Geography and Affect.

Urban Spatial Analytics

Pratik P. Watwani
(ppw220)

https://github.com/pratikwatwani/Spatial-Affect-Analysis

Introduction

This project aims to test the hypothesis if geography has an impact on anxiety of an individual. In this particular situation, the spatial analysis is conducted in the Metrotech Commons are, an urban plaza which houses school buildings, cafes, restaurants, as well as industrial buildings. Any or all of these could be a factor in influencing such emotions in an individual. Furthermore, the project performs spatial autocorrelation using Morans I, thus calling for a good understanding of the underlying theory.

Data

Procedurally the data observations were collected only during the day when there was a significant proportion of the population in the MetroTech Commons, considering that most of the classes run during the day time, it was much more efficient to thus collect the data at that time.

There were 5 distinct data observations that were taken on Tuesdays, Wednesdays and Thursdays for two weeks for one particular hour (9:00 AM - 10:00 AM) that resulted in a significant number of observational points. A few data observations were also collected during night time and also during Thanksgiving and Black Friday, however, the data was not sufficient enough, as it lacked observational points.Thus this data was not considered during the plotting and testing phase of this project, as it would significantly change results and testing results.

Of all the data, only a small portion was then used in the analysis, this is due to the fact that this project considers facial expressions as an underlying observational limitations. Thus, only those individuals are considered which demonstrated a facial expression which could help me to determine if they feel anxious or not.

Photos, as well as video recordings, were taken to keep a track of the individuals and their locations, with the help of iPhone, with 0.5x wide angle lens, thus allowing to collect more data from eleven positions in the metrotech commons. However, not all individuals in the plots carry a geocoordinate, they are added to the data set relative to their positioning to the created grid structure in ArcMap. The data collection required a significant period of observation to get a gist of the current movement period.

Location.csv Events

	Latitude	Longitude	Name	Shape *
▸	40.694449	-73.98603	Library	Point
	40.694192	-73.98617	MetrotechCommon-MidLeft	Point
	40.693844	-73.9862	MetrotechCommon-LowerLeft	Point
	40.693878	-73.98602	MetrotechCommon-Triangle-LowerLeft	Point
	40.69385	-73.985346	MetrotechCommon-Triangle-LowerRight	Point
	40.693938	-73.985684	MetrotechCommon-Triangle-Center	Point
	40.694118	-73.98567	MetrotechCommon-Triangle-Peak	Point
	40.694286	-73.985757	MetrotechCommon-Walk-UpperCenter	Point
	40.694265	-73.985267	MetrotechCommon-Walk-UpperRight	Point
	40.694332	-73.985081	MetrotechCommon-Wasserman	Point
	40.693815	-73.985191	MetrotechCommon-MyrtleAve	Point

Geocoordinate locations for image observations.

The below figure projects the path taken to collect video observations, maximizing observation and minimizing maneauver in the area.

Direction flow and starred locations for video and image collection.

