
Adding real-world GIS and census data to agent-based modelling for social scientists

Dr. J. Kasmire
UK Data Service



30 January 2020

UK Data Service



ABM for social scientists – webinar series

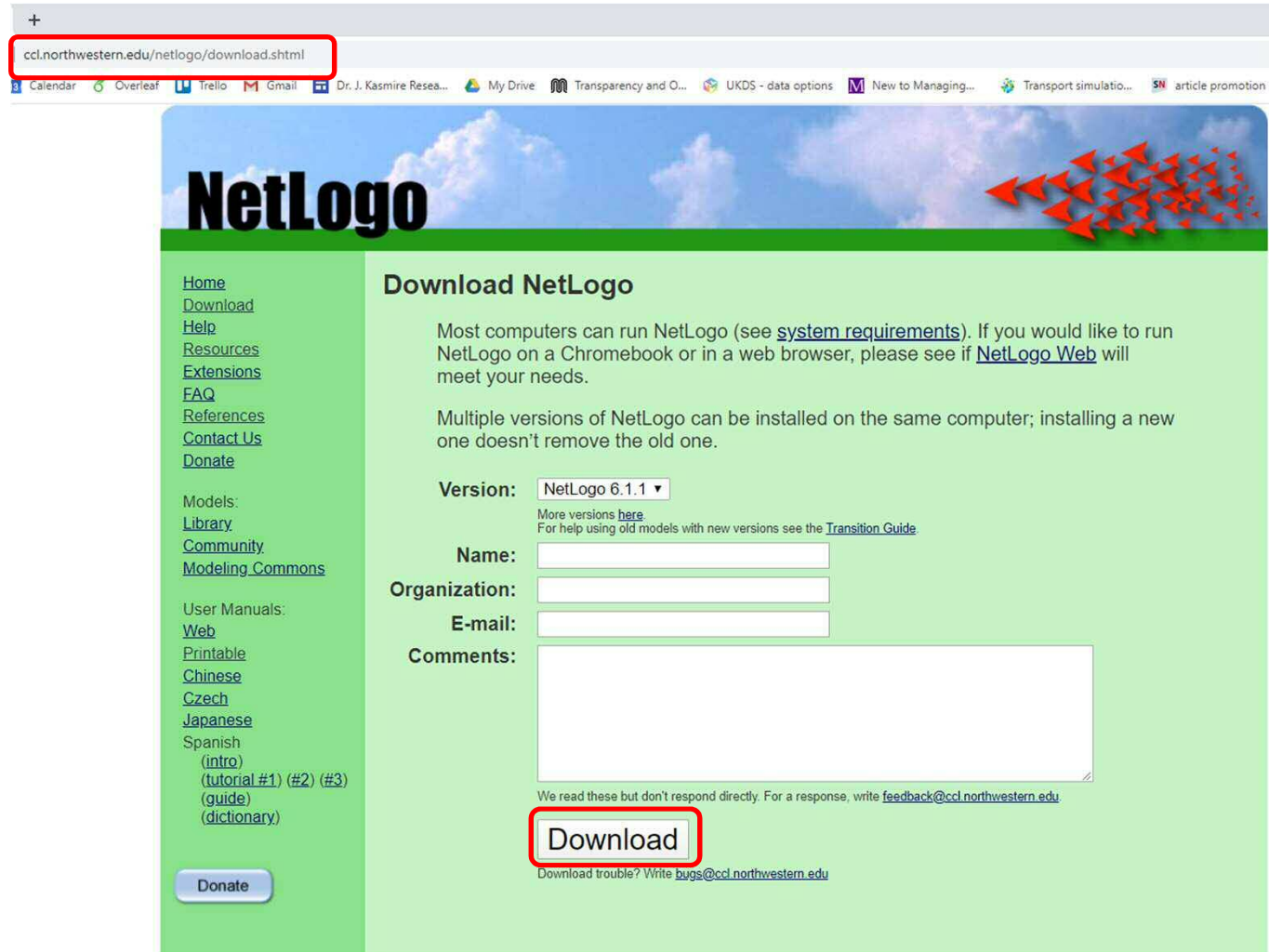
- ABM: An Intro
 - – Jan 16, 2020, recording available
- **ABM: Adding Data**
 - – Jan 30, 2020
- ABM: Experiments and Output
 - – Feb 13, 2020



Table of Contents

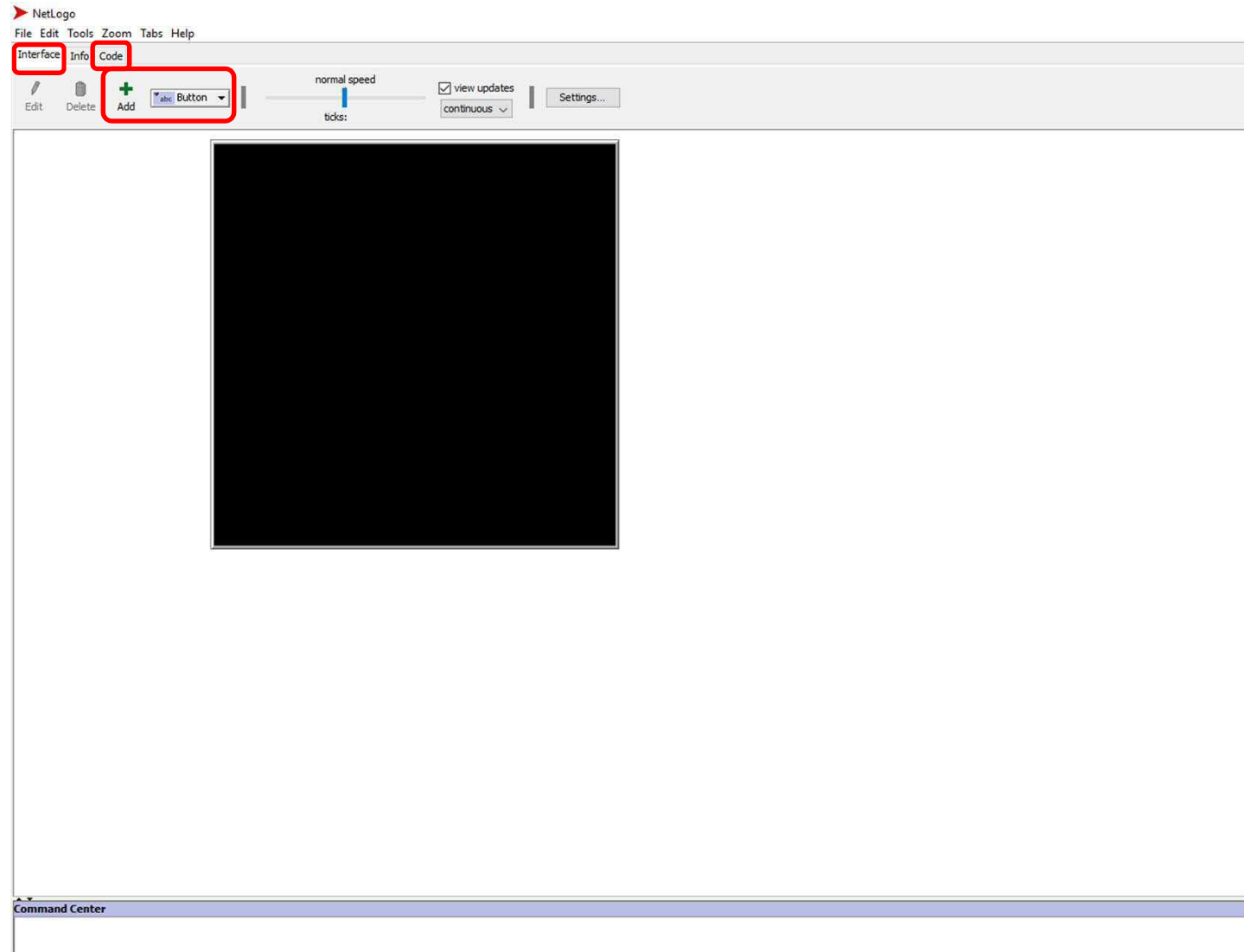
- Opening models in Netlogo – Model Library, Code Examples and .nlogo files
- Set up and run a random "Tram commute model"
- Download and prepare real-world GIS and census data files
- Import Greater Manchester shapefiles into code and load them into the model
- What else could I do with this?
- Summary
- Questions

Download NetLogo

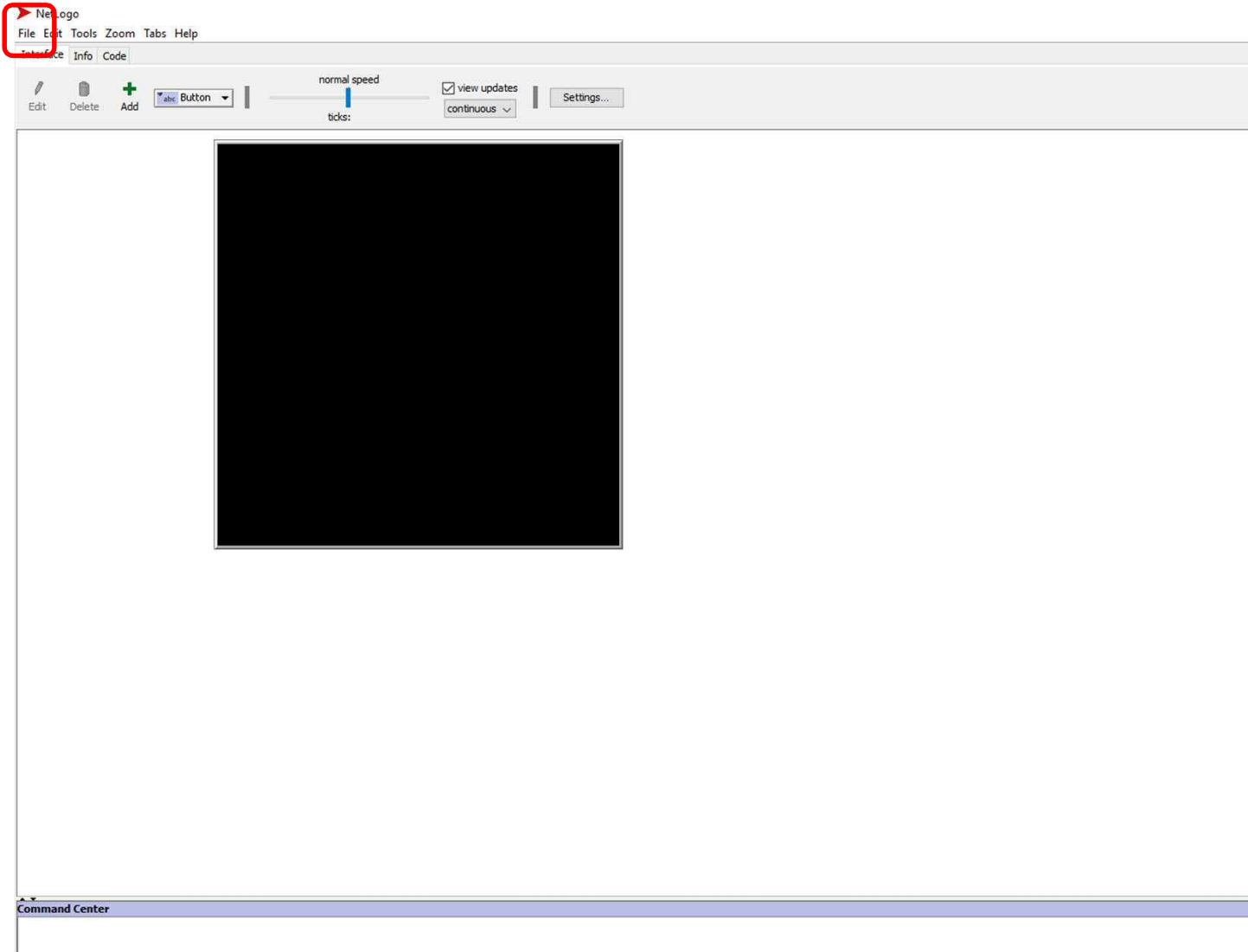


The screenshot shows a web browser window with the address bar containing `ccl.northwestern.edu/netlogo/download.shtml`, which is highlighted with a red box. The browser's tab bar shows several open tabs, including 'Calendar', 'Overleaf', 'Trello', 'Gmail', and others. The main content area of the browser displays the NetLogo website. The header features the 'NetLogo' logo and a decorative graphic of red arrows. A left sidebar contains a list of links: Home, Download, Help, Resources, Extensions, FAQ, References, Contact Us, and Donate. Below these are sections for 'Models' (Library, Community, Modeling Commons) and 'User Manuals' (Web, Printable, Chinese, Czech, Japanese, Spanish intro, tutorial #1, #2, #3, guide, dictionary). A 'Donate' button is at the bottom of the sidebar. The main content area is titled 'Download NetLogo' and contains text explaining that most computers can run NetLogo, with links to system requirements and NetLogo Web. It also states that multiple versions can be installed on the same computer. Below this text are form fields for 'Version' (set to NetLogo 6.1.1), 'Name', 'Organization', and 'E-mail', followed by a large 'Comments' text area. A 'Download' button is highlighted with a red box. At the bottom of the form, there is a note: 'We read these but don't respond directly. For a response, write feedback@ccl.northwestern.edu.' and a link for download trouble: 'Download trouble? Write bugs@ccl.northwestern.edu'.

