

The Strava challenge

11.6.2018

ITT8, SAMba

Alex Noyvirt

The Challenge

- Can you model the physical activities of the Strava users? What makes people become more physically active?
- Potential additional questions : Where/when/how far/how often people choose to go running/cycling? How popular a particular running/cycling route will be (route choice model)? What interventions could increase the amount of physical activity taken by certain groups of people in certain areas?
- Even more questions: Is there correlation between physical activity and household income?

Background information

What is Strava?

- More than 1M users worldwide some of them premium
- Cyclists and runners are the typical users
- Users can "follow" each other and activities are automatically grouped together when they occur at the same time and place
- Organised challenges
- Logged data includes:
 1. route (plan view)
 2. elevation (net and unidirectional)
 3. speed (average, min/max)
 4. timing (total and moving time)
 5. power/energy



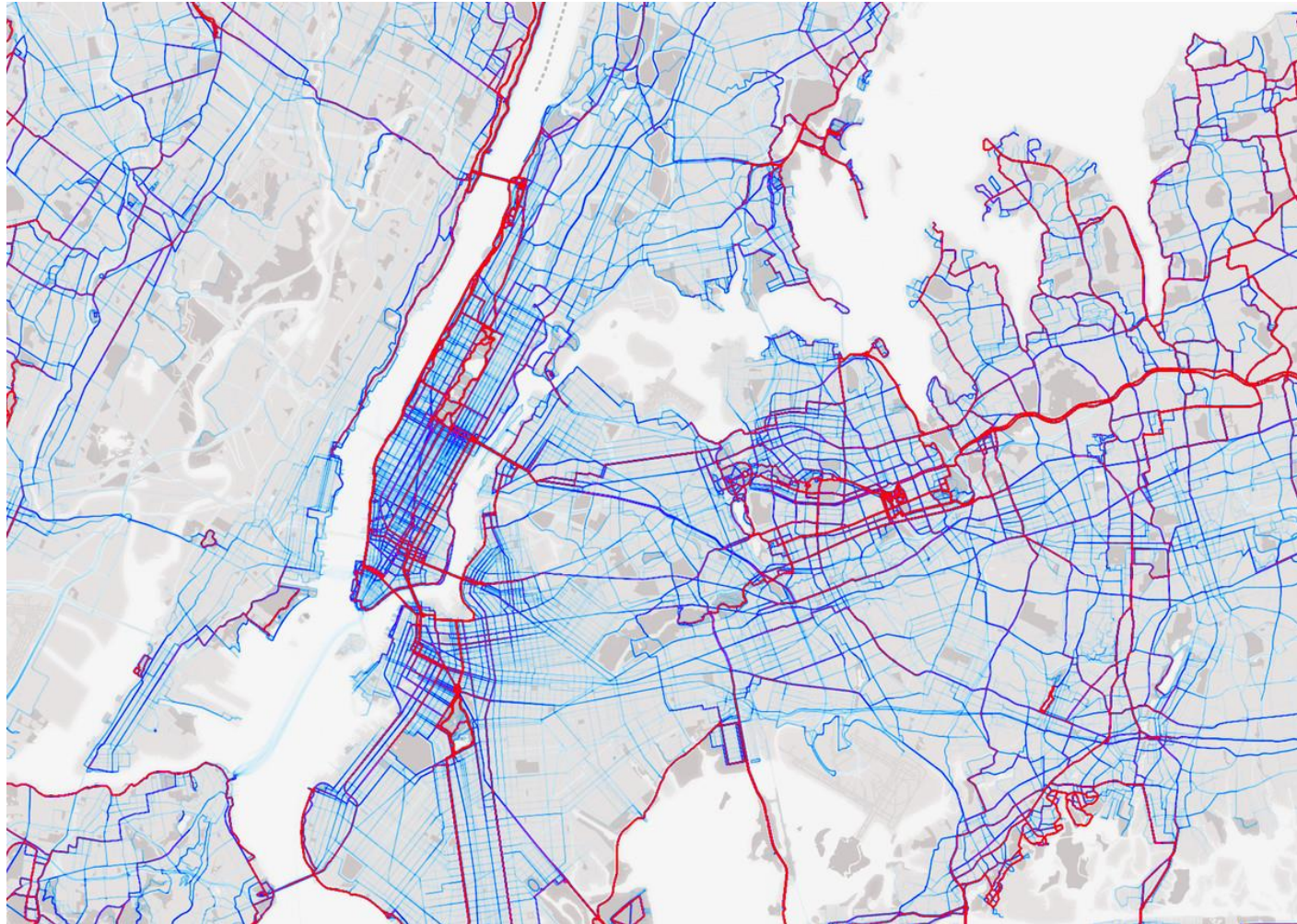
Strava Metro Core Data

- We have access to Strava Metro for Manchester 2017 – many thanks to [UBDC](#) .
- Street-level database file (sql or csv) with minute-to-minute cycling/pedestrian information on every street for the time period of the delivery.
- Intersection database file (sql or csv) with minute-to-minute cycling/pedestrian information on nodes (street intersections) for the time period of the delivery.
 - Origin/Destination file representing pairs for all trips during the time period of the delivery.

Stava Metro roll-ups

- Street/Intersections: each core dataset will have a set of summarised views that show monthly use, weekend/weekday, seasonality, hours groupings and total counts. This typically results in over 50 rolled-up views for the cycling data.

and it looks like this...



User guide with data dictionary

http://metro.strava.com/wp-content/uploads/2017/04/StravaMetro_5.01_User_Guide.pdf

Some ideas about linking with other datasets

- [OS Open Green Spaces](#) - covering a range of greenspaces in urban and rural areas including playing fields, sports' facilities, play areas and allotment.
- Open data greater Manchester - <http://open.manchester.gov.uk/>
- Smart cities IoT data - <https://portal.bt-hypercat.com/index.php/data-catalogue/>
- Is the weather/temperature a factor in physical activity? - Openweathermap.org
- Open government data - data.gov.uk (tip: search for Manchester)
- Open police data - data.police.uk/data/
- road surface quality – image processing using the [Google StreetView API](#)

Currently available dataset

- Coverage : Manchester 2017
- Format: esi.shp, CSV
- Size: ~700MB
- Any participant needs to sign an end-user agreement before receiving the data:
 - Online: <https://ubdc.gla.ac.uk/itt8-bath> user: **itt8-user** password: **ittbD@t@**
 - Or paper based forms (ask me after the presentation)

The Challenge

- Can you model the physical activities of the Strava users? What makes people become more physically active?
- Potential additional questions : Where/when/how far/how often people choose to go running/cycling? How popular a particular running/cycling route will be (route choice model)? What interventions could increase the amount of physical activity taken by certain groups of people in certain areas?
- Even more questions: Is there correlation between physical activity and household income?

Questions?

Some links :

Bike Life Greater Manchester 2017 report - <https://www.tfgm.com/cycling/bike-life-2017>

Visualisation of London cycling heatmap: <https://www.youtube.com/watch?v=QZ5DuQTUPqk>