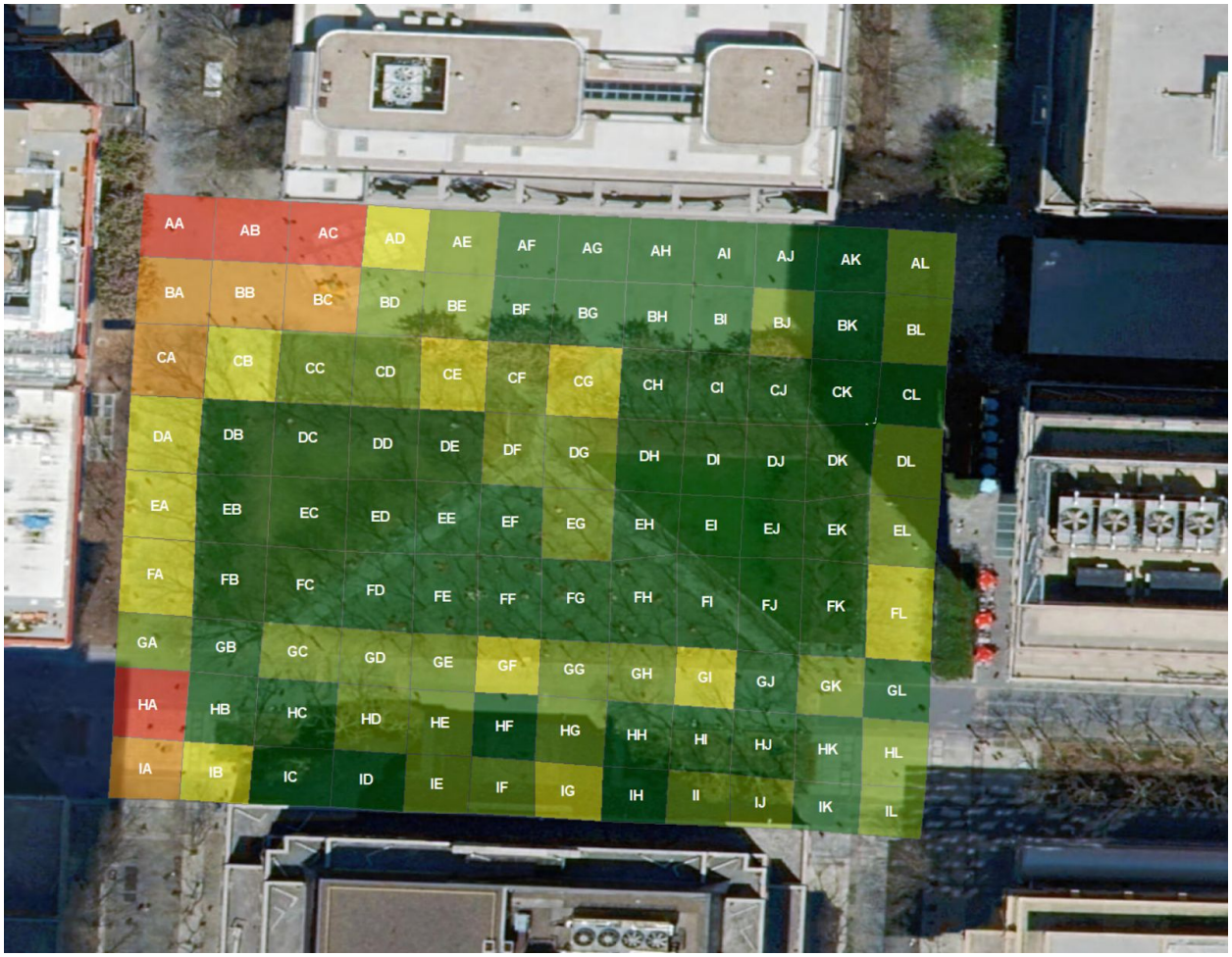
Methodology

The data collection method involved the recording of geo-coordinates, imaging as well as video recording which was then referenced to plot individual entities in ArcMap.

Sr. No.	Layer Name	Type
1	Image Observations	Points/Multi points
2	Video Observations	Polyline
4	Grid Pattern	Polygon
5	Location	Events txt

Data that specified XY (longitude and latitude location) was imported via Add Data Set and referenced accordingly.

A 9x12 grid pattern was layered or fused over the satellite basemap provided by Arc organization. This grid pattern helps to perform a spatial correlation analysis with value that each box in the grid.



A 9x12 grid structure fused on top of metrotech commons.

To be more precise about what each box in the grid holds can be understood by the image below which represents the data table of the above shown grid pattern. Although there are more than 100 boxes in the pattern, not all are shown in the image below.

