

# Preferred citation style for this presentation

---

Michael VAN EGGEMOND, Haohui CHEN, Alexander ERATH, Manuel CEBRIAN (2014), Investigating the potential of social network data for transport demand models, Washington DC, January 2015.

# Investigating the potential of social network data for transport demand models

Michael van Eggermond  
Haohui Chen  
Alexander Erath  
Manuel Cebrian  
TRB 2015

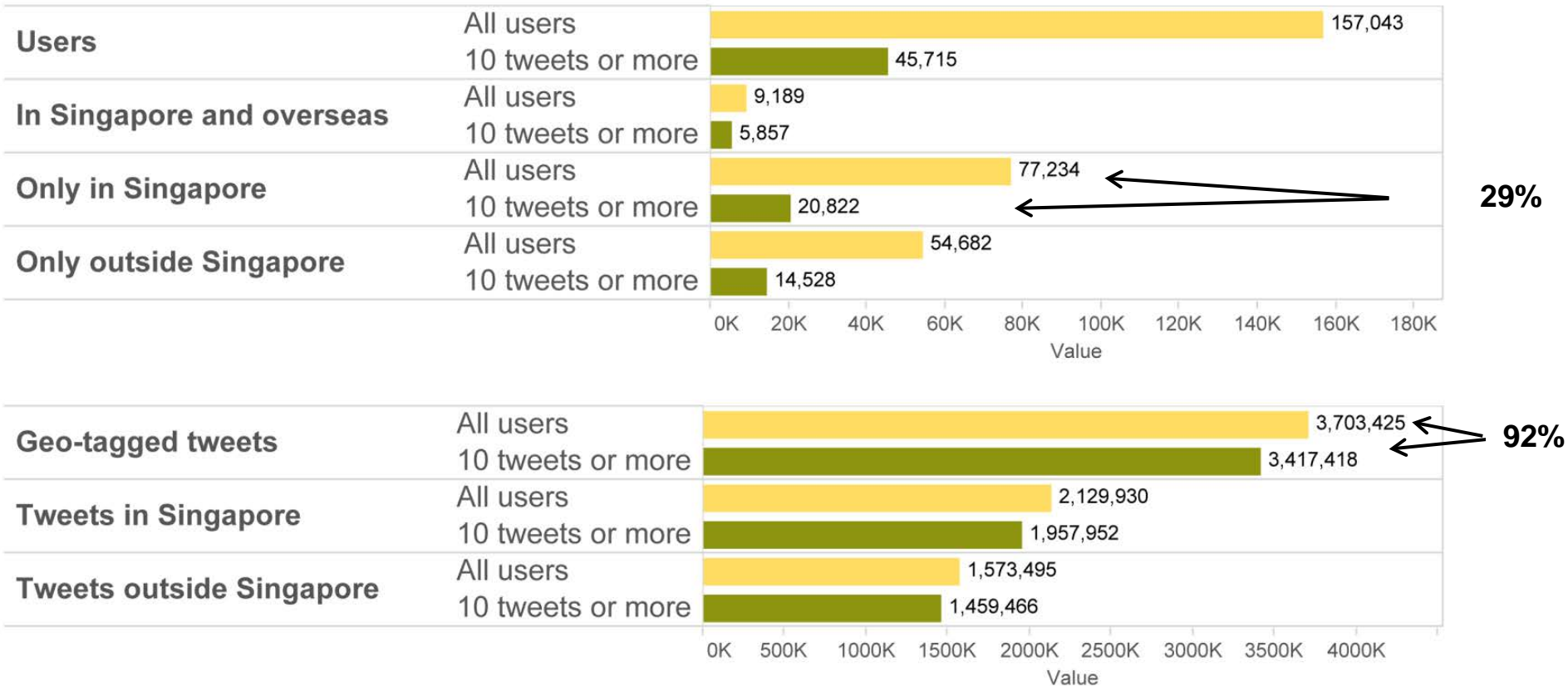
(FCL) FUTURE CITIES  
LABORATORY 未来城市实验室

