Department of Decision Science

Faculty of Business

University of Moratuwa

DA3480 - Spatial Data Analysis

Semester 06

Individual Assignment

Exploring Spatial Autocorrelation and Moran's Index Analysis of HotelLocation_ColomboGT300Reviews

M. W. S. A. U. Silva

206121X

Introduction

The "HotelLocation_ColomboGT300Reviews.csv" file is a collection of data related to hotel properties situated in Colombo. This dataset has been prepared for use with GeoDa's analysis of spatial data. With a focus on its geographic locations and customer reviews, it provides insightful information on many aspects of these hotel properties. This dataset offers a snapshot of the Colombo hotel landscape with a total of 35 records.

The dataset includes several key columns, each of which offers particular information:

ID: A specific identifier that can be used to refer to each hotel facility.

Property: Each hotel's name or designation

Latitude and longitude: the geographical coordinates that identify each hotel's precise location on the surface of the Earth.

Review Score: A numerical representation of each hotel's overall customer satisfaction rating that provides a measure of the level of service provided.

Street, Town, Country: This section of the address specifies the street, town, and country information for the hotel's Colombo location.

Expenditure ($): The dataset's financial component represents the costs related to each hotel facility
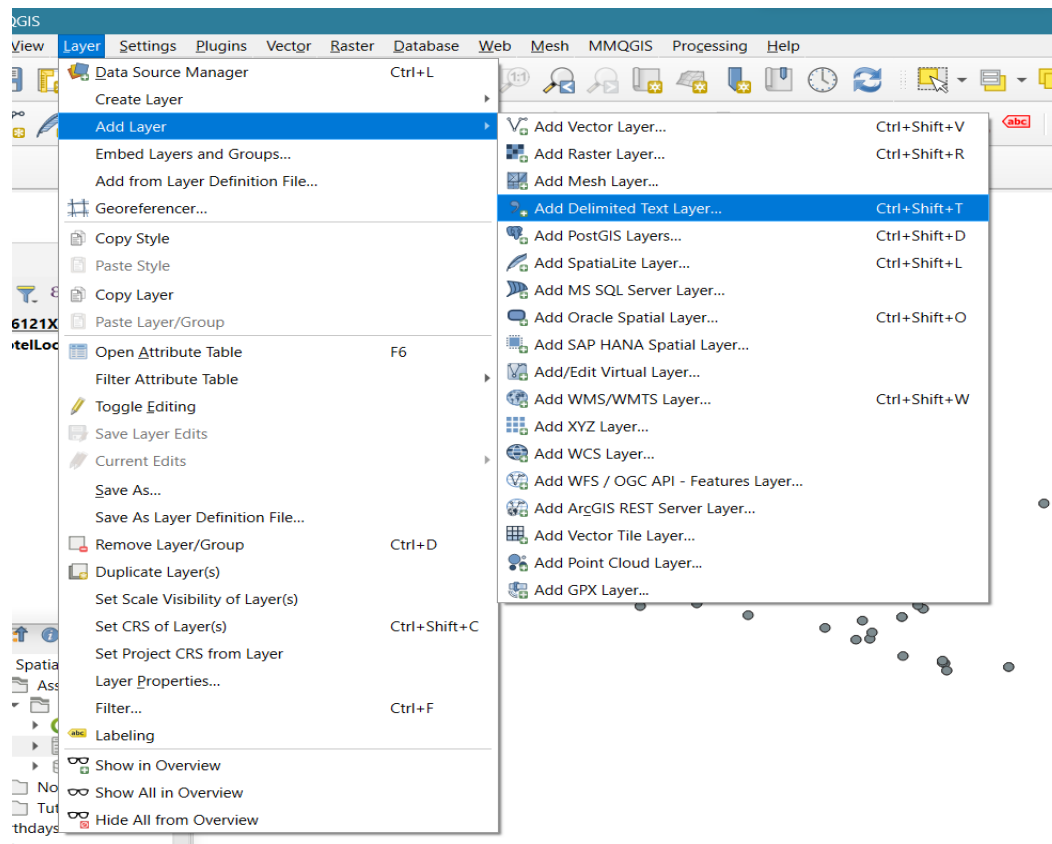
1. Before Analysis Setup:

1.1. What is the purpose of setting a seed for randomization?

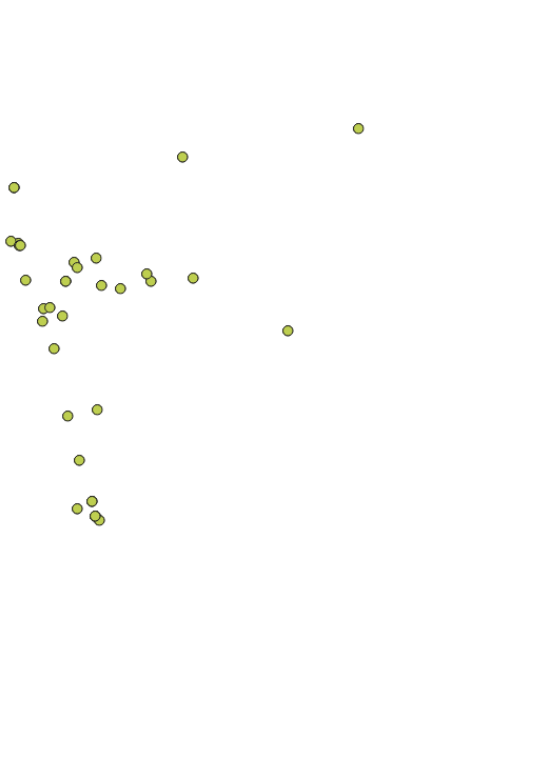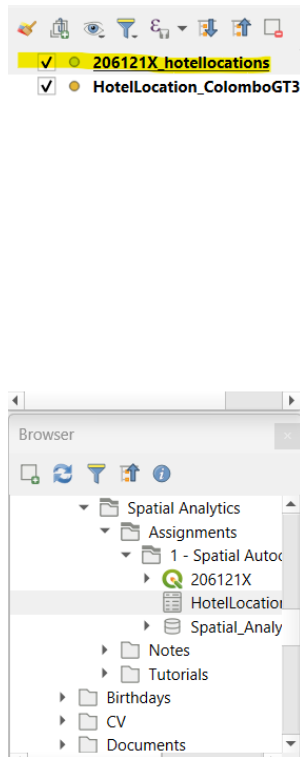To ensure the reproducibility on a particular machine.

2. Creating Distance-Band Weights:

2.1. How do you reproject the provided dataset to UTM projection?

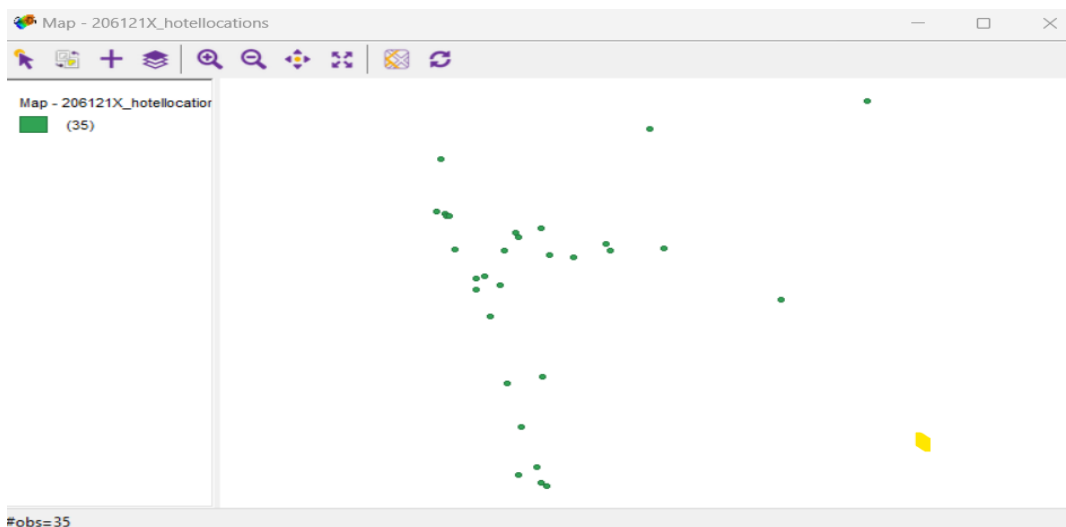- First, we have to add the "HotelLocation_ColomboGT300Reviews.csv" to QGIS as a delimited text layer.

