	Basement
Traffic	Kitchen	Windows	Windows
Neighbors	Towels	Basement	Streets
Door	Basement	Temperature	Entrance
Construction	Hallway	Floor	Blinds
Bar	Mold	Upstairs	Bulbs
Fan	Cat	Heating	Lamps

# Findings: Thermal Complaints by Season

Percent of Comments per Season with Heat-Related Complaints

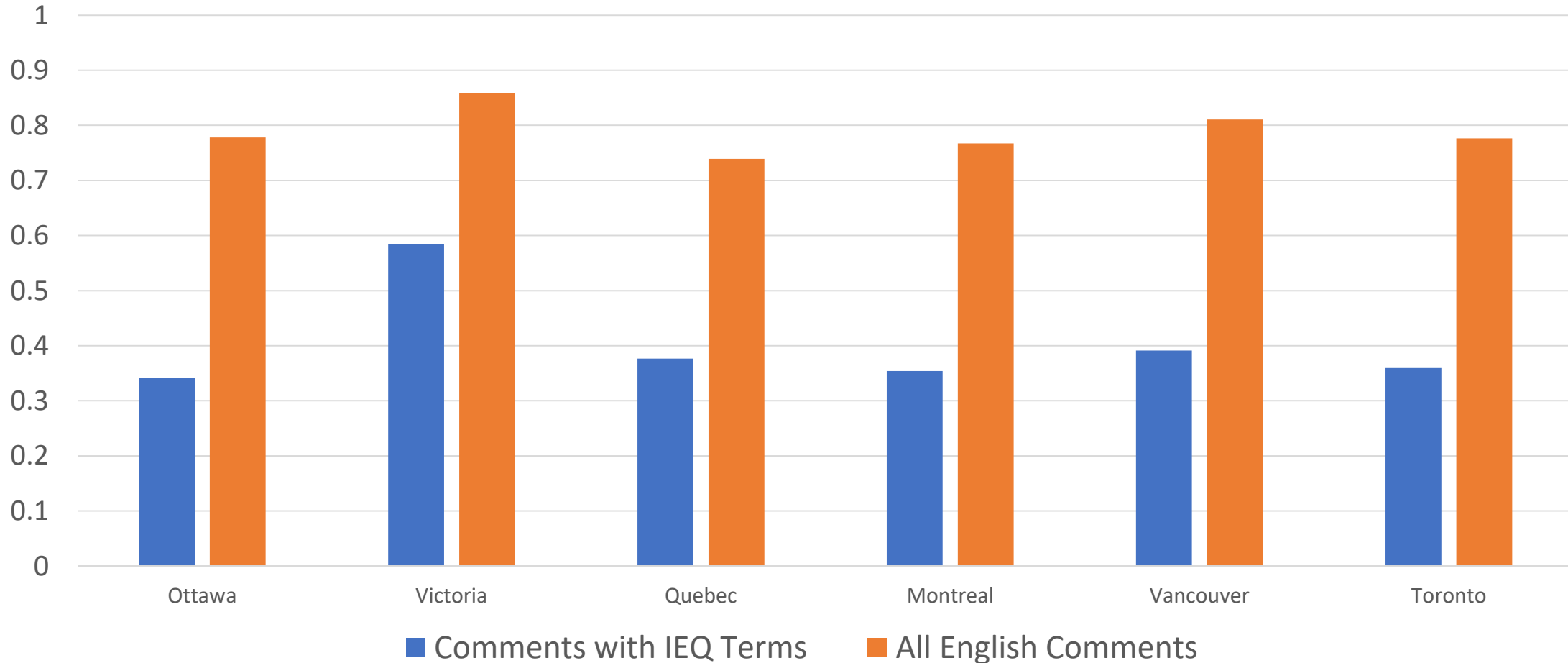


Percent of Comments Per Season with Cold-Related Complaints



# Findings: Effect of IEQ complaints on sentiment scores

## Average Sentiment Scores



# Some Interesting Comments

-And why qualitative analysis is important too!

