



## Module INM433 – Visual Analytics

### Practical 04

# Interactive visual analysis of OD movement data

given by

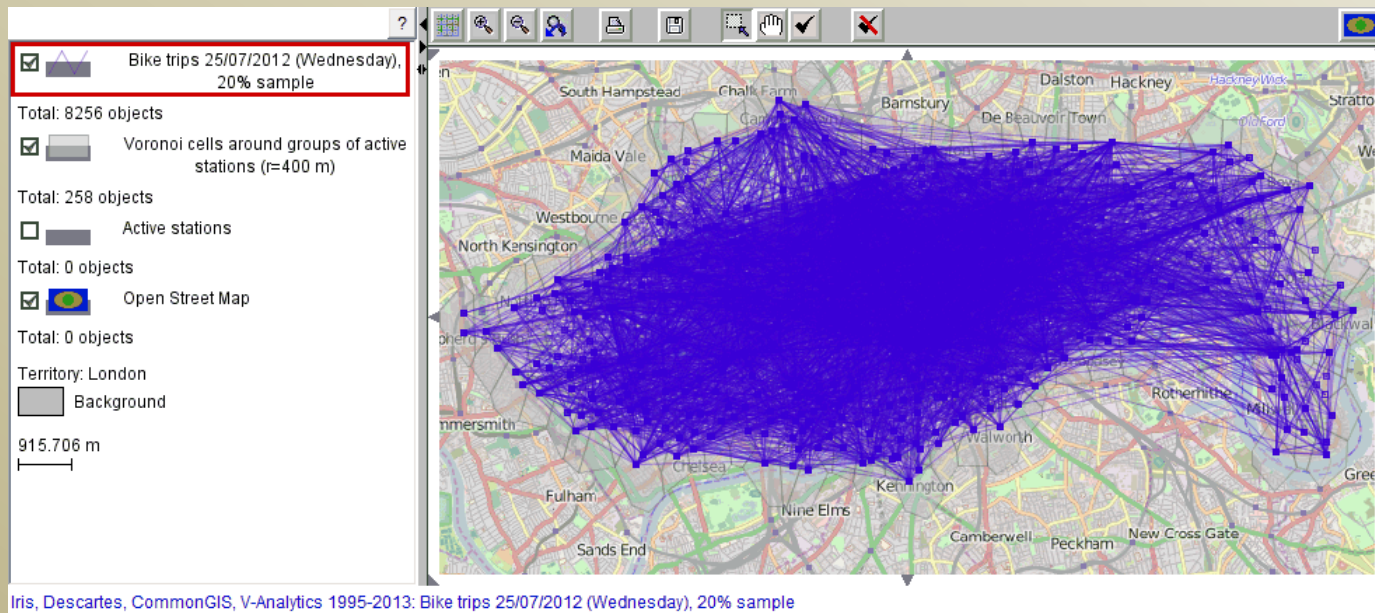
prof. Gennady Andrienko and

prof. Natalia Andrienko



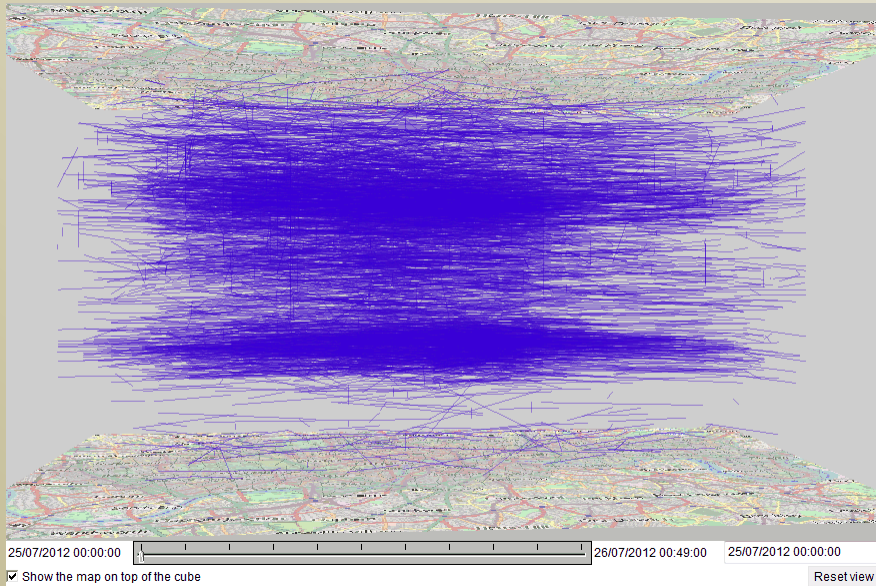
# Data

- A 20% random sample of the London bike trips from 25/07/2012 (Wednesday)
- Load the project file **trips\_Wednesday.app** from folder exercises/OD\_moves

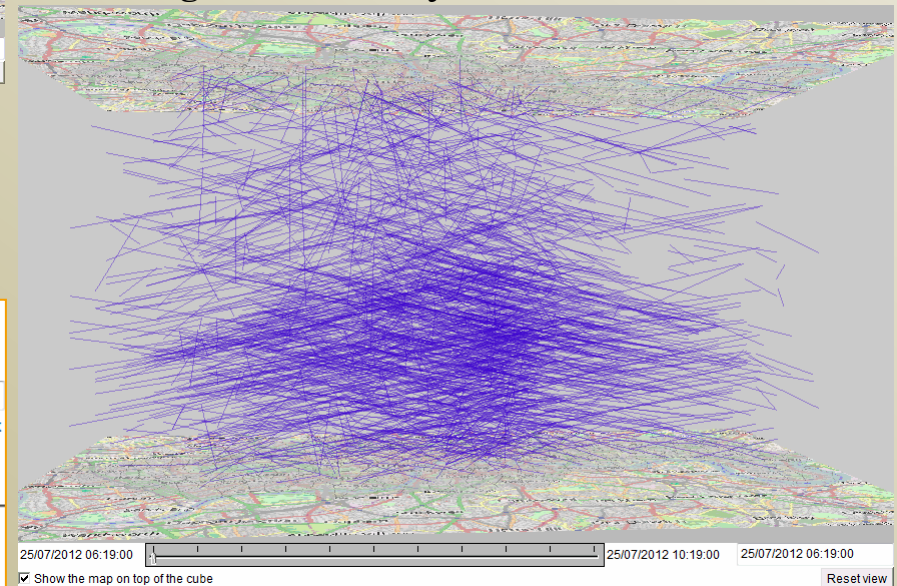
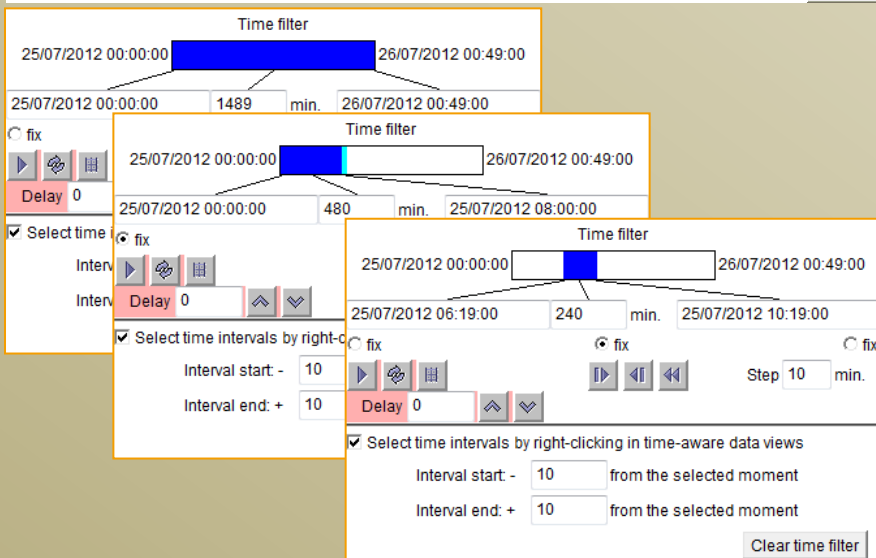




