Profiling the sensorial, emotional, and ironic life of a city

[ROSSANO SCHIFANELLA]

PAN @ CLEF 2017

September 12st, 2017

Few words about me

DISTRIBUTED SYSTEMS
RECOMMENDER SYSTEMS

NETWORK SCIENCE
HUMAN COMPUTING

CROWDSOURCING

URBAN COMPUTING
MACHINE LEARNING

BEHAVIOURAL STUDIES
COMPUTATIONAL *

URBAN COMPUTING
COMPUTATIONAL SOCIAL SCIENCE

[World Urbanization Prospects: The 2014 Revision @United Nations]

HUMANITY IS URBAN

30%	54%	66%
1950	2014	2050

INFORMATICS ARE PERVASIVE

[Shakespeare, Coriolanus]

“ What is a city but people? ”

EFFICIENT.SUSTAINABLE.SMART

+

SOCIAL.HEALTHY.HAPPY


HOME
PROJECTS
DATA
TEAM
JOIN US
good CITY life
SUBSCRIBE
JOIN US
SPONSOR

How digital data can be used to

1. Study urban phenomena at scale
2. Profiling


SOCIAL
MEDIA
BLOG
SMS
BLOG
SMS
SMS
SMS
SMS
BLOG

WHERE

coordinates

WHEN

timestamp

WHAT

text+visual+audio

WHO

author

WHY

intent, context

Sensing

1

HAPPY MAPS

How visually pleasant is a city?

2

SMELLY MAPS

How does a city smell?

3

CHATTY MAPS

How does a city sound?

Happy Maps

Get the short and **pleasant** route

HyperText 2014

SHORTEST

SHORT and PLEASANT

HOW?

- Collect urban perceptions

UrbanGems: Crowdsourcing Quiet, Beauty and Happiness

Change Question

Which place do you find more beautiful?

Progress: 0/10

Picture Info

Picture Info

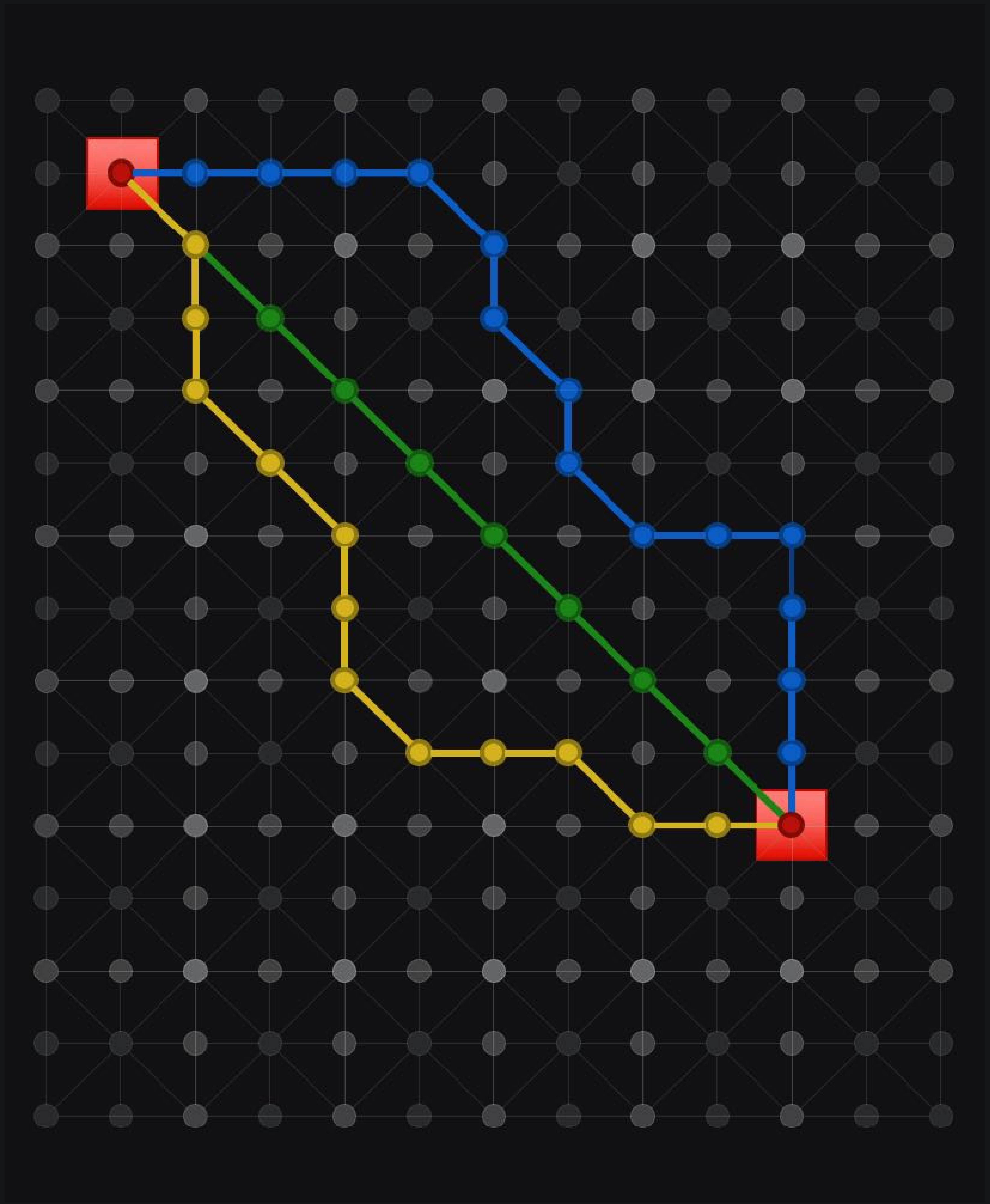
Can't Tell

HTTP://URBANGEMS.ORG/

HOW?

COLLECT URBAN PERCEPTIONS

GENERATE EMOTIONALLY-AWARE ROUTES



HOW?

- Collect urban perceptions
- Generate emotionally-aware routes
- Evaluate

- Survey in London
  - Path from Euston Square and Tate Modern
  - 3 situations (happy, quiet, beauty scenarios)
  - 4 paths to vote on a Likert scale (paths are unlabeled)

HOW?

Collect urban perceptions

Generate emotionally-aware routes

Evaluate

Model aesthetics with social media data

Data

7M geolocated photos in London

Method

- For each street segment we extract:
  - Number of **pictures** (density), number of **views**, of **favorites**, of **comments**, and **tags**
  - Tags (**LIWC** dictionary, 72 categories)
- Extract features that are **significantly correlated with beauty scores**
  - Example: density, 'posemo', 'negemo', 'swear', 'anx' (anxiety), 'sad', and 'anger' LIWC categories
- Build a **model** to predict beauty

Longfellow Bridge
Cambridge St
Land Blvd
Embankment Rd
Phillips St
Revere St
Myrtle St
Pinckney St
Mt Vernon St
Chestnut St
Beacon St
Charles St
W Cedar St
Garden St
Irving St
S Russell St
Joy St
Hancock St
Ridgeway Ln
Temple St
Bowdoin St
Somerset St
Cambridge St
Boston
Sudbury St
Congress St
North St
State St
Court Sq
Devonshire St
India St
Broad St
Milk St
Franklin St
Pearl St
Oliver St
Federal St
Congress St
Arch St
High St
Atlantic Ave
John F Fitzgerald Surface Rd
John F Fitzgerald Expy
Seaport Blvd
Charles River
Charles River Esplanade
Esplanade
Dr Paul Dudley White Path
Storrow Dr
Back St
Beacon St
Public Alley 421
Marlborough St
Public Alley 422
Commonwealth Ave
Public Alley 437
Public Alley 438
Public Alley 426
Public Alley 433
Public Alley 441
Boylston St
Newbury St
Exeter St
Dartmouth St
Clarendon St
Berkeley St
Arlington St
St James Ave
Stuart St
Columbus Ave
Boston Public Garden
The Lagoon
Boston Common
Tremont St
Mason St
Downtown Crossing
Washington St
Chauncey St
Church Green
Purchase St
Essex St
Dock Square
Summer St
Chinatown
Piedmont St
Winchester St
Melrose St
Church St
Charles St S
Harrison Ave
Tyler St
Harvard St
Beach St
South St
Kneeland St
Dorchester Ave
Boston Harborwalk
Melcher St
Cortes St

Smelly Maps

How does a city smell?

ICWSM 2015

Science 2014

Humans discriminates millions of odors

Yet, city planning can discriminate only a few bad odors

Why this negative perspective?


