Geocomputation
Complex Visualisations

Module outline

W1   Reproducible Spatial Analysis

W2   Spatial Queries and Geometric Operations

W3   Point Data Analysis

W4   Spatial Autocorrelation

W5   Spatial Models

W6   Raster Data Analysis

W7   Geodemographic Classification

W8   Accessibility Analysis

W9   Beyond the Choropleth

W10   Complex Visualisations

Core Spatial Analysis

Applied Spatial Analysis

Data Visualisation

This week

- Some pointers for the exam.

- Package and code management.

- Tidy data.

- Grammar of graphics.

Before we start

- Go to www.menti.com

- Use code: XXXX XXXX

Exam

Some pointers for the exam

Some pointers for the exam

- Two-hour timed exam scheduled for TBC.

- Two essays (1,000-word limit per question), aim to spend one hour on each

- Six questions to choose from a range of topics covered in the lectures

- No questions on coding or programming: focus on theory, underlying principles, methods and applications

- Q&A Session on TBC.

Some pointers for the exam

- Construct a clear argument: Ensure your essay includes a well-defined introduction, discussion, and conclusion.

- Maintain coherence: Structure your argument logically, progressing systematically from one point to the next.

- Provide evidence: Draw from readings and case studies to substantiate your argument, using books as technical references and articles to illustrate case studies.

Example question #1

- Both Thiessen Polygons and Inverse Distance Weighting (IDW) can be used for spatial interpolation. Explain how these techniques work and, with the use of a real-world example, elaborate on their strengths and weaknesses.

Example question #2

- The NHS wants to improve their ambulance response times and they have asked you to assess their current response times in London. Outline the data that you would request, the analytical steps you might take to complete this task and the final maps you could produce in response.

Some pointers for the exam

Marking Criteria for Exams[3]

	Inadequate	Adequate	Fair	Good	Excellent	Outstanding	Exceptional
	1-19: 5 inadequate 1-5 20-29: 4 inadequate 1-5 30-39: 3 inadequate 1-5	40-49: Majority of 1-5 Adequate or higher	50-59: Majority 1-5 Fair or higher	60-69: Majority 1-5 Good or higher	70-79: Majority 1-5 Excellent or higher	80-89: Majority 1-5 Outstanding or higher	90+: Majority 1-5 Exceptional
1. Response to the task set	Either no argument or argument is inept & irrelevant. Conclusions absent or irrelevant.	An indirect response to the task set, with gesture towards a relevant argument & conclusions.	A reasonable response with a limited sense of argument & partial conclusions.	A sound response with a reasonable argument & straightforward, logical conclusions.	A distinctive response with a clear argument & sensible conclusions; evidence of nuance.	Impressive response with nuanced argument, presenting significant & nuanced insights.	Exceptional response with very sophisticated argument ending in subtle conclusions.
2. Grasp of relevant issues	General misunderstanding of issues discussed.	Rudimentary, inter-mittent grasp of issues with confusions.	Reasonable grasp of issues & their broader implications.	Sound understanding, with insights into broader implications.	Thorough grasp w/ some sophisticated insights.	Striking grasp of complexities & significance of issues.	Exceptional grasp of complexities & issues' significance.
3. Reflection, thought, & conceptual framework	Erroneous or un-conceptual analysis. Poor understanding of basics of the taught materials.	Analysis relies on partial reproduction of taught materials. Some concepts absent or wrongly used.	Reasonable reproduction of taught materials. Rudimentary use of concepts.	Evidence of student's own analysis. Concepts effectively defined & used systematically.	Innovative analysis. Concepts deftly defined & used with some theoretical context.	Impressive thought, insights & analysis. Concepts deftly defined & used w/ strong context.	Engaged / cutting edge work. Sophisticated conceptual framework used in context.
4. Knowledge of literature & empirical topic	No evidence of, or largely inaccurate use of, conceptual literature & empirical material.	Rudimentary knowledge of required reading & empirical aspects of topic, with inaccuracies.	Limited knowledge of required reading & empirical aspects of the topic, with inaccuracies.	Sound knowledge of required reading & empirical aspects, with occasional inaccuracies.	Thorough knowledge of relevant reading & empirical aspects of the topic.	Impressive knowledge of relevant literature & empirical aspects of the topic.	Comprehensive knowledge of relevant literature & empirical aspects of the topic.
5. Evidence to support claims	No effective use of evidence to support any claims made.	Evidence rarely or ineffectively used to support claims.	A few claims warranted by evidence.	The essay's most crucial claims are supported by relevant evidence.	The essay's claims are mostly supported by appropriate evidence.	The essay's claims are warranted by apt, accurate evidence.	Claims are warranted by accurate, up-to-date & detailed evidence.
6. Structure & planning	Structure not discernible; minimal progression. (-3)	Structure discernible, but frequently absent. (-2)	Structure is apparent; frequent digression. (-1)	Conventional structure w/ rare digressions. (0)	Logical, coherent structure. (1)	Outstanding structure adds to overall effect. (2)	Exceptional structure crucial to argument. (3)
7.Writing style[4]	Style is consistently unclear. Inappropriate word choices. (-3)	Style & word choice lacks fluency & argument is only clear in parts. (-2)	Style generally clear but errors in use of jargon, grammar & spelling. (-1)	Style largely clear & fluent. Use of jargon is generally accurate. (0)	Style consistently clear & fluent with accurate use of terms. (1)	Style is elegant & precise with accurate use of jargon. (2)	Sophisticated style w/ impeccable spelling, grammar & jargon. (3)

