Grant agreement no: 265114

Seventh Framework Programme

Project acronym: URGENCHE

Project full title: Urban reduction of GHG Emissions in China and Europe

Deliverable: 3.2

Title: Prototype GIS System

Author name(s)/Affiliations: Dr. Zena Wood, Dr. Clive Sabel (University of Exeter), Niko Karvosenoja (SYKE) and Ville-Veikko Paunu (SYKE)

April 2013

Start date of project: 1 September 2011

Duration: 36 months

1. Introduction

Deliverable 3.2 involved the development of a prototype Geographical Information System (GIS), a system that comprises a set of tools that could be used by the cities within the project to collect, store and explore their datasets; the developed tools needed to allow each city to interactively explore their datasets. The GIS needed to be able to handle health and environmental data with core datasets including data relating to population, socio-economic status, topography, regional climate, infrastructure and administrative areas. The GIS needed to be web-based and build upon existing open-source software.

2. The Prototype System: Requirements

For development purposes the prototype was split into three components: uploading, processing and visualisation.

2.1. Uploading

Each dataset that was going to be analysed within the web-based GIS needed to be uploaded. Originally the aim was to have a geodatabase that would be a central repository for all datasets required within the project (i.e., by each and every each city). Two possible methods were identified for uploading the data: use a central systems administrator or allow each city to upload their own datasets. Consultation with the cities suggested that the latter option was preferable. A prototype interface was developed as depicted in figure 1. Since the GIS tools were to be provided on a city level, and different cities may need different tools, the user would need to indicate what city the dataset related to.

Figure 1: A prototype interface for uploading datasets

After further discussions it became clear that a central geodatabase would not be needed as individual cities are carrying out the analyses locally, with support my the methodological WPs. However, the feedback gained from the prototype interface was still important since any web-based tool would require the datasets to be uploaded.

2.2. Processing

It was important that the GIS tools that were to be developed would be useful to the users of the GIS. Consultation with the cities indicated the following types of analysis would be useful within an online environment.

- Filtering:
    - Data is filtered to only display those matches a specific criterion (e.g., where values are greater than, equal to or less than a given value).

- Proximity analysis:
    - Identify features that are near to another feature.

- Point density maps:
    - Splits the map into a neighbourhood and calculates the density of points per cell where a point will represent a given feature.

- Clipping tool
    - Allows a user to focus on a target region.

2.3. Visualisation

The GIS was required to provide a mechanism that would allow users to interactively explore their datasets and, therefore, the visualisation that the GIS would provide were very important. A GIS should allow a user to understand their data and quickly extract meaningful information (e.g., patterns, anomalies). There are many common visualisation methods that are used for geographical datasets including a range of charts, plots and graphs. Since each user should be able to interact with the visualisations, it is important that a variety of options are available to them within the GIS tool and where possible each visualisation should be clearly linked so patterns and impacts can become clear to a user; this would also aid scenario analyses.

3. The Prototype System: Implementation

The web-based GIS had to build upon existing open-source tools. Initially three technologies were identified as being potential solutions: *d3*, *Python* and *R*. d3, a JavaScript library, has been recently developed as a web-based technology that focuses on the visualisation of data allowing documents to be driven by the data on which they are based. Python is a scripting language that is becoming increasingly popular with GIS developers; a number of libraries have been developed specifically for use in GIS applications. Applications need not be web-based if developed in Python. R is a programming language that has been developed for statistics and graphics. Like Python, libraries have been developed for R that are specific to GIS applications. Discussions with potential users suggested that the interactivity and powerful visualisations that *d3* offered made it the most suitable option. Therefore, the prototype GIS tool was developed using *d3* and the necessary web technologies (e.g, *php*).

The remainder of this section will describe and give evidence of the functionality of the present GIS according to the requirements listed in section 2. To illustrate the functionality of the GIS a single sample dataset will be used throughout; the dataset has been chosen because it allows all of the capabilities of the GIS to be shown and incorporates a variety of information relating to population, socio-economic status, and administrative areas.