# Exercise 1: Visual representation of OD moves as spatio-temporal objects



1. Open a space-time cube view with the bike trip data (*Display > Space-time cube*).
2. For temporal zooming in the STC, use the time filter (*Filter > Time filter*).
  - Manipulate the length and position of the time slider.
3. Pay attention to the differences in the line slopes. What do they mean?
4. Observe the variation of the line density throughout the day.







# Find vectors representing round trips

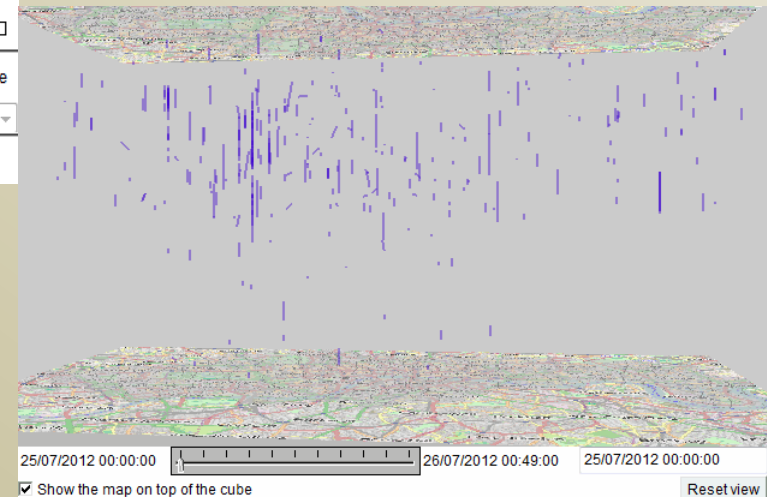
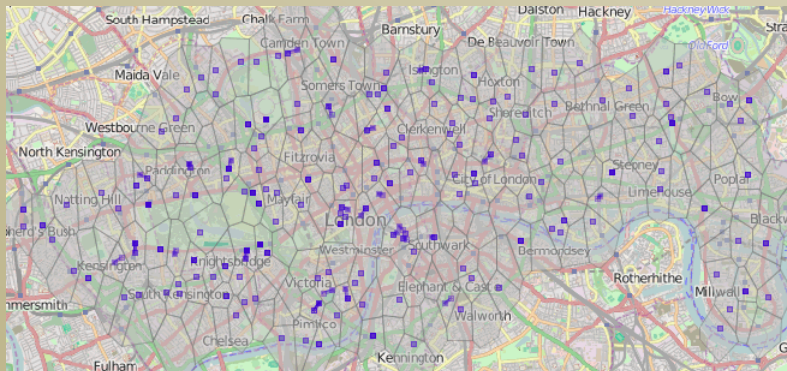
- Use attribute-based filter
  - Filter > Attribute-based filter > Select table “Bike trips 25/07/2012...” > Select attribute “Track length” > set the upper bound to 0.2 km
- Clear the time filter
- To better see the vectors in the STC, increase the line thickness
  - For changing the drawing settings, click on the layer’s icon in the map legend.

Dynamic Query for Bike trips 25/07/2012, 20% sample: general data

	yes	no	Track length		
	<input checked="" type="radio"/>	<input type="radio"/>	0.00	12.00	3.6% : 300 from 8256
	<input type="radio"/>	<input checked="" type="radio"/>	0.00	0.20	3.6% : 300 from 8256

☒ Filter out missing values  ☒ Display statistics ☐ Dynamic update

[Query results -> Qualitative attribute](#)





# Question to exercise 1

- What characteristics of the trips are represented by the following properties of the lines (vectors)?
  - length in the spatial dimension
  - length (height) in the temporal dimension
  - inclination



## Exercise 2: Density-based clustering of OD moves by the spatial positions of their starts and ends

- Invert the attribute-based filter (i.e., filter out the round trips)
- Activate the clustering tool:
  - Menu “Analyse” > “OD moves: density-based clustering” > a dialog appears; the layer with the moves is pre-selected > press OK
- Set the clustering parameters
  - Uncheck the check box “start and/or end times” (the times will not be taken into account in this exercise).
  - Set the spatial distance threshold to 300m.
  - Set the minimal size of the clusters you are interested in to 10 (in the text field following the check box “Ignore clusters with less than” at the bottom of the dialog).
  - Pressing OK starts the clustering.
- After the clustering finishes, the system shows the results by colouring the lines on the map and in the STC according to the cluster membership.

Dimensions of the trip set

X-extent: 15359.13 m  
Y-extent: 6559.44 m  
Time span: 25/07/2012 00:00:40 .. 26/07/2012 00:48:03  
Number of active trips: 7956 (96.37% of the total 8256 trips)

The trips to be clustered by

☒ start positions  
☒ end positions  
☐ start and/or end times ☒ absolute ☐ time of the day

Define the spatial and temporal neighbourhood of a trip:

Spatial distance threshold: 300 m  
Temporal distance threshold: 900 seconds

☐ Use additional attributes of the trips

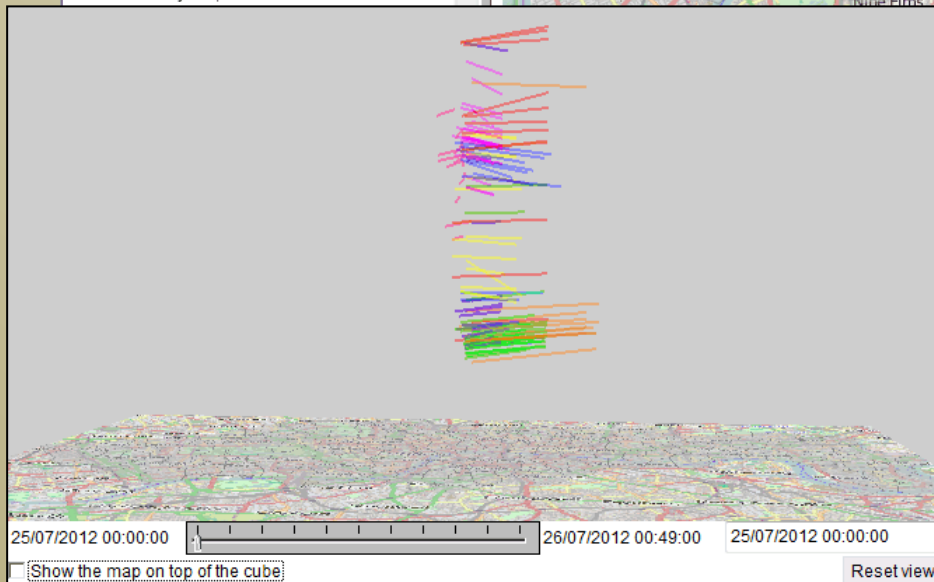
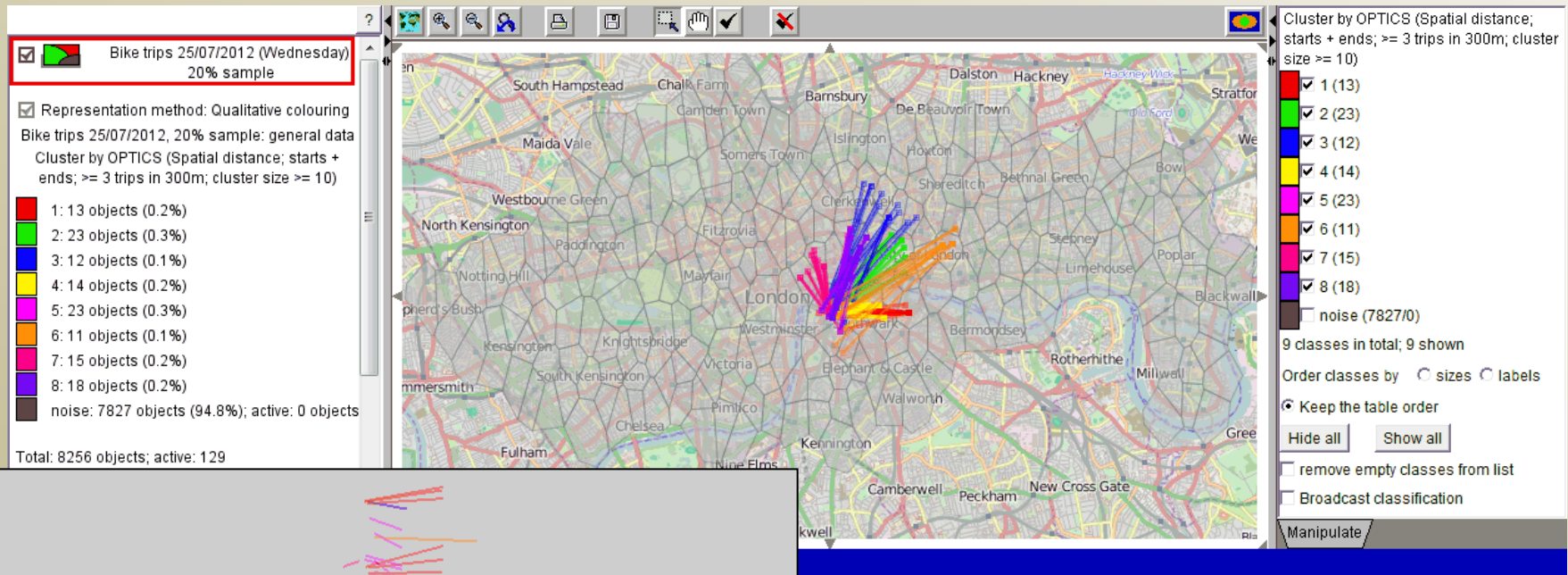
Minimal number of trips in the neighbourhood \*: 3  
\* required for a trip to be in cluster core

☒ Ignore clusters with less than 10 trips

OK Cancel



# Exploration of the clustering results



Switch off the “noise” and observe the spatial characteristics of the trip clusters on the map: positions of the trip origins (hollow squares) and destinations (filled squares), movement directions, and displacement distances (*use map zooming when needed*).

Look also in the STC: are there trip clusters that mostly occurred in certain time periods (morning, midday, afternoon)?





# Testing the impact of the distance threshold

- Switch on the “noise”, to have all moves visible.
- Move the current visualisation to another window:
  - “Display” > “Move the map to another window”
- Start again the clustering tool as previously
- In the dialog for setting the parameters, uncheck again the check box “start and/or end times” and change the spatial distance threshold to 350 m instead of the previous 300 m.
- Run the clustering algorithm and visually explore the results. Compare with the results of the previous clustering.

Dimensions of the trip set

X-extent: 15359.13 m  
Y-extent: 6559.44 m  
Time span: 25/07/2012 00:00:40 .. 26/07/2012 00:48:03  
Number of active trips: 7956 (96.37% of the total 8256 trips)

The trips to be clustered by

☒ start positions  
☒ end positions  
☐ start and/or end times ☒ absolute ☐ time of the day

Define the spatial and temporal neighbourhood of a trip:

Spatial distance threshold: 350 m  
Temporal distance threshold: 900 seconds

☐ Use additional attributes of the trips

Minimal number of trips in the neighbourhood \*: 3  
\* required for a trip to be in cluster core

☒ Ignore clusters with less than 10 trips

OK Cancel





# Questions to Exercise 2

- Describe the differences between the results of the clustering with the different distance thresholds. Consider the following aspects:
  - Number of clusters
  - Sizes of the clusters
  - Amount of “noise”: how many moves that were earlier in the “noise” have been included in clusters?
    - Hint: switch off the “noise” in the new map and look in the legend of the older map.
  - Spatial properties of the additional clusters.
  - Internal variance within the clusters.
    - Hint: select several largest clusters one by one after ordering the clusters by the sizes.



# Exercise 3: Density-based spatio-temporal clustering of OD moves

- Switch on the visibility of all clusters and “noise”.
- Move the current visualisation to another window:
  - “Display” > “Move the map to another window”
- Start again the clustering tool as previously
- In the dialog for setting the parameters:
  - The check box “start and/or end times” must be checked – now the times will be taken into account.
  - Set the temporal distance threshold to 30 minutes.
  - Set the minimal cluster size to 5.
- Run the clustering algorithm and visually explore the results using the map and the space-time cube.