(SEC) SINGAPORE-ETH CENTRE 新加坡-ETH 研究中心



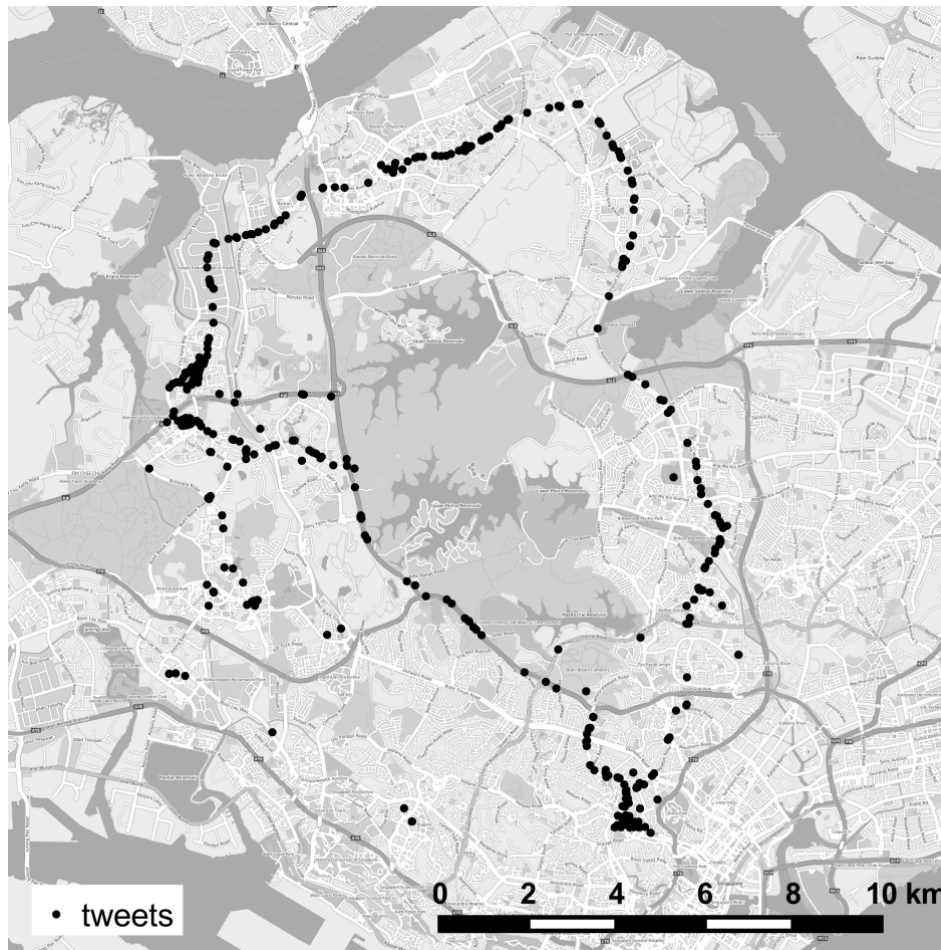
# Motivation

## Twitter - September, 2013 to February, 2014



# Methodology & Data