FID	Shape *	Id	BoxLabel	NonAnxious	Anxious
0	Polygon	0	DA	4	3
1	Polygon	0	EA	8	3
2	Polygon	0	FA	12	3
3	Polygon	0	DL	2	2
4	Polygon	0	EL	10	2
5	Polygon	0	GA	6	2
6	Polygon	0	DB	0	0
7	Polygon	0	DC	0	0
8	Polygon	0	DD	2	0
9	Polygon	0	DE	0	0
10	Polygon	0	DF	2	1
11	Polygon	0	DG	1	2
12	Polygon	0	DH	0	0
13	Polygon	0	DI	0	0
14	Polygon	0	DJ	0	0
15	Polygon	0	DK	0	0
16	Polygon	0	EK	0	0
17	Polygon	0	EJ	0	0
18	Polygon	0	EI	0	0
19	Polygon	0	EB	0	0
20	Polygon	0	EC	0	0
21	Polygon	0	ED	1	0
22	Polygon	0	EE	0	0
23	Polygon	0	EF	1	0
24	Polygon	0	EG	1	2
25	Polygon	0	EH	3	0
26	Polygon	0	FB	0	0
27	Polygon	0	FC	3	0
28	Polygon	0	FD	3	0
29	Polygon	0	FE	2	0
30	Polygon	0	FF	1	0
31	Polygon	0	FG	2	0
32	Polygon	0	FH	2	0
33	Polygon	0	FI	1	0
34	Polygon	0	FJ	0	0
35	Polygon	0	FK	0	0
36	Polygon	0	FL	8	3
37	Polygon	0	GB	2	0
38	Polygon	0	GC	2	1
39	Polygon	0	GD	0	1
40	Polygon	0	GE	0	2
41	Polygon	0	GF	0	3
42	Polygon	0	GG	4	2
43	Polygon	0	GH	5	1
44	Polygon	0	GI	2	3
45	Polygon	0	GJ	2	0
46	Polygon	0	GK	4	1
47	Polygon	0	GL	10	0
48	Polygon	0	CA	2	5
49	Polygon	0	CB	1	3
50	Polygon	0	CC	0	2
51	Polygon	0	CD	0	2
52	Polygon	0	CE	0	4
53	Polygon	0	CF	1	2
54	Polygon	0	CG	5	3
55	Polygon	0	CH	0	0
56	Polygon	0	CI	1	0
57	Polygon	0	CJ	2	0
58	Polygon	0	CK	6	0
59	Polygon	0	CL	5	0

Data table for the grid pattern.

The table carries three new columns, i.e, “BoxLabel”, “NonAnxious”, and “Anxious.” The BoxLabel carries a distinct name of each box in the grid, it spans from A to I in rows and A to L across columns, resulting in a 108 box grid pattern. The image and video observations consisted of a lot of individuals, however, this project is constructed on a particular limitation of merely using those individuals in the data set that exhibited emotions of stress or anxiety via facial expressions or body movements, which thus resulted in a smaller data set of the whole.

These individuals are then counted and then accordingly inserted into respective boxes with columns, “Anxious” carrying the number of people in that box who exhibited expressions of anxiety or stress, and “NonAnxious” who were on the opposite side of the expression by exhibiting facial expressions of smiles, laughs, etc. The images below help understand the consideration of the data:

Facial expressions and body movements were used to determine if the individual was anxious or not.

The classification decision was a binary approach instead of a scale. So, it was either a person was classified as anxious or not anxious, nothing in between. Another example to demonstrate how a person was classified as anxious is the image shown below, the man was observed as to be pacing around the street lamp continuously.

A man is seen to be continuously be pacing around a lamp post.

Simply put, any expression, activity or body movement that hinted anything about anxiety was used as an aid in considering which individuals to be picked for the analysis - Friends in sweats casually walking and running around, or a donation center, individual entering the work building, has eyebrows lifted, walking the dog, a couple sharing laughing talks etc

Observation

Grid pattern symbolized with respect to anxiety values as a percent of whole.

The Moran's I statistic is used to measure the autocorrelation for the data generated in this project. The algorithm or the function returns a z-value that allows us to make a decision as to whether we can accept or reject the null hypothesis. A z-value is a statistical testing procedure that helps to make a decision if we should accept or reject the formed Null Hypothesis. The graphical representation output from the Moran's I index function aids us to determine if the data we are testing has a relation, i.e if it forms a cluster or not, visually.

