空间统计分析方法简介

空间自相关的含义

- **空间自相关（spatial auto correlation）**
- 是测试空间某点的观测值是否与其相邻点的值存在相关性的一种分析方法。
- 如果某一位置变量值高，其附近位置上该变量值也高，则为正空间自相关，反之，则为负空间自相关。

全局莫兰指数计算公式

- 全局莫兰指数（Moran’s I）
- 该指标可以指出区域属性值的分布是否是聚集，离散或者随机分布模式。
- 莫兰指数的值域为[-1，1]，取值为-1表示完全负相关，取值为1表明完全正相关，而取值为0表示不相关。

$$I=\frac{n}{\sum_i\sum_j w_{i,j}}\frac{\sum_i\sum_j w_{i,j}(x_i-\bar{x})(x_j-\bar{x})}{\sum_i(x_i-\bar{x})^2}$$

空间权重矩阵

- 通常用一个二元对称空间权重矩阵W来表达n个位置的区域的邻近关系，其中，wij为区域i与j的邻近关系。

$$W=\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

- 简单的二进制邻接矩阵：

$$w_{ij}=\begin{cases}1 & \text{当区域}i\text{和}j\text{相邻接}\\ 0 & \text{其他}\end{cases}$$

- 基于距离的二进制空间权重矩阵：

$$w_{ij}=\begin{cases}1 & \text{当区域}i\text{和}j\text{的距离小于}d\text{时}\\ 0 & \text{其他}\end{cases}$$

显著性检验

- Z值检验

$$Z=\frac{Moran'sI-E(I)}{\sqrt{VAR(I)}}$$

$$E(I)=-\frac{1}{n-1}$$

$$s^2d=VAR(I)=\frac{n^2w_1+nw_2+3w_0^2}{w_0^2(n^2-1)}-E^2(I)$$

$$w_0=\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}$$

$$w_1=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{ij}+w_{ij})^2$$

$$w_2=\sum(w_{i\cdot}+w_{\cdot j})^2$$

局部莫兰指数计算公式

$$I_i=\frac{n^2}{\sum_i\sum_j w_{ij}}\times\frac{(x_i-\bar{x})\sum_j w_{ij}(x_j-\bar{x})}{\sum_j(x_j-\bar{x})^2}$$

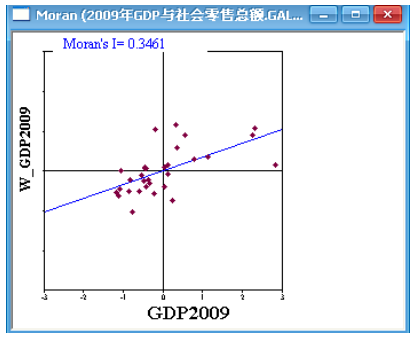
全局Moran's I和Local Moran's I指数之间存在一下关系：

$$\sum_{i=1}^{n}I_i=n\times I$$

- Local Moran’s I统计量它是Global Moran’s I的分解形式。
- 局部莫兰指数高值表明有相似变量值的面积单元在空间集聚（高值或低值），低值表明不相似变量值的面积单元在空间集聚。