[3] For an incomplete answer, apply the criteria for which any evidence has been provided (e.g., an essay plan as evidence of structure). An incomplete answer will not necessarily constitute a fail.

[4] The clarity of handwriting is not a marking criteria, however where handwriting is illegible work may have to be transcribed at student expense.
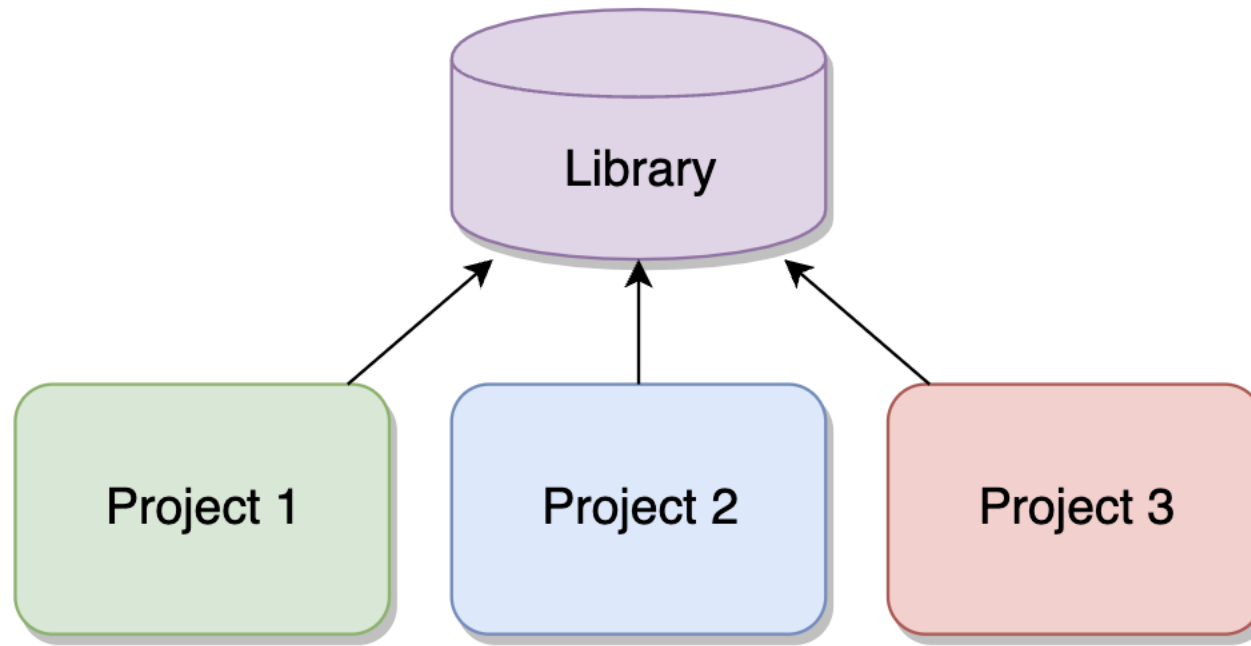
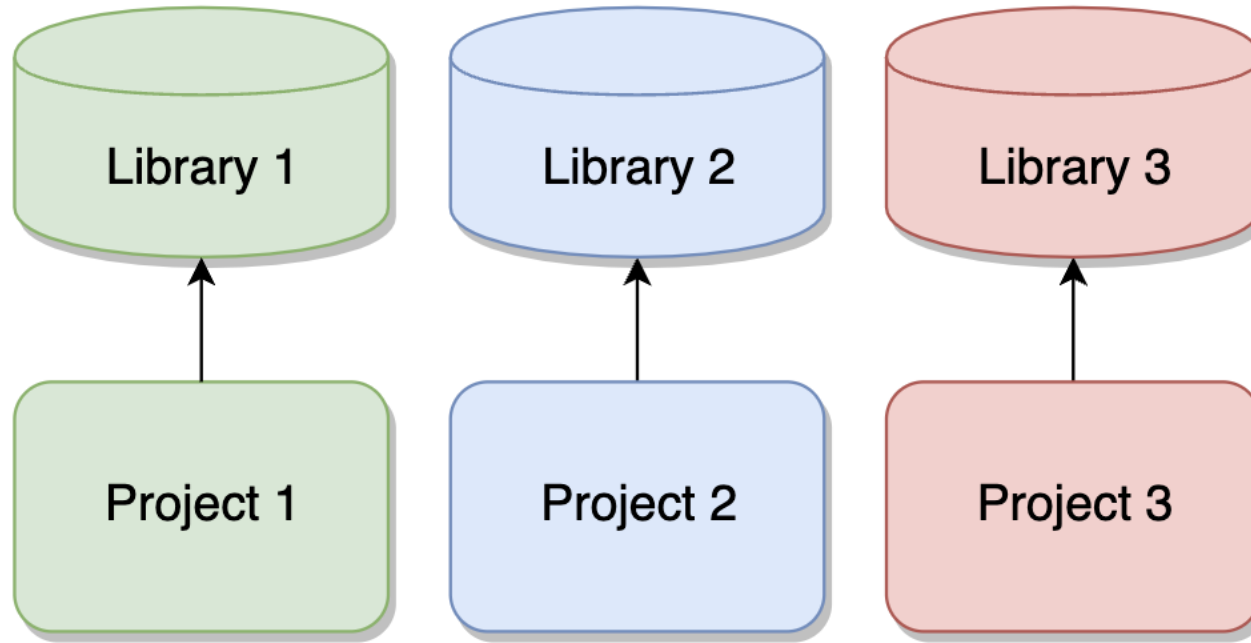# Package and data management