Dimensions of the trip set

X-extent: 15359.13 m  
Y-extent: 6559.44 m  
Time span: 25/07/2012 00:00:40 .. 26/07/2012 00:48:03  
Number of active trips: 7956 (96.37% of the total 8256 trips)

The trips to be clustered by

☒ start positions  
☒ end positions  
☒ start and/or end times ☐ absolute ☐ time of the day

Define the spatial and temporal neighbourhood of a trip:

Spatial distance threshold: 350 m  
Temporal distance threshold: 30 minutes  
☐ Use additional attributes of the trips

Minimal number of trips in the neighbourhood \*: 3  
\* required for a trip to be in cluster core

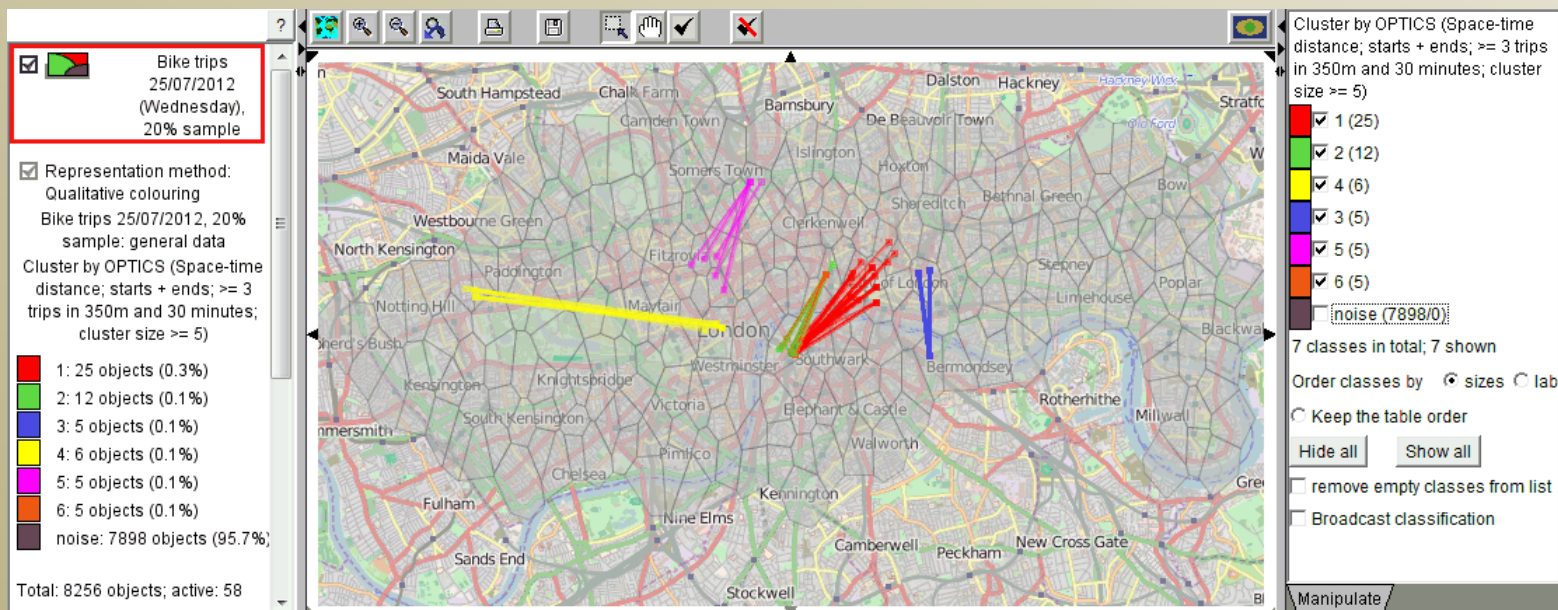
☒ Ignore clusters with less than 5 trips

OK Cancel



# Questions to exercise 3

- Describe the spatial and temporal characteristics of the spatio-temporal clusters of OD moves.
- How are the largest spatio-temporal clusters (size $\geq 10$ ) related to the earlier obtained spatial clusters (results of the previous run)?





# Exercise 4: Spatio-temporal aggregation of OD moves

- Cancel all filters (cluster selection and attribute-based filter).
- Close all additional windows (leave only the main window) and clean the main map (remove the cluster visualisation).
- Start the aggregation tool:
  - “Analyse” > “OD moves: spatio-temporal aggregation” > a dialog appears with two pre-selected map layers (with the moves and with the space compartments) > press “OK”
  - A time division dialog appears; the temporal resolution “hour(s)” and the interval length “1” are proposed at the bottom > Press “Divide” > A dialog showing the number of breaks appears > press “Yes” > The list of breaks appears in the list box > Press “OK” > A dialog asking about finding useless breaks appears > Press “Yes” > The tool informs about removing a useless break > Press OK
  - Press “OK” in each of the following dialogs (i.e., agree to the use of the default settings) .

1 breaks were removed as useless (not in the time range of the data); pre

Date/time specification scheme: dd/mm/yyyy hh:tt:ss  
(‘t’ stands for minutes; separator symbols are irrelevant)

Divide the time period

from 25/07/2012 00:00:00 to 26/07/2012 00:48:03  
Actual data range: from 25/07/2012 00:00:40 to 26/07/2012 00:48:03

Wednesday Thursday

as ☒ time line ☐ cycle of hours in week

by breaks:

25/07/2012 01:00:00	Remove
25/07/2012 02:00:00	
25/07/2012 03:00:00	
25/07/2012 04:00:00	
25/07/2012 05:00:00	
25/07/2012 06:00:00	Remove all

New break: Add

Automatic division:

- into 2 equal intervals Divide

- into intervals of length 1 hour(s) Divide

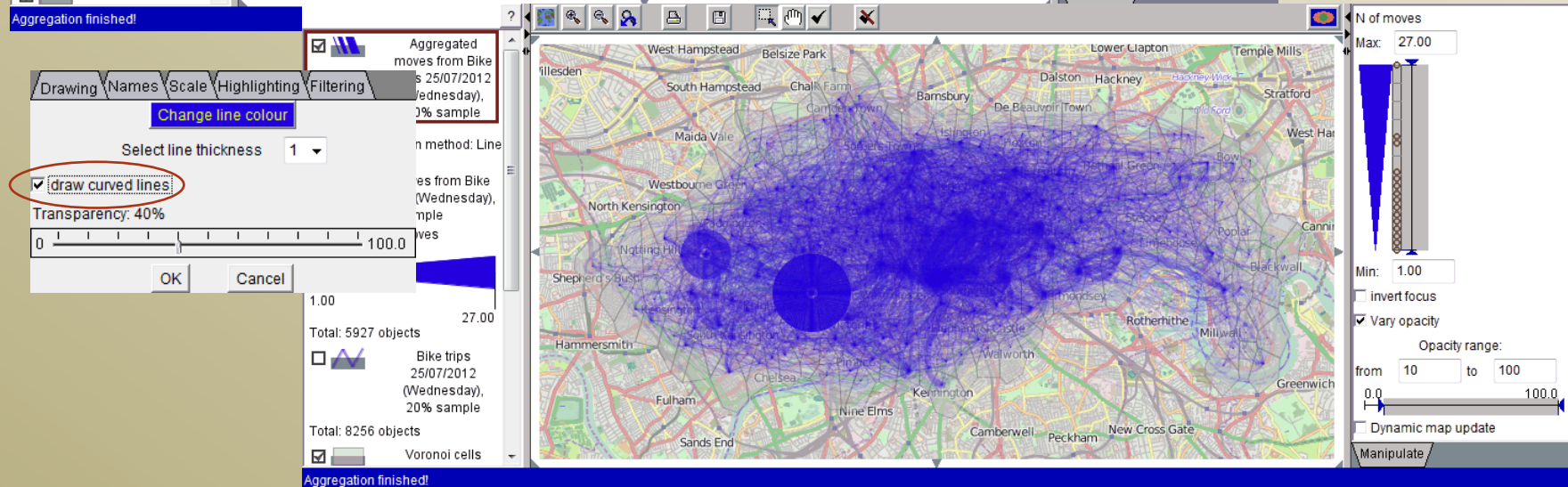
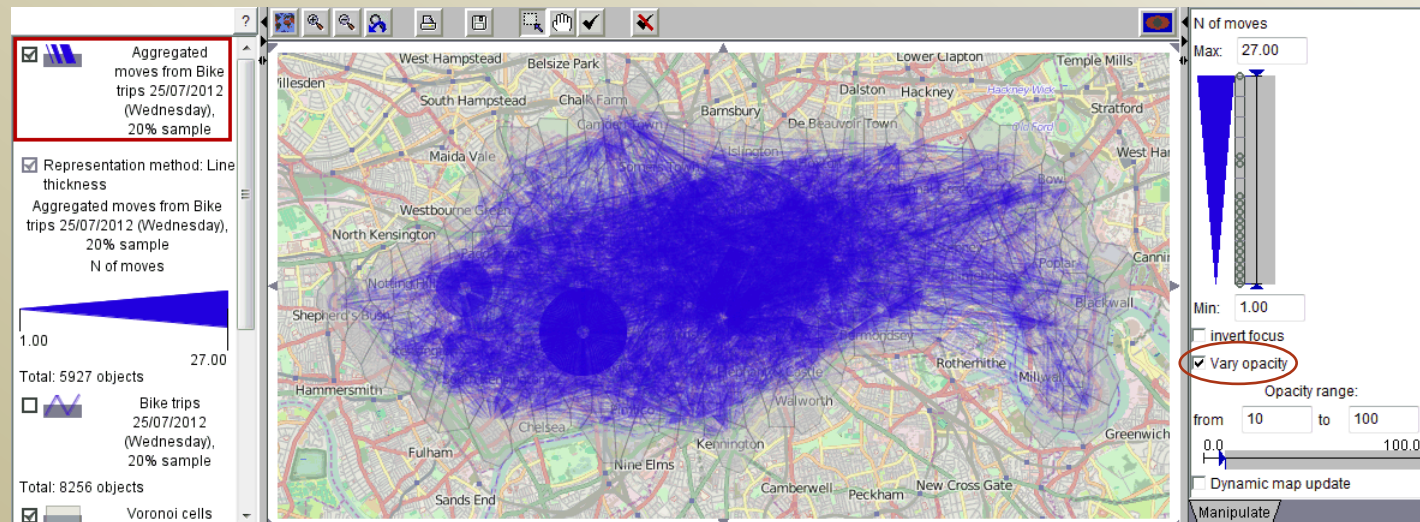
☒ Adjust the breaks to the selected unit

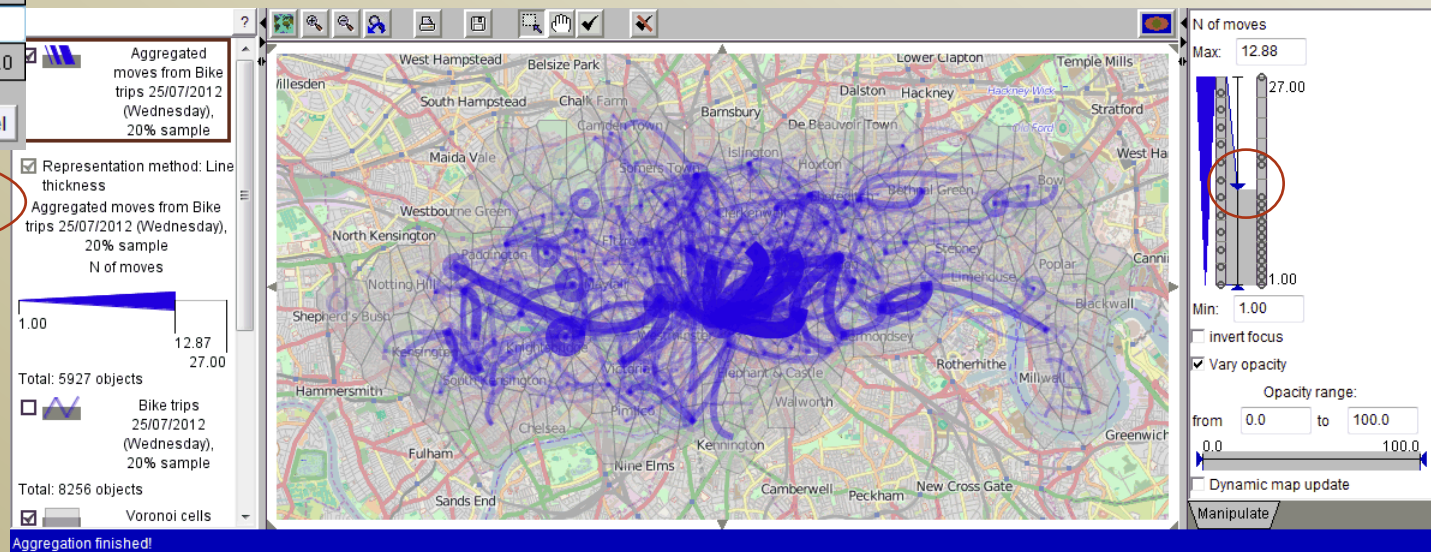
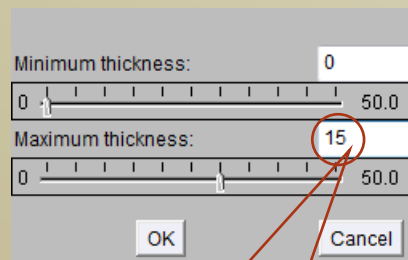
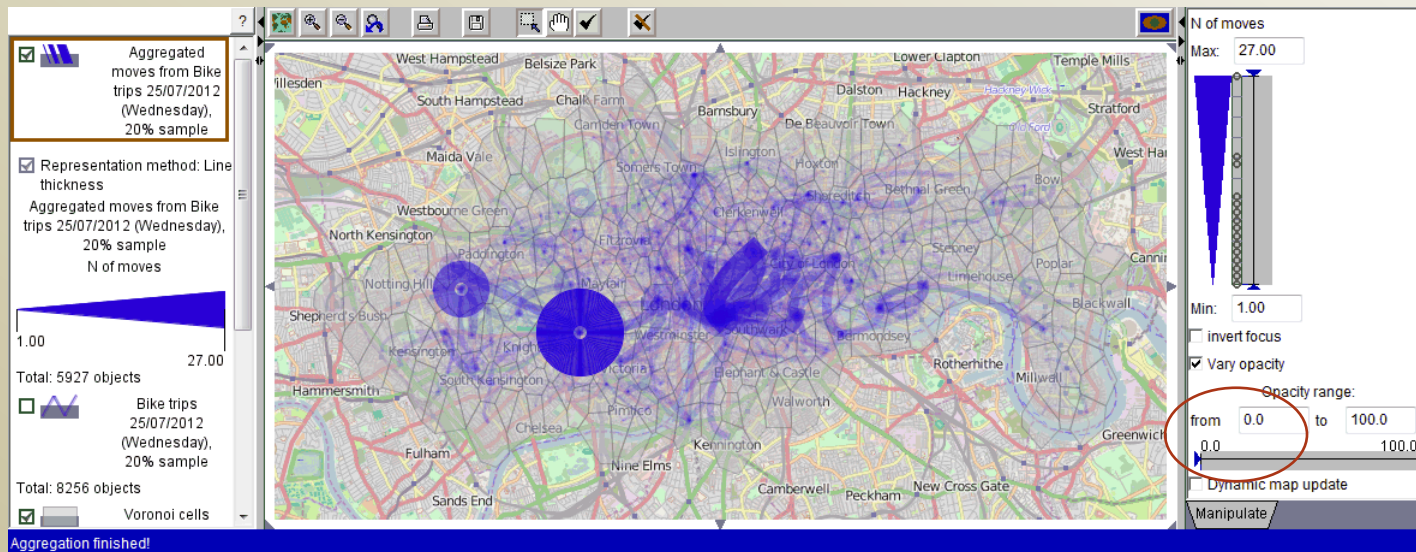
OK Cancel