- Then we need to export the layer as a shape file as "206121X_hotellocations.shp"

2.2.What are the steps for creating a weights matrix in GeoDa?

- First add the shape file to the GeoDa

- Then, click "tools" tab in the GeoDa and select "Weights Manager". Click "Create" button on the Weights Manager dialog box and it will open another dialog box called "Weights File Creation". Set ID variable as "ID".
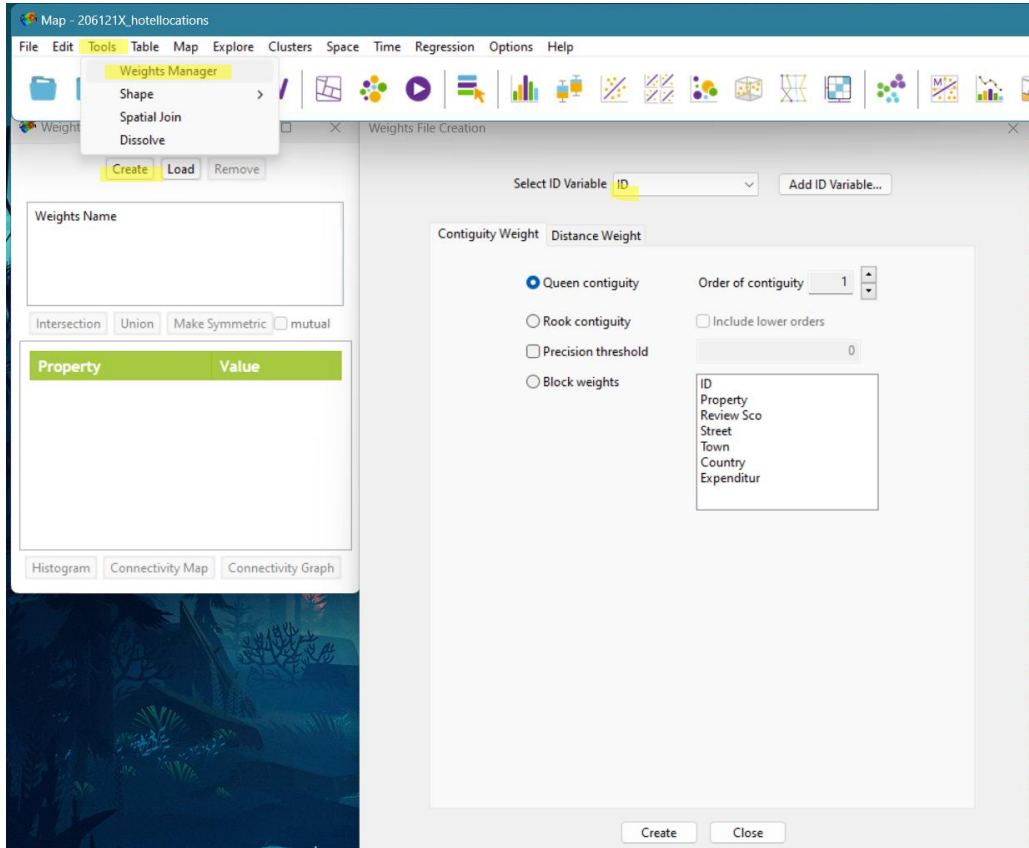
- Then go to "Distance Weight" tab and set Longitude as X-coordinate variable and Latitude as Y-coordinate variable. Then click "Create" and save the file as "202161X_weights.gwt".

- Finally, "Distance-band weights" shown as follows.

Weights Manager — □ ✕

Create Load Remove

Weights Name

206121X_weights

Intersection Union Make Symmetric ☐ mutual

Property	Value
type	threshold
symmetry	symmetric
file	206121X_weights.gwt
id variable	ID
distance metric	Euclidean
distance vars	Long, Lat
distance unit	degree
threshold value	0.0372186
# observations	35
min neighbors	1
max neighbors	31
mean neighbors	19.20
median neighbors	22.00
% non-zero	54.86%

Histogram Connectivity Map Connectivity Graph

Distance Metric Questions

1. Why did you choose Euclidean distance when creating the Distance Metric?

   - Since the data is projected we can use Euclidean distance to create Distance Metric. Even though Euclidean distance ignore the curvature of the earth, since this dataset is a low-dimensional dataset and also covers only a small area (Colombo district), Euclidean distance is appropriate when creating the Distance Metric.

