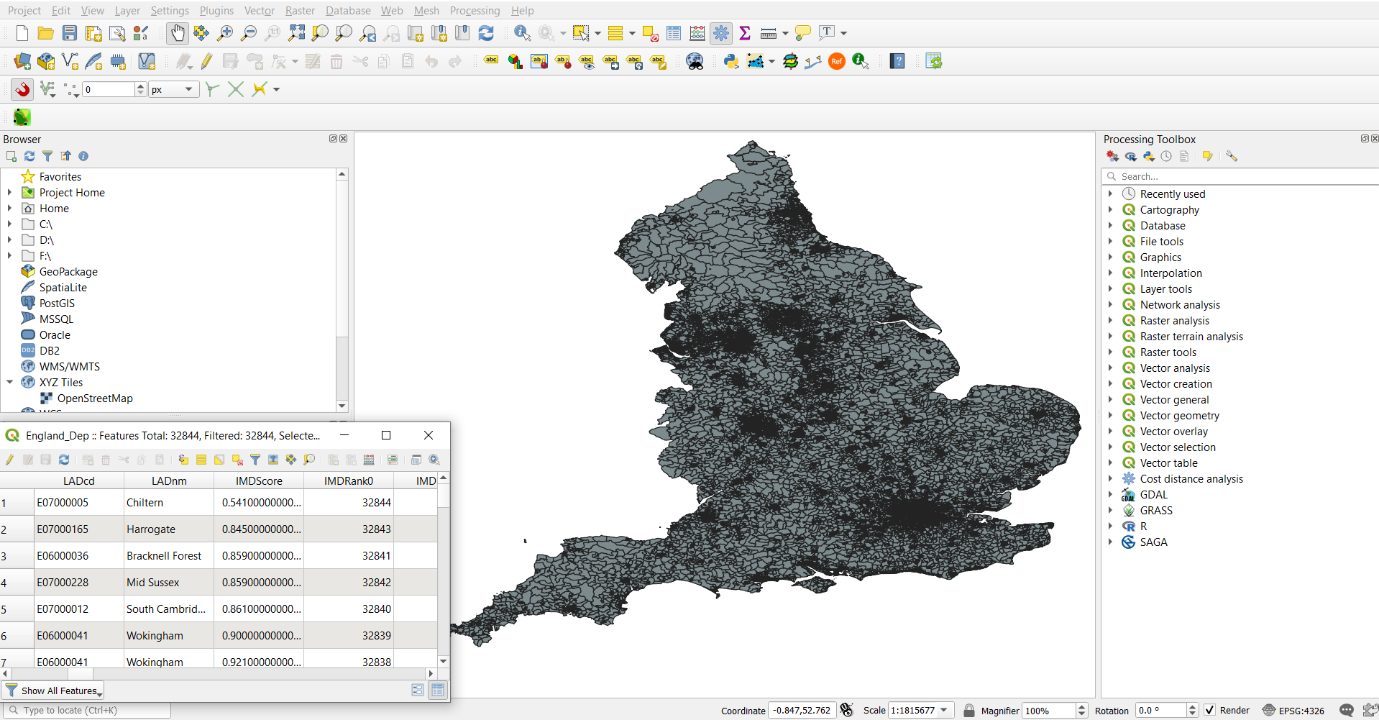
**R & QGIS: Integrating Statistical and Spatial Data Analysis**

In this practical, we will work through some methods together, utilising both QGIS and R.