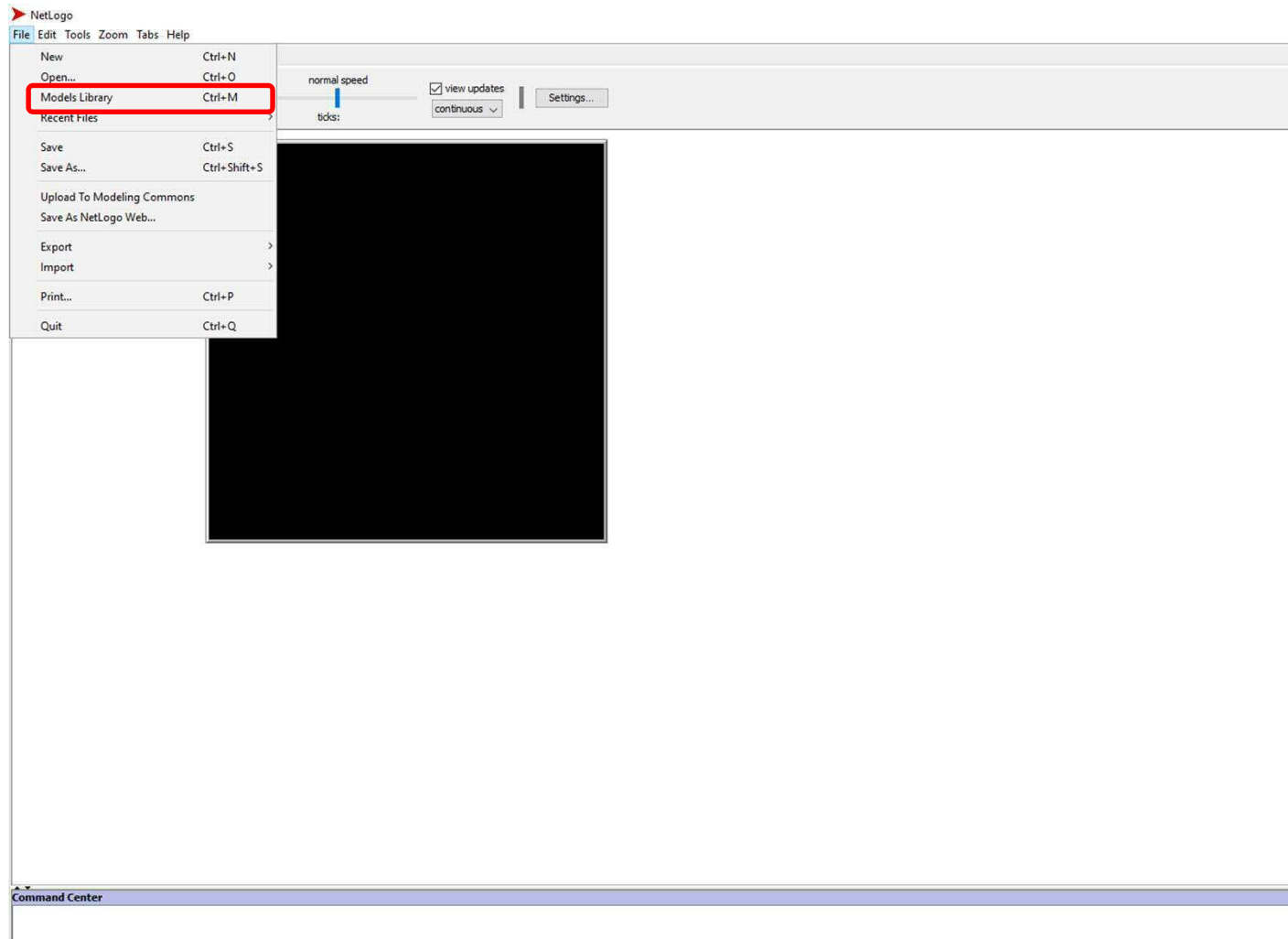
NetLogo opening view



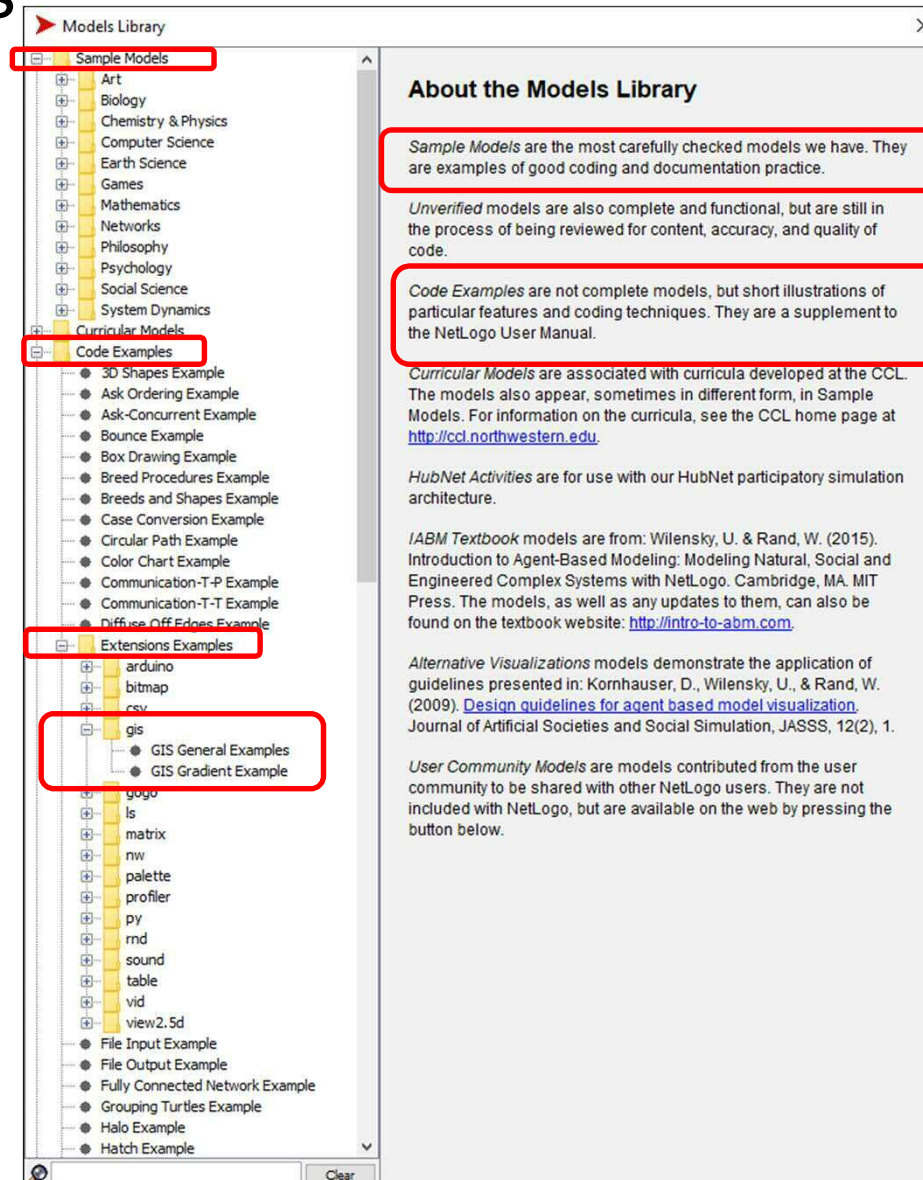
Open a NetLogo model



File -> Models Library



Models Library: demonstrate concepts, code, extensions



Models Library: so many option

The screenshot shows the NetLogo Models Library window. On the left, a tree view lists various model categories. On the right, a text panel titled 'About the Models Library' provides information about different model types. Red boxes and arrows highlight specific features:

- Sample Models:** A list of categories including Art, Biology, Chemistry & Physics, Computer Science, Earth Science, Games, Mathematics, Networks, Philosophy, Psychology, Social Science, and System Dynamics. Sub-categories under System Dynamics include Exponential Growth, Logistic Growth, Wolf Sheep Predation (Docked Hybrid), and Wolf Sheep Predation (System Dynamics).
- Curricular Models:** A category containing models like BEAGLE Evolution, Connected Chemistry, EACH, epiDEM, GasLab, GenEvo, Lattice Land, MaterialSim, Mind the Gap, ModelSim, NIELS, PNoM, ProbLab, and Urban Suite.
- Code Examples:** A category containing HubNet Activities, IABM Textbook, and Alternative Visualizations.
- About the Models Library:** A text panel explaining the library's structure and the types of models available.

About the Models Library

Sample Models are the most carefully checked models we have. They are examples of good coding and documentation practice.

Unverified models are also complete and functional, but are still in the process of being reviewed for content, accuracy, and quality of code.

Code Examples are not complete models, but short illustrations of particular features and coding techniques. They are a supplement to the NetLogo User Manual.

Curricular Models are associated with curricula developed at the CCL. The models also appear, sometimes in different form, in Sample Models. For information on the curricula, see the CCL home page at <http://ccl.northwestern.edu>.

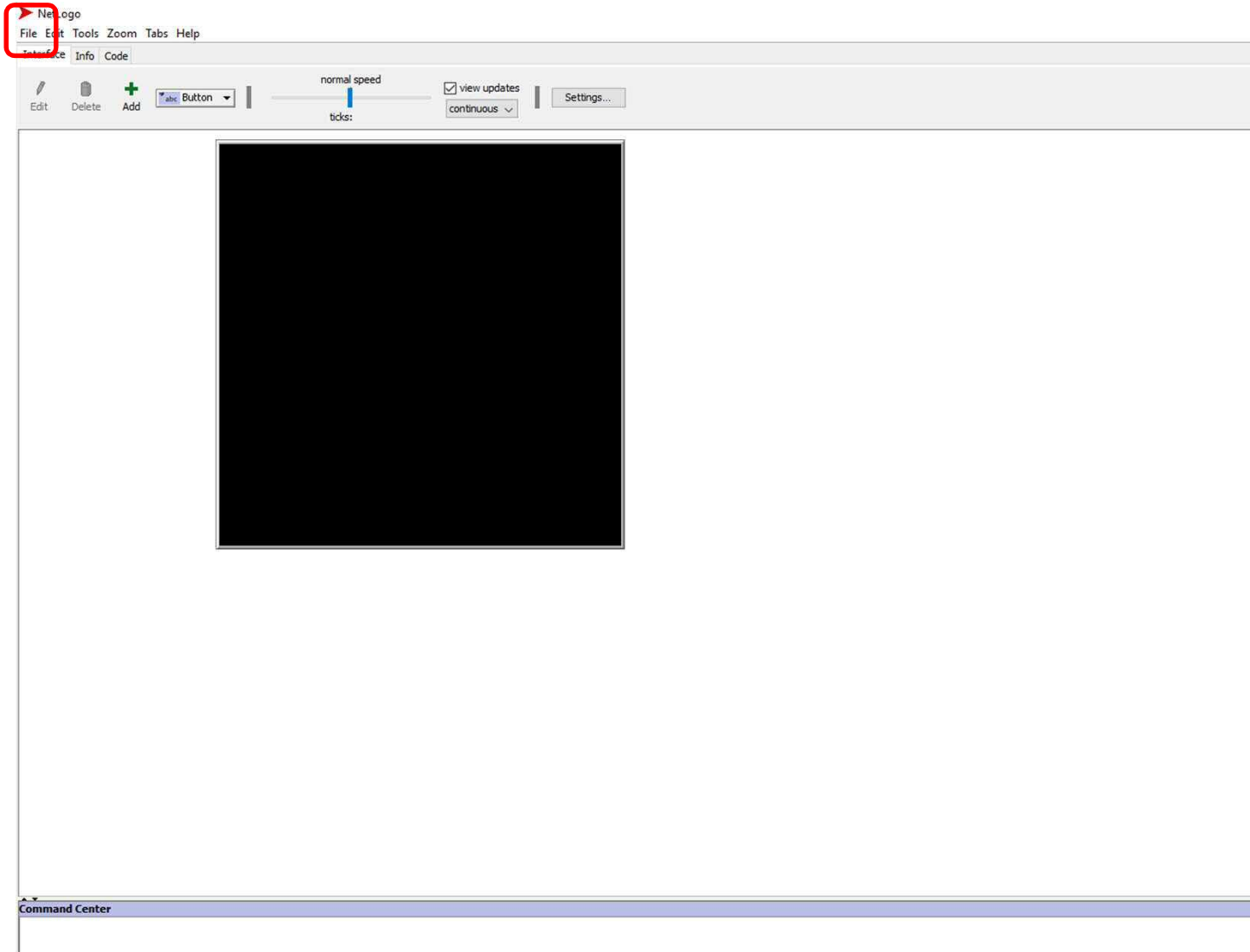
HubNet Activities are for use with our HubNet participatory simulation architecture.