3.1. Uploading

A web-based interface, shown in figure 2, has been developed that allows users to upload a dataset for analysis and processing within the GIS. A user must select what city they are from and choose which file they would like to upload. The system can currently process *.csv* (comma-separated values files), *.json* (JavaScript Object Notation files) but will soon be updated to include .shp (shapefiles); the uploading tool will automatically identify what type of file has been uploaded and alert the user if the file is not compatible.
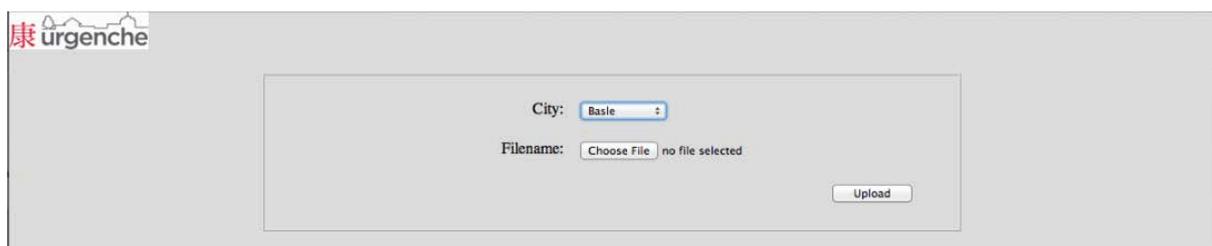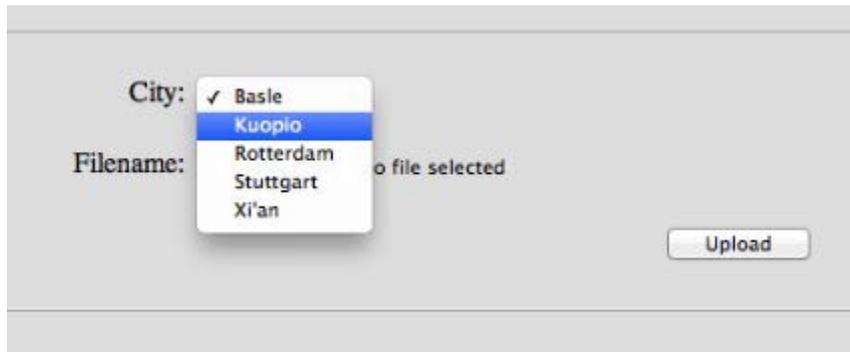

Figure 2: The interface for uploading a dataset

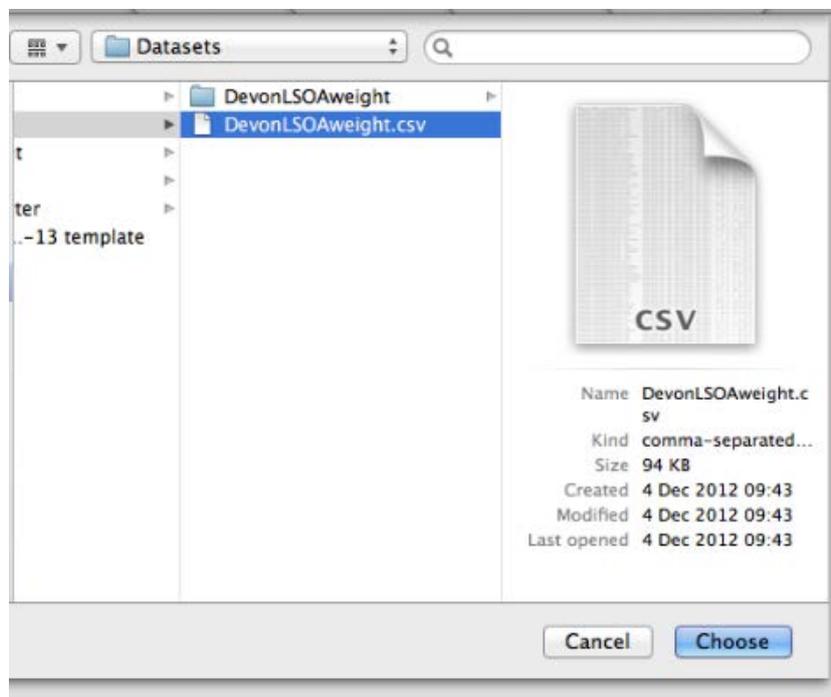**Figure 3: The user must select what city they are from**