2. What is the critical distance for your point data?

   - 0.0372186

3. What is the meaning of "critical distance"?

   - Critical distance shows the distance at which spatial relationships between data points become negligible or less relevant.

4. Explain the importance of the "threshold value".

   - Threshold value is important when determining the spatial relationships and patterns that can be considered as significant and relevant.

   - The results of spatial analyses, the identification of spatial outliers or clusters, and the outcome of decisions including spatial autocorrelation and pattern recognition can all be affected by the selection of threshold values properly. As a result, choosing a suitable threshold value is a crucial step in carrying out a precise and accurate spatial data analysis.

5. Create the Moran's index using the Review Score as the spatial variable.

Moran's I: 0.061
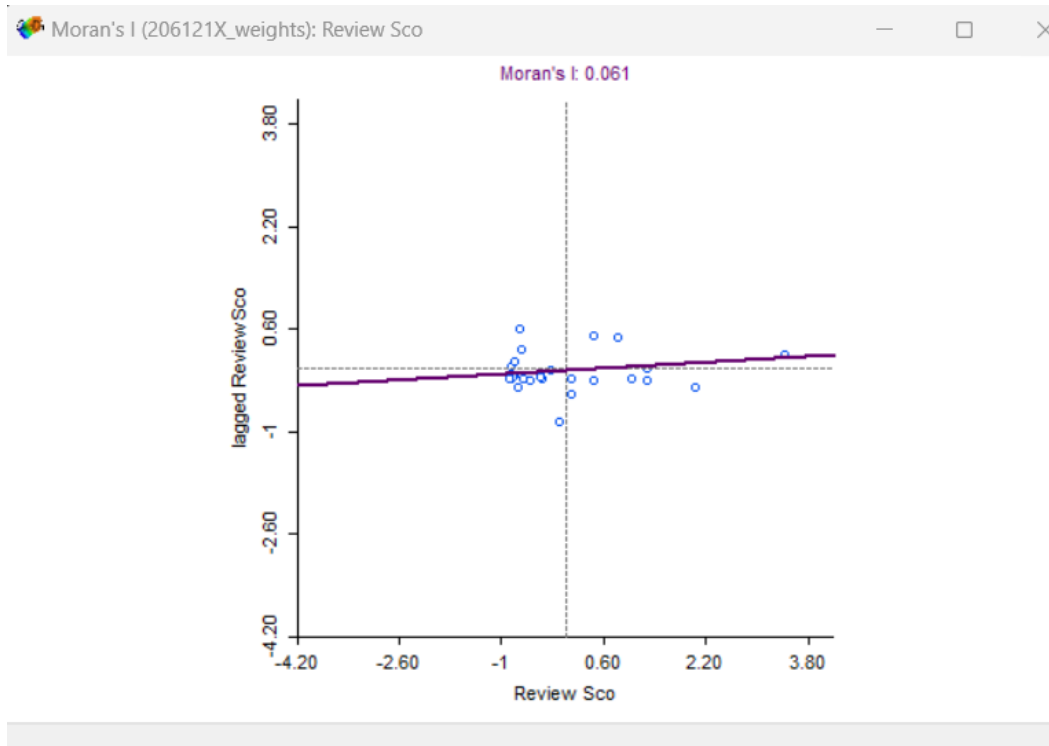
Interpretation Questions

1. What can you tell about the shape of the point cloud?

   - There are multiple outliers on the right end, which affects the shape of the point cloud.

2. Identify locations in the map (or, in any other open view) associated with each of the four types of spatial autocorrelation. What interpretations can you make using the available information?

   - Since points are in each of the four quadrants, it shows that the data has some degree of spatial autocorrelation, both positive and negative.

   - Since the points are very close to the center, suggests that the spatial autocorrelation is not very significant or noticeable.

- The pattern of the right-side points being relatively farther apart than the left-side ones indicates that the data reflects wave-like movement.

- The presence of a spatial wave signifies that the variable values may gradually shift or transition throughout location.