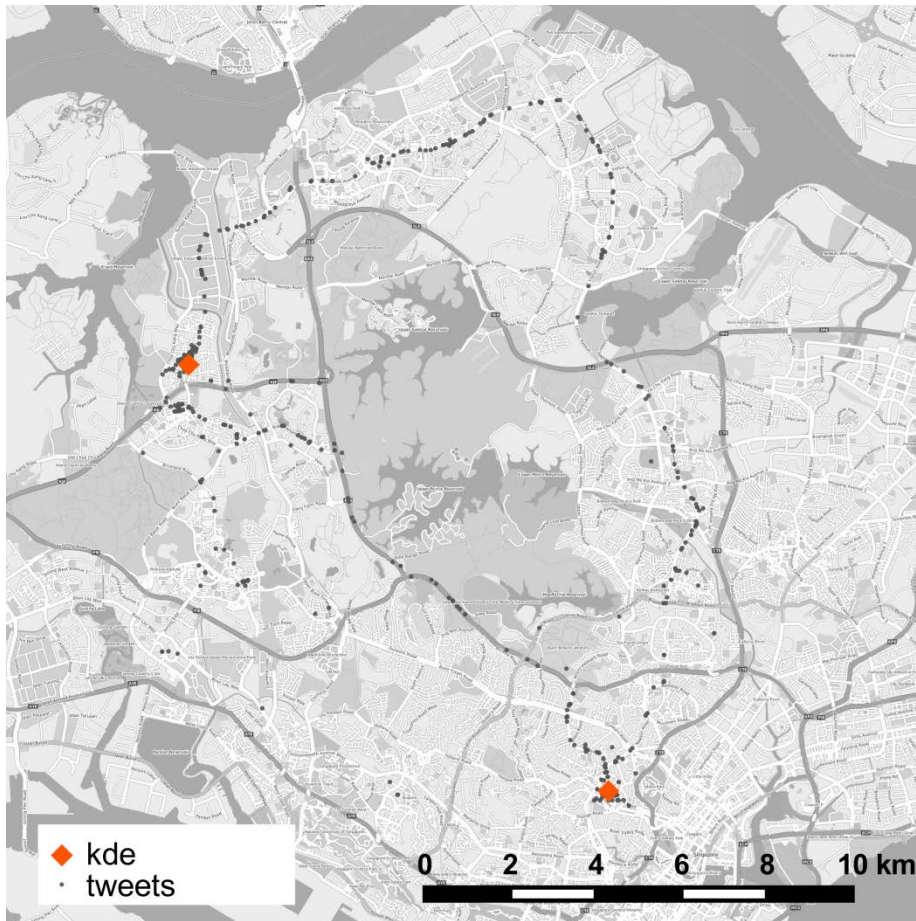
## Tweets of a random user (1,405)



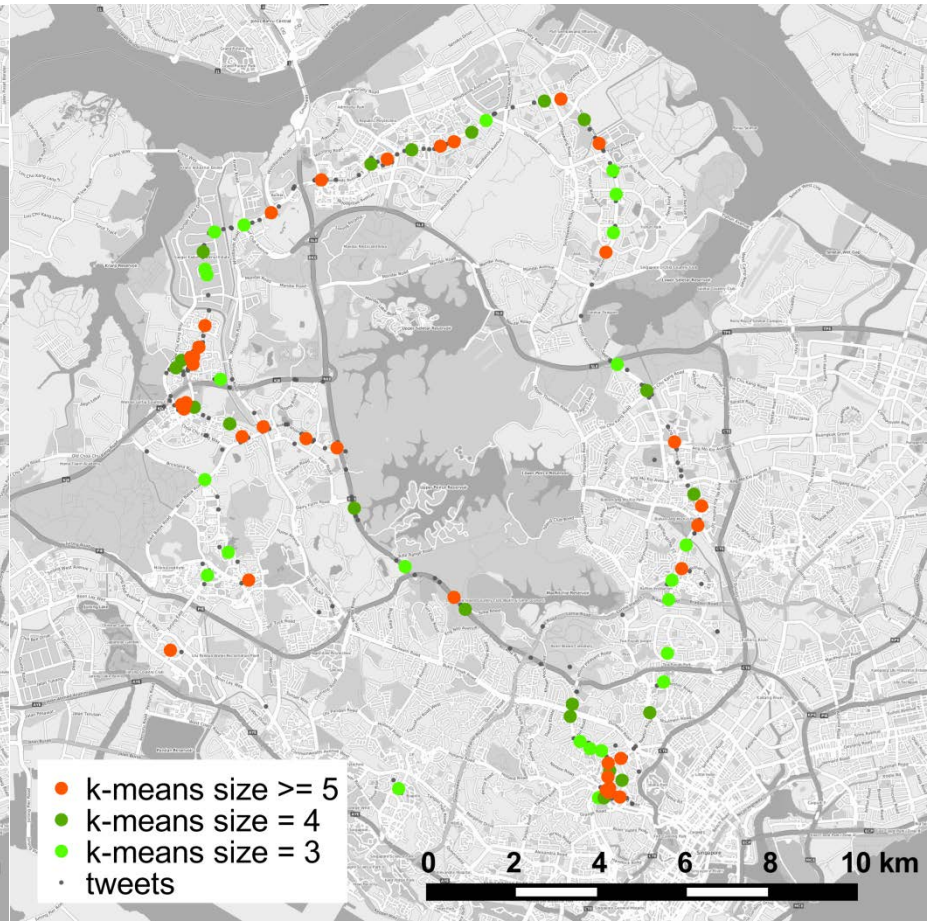
Background map: <http://openstreetmap.org>.

