University of North Texas
Department of Geography and the Environment

GEOG 5190

LAB 2: SPATIAL PATTERN AND SPATIAL AUTOCORRELATION

Part A: Nearest Neighbour Analysis

Scenario: you are a physical geographer studying the pattern of debris laid down by a glacier in Glacier National Park, Montana. Based on a thorough understanding of the relevant research literature, you have developed a model that suggests that there is a process at work with your particular glacier that should make the debris pattern you observe not random, but tending towards uniformity.

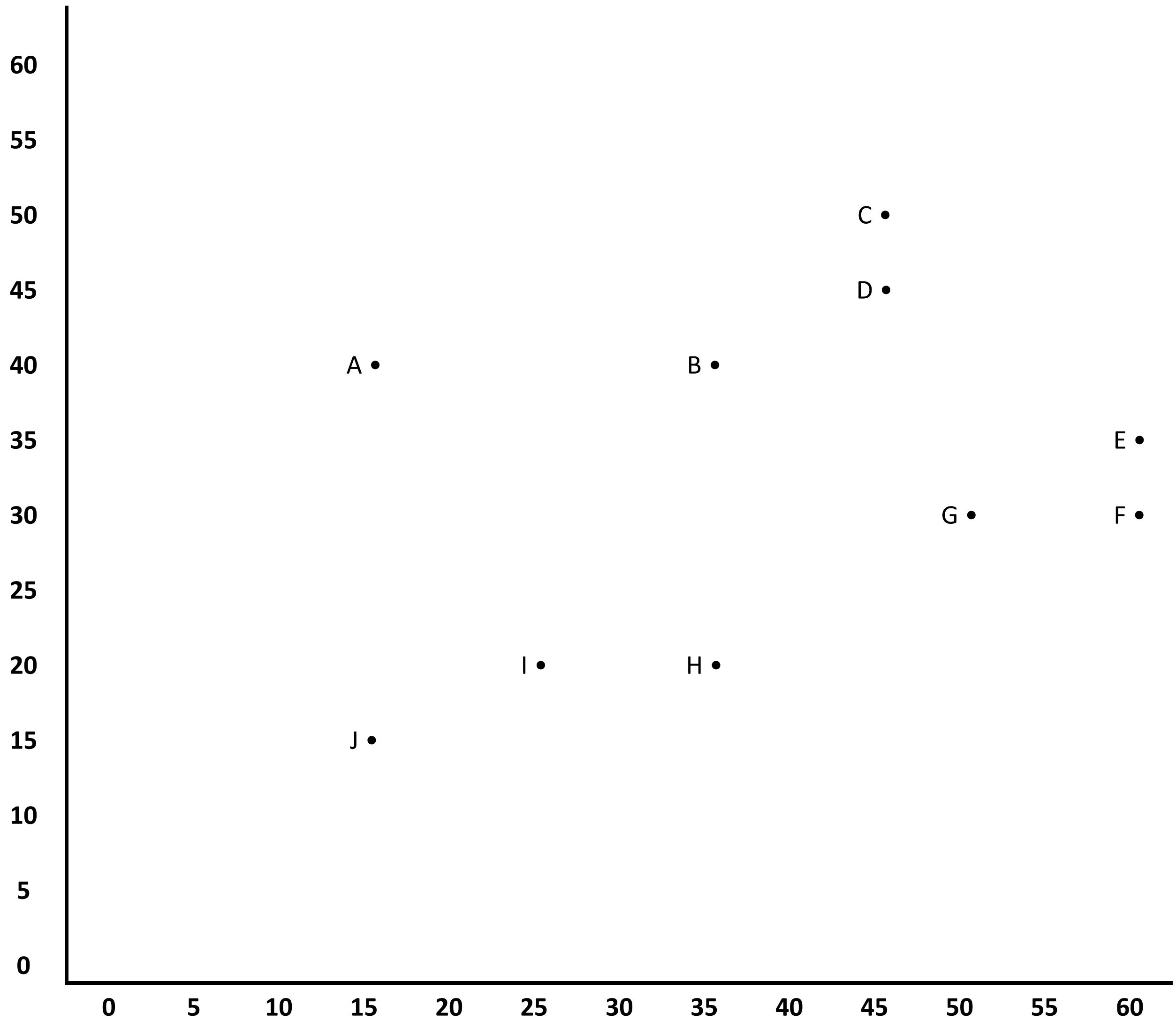
You want to test your model using debris data you have collected from the glacier. To test your model, you plan to do a *nearest neighbour analysis* on the point pattern data you have generated from your observations. The debris map included on the following page graphically represents the spatial distribution of debris in your dataset.

The distance matrix for the 10 debris observations included on this map is as follows (all numbers are in meters):

	A	B	C	D	E	F	G	H	I	J
A	0	20	32	31	47	48	38	30	23	25
B	20	0	14	12	21	28	19	21	23	33
C	32	14	0	5	21	25	21	32	31	42
D	31	12	5	0	17	22	16	22	32	43
E	47	21	21	17	0	6	11	30	38	50
F	48	28	25	22	6	0	10	23	28	48
G	38	19	21	16	11	10	0	19	27	39
H	30	21	32	22	30	23	19	0	10	20
I	23	23	31	32	38	28	27	10	0	11
J	25	33	42	43	50	48	39	20	11	0

The size of the rectangle that actually contains these debris observations (*i.e.* the "enclosing rectangle" we discussed in class) is 35 by 45 meters.