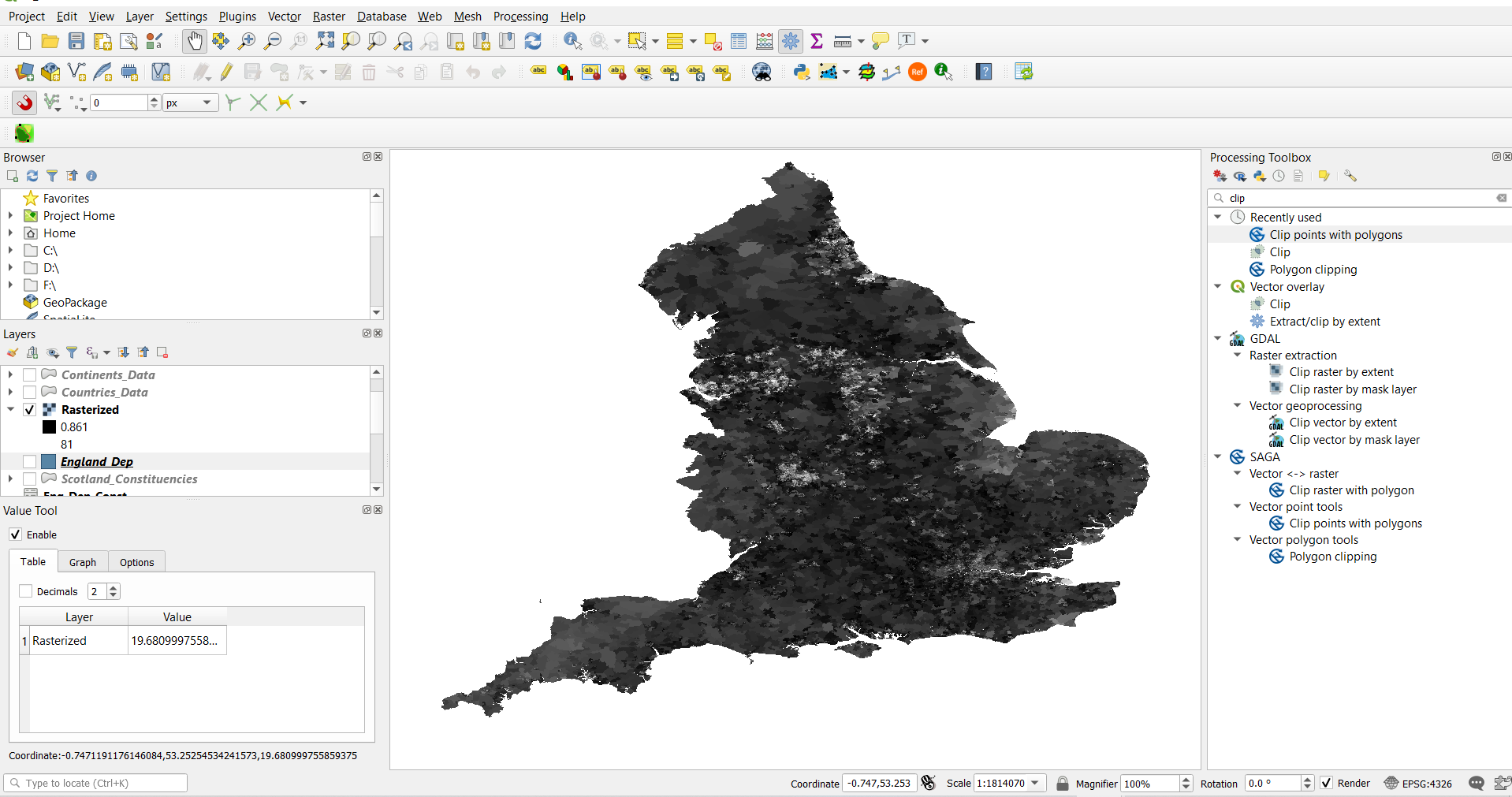
We will use the data we have used so far for the practicals in order to find and visualise links between deprivation and voting in the UK 2015 election in England.

We have two shapefiles, one the familiar GB constituencies map and the second the IMD for England. These have different polygons as constituencies do not line up exactly with administrative boundaries used to calculate IMD. However, there is an overlap.

