EFFECTS OF VARIOUS SOCIO-ECONOMIC INDICATORS ON THE SPATIAL DISTRIBUTION OF AVERAGE INCOME IN HELSINKI REGION.

BY: OYEDAYO OYELOWO

STUDENT NUMBER : 014717208

Variable Names:

HR_KTU - Average income of inhabitants, 2014 (HR). --------- RESPONSE VARIABLE

HE_KIKA: Average age of inhabitants, 2015 (HE) ---------- PREDICTORS

HR_OVY: Accumulated purchasing power of inhabitants, 2014 (HR)

KO_KOUL: With education, total, 2014 (KO)

PT_TYOTT: Unemployed, 2014 (PT)

TP_PALV_GU : Services, 2014 (TP)

TP_O_JULK : Public administration and defence; compulsory social security, 2014 (TP)

KO_YL_KORK : Academic degree - Higher level university degree, 2014 (KO)

KO_AMMAT : Vocational diploma, 2014 (KO)

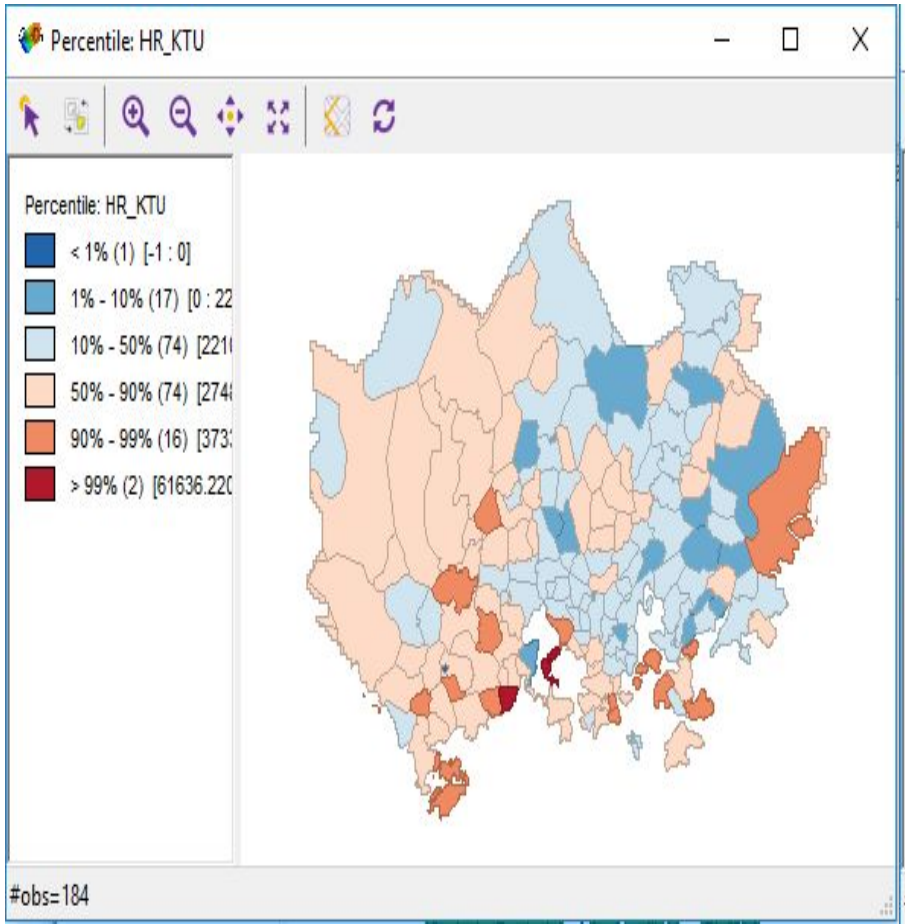
HE_MIEHET : Males, 2015 (HE)