# Clusters of one individual

## Kernel density estimation



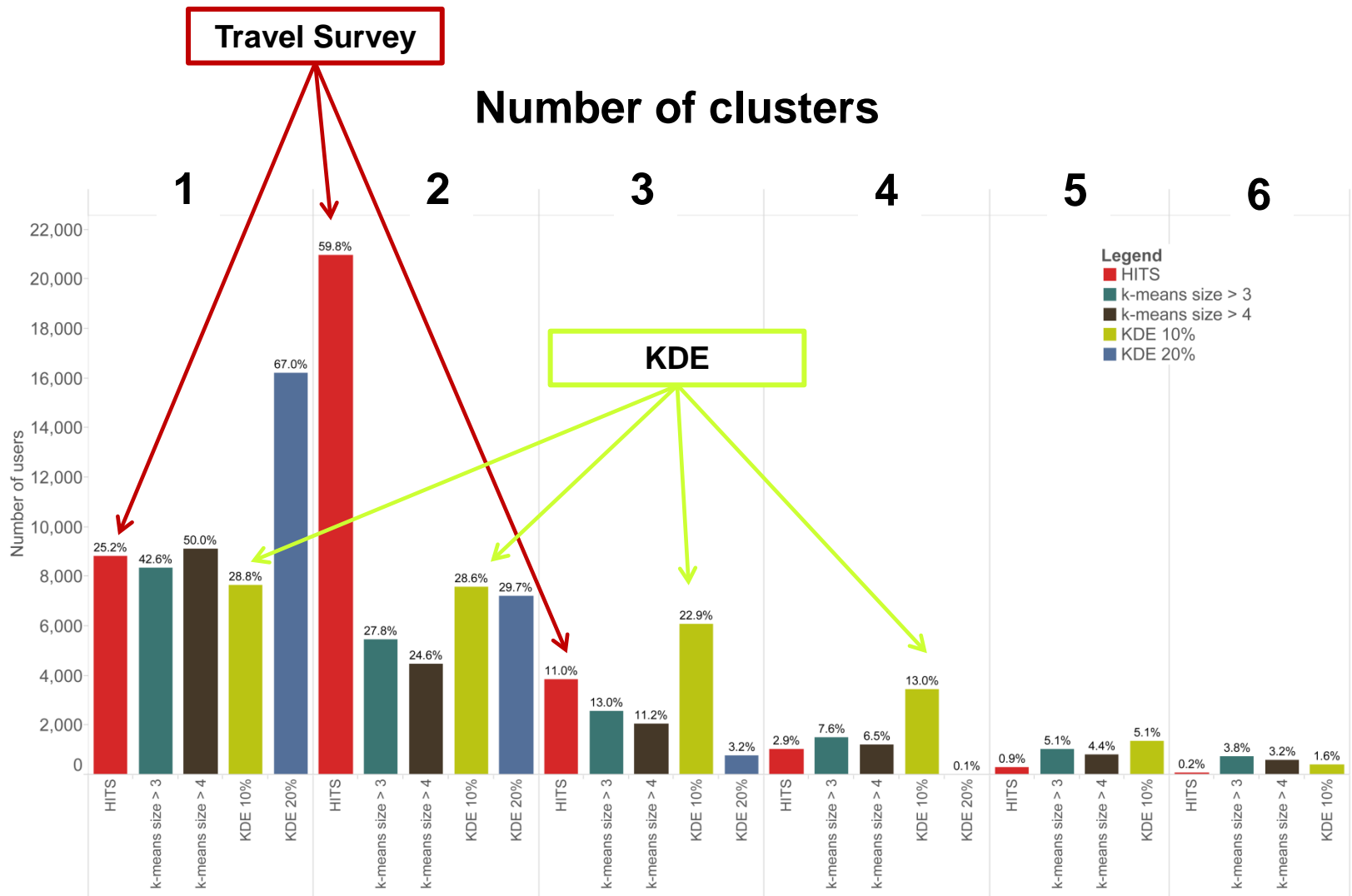
## K-means clustering



Background map: <http://openstreetmap.org>.