60
2
4
7
15
30

Smell Walks

Amsterdam, Pamplona, Glasgow, Edinburgh, Newport, Paris, New York, Singapore

DATA

London + Barcelona

17M

PHOTOS

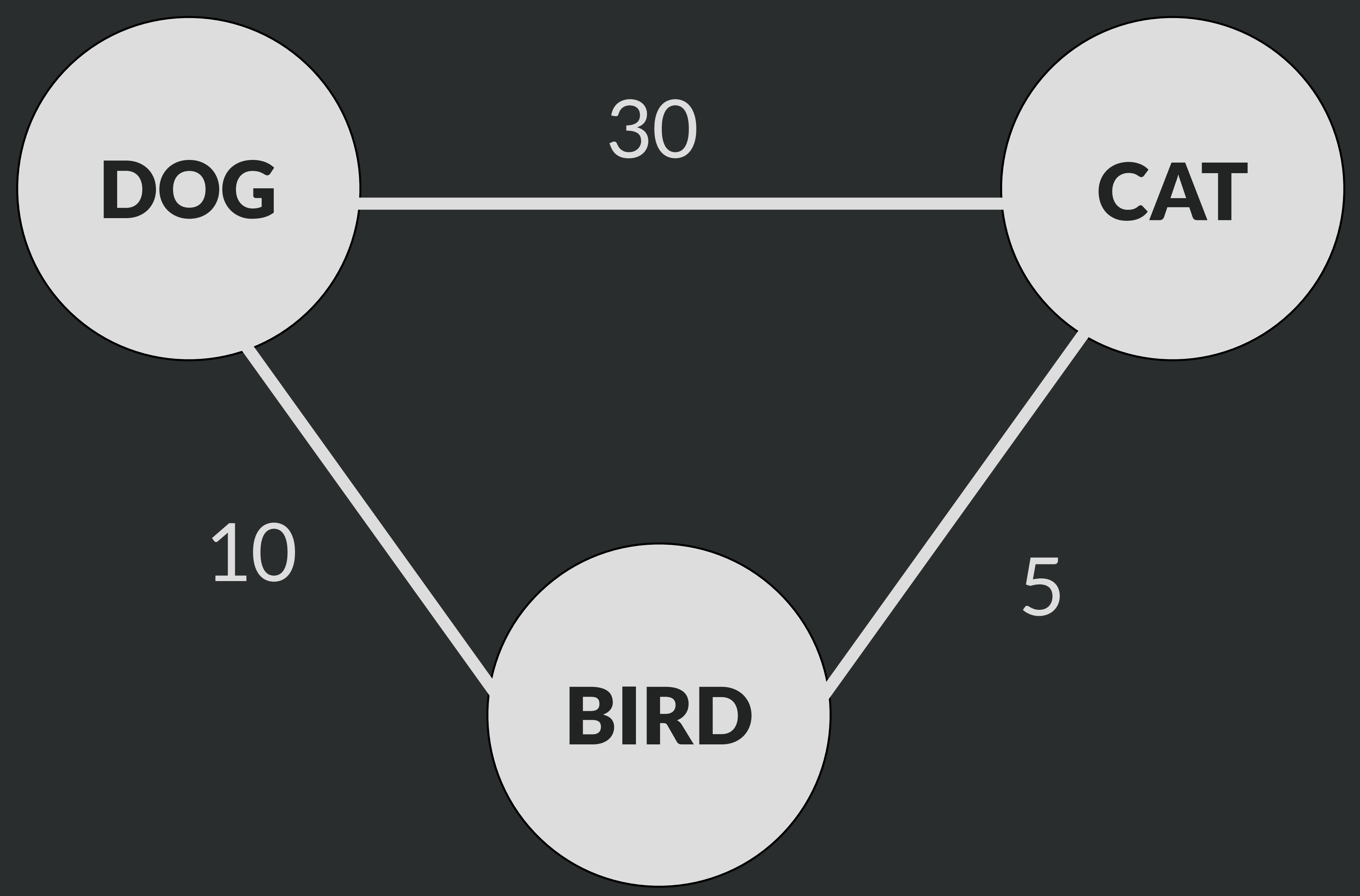
436K

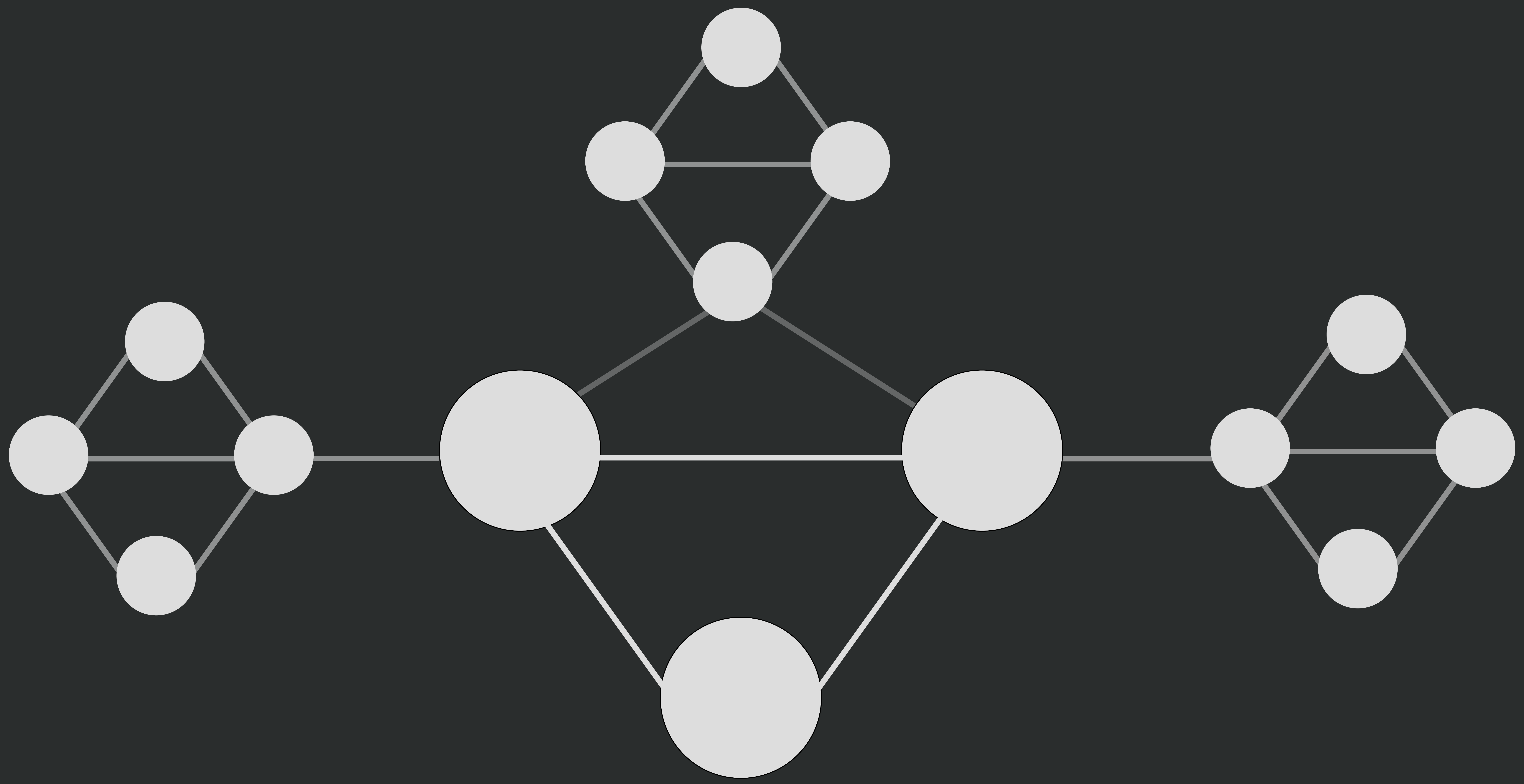
PHOTOS

1.7M

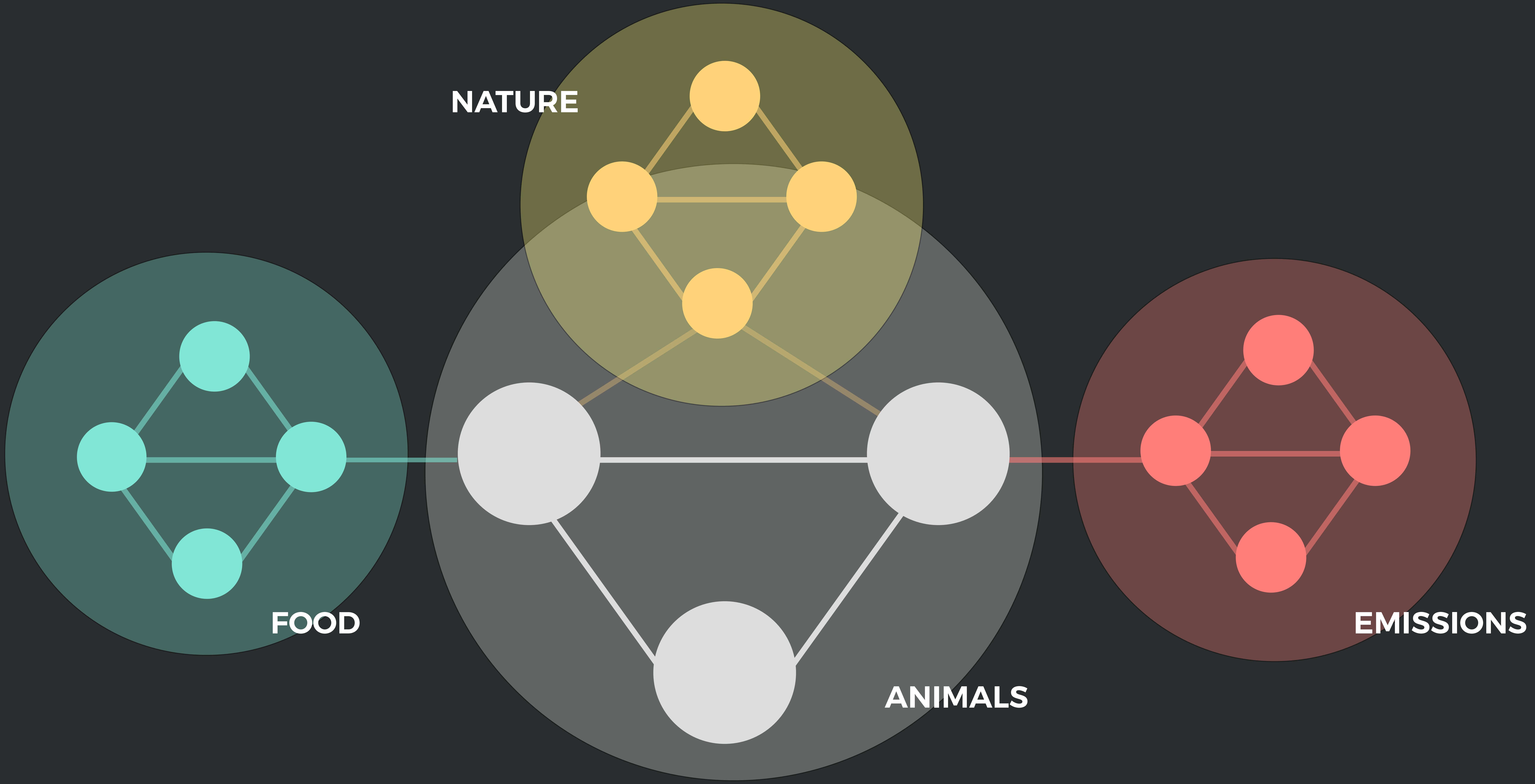
TWEETS

CO-OCCURENCE NETWORK

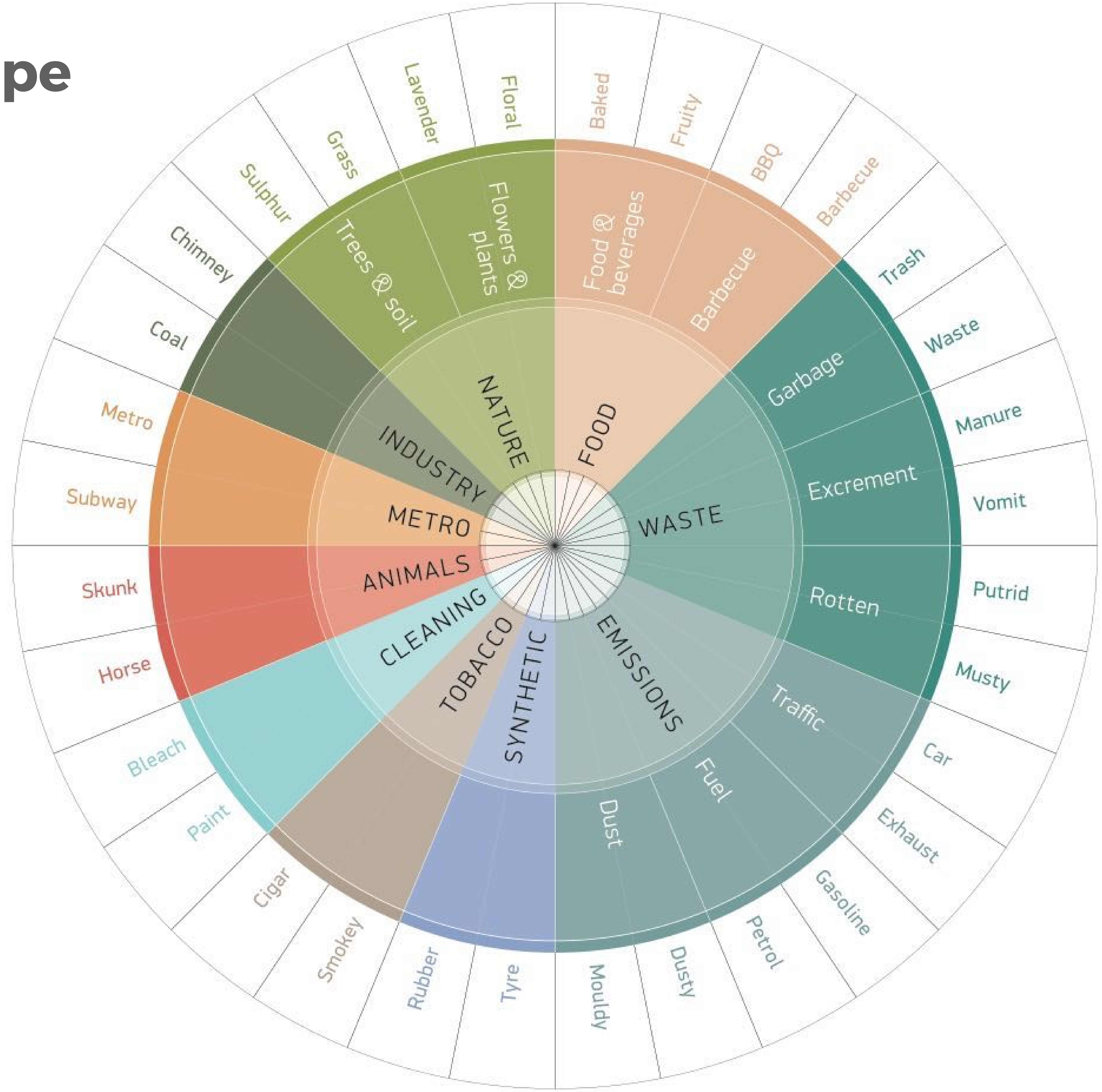




EMERGENCE OF CLUSTERS



Urban Smellscape
Aroma Wheel




NATURE
EMISSIONS
Leaflet | Mapbox Terms & Feedback, CartoDB attribution

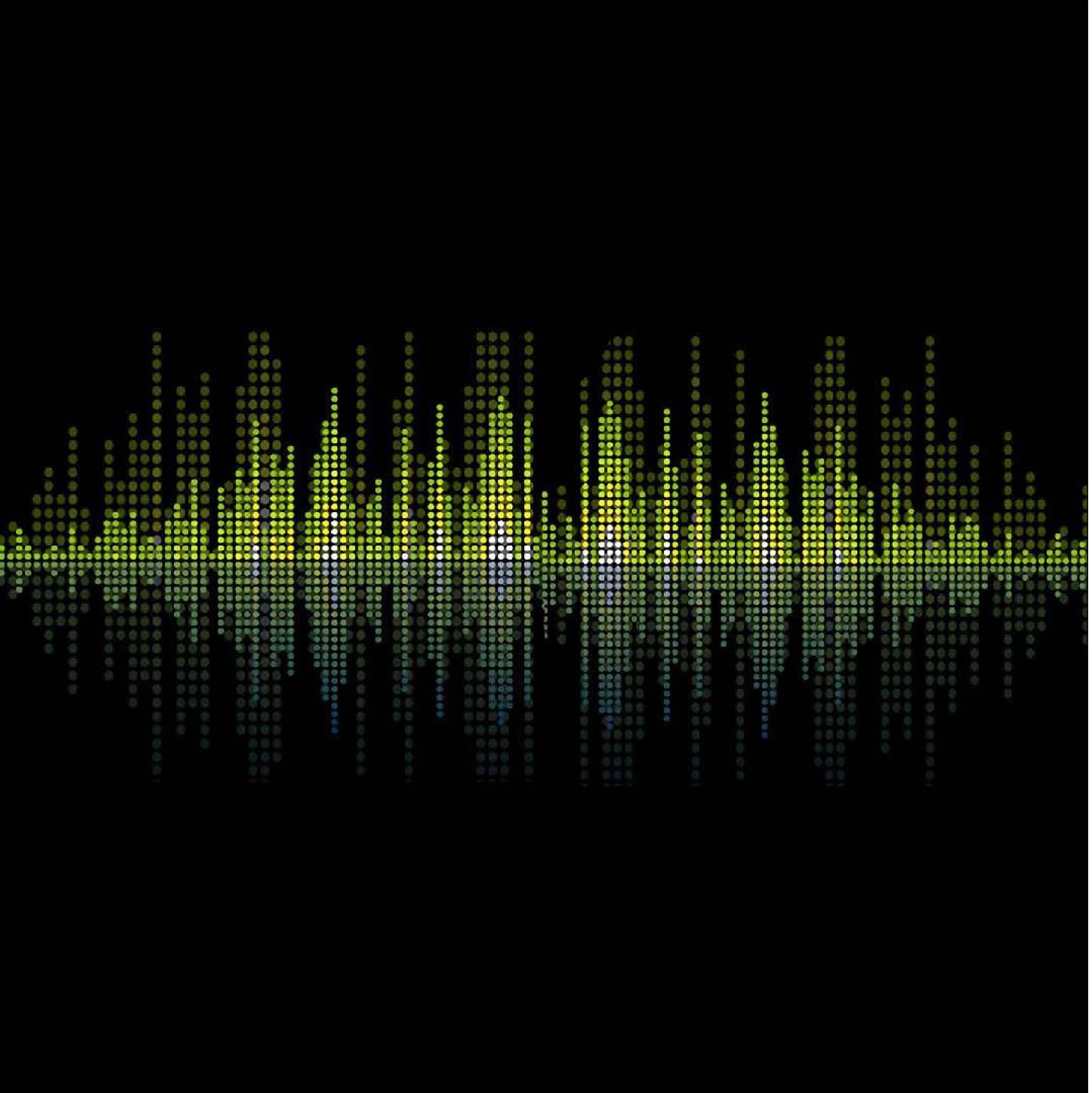
Chatty Maps

EXPLORER PROJECT GOODCITYLIFE

DEMO

GOODCITYLIFE.ORG

How does the urban smellscape change through time and space?

