# Hosts' listing descriptions and guest reviews of Airbnb in Copenhagen

Carl H. Marcussen, senior researcher

Centre for Regional and Tourism Research, Bornholm, Denmark, www.crt.dk

# Introduction

Research question:

How are Airbnb accommodations and the destination of

Copenhagen described by Airbnb hosts and their guests?

Method/approach: Text mining followed by quantitative analyses.

# Airbnb in Denmark – Copenhagen and provinces

| Airbnb | 2017 | 2018 | 2017-18 | 2018 | 2018 | 2018 | |
|---|---|---|---|---|---|---|---|
| | Arrivals | Arrivals | Change | International | Ave.nights | Nigths | |
| Copenhagen | 508.360 | 520.000 | 2,3% | 92% | 3 | 1,6 | mill. |
| Provinces | 390.020 | 517.000 | 32,6% | 58% | 3,8 | 2,0 | mill. |
| Denmark | 898.380 | 1.037.000 | 15,4% | 75% | 3,4 | 3,5 | mill. |
| | | | | | | | |
| Total registered number of overnights in Denmark, 2018 | | | | | | 53,8 | mill. |
| Overnights according to Statistics Denmark + Airbnb nights | | | | | | 57,4 | mill. |
| Airbnb's share overnights in Denmark overall | | | | | | 6,1% | |

# Airbnb in Copenhagen - 2018

| 2018 | Hotels + hostels | Airbnb nights | Total nights | Airbnb's share |
|---|---|---|---|---|
| Denmark | 3,4 | 0,1 | 3,5 | 3,5% |
| International | 6,0 | 1,4 | 7,5 | 19,2% |
| Total | 9,5 | 1,6 | 11,0 | 14,2% |
| | million | million | million | |
| Sources: For Hotel+hostels, based on Statistics Denmark, | | | | |
| Central Copenhagen + sububs. | | | | |
| For Airbnb: Airbnb/Visitdenmark+own calculations, | | | | |
| assuming 3 nights per Airbnb stay in Copenhagen. | | | | |

CRT
CENTRE for REGIONAL
& TOURISM RESEARCH

# Text mining: The data matrix – from text to numbers

Section A: Background information – per observation

Section B: A set of text variables – then frequencies

Section C: A set of numeric variables, one per top word

Analyses: Standard analytical tools – factor analysis ..

# Getting the data

- Google Chrome extension – Web scraper
- Python, flexible web scraper tool
- Airbnb web-scrapes – Thanks to Murray Cox

    (also Tom Slee till end of 2017)

# Some results – from <u>guests'</u> reviews

# reviews included: ~350.000

Happy smiley:           15952

Sour smiley:            5 (*)

Host cancelled:         10677

Note: *) This low number indicates that negative reviews are rarely published.

# Some results – from guests' reviews

All five (!) sour smileys:

1) Parking in Copenhagen
2) Host required silence
3) Problem with the shower
4) The guest only met one of the hosts, not the second
5) The guest forgot to hand back the key,
        …. and the guest was sorry about that.

Comment: Only #3 is (marginally) negative.

# Most frequently used words – in <u>guests'</u> reviews (93 of 100)

and a the was is to in apartment very place great location for stay with we of nice host at clean Copenhagen had I really perfect everything us it our good lovely you close easy from located city comfortable well were as room helpful beautiful an

time wonderful friendly home but super all house so spacious cozy (cosy) on this quiet are area central my has enjoyed recommend just her center amazing metro excellent station stayed walk she have by there that fantastic stylish communication staying neighborhood bus not welcoming space restaurants kind loved

CRT
CENTRE for REGIONAL & TOURISM RESEARCH

# Running a factor analyses (CPA)
## – 37 groups of (the 93) words – guests' reviews

1) had a wonderful time at …
2) host (or host's first name) friendly helpful welcoming
3) close to metro restaurants 4) the city center (centre)
5) walk from central station
6) clean comfortable and spacious room 7) she was us her 8) our stay
9) I stayed my 10) in Copenhagen 11) fantastic great location 12) well located 13) that it is 14) perfect for ….