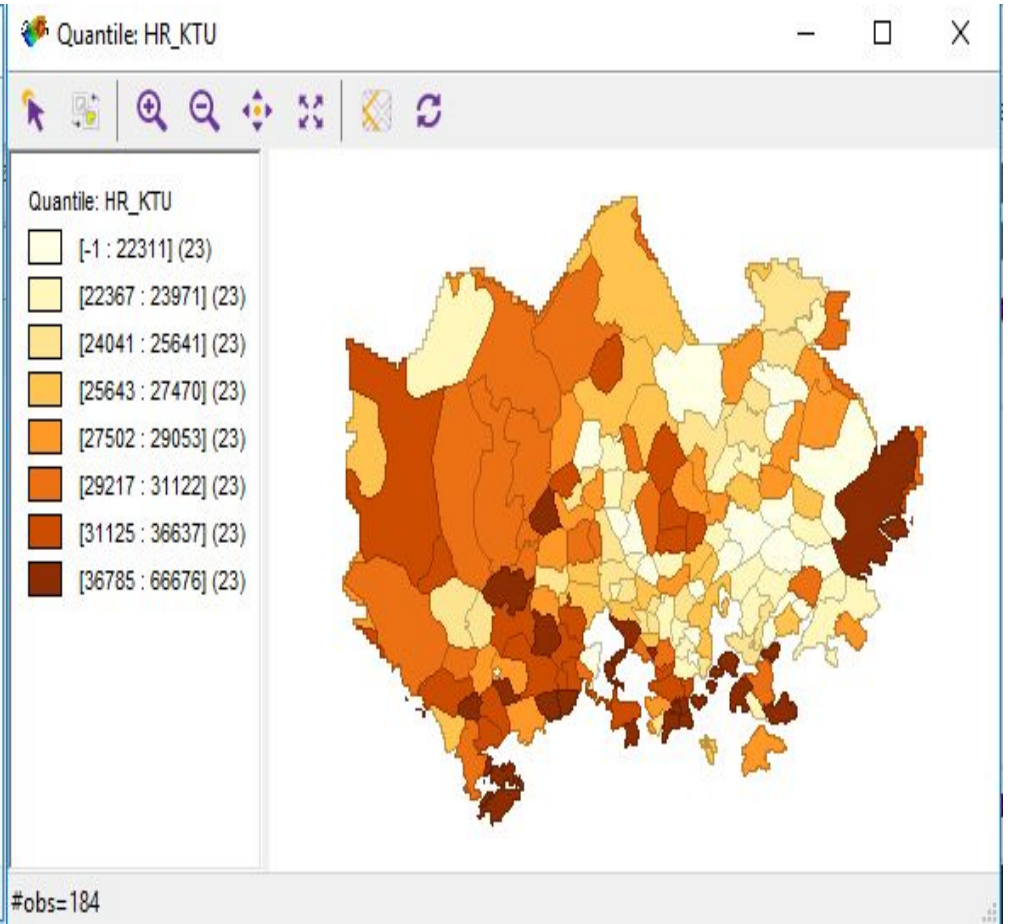
KO_PERUS: Basic level studies, 2014 (KO)

Research Questions

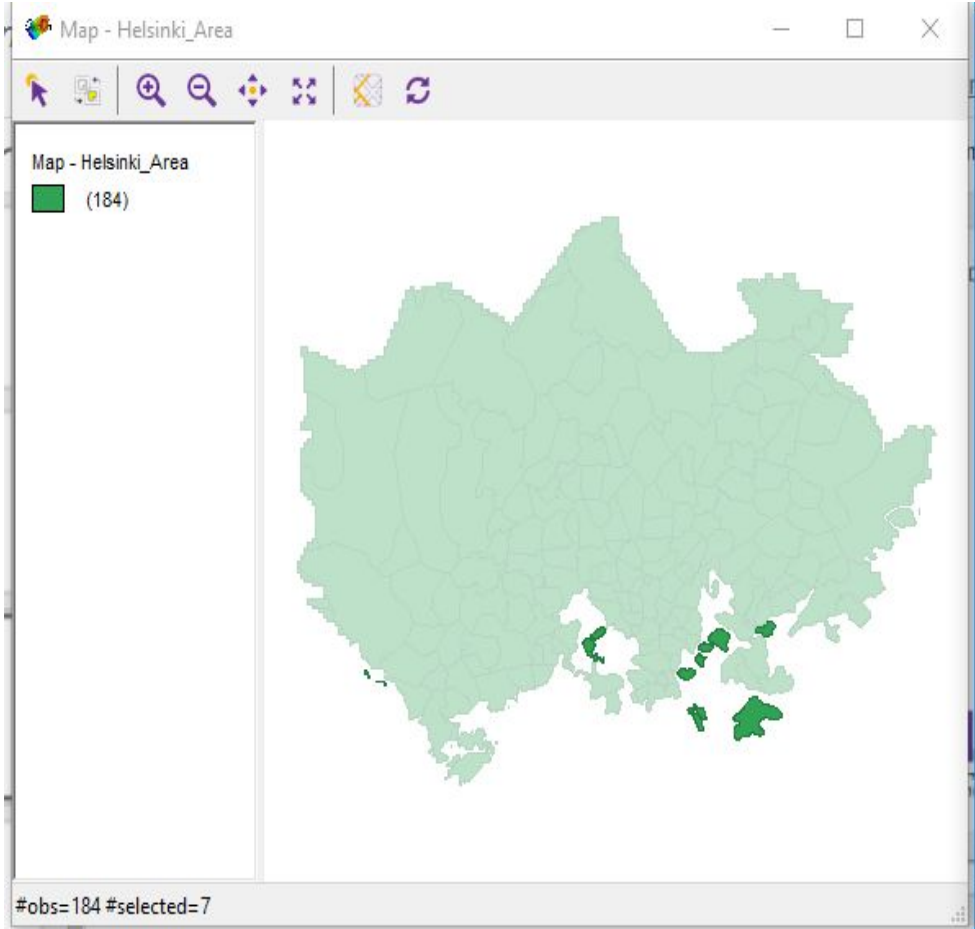
- Does Average income vary spatially in Helsinki region?
- If so, how much does it vary?
- What are the socioeconomic variables affecting average income in the region.
- How do their effects vary across the region?

