Department for Transport

Mapping WITA Outputs

User Manual

April 2022

Department for Transport
Great Minster House
33 Horseferry Road
London SW1P 4DR

This publication is also available on our website at
www.gov.uk/government/organisations/department-for-transport

Any enquiries regarding this publication should be sent to us at
www.gov.uk/government/organisations/department-for-transport

Contents

1. Preliminaries

Introduction

WITA (Wider Impacts in Transport Appraisal) is a software application developed by Atkins Limited for the Department for Transport to estimate wider impacts of transport schemes. The purpose of WITA is to undertake the estimation of wider impacts in accordance with the Department's guidance as set out in TAG Unit A2.1 "Wider Economic Impacts Appraisal" and associated Units A2.2 to A2.4. WITA estimates the following types of wider impacts: induced investment, employment effects and productivity impacts.

The Mapping WITA Output tool maps the outputs calculated by WITA in RStudio. The mapping tool can be used to produce national and regional choropleth maps of wider impacts which are customisable according to the user's needs. For example, the data mapped can be aggregated over all impacts across the entire appraisal period or disaggregated to a specific wider impact in a specified year. Aesthetic changes to the map can also be made. The mapping tool can map data for the agglomeration and employment impacts estimated by WITA.

This instructions manual provides an overview of the files required to run the code and a guide on how to run the code including explanations of modifications the user can make.

User Agreement

The Department for Transport assumes no liability for the use of the Mapping WITA Outputs tool or any outputs it produces.

The user acknowledges that they run the Mapping WITA Outputs tool at their own risk. The user also acknowledges that any outputs produced by the tool are used at their own risk and that the user is liable for the data displayed on outputs.

The Mapping WITA Outputs tool is designed display WITA data on maps according to the parameters outlined in this user manual. Such is the nature of the R, the Mapping WITA Outputs R script is open source. The user is liable for any amendments that they make to the code.

The use of outputs from the Mapping WITA Outputs tool is at the discretion of the user. The Department for Transport assumes no liability for the use of outputs from the tool.

Licencing

Mapping WITA Output is covered by the **The MIT License**. This is stated below.

Copyright (c) 2022 Crown Copyright (Government Digital Service)

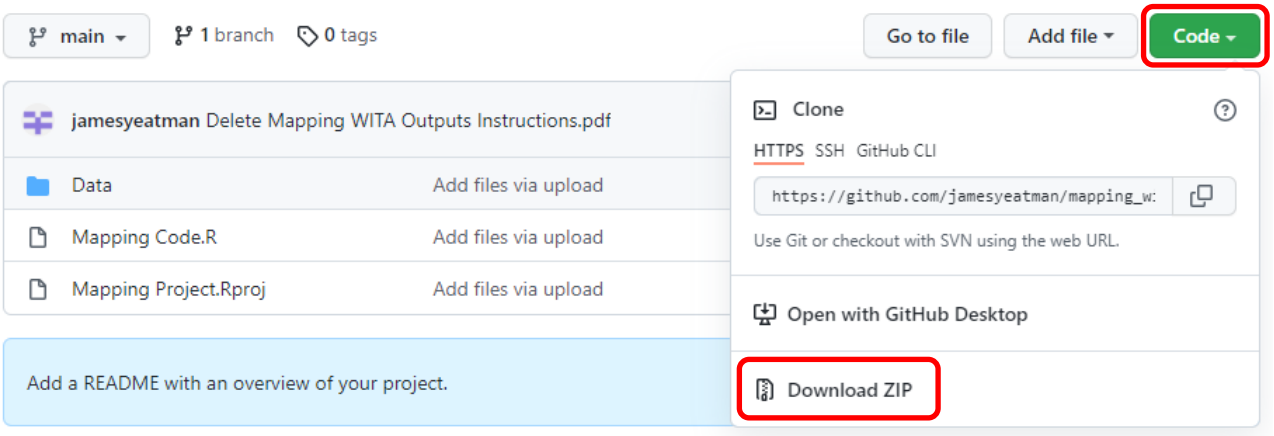
Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Access from GitHub

The Mapping WITA Output tool and supporting files are available to download from the Department for Transport GitHub page. To do this, navigate to the ”Code" section of the "mapping_wita_outputs" repository. Here, select "Download ZIP", found under the green "Code" dropdown menu. This will download a ZIP file including the entire contents of the repository which includes everything needed to run the Mapping WITA Outputs tool.