IABM Textbook models are from: Wilensky, U. & Rand, W. (2015). Introduction to Agent-Based Modeling: Modeling Natural, Social and Engineered Complex Systems with NetLogo. Cambridge, MA. MIT Press. The models, as well as any updates to them, can also be found on the textbook website: <http://intro-to-abm.com>.

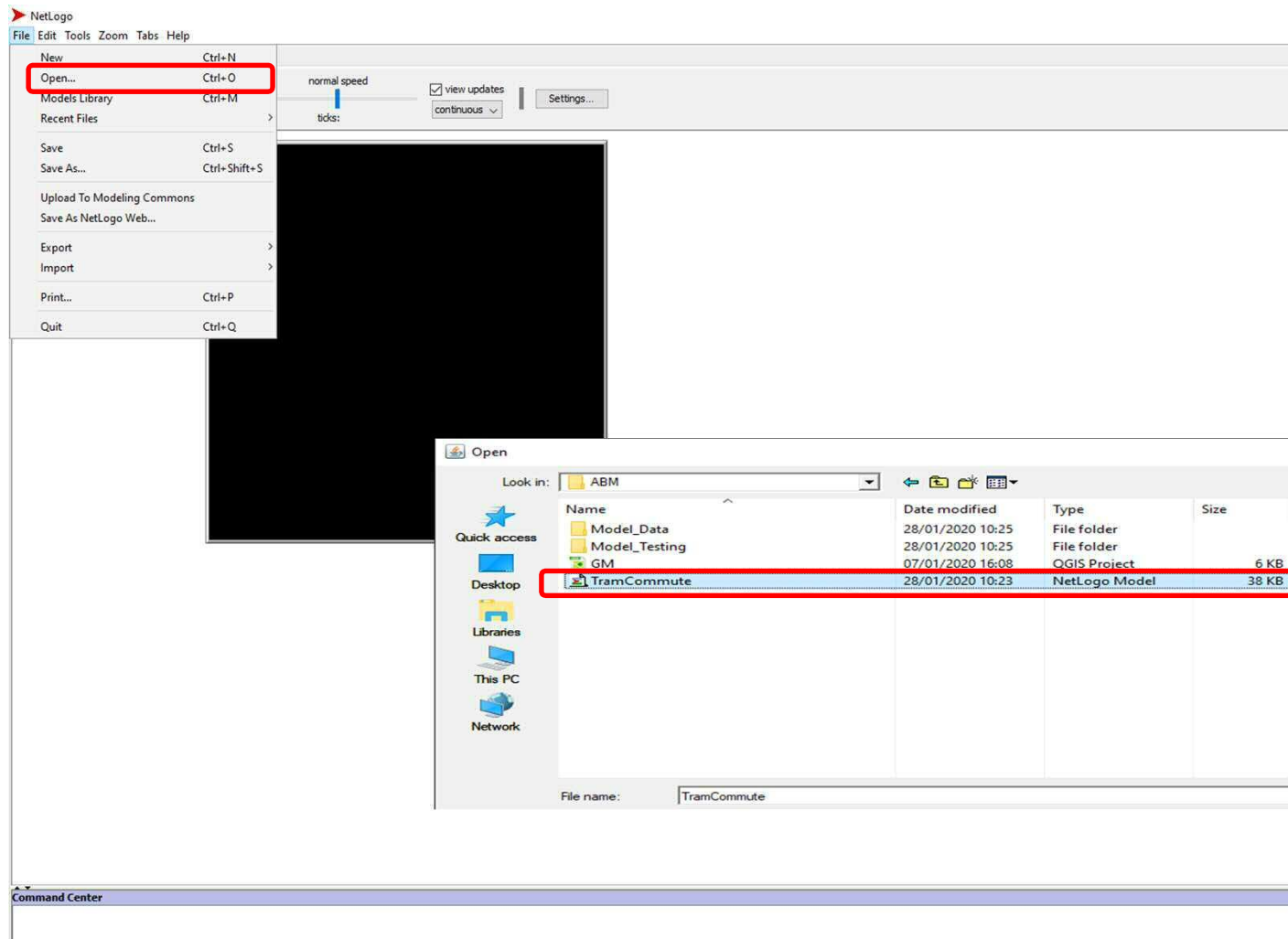
Alternative Visualizations models demonstrate the application of guidelines presented in: Kornhauser, D., Wilensky, U., & Rand, W. (2009). [Design guidelines for agent based model visualization](#). Journal of Artificial Societies and Social Simulation, JASSS, 12(2), 1.

User Community Models are models contributed from the user community to be shared with other NetLogo users. They are not included with NetLogo but are available on the web by pressing the

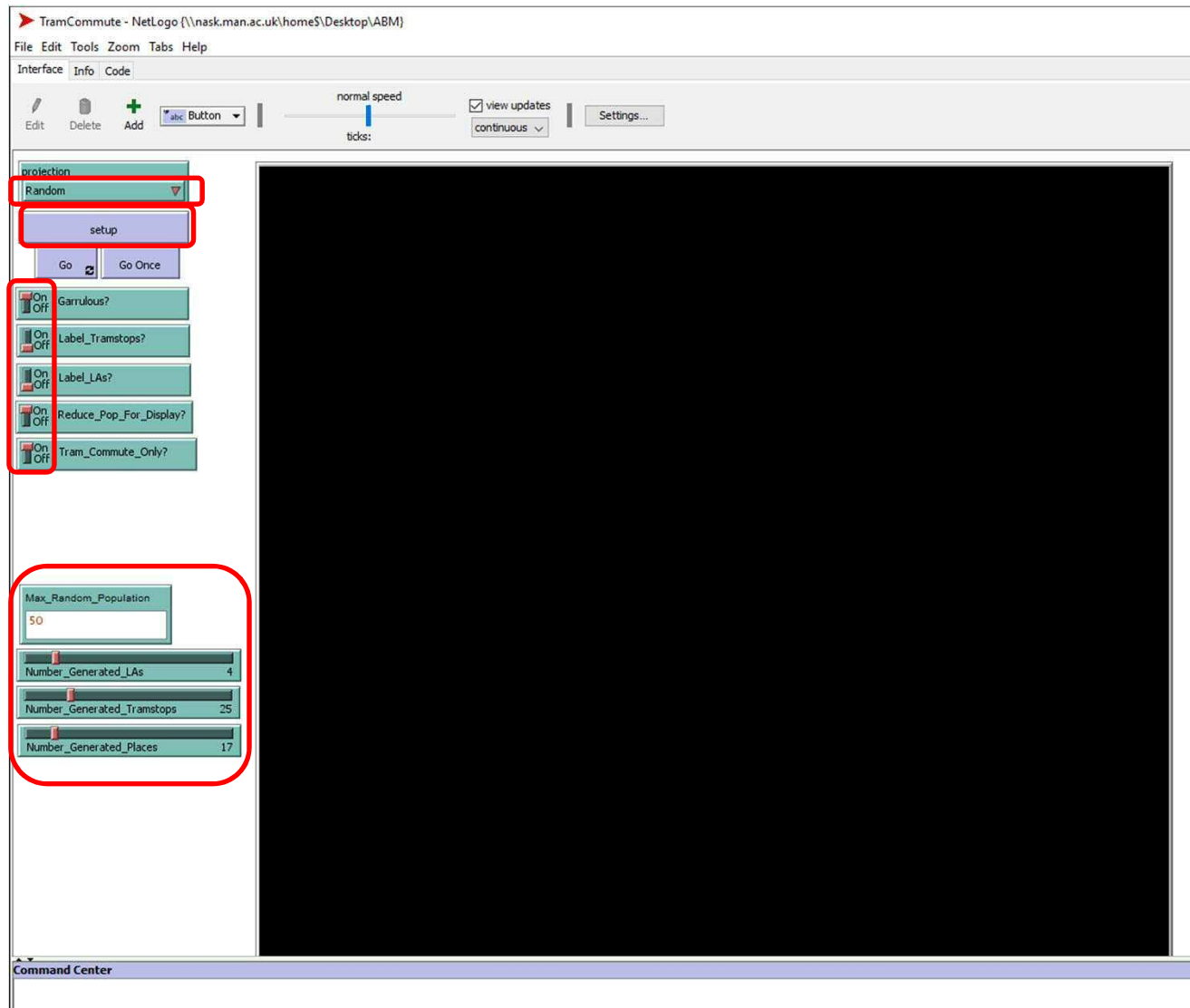
Open an existing NetLogo model



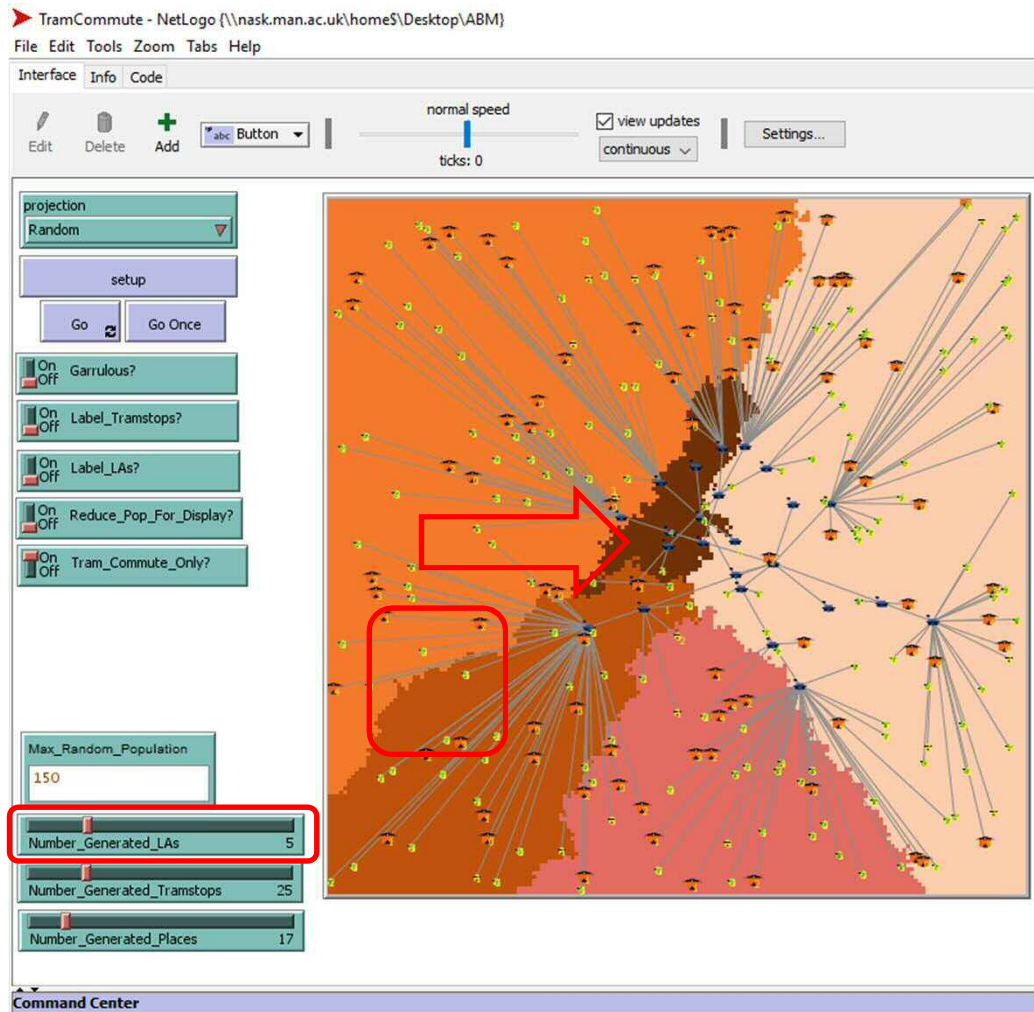
File -> Open -> Navigate to TramCommute.nlogo