Package management

- Package management is the process of handling the many and varied dependencies and artifacts for your servers, applications, and developers.
- Toolkit used to manage project-local libraries.
- Combination with version management through Git.

Package management

Package management

Code management

- Version control: The practice of tracking and managing changes to software code.

- Example: Git (GitHub, GitLab).

- "Track changes" on steroids.

- Beyond the scope of this module but you can still set up your own small version control system (`"_v0.1"`, `"v_0.2"`, `"v_1.0"`).

Tidy data

Managing data

- Wickham 2014. 80 percent of your time goes to data cleaning and preparation ('data wrangling').
- Tidy data refers to the structure and organisation of your data set.
- The idea boils down to three principles.
- Brought together in the tidyverse.

Tidy data

Each variable must have its own column

Tidy data

Each observation must have its own row

Tidy data

Each value must have its own cell

Tidy data

	A	B	C	D	E	F	G	H	I
1	**MYE2: Persons by single year of age and sex for local authorities in England and Wales, mid-2022**								
2	This worksheet contains one table. Freeze panes are turned on.								
3	To turn off freeze panes select the 'View' ribbon then 'Freeze Panes' then 'Unfreeze Panes' or use [Alt W, F]								
4	Please choose from the links presented in the cells below to e-mail us your opinion on this table:								
5	This met my needs, please produce it next year								
6	I need something slightly different (please specify)								
7	This is not what I need at all (please specify)								
8	Code	Name	Geography	All ages	0	1	2	3	4
9	K04000001	ENGLAND AND WALES	Country	60,238,038	625,535	621,259	645,971	658,620	668,5
10	E92000001	ENGLAND	Country	57,106,398	596,306	592,565	615,537	627,205	635,7
11	E12000001	NORTH EAST	Region	2,683,040	25,453	25,572	26,261	27,574	28,1
12	E06000047	County Durham	Unitary Authority	528,127	4,649	4,696	4,876	5,020	5,1
13	E06000005	Darlington	Unitary Authority	109,469	1,066	1,066	1,131	1,129	1,1
14	E06000001	Hartlepool	Unitary Authority	93,861	905	941	1,021	1,040	1,0
15	E06000002	Middlesbrough	Unitary Authority	148,285	1,775	1,683	1,764	1,857	1,8
16	E06000057	Northumberland	Unitary Authority	324,562	2,555	2,744	2,708	2,919	3,0
17	E06000003	Redcar and Cleveland	Unitary Authority	137,175	1,323	1,196	1,323	1,378	1,4
18	E06000004	Stockton-on-Tees	Unitary Authority	199,966	1,971	1,948	2,149	2,222	2,2
19	E11000007	Tyne and Wear (Met County)	Metropolitan County	1,141,795	11,209	11,298	11,289	12,009	12,2
20	E08000037	Gateshead	Metropolitan District	197,722	1,921	1,942	1,958	2,036	2,0
21	E08000021	Newcastle upon Tyne	Metropolitan District	307,565	3,107	3,084	3,070	3,230	3,2
22	E08000022	North Tyneside	Metropolitan District	210,487	2,026	1,957	2,057	2,346	2,3
23	E08000023	South Tyneside	Metropolitan District	148,667	1,403	1,560	1,480	1,574	1,6
24	E08000024	Sunderland	Metropolitan District	277,354	2,752	2,755	2,724	2,823	2,9
25	E12000002	NORTH WEST	Region	7,516,113	78,957	78,368	81,625	83,737	84,0
26	E06000008	Blackburn with Darwen	Unitary Authority	155,762	1,904	1,960	2,018	2,096	2,0