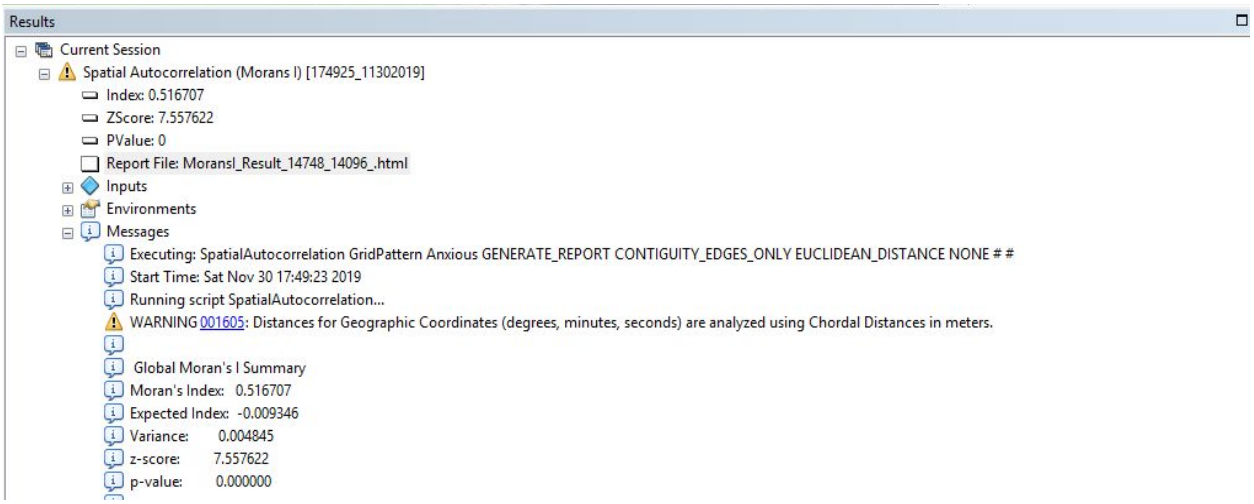
In order to make a decision if we can accept or reject the Null Hypothesis, we observe the Z-Score in the output produced by ArcMap. If the value of the Z-Scores are either too large, or too small in a significance interval, then it is safe to reject the Null Hypothesis.

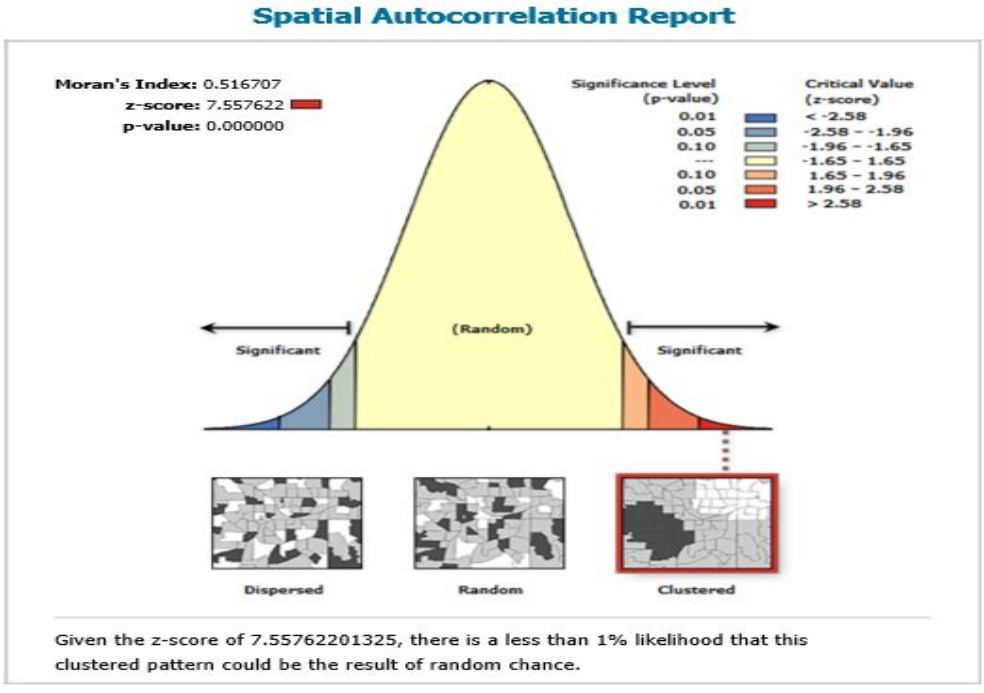
Considering a spectrum that ranges in values from -1 to 1, the extreme ends, say the negative points out that there is no correlation, and thus there is no existence of clusters in the data, however, if the value takes +1.0 then the data holds a high relation and thus is clustered. This

value is a Moran's I index value that we refer to after making a decision about the Null Hypothesis.

A visual assessment of the above grid pattern which is symbolized with respect to number of individuals as a percent of whole having anxiety suggests a hint of clustering in the data, thus in order to test the hypothesis we perform global correlation test, that is Spatial Autocorrelation (Morans I).