Map of Glacier Debris Study Area
Nearest Neighbour Analysis



As in lab 1, follow the six steps of a formal statistical test to complete your nearest neighbour analysis. The following walks you through each of the six steps, along with some items for you to consider for each step:

1. **Formulate hypotheses (H_0 and H_1)**
    a. State each of these hypotheses clearly. Both are necessary.
    b. Choose an appropriate H_1 (either one- or two-tailed), based on the information provided in this lab.
    c. Be sure to explain why the form of alternative hypothesis you have chosen is the best choice for this situation:
        i. If you choose a one-tailed test: you will need to explain why a one-tailed test is most appropriate. Also, you will need to specify a direction and again justify why this direction is most appropriate.
        ii. If you choose a two-tailed test: you will need to explain why this is the most appropriate choice. This is a simple case, so no further discussion is needed beyond the basic justification if you choose the two-tailed option.

2. **Select a test statistic**
    a. You know you are using a nearest neighbour test, so you can simply state this here (please do that for a complete record).

3. **Select your significance level and whether to use a one- or two-tailed test**
    a. Please use the 0.05 level (again, please state that here for completeness).

4. **Look up the critical value for acceptance/rejection of your H_0**
    a. Use the critical value table provided below, and note the correct critical value for your situation:

Critical Values of a Standard Normal Deviate z

	Significance level (one-tailed)				
	0.1	0.05	0.01	0.005	0.001
z	1.282	1.645	2.326	2.576	3.090
−z	−1.282	−1.645	−2.326	−2.576	−3.090

	Significance level (two-tailed)				
	0.1	0.05	0.01	0.005	0.001
z	1.645	1.960	2.576	2.813	3.291
−z	−1.645	−1.960	−2.576	−2.813	−3.291

5. **Compute your test statistic from your data**
   a. For this nearest neighbour analysis, follow these steps:

      1. Show the nearest neighbour pairs in the matrix and link them up graphically on the map included with this lab. Draw the "enclosing rectangle" on the map as well. Use clear labels.

      2. Calculate the value of $\bar{d}_{obs}$ (the average inter-point distance) for these point pattern data. Show your calculation, including the original equation for $\bar{d}_{obs}$.

         Please remember: every time you are asked to show your calculations in this course, please show the original equation and all of your work.

      3. Calculate the value of $\bar{d}_{ran}$ (the average inter-point data we would expect if a random process was at work) for these data. Show the equation and your calculation.

      4. Calculate the value of *R* (the nearest neighbour index). What does this value (on its own) tell you about the point pattern? Show the equation, your calculation, and provide your interpretation of the result.