Common errors

- Column headers are values rather than variable names.

- Multiple variables are stored in one column.

- Variables are stored in both rows and columns.

- Multiple observational units are stored in the same column.

- A single observation is stored in multiple tables.

Tidy ?

country	year	type	count
Afghanistan	2019	cases	745
Afghanistan	2019	population	19 987 071
Afghanistan	2020	cases	2 666
Afghanistan	2020	population	20 595 360
Brazil	2019	cases	3,7737
Brazil	2019	population	172 006 362
Brazil	2020	cases	80 488
Brazil	2020	population	174 504 898
China	2019	cases	212 258
China	2019	population	1 272 915 272
China	2020	cases	213 766
China	2020	population	1 280 428 583

Tidy ?

country	year	rate
Afghanistan	2019	745 / 19,987,071
Afghanistan	2020	2,666 / 20,595,360
Brazil	2019	3,7737 / 172,006,362
Brazil	2020	80,488 / 174,504,898
China	2019	212,258 / 1,272,915,272
China	2020	213,766 / 1,280,428,583

Tidy ?

Cases

country	2019	2020
Afghanistan	745	2 666
Brazil	3,7737	80 488
China	212 258	213 766

Population

country	2019	2020
Afghanistan	19 987 071	20 595 360
Brazil	172 006 362	174 504 898
China	1 272 915 272	1 280 428 583

Tidy ?

country	year	cases	population
Afghanistan	2019	745	19 987 071
Afghanistan	2020	2 666	20 595 360
Brazil	2019	3,7737	172 006 362
Brazil	2020	80 488	174 504 898
China	2019	212 258	1 272 915 272
China	2020	213 766	1 280 428 583

Grammar of graphics

Grammar of graphics

- Graphics are constructed by layering multiple elements of data.

- Values in a dataset serve as aesthetics: attributes that can be visually represented in a graphic.

- Data, scales, coordinate systems, and plot annotations are layered on top of these aesthetics to create the final graphic.