Dependencies

The Mapping WITA Outputs tool is dependent on the user having access to RStudio. RStudio is freely available to download here.

The Mapping WITA Outputs tool is dependent on six R packages freely available to download from CRAN. Lines 16 to 21 of the Mapping WITA Outputs R Script automatically download these packages. The relevant packages are:

- tidyverse
- rgeos
- rgdal
- maptools
- readxl
- stringr

Files Required

The folder “Mapping WITA Outputs” contains all the files required to run the mapping tool. Inside this folder are the following:

- “April 2022 Mapping WITA Outputs.PDF” – this instruction manual.
- “Mapping Code.R” – the code to run.
- “Mapping Project.RPROJ” – the project in which the code sits.
- “Data” – a folder with relevant data and data templates inside.

The folder “Data” contains the relevant data and data templates to run the code. Inside this folder are:

- “LAD SHP 2015 SG” – this is a folder with the data used to create the map plot. This folder does not need altering at any point.
- “Templates” – this folder contains the “Data Template” and “Zone Template” files.
- “LAD to Region.csv” – this file details which regions LAD are in. This file does not need altering at any point.
- “Employment Forecasts.csv” – this file contains employment forecasts used for the per employee option. this file contains employment forecasts used for the per employee option. This file does not need altering at any point.

The folder “Templates” contains the csv files:

- “Data Template.csv” – this is where the WITA outputs should be pasted.
- “Zone Template.csv” – this is where the WITA zone to LAD matrix used to run WITA should be pasted.
- “Data Template Example.csv” – this is an example using random numbers for the purpose of this manual.
- “Zone Template Example.csv” – this is an example for the purpose of this manual.

Queries

For queries regarding the Mapping WITA Output tool, please contact the DfT's TASM Mailbox using the email address TASM@dft.gov.uk.

2. A Guide to the Code

Opening the Code

When accessing the code, download the entire “Mapping WITA Outputs” folder and save in an appropriate place on your computer. Do not change the location of files within this folder.

To open the project, open RStudio, go to “File/Open Project…” and navigate to the “Mapping Project.Rproj” file saved in the “Mapping WITA Outputs” folder.

To open the code, select the “Mapping Code.R” file within the “Files” directory panel in RStudio.

Data Templates

Open the file “Data Template.csv”. From cell A2, paste in the data produced by WITA and found in the “detail” csv file generated. Double check that the columns line up appropriately and that you have copied the entirety of your data into the template file. Save your template with your appropriate name, in the same location at the original file “Data Template.csv”. For example, if your appraisal is for the ABC road scheme, then name it “Data Template ABC.csv”.

Open the file “Zone Template.csv”. From cell A2, paste in the WITA zone to LAD matrix which was used as an input for WITA. If your matrix is in a text file, then use the Excel “From Text/CSV” function in the “Data” ribbon to import your data. Double check that the columns line up appropriately and that you have copied the entirety of your data into the template file. Save your template with your appropriate name, in the same location at the original file “Zone Template.csv”. For example, if your appraisal is for the ABC road scheme, then name it “Zone Template ABC.csv”.

Defining Variables

```
23 # ## USER INPUT ## Define variables  ------------------------------------
24
25   #choose "sum_of_appraisal_period" or a number corresponding to a year in the appraisal period, eg. 2040
26   forecast_year <- "sum_of_appraisal_period"
27
28   #choose "agglomeration_manufacturing", "agglomeration_construction", "agglomeration_consumer_services",
29   #       "agglomeration_producer_services", "labour_supply_impact", "M2MLPJs", "total_agglomeration" or
30   #       "total_wider_economic_impact"
31   forecast_data <- "total_wider_economic_impact"
32
33   #choose "yes" or "no"
34   per_employee <- "no"
35
36   #choose "2011", "2016", "2021", ..., "2081"
37   employment_forecast <- "2021"
38
39   #choose "Great Britain", "East Midlands", "East of England", "London", "North East", "North West",
40   #       "Scotland", "South East", "South West", "Wales", "West Midlands", "Yorkshire and The Humber"
41   data_region <- "Great Britain"
42
```

Lines 23 to 42 define the data variables for which you want to map.

Data Period

Line 26 defines the period of data mapped. After the arrow on line 26, you can type the year in your appraisal period that you wish to map the data for, or "sum_of_appraisal_period" to map the data summed over the entire appraisal period.