GLOBAL CORRELATION
For Anxious:
Analysis and Report





Global Moran's I Summary

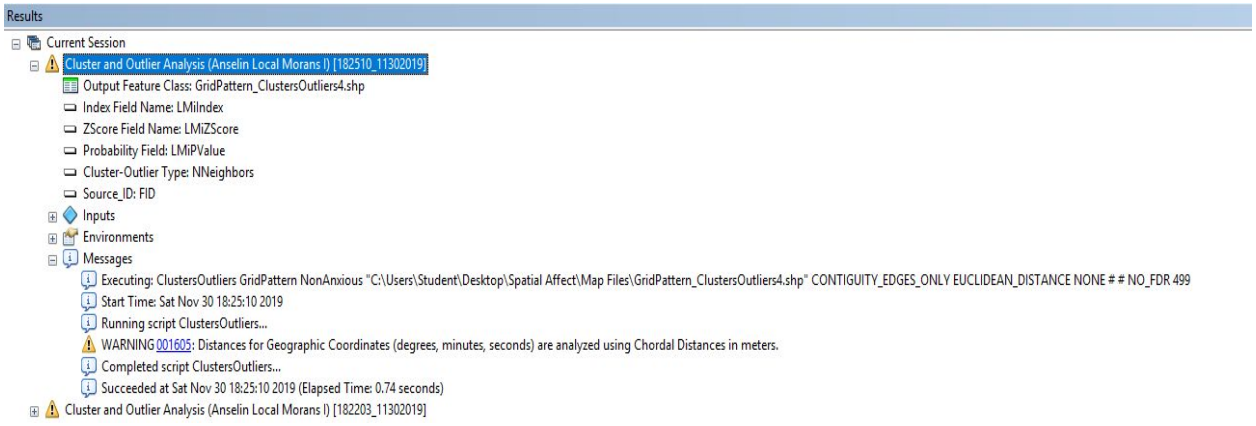
Moran's Index:	0.516707
Expected Index:	-0.009346
Variance:	0.004845
z-score:	7.557622
p-value:	0.000000

Dataset Information

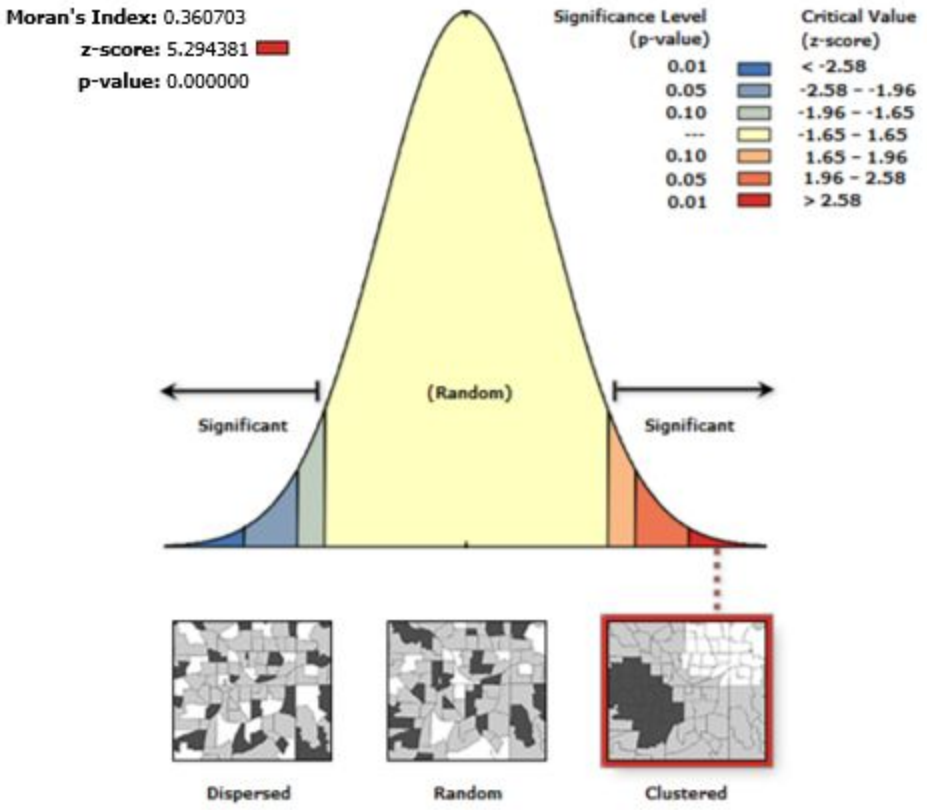
Input Feature Class:	GridPattern
Input Field:	ANXIOUS
Conceptualization:	CONTIGUITY_EDGES_ONLY
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	None
Weights Matrix File:	None
Selection Set:	False