15) we were 16) communication easy with 17) really enjoyed 18) good very nice
19) an amazing 20) has everything you 21) beautiful apartment (or place) 22) loved staying
23) just as 24) lovely area 25) but not …. 26) recommend this
27) quiet neighbourhood 28) there are
29) cozy stylish super 30) home 31) all 32) excellent 33) by bus 34) house 35) of space 36) have
37) so kind 37x) on (not when using "so kind").

# Running a cluster analysis
## – 37 groups of (the 93) words – guests' reviews

1) the apartment was very cozy comfortable and
       clean  - it is well located
2) had wonderful time
3) close to city center (centre) metro by bus
4) in … area of Copenhagen
5) enjoyed our stay at … place
6) great … good location 7) perfect for
8) communication … easy with
9) we were
10) really nice
11) host (or first name) she helpful friendly kind welcoming us … her …
12) I … my stayed 13) has everything you 14) lovely
15) walk from central station   16) just as
17) room 18) beautiful 19) an amazing … excellent
20) home 21) but not …
22) super 23) all 24) house 25) so 26) spacious 27) on 28) recommend
this 29) quiet neighborhood 30) there are 31) have 32) that 33) fantastic
34) stylish 35) staying 35) loved 36) space 37) restaurants

CRT
CENTRE for REGIONAL
& TOURISM RESEARCH

# Top 30 – of the 93 words used most frequently by Airbnb hosts in CPH

| | | |
|---|---|---|
| 1 | apartment | 78% |
| 2 | in | 68% |
| 3 | the | 62% |
| 4 | and_ | 55% |
| 5 | Copenhagen | 42% |
| 6 | a | 38% |
| 7 | of | 37% |
| 8 | to_ | 35% |
| 9 | with_ | 35% |
| 10 | cozy | 34% |

| | | |
|---|---|---|
| 11 | is | 27% |
| 12 | located | 24% |
| 13 | room | 21% |
| 14 | close | 21% |
| 15 | center | 17% |
| 16 | central | 16% |
| 17 | two_2 | 16% |
| 18 | city | 15% |
| 19 | heart | 14% |
| 20 | bright | 14% |

| | | |
|---|---|---|
| 21 | for | 13% |
| 22 | spacious | 12% |
| 23 | Nørrebro | 12% |
| 24 | from | 12% |
| 25 | nice | 11% |
| 26 | charming | 11% |
| 27 | lovely | 11% |
| 28 | minutes | 10% |
| 29 | area | 10% |
| 30 | balcony | 9% |

N=25100 (number of listings)

CRT
CENTRE for REGIONAL
& TOURISM RESEARCH

# Top 10 out of 37 groups of 89 words – used by Airbnb hosts in CPH

| Group | Words 1-29 | Loadings |
|---|---|---|
| 1 | minutes | 0,79 |
| | from | 0,66 |
| | ten_10 | 0,60 |
| | walk | 0,58 |
| | only | 0,43 |
| | five_5 | 0,36 |
| 2 | heart | 0,76 |
| | of | 0,75 |
| | the | 0,60 |
| | in | 0,41 |

| | | |
|---|---|---|
| 3 | to_ | 0,83 |
| | close | 0,83 |
| 4 | newly | 0,93 |
| | renovated | 0,92 |
| 5 | house | 0,73 |
| | garden | 0,69 |
| OR | apartment | -0,58 |
| 6 | center | 0,80 |
| | city | 0,73 |
| 7 | friendly | 0,87 |
| | family | 0,84 |

| | | |
|---|---|---|
| 8 | living | 0,67 |
| | room | 0,54 |
| | kitchen | 0,51 |
| | large | 0,44 |
| 9 | perfect | -0,78 |
| | for | -0,77 |
| 10 | is | 0,75 |
| | located | 0,66 |

N=25100 (number of listings)

CRT
CENTRE for REGIONAL
& TOURISM RESEARCH

# Running a factor analyses (CPA)
## – 37 groups of words – hosts' descriptions

(1) only five_5 ten_10 minutes walk from (2) in the heart of (3) close to (4) newly renovated (5) house garden OR apartment (6) city center (7) family friendly (8) living room large kitchen (9) perfect for (10) is located (11) near near metro beach