Note: you do not have to use quotation marks when defining a numerical year, but you do when specifying the sum.

Data Type

Line 31 defines the data mapped. The options here align to the WITA outputs found in the “detail” file. Depending on the data you wish to map, after the arrow on line 31, type the appropriate option from the list below (note that quotation marks here are important):

- "agglomeration_manufacturing"
- "agglomeration_construction"
- "agglomeration_consumer_services"
- "agglomeration_producer_services"
- "labour_supply_impact"
- "M2MLPJs"
- “total_agglomeration”
- “total_wider_economic_impact”

Note: If a data series is fully populated by zeros, then the code will display an error message.

Per Employee Option

Line 34 instructs the code whether you would like the data mapped to be displayed in per employee units or not. To display in per employee units, type “yes” after the arrow on line 34. Otherwise, type “no” after the arrow on line 33.

Line 37 defines the forecast year the employment data should be taken from. For example, if you wish to use the employment data from 2021, then “2021” should be typed after the arrow on line 37. The time scale of the employment data is every 5 years from 2011 to 2081. The source of the data is the July 2021 version Wider Impacts Dataset.

Note: Lines 111 to 116 fix the population of low population LADs to that of a higher population neighbour. The LADs altered are the Isles of Scilly (E06000053), the Orkney Islands (S12000023), Na h-Eileanan Siar (S12000013), and the Shetland Islands (S12000027). This is done so that low population LADs do not have very high benefits per employee which then distort the colour gradient on the map produced. Lines 110 to 115 can be deleted if you wish to display benefits per employee without these adjustments.

Region

Line 41 defines the region for which the LAD should be mapped. Depending on the region you wish to map, after the arrow on line 41, type the appropriate option from the list below (note that quotation marks are important):

- “Great Britain”
- “East Midlands”
- “East of England”
- “London”
- “North East”
- “North West”
- “Scotland”
- “South East”
- “South West”
- “Wales”
- “West Midlands”
- “Yorkshire and The Humber”

The “Great Britain” option displays the entire map.

If you wish to map more than one region on the same map, then a vector should be written after the arrow on line 41 listing all regions you wish to map. For example, to map Scotland, the North East and the North West, the code on line 41 would read:

```
data_region <- c("Scotland", “North East”, “North West”)
```

Note: combining regions lowers the quality of the map. It may be that small white gaps appear between LADs.

Defining Map Aesthetics

```
43 # ## USER INPUT ## Define aesthetics for map plot  ------------------------
44
45   high_colour <- print("darkgreen") #choose colour representing a high value
46
47   mid_colour <- print("grey95") #choose colour representing a zero value
48
49   low_colour <- print("red3") #choose colour representing a low value
50
51   title <- print("black") #choose "black" to display the title, or "white" to not
52
53   legend <- print("right") #choose "right" to display the legend, or "none" to not
54
```

Lines 43 to 54 define variables which determine the aesthetics of the map.

Colours

Lines 45, 47 and 49 define the high, midpoint and low colours used for the scale gradient, respectively. Here you can change these colours. Colours must be written in quotation marks within the print() function. A useful guide of colours used in R can be found here.

Title

Line 51 defines whether the title is displayed on the plot. If you want the title to be displayed, then line 51 should read:

```
title <- print("black")
```

If you do not want the title to be displayed, then line 51 should read:

```
title <- print("white")
```

Legend

Line 53 defines whether the legend is displayed on the plot. If you want the legend to be displayed, then line 53 should read:

```
legend <- print("right")
```

If you do not want the legend to be displayed, then line 53 should read:

```
legend <- print("none")
```

Displaying LAD Borders

```
184 # ##USER INPUT## Plot map ------------------------------------------------
185
186   plot <- ggplot(data = wita_map, aes(x = long, y = lat, group = group, fill = eval(parse(text = forecast_data)))) +
187     #define geom_polygon(size = 0.1, colour = "black") to include LAD borders, or geom_polygon() to remove them
188     geom_polygon() +
189     coord_equal() +
190     theme_void() +
```

Lines 184 to 207 define the format of the map plot and instruct R to produce the map.

If you wish to produce a map without black LAD borders, then line 188 should read:

```
geom_polygon() +
```

If you wish to include the black LAD borders, then line 188 should read:

```
geom_polygon(size = 0.1, colour = "black") +
```

Note that a non-white colour (such as “grey90”) should be used for the midpoint colour in line 47 if LAD borders are not used.