3. Inference:

3.1. How do you perform randomization with a specified number of permutations?

- Right click on the Moran's scatter plot – Review Sco. Then select "Randomization" and select the specified number of permutations (e.g.: 999 permutations).
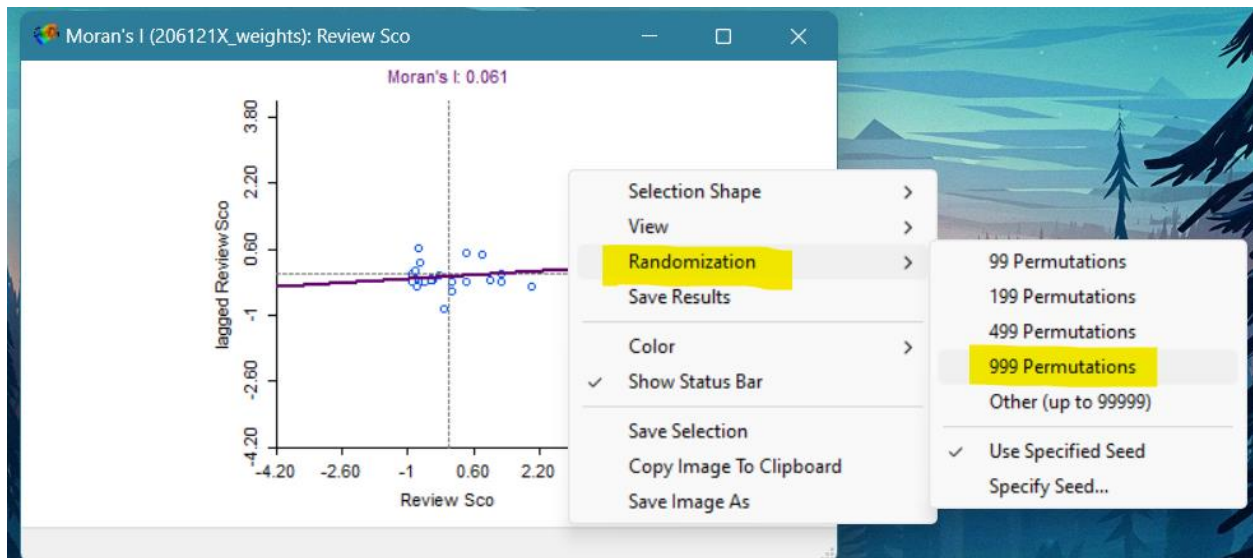
- Result of permutation operation:

permutations: 999
pseudo p-value: 0.048000

I: 0.0605 E[I]: -0.0294 mean: -0.0284 sd: 0.0459 z-value: 1.9362
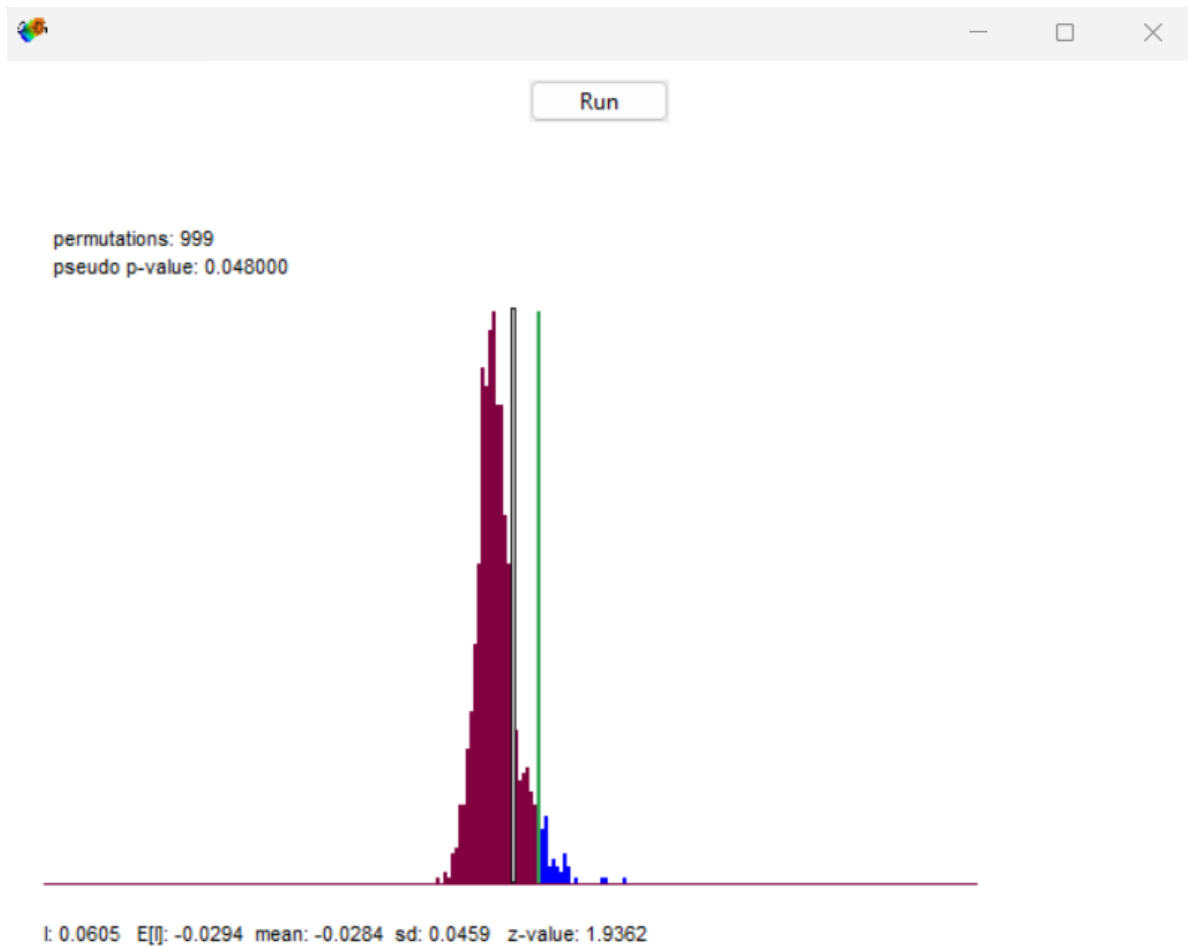
What information can you draw from the distribution?

- Green line = 0.0605 the value of statistic for the actual data.

- Number of permutations = 999

- Pseudo p-value = 0.048

- Theoretical expected value = -0.0294

- Mean = -0.0284

- Standard deviation = 0.0459

3.2. What is the purpose of setting the number of permutations?

- It can be used to identify where the observed Moran's I value falls when compared to the reference distribution. GeoDa creates a histogram of the values of Moran's I when compared to the observed Moran's I after determining the number of permutations.

3.3. What does the reference distribution for the statistic depict, and how is it represented?

3.4. How can you use the distribution to draw conclusions about the analysis?

- An understanding of the characteristics and patterns of the dataset can be gained from the distribution of the data. We can determine the nature of the data, find spatial patterns, identify outliers, and make accurate decisions in spatial analysis by looking at measures of central tendency, spread, skewness, and visual representations. The distribution is an important point to start when conducting more research and testing hypotheses in spatial analysis.

4. Moran Scatter Plot Options:

4.1. What options can be accessed by right-clicking on the Moran scatter plot?
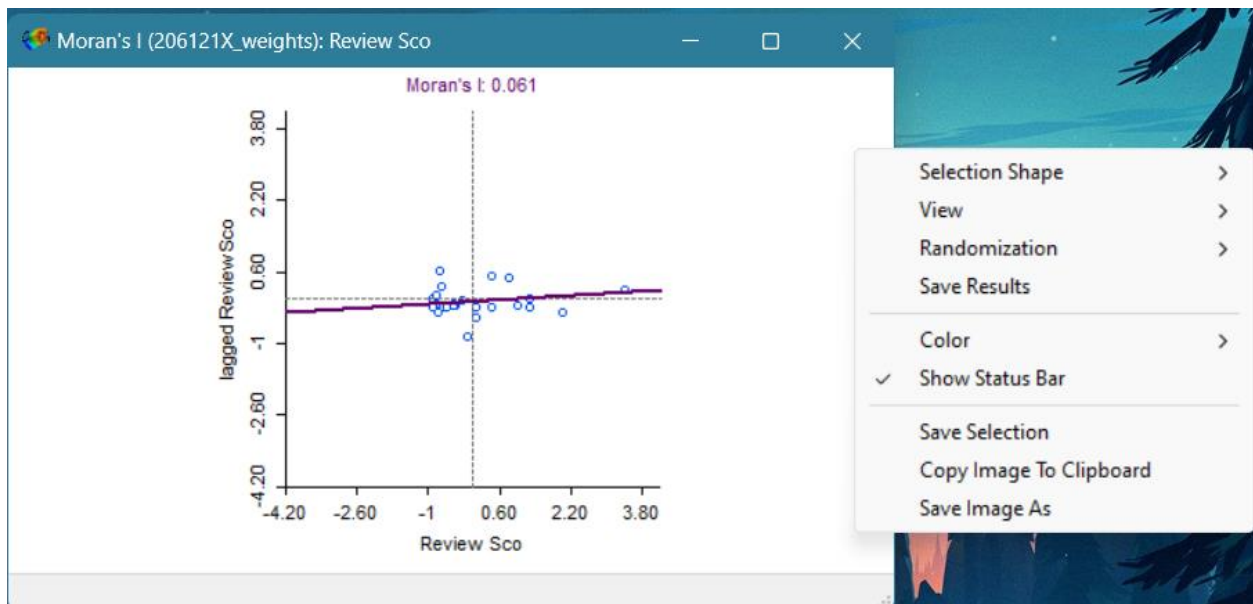
4.2. What steps are involved in computing a standardized version of the Review Score and its spatial lag?