In the testing we can observe a p-value score of 0 and a z-score value of 7.5576622 that simply implies that is not by random that we observed a clustering in the geographical area with respect to anxiety, with Moran's Index = 0.516707 we can consider the data to be correlated, however, not perfectly but yes to some degree.

For NonAnxious
Analysis and Report



Spatial Autocorrelation Report



Given the z-score of 5.29438139875, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

Global Moran's I Summary

Moran's Index:	0.360703
Expected Index:	-0.009346
Variance:	0.004885
z-score:	5.294381
p-value:	0.000000

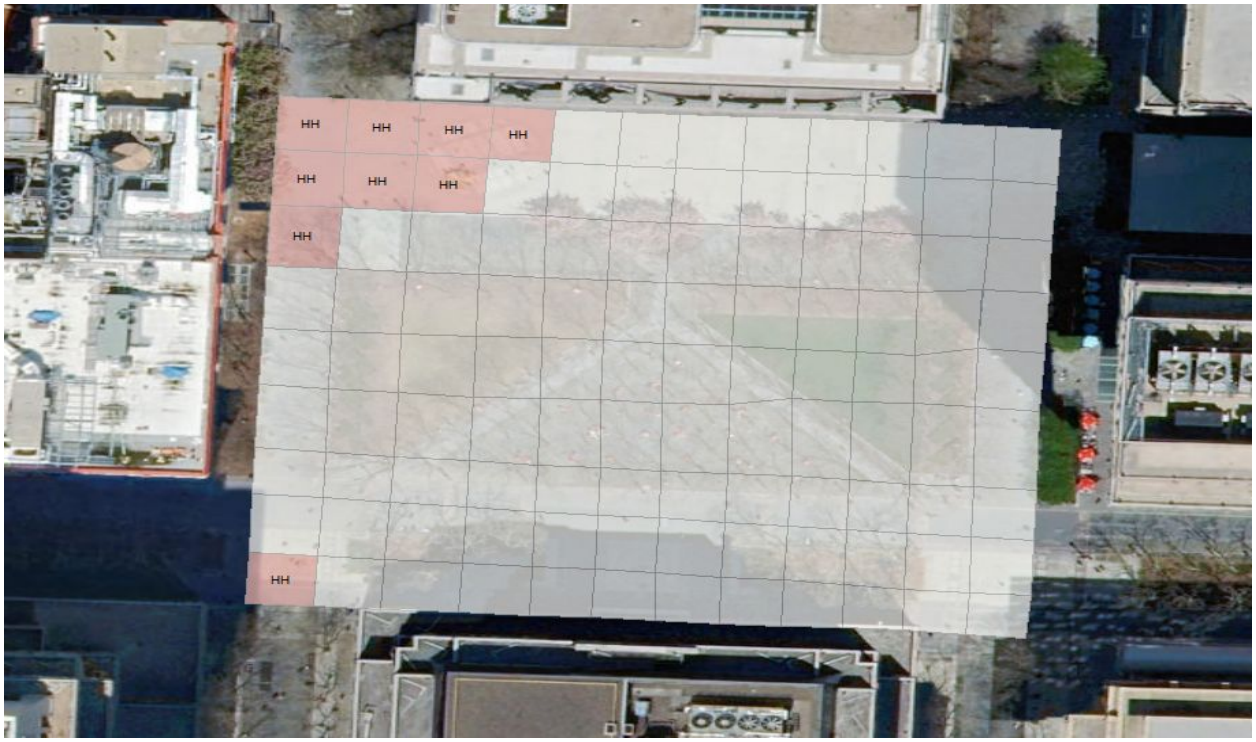
Dataset Information

Input Feature Class:	GridPattern
Input Field:	NONANXIOUS
Conceptualization:	CONTIGUITY_EDGES_ONLY
Distance Method:	EUCLIDEAN
Row Standardization:	False
Distance Threshold:	None
Weights Matrix File:	None
Selection Set:	False

There is the same case with non anxious individuals, preliminary observations suggested clustering, however not as strong as those individuals considered as anxious. The Morans I testing yields a z-score value of 5.294381 which is somewhat less than the previous case, this could be because of a lot of unconsidered factors in data labeling. However, we still see a clustering pattern with p = 0.