Loading Data

Loading WITA Data

```
60 # ##USER INPUT## Load WITA output data ----------------------------------
61
62   #Change file name where appropriate
63   wita_output <- read.csv("Data/Templates/Data Template Example.csv")
64
```

Lines 60 to 64 load the data you previously populated the data template with. In line 63, you will need to specify the name of the data template file you generated. It is important to include the file path “Data/Templates/” before your file name.

For example, if you named your file “Data Template Example.csv”, then line 63 should read as in the screenshot above. Alternatively, if you named your file “Data Template ABC.csv”, then line 62 should read:

```
wita_output <- read.csv("Data/Templates/Data Template ABC.csv")
```

Loading Zone Matrix

```
65 # ##USER INPUT## Load WITA zone to LADs matrix ----------------------------
66
67   #Change file name where appropriate
68   wita_zones <- read.csv("Data/Templates/Zone Template Example.csv")
69   wita_zones <- as.data.frame(wita_zones)
70
```

Lines 65 to 70 load the data you previously populated the zone template with. In line 68, you will need to specify the name of the zone template file you generated. It is important to include the file path “Data/Templates/” before your file name.

For example, if you named your file “Zone Template Example.csv”, then line 68 should read as in the screenshot above. Alternatively, if you named your file “Zone Template ABC.csv”, then line 68 should read:

```
wita_zones <- read.csv("Data/Templates/Zone Template ABC.csv")
```

3. Examples

Example 1

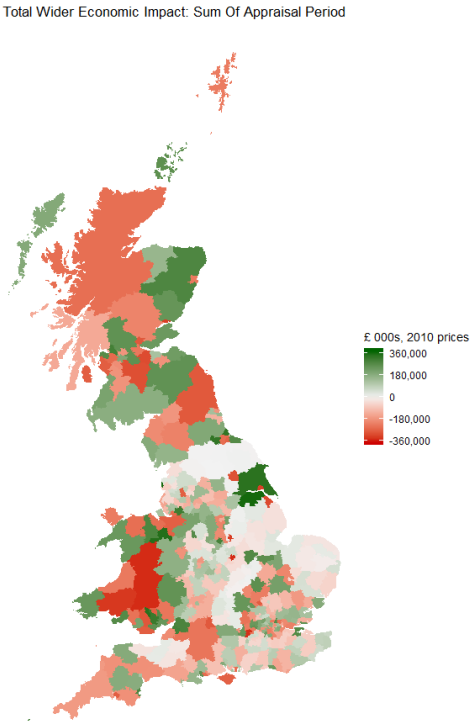
Example 1: User Requirements

```
# ## USER INPUT ## Define variables  ---------------------------------------

  #choose "sum_of_appraisal_period" or a number corresponding to a year in the appraisal period, eg. 2040
  forecast_year <- "sum_of_appraisal_period"

  #choose "agglomeration_manufacturing", "agglomeration_construction", "agglomeration_consumer_services",
  #       "agglomeration_producer_services", "labour_supply_impact", "M2MLPJs", "total_agglomeration" or
  #       "total_wider_economic_impact"
  forecast_data <- "total_wider_economic_impact"

  #choose "yes" or "no"
  per_employee <- "no"

  #choose "2011", "2016", "2021", ..., "2081"
  employment_forecast <- "2021"

  #choose "Great Britain", "East Midlands", "East of England", "London", "North East", "North West",
  #       "Scotland", "South East", "South West", "Wales", "West Midlands", "Yorkshire and The Humber"
  data_region <- "Great Britain"

# ## USER INPUT ## Define aesthetics for map plot  ------------------------

  high_colour <- print("darkgreen") #choose colour representing a high value

  mid_colour <- print("grey95") #choose colour representing a zero value

  low_colour <- print("red3") #choose colour representing a low value

  title <- print("black") #choose "black" to display the title, or "white" to not

  legend <- print("right") #choose "right" to display the legend, or "none" to not
```