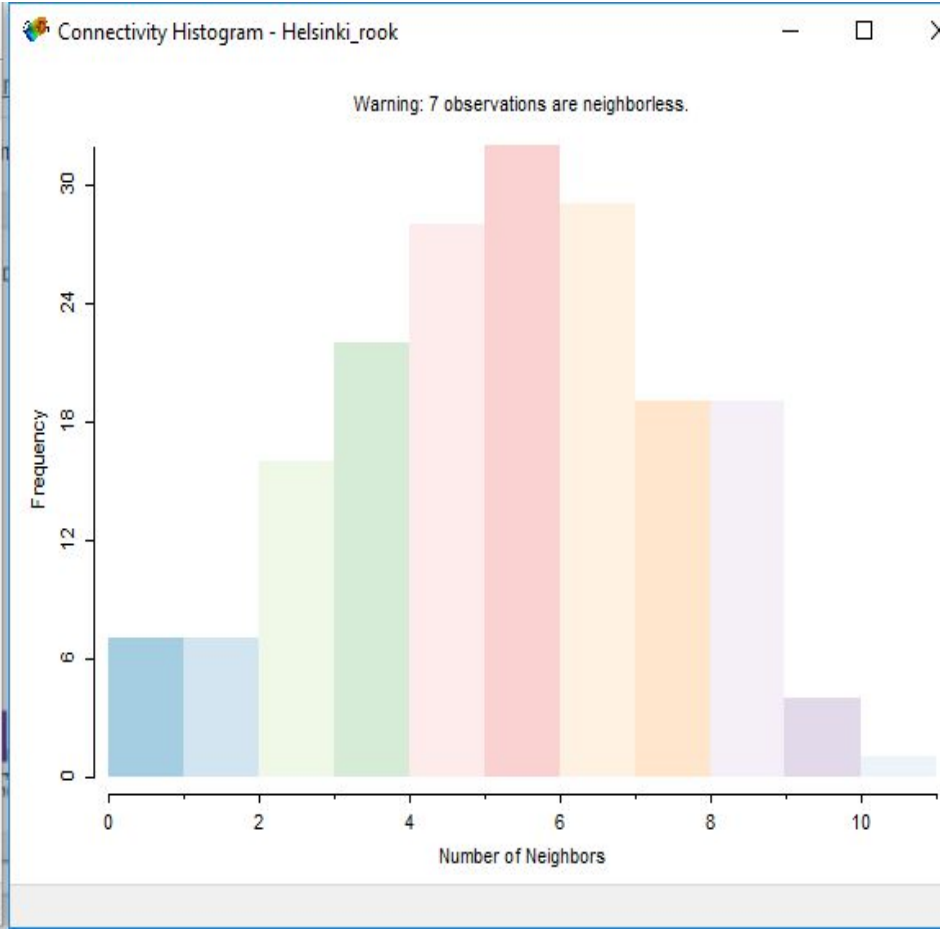
Average Income of Inhabitants, 2014



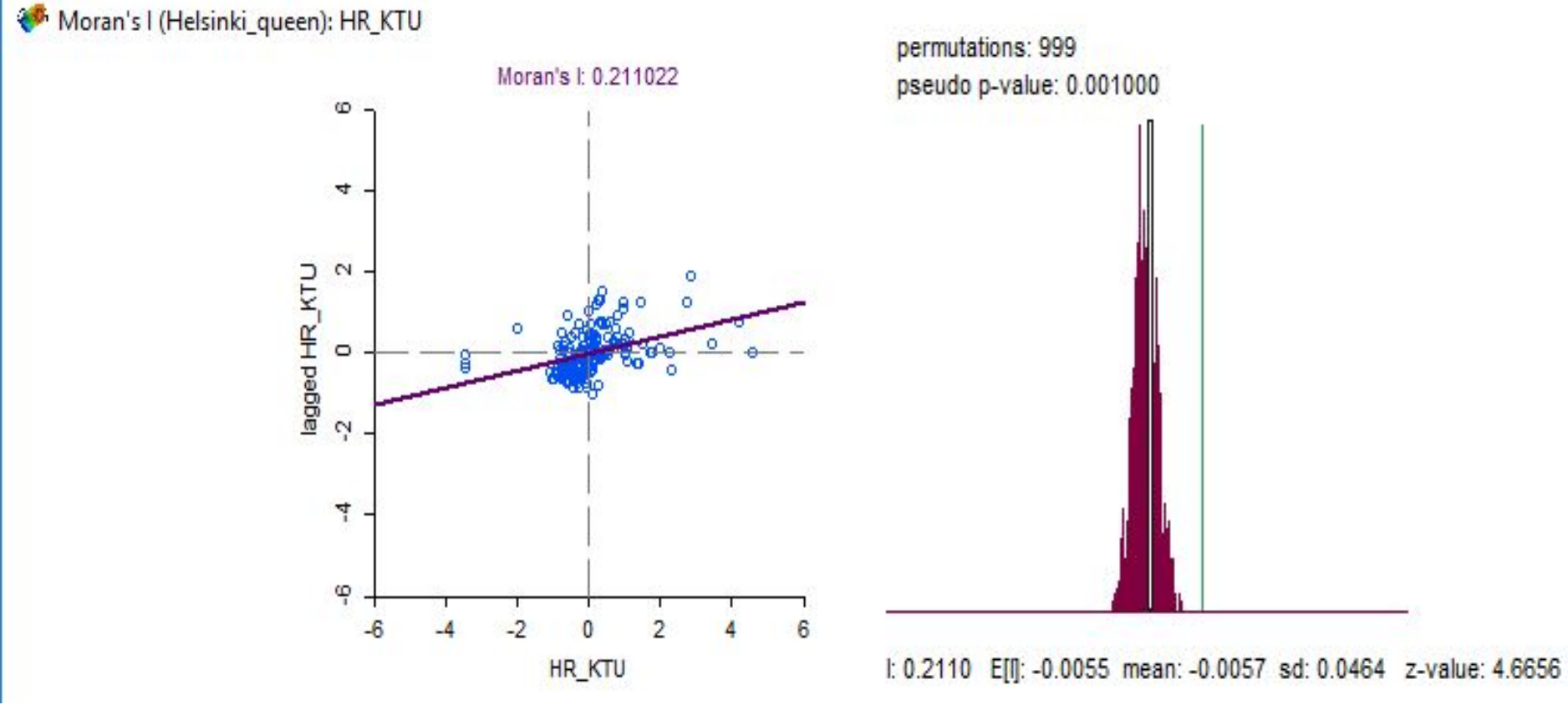


Queen Neighbourhood

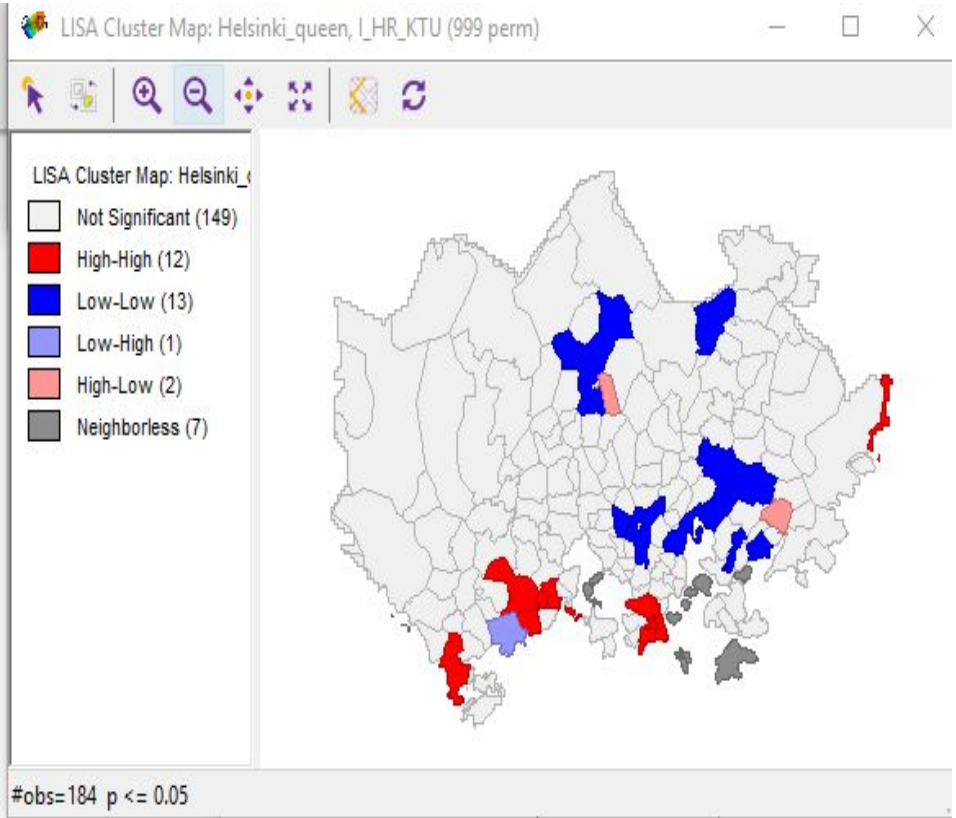


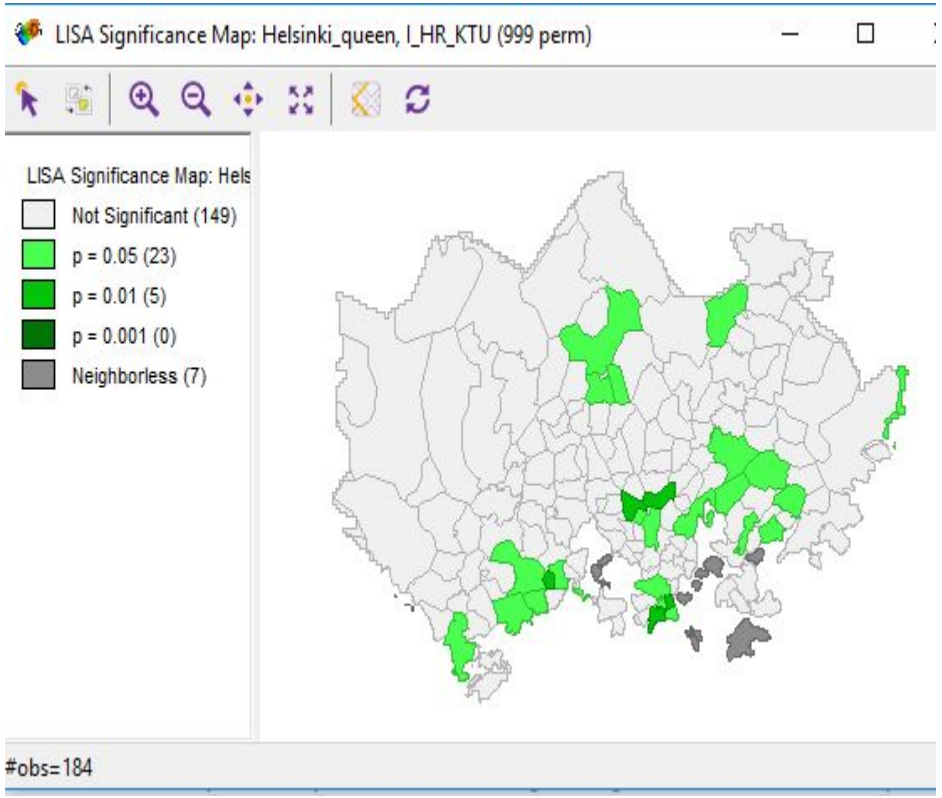


The Spatial Autocorrelation: Global Moran's I



Spatial Autocorrelation





Regression Report

```
>>02/21/18 18:02:42
REGRESSION
----------
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            :  Helsinki_Area
Dependent Variable  :      HR_KTU  Number of Observations:   184
Mean dependent var  :     28574.5  Number of Variables   :    11
S.D. dependent var  :     8334.06  Degrees of Freedom    :   173

R-squared           :    0.688250  F-statistic           :     38.1932
Adjusted R-squared  :    0.670230  Prob(F-statistic)     :9.80587e-039
Sum squared residual:3.98417e+009  Log likelihood        :    -1815.03
Sigma-square        :2.30299e+007  Akaike info criterion :     3652.05