(12) welcome our home (13) with sunny balcony (14) on … floor (15) right next (16) bright and spacious (17) great location best (18) Copenhagen central station (19) quiet neighborhood (20) two_2 rooms (21) a view (22) trendy Vesterbro (23) hip area (24) it has (25) cozy small OR by (26) one_1 bedroom (27) very nice (28) beautiful Frederiksberg (29) charming OR private (30) this you OR m2 (31) my (32) new modern (33) three_3 big (34) Østerbro OR Nørrebro (35) lovely <-> penthouse (36) light (37) an OR at

# Cluster analysis approach to the hosts' descriptions of their Copenhagen Airbnb accommodations

(1) cozy apartment (2) in the heart of Copenhagen (3) bright and spacious (4) a quiet neighborhood (5) close to city center near metro beach (6) sunny balcony with view (7) my … is located … (8) living room  one_1 large bedroom kitchen (9) central station (10) two_2 rooms … three_3 (11) perfect for family friendly (12) hip Nørrebro area (13) only 5-10 minutes walk from (14) very nice (15) charming (16) lovely

(17) on … floor (18) welcome .. our home (19) beautiful (20) you (21) trendy Vesterbro (22) big (23) great location best (24) new modern (25) light (26) at (27) Frederiksberg (28) m2 (29) house private garden (30) newly renovated (31) by (32) Østerbro (33) small (34) it has (35) right next (36) an (37) penthouse

# Differences and overlap – words in reviews vs. listings

| in reviews | in listings | | |
| --- | --- | --- | --- |
| | No | Yes | Total |
| No | 0 | 43 | 43 |
| Yes | 47 | 46 | 93 |
| Total | 47 | 89 | 136 |

# Conclusions

- Traditional analytical tools can be used to analyse text, for example text related to Airbnb accommodations
- It's a matter of taste if factor OR cluster analysis is used
- Overlap and differences in words used hosts and guests
- Descriptions provided by Airbnb hosts and the reviews of stays provided by Airbnb guests can be used to develop *profiles or images of destinations/districts*
- Other applications of text mining: For example reviews on Tripadvisor