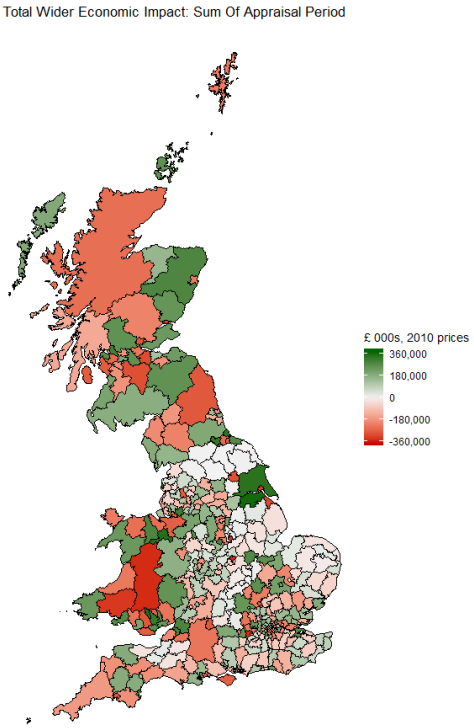
Example 1: Data

```
# ##USER INPUT## Load WITA output data ------------------------------------

  #Change file name where appropriate
  wita_output <- read.csv("Data/Templates/Data Template Example.csv")

# ##USER INPUT## Load WITA zone to LADs matrix ----------------------------

  #Change file name where appropriate
  wita_zones <- read.csv("Data/Templates/Zone Template Example.csv")
  wita_zones <- as.data.frame(wita_zones)

```

Example 1: Map without LAD Boundaries*



Example 1: Map with LAD Boundaries*



*The data used in these examples is fabricated and not based on any real-life scheme. To include LAD boundaries, see the Displaying LAD Boundaries section of this user manual.

Example 2

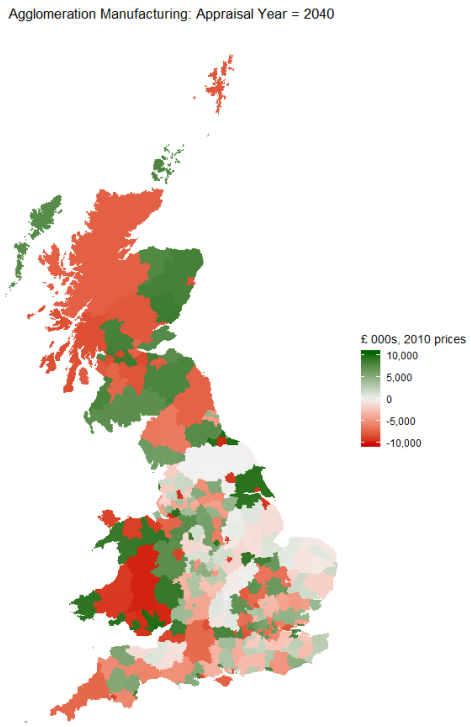
Example 2: User Requirements

```
# ## USER INPUT ## Define variables  ------------------------------------

  #choose "sum_of_appraisal_period" or a number corresponding to a year in the appraisal period, eg. 2040
  forecast_year <- 2040

  #choose "agglomeration_manufacturing", "agglomeration_construction", "agglomeration_consumer_services",
  #       "agglomeration_producer_services", "labour_supply_impact", "M2MLPJs", "total_agglomeration" or
  #       "total_wider_economic_impact"
  forecast_data <- "agglomeration_manufacturing"

  #choose "yes" or "no"
  per_employee <- "no"

  #choose "2011", "2016", "2021", ..., "2081"
  employment_forecast <- "2021"

  #choose "Great Britain", "East Midlands", "East of England", "London", "North East", "North West",
  #       "Scotland", "South East", "South West", "Wales", "West Midlands", "Yorkshire and The Humber"
  data_region <- "Great Britain"

# ## USER INPUT ## Define aesthetics for map plot  ------------------------

  high_colour <- print("darkgreen") #choose colour representing a high value

  mid_colour <- print("grey95") #choose colour representing a zero value

  low_colour <- print("red3") #choose colour representing a low value

  title <- print("black") #choose "black" to display the title, or "white" to not

  legend <- print("right") #choose "right" to display the legend, or "none" to not
```