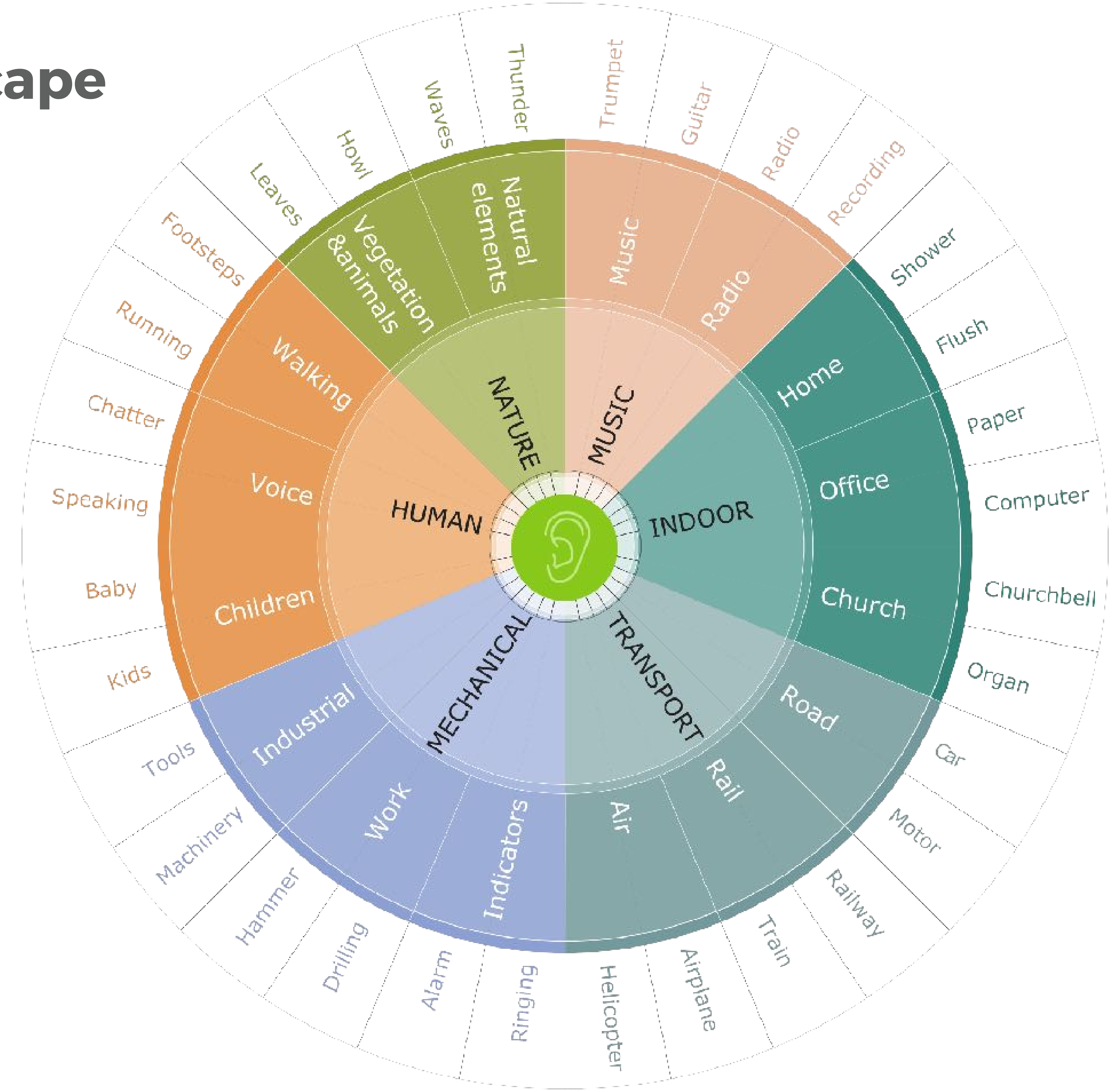


Chatty Maps

How does a city sound?

RSOS 2016

Urban Soundscape
Wheel



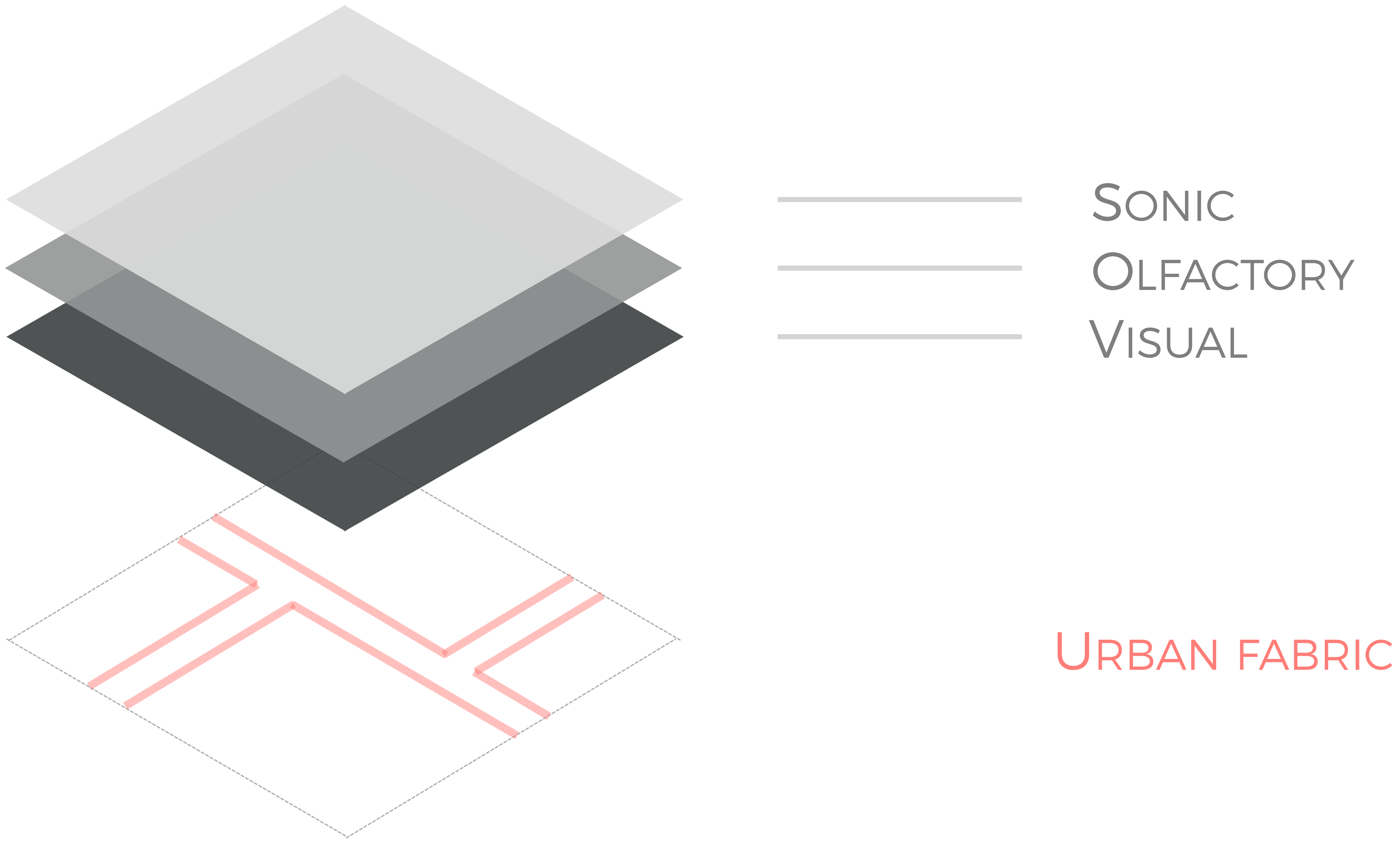
VALIDATION

CITY OFFICIALS

Air quality indicators

OPEN STREET MAPS

Presence of nature, food, etc. tags


Sonic
Olfactory
Visual
Urban fabric

Socio-economic Indicators

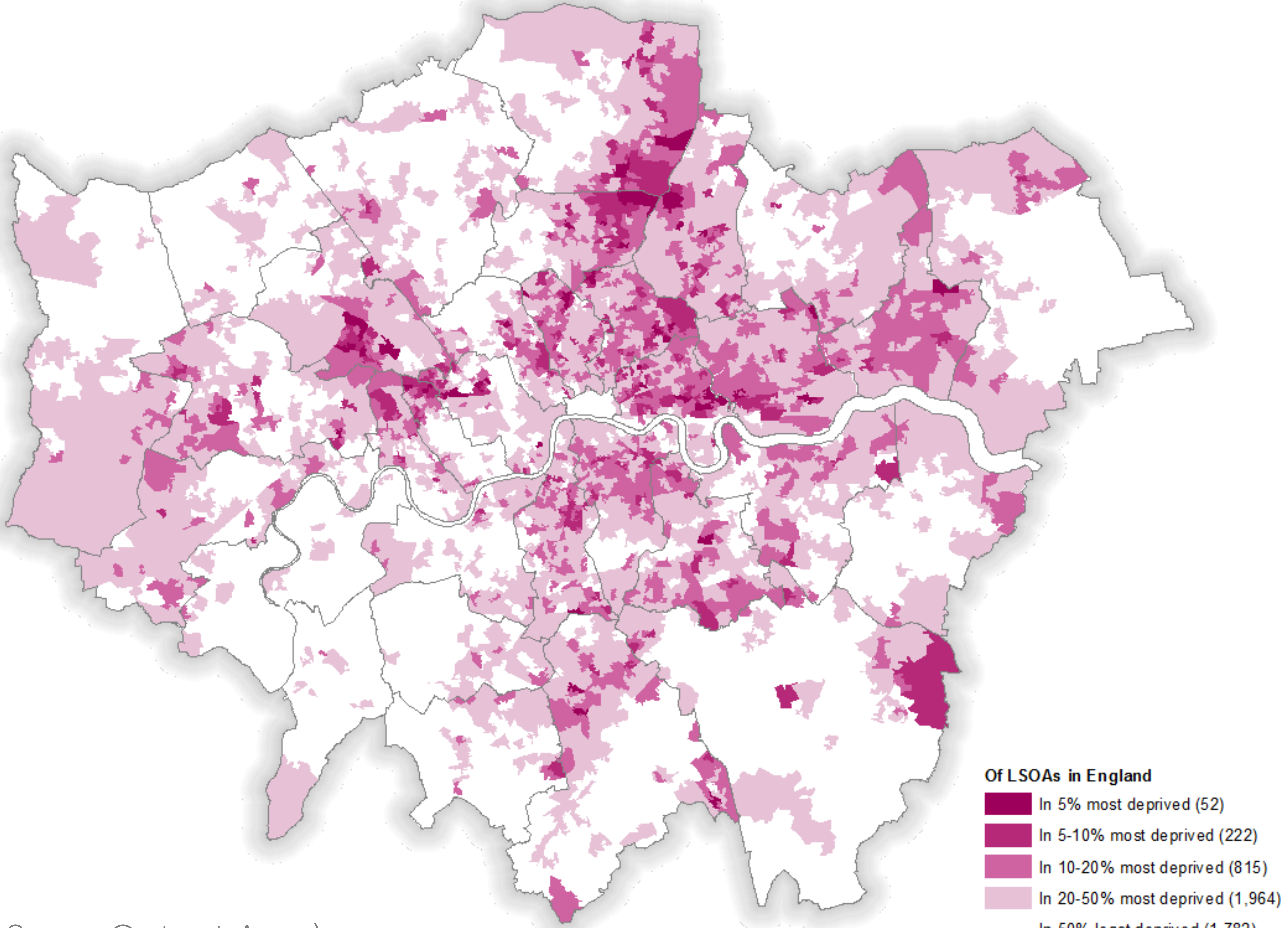
London

New York

London

IMD: Index of Multiple Deprivation

- **Income** deprivation
- **Employment** deprivation
- **Health** deprivation and disability
- **Education**, skills and training deprivation
- Barriers to **housing** and services
- **Living environment** deprivation
- **Crime**



LSOA (Lower Layer Super Output Area)

Smell (London)

IMD	nature	0.24***
	animals	0.16***
	emissions	-0.16***
	waste	-0.26***
	food	-0.1***
	cleaning	-0.19***
	industry	-0.2***
	smoke	-0.15***

Smell (London)

Living Environment	nature	0.29***
	animals	0.17***
	emissions	-0.23***
	waste	-0.35***
	food	-0.4***
	cleaning	-0.35***
	industry	-0.24***
	smoke	-0.3***
	synthetic	-0.15***

Smell (London)

LIVING ENVIRONMENT	animals	0.12***
	nature	0.2***
INCOME		
	emissions	-0.1***
	waste	-0.18***
	cleaning	-0.12***
	industry	-0.15***

Smell (London)

Living Environment	animals	0.12***
	nature	0.21***
Income		
	waste	-0.23***
Health	food	-0.14***
	cleaning	-0.17***
	industry	-0.18***
	smoke	-0.12***

Smell (London)

LIVING ENVIRONMENT	animals	0.1***
INCOME		
	waste	-0.19***
HEALTH	cleaning	-0.14***
CRIME		

Smell (London)

LIVING ENVIRONMENT

INCOME

HEALTH

CRIME

HOUSING

animals	0.14***
nature	0.17***
emissions	-0.15***
waste	-0.19***
industry	-0.16***
smoke	-0.12***

Sound (London)

IMD	human	0.11***
	nature	0.11***
	mechanical	-0.14***
	motorised	-0.17***
	music	-0.17***

Sound (London)

Living Environment	nature	0.12***
	mechanical	-0.27***
	motorized	-0.22***
	music	-0.36***
	indoor	-0.31***

Crime 2008-2016 (London)

nature	-0.38
animals	-0.24
emissions	0.43
waste	0.35
metro	0.35
cleaning	0.32
industry	0.3
smoke	0.27
food	0.19
synthetic	0.19

Crime 2008-2016 (London)