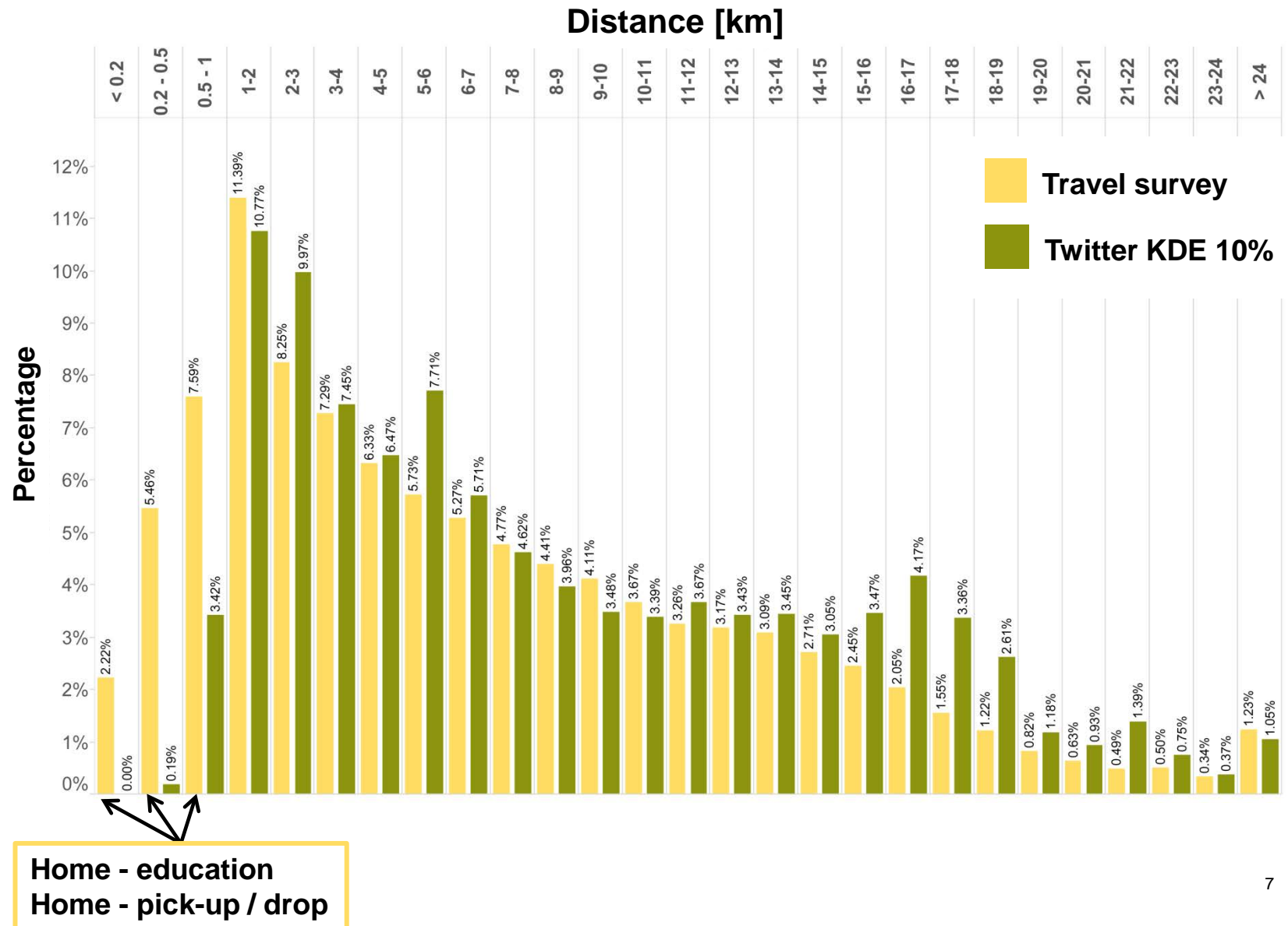


# Clustering for all users



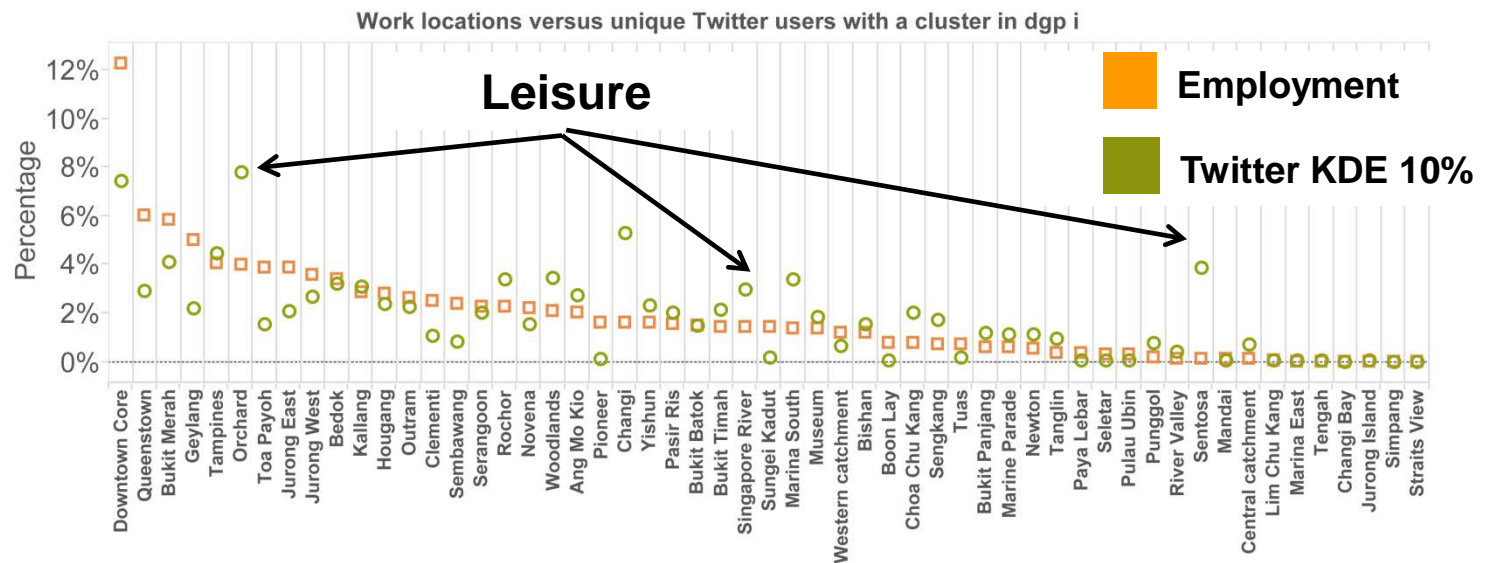