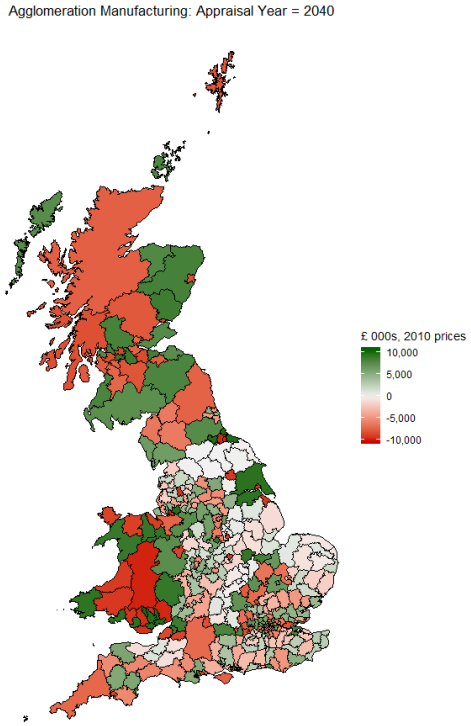
Example 2: Data

```
# ##USER INPUT## Load WITA output data ------------------------------------

  #Change file name where appropriate
  wita_output <- read.csv("Data/Templates/Data Template Example.csv")

# ##USER INPUT## Load WITA zone to LADs matrix ----------------------------

  #Change file name where appropriate
  wita_zones <- read.csv("Data/Templates/Zone Template Example.csv")
  wita_zones <- as.data.frame(wita_zones)

```

Example 2: Map without LAD Boundaries*



Example 2: Map with LAD Boundaries*



*The data used in these examples is fabricated and not based on any real-life scheme. To include LAD boundaries, see the Displaying LAD Boundaries section of this user manual.

Example 3

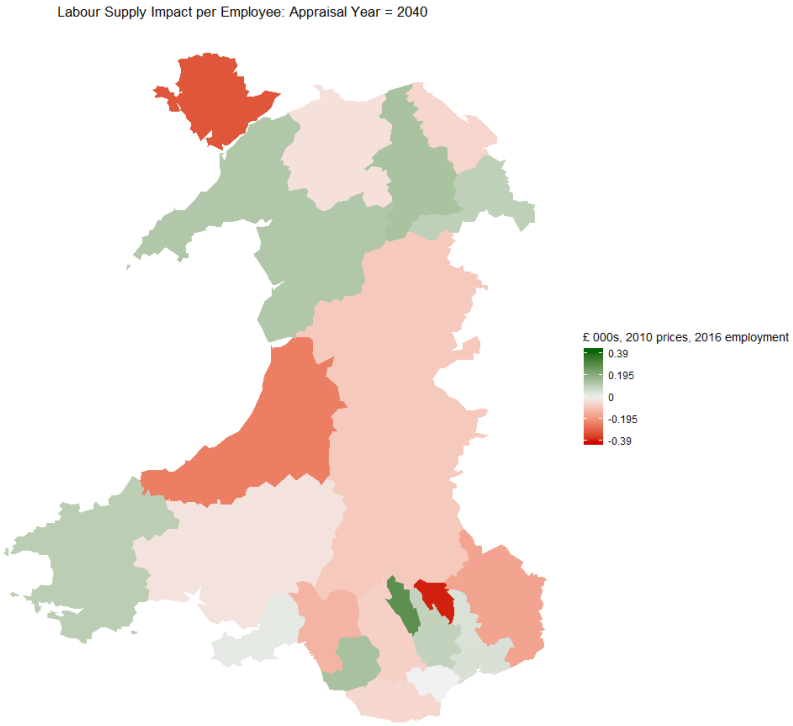
Example 3: User Requirements

```
# ## USER INPUT ## Define variables  ---------------------------------------

  #choose "sum_of_appraisal_period" or a number corresponding to a year in the appraisal period, eg. 2040
  forecast_year <- 2040

  #choose "agglomeration_manufacturing", "agglomeration_construction", "agglomeration_consumer_services",
  #       "agglomeration_producer_services", "labour_supply_impact", "M2MLPJs", "total_agglomeration" or
  #       "total_wider_economic_impact"
  forecast_data <- "labour_supply_impact"

  #choose "yes" or "no"
  per_employee <- "yes"

  #choose "2011", "2016", "2021", ..., "2081"
  employment_forecast <- "2016"

  #choose "Great Britain", "East Midlands", "East of England", "London", "North East", "North West",
  #       "Scotland", "South East", "South West", "Wales", "West Midlands", "Yorkshire and The Humber"
  data_region <- "Wales"

# ## USER INPUT ## Define aesthetics for map plot  ------------------------

  high_colour <- print("darkgreen") #choose colour representing a high value

  mid_colour <- print("grey95") #choose colour representing a zero value

  low_colour <- print("red3") #choose colour representing a low value

  title <- print("black") #choose "black" to display the title, or "white" to not

  legend <- print("right") #choose "right" to display the legend, or "none" to not
```