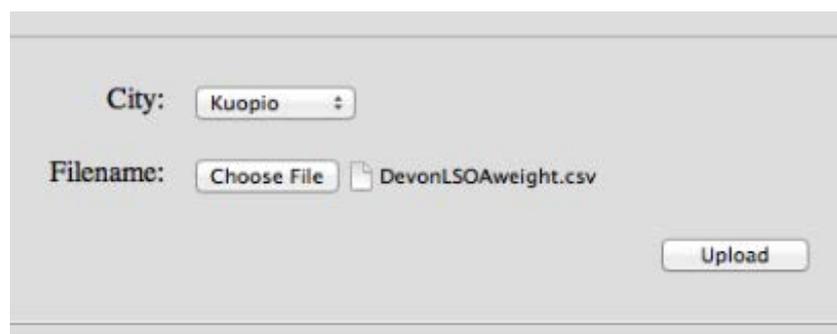**Figure 4: The user selects which dataset they would like to upload**


Figure 5: Once the file has been chosen the user can upload their dataset

3.2. Processing

Consultation with the cities indicated that four types of analysis would be useful in an online tool. The prototype GIS includes elements of filtering, proximity analysis and clipping; point density maps are currently being added. The technology that has been used to develop the prototype, *d3* is a powerful visualisation tool that allows users to interactively explore their datasets. It was established that the prototype GIS tool would focus on this element of the technology developing a collection of visualisation tools that would allow a user to identify patterns and links within their dataset.

Once a dataset has been successfully uploaded the GIS will identify the categories within the dataset that could be analysed. The extracted categories are displayed to the user who can then select which ones they may wish to include in later analysis (see figure 6).

Message informing user of successful upload
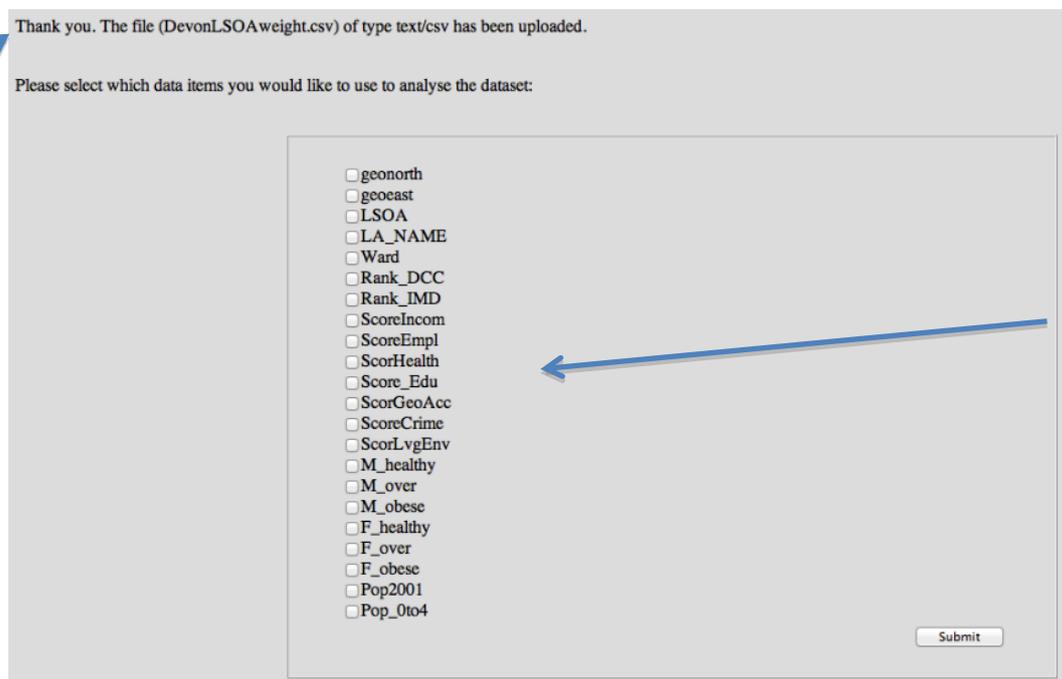
Thank you. The file (DevonLSOAweight.csv) of type text/csv has been uploaded.