# Visual representation of flows resulting from the aggregation

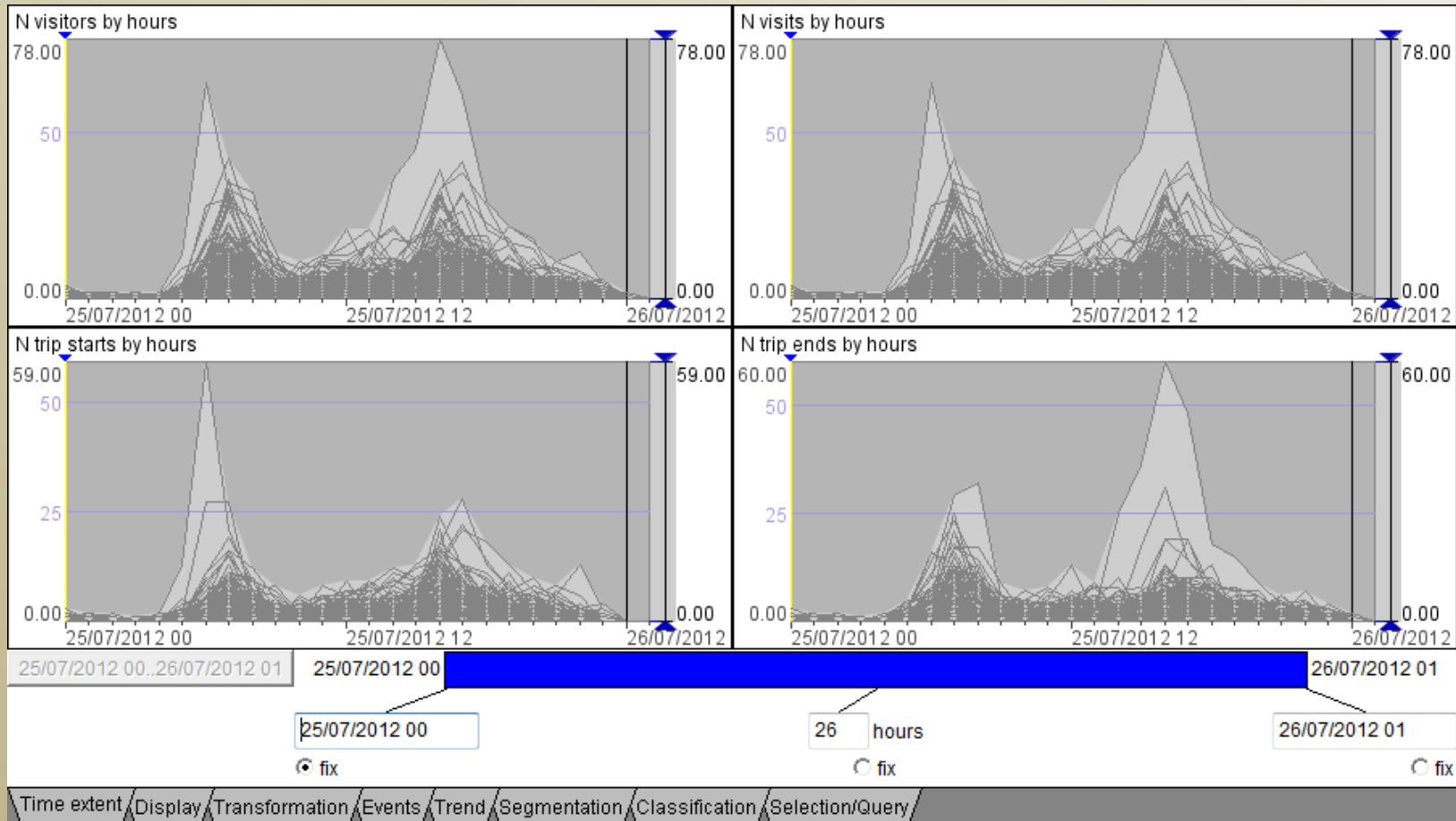




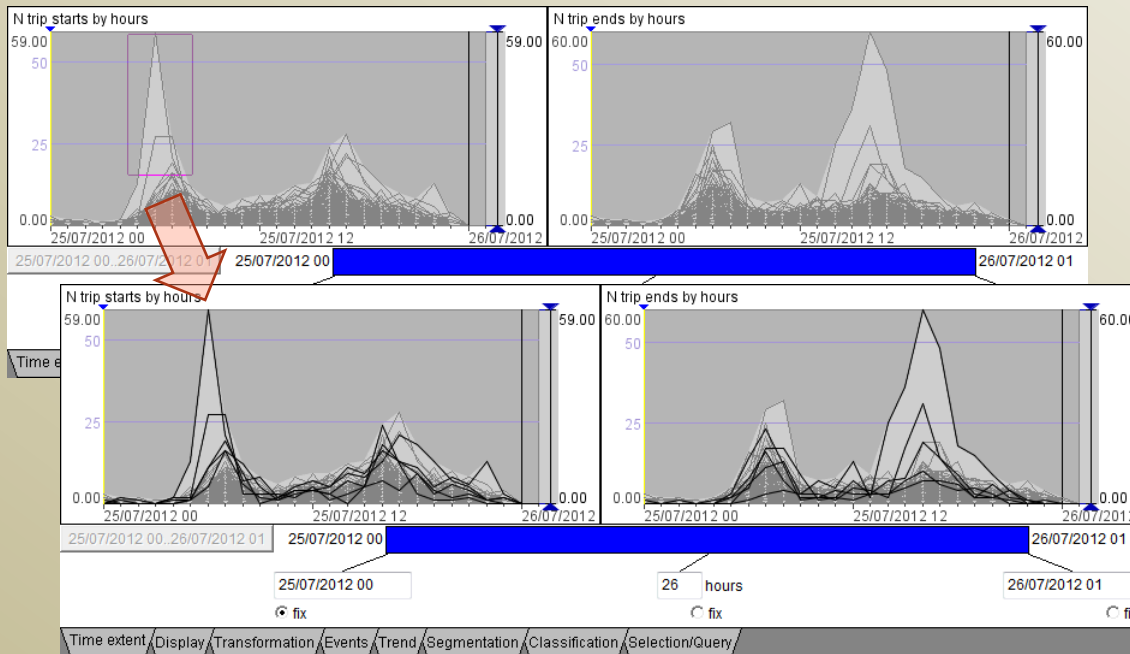




# Place-referenced time series

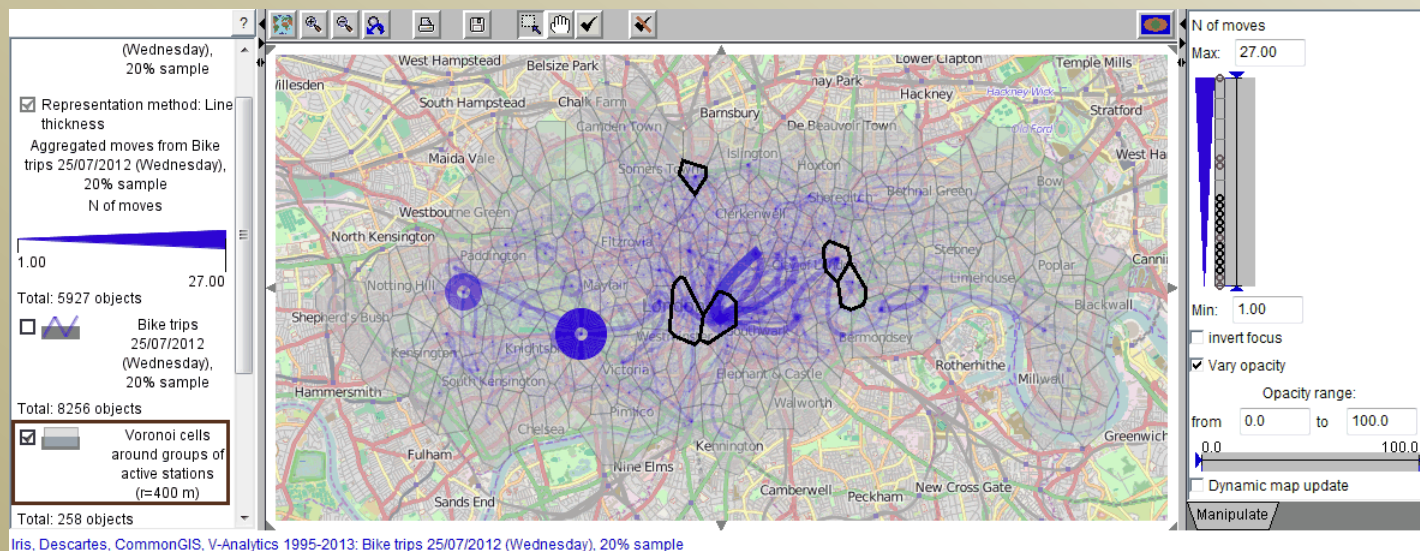


Display > Display wizard > Select table “Voronoi cells ...” > Select time-variant attributes > Select “Time graph”



