MA22019 2025 - Problem Sheet 8

Analysis of point pattern and lattice/areal data

Overview

This week's problem sheet focuses on the methods for analyzing point pattern and lattice/areal unit data in Sections 4.5-4.7 in the lecture notes. Exercises 1-2 help you with revising the content of the lecture in Week 9, and Tutorial Questions 1 and 2 provide some more advanced questions. You can check your answers for Exercises 1 and 2 using the quiz provided on Moodle.

Your answer to the Homework Question can be submitted on Moodle to your tutor for feedback. The submission deadline is 17:00 on Thursday 24 April 2025. You should submit a single PDF or Word file that provides your R code, any created R output and all your comments.

You may want to load the following packages before starting the exercise:

```
library( dplyr )
library( ggplot2 )
library( tidyr )
library( sf )
library( ggspatial )
library( prettymapr )
library( spatstat )
```

When working on a University PC, you will have to first install some of these packages.

Exercise 1 - Swine fever outbreaks across Europe

The file "Outbreaks.csv" contains the data for animal disease outbreaks between September 2015 and August 2017 as provided by the EMPRES Global Animal Disease Information System. Each entry in the data set specifies the location, date and type of disease for an individual outbreak. In this question we want to focus on the African swine fever that has affected large parts of Europe in the past years. Address the following three research tasks / questions and then complete the Moodle quiz.

a. Which three countries were the worst affected in terms of the total number of outbreaks of African swine fever?

b. Visualize the locations of outbreaks of African swine fever for the Baltic states (Estonia,

Latvia and Lithuania). What do you conclude?

c. Latvia has recorded a high number of outbreaks. Identify the parts of Latvia that observed a high frequency of swine fever outbreaks using quadrat counting or the kernel smoothed intensity function. Is it reasonable to conclude that the point process describing outbreaks of African swine fever across Latvia is homogeneous?

Exercise 2 - Crime rates across France

The file "Crimes France.csv" contains crime statistics from 2015 for all French départments (except Corsica). The file uses the UTF-8 encoding and you should ensure to specify this when loading the file. A shapefile for France with the département boundaries is provided in the file "gadm41_FRA_2.shp". Use the techniques for areal unit data to complete the following tasks and then go to Moodle and answer the quiz questions.

a) Import the shapefile and create a map for France which includes the boundaries for the départments.

b) Visualize the rate of violence per 1000 people per département. Which areas should we avoid when we are concerned about violence?

c) Visualize the rate of burglaries per 1000 people per département. What do you conclude?

Tutorial Question 1 - Crime across Utopia

Utopia's police department needs your help with analyzing their 2015-2021 data regarding certain crimes. The data is provided in the file "UtopiaCrimes.csv". Note, longitude and latitude coordinates are only available for drug possession offences across District 44.

Utopia consists of 59 districts and a shapefile of Utopia is provided as "UtopiaShapefile.shp". To hide Utopia's location, constants have been added to the latitude and longitude coordinates, but the shapes they define are correct. The population for each district is provided in the file "UtopiaPopulation.csv".

a) Identify the three most common crimes in Utopia for the period 2015-2021.

b) Visualize the rate per 1,000 population for the most common crime for 2015-2021 for the different districts. What do you conclude from your plot?

c) You are told that District 44 is notorious for drug possession. The police is planning to conduct a raid to tackle the issue, but they are unsure which areas of the district are the most seriously affected. Use spatial data analysis techniques to identify the parts of District 44 they should be focusing on. **Hint:** You may want to consider the function **st_reverse()** should you get an error message regarding orientation/direction.

Tutorial Question 2 - Cost of living crisis in Texas

As many other countries around the world, the United States have seen high inflation over the past years. In the following we want to analyze socioeconomic data for the state of Texas. We have access to two data files:

- "Cost_Texas.csv" gives the cost for essentials (food, rent, healthcare,etc.) in 2023 for a single household for each county in Texas based on the Family Budget Calculator by the Economic Policy Institute.
- "Income_Texas.csv" gives the median income per capita for each county in Texas as reported by the Bureau of Economic Analysis of the US government for 2021.

Perform the following tasks and answer the research questions:

a) Load the two data files and calculate the difference between median income and cost of living for each county.

b) The file "Texas.Rdata" provides a shapefile of Texas, called **Texas**, with all county boundaries. Use the shapefile to create a map which visualizes the variable calculated in part a). What do you conclude?

c) How reliable are our results found in part b)? **Hint:** You may want to extract the two counties with the highest median income and see what you find out about them on Wikipedia.

Homework Question - Tracking grey squirrels across the UK

The grey squirrel is classified as an invasive species in the UK, and it has displaced the native red squirrel across large parts of the UK. A wildlife conservation charity has collected data on reported sightings of grey squirrels for 2020-2022. The data they provided includes the following:

GreySquirrels.csv: Sightings of grey squirrels reported to the wildlife conservation charity for 2020-2022:

- Year: Year the sighting happened
- Lon: Longitude coordinate of the sighting
- Lat: Latitude coordinate of the sighting
- Admin: Name of the administrative area, where the grey squirrel was sighted

UK Shapefile: Folder containing shapefiles for the UK

- UK.shp: Shapefile without administrative boundaries
- UK admin.shp: Shapefile with boundaries for the administrative areas named in the file "GreySquirrels.csv"

The charity assured us that the data is representative of the spatial distribution of squirrels across Great Britain for all years. They ask you to use the data to investigate the following aspects:

a) What can we say about the spatial distribution of grey squirrels across Great Britain in 2022?

b) Are there any areas of Great Britain that saw a notable change in the number of grey squirrels when we compare the data for 2020 and 2022?