- **Acoustic/IAQ:** "A drunk sang down the alley at 2am. The trash trucks came at 6 to service the restaurant next door. The seagulls are many and loud in the morning. You may walk out the door into a cloud of BC's smoke-able drug of choice."
- **IAQ:** "... The basement quickly became known to us as the onion chamber. Everything was clean, but man oh man, the smell of onions was potent. I mean this smell was thick, you walk to the basement and bam it was like walking into a rainforest of pure onion aroma."
- **Thermal/Acoustic:** "...There is some city noise when the windows are open as to be expected. If it's hot out , having a fan to keep cool and drown out the noise from the seagulls and city noise is key."

# Thank You

## Questions?

*[hannahvilleneuve@cmail.carleton.ca](mailto:hannahvilleneuve@cmail.carleton.ca)*



# References

Airbnb. "New Data: The Airbnb Advantage." *Airbnb*, <https://press.airbnb.com/new-data-the-airbnb-advantage/>. Accessed 12 June 2019.

Bridges, Judith, and Camilla Vásquez. "If nearly all *Airbnb* reviews are positive, does that make them meaningless?". *Current Issues in Tourism*, vol. 21, no.18, 11 Dec. 2016, pp. 2057-2075. *Taylor and Francis Online*, doi:[10.1080/13683500.2016.1267113](https://doi.org/10.1080/13683500.2016.1267113).

# References – Images

"About Inside Airbnb." *About Inside Airbnb*, Inside Airbnb, <http://insideairbnb.com/about.html>.

"Clipart map Canada." *Download Free Clipart*, WebStockReview, <https://webstockreview.net/picfile/get>.

"Woman Covering Ears." *Freepik*, [https://www.freepik.com/free-photo/woman-covering-ears\\_2341523.htm](https://www.freepik.com/free-photo/woman-covering-ears_2341523.htm).

Bassig, Migs. "Airbnb Reviews Example." *A Guide to How Airbnb Reviews Work*, reviewtrackers, 01 June 2018, <https://www.reviewtrackers.com/airbnb-reviews/>.

Ehrlich, Paul, M.D., and Larry Chiaramonte, M.D.. "Girl Sneezing". *Does Your Child Have Allergies or Asthma*, San Diego Family, <https://www.sandiegofamily.com/health-guide/does-your-child-have-allergies-or-asthma>.

JaJaWa. "Airbnb's logo since July 2014." *Wikimedia Commons*, Wikipedia, July 2014, [https://commons.wikimedia.org/wiki/File:Airbnb\\_Logo\\_B%C3%A9lo.svg](https://commons.wikimedia.org/wiki/File:Airbnb_Logo_B%C3%A9lo.svg).

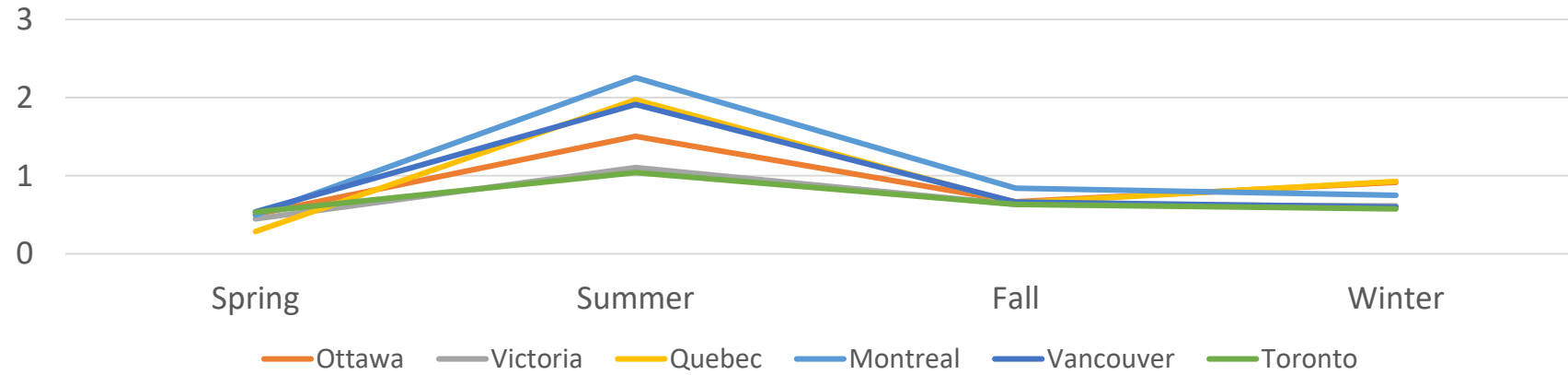
Kayne, R. "Privacy Miniblinds." *What are Privacy Miniblind*, wiseGeek, 25 June 2019, <https://www.wisegeek.com/what-are-privacy-mini-blinds.htm>.