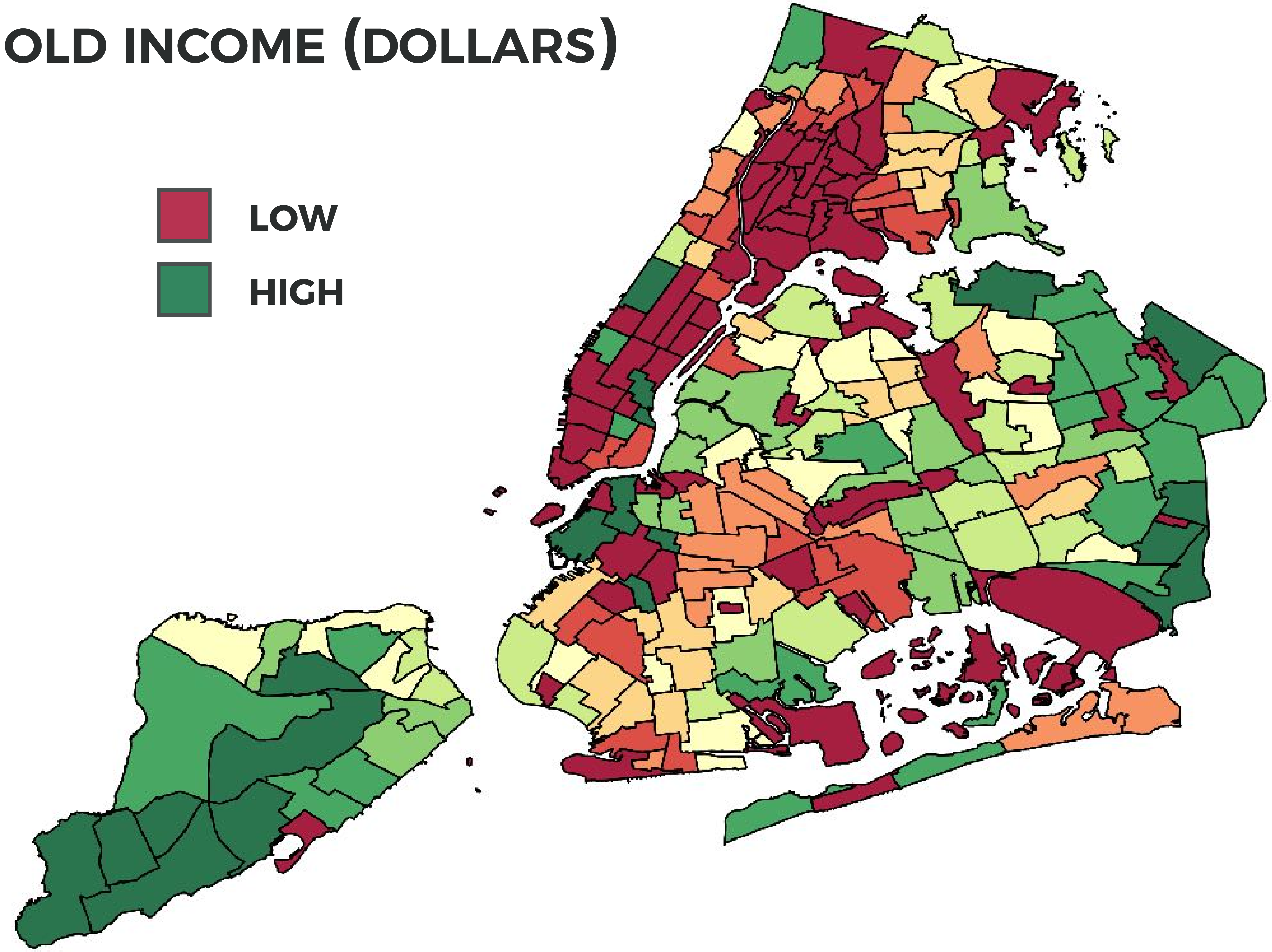
nature	-0.21
humans	-0.16
motorised	0.36
mechanical	0.17
music	0.3

New York (NTA level)

Census Bureau ACS Economic Profile

- **Employment** status
- **Commuting to work**
- **Occupation**
- **Industry**
- **Class of worker**
- **Health Insurance** coverage
- **Poverty** level

MEDIAN HOUSEHOLD INCOME (DOLLARS)



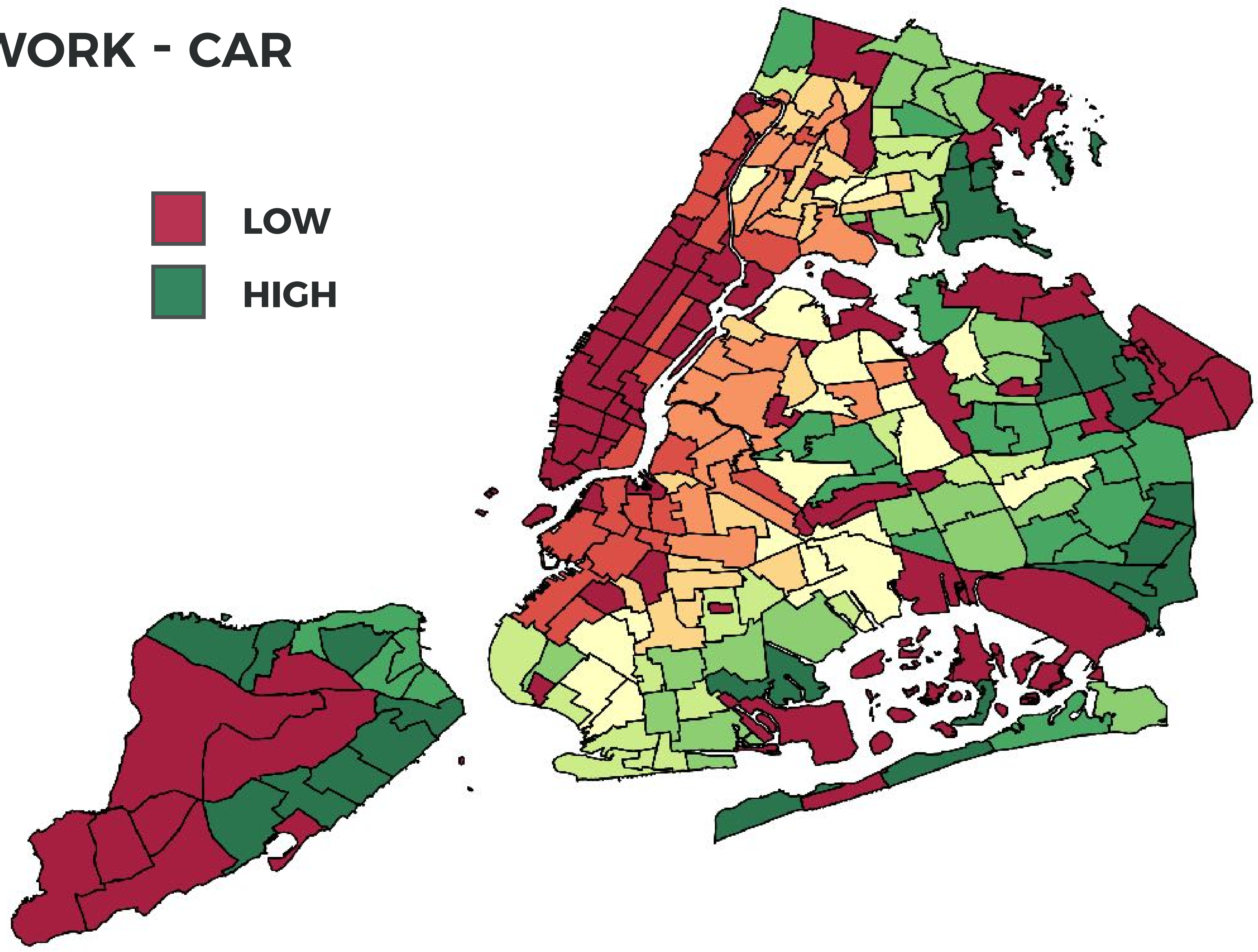
Smell (NYC)

Median household income	nature	0.38***
	food	0.16***
	metro	-0.43***
Median non-family income	+tobacco	0.22**

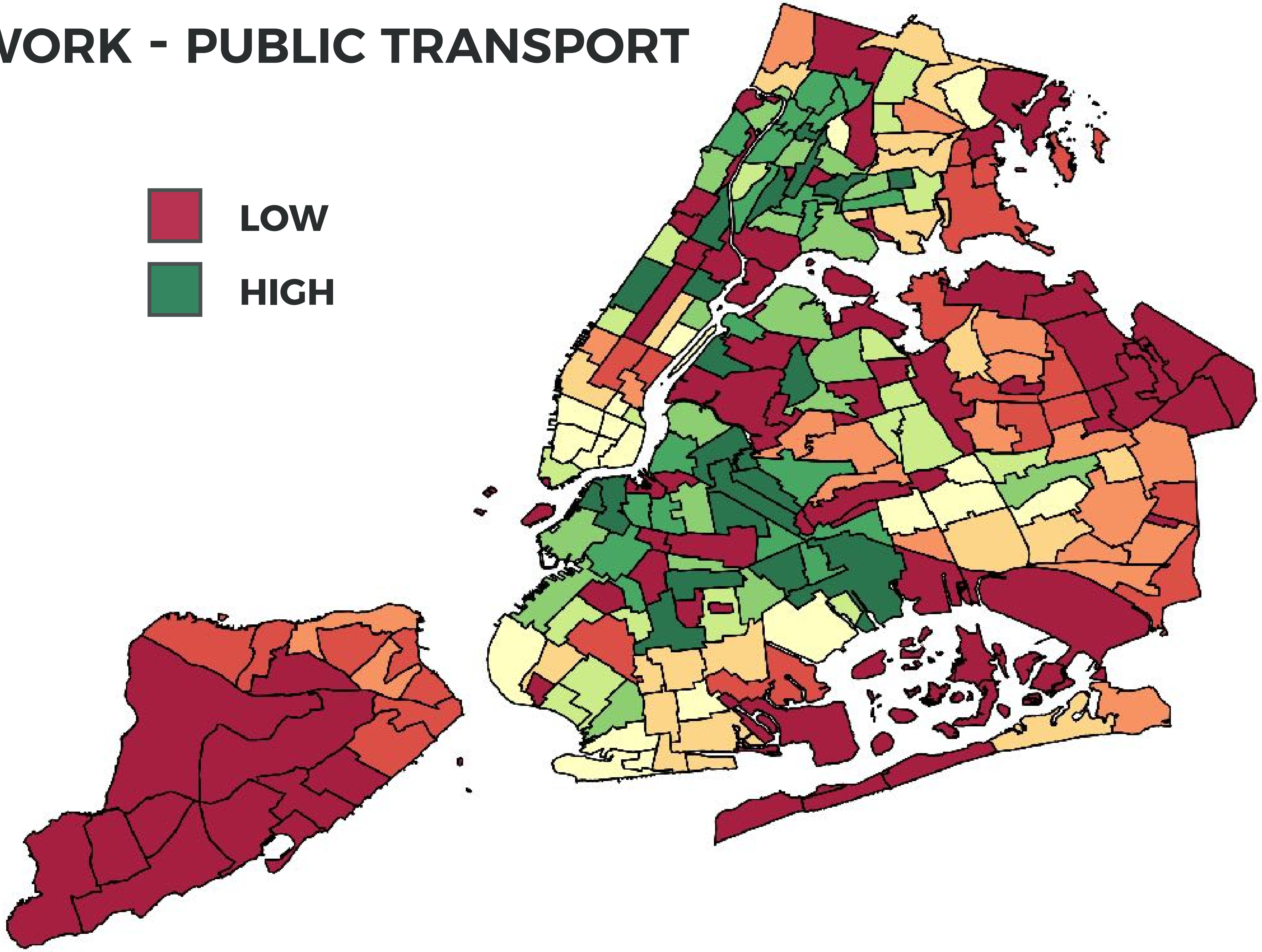
Income per household (NYC)

	emissions	nature	waste, industry, synthetic, smoke	metro
<10K	0.16	-0.39		0.45
10K-15K	0.15	-0.36		0.44
15K-25K		-0.38	0.15 (waste)	0.44
25K-35K		-0.29		0.31
35K-50K		-0.2	-0.25 (smoke)	0.26
50K-75K	-0.2	0.21	-0.27 (industry), -0.25 (smoke)	-0.15
75K-100K	-0.21	0.41		-0.52
100K-150K		0.41		-0.54
150K-200K		0.36	0.15 (smoke)	-0.45
>200K		0.3	0.25 (smoke)	-0.35

COMMUTING TO WORK - CAR



COMMUTING TO WORK - PUBLIC TRANSPORT



Commuting (NYC)

Car	nature (0.3) waste (-0.15) cleaning (-0.25) emissions (-0.2) food (-0.2) metro (-0.49) synthetic (-0.24) smoke (-0.38)
Public transportation	nature (-0.32) waste (0.18) cleaning (0.24) industry (0.17) metro (0.54) smoke (0.22)
Walked	nature (-0.25) food (0.15) industry (0.27) metro (0.41) synthetic (0.32) smoke (0.28)
Worked at home	cleaning (0.16) emissions (0.17) industry (0.2) metro (0.15) synthetic (0.32) smoke (0.36)

Sound (NYC)

MEDIAN HOUSEHOLD INCOME	nature	0.3***
	motorized	-0.43***
MEDIAN NON-FAMILY INCOME	+music	0.17**

Income per household (NYC)

	human	nature	motorized
<10K	0.3	-0.29	0.24
10K-15K	0.29	-0.27	0.29
15K-25K	0.29	-0.3	0.34
25K-35K	0.27	-0.26	0.29
35K-50K	0.3	-0.19	0.29
50K-75K	0.28	-0.22	-
75K-100K	-0.33	0.36	-0.28
100K-150K	-0.18	0.37	-0.36
>200K	0.17	0.24	-0.37

Commuting (NYC)

Car	nature (0.37) human (-0.45) mechanical (-0.17) music (-0.5)
Public transportation	nature (-0.36) human (0.28) mechanical (0.23) music (0.28)
Walked	nature (-0.3) human (0.39) music (0.41)
Worked at home	nature (-0.25) motorised (-0.17) human (0.28) indoor (0.29) music (0.35)