S.E. of regression  :     4798.94  Schwarz criterion     :     3687.41
Sigma-square ML     :2.16531e+007
S.E of regression ML:     4653.29

--------------------------------------------------------------------------------
      Variable      Coefficient      Std.Error    t-Statistic    Probability
--------------------------------------------------------------------------------
      CONSTANT        4346.09         2534.63        1.71468        0.08819
       KO_KOUL       -5.36512         1.44481       -3.71337        0.00028
        HR_OVY   0.000249703   2.59412e-005        9.62573        0.00000
       HE_KIKA        646.674         65.4462        9.88101        0.00000
       PT_TYOTT        16.8955         6.12208        2.75977        0.00641
     TP_PALV_GU      -0.298818        0.122209       -2.44514        0.01548
      TP_O_JULK         2.2154        0.852248        2.59948        0.01014
     KO_YL_KORK       -11.6573         3.60801       -3.23095        0.00148
```

Regression Report

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	4346.09	2534.63	1.71468	0.08819
KO_KOUL	-5.36512	1.44481	-3.71337	0.00028
HR_OVY	0.000249703	2.59412e-005	9.62573	0.00000
HE_KIKA	646.674	65.4462	9.88101	0.00000
PT_TYOTT	16.8955	6.12208	2.75977	0.00641
TP_PALV_GU	-0.298818	0.122209	-2.44514	0.01548
TP_O_JULK	2.2154	0.852248	2.59948	0.01014
KO_YL_KORK	-11.6573	3.60801	-3.23095	0.00148
KO_AMMAT	-8.85584	2.66008	-3.32916	0.00106
HE_MIEHET	5.29272	1.52869	3.46226	0.00067
KO_PERUS	-9.05561	2.18957	-4.13579	0.00006

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER 76.038414
TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	433.9371	0.00000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	10	73.6672	0.00000
Koenker-Bassett test	10	17.0145	0.07404

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER 76.038414
TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	433.9371	0.00000