Please select which data items you would like to use to analyse the dataset:

- geonorth
- geoeast
- LSOA
- LA_NAME
- Ward
- Rank_DCC
- Rank_IMD
- ScoreIncom
- ScoreEmpl
- ScorHealth
- Score_Edu
- ScorGeoAcc
- ScoreCrime
- ScorLvgEnv
- M_healthy
- M_over
- M_obese
- F_healthy
- F_over
- F_obese
- Pop2001
- Pop_0to4

Submit

The GIS identifies the elements in the dataset that could be analysed

Figure 6: The interface that allows a user to select the categories that they would like to analyse

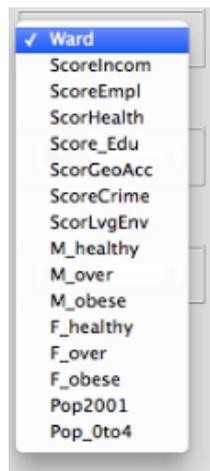The options that the user has selected are taken forward for further analysis.

Figure 7: The populated list of categories chosen by the user

3.3. Visualisation and Analysis

A range of visualisations has been developed within the GIS that allow a user to interactively explore their datasets. These visualisations include different charts: grouped bar, stacked bar, doughnut, scatter and area. A user can select different categories to compare and display, focusing on a small section of the dataset or the entire dataset (i.e., filtering). Up to three charts can be linked to help identify patterns and anomalies. Linked charts are viewed in the browser at the same time and any interaction with one chart will update the linked charts accordingly. For example, the user could identify a subsection of the data that they would like to concentrate on or the categories that are being compared.

3.3.1. Examples of Charts

Figure 8: Categories being compared according to proportion. In this example, the weight of the population is considered according to gender and the categories: healthy, overweight and obese.
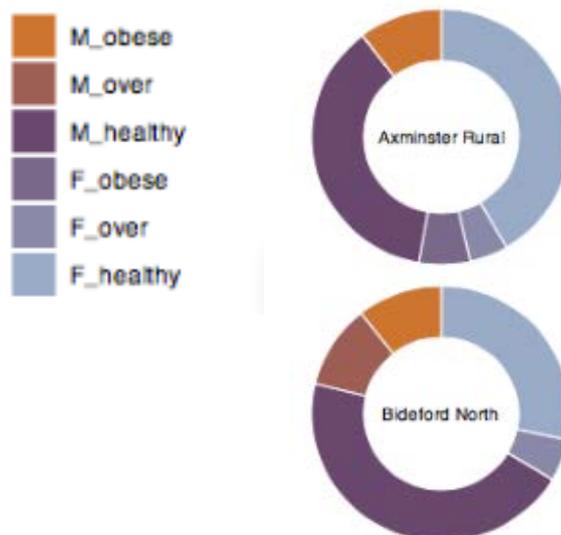