- **1ˢᵗ way** → Right click on the Moran scatter plot and click "Save Results". It will open a dialog box with default names for Standardized Data and Spatial Lag. Then click OK to add the two variables to the data table.

- **2nd way** → we can use calculator option in table tab to verify the values we got by save results. First, go to Univariate tab and add variable "MST" for review sco. Then select "STANDERDIZED (Z)" as the operator and "Review Sco" as Constant.

Calculator   ✕

Special Univariate Bivariate Spatial Lag Rates   Date/Time

                                    Operator   Variable / Constant
Result   Add Variable   STANDARDIZED (Z) ∨   Review Sco   ∨
                          =
MST   ∨
                               MST = standardized dev from mean of Review Sco

                                      Apply   Close

- Then go to Spatial Lag tab and add a new variable as "MLAG" for Spatial Lag. Then select "206121X_weights" file from Weights list and choose "MST" as variable.

Special Univariate Bivariate Spatial Lag Rates   Date/Time

                                              Weight W
                          206121X_weights   ∨
                                                    Variable
Result   Add Variable   MST   ∨
MLAG   ∨   =   MLAG = 206121X_weights * MST

                                   ☑ Use row-standardized weights
                                   ☐ Include diagonal of weights matrix
                                   ☐ Median spatial lag

                                      Apply   Close

- Finally, click "Apply" button and this will apply these two variables to the data table. So, there will be 4 new columns in the data table as follows.