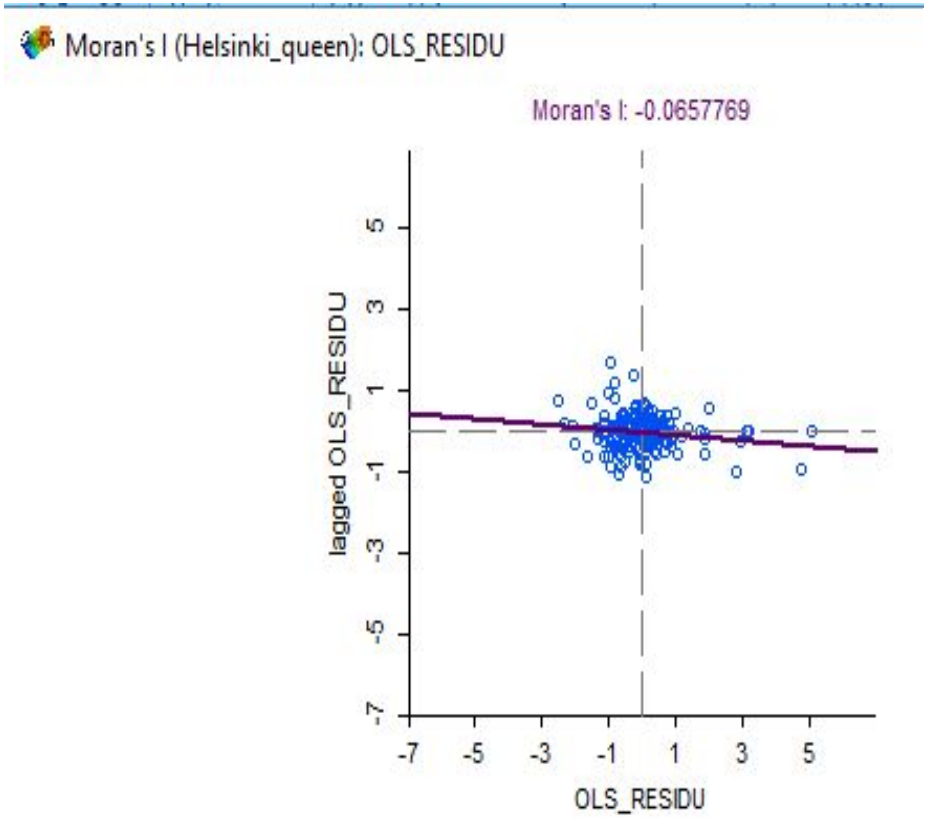
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS

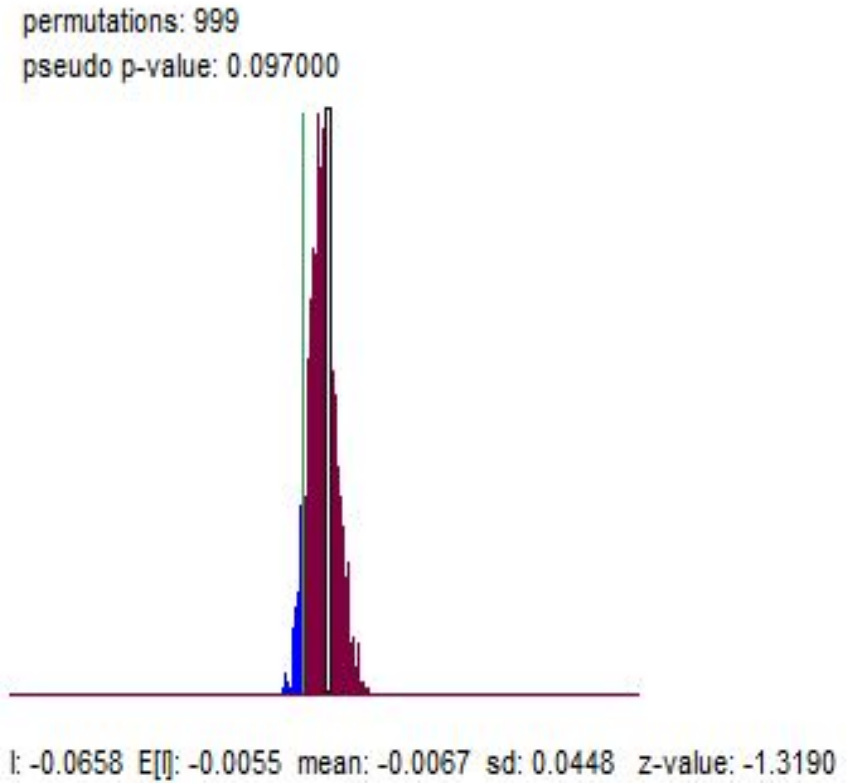
TEST	DF	VALUE	PROB
Breusch-Pagan test	10	73.6672	0.00000
Koenker-Bassett test	10	17.0145	0.07404

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX : Helsinki_queen
(row-standardized weights)