局部莫兰指数的计算结果

GDP2009	XIAOFEI	STD_GDP2009	LAG_GDP2009	I_GDP2009	CL_GDP2009	PVAL_GDP2009
8587.000000	3401.800000	-0.328557	-0.336522	0.110566	0.000000	0.360000
9740.250000	2855.300000	-0.210037	-0.592117	0.124366	2.000000	0.040000
4277.050000	1177.500000	-0.771493	-1.042841	0.804545	2.000000	0.002000
7278.750000	2957.300000	-0.463007	-0.269297	0.124686	0.000000	0.444000
15212.490000	5812.600000	0.352349	0.560253	0.197405	0.000000	0.170000
3387.560000	1183.000000	-0.862907	-0.546928	0.471948	2.000000	0.026000
17235.480000	5764.900000	0.560253	0.887603	0.497282	0.000000	0.076000
12153.030000	5309.900000	0.037926	0.061115	0.002318	0.000000	0.372000
7358.310000	2809.000000	-0.454830	0.069834	-0.031763	0.000000	0.408000
7521.850000	2430.800000	-0.438023	0.037926	-0.016613	0.000000	0.362000
8169.800000	2699.700000	-0.371433	-0.248058	0.092137	0.000000	0.192000
1353.310000	339.300000	-1.071968	-0.481459	0.516108	0.000000	0.172000
1081.270000	300.500000	-1.099926	-0.639200	0.703073	2.000000	0.038000
33896.650000	12363.000000	2.272533	0.876123	1.991017	1.000000	0.022000
441.360000	156.600000	-1.165690	-0.551278	0.642619	0.000000	0.086000
19480.460000	6746.400000	0.790971	0.278071	0.219946	0.000000	0.250000
34457.300000	11484.100000	2.330151	1.047824	2.441588	0.000000	0.088000
10962.830000	3527.800000	-0.176886	1.018062	-0.180081	3.000000	0.014000



确定散点图的数据

GDP2009	XIAOFEI	STD_GDP2009	LAG_GDP2009	I_GDP2009	CL_GDP2009	PVAL_GDP2009
8587.000000	3401.800000	-0.328557	-0.336522	0.110566	0.000000	0.360000
9740.250000	2855.300000	-0.210037	-0.592117	0.124366	2.000000	0.040000
4277.050000	1177.500000	-0.771493	-1.042841	0.804545	2.000000	0.002000
7278.750000	2957.300000	-0.463007	-0.269297	0.124686	0.000000	0.444000
15212.490000	5812.600000	0.352349	0.560253	0.197405	0.000000	0.170000
3387.560000	1183.000000	-0.862907	-0.546928	0.471948	2.000000	0.026000
17235.480000	5764.900000	0.560253	0.887603	0.497282	0.000000	0.076000
12153.030000	5309.900000	0.037926	0.061115	0.002318	0.000000	0.372000
7358.310000	2809.000000	-0.454830	0.069834	-0.031763	0.000000	0.408000
7521.850000	2430.800000	-0.438023	0.037926	-0.016613	0.000000	0.362000
8169.800000	2699.700000	-0.371433	-0.248058	0.092137	0.000000	0.192000
1353.310000	339.300000	-1.071968	-0.481459	0.516108	0.000000	0.172000
1081.270000	300.500000	-1.099926	-0.639200	0.703073	2.000000	0.038000
33896.650000	12363.000000	2.272533	0.876123	1.991017	1.000000	0.022000
441.360000	156.600000	-1.165690	-0.551278	0.642619	0.000000	0.086000
19480.460000	6746.400000	0.790971	0.278071	0.219946	0.000000	0.250000
34457.300000	11484.100000	2.330151	1.047824	2.441588	0.000000	0.088000
10962.830000	3527.800000	-0.176886	1.018062	-0.180081	3.000000	0.014000

散点图的含义

LH区域自身福利水平较低，周边地区较高，二者的空间差异程度较大，较强的空间负相关，即异质性突出。

HH区域自身和周边地区的福利水平均较高，二者的空间差异程度较小，存在较强的空间正相关，即为热点区。

LL区域自身和周边地区的福利水平均较低，二者的空间差异程度较小，存在较强的空间正相关，即为盲点区。

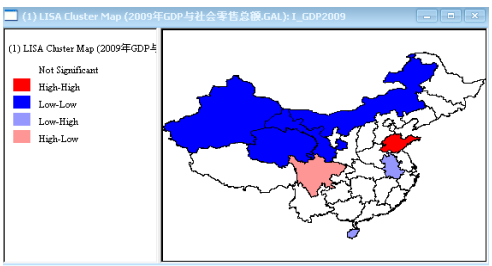
HL区域自身福利水平较高，周边地区较低，二者的空间差异程度较大，较强的空间负相关，即异质性突出。