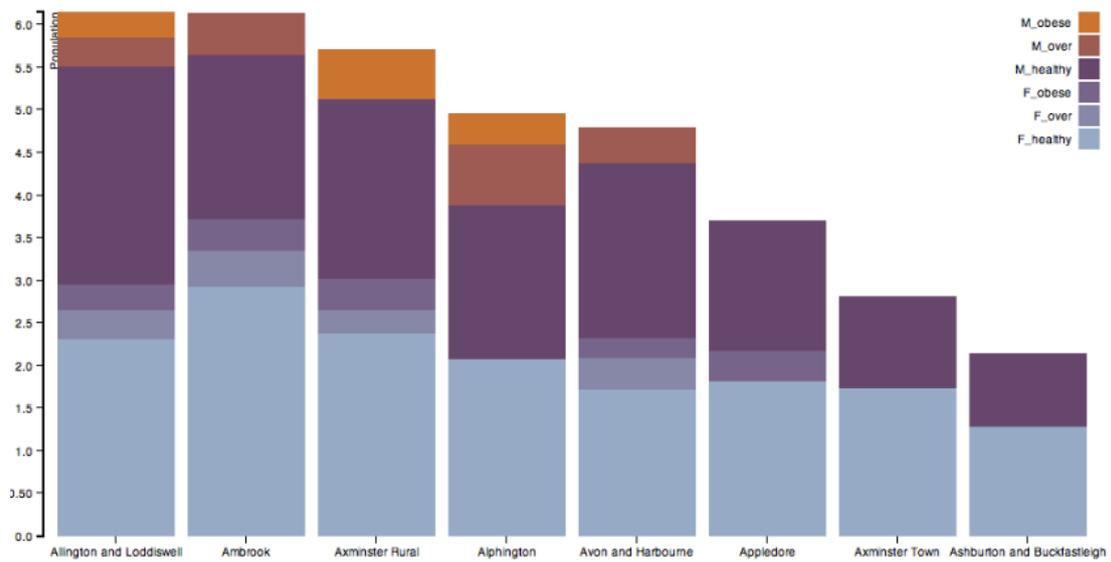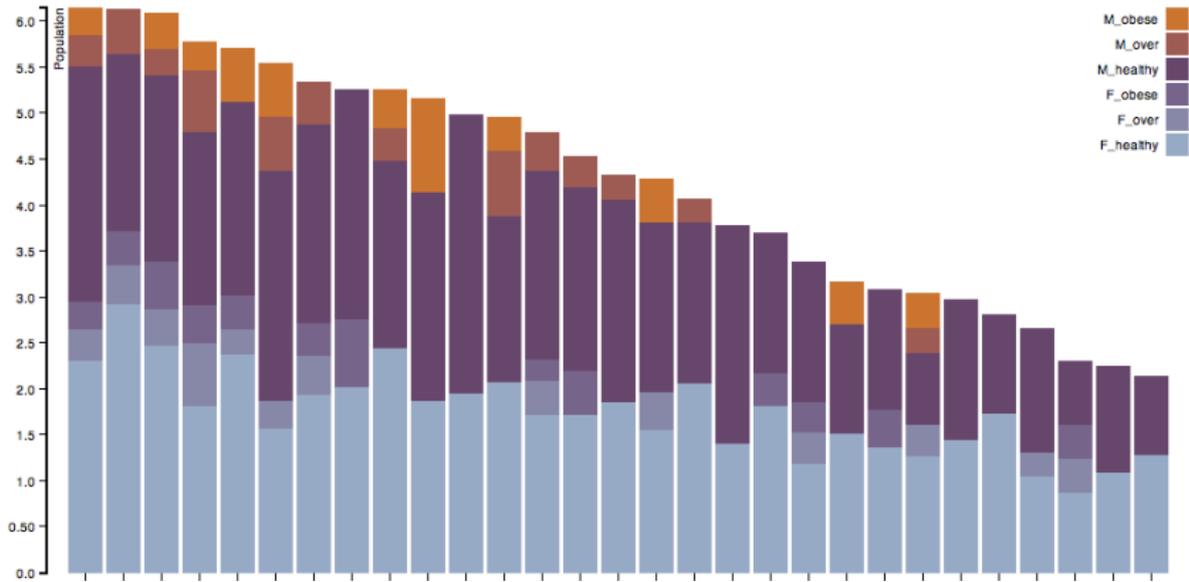**Figure 9: Two sample wards from figure 8**