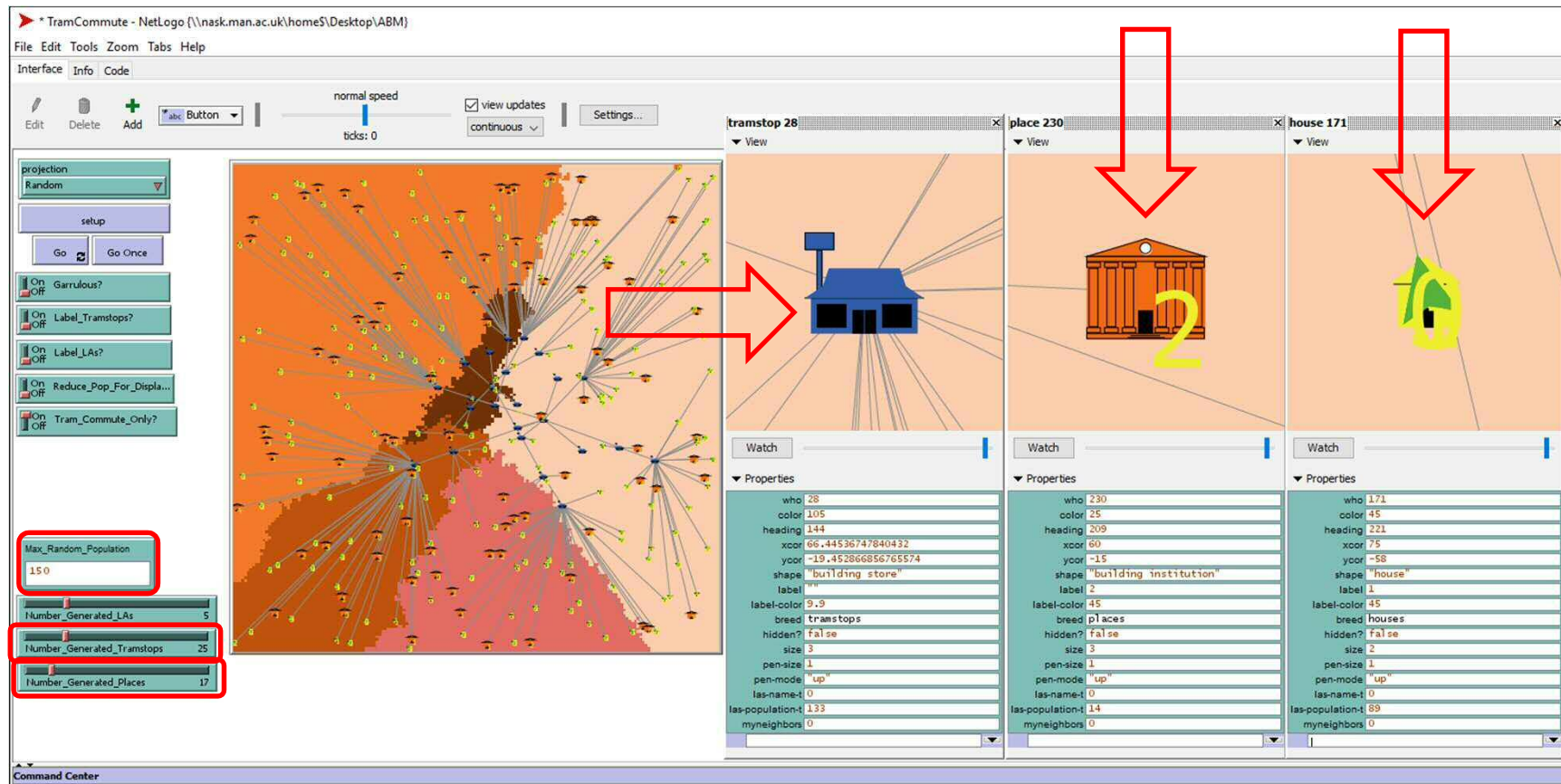
Setup a random "Tram Commute Model"



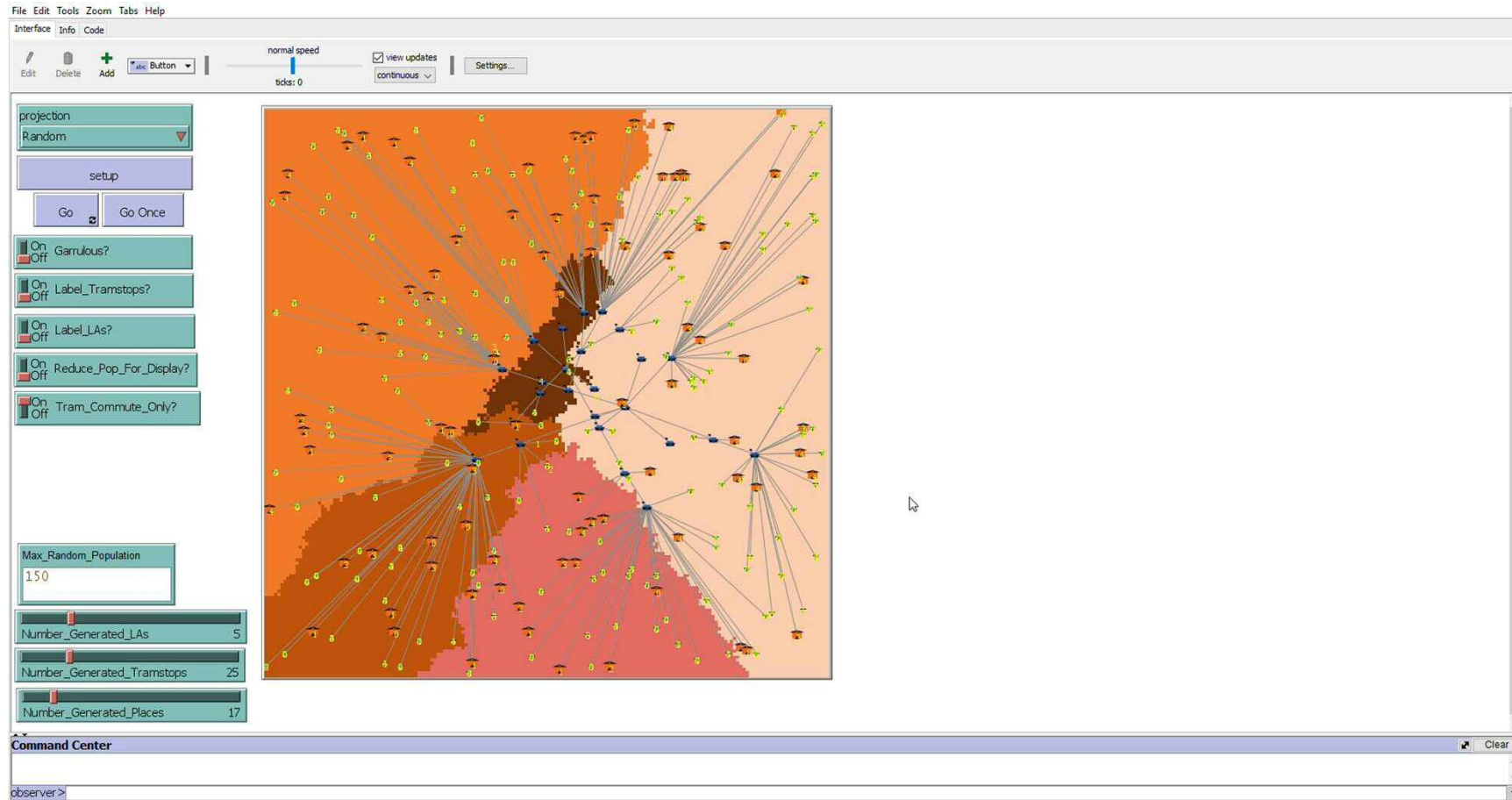
A random "Tram Commute Model" when setup



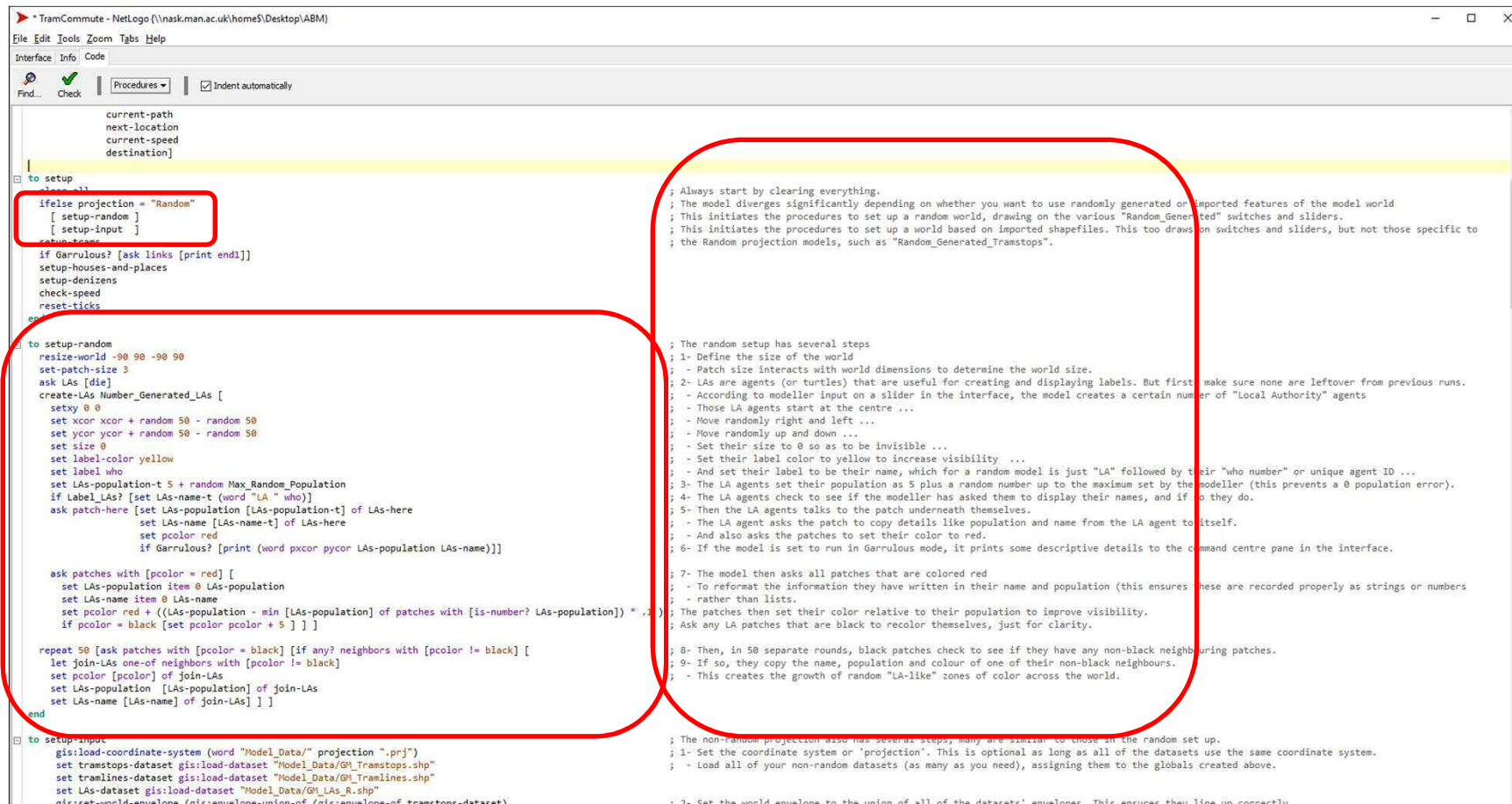
A random "Tram Commute Model" when setup



Running the random "Tram Commute Model"



Create Random Setup



```

current-path
next-location
current-speed
destination]

to setup
  clear-all
  ifelse projection = "Random"
  [ setup-random ]
  [ setup-input ]
  setup-tram
  if Garrulous? [ask links [print end]]
  setup-houses-and-places
  setup-denizens
  check-speed
  reset-ticks
end

to setup-random
  resize-world -90 90 -90 90
  set-patch-size 3
  ask LAs [die]
  create-LAs Number_Generated_LAs [
    setxy 0 0
    set xcor xcor + random 50 - random 50
    set ycor ycor + random 50 - random 50
    set size 0
    set label-color yellow
    set label who
    set LAs-population-t 5 + random Max_Random_Population
    if label_LAs? [set LAs-name-t (word "LA " who)]
    ask patch-here [set LAs-population [LAs-population-t] of LAs-here
      set LAs-name [LAs-name-t] of LAs-here
      set pcolor red
      if Garrulous? [print (word pcolor pycor LAs-population LAs-name)]]

    ask patches with [pcolor = red] [
      set LAs-population item 0 LAs-population
      set LAs-name item 0 LAs-name
      set pcolor red + ((LAs-population - min [LAs-population] of patches with [is-number? LAs-population]) * .5)
      if pcolor = black [set pcolor pcolor + 5 ] ]

  repeat 50 [ask patches with [pcolor = black] [if any? neighbors with [pcolor != black] [
    let join-LAs one-of neighbors with [pcolor != black]
    set pcolor [pcolor] of join-LAs
    set LAs-population [LAs-population] of join-LAs
    set LAs-name [LAs-name] of join-LAs ] ]
end

to setup-input
  gis:load-coordinate-system (word "Model_Data/" projection ".prj")
  set tramstops-dataset gis:load-dataset "Model_Data/GM_Tramstops.shp"
  set tramlines-dataset gis:load-dataset "Model_Data/GM_Tramlines.shp"
  set LAs-dataset gis:load-dataset "Model_Data/GM_LAs_R.shp"
  create-world-environment (intersection-of (intersection-of tramstops-dataset)

```

Always start by clearing everything.
 The model diverges significantly depending on whether you want to use randomly generated or imported features of the model world
 This initiates the procedures to set up a random world, drawing on the various "Random_Generated" switches and sliders.
 This initiates the procedures to set up a world based on imported shapefiles. This too draws on switches and sliders, but not those specific to the Random projection models, such as "Random_Generated_Tramstops".