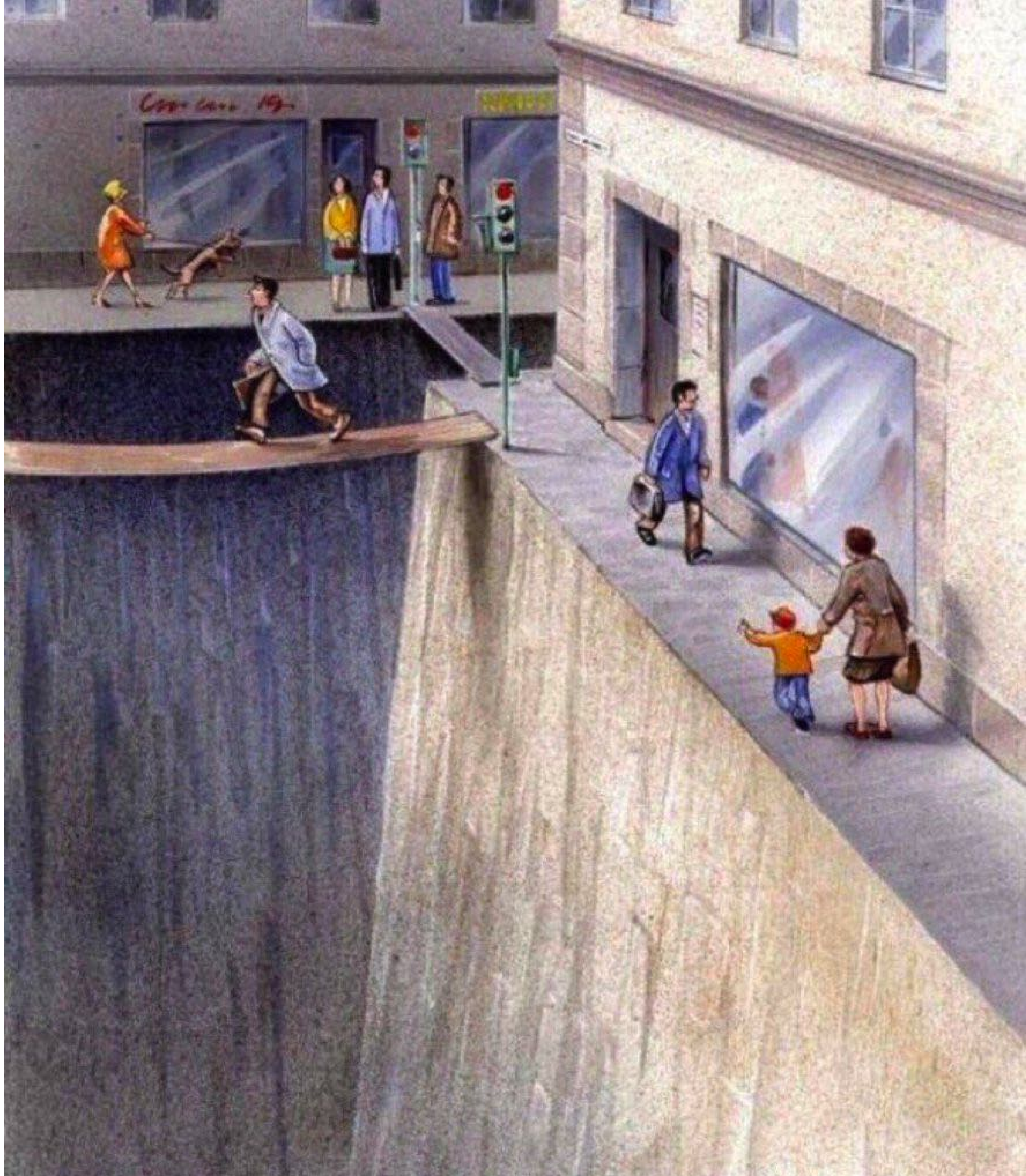
Walkability+Activities+Ambiance

Is walkability quantifiable?

Public space surrendered to cars

"The General Theory of Walkability explains how, to be favored, a walk has to satisfy four main conditions: it must be **useful**, **safe**, **comfortable**, and **interesting**. Each of these qualities is essential and none alone is sufficient."

WALKABLE CITY
Jeff Speck



Questions (safety)

- **Can safe streets be identified by night activity?**
  - Safe streets are photographed not only during the day but also at night, while unsafe ones mostly during the day
- **Can safe streets be identified by activity segmented by gender or age?**
  - Safe streets are predominantly visited by a **male** population ($r = 0.58$)
  - Safe streets are predominantly visited by an **adult** population ($r = 0.32$)

Questions (walkability)

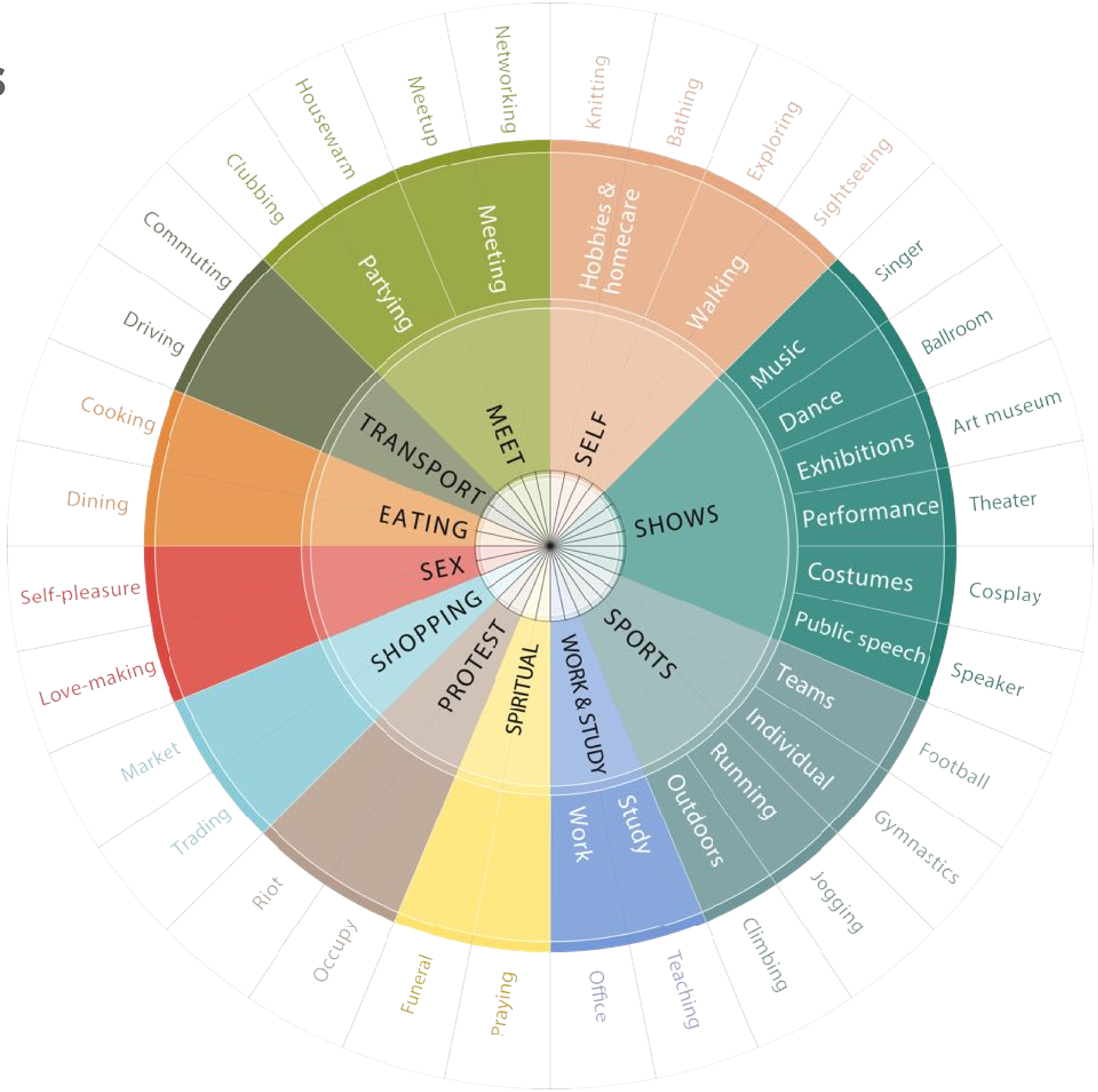
- **Can walkable streets be identified by the presence of specific types of places?**
- **Can walkability be predicted?**
  - **yes!**

Activities

Profiling urban activities

- **Identifying activity words**
  - From Flickr
  - From web documents
  - Expansion of activity words
- Focus on **private activities**
  - **indoor** vision tag
- **Clustering** of activity words in a hierarchical taxonomy

Urban Activities
Wheel

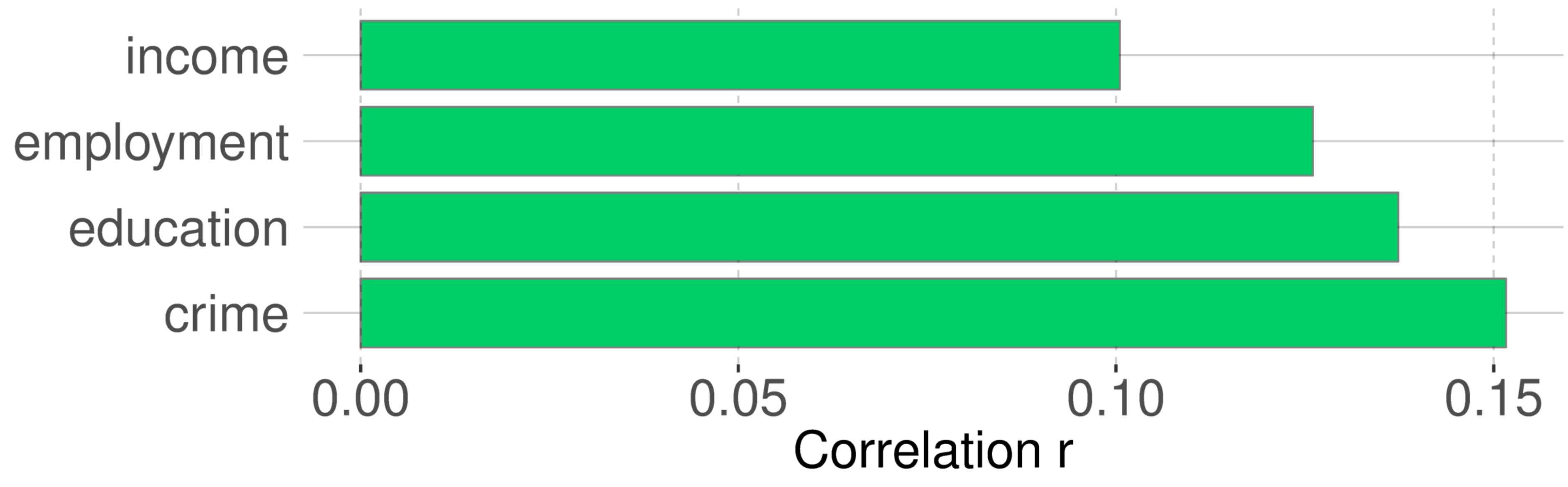


Results (some)

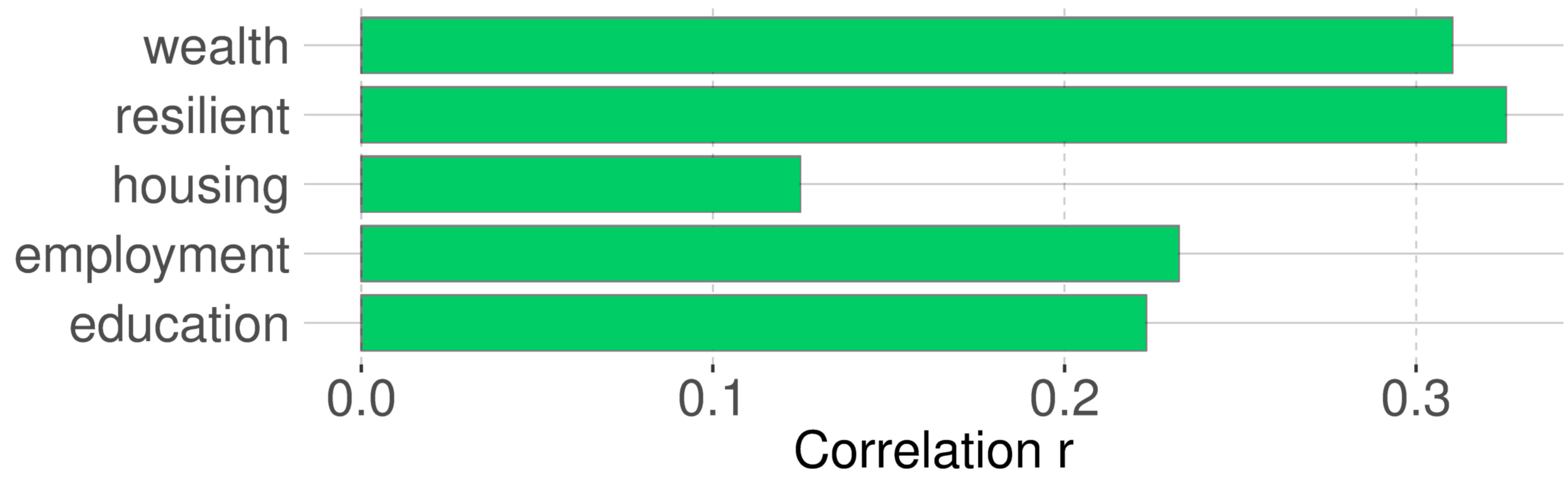
+work&study	+education(NYC) +housing (L)
+protest	-education (L), -income (L/NYC)
+self	-income
+show	-crime (L), -education, -employment, -income (NYC)