Table - 206121X_hotellocations

#	MLAG	MST	ID	Property	Lat	Long	Review Sco	Street	Town	Country	Expenditur	MORAN_STD	MORAN_LAG
1	0.502242	0.426524	1	Barefoot Garden Cafe	6.866487	79.862962	1569	704 Galle Rd	Colombo	Sri Lanka	172	0.426524	0.502242
2	0.271366	-0.713376	2	Beach Wadiya	6.875043	79.860350	467	2 Station Ave	Colombo	Sri Lanka	86	-0.713376	0.271366
3	-0.145796	-0.867501	3	Cafe Francais	6.916498	79.859293	318	Park Street	Colombo	Sri Lanka	89	-0.867501	-0.145796
4	-0.167297	-0.848882	4	Cafe Shaze	6.911066	79.869020	336	56 Horton Pl	Colombo	Sri Lanka	54	-0.848882	-0.167297
5	-0.108515	-0.414437	5	Capital Bar & Grill	6.920400	79.847542	756	1 Galle Face	Colombo	Sri Lanka	55	-0.414437	-0.108515
6	0.462832	0.820628	6	Chutneys At Cinnamon Grand Colombo	6.863257	79.863676	1950	77 Galle Road	Colombo	Sri Lanka	149	0.820628	0.462832
7	-0.843710	-0.118601	7	Cricket Club Cafe	6.902172	79.903944	1042	12 Flower Road	Colombo	Sri Lanka	65	-0.118601	-0.843710
8	-0.166441	-0.372027	8	Curry Leaf	6.912569	79.857540	797	2 Sir Chittampalam A Gardiner Mawatha	Colombo	Sri Lanka	99	-0.372027	-0.166441
9	0.001844	-0.872673	9	Dolce Italia	6.885726	79.864017	313	5 Skelton Road	Colombo	Sri Lanka	67	-0.872673	0.001844
10	-0.168831	-0.859226	10	Flamingo House	6.911713	79.865022	326	58A Horton Place	Colombo	Sri Lanka	67	-0.859226	-0.168831
11	-0.185750	0.091381	11	FLOW	6.917385	79.863779	1245	200 Union Place	Colombo	Sri Lanka	62	0.091381	-0.185750
12	0.502242	0.426524	12	Green Cabin Kollupitiya	6.866487	79.862962	1569	704 Galle Rd	Colombo	Sri Lanka	131	0.426524	0.502242
13	-0.192508	1.265417	13	King of the Mambo	6.920857	79.845978	2380	Galle Face Hotel	Colombo	Sri Lanka	113	1.265417	-0.192508
14	-0.167069	1.017163	14	Lagoon	6.932027	79.846783	2140	117 Sir Chittampalam A Gardiner Mawatha	Colombo	Sri Lanka	174	1.017163	-0.167069
15	0.093450	-0.801300	15	Menya Hanabi Sri Lanka	6.944549	79.918653	382	No. 18 19 Level 03	Colombo	Sri Lanka	86	-0.801300	0.093450
16	0.203096	3.417988	16	Ministry of Crab	6.862575	79.864529	4461	Old Dutch Hospital	Colombo	Sri Lanka	101	3.417988	0.203096
17	-0.125888	-0.817850	17	Nihonbashi	6.920025	79.847739	366	Galle Face Terrace	Colombo	Sri Lanka	95	-0.817850	-0.125888
18	-0.040623	-0.242728	18	Noodles Restaurant	6.884390	79.857869	922	Galle Road A2	Colombo	Sri Lanka	61	-0.242728	-0.040623
19	0.462832	0.820628	19	Nuga Gama	6.863257	79.863676	1950	77 Galle Road	Colombo	Sri Lanka	119	0.820628	0.462832
20	-0.212460	0.426524	20	Palmyrah Restaurant	6.912759	79.849173	1569	704 Galle Rd	Colombo	Sri Lanka	132	0.426524	-0.212460
21	-0.014561	1.277830	21	Paradise Road The Gallery Cafe	6.898470	79.855083	2392	2 Alfred House Road	Colombo	Sri Lanka	105	1.277830	-0.014561
22	-0.147133	-0.835435	22	Park Street Mews Restaurant	6.915422	79.859962	349	50/1 Park Street	Colombo	Sri Lanka	80	-0.835435	-0.147133
23	-0.167069	1.017163	23	Royal Thai - Cinnamon Lakeside	6.932027	79.846783	2140	117 Sir Chittampalam A Gardiner Mawatha	Colombo	Sri Lanka	123	1.017163	-0.167069
24	-0.166441	-0.372027	24	Sea Fish Restaurant	6.912569	79.857540	797	2 Sir Chittampalam A Gardiner Mawatha	Colombo	Sri Lanka	56	-0.372027	-0.166441
25	0.618197	-0.733030	25	Shanmugas	6.864802	79.859882	448	06 53/3 Ramakrishna Road Colombo	Colombo	Sri Lanka	66	-0.733030	0.618197
26	-0.144047	-0.884051	26	The Barnesbury	6.912527	79.875354	302	91 Barnes Place	Colombo	Sri Lanka	94	-0.884051	-0.144047
27	-0.144225	-0.414437	27	The Bavarian German Restaurant and Pub	6.919954	79.848134	756	1 Galle Face	Colombo	Sri Lanka	62	-0.414437	-0.144225
28	-0.178768	-0.883017	28	The Cafe on the 5th	6.905308	79.856730	303	5th Lane	Colombo	Sri Lanka	94	-0.883017	-0.178768
29	0.462832	0.820628	29	The London Grill	6.863257	79.863676	1950	77 Galle Road	Colombo	Sri Lanka	105	0.820628	0.462832
30	-0.299145	2.006042	30	The Mango Tree	6.906683	79.852765	3096	No 10 Deal Place A	Colombo	Sri Lanka	125	2.006042	-0.299145