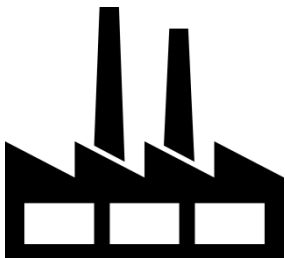
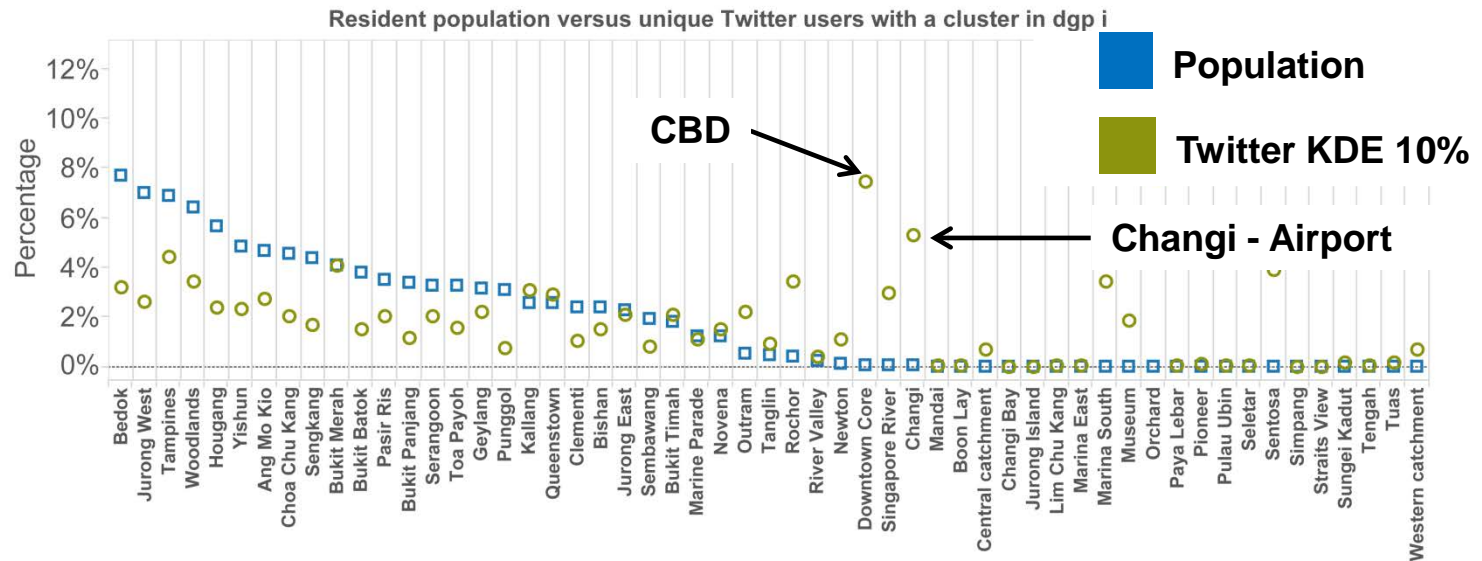


