



Master of Science in Economics
ECON6027 Spatial Econometrics and Data Analysis
10 November 2022
Final Examination

Instructor: Dr. Shew Fan Liu

Student's Name: _____ (in full)

Student's ID Number: _____

READ ALL INSTRUCTIONS CAREFULLY

- 1 For each question, please provide separate submissions of the following:
 - a. Html outputs
 - b. R script file
 - c. Any other files that you create (eg: .txt, .GAL, etc.)
 - d. File name to follow the question number (eg: "Q1.html", "Q1.Rscript", etc.)
- 2 Do not compress your submissions.
- 3 In case you have trouble compiling, copy+paste the outputs into a word document and prepare for submission.
- 4 Label the answers to the sub-questions clearly.
- 5 This paper consists of 5 printed pages including this cover page.
- 6 The examination has a 40% component in the course assessment. Please answer **all** questions. The marks for each part are indicated in brackets.
- 7 You have **three hours** to complete the paper.

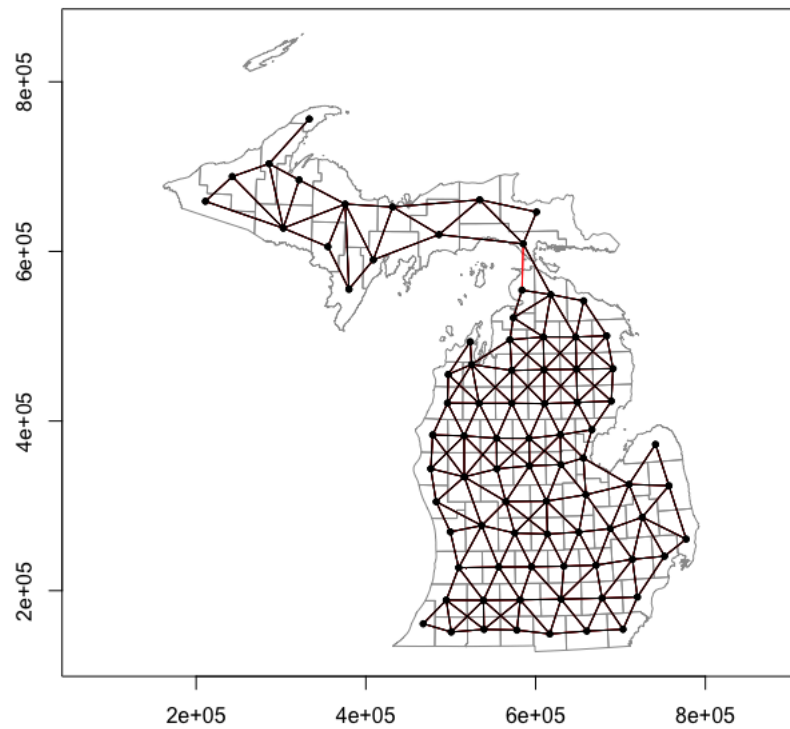
Question 1: Areal Data (40 marks)

Shared with you are two files:

1. “covid.Rdata” gives various covid related observations aggregated at county level in Michigan taken from [New York Times](#). The variables are:
 - i. TOT_CASES: total cases
 - ii. CASES_100K: cases per 100,000 population
 - iii. TOT_DEATHS: total deaths
 - iv. DEATHS_100K: deaths per 100,000 population
 - v. PCT_VACC: vaccination take-up rate.
2. “mi_county” shapefile for the boundaries. Michigan is the only state to consist of two peninsulas. The Lower Peninsula is shaped like a mitten. The Upper Peninsula (often called “the U.P.”) is separated from the Lower Peninsula by the Straits of Mackinac, a five-mile (8 km) channel that joins Lake Huron to Lake Michigan.