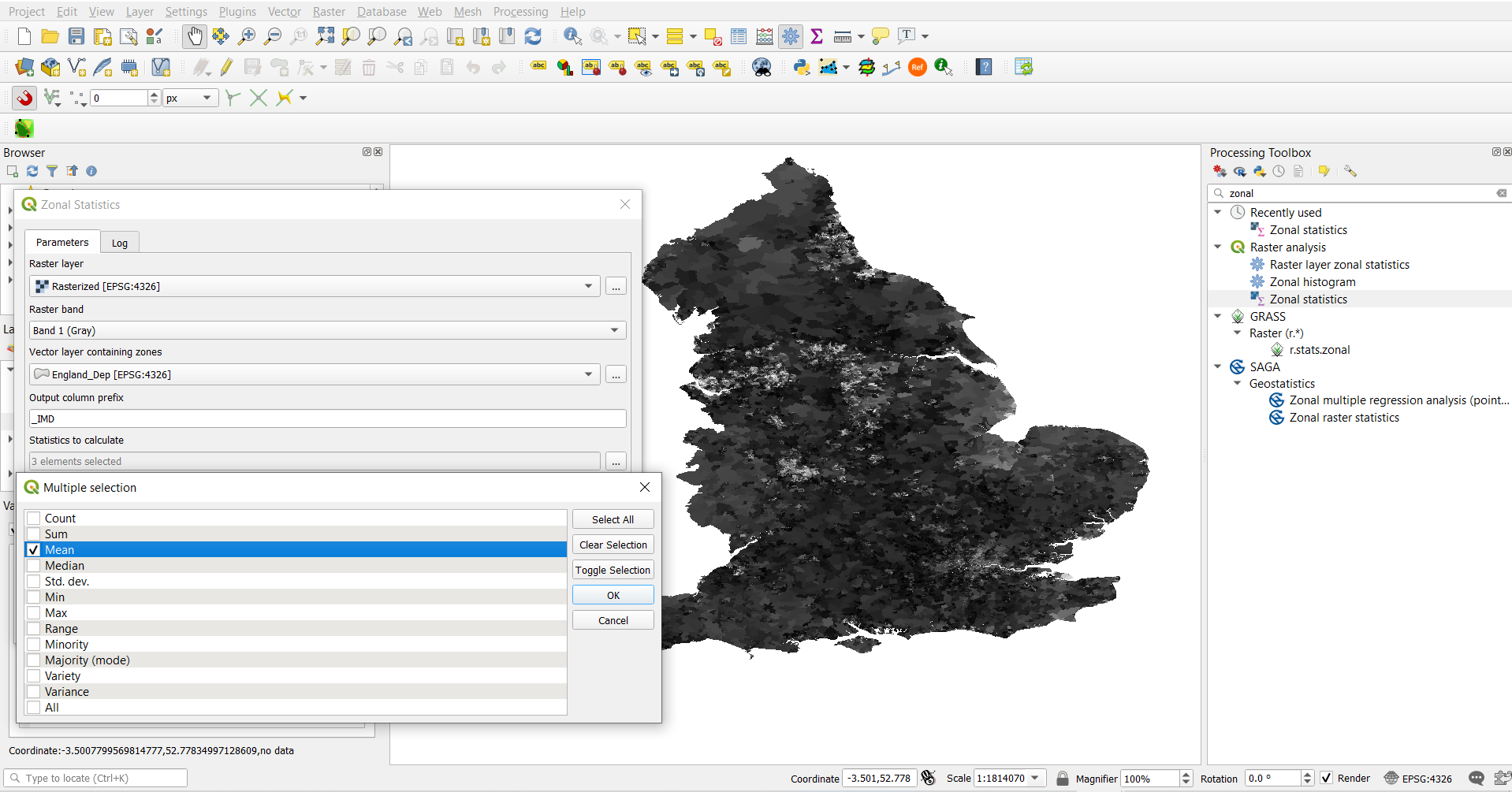
First we need to isolate Constituencies.shp to include only the English constituencies (IMD is calculated differently in each country of the UK, so UK wide analysis is unhelpful).



We can add the mean IMD score from Eng\_Dep.shp to the English Constituencies.shp. Eng\_Dep.shp has to be rasterised first, and then zonal statistics can be used to calculate the mean (or a variety of additional functions).



**(The ‘value tool’ plugin is extremely helpful for quickly obtaining exact raster values in a given cell)**

****

With zonal statistics the selected statistics from the specified raster can be added to the polygon. We can then export this data as a .csv and perform some further analysis in R.

Using group\_by we can calculate the mean IMD for each parties constituencies.

Eng\_Dep\_Con\_Mean <- Eng\_Dep %>%

  group\_by(frst\_pr) %>%

  summarise\_at(vars(IMDmean),

               list(IMD\_Mean = mean))