# Distance between clusters





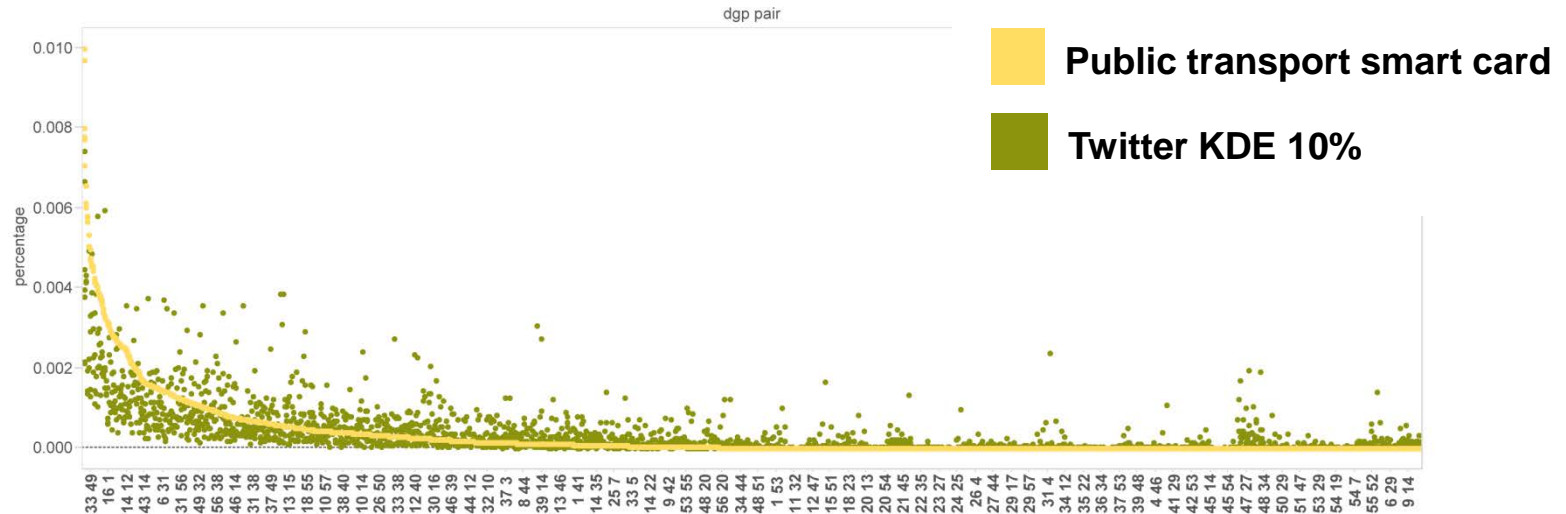
# Population & employment statistics



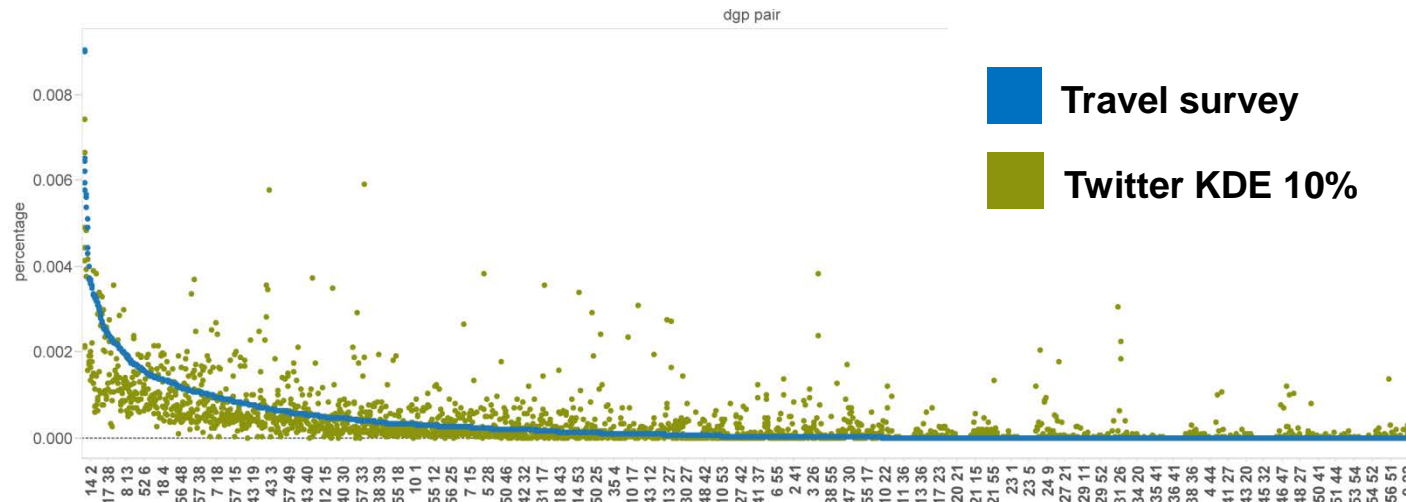


# OD relations

## Public transport smart card vs Twitter



## Travel Survey vs Twitter





# Summarizing

---



## **Location-based social network data offers:**

Insight in activity locations

Separation between clusters



## **Differences:**

Due to population sub-sample? Probably yes.

Is social media used more at certain locations? Probably yes.

Scaling?



## **And:**

Not as rich as travel survey -> activity duration, start times

Definition of cluster

Activity definition?

# Questions

---

Michael van Eggermond  
eggermond@ethz.ch

# Literature

---

Lichman, M. and P. Smyth (2014) Modeling Human Location Data with Mixtures of Kernel Densities, paper presented at the *KDD '14*, New York.

# Acknowledgements

---

The research conducted at the Future Cities Laboratory is co-funded by the Singaporean National Research Fund and the ETH Zurich, and located at the Campus for Research Excellence And Technological Enterprise (CREATE). NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy 1 and the Australian Research Council through the ICT Centre of Excellence program. The authors would like to thank Professor Kay W. Axhausen for his suggestions and comments. We wish to express our gratitude to the Land Transport Authority for providing us invaluable data sets on transport in Singapore. Also we are very thankful to the Urban Redevelopment Authority for providing us with a wide range of data sets.

# Appendix

---



# Study area



## Data (2)

---

---

### **Twitter (September 2010, 2013 to February 27, 2014)**

---

---

### **Household Interview Travel Survey 2008**

---

Number of households	10,641
Number of persons	36,978

---

---

### **Public transport smart card data**

---

Number of card identifiers	3,475,574
Number of journeys over 7 days	23,994,771

---

---

### **Singapore statistics**

---

Total population	5,319,000
Total resident population	3,825,000
Land-area 2013 [km2]	716.1

---