The random setup has several steps
 1- Define the size of the world
 2- Patch size interacts with world dimensions to determine the world size.
 3- LAs are agents (or turtles) that are useful for creating and displaying labels. But first make sure none are leftover from previous runs.
 4- According to modeller input on a slider in the interface, the model creates a certain number of "Local Authority" agents
 5- Those LA agents start at the centre ...
 6- Move randomly right and left ...
 7- Move randomly up and down ...
 8- Set their size to 0 so as to be invisible ...
 9- Set their label color to yellow to increase visibility ...
 10- And set their label to be their name, which for a random model is just "LA" followed by their "who number" or unique agent ID ...
 11- The LA agents set their population as 5 plus a random number up to the maximum set by the modeller (this prevents a 0 population error).
 12- The LA agents check to see if the modeller has asked them to display their names, and if so they do.
 13- Then the LA agents talk to the patch underneath themselves.
 14- The LA agent asks the patch to copy details like population and name from the LA agent to itself.
 15- And also asks the patches to set their color to red.
 16- If the model is set to run in Garrulous mode, it prints some descriptive details to the command centre pane in the interface.

7- The model then asks all patches that are colored red
 8- To reformat the information they have written in their name and population (this ensures these are recorded properly as strings or numbers rather than lists.
 9- The patches then set their color relative to their population to improve visibility.
 10- Ask any LA patches that are black to recolor themselves, just for clarity.

8- Then, in 50 separate rounds, black patches check to see if they have any non-black neighbouring patches.
 9- If so, they copy the name, population and colour of one of their non-black neighbours.
 10- This creates the growth of random "LA-like" zones of color across the world.

The non-Random projection also has several steps, many are similar to those in the random set up.
 1- Set the coordinate system or 'projection'. This is optional as long as all of the datasets use the same coordinate system.
 2- Load all of your non-random datasets (as many as you need), assigning them to the globals created above.
 3- Set the world envelope to the union of all of the datasets' envelopes. This ensures they line up correctly

Download Boundary Data – Local Authority shapes

<https://tinyurl.com/tek2xhs>

The screenshot shows a web browser window with the URL `borders.ukdataservice.ac.uk/bookmark.html?userDirectory=3AE48D0D5F1F354211157668198666316%2F15766819869431671623156483231789&service=BOUNDARIES`. The website is the UK Data Service Census Support portal. On the left, under 'QUICK ACCESS TO', the 'Boundary Data Selector' is highlighted. The main content area shows search results for 'InFuse Local Authorities, 2011'. The 'Data Format' is 'ESRI Shape File' and the 'Archive Method' is 'zip'. Below this, a section titled 'FILES EXTRACTED' provides instructions on downloading the files. A table lists the extracted file: 'BoundaryData.zip', which is a 'Zip file containing your data' and is 601 KB in size.

UK Data Service
Census Support

QUICK ACCESS TO

- Easy Download
- Boundary Data Selector**
- Thematic Mapper
- Postcode Directory
- Postcode Data Selector

Data available News Accessibility Statement Privacy Statement

Home / UK Data Service Data Download Bookmark facility

Search Summary: InFuse Local Authorities, 2011

Data Format: ESRI Shape File

Archive Method: zip

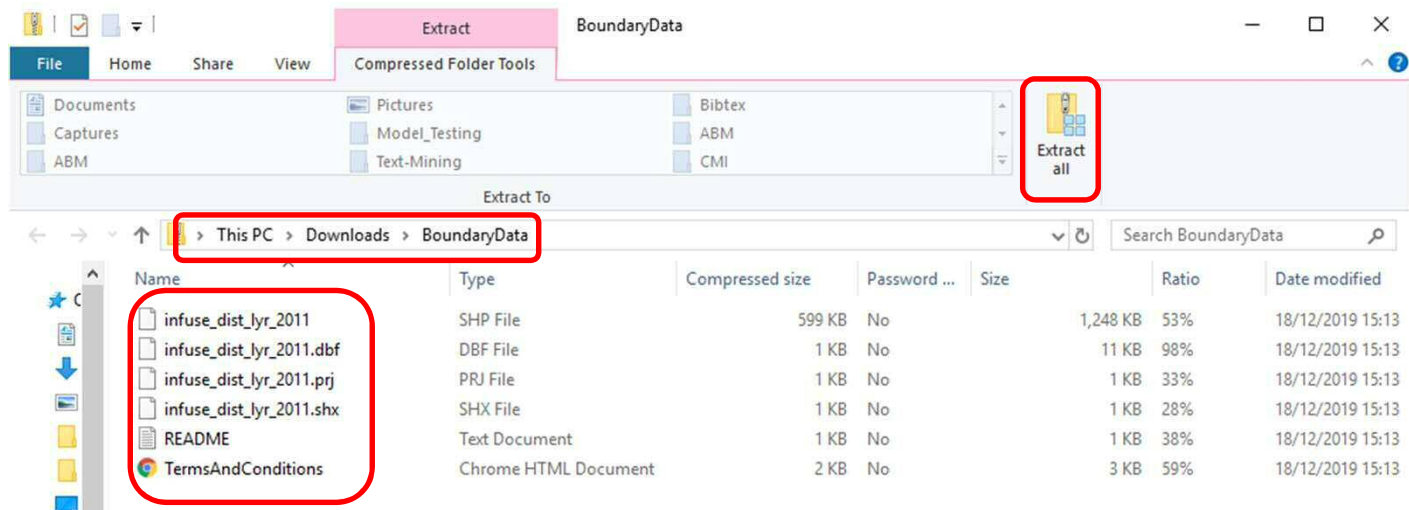
FILES EXTRACTED

Here are the files containing the data you have requested from Census Support, together with their type and size. The files in this directory will be automatically deleted within seven days of being generated. You can download an individual file using your web browser by clicking on its name in the table. If you use Firefox, the best way to do this is to click with the right mouse button and go to **Save Link As**. With Internet Explorer you should go to **Save Target As** after clicking with the right mouse button.

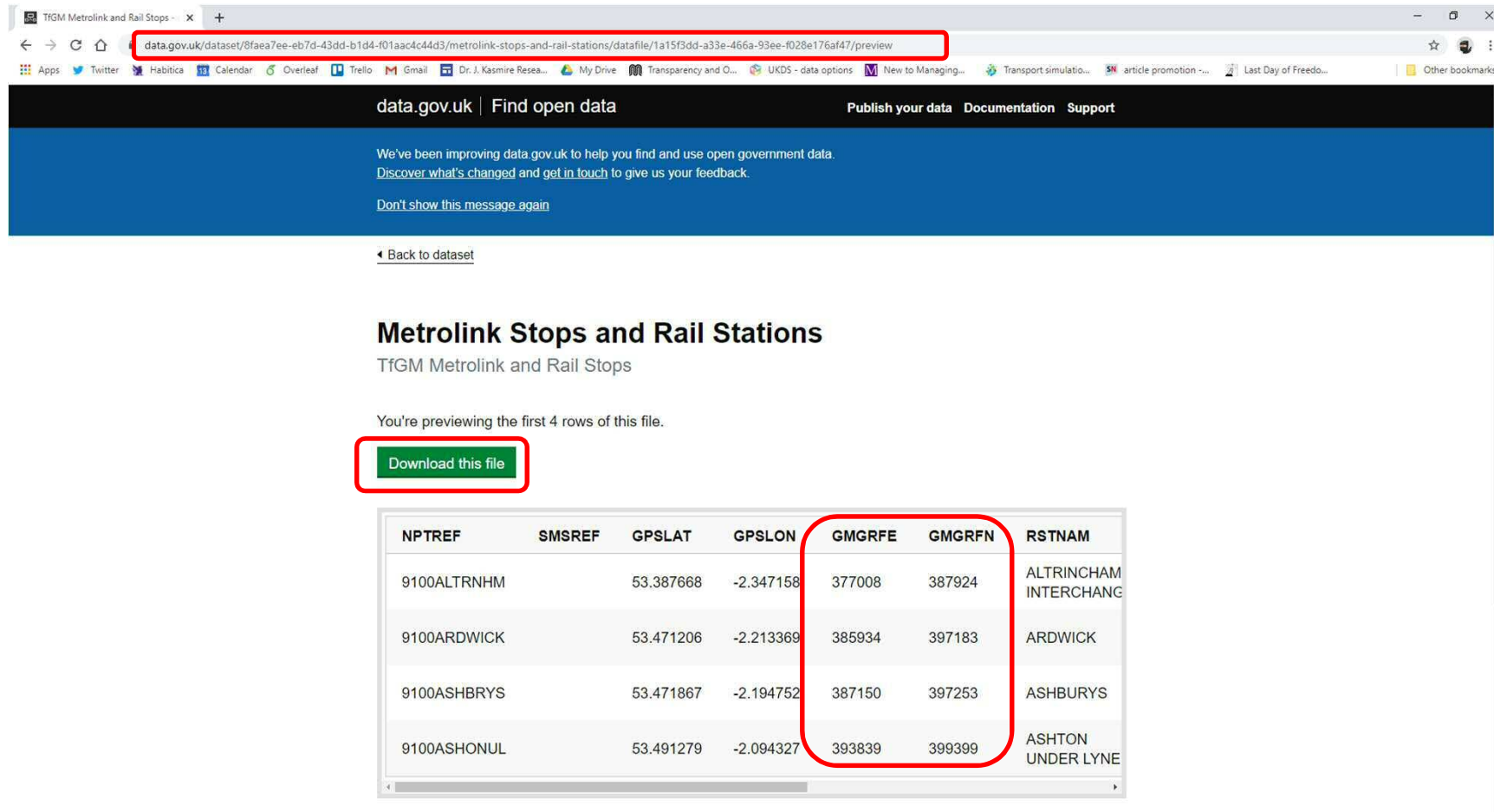
File Name	Description	Size
BoundaryData.zip	Zip file containing your data	601 KB

© Copyright 2018 UK Data Service. All rights reserved. Powered by UKBORDERS from EDINA @ University of Edinburgh

Extract LA Boundary Data



Download GIS Location Data – Metrolink stops



data.gov.uk | Find open data | Publish your data | Documentation | Support

We've been improving data.gov.uk to help you find and use open government data.
[Discover what's changed](#) and [get in touch](#) to give us your feedback.
[Don't show this message again](#)

[Back to dataset](#)