Grammar of graphics

Dataframe

x	y	Shape
2	4	a
1	1	a
4	15	b
9	80	b
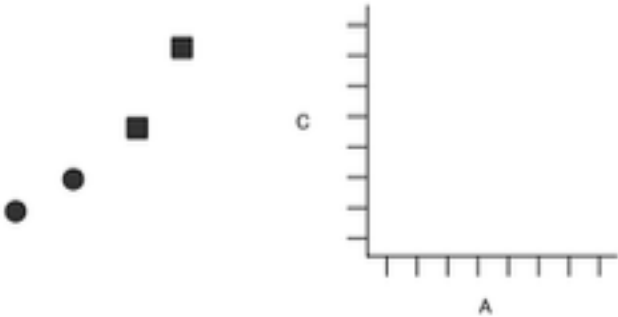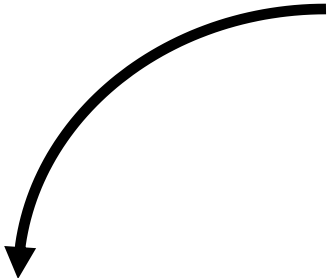
Dataframe
values
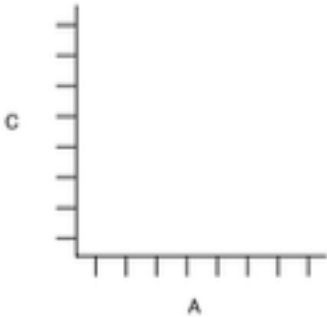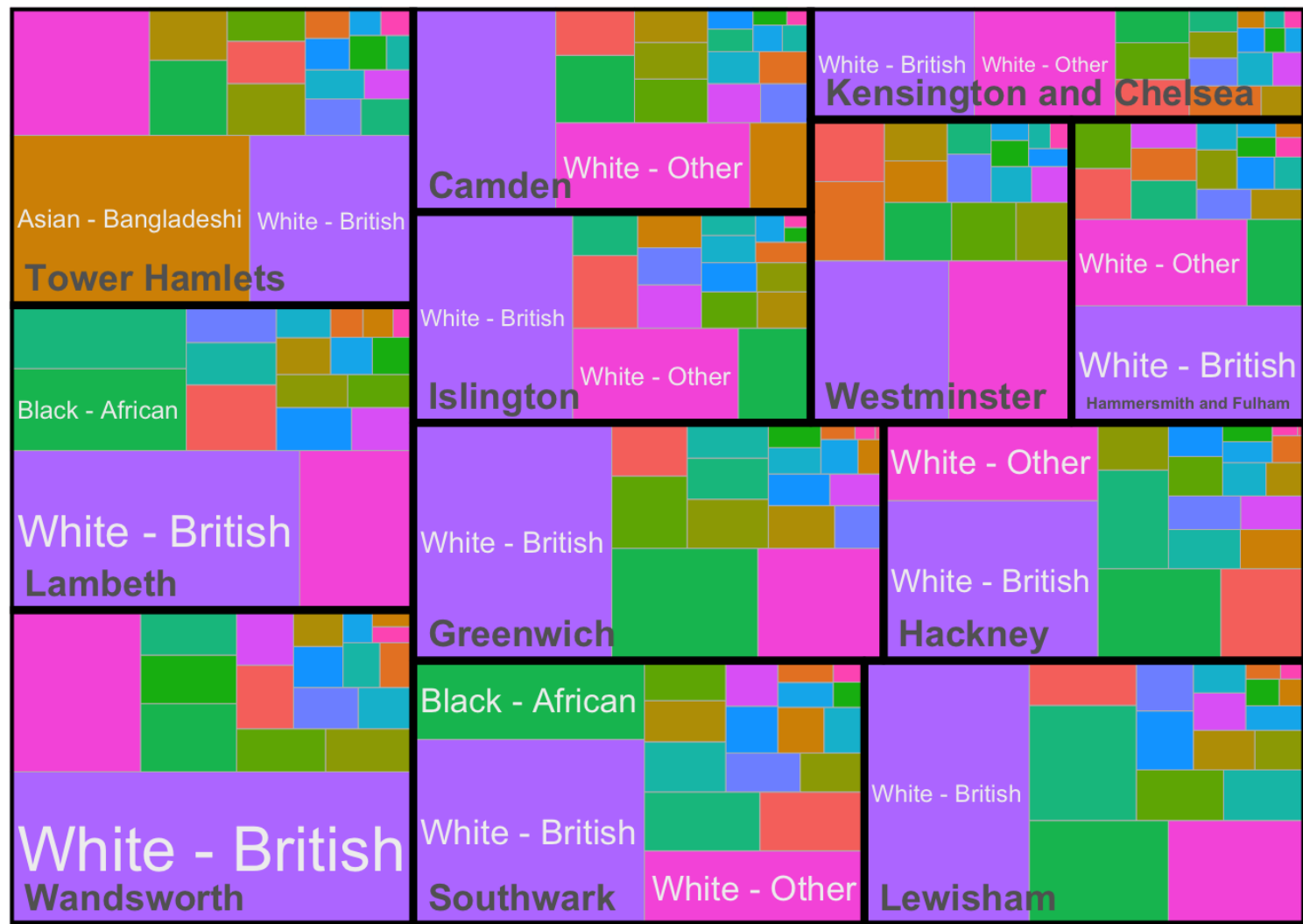
Dataframe
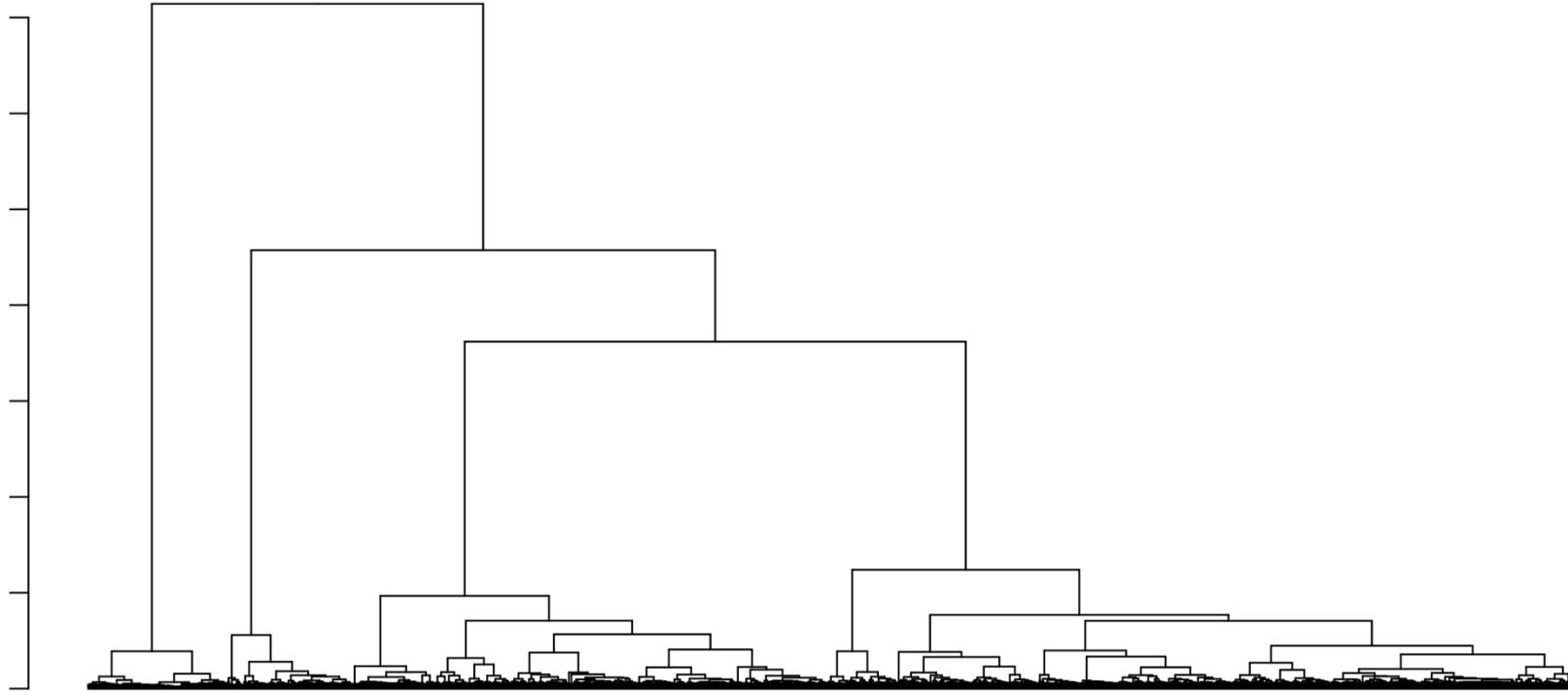scale