Moran散点图空间涵义解释

空间集聚情况数据表

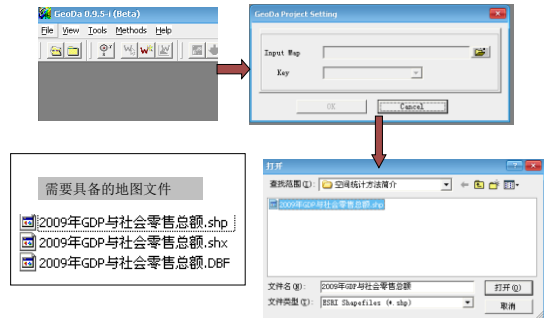
GDP2009	XIAOFEI	STD_GDP2009	LAG_GDP2009	I_GDP2009	CL_GDP2009	PVAL_GDP2009
8587.000000	3401.800000	-0.328557	-0.336522	0.110566	0.000000	0.360000
9740.250000	2855.300000	-0.210037	-0.592117	0.124366	2.000000	0.040000
4277.050000	1177.500000	-0.771493	-1.042841	0.804545	2.000000	0.002000
7278.750000	2957.300000	-0.463007	-0.269297	0.124686	0.000000	0.444000
15212.490000	5812.600000	0.352349	0.560253	0.197405	0.000000	0.170000
3387.560000	1183.000000	-0.862907	-0.546928	0.471948	2.000000	0.026000
17235.480000	5764.900000	0.560253	0.887603	0.497282	0.000000	0.076000
12153.030000	5309.900000	0.037926	0.061115	0.002318	0.000000	0.372000
7358.310000	2809.000000	-0.454830	0.069834	-0.031763	0.000000	0.408000
7521.850000	2430.800000	-0.438023	0.037926	-0.016613	0.000000	0.362000
8169.800000	2699.700000	-0.371433	-0.248058	0.092137	0.000000	0.192000
1353.310000	339.300000	-1.071968	-0.481459	0.516108	0.000000	0.172000
1081.270000	300.500000	-1.099926	-0.639200	0.703073	2.000000	0.038000
33896.650000	12363.000000	2.272533	0.876123	1.991017	1.000000	0.022000
441.360000	156.600000	-1.165690	-0.551278	0.642619	0.000000	0.086000
19480.460000	6746.400000	0.790971	0.278071	0.219946	0.000000	0.250000
34457.300000	11484.100000	2.330151	1.047824	2.441588	0.000000	0.088000
10962.830000	3527.800000	-0.176886	1.018062	-0.180081	3.000000	0.014000