Metrolink Stops and Rail Stations

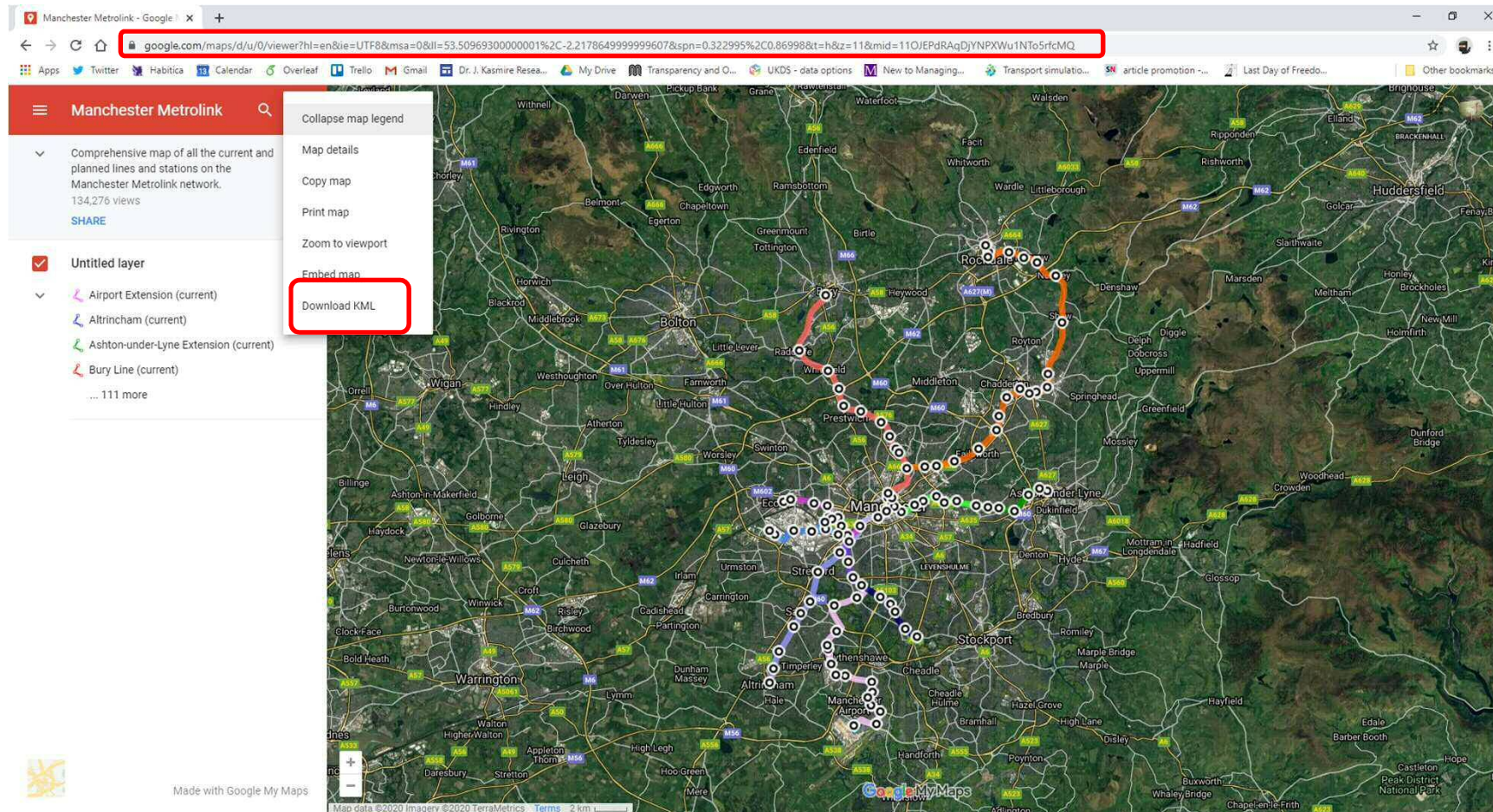
TfGM Metrolink and Rail Stops

You're previewing the first 4 rows of this file.

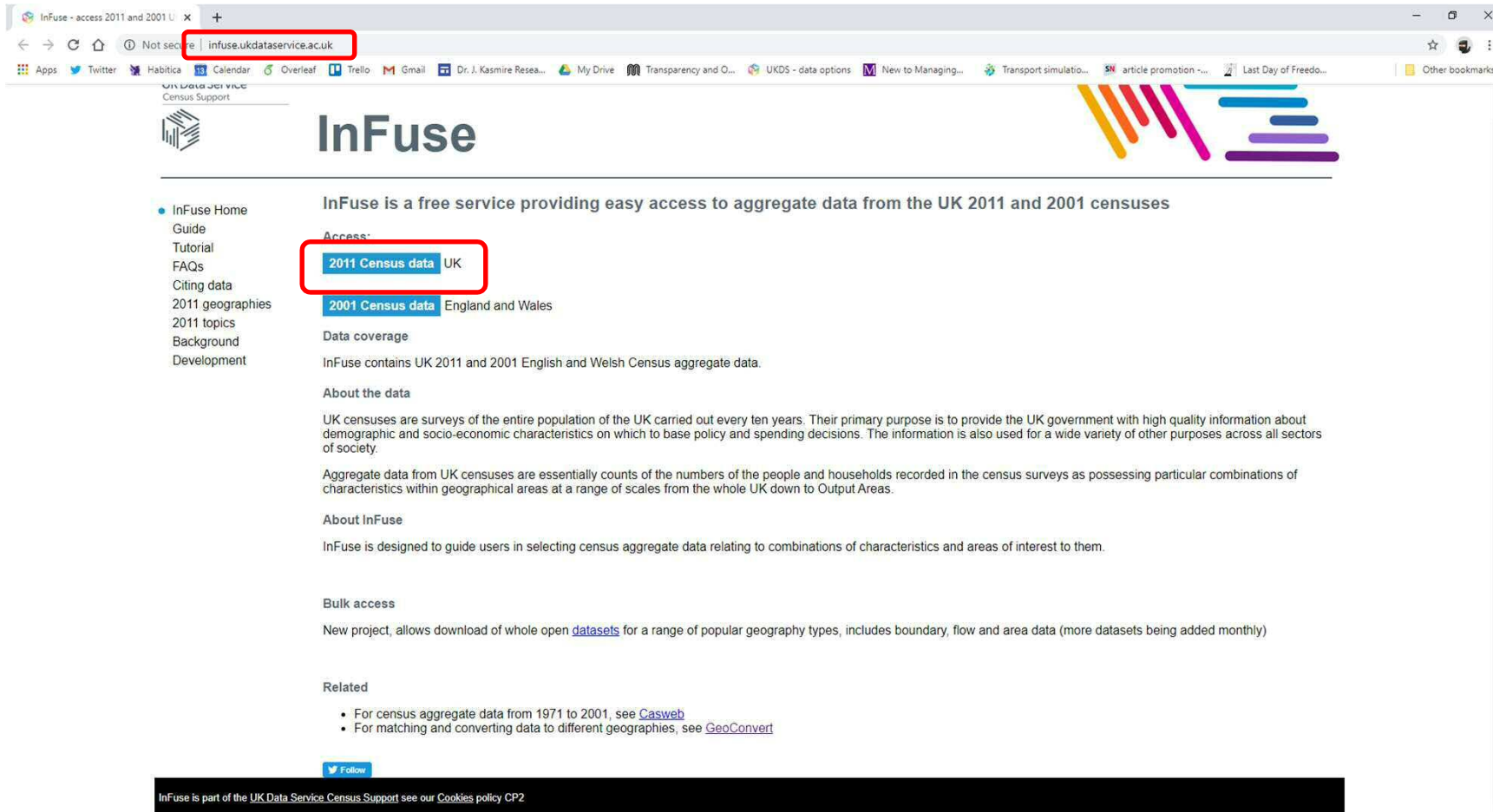
[Download this file](#)

NPTRF	SMSREF	GPSLAT	GPSLON	GMGRFE	GMGRFN	RSTNAM
9100ALTRNHM		53.387668	-2.347158	377008	387924	ALTRINCHAM INTERCHANG
9100ARDWICK		53.471206	-2.213369	385934	397183	ARDWICK
9100ASHBRY		53.471867	-2.194752	387150	397253	ASHBURY
9100ASHONUL		53.491279	-2.094327	393839	399399	ASHTON UNDER LYNE

Download GIS Network Data – Metrolink lines



Download Census Data - population



The screenshot shows a web browser window with the URL infuse.ukdataservice.ac.uk in the address bar. The page title is "InFuse". The main heading reads: "InFuse is a free service providing easy access to aggregate data from the UK 2011 and 2001 censuses". Under the "Access:" section, the link "2011 Census data" is highlighted with a red box, followed by "UK". Below it, "2001 Census data" is followed by "England and Wales". The left sidebar contains a menu with items: InFuse Home, Guide, Tutorial, FAQs, Citing data, 2011 geographies, 2011 topics, Background, and Development. The main content area includes sections for "Data coverage", "About the data", "About InFuse", "Bulk access", and "Related" links. The footer states: "InFuse is part of the UK Data Service Census Support see our Cookies policy CP2".

InFuse - access 2011 and 2001 U x +

Not secure | infuse.ukdataservice.ac.uk

Apps Twitter Habica 17 Calendar Overleaf Trello Gmail Dr. J. Kasmire Resea... My Drive Transparency and O... UKDS - data options M New to Managing... Transport simulation... article promotion ~... Last Day of Freed... Other bookmarks

UK Data Service
Census Support

InFuse

InFuse is a free service providing easy access to aggregate data from the UK 2011 and 2001 censuses

Access:

2011 Census data UK

2001 Census data England and Wales

Data coverage

InFuse contains UK 2011 and 2001 English and Welsh Census aggregate data.

About the data

UK censuses are surveys of the entire population of the UK carried out every ten years. Their primary purpose is to provide the UK government with high quality information about demographic and socio-economic characteristics on which to base policy and spending decisions. The information is also used for a wide variety of other purposes across all sectors of society.

Aggregate data from UK censuses are essentially counts of the numbers of the people and households recorded in the census surveys as possessing particular combinations of characteristics within geographical areas at a range of scales from the whole UK down to Output Areas.

About InFuse

InFuse is designed to guide users in selecting census aggregate data relating to combinations of characteristics and areas of interest to them.

Bulk access

New project, allows download of whole open [datasets](#) for a range of popular geography types, includes boundary, flow and area data (more datasets being added monthly)

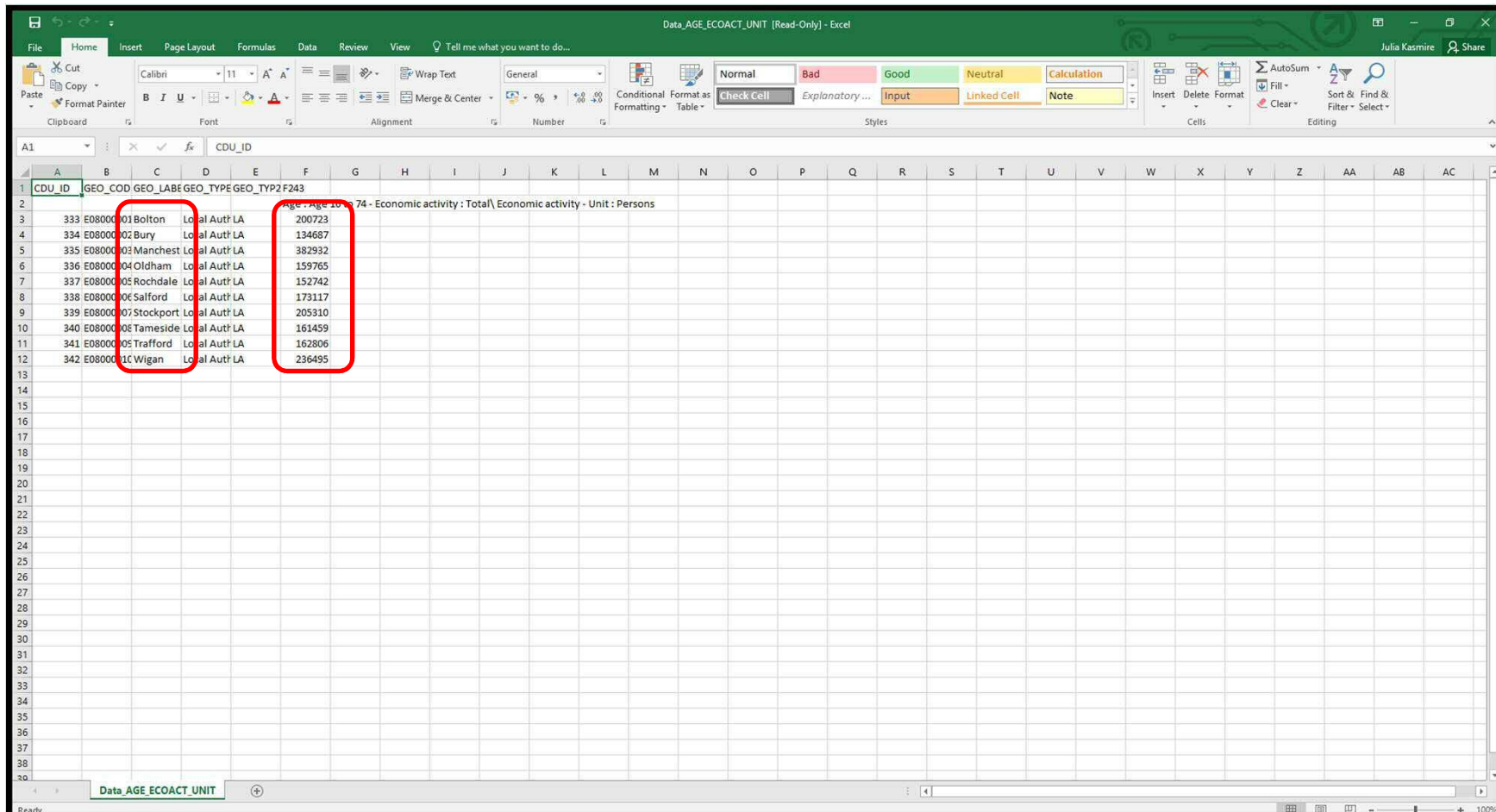
Related

- For census aggregate data from 1971 to 2001, see [Casweb](#)
- For matching and converting data to different geographies, see [GeoConvert](#)