Diversity

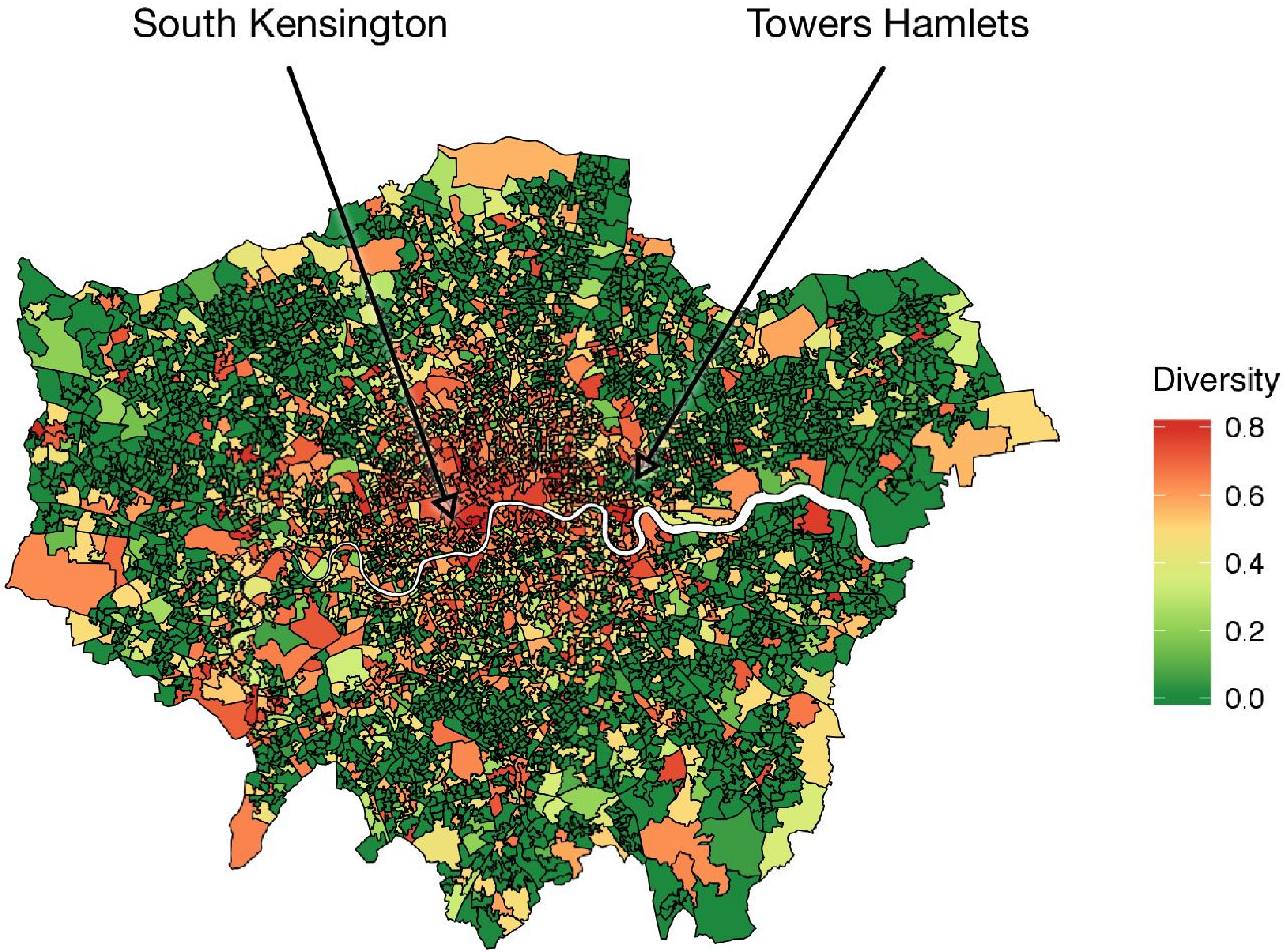
London

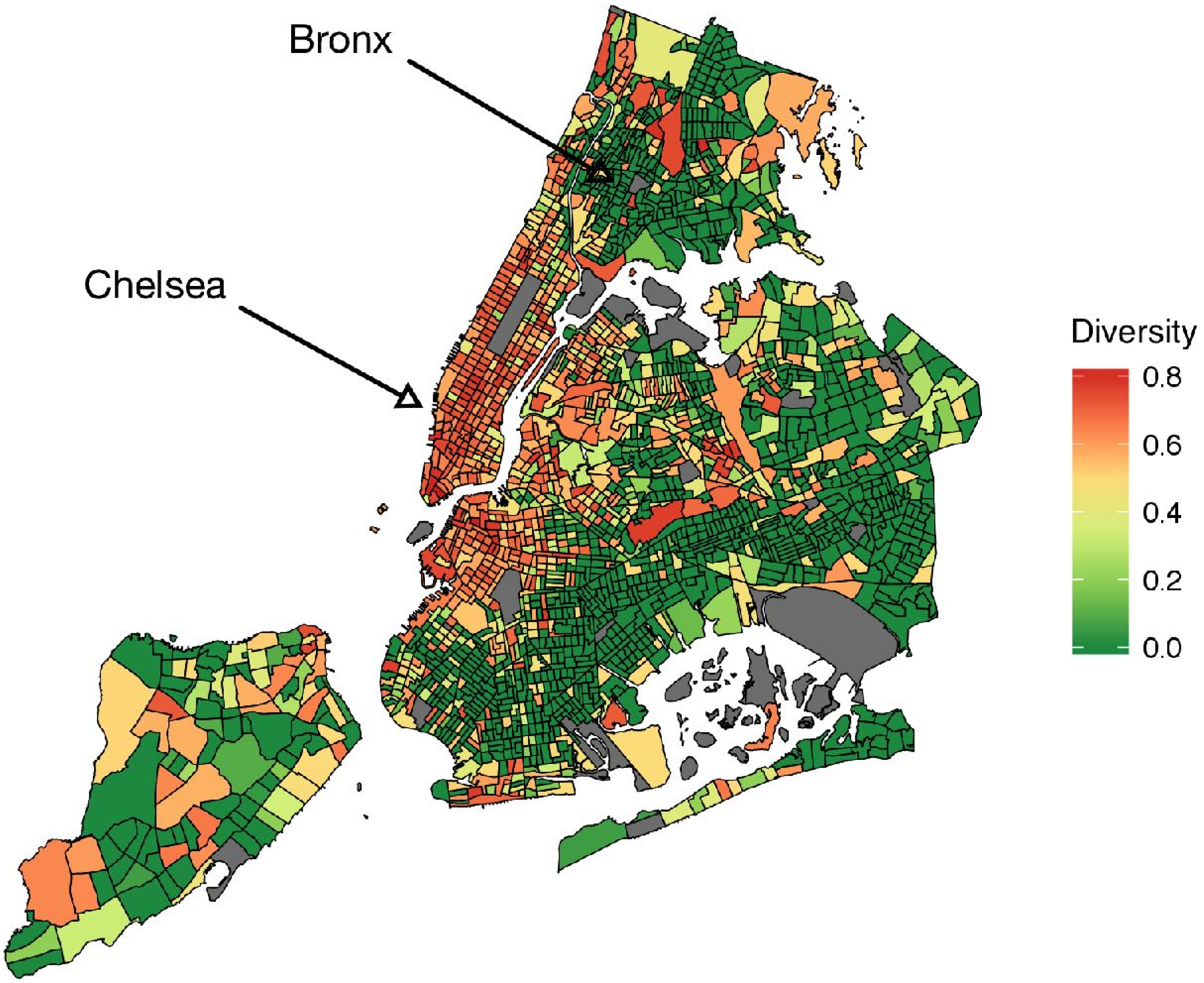


New York



economic development is associated with activity diversity


South Kensington
Towers Hamlets
Diversity
0.8
0.6
0.4
0.2
0.0



Bronx
Chelsea
Diversity
0.8
0.6
0.4
0.2
0.0

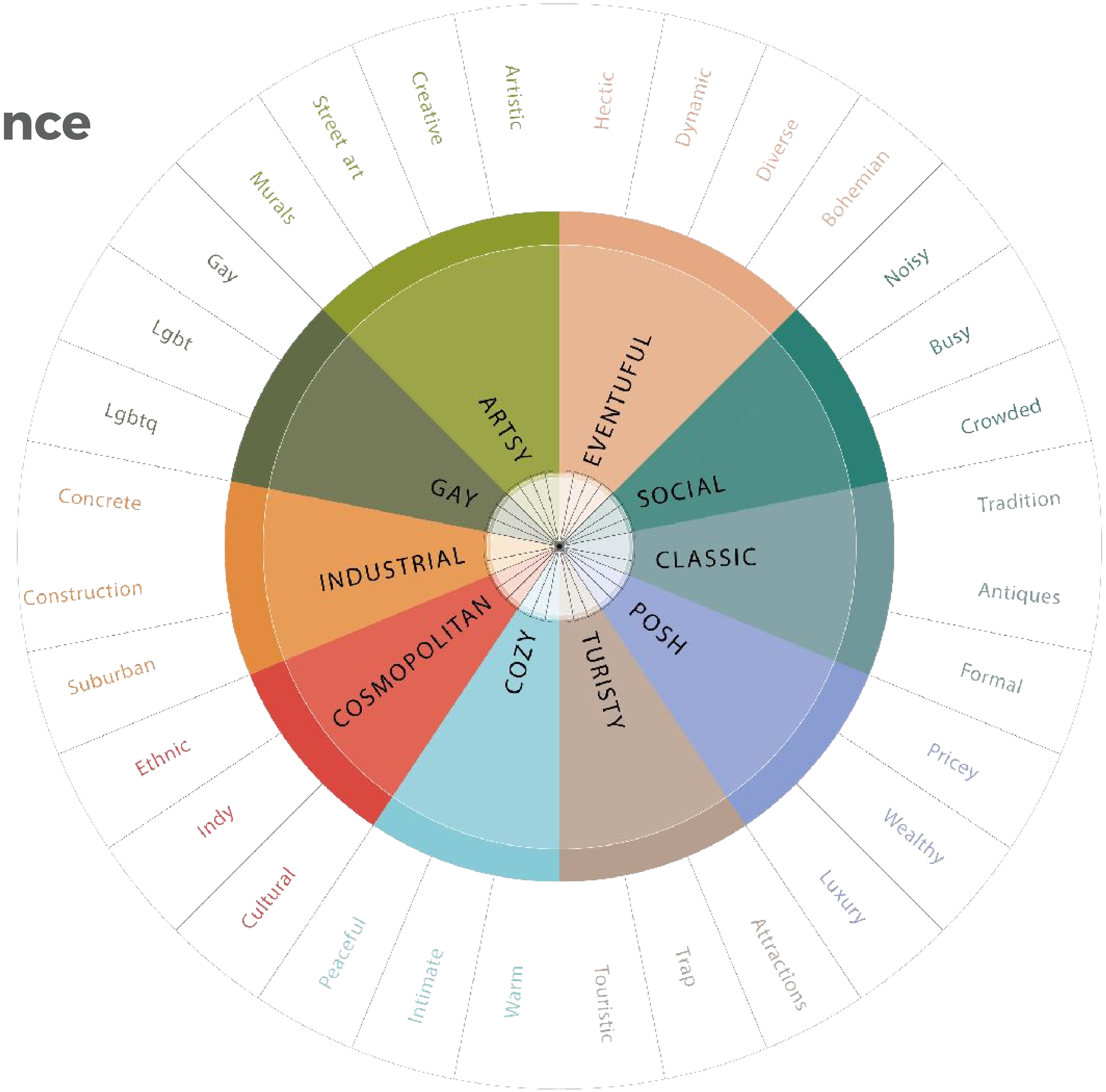
Limitations

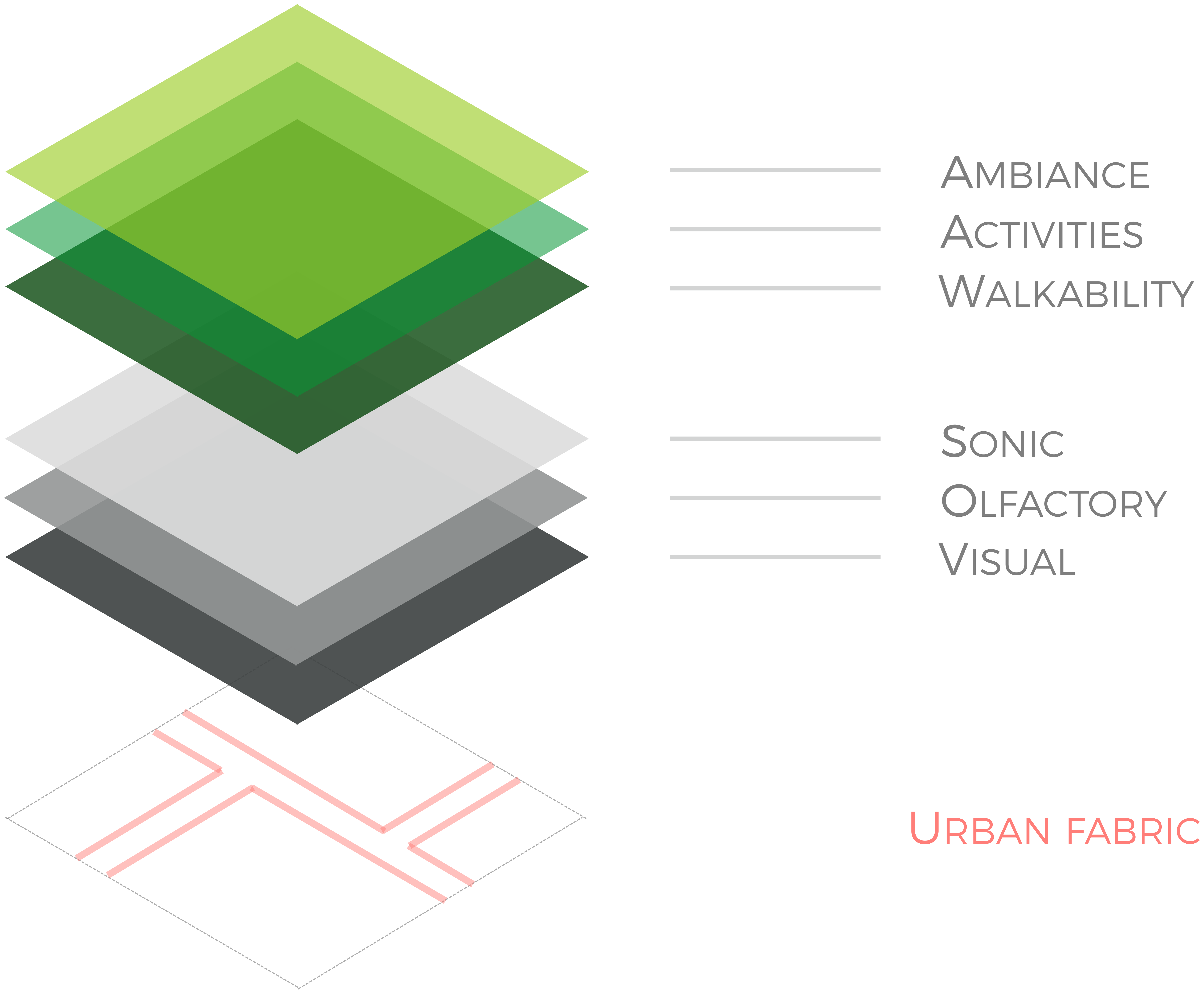
- **Not exhaustive** list of activities
- **Population-demographic** bias
- **Self-selection** bias
  - well-to-do areas might be over-represented
- Results do **not** speak to **causality**

AMBIENCE

Can the ambience of a place be predicted from pictures?