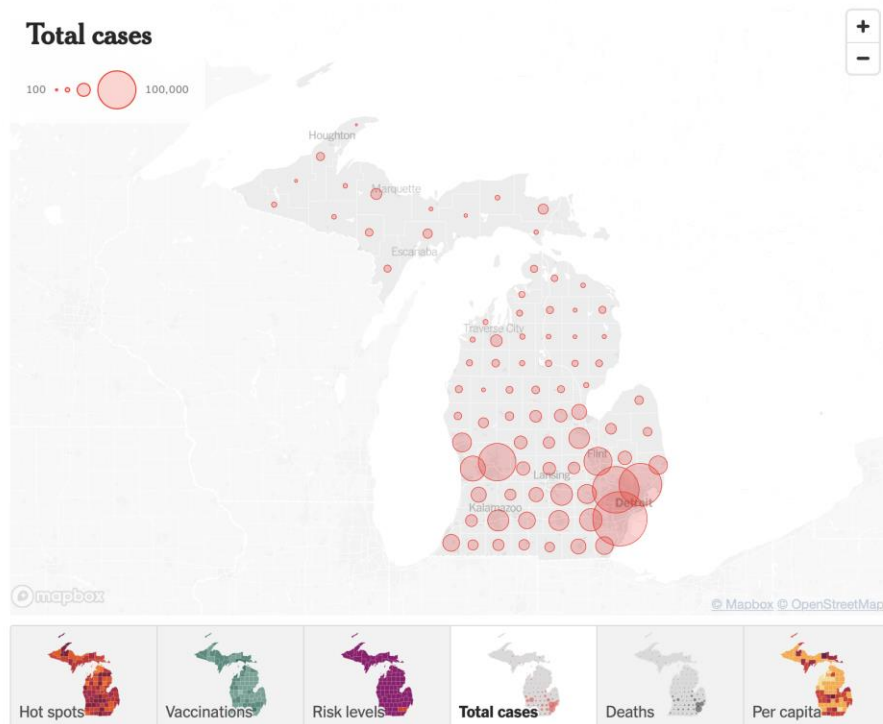
Answer the following questions:

- a) Load the Michigan county shapefile and give a “nb” connectivity plot based on the queen contiguity criterion. (4 marks)
- b) Notice that the “nb” object does not connect the counties “**Mackinac**” and “**Emmet**”. Make this connection. Your final “nb” object to be used in this question should be as given in Figure 1.1. This plot must be given along with the summary of the “nb” object. (4 marks)
- c) Load the “covid.Rdata” file and give it spatial awareness using the Michigan county shapefile. (*Hint: You will need to manipulate the “by” argument when mutating joins.*) (4 marks)
- d) Using a Moran’s I statistic, give a comparison of the spatial autocorrelation of the variables “**CASES_100K**” and “**DEATHS_100K**” in each county. (4 marks)
- e) Construct an appropriate plot to identify any significant local cluster(s) of counties with strong spatial spill-over effects in the “**TOT_CASES**”. What is the geographic/economic significance of the cluster(s) you observe? (4 marks)
- f) How many spatial outliers are there for the variable “**TOT_DEATHS**” per county? (4 marks)
- g) What is the most suitable lag order for the variable “**DEATHS_100K**”. (4 marks)
- h) Suppose you wish to **regress** the variable “**DEATHS_100K**” on “**PCT_VACC**”.
 - i. Justify the use of spatial regression models in this case. (4 marks)
 - ii. Compare suitable (spatial) regression models and identify the most suitable model. Use a 5% level of significance in the tests. (4 marks)
- i) Figure 1.2 gives a screen grab of an interactive map as given in The New York Times (it is still there!). Re-create a similar interactive map. (*A legend is not required.*) (4 marks)

Figure 1.1**Queen Contiguity with the coerced connection in Red****Figure 1.2**

The New York Times

U.S. | Coronavirus in the U.S.: Latest Map and Case Count



Question 2: Geospatial Data (30 marks)

Shared with you are two files:

1. “sg_air.Rdata” contains various air quality indicators from 11 location in Singapore.
2. “sg_regions.geojson” contains the region boundaries of Singapore.