CRT

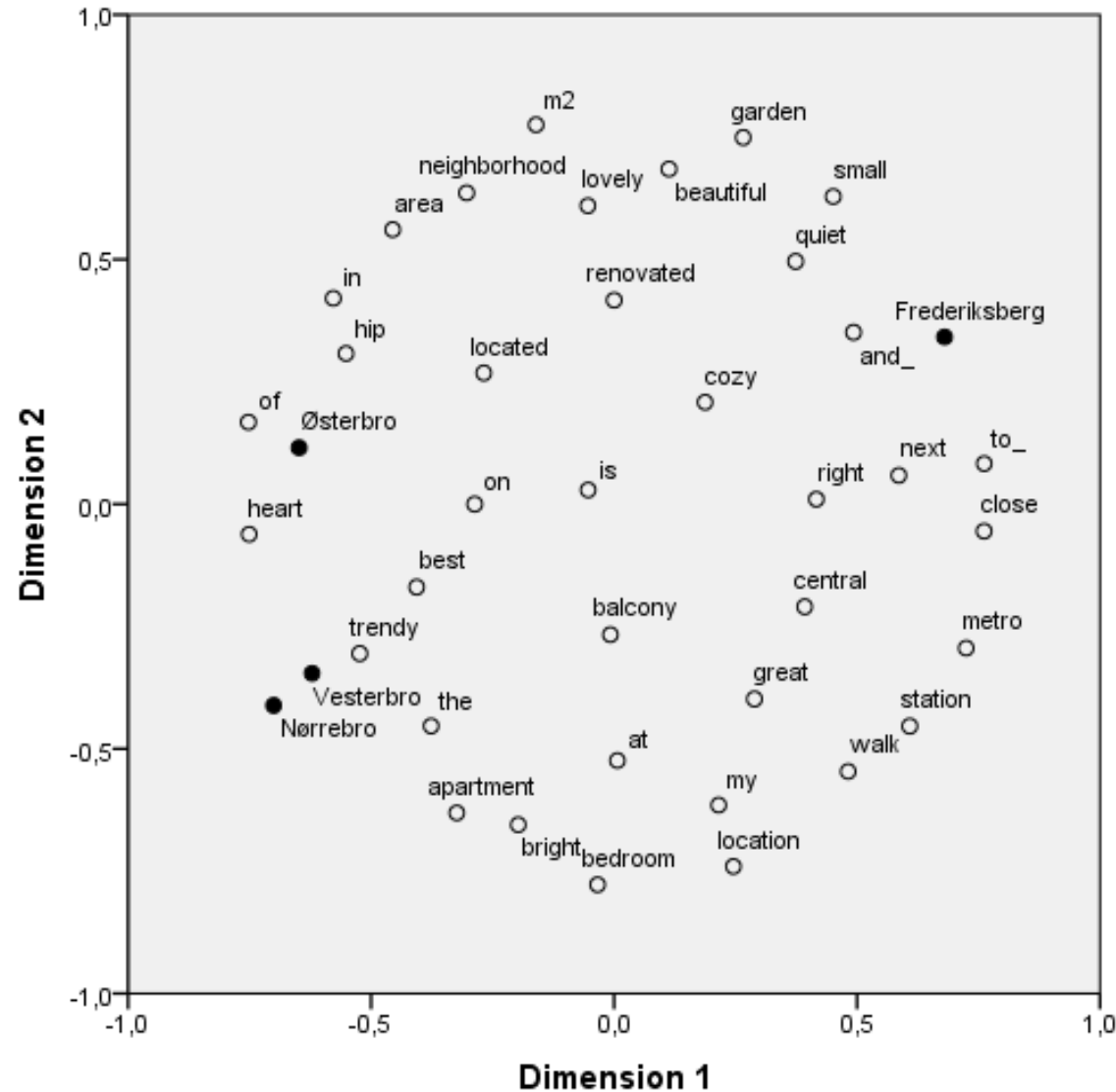CENTRE for REGIONAL & TOURISM RESEARCH

# Thank you for your attention!

Questions – Comments - Discussion

# Words signifying districts in CPH (~image) based on Airbnb listing descriptions

| #Nørrebro | #Vesterbro | Frederiks-#berg | #Østerbro |
|-----------|------------|-----------------|-----------|
| 1 hip | 1 trendy | 1 metro | 1 located |
| 2 in | 2 hip | 2 beautiful | 2 at |
| 3 trendy | 3 heart | 3 in | 3 in |
| 4 heart | 4 in | 4 located | 4 lovely |
| 5 of | 5 of | 5 garden | 5 apartment |
| 6 cozy | 6 central | 6 lovely | 6 on |
| 7 apartment | 7 the | 7 apartment | 7 neighborhood |
| 8 at | 8 located | 8 next | 8 cozy |
| 9 the | 9 station | 9 quiet | 9 bright |
| 10 located | 10 apartment | 10 at | 10 m2 |
| 11 area | 11 on | 11 to_ | 11 area |
| 12 on | 12 area | 12 on | 12 quiet |
| 13 bright | 13 at | 13 close | 13 bedroom |
| 14 my | 14 balcony | 14 central | 14 heart |
| 15 lovely | 15 neighborhood | 15 walk | 15 renovated |
| 16 is | 16 walk | 16 heart | 16 friendly |
| 17 neighborhood | 17 best | 17 of | 17 light |
| 18 great | 18 cozy | 18 and_ | 18 family |
| 19 small | 19 right | 19 location | 19 beautiful |

Note:
Based on simple correlations (r). All correlations are significant at the 0.05 level except for words in grey.

CRT
CENTRE for REGIONAL & TOURISM RESEARCH

# MDS diagram (image diagram) with 37 words related to 4 districts in Copenhagen CPH (based on Airbnb listing descriptions)



Note:
Nørrebro and Vesterbro are apparently perceived as rather similar, whereas Frederiksberg is distinctly different.