Dataframe
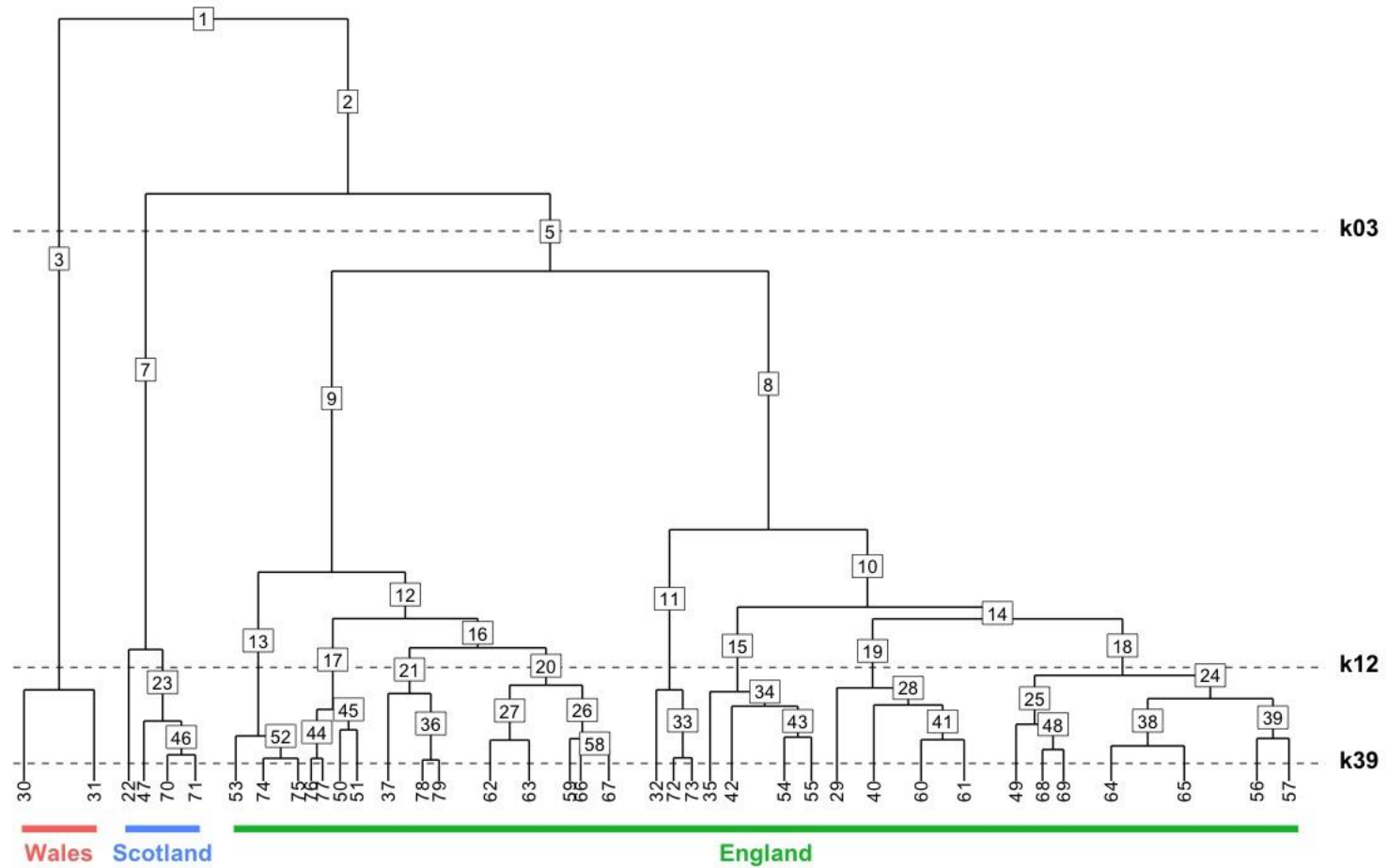annotations
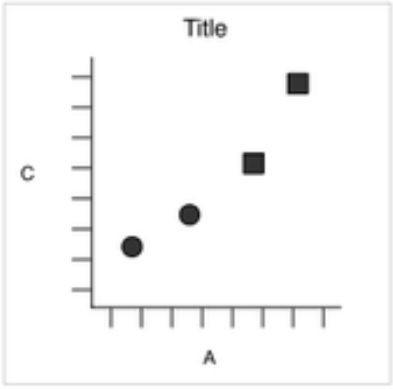
Final
Graphic