空间集聚图形

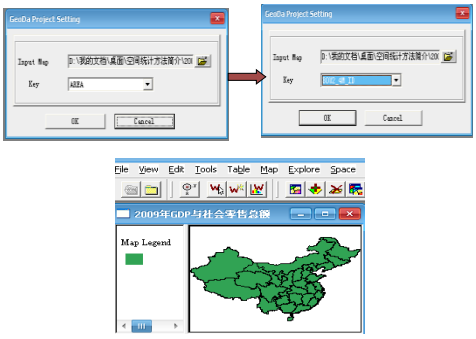


空间统计分析软件
——Geoda的使用

打开地图文件

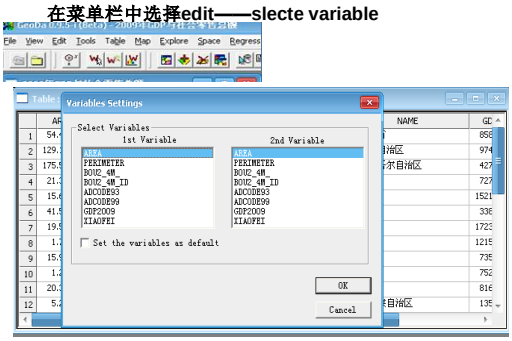


打开地图文件



打开数据窗口

在菜单栏中选择edit——slecte variable



数据录入

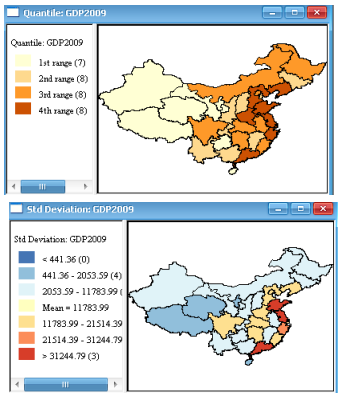
Table : 2009年GDP与社会零售总额

	BOU2_4M_	BOU2_4M_ID	ADCODE93	ADCODE99	NAME	GDP2009	XIAOFEI
1	2	23	230000	230000	黑龙江省	8587.000000	3401.800000
2	3	15	150000	150000	内蒙古自治区	9740.250000	2855.300000
3	4	65	650000	650000	新疆维吾尔自治区	4277.050000	1177.500000
4	5	22	220000	220000	吉林省	7278.750000	2957.300000
5	6	21	210000	210000	辽宁省	15212.490000	5812.600000
6	7	62	620000	620000	甘肃省	3387.560000	1183.000000
7	8	13	130000	130000	河北省	17235.480000	5764.900000
8	9	11	110000	110000	北京市	12153.030000	5309.900000
9	27	14	140000	140000	山西省	7358.310000	2809.000000
10	29	12	120000	120000	天津市	7521.850000	2430.800000
11	54	61	610000	610000	陕西省	8169.800000	2699.700000
12	65	64	640000	640000	宁夏回族自治区	1353.310000	339.300000