By selecting lines in the time graphs of the counts of trip starts and ends and looking at the map, find the most popular places of trip origins and destinations in the morning and in the afternoon.

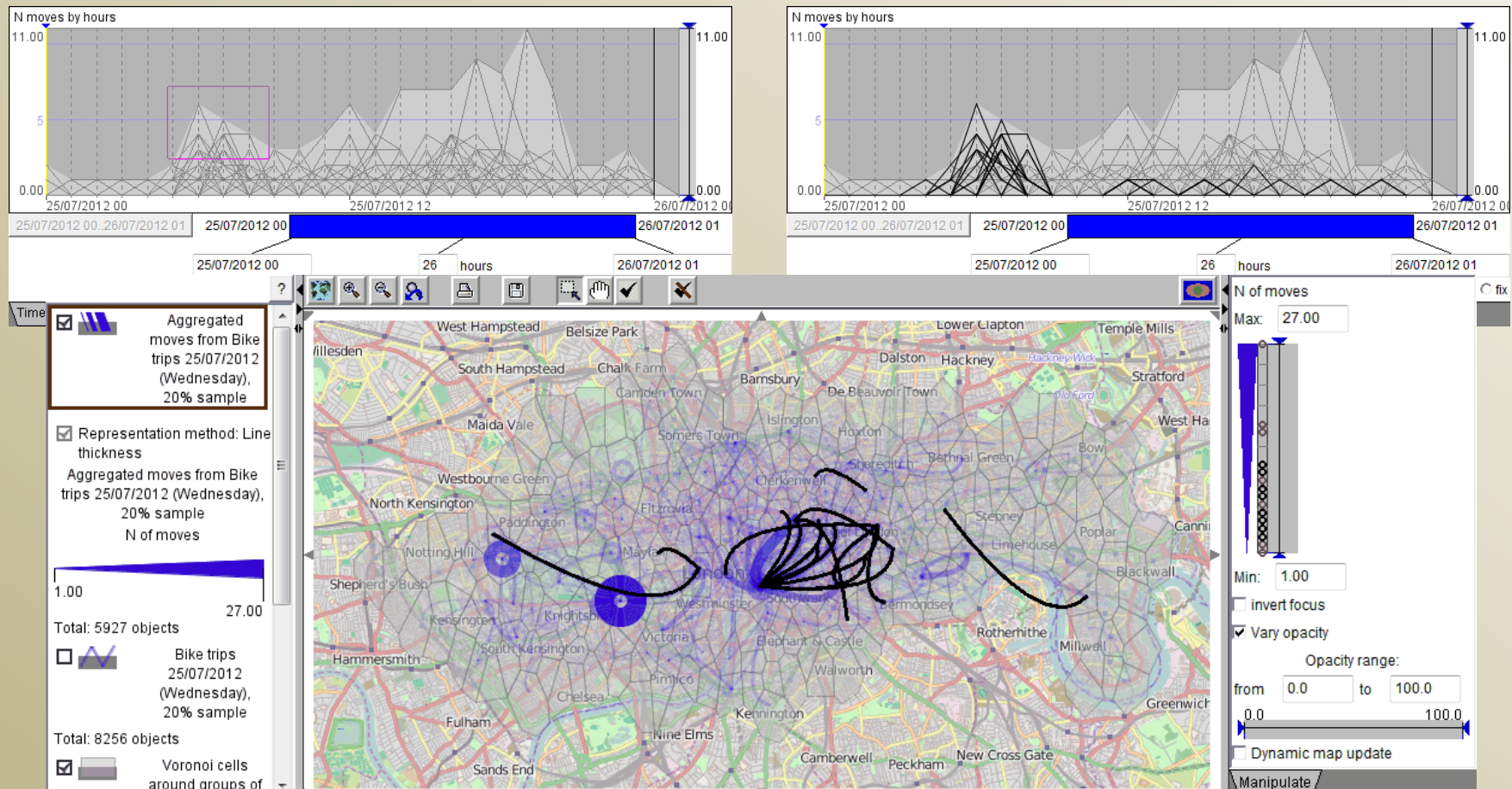
Important note: the layer “Voronoi cells ...” must be active in the map (marked by a red frame in the map legend). Click on the layer name in the legend to make it active.