[Follow](#)

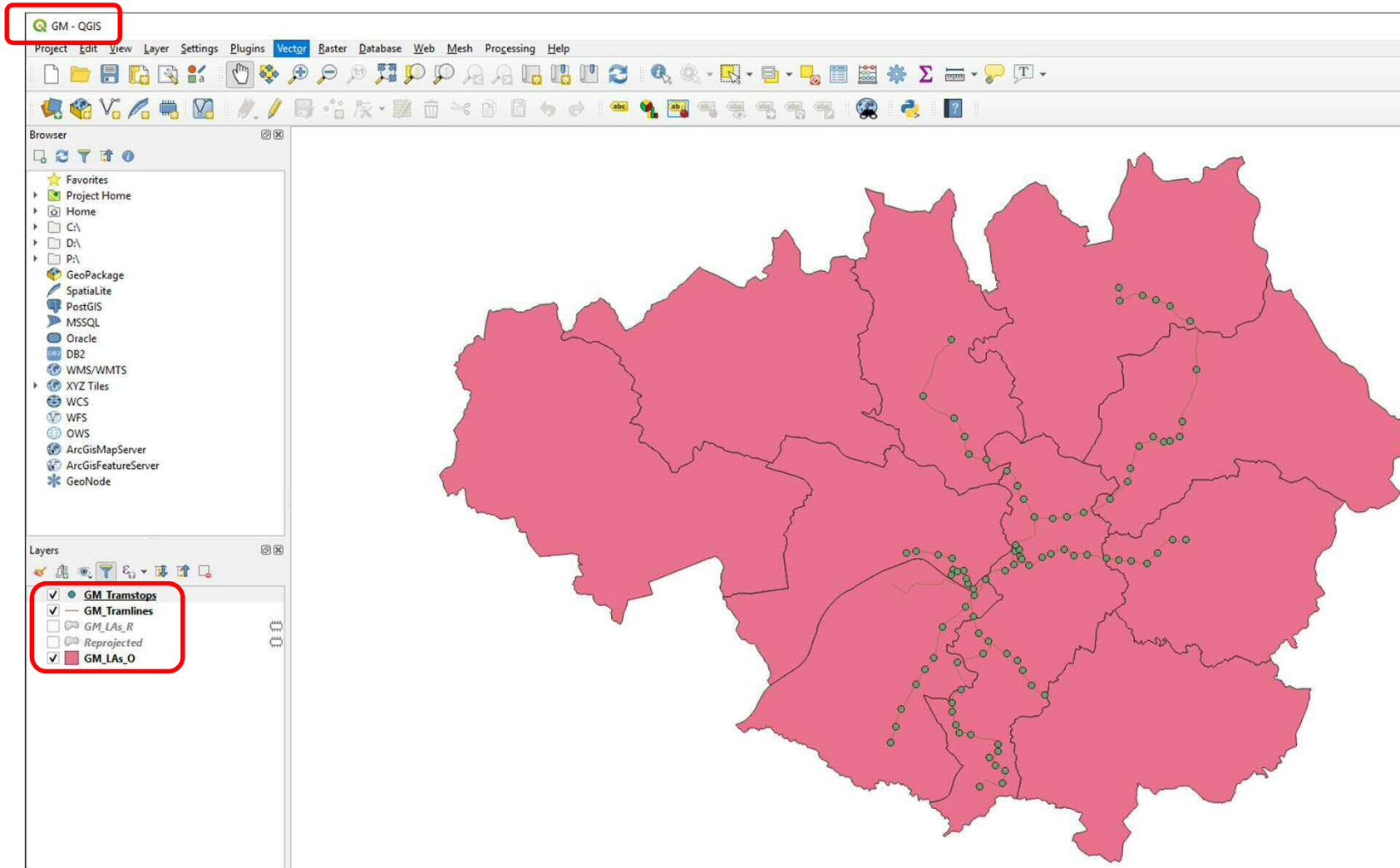
InFuse is part of the UK Data Service Census Support see our [Cookies policy CP2](#)

Viewing Census Data - population

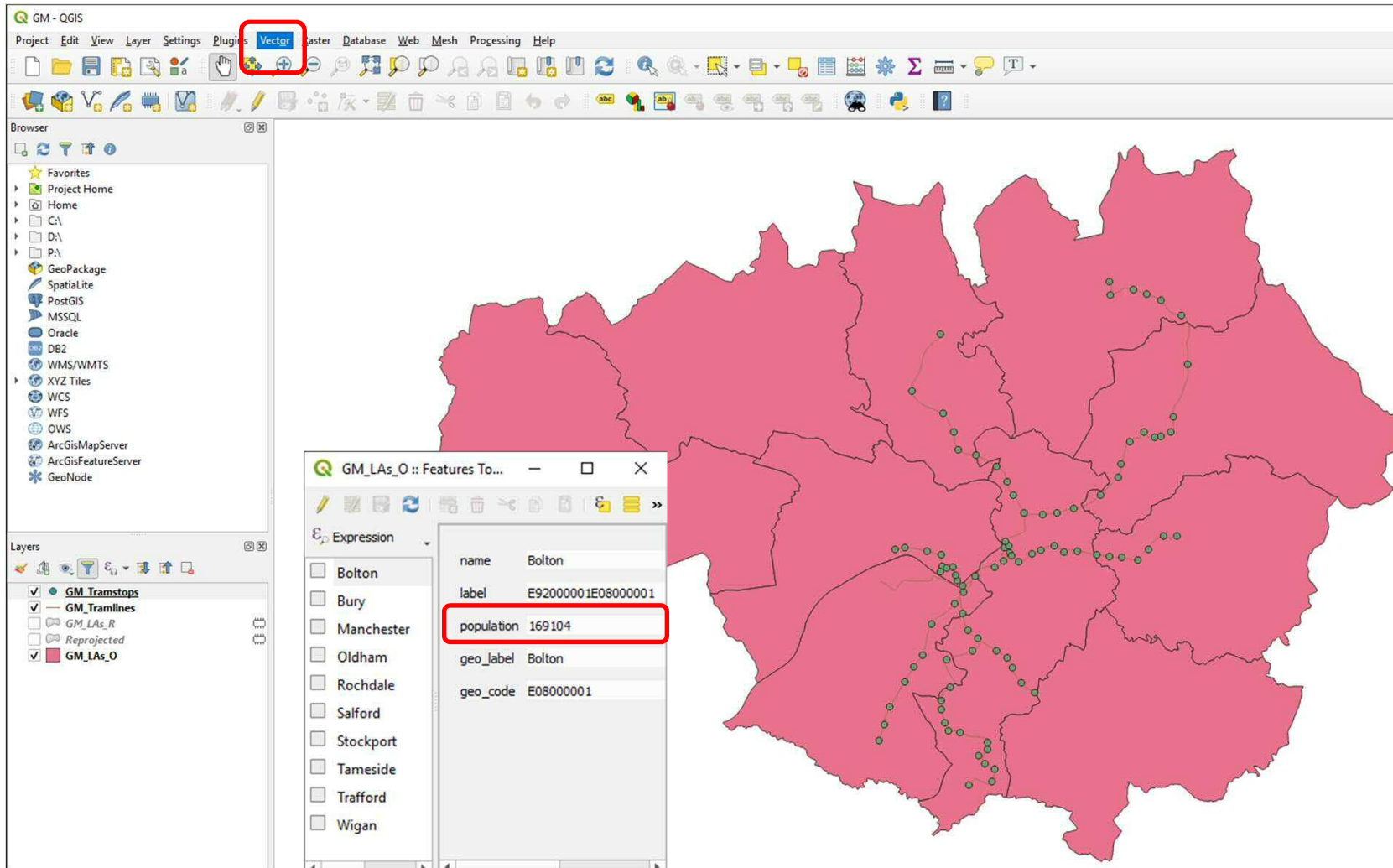


CDU_ID	GEO_COD	GEO_LAB	GEO_TYPE	GEO_TYP2	Age - Age 20-74 - Economic activity : Total	Economic activity - Unit : Persons
333	E08000001	Bolton	Local Auth	LA	200723	
334	E08000002	Bury	Local Auth	LA	134687	
335	E08000003	Manchest	Local Auth	LA	382932	
336	E08000004	Oldham	Local Auth	LA	159765	
337	E08000005	Rochdale	Local Auth	LA	152742	
338	E08000006	Salford	Local Auth	LA	173117	
339	E08000007	Stockport	Local Auth	LA	205310	
340	E08000008	Tameside	Local Auth	LA	161459	
341	E08000009	Trafford	Local Auth	LA	162806	
342	E08000010	Wigan	Local Auth	LA	236495	