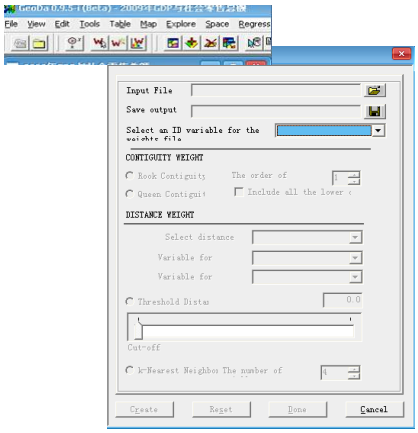
点击右键——新建或删除变量

输入数据

制图

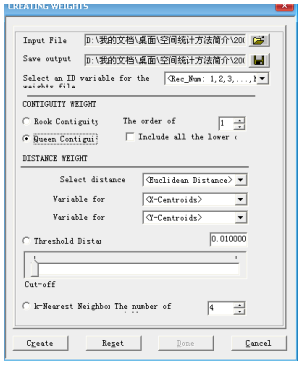


等级图

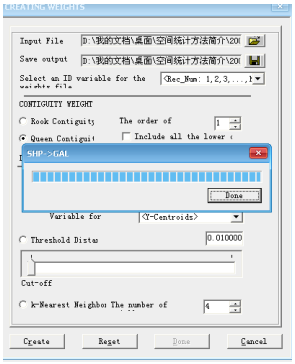


Tools_weights

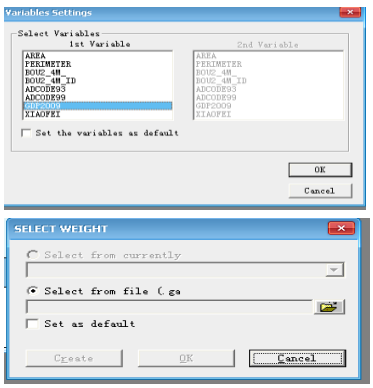
制作空间权重矩阵



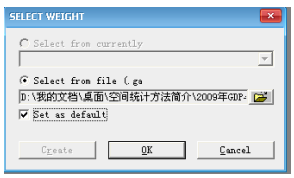
空间权重矩阵的确定方式

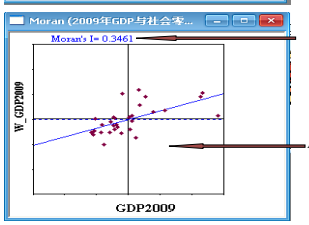


空间权重矩阵的确定方式



制作散点图 计算莫兰指数

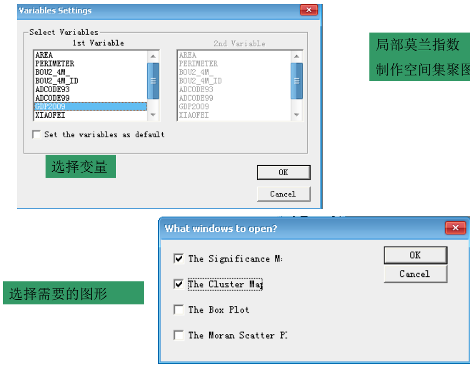




莫兰指数

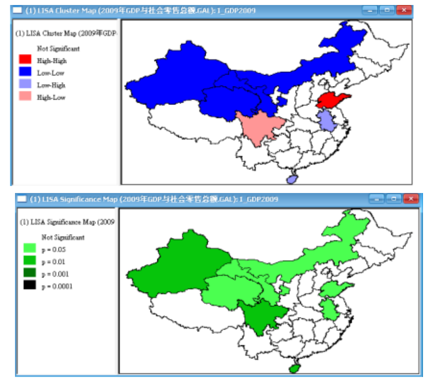
点击右键可以保存结果

局部莫兰指数
制作空间集聚图



选择变量

选择需要的图形



空间集聚图

点击右键选择保存结果

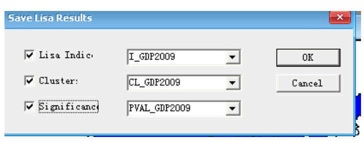
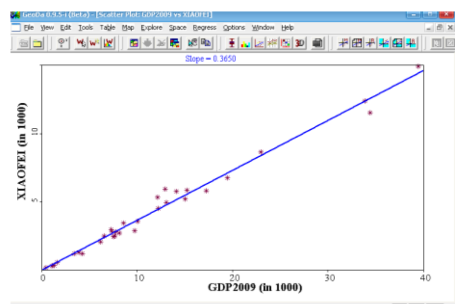


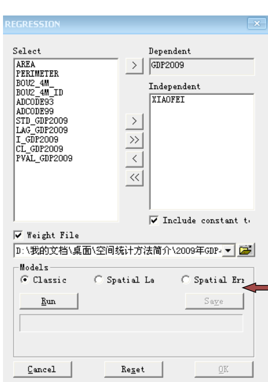
Table : 2009年GDP与社会零售总额

	GDP2009	XIAOFEI	STD_GDP2009	LAG_GDP2009	I_GDP2009	CL_GDP2009	PVAL_GDP2009
1	8587.000000	3401.800000	-0.328557	-0.336522	0.110566	0.000000	0.354000
2	9740.250000	2855.300000	-0.210037	-0.592117	0.124366	2.000000	0.032000
3	4277.050000	1177.500000	-0.771493	-1.042841	0.804545	2.000000	0.004000
4	7278.750000	2957.300000	-0.463007	-0.269297	0.124686	0.000000	0.372000
5	15212.490000	5812.600000	0.352349	0.560253	0.197405	0.000000	0.170000
6	3387.560000	1183.000000	-0.862907	-0.546928	0.471948	2.000000	0.040000
7	17235.480000	5764.900000	0.560253	0.887603	0.497282	0.000000	0.076000
8	12153.030000	5309.900000	0.037926	0.061115	0.002318	0.000000	0.384000
9	7358.310000	2809.000000	-0.454830	0.069834	-0.031763	0.000000	0.398000
10	7521.850000	2430.800000	-0.438023	0.037926	-0.016613	0.000000	0.380000
11	8169.800000	2699.700000	-0.371433	-0.248058	0.092137	0.000000	0.242000
12	1353.310000	339.300000	-1.071968	-0.481459	0.516108	0.000000	0.164000

