

CALCULATING ROAD DISTANCES BETWEEN LOCATIONS

The Stata command **georoute**¹ is a great shortcut for reliably calculating routing distances and travel time between either two geographical coordinates or two addresses. This command is particularly useful as it calculates the driving distance and travel time (by car, under usual traffic conditions) as opposed to straight-line distances. This is important in socio-economic research as it enables the user to obtain a measure of proximity or isolation of locations relative to others, such as urban centres, employment centres, ports or airports, local amenities, etc.

Additionally, it does not require loading context-specific maps and shapefiles into Stata or even outsourcing this step to another software package. **georoute** calculates the routing distance between two geographical points, identified by their coordinates. The command draws on mapping data by a commercial provider called HERE, which is of high quality even for remote contexts, but requires you to set up a free account for developers (see below). Once this is set up, the HERE ID and HERE CODE are used directly in the code.

The simple example below illustrates this in rural Kenya, where the variable of interest is the distance from the survey respondent's village to the major local city, Kisumu. The code can be installed using the command:

```
ssc install georoute
```

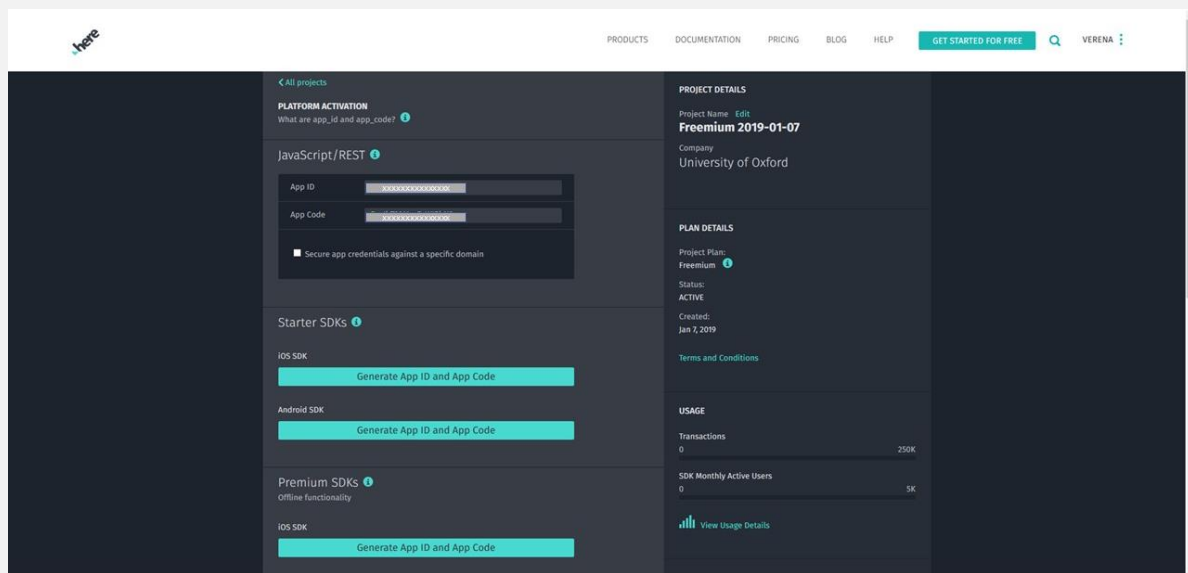
You also need to install the two packages **insheetjson** and **libjson** in order to run **georoute**.

The syntax for the command is:

```
georoute, hereid(here_id) herecode(here_code) startxy(village_latitude  
village_longitude) endxy(kisumu_latitude kisumu_longitude) km  
distance(roaddistance_kisumu) time(ttime_kisumu)
```

The first options specified in the syntax are **hereid()** and **herecode()**. As mentioned, in order to use this package, you have to set up an account on [HERE](https://developer.here.com/) and fill in these two options based on your account's ID and Code – these will be two long strings, a combination of letters, numbers and symbols. Once you have your account set up and are logged in, click “get started for free” to set up your own developer plan (called Freemium). This plan can be used for up to 250,000 transaction a month. Start a new project and generate your APP ID and APP Code under JavaScript/REST (see screenshot below). Both the ID and the Code are a series of 20 and 22 characters.

¹ **georoute** is a user-written Stata command by Weber and Péclat (2017).



Following the ID and code, the user has the option to specify start point and end point of the distance using either the options `startxy(varlist)` and `endxy(varlist)` or `startaddress(varlist)` and `endaddress(varlist)`. If using the first method to specify the start and finish points, your varlist should specify the coordinates in decimal degrees; see the example above, which specifies `village_latitude` and `village_longitude` for the start point and `kisumu_latitude` and `kisumu_longitude` as the end point. The former was recorded using GPS equipment for each village, whereas the latter was obtained from open data sources. Note the latitude must be between -90 and 90 and the longitude between -180 and 180.

`georoute` also has the capability to calculate the distance between two points defined by their addresses using the `startaddress(varlist)` and `endaddress(varlist)` options. Let's say we want to calculate the distance between a plant of that produces the popular Pick N'Peel fruit juice in Thika, Kenya and several markets across the country. The address can be saved in a single string variable or multiple variables (i.e. one specifying the road, one the postcode and one the town).

As an intermediate step the programme first finds the geographical coordinates of the specified addresses before returning the usual output. In case we want to keep the coordinates from the intermediate step, we can ask the programme to do so using the option `coordinates(str1 str2)`. The option requires us to specify one prefix for the coordinates of the start address and one for the end address (`factorylocation` and `marketlocation` in the example below).

The syntax for the above example finding the address of Pick N'peel fruit juice is:

```
georoute, hereid(here_id) herecode(here_code) startad(Piock N'peel fruit
juice, Thaika, Kenya) endad(markettown, country) km
distance(factory_market_distance) time(ttime_factory_market)
coordinates(factorylocation marketlocation)
```

By default, the command generates three variables; `travel_distance`, `travel_time` and `georoute_diagnostic`. The first is road distance, the second is the travel time and the last is an indicator of any issues that the program ran into for an observation, such as "no route found". The names of these variables can be changed from default by using the options `distance()`, `time()` and `diagnostic()`, respectively. By default, the distance is calculated in miles, unless the option `km` is specified. The travel time is given in minutes. The command help file provides details of other options available.