Figure 10: Categories can be compared in interactive bar charts. Users can scroll through datasets in this format when there is too much to fit on the screen. Bar charts can also be sorted according to a category.
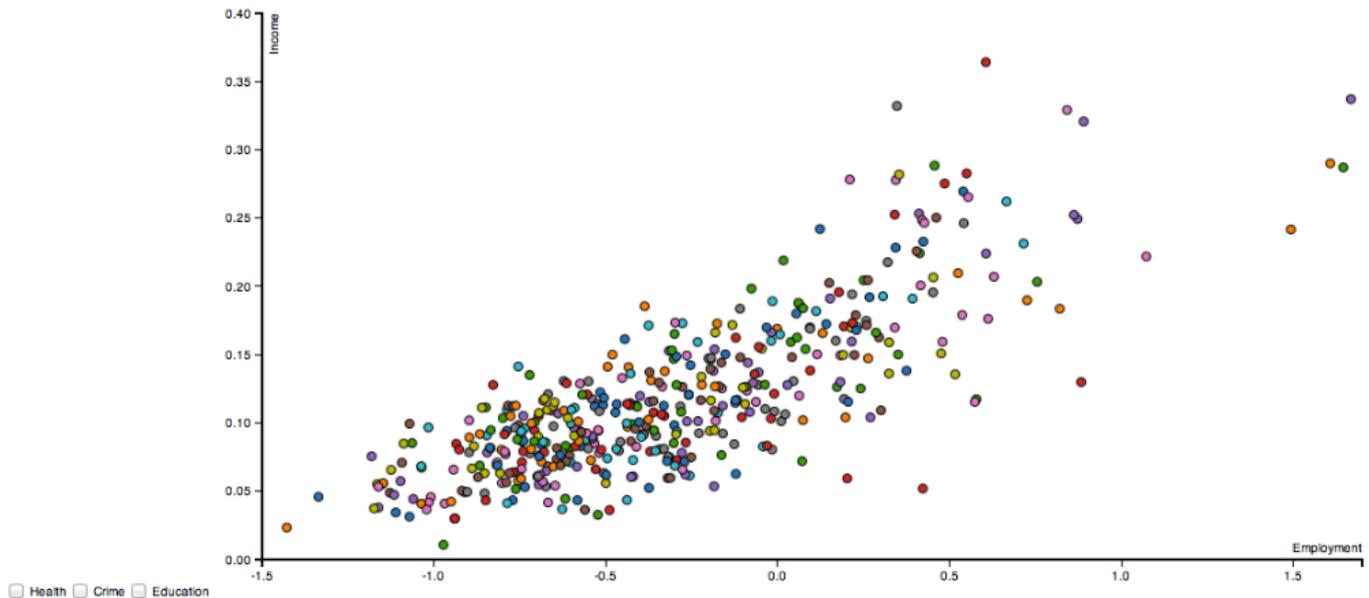
Figure 11: Interactive scatterplots allows users to choose which categories they compare

4. Discussion

The developed prototype GIS tool comprises a set of tools and methods that can be used by each of the cities within the project. The GIS is currently being updated to allow the visualisations to be linked to a map when possible (i.e., there is sufficient data within the dataset that would allow the map to be drawn). Currently the user must go through a drop down list; a map would be more visually pleasing and ensure the user understands the topography that is being considered.

It has become apparent that many of the cities are using their own GIS tools; these tools have a much greater level of functionality, much more than could be included in a prototype web-based system. Discussions with the cities have identified that a tool is needed that would allow them to disseminate the results of their analysis to policy makers. The feedback received regarding the developed prototype GIS tool, and from the development of the tool, has shown that some of the technologies that have been used could be adapted to form the necessary dissemination tool.

5. Conclusion

Deliverable 3.2 involved the development of a prototype GIS system that would allow cities to collect, store and interactively explore their datasets. The system has to be web-based and built upon existing open-source software. Discussions with the cities led to a list of requirements being developed. From these requirements suitable technologies were identified and a prototype GIS system developed in *d3.* Although the final prototype fulfils many of the given requirements it has become clear that

each city has access to GIS tools that provide a greater level of functionality. Therefore, the developed prototype will be adapted to provide a GIS tool that would allow a city to communicate the results of their analyses to policy makers.