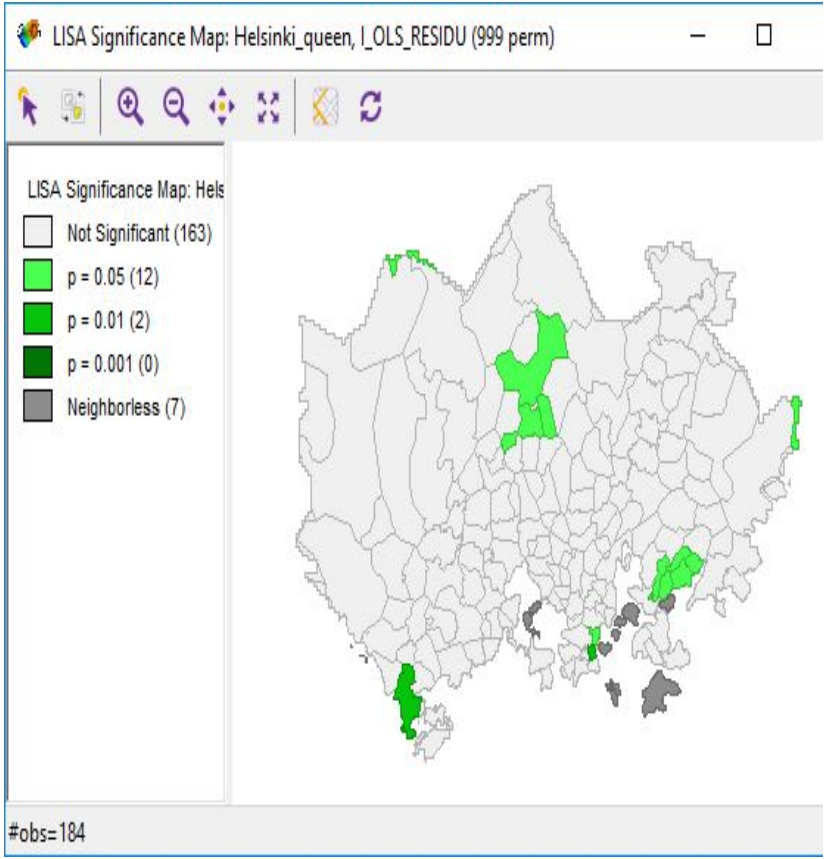
TEST	MI/DF	VALUE	PROB
Moran's I (error)	-0.0658	-1.0330	0.30163
Lagrange Multiplier (lag)	1	9.4458	0.00212
Robust LM (lag)	1	7.6527	0.00567
Lagrange Multiplier (error)	1	1.7933	0.18052
Robust LM (error)	1	0.0001	0.99132
Lagrange Multiplier (SARMA)	2	9.4460	0.00889