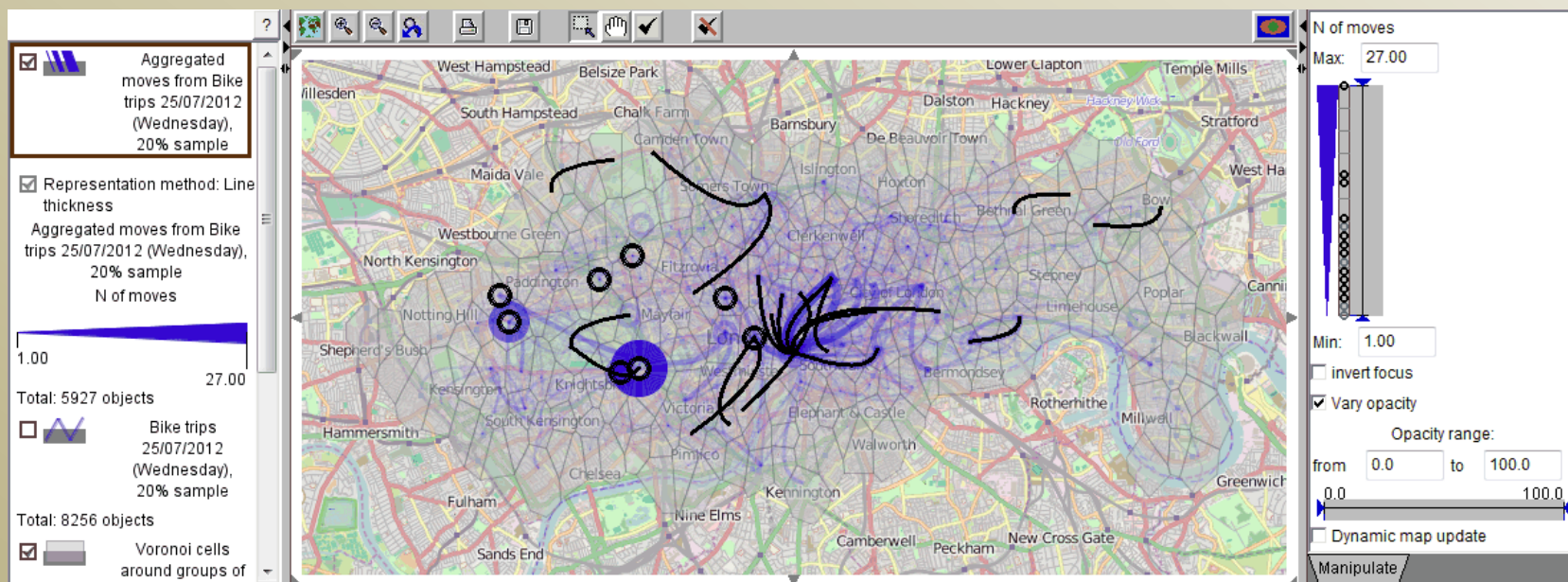
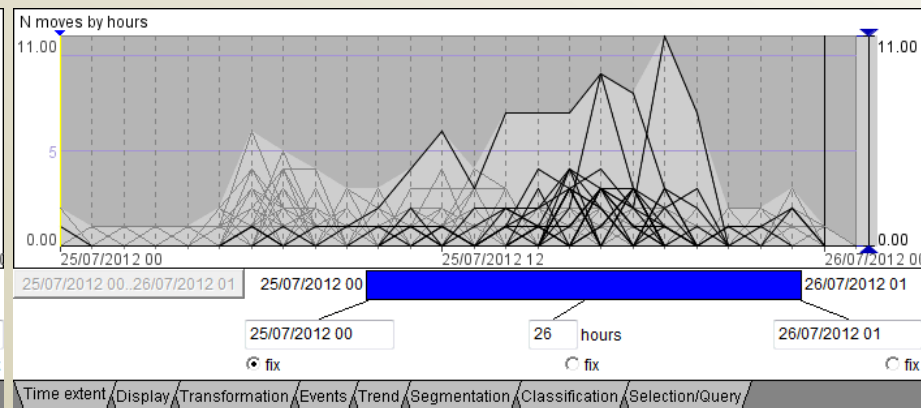
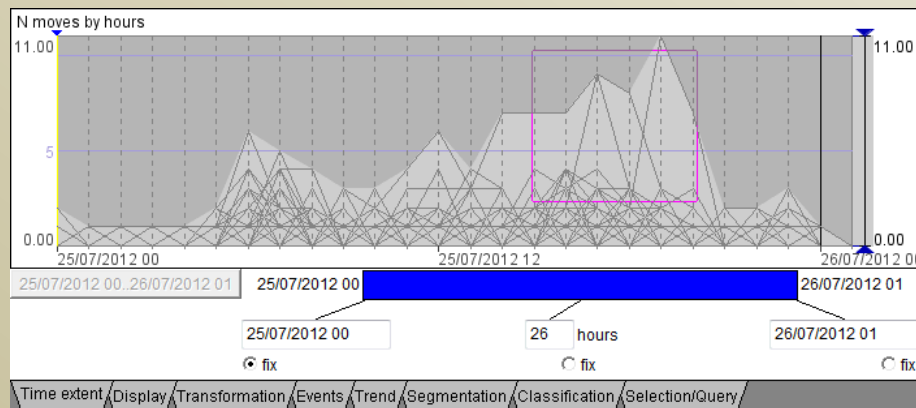




# Link-referenced time series



Create a time graph for table “Aggregated moves from bike trips ...” and attribute “N moves by hours”. Select the highest morning flows, then the highest afternoon flows. Before that, make sure that the layer “Aggregated moves ...” is active in the map.



Iris, Descartes, CommonGIS, V-Analytics 1995-2013: Bike trips 25/07/2012 (Wednesday), 20% sample



# Questions to Exercise 4

- Describe the main features of the collective movement behaviour that could be learned by exploring the aggregated data:
  - Popular trip origins and destinations in different times of the day
  - Frequencies of trips within the same areas (represented by rings)
  - Major flows in different times of the day
    - origins, destinations, directions



# Try on your own

- Bike trip data from one week (03-09/09/2012) aggregated by hourly intervals
  - Project file `aggr_hours_week_2012_09_03-09.app`
- Try partition-based clustering for place-based and link-based time series
  - The same operations as for the aggregated events in practical 3 can also be used for these data.
  - Generally, the partition-based clustering tools are used in standard ways for spatial time series of any origin.