Grammar of graphics
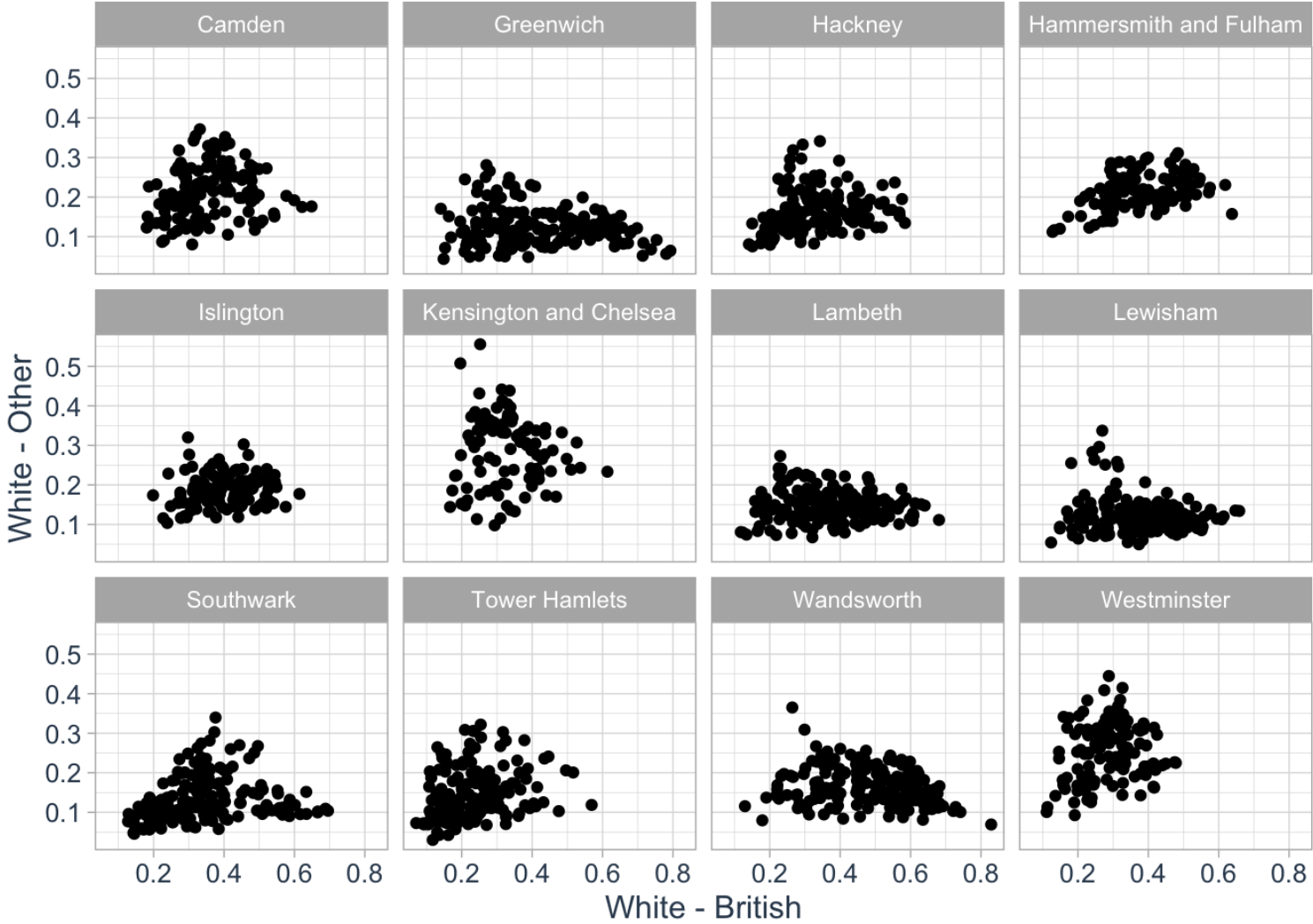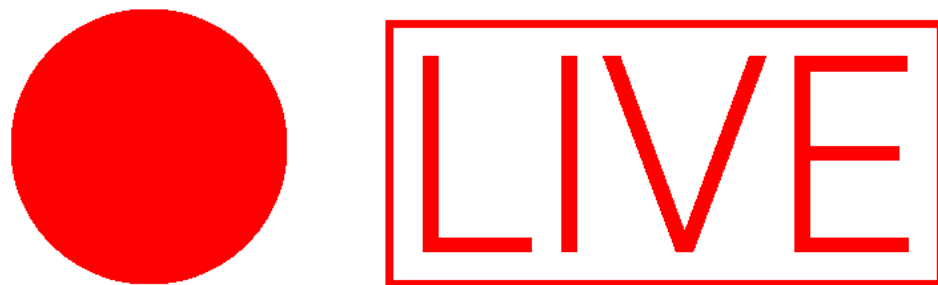
Grammar of graphics

Grammar of graphics

Grammar of graphics

RStudio

Everything we covered

Everything we covered

1) Geocomputation as a GIS 2.0: working with geographic data in a computational way, focusing on code, reproducibility and modularity.

2) Spatial queries and geometric operations: the core of spatial analysis.

3) Working with point event data: special attention to clustering and visualisation of these using DBSCAN and Kernel Density Estimation.

4) The First Law of Geography in action: measuring spatial autocorrelation.

5) Dealing with spatial autocorrelation and non-stationarity: spatial models

Everything we covered

6) Dealing with raster data: spatial data interpolation.

7) Geodemographics: analysis of people by where they live.

8) Measuring accessibility: working with a digital network.

9) Positioning the map: Coordinate Reference Systems.

10) Complex Visualisations: using ggplot2 to build a graphic layer by layer.

Questions

Justin van Dijk

j.t.vandijk@ucl.ac.uk

Thank you!