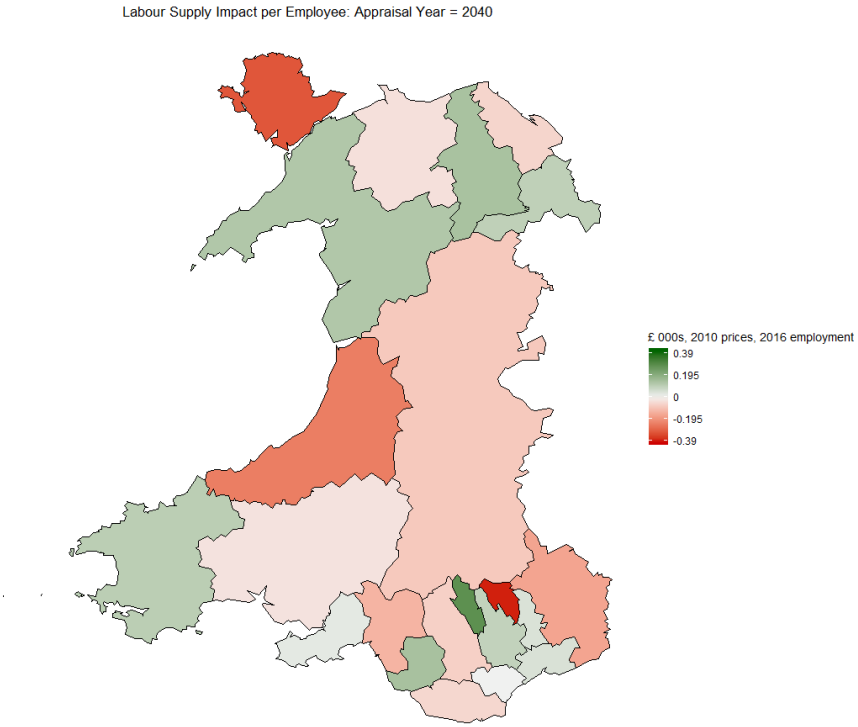
Example 3: Data

```
# ##USER INPUT## Load WITA output data ---------------------------------------

  #Change file name where appropriate
  wita_output <- read.csv("Data/Templates/Data Template Example.csv")

# ##USER INPUT## Load WITA zone to LADs matrix ----------------------------

  #Change file name where appropriate
  wita_zones <- read.csv("Data/Templates/Zone Template Example.csv")
  wita_zones <- as.data.frame(wita_zones)

```

Example 3: Map without LAD Boundaries*



Example 3: Map with LAD Boundaries*



*The data used in these examples is fabricated and not based on any real-life scheme. To include LAD boundaries, see the Displaying LAD Boundaries section of this user manual.