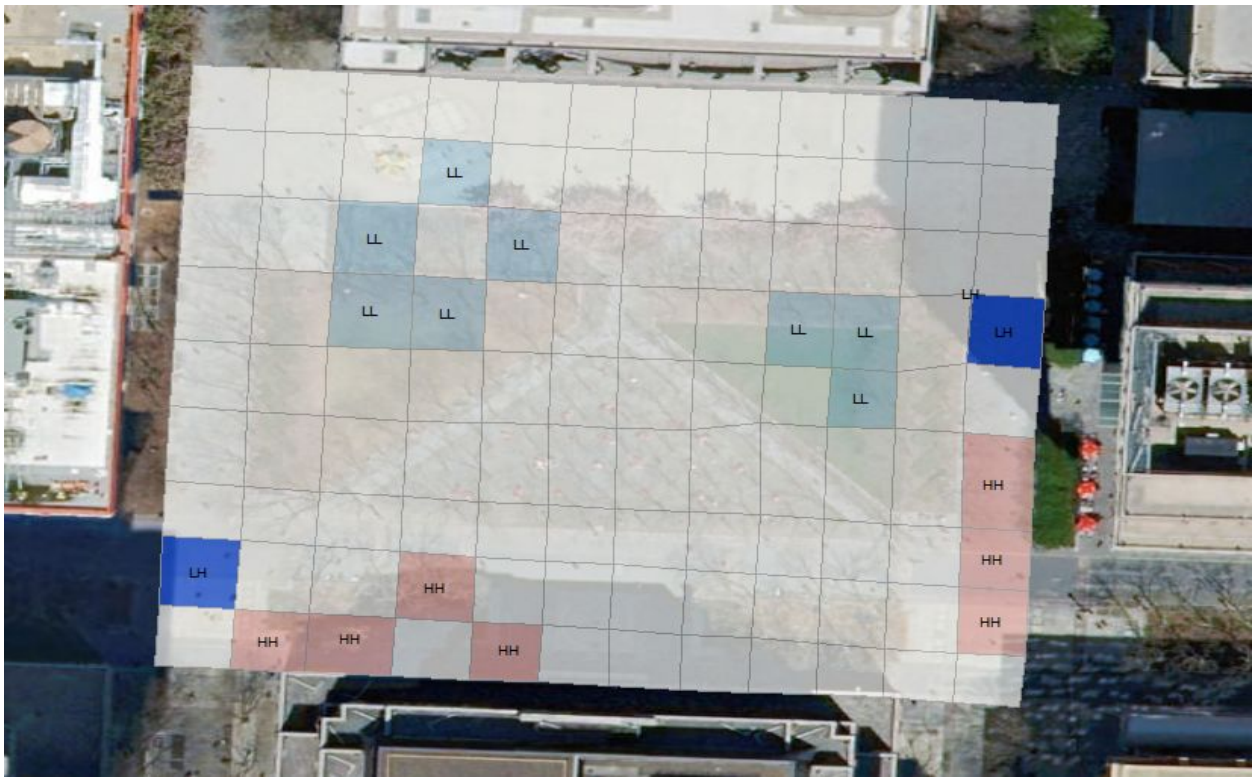
LOCAL CORRELATION

Anxious

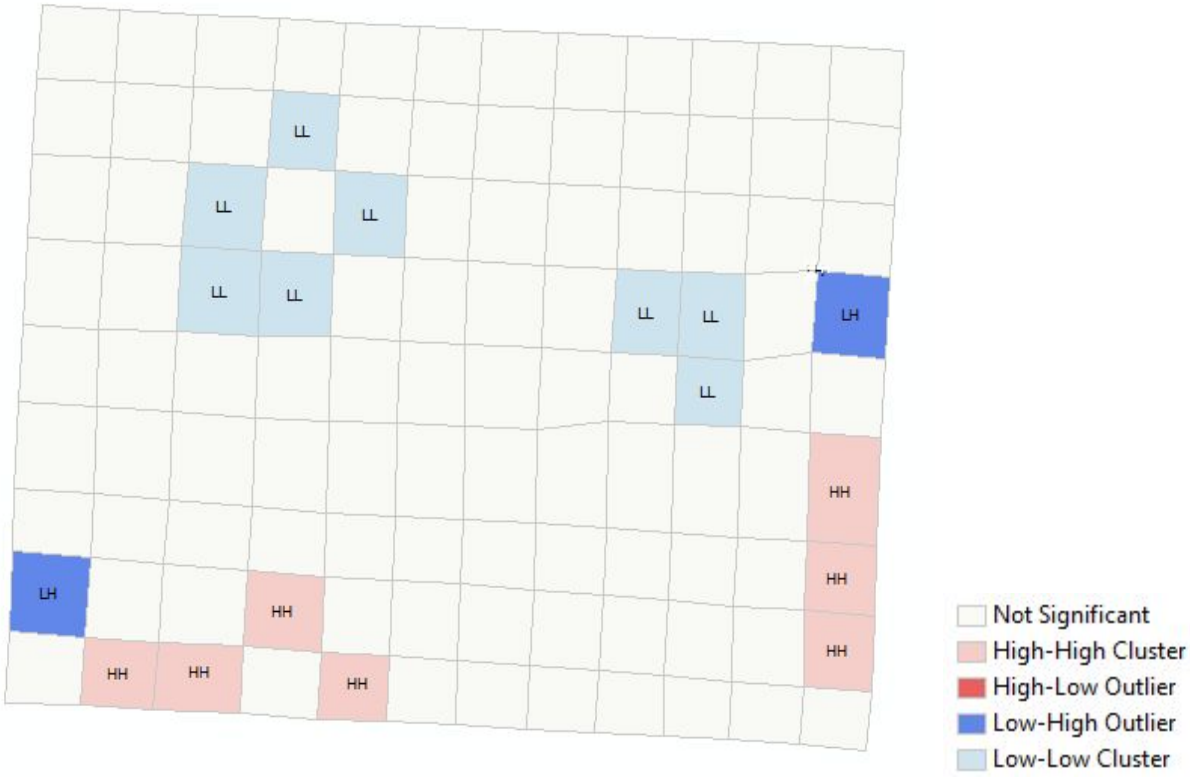


Local correlation analysis on anxious values.

NonAnxious



Local correlation analysis on nonanxious values.



To perform outlier and local correlation, Anselin Local Morans I statistic tool was used. The analysis of the test yields clustering analysis and representation high-low clusters. In the case of Anxious testing, we observe only High-High results and no other categories. However, in the case of non anxious individuals, we observe Low-Low, Low-High, High-High values, say for cell in bottom left, we have LH values, cross checking with the data in the data table:

72	Polygon	0	HA	2	7

Data entry for cell HA, polygon number 72, having Low-High Values in cluster.

We have 2 non anxious individuals and 7 anxious individuals. One of the issues that we face here is a number of boxes or cells contain Null values, these are positions that cannot encompass individuals like columns of tree plantations, christmas tree cell etc. thus this can produce some differences in analysis.

In conclusion, we can see clustering that has been verified statistically using Morans I, and to talk about the observation it is probable that individuals that fall in the cells that lie around the school building can be more anxious than others who may be sitting in the triangular section of the area, or someone who is talking a walk with a dog. The triangular area sees a cooler effect in terms of anxiety, this is straightforward, as a student if I am stressed about something, I'd rather go straight to the library to get my work done, instead of taking a shortcut through the

central part of the area. The other factors that might influence the considered assumption of this project on emotional expressions is the directional flow of an individual in the grid, are they leaving the building or entering it, which is thus not considered, and thus can be extended beyond the scope of this project.

Sources:

Geo Tracker - GPS Tracker, Android application for coordinate recording, https://play.google.com/store/apps/details?id=com.ilyabogdanovich.geotracker&hl=en_US
Google Satellite Imagery, last accessed Thursday, 28th November, 5:35 PM https://maps.google.com
ArcGIS for Desktop Help, Spatial Autocorrelation (Global Moran's I) (Spatial Statistics)
ArcGIS for Desktop Help, Cluster and Outlier Analysis (Anselin Local Moran's I) (Spatial Statistics)
The base map used is “Streets”, provided by ArcMap Desktop