计算结果

拟合曲线（GDP与消费总额）





回归分析

选择回归分析方式

普通最小二乘法

```
REGRESSION
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : 2009ÄêGDPÓëÉç»áÁãÊÛ×Ü¶î
Dependent Variable  :     GDP2009  Number of Observations:   31
Mean dependent var  :       11784  Number of Variables   :    2
S.D. dependent var  :     9572.17  Degrees of Freedom    :   29

R-squared           :    0.984775  F-statistic           :     1875.77
Adjusted R-squared  :    0.984250  Prob(F-statistic)     :6.56379e-028
Sum squared residual:4.32452e+007  Log likelihood        :    -263.287
Sigma-square        :1.49121e+006  Akaike info criterion :     530.575
S.E. of regression  :     1221.15  Schwarz criterion     :     533.443
Sigma-square ML     :  1.395e+006
S.E of regression ML:      1181.1
-----------------------------------------------------------------------
    Variable    Coefficient       Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT      235.3303        345.2619      0.6815994     0.5008997
     XIAOFEI      2.698313      0.06230203        43.3102     0.0000000
-----------------------------------------------------------------------
```

空间滞后模型

$$Y = \rho WY + X\beta + \varepsilon$$

```
REGRESSION
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : 2009ÄêGDPÓëÉç»áÁãÊÛ×Ü¶î
Spatial Weight      : 2009ÄêGDPÓëÉç»áÁãÊÛ×Ü¶î.GAL
Dependent Variable  :      GDP2009  Number of Observations:   31
Mean dependent var  :        11784  Number of Variables   :    3
S.D. dependent var  :      9572.17  Degrees of Freedom    :   28
Lag coeff.   (Rho)  :    0.0791261

R-squared           :     0.986290  Log likelihood        :    -261.693
Sq. Correlation     : -             Akaike info criterion :     529.386
Sigma-square        :1.25617e+006   Schwarz criterion     :     533.688
S.E of regression   :      1120.79

----------------------------------------------------------------------
    Variable     Coefficient      Std.Error     z-value    Probability
----------------------------------------------------------------------
   W_GDP2009     0.07912609     0.04126981     1.917288    0.0552013
    CONSTANT      -306.5877       424.3664   -0.7224599    0.4700116
     XIAOFEI       2.618238     0.07096619     36.89416    0.0000000
----------------------------------------------------------------------
```

空间误差模型

$$y = X\beta + \varepsilon$$
$$\varepsilon = \lambda W\varepsilon + \mu$$

```
REGRESSION
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : 2009ÄêGDPÓëÉç»áÁãÊÛ×Ü¶î
Spatial Weight      : 2009ÄêGDPÓëÉç»áÁãÊÛ×Ü¶î.GAL
Dependent Variable  :      GDP2009  Number of Observations:   31
Mean dependent var  :11783.990000   Number of Variables   :    2
S.D. dependent var  : 9572.173734   Degree of Freedom     :   29
Lag coeff. (Lambda) :    -0.341708

R-squared           :     0.986435  R-squared (BUSE)      : -
Sq. Correlation     : -             Log likelihood        : -262.057318
Sigma-square        :1242949.038647  Akaike info criterion :     528.115
S.E of regression   :      1114.88  Schwarz criterion     :  530.982611

----------------------------------------------------------------------
    Variable     Coefficient      Std.Error     z-value    Probability
----------------------------------------------------------------------
    CONSTANT       120.0422       264.2762    0.4542299    0.6496634
     XIAOFEI       2.727504     0.05080515     53.68557    0.0000000
      LAMBDA     -0.3417083      0.2020726    -1.691018    0.0908334
----------------------------------------------------------------------
```

谢谢！