Example 4

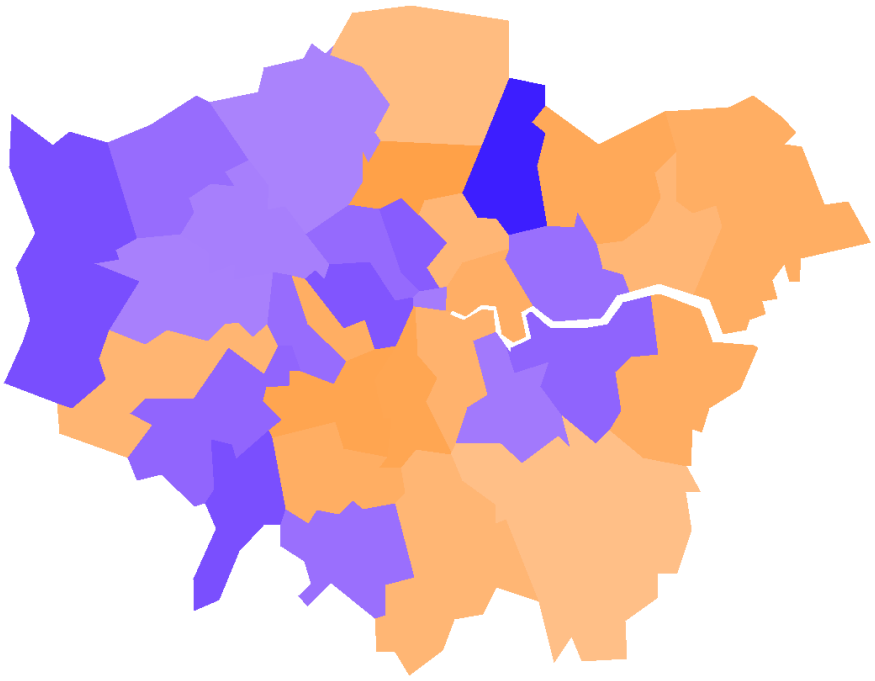
Example 4: User Requirements

```
# ## USER INPUT ## Define variables  ------------------------------------

  #choose "sum_of_appraisal_period" or a number corresponding to a year in the appraisal period, eg. 2040
  forecast_year <- "sum_of_appraisal_period"

  #choose "agglomeration_manufacturing", "agglomeration_construction", "agglomeration_consumer_services",
  #       "agglomeration_producer_services", "labour_supply_impact", "M2MLPJs", "total_agglomeration" or
  #       "total_wider_economic_impact"
  forecast_data <- "total_agglomeration"

  #choose "yes" or "no"
  per_employee <- "no"

  #choose "2011", "2016", "2021", ..., "2081"
  employment_forecast <- "2021"

  #choose "Great Britain", "East Midlands", "East of England", "London", "North East", "North West",
  #       "Scotland", "South East", "South West", "Wales", "West Midlands", "Yorkshire and The Humber"
  data_region <- "London"

# ## USER INPUT ## Define aesthetics for map plot  ------------------------

  high_colour <- print("blue") #choose colour representing a high value

  mid_colour <- print("grey95") #choose colour representing a zero value

  low_colour <- print("darkorange") #choose colour representing a low value

  title <- print("white") #choose "black" to display the title, or "white" to not

  legend <- print("none") #choose "right" to display the legend, or "none" to not
```

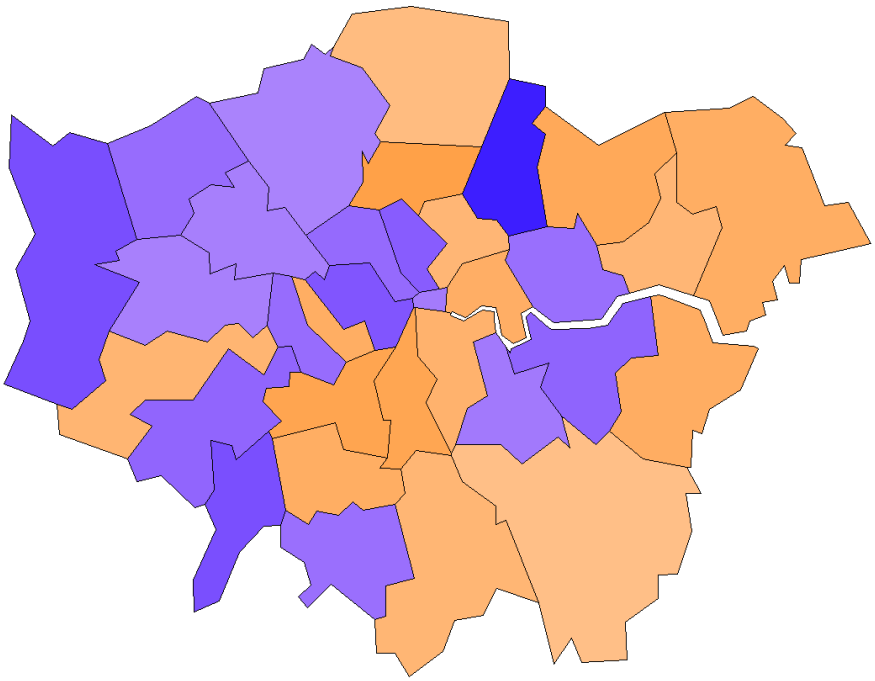
Example 4: Data

```
# ##USER INPUT## Load WITA output data ------------------------------------

  #Change file name where appropriate
  wita_output <- read.csv("Data/Templates/Data Template Example.csv")

# ##USER INPUT## Load WITA zone to LADs matrix ----------------------------

  #Change file name where appropriate
  wita_zones <- read.csv("Data/Templates/Zone Template Example.csv")
  wita_zones <- as.data.frame(wita_zones)

```

Example 4: Map without LAD Boundaries*



Example 4: Map with LAD Boundaries*



*The data used in these examples is fabricated and not based on any real-life scheme. To include LAD boundaries, see the Displaying LAD Boundaries section of this user manual.