Moler, Cleve. "The MathWorksLogo is an Eigenfunction of the Wave Equation." *MATLAB & Simulink*, MathWorks, 2003, <https://www.mathworks.com/company/newsletters/articles/the-mathworks-logo-is-an-eigenfunction-of-the-wave-equation.html>.

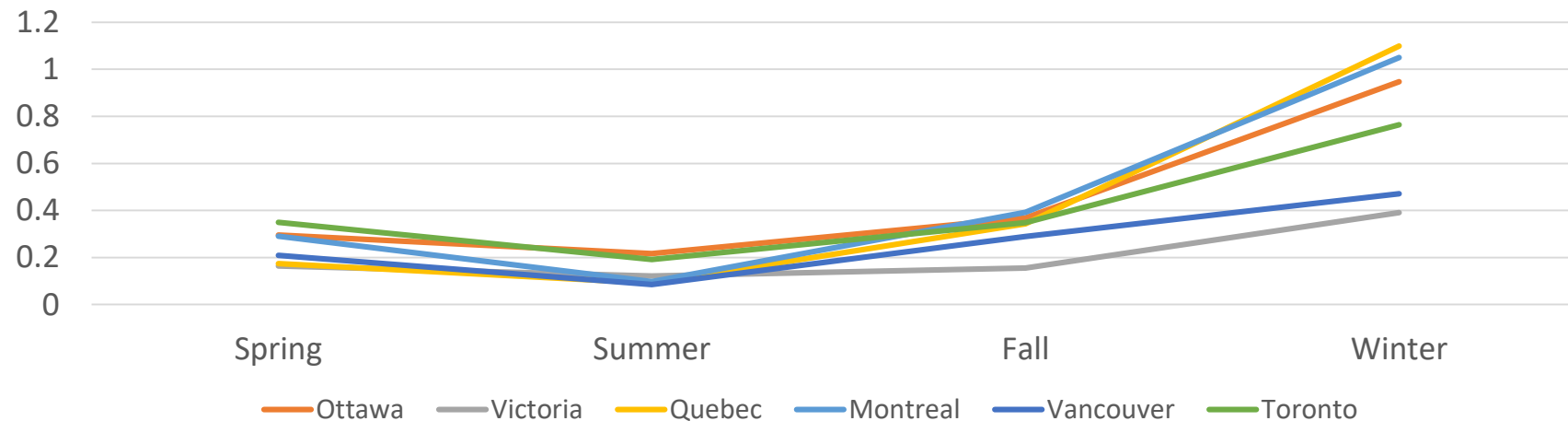
Nordqvist, Christian. "If people cannot sweat properly, they cannot cool." *MedicalNewsToday*, 18 October 2017, <https://www.medicalnewstoday.com/articles/266427.php>.

# Findings: Updated Thermal Complaints by Season?

More Accurate Percent of Comments per Season with Heat-Related Complaints?



More Accurate Percent of Comments per Season with Cold-Related Complaints?



# Findings: Acoustic Complaints?

Percent of Comments per Season with Acoustic Complaints