Consolidate GIS and Census Data



Consolidate GIS and Census Data



Import Greater Manchester Shapefiles

```

TramCommuter - NetLogo (\\nask.man.ac.uk\home\Desktop\ABM)
File Edit Tools Zoom Tabs Help
Interface Info Code
Find... Check Procedures Indent automatically

; The non-random projection also has several steps, many are similar to those in the random set up.
; 1- Set the coordinate system or 'projection'. This is optional as long as all of the datasets use the same coordinate system.
; - Load all of your non-random datasets (as many as you need), assigning them to the globals created above.

; 2- Set the world envelope to the union of all of the datasets' envelopes. This ensures they line up correctly.

; 3- As with the Random projection, clear any agents that may be around.
; 4- Set the drawing color to white.
; - Draw the polygon data from the shapefile.
; 5- Technical processes of identifying features from the shapefile and loading them into temporary values.

; 6- The middle of each polygon is identified and added to a list (but not if it lies outside the world as defined).
; 7- If the centroid list is not empty,
; - Then create an LA agent and ...
; - Move it to the right position (right/left)
; - Move it to the right position (up/down)
; - Set their size to 0 so as to be invisible ...
; - Set their label color to yellow to increase visibility ...
; - Set their label color to yellow to increase visibility ...
; - Set their label to be their name, which is drawn from the imported shapefile ...
; - And copy that name to turtles-own feature.
; 7- Then the LA agents talk to the patch underneath themselves.
; - The LA agent asks the patch to copy details like population and name from the LA agent to itself.
; - And also asks the patches to set their color to red.

; 8- Pass the population feature from the LA to the patches within the LA
; - Also pass the name feature from LA to patches.
; 9- The patches then set their color relative to their population to improve visibility.

; - Ask any LA patches that are black to recolor themselves, just for clarity.

; Set up locations for tramstops (and create links between them) based on random generation and user input or uploaded shapefiles

; Drawing point data from a shapefile, creating tramstop-turtles at those points

to setup
  clear-all
  reset-ticks

  ; Load coordinate system (word "Model Data/" "projection" "pcr")
  ; Load datasets
  set tramlines-dataset gis:load-dataset "Model_Data/GM_Tramlines.shp"
  set LAS-dataset gis:load-dataset "Model_Data/GM_LAS_R.shp"
  set world-envelope (gis:envelope-union-of (gis:envelope-of tramlines-dataset)
                                           (gis:envelope-of LAS-dataset))

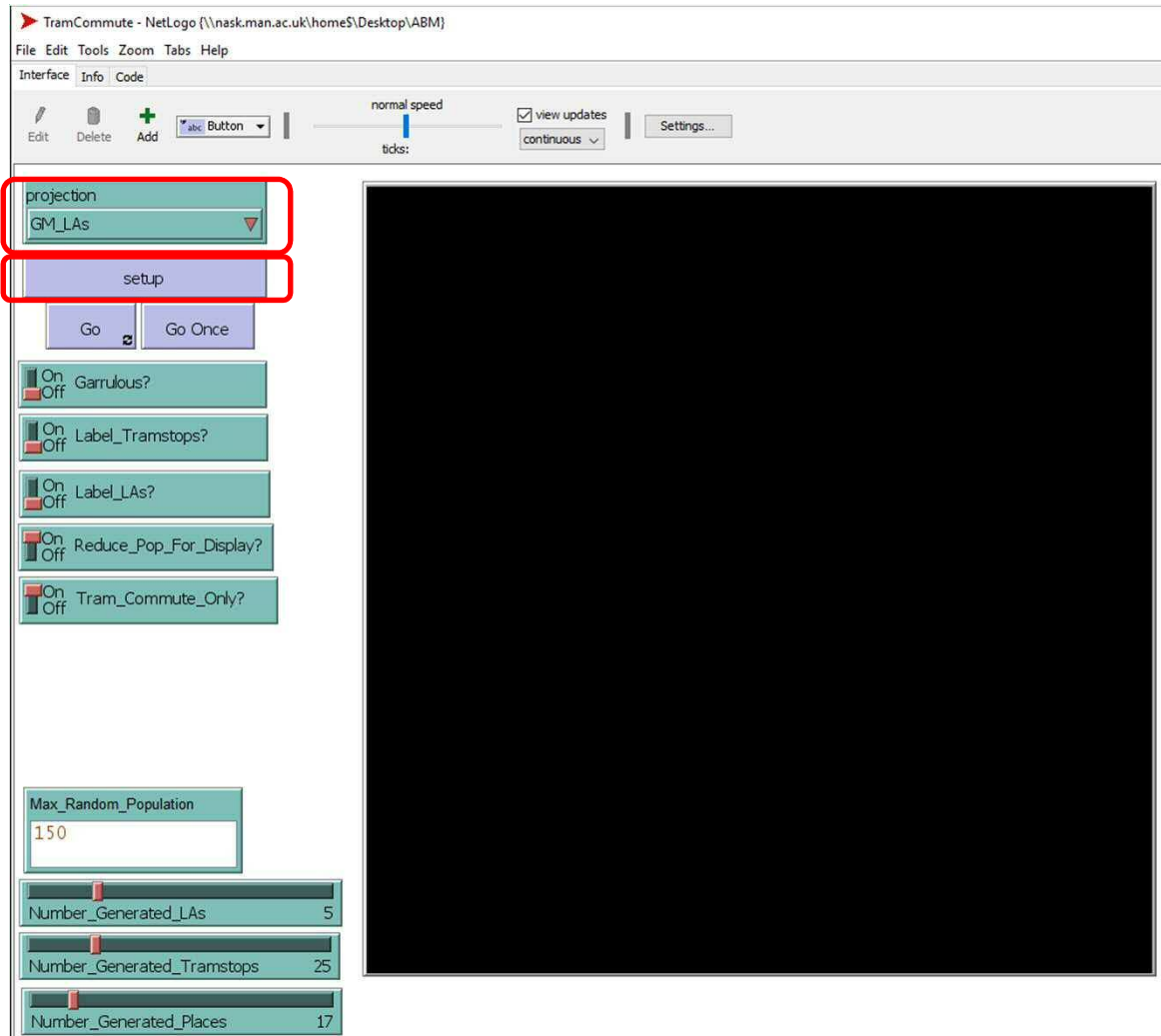
  ask LAS [ die ]
  gis:set-drawing-color white
  gis:draw LAS-dataset 1
  let i 1
  foreach gis:feature-list-of LAS-dataset [ vector-feature ->
    let centroid-y gis:location-of gis:centroid-of vector-feature
    if not empty? centroid-y
    [ create-LAS 1
      [ set xcor item 0 centroid-y
        set ycor item 1 centroid-y
        set size 0
        set label-color yellow
        if label_LAS? [set label gis:property-value vector-feature "name"]
        set LAS-population-t gis:property-value vector-feature "population"
        set LAS-name-t gis:property-value vector-feature "name"
        ask patch-here [set LAS-population [LAS-population-t] of LAS-here
          set LAS-name [LAS-name-t] of LAS-here
          set pcolor red] ] ]
      set i i + 1 ]

  gis:apply-coverage LAS-dataset "POPULATION" LAS-population
  gis:apply-coverage LAS-dataset "NAME" LAS-name
  let min-pop min [read-from-string LAS-population] of patches with [is-string? LAS-population]
  ask patches with [is-string? LAS-population] [
    set pcolor red + ((read-from-string LAS-population - min-pop) * .1)
    if pcolor = black [set pcolor pcolor + 5 ] ]

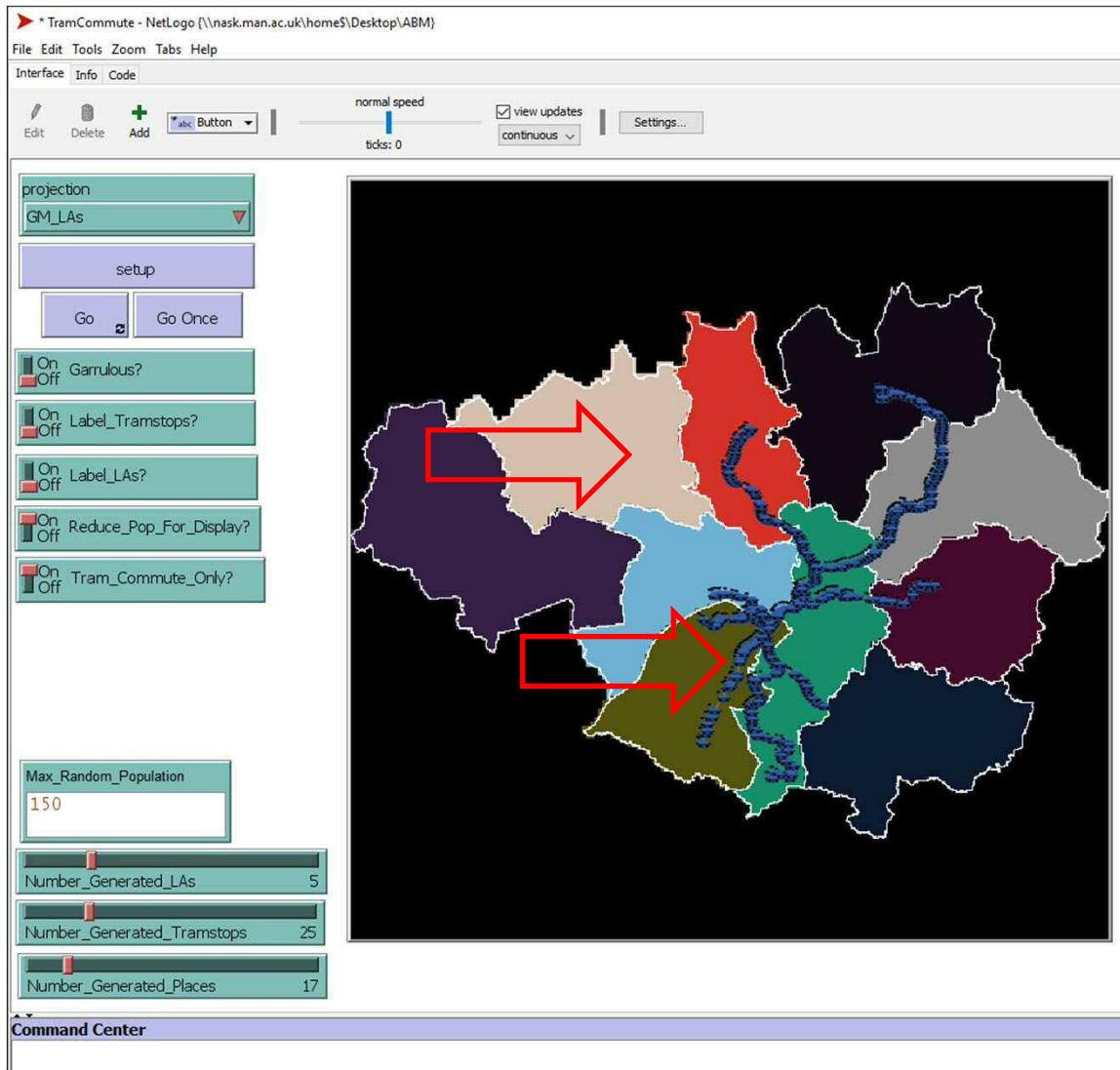
  ; Set up trams
  ask tramstops [ die ]
  set-default-shape tramstops "building store"
  ifelse projection = "Random"
  [ nw:generate-preferential-attachment tramstops links Number_Generated_Tramstops 1 [
    move-to one-of patches with [pcolor != black]
    if label_Tramstops? [set label (word LAS-name-t who) ] ] ]
  repeat 50 [ layout-spring tramstops links 0.2 11 1 ]
  [ gis:set-drawing-color cyan
    foreach gis:feature-list-of tramlines-dataset [ ?1 ->
      foreach gis:vertex-lists-of ?1 [ for the tramline feature, get the list of vertices
        [ ??1 ->
          let previous-node-pt nobody
          foreach ??1 [ for each vertex in road segment feature
            [ ???1 ->
              let location gis:location-of ???1
              if not empty? location

```

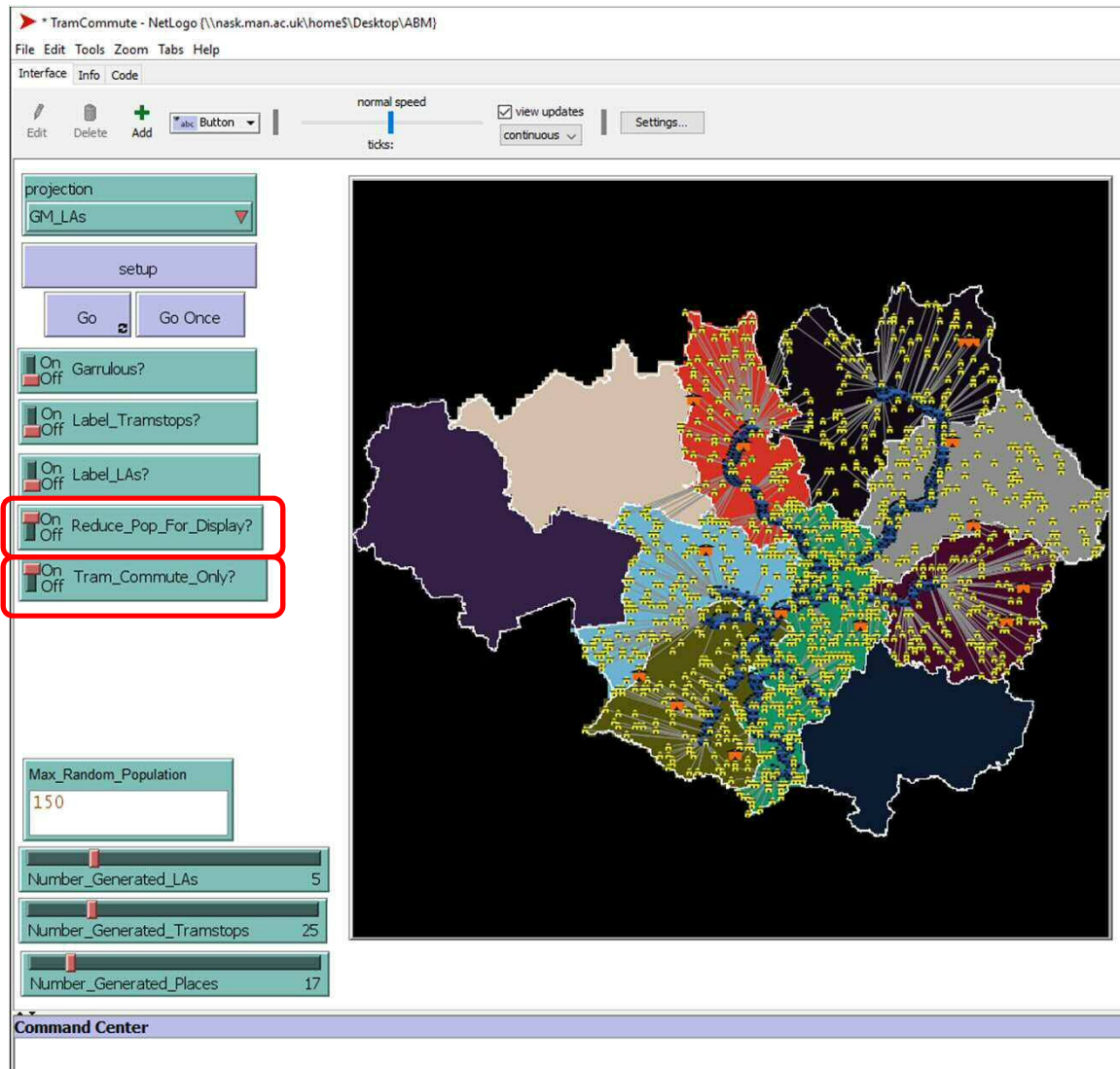
Import Greater Manchester Shapefiles



Import Greater Manchester Shapefiles



Import Greater Manchester Shapefiles



What can I do with this?

- Use different boundaries?
- Use data to guide behaviour data?
- Add average income/education level/health indicators/etc. data to residences?
- Community interaction?
- Add other ways to travel?
- Create relevant output?



Summary

- There are lots of verified, functional, documented models to explore
- Random generated models are very useful for testing
 - Check the model logic quickly
 - Rapid build-test cycles
 - Test code and data separately
 - Get sample output to check the format
 - Etc.
- There is a wealth of real-world GIS, census and other data to use.
- Real-world data needs a few steps to acquire, clean, combine, etc.
- Actually importing the real-world data into NetLogo models is easy

Questions

Dr. J. Kasmire

julia.kasmire@manchester.ac.uk

 @JKasmireComplex

ukdataservice.ac.uk/help/

Subscribe to UK Data Service news at

<https://www.jiscmail.ac.uk>

 @UKDataService

 UKDataService