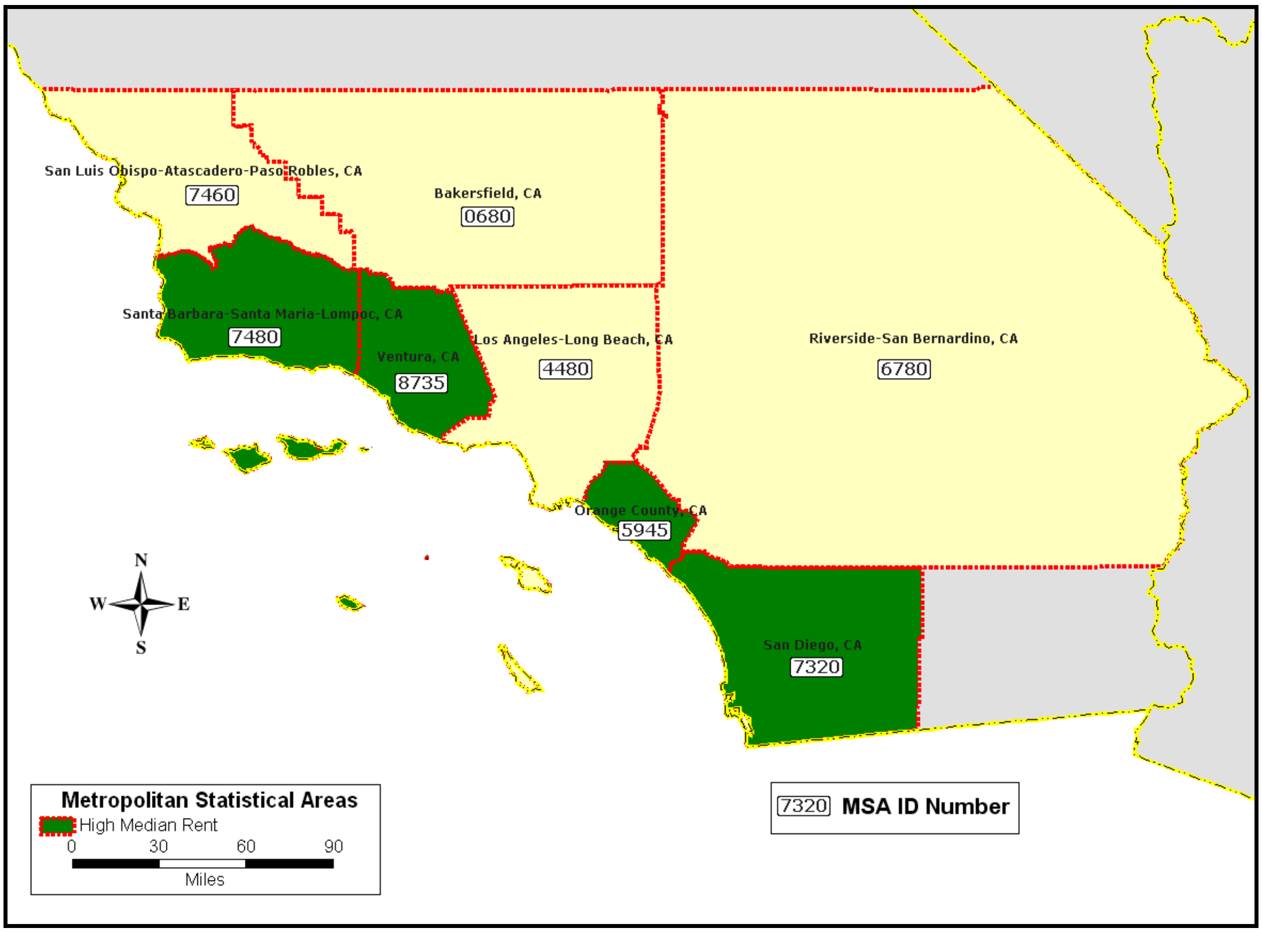
      5. Calculate *Geary's C*, the value of the test statistic for this nearest neighbour analysis. Show the equation and all steps of all of your calculations (including your calculation of the standard error of the average distance, $SE_{\bar{d}}$).

6. **Make a decision: accept/reject H_0?**
   a. What can you say now about the map pattern? Be clear about your conclusion, and explain it well.

Overall, please include full (but not necessarily lengthy) explanations for every judgment and decision you make as you complete the six steps.

Part B: Join Count Statistics

Scenario: as a geographer with expertise in spatial statistics, you have been hired by the state of California to do research into the economic characteristics of southern California. You have collected statistics for each of the metropolitan areas that make up the southern California region. In the map below, shaded areas (such as San Diego, MSA ID number 7320) represent metropolitan areas with <u>high rents</u>, while light areas (such as Bakersfield, MSA ID number 0680) represent <u>low rents</u> (unlabelled areas to the north and east are non-metropolitan, and fall outside of the study).



Your intent is to <u>measure the spatial autocorrelation in the map</u> in order to quantify the degree of clustering or dispersion present. Your expectation is that the arrangement is <u>not random</u>. Given this scenario and your data as described above, you've chosen to employ *join count statistics* to achieve your goal of analyzing the spatial pattern in your dataset.

Follow the six steps of a formal statistical test to complete your join count analysis. The following walks you through the six steps, along with some items for you to consider for each step:

1. **Formulate hypotheses (H_0 and H_1) and decide whether to use a one- or two-tailed test**
    b. State each of these hypotheses clearly. Both are necessary.
    c. Please state and explain your tail selection.

2. **Select a test statistic**
    a. You know the join count part, but which option is most appropriate for you to employ in this situation, <u>free sampling</u> or <u>non-free sampling</u>? Explain fully (but this explanation does not need to be lengthy).

3. **Select your significance level**
    a. Please use the 0.05 level for your analysis.

4. **Look up the critical value for acceptance/rejection of your H_0**
    a. Please the use the critical value table provided in part A of this lab.

5. **Compute your test statistic from your data**
    a. Determine the value of O_{BW} from the map. Note: it might be difficult to see on the metropolitan area map, but <u>please assume that MSAs 7480 and 0680 do not join.</u> Also, for this analysis, please assume that all corner-to-corner joins do not count (*i.e.* you are implementing the "<u>rook's case</u>" discussed in class).
    b. Calculate the appropriate value of E_{BW} for either <u>free sampling</u> or <u>non-free sampling</u> (following from your choice made in step 2 above; note that your class notes refer to the non-free sampling calculations, but your reading includes the equations for both free sampling and non-free sampling). Show the E_{BW} equation and all steps of your calculation.
    c. Calculate the value of z. Show the equation and your calculation. **You are given that σ_{BW} = 1.42**: although you have all the information you need to do this calculation, <u>please do not do this σ_{BW} calculation</u> – please simply use this value and save yourself a lot of time.

6. **Make a decision: accept/reject H_0?**
    a. Be clear about your conclusion. Explain the conclusion itself clearly, as well as what it means.

Overall, please include full (but not necessarily lengthy) explanations for <u>every judgment and decision</u> you make as you complete the six steps.