Urban Ambiance
Wheel

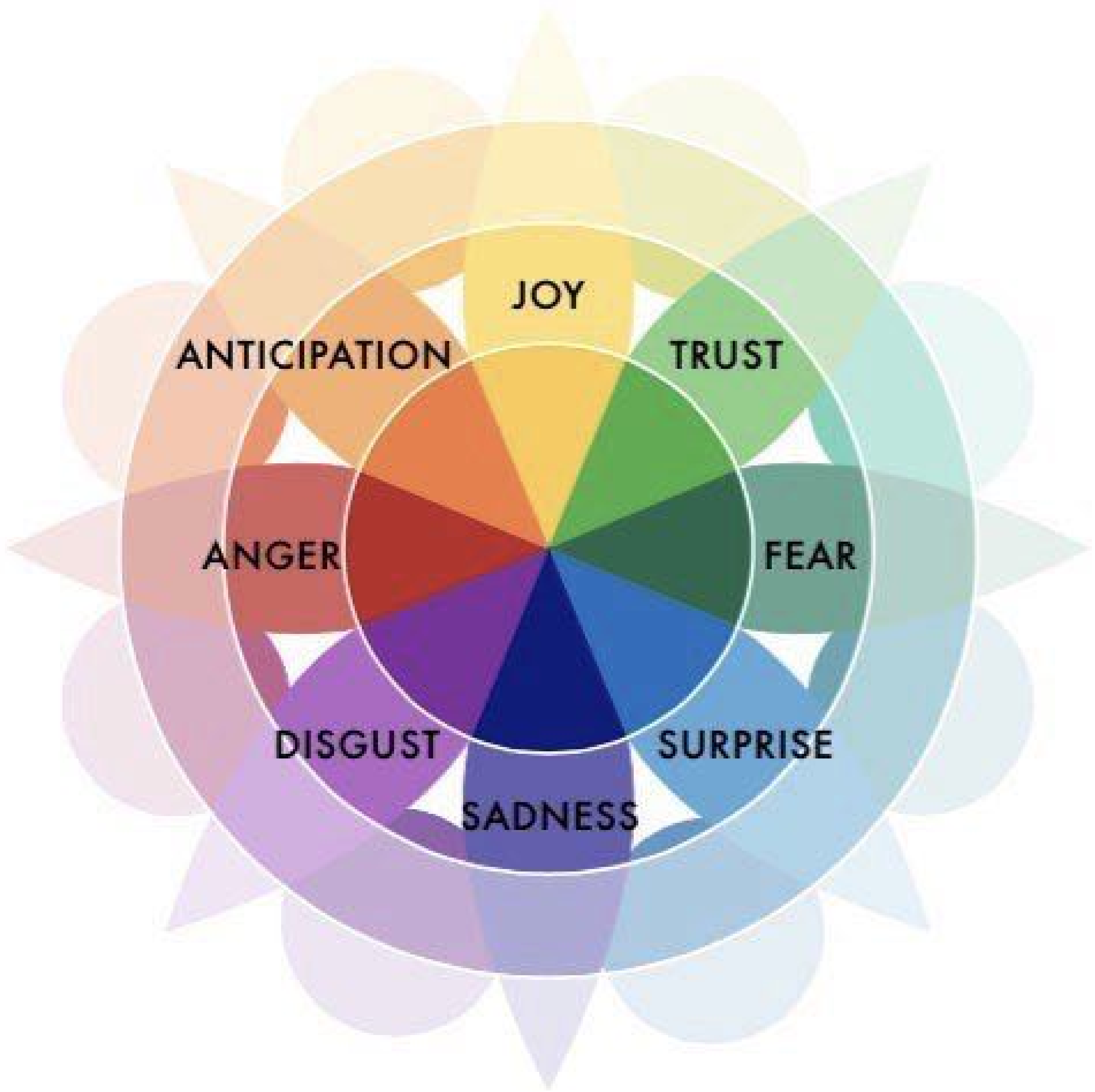



AMBIANCE
ACTIVITIES
WALKABILITY
SONIC
OLFACTORY
VISUAL
URBAN FABRIC

Emotions

EMOTIONS

To model sentiment we adopt the EmoLex lexicon that follows the 8 primary emotions from Plutchik's psychoevolutionary theory.



EMISSIONS

WASTE

FOOD

NATURE

CORRELATION BETWEEN EMOTIONS AND SMELLS

TRAFFIC

MUSIC

HUMAN

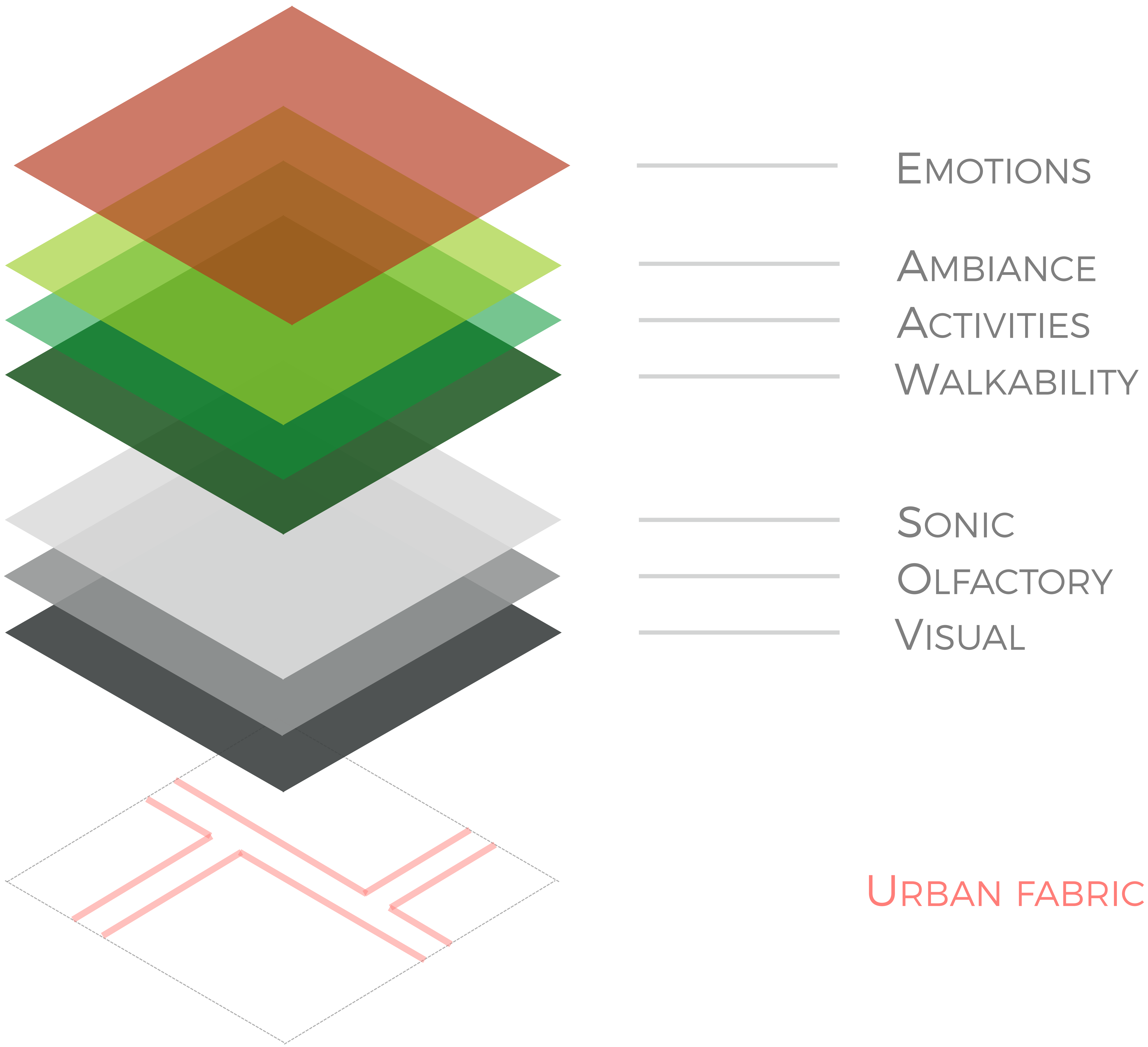
Correlation between emotions and sounds

MUSIC

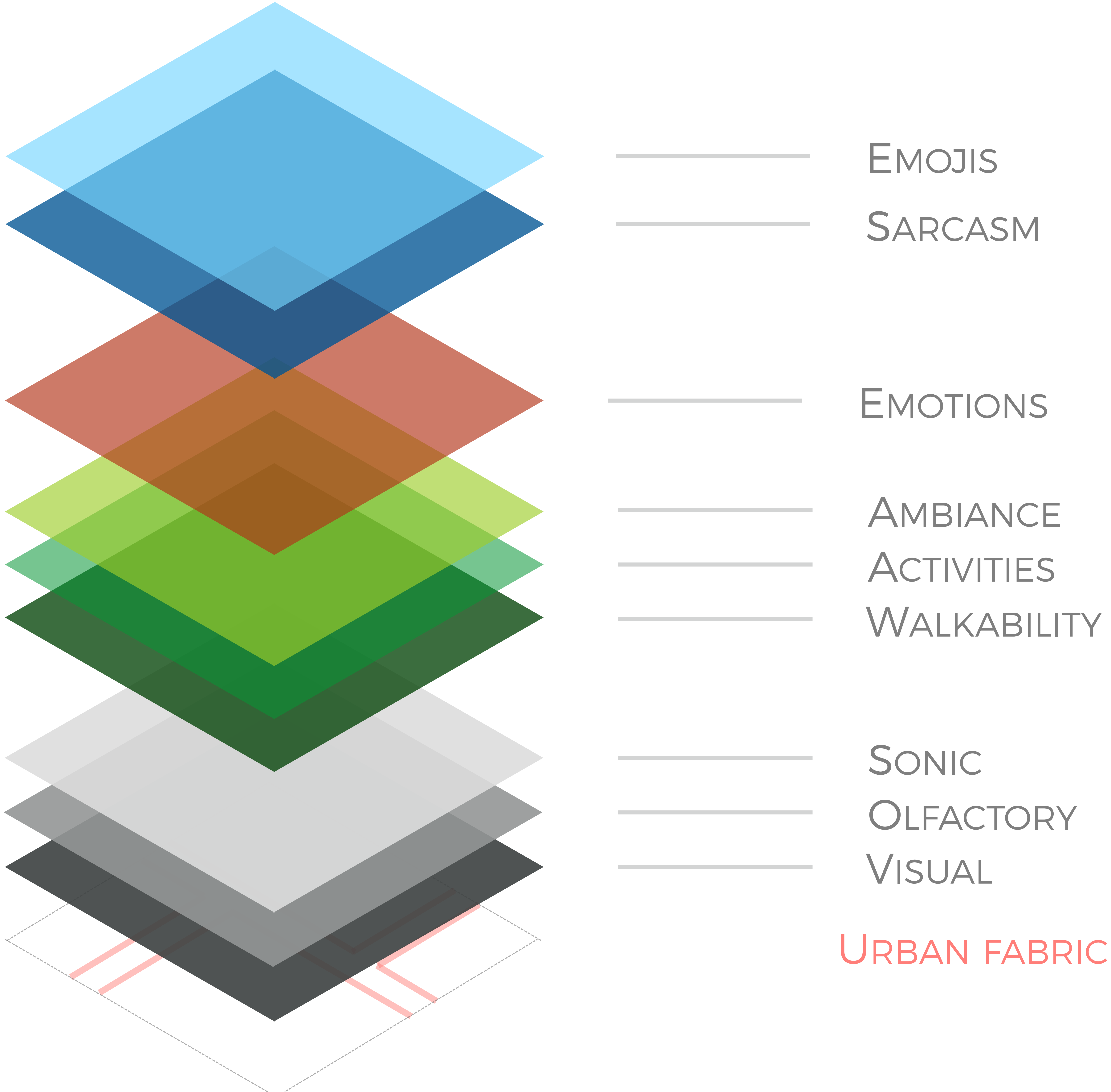
Music triggers both **joy** and **sadness**

Example (London)

sadness	-income, -employment, -health, -crime, -housing, -living environment
negative	-income, -health, -education, -employment
joy	+education, +housing