Figure 2.1



Figure 2.1 shows the live map taken from <https://www.iqair.com/> showing the Air Quality Index at 11 monitoring stations in Singapore on 2nd of November 2022. Three other Malaysian monitoring locations closer to Singapore border can also be seen on the map. “sg_air.Rdata” contains the dataset related to these observations. The main indicator is an attribute called US-AQI (click [here](#) to learn more about Air Quality Indices) which is an index computed using the usual pollutants such as ozone, nitrogen dioxide, sulphur dioxide, etc. Your task is to recreate the surface of the US-AQI using geospatial analysis methods.

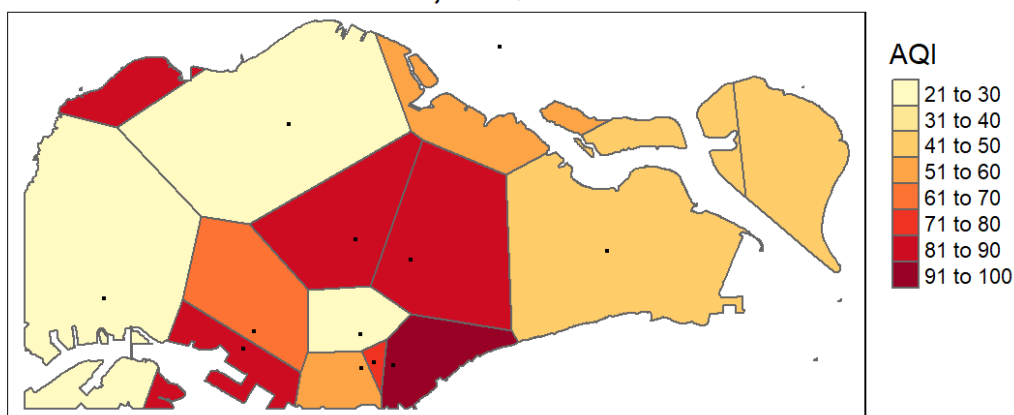
Answer the following questions:

- a) Give spatial awareness to the “sg_air.Rdata” dataset. Prepare your dataset for analysis and give a basic plot of your dataset (5 marks)
- b) Re-create the Voronoi Tessellation given in Figure 2.2. (5 marks)
- c) Reconstruct the surface using the inverse distance weighting method where the “idp=1” and “idp=3”. Your output must show the head of the interpolated estimates and plots of the estimates. Which idp is more appropriate? (5 marks)
- d) Use the `autofitVariogram()` function to find the most suitable parametric model fit for the empirical semi-variogram. Use this fitted model to predict the US-AQI using ordinary kriging method. Your final answer should include (5 marks):
 - i. Plot of the fitted variogram.
 - ii. Summary of the predictions.
 - iii. Plot of the reconstructed US-AQI indicator over Singapore.

- e) Compare the variogram fits using the models: spherical, wave, Matern and pentaspherical. (5 marks)
- Select the model fit that includes a “nugget effect”.
 - Re-do the answers to part (d) using the new parametric fit.
- f) Compare the variances of the estimates in parts (d) and (e) and hence identify the best estimate. Your answer may include comparisons of maps, histograms, descriptive statistics, etc. (5 marks)

Figure 2.2

Voronoi Tessellation, AQI



Question 3: Point Data (30 marks)

Shared with you are two geojson files.

- “empty.geojson” gives the locations of empty homes in Plymouth, UK and
- “plymouth.geojson” is a file containing the neighbourhood boundary of Plymouth, UK.

You are required to prepare a report conducting a comprehensive point pattern analysis of the pattern of empty homes in Plymouth.

Label your outputs clearly and remember to give your interim conclusion at each stage and your overall conclusion. Your analysis can include various tests and plots to support your conclusions.