Emotions
Ambiance
Activities
Walkability
Sonic
Olfactory
Visual
Urban fabric

Ongoing work


EMOJIS
SARCASM
EMOTIONS
AMBIANCE
ACTIVITIES
WALKABILITY
SONIC
OLFACTORY
VISUAL
URBAN FABRIC

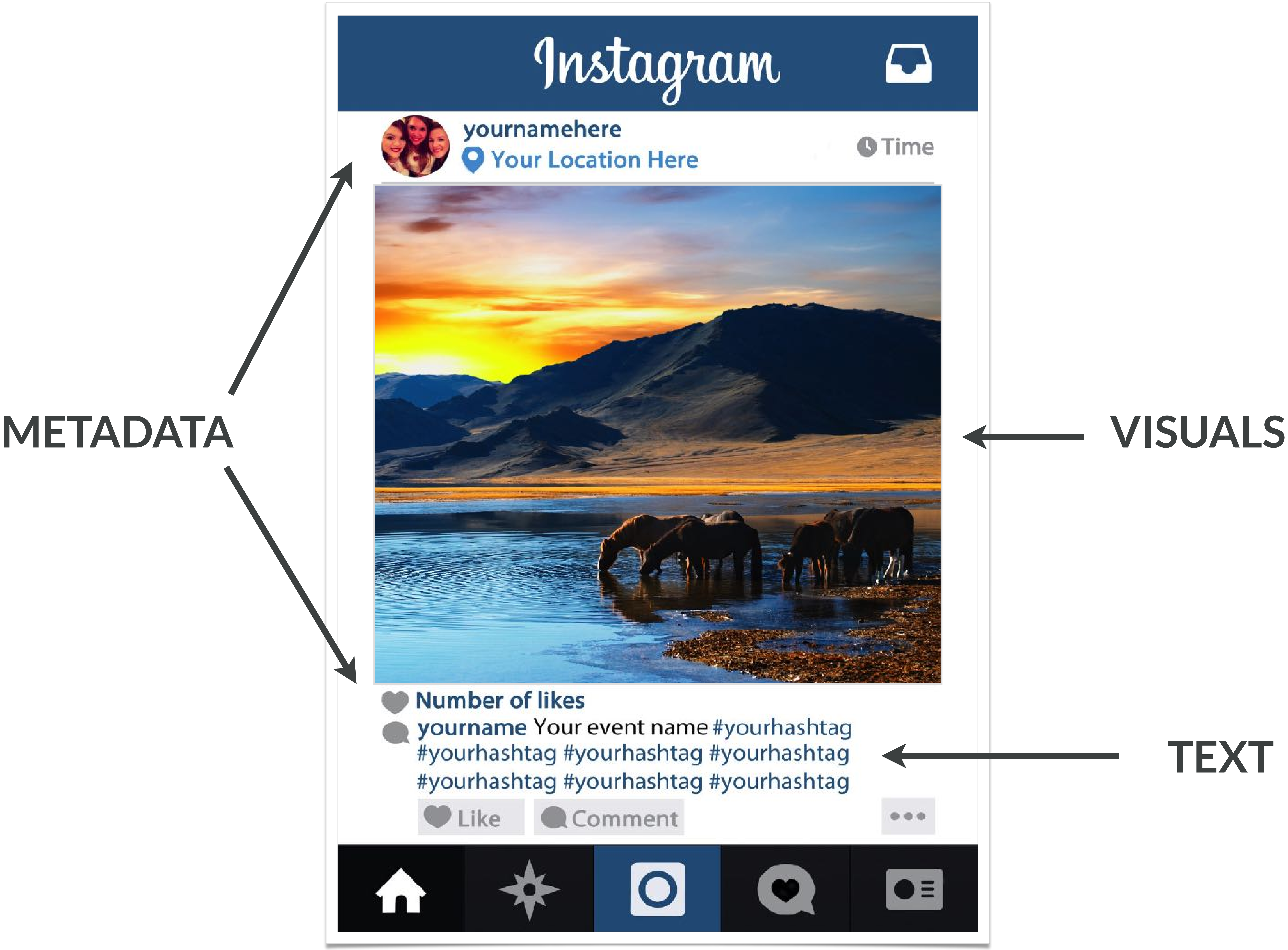
SARCASM

LITERAL ≠ INTENDED

Some previous work

- Lexical and linguistic markers
- Context
  - hashtags, emojis
  - previous posts
  - author profile, propensity to sarcastic utterances

SOCIAL MEDIA IS MULTIMODAL


METADATA
VISUALS
TEXT
Instagram
yournamehere
Your Location Here
Time
Number of likes
yourname Your event name #yourhashtag
#yourhashtag #yourhashtag #yourhashtag
#yourhashtag #yourhashtag #yourhashtag
Like
Comment


Great day today


Great day today

Text+Image

Image as a **contextual** clue

[POSTS CONTAINING #SARCASM OR #SARCASTIC]

DATA

517K	**63K**	**20K**
99% TEXT+IMAGE	40% TEXT+IMAGE	7.56% TEXT+IMAGE

1

CHARACTERISE THE ROLE OF IMAGES

Study of the interplay between textual and visual components

1

CHARACTERISE THE ROLE OF IMAGES

Study of the interplay between textual and visual components

2

COLLECT A GROUND TRUTH FOR SARCASM

A. Evaluate the impact of visuals as a source for context
B. Identify sarcastic posts with a high level of agreement

ASK THE CROWD!

1K POSTS

5 JUDGEMENTS

FIRST EXPERIMENT

Show only the textual component of a post

Our beautiful, balmy, sunny, summer holiday

Is this text sarcastic?

- Yes
- No
- I don't know

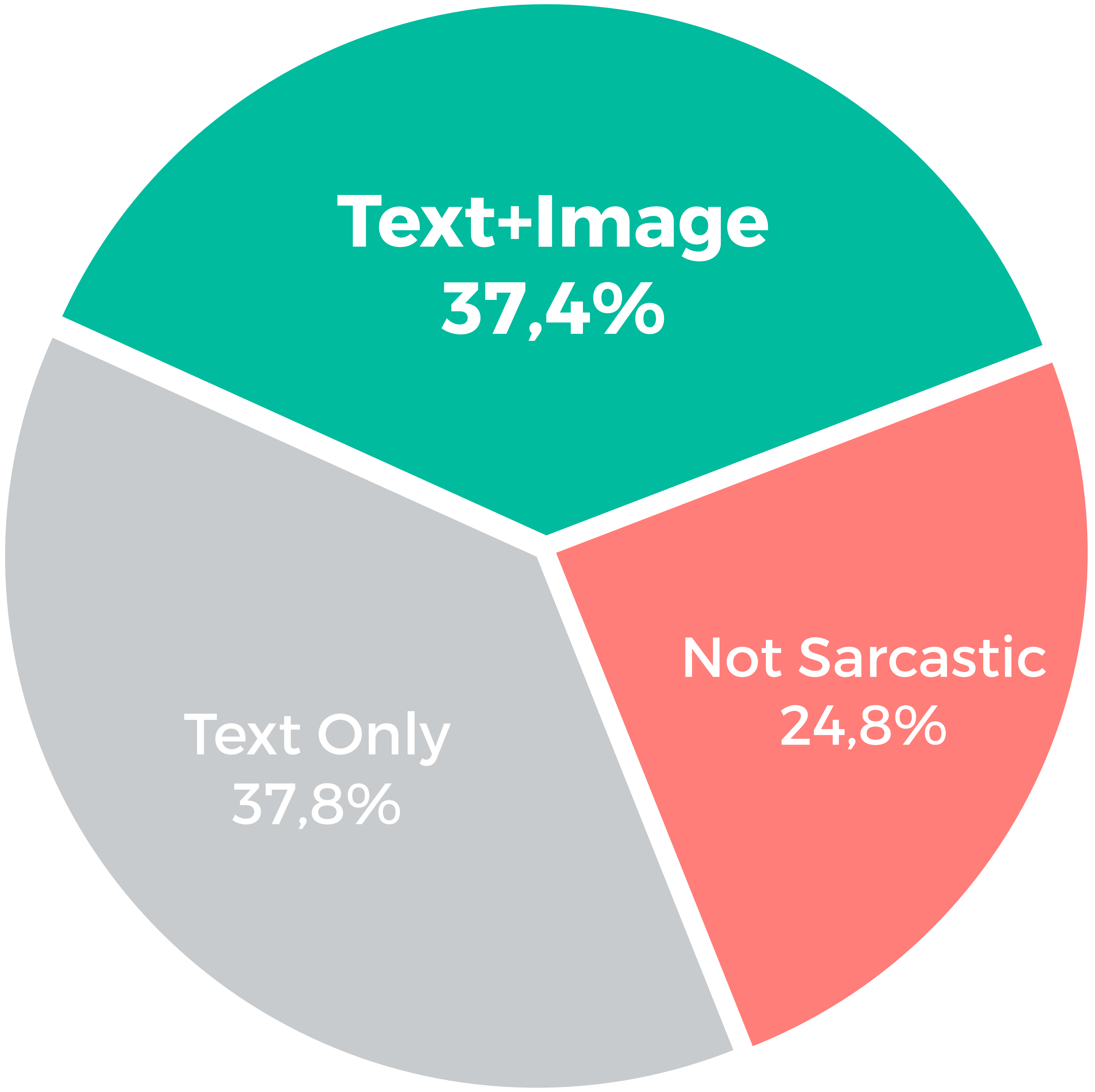
SECOND EXPERIMENT

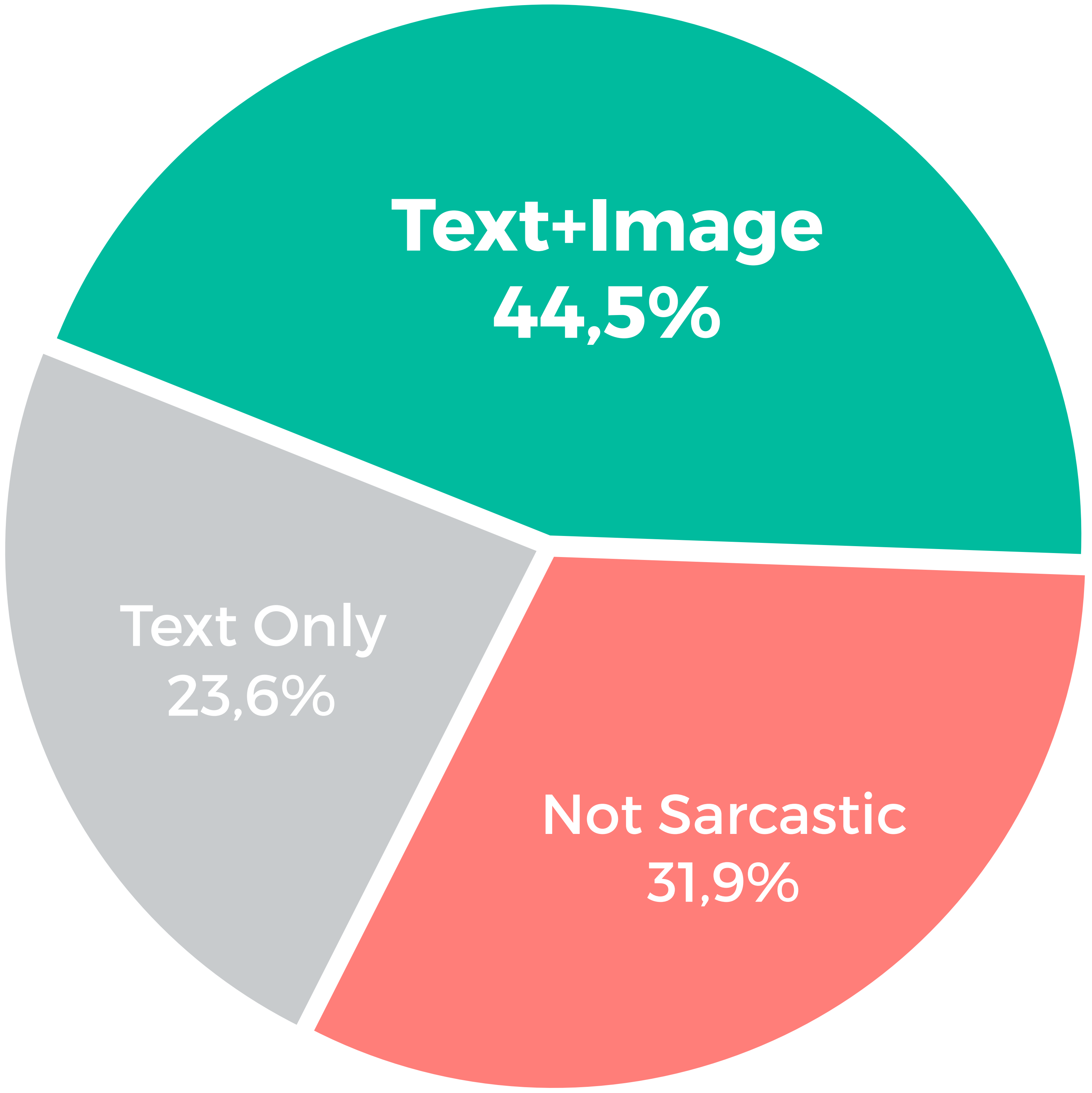
For all the posts that are judged **not** sarcastic in the previous step, show the text **and** the image

Our beautiful, balmy, sunny, summer holiday

Is this post sarcastic?

- Yes
- No
- I don't know


Text+Image
37,4%
Not Sarcastic
24,8%
Text Only
37,8%


Text+Image
44,5%
Text Only
23,6%
Not Sarcastic
31,9%

1

CHARACTERISE THE ROLE OF IMAGES

Study of the interplay between textual and visual components

2

COLLECT A GROUND TRUTH FOR SARCASM

A. Evaluate the impact of visuals as a source for context
B. Identify sarcastic posts with a high level of agreement

3

DETECT SARCASM

SVM Fusion+Deep learning fusion approaches

How can we detect sarcasm in multimodal posts?

- **Different fusion approaches**
  - SVM based
  - Deep learning
- **Open questions:**
  - Does the use of figurative language change according to socio-demographic variables?
  - Does the use of figurative language change in different areas of the city?

Questions?

@rschifan

schifane@di.unito.it

http://www.di.unito.it/~schifane