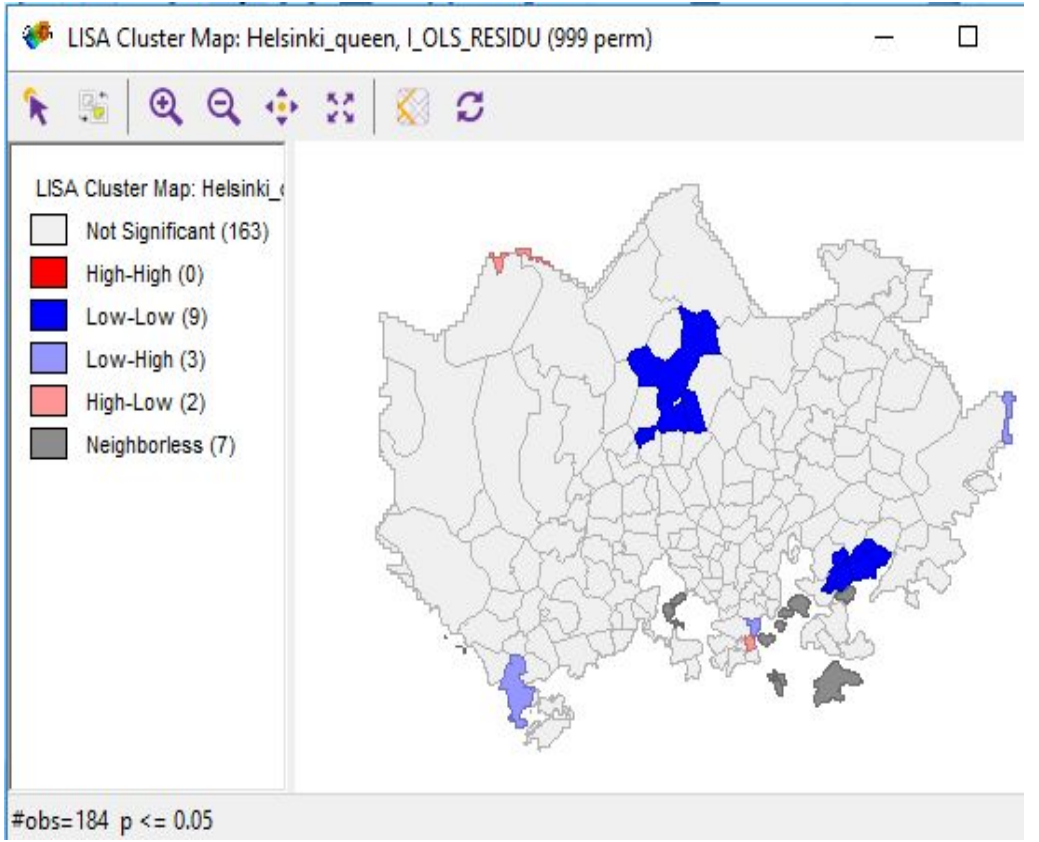
Spatial Autocorrelation of error





Cluster of the Residuals





Report: Spatial Lag Model

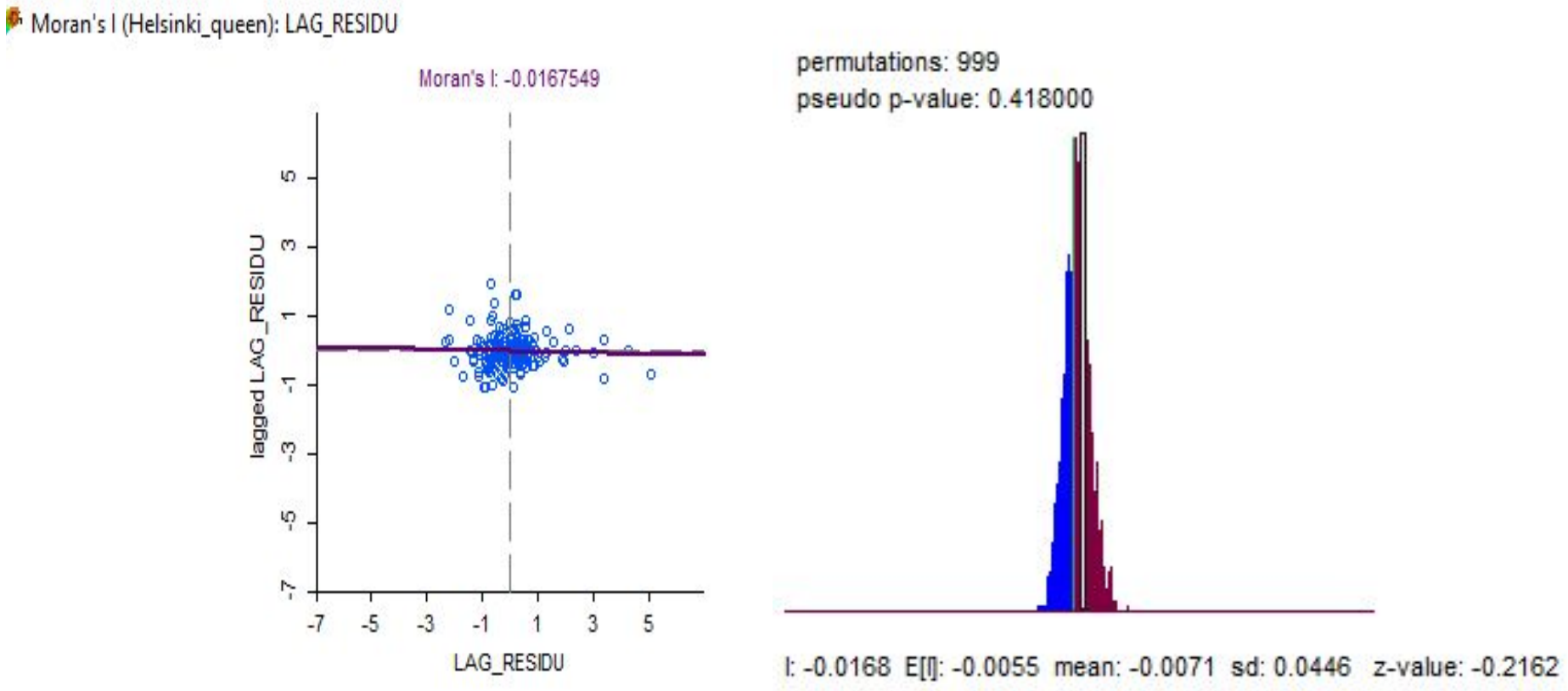
```
>>02/21/18 18:40:06
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : Helsinki_Area
Spatial Weight      : Helsinki_queen
Dependent Variable  :      HR_KTU  Number of Observations:  184
Mean dependent var  :     28574.5  Number of Variables   :   12
S.D. dependent var  :     8334.06  Degrees of Freedom    :  172
Lag coeff.   (Rho)  :   -0.161727

R-squared           :    0.707028  Log likelihood        :   -1809.77
Sq. Correlation     : -            Akaike info criterion :    3643.53
Sigma-square        :2.03488e+007  Schwarz criterion     :    3682.11
S.E of regression   :     4510.97


REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                 DF       VALUE         PROB
Breusch-Pagan test                   10        72.2687      0.00000

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : Helsinki_queen
TEST                                 DF       VALUE         PROB
Likelihood Ratio Test                1         10.5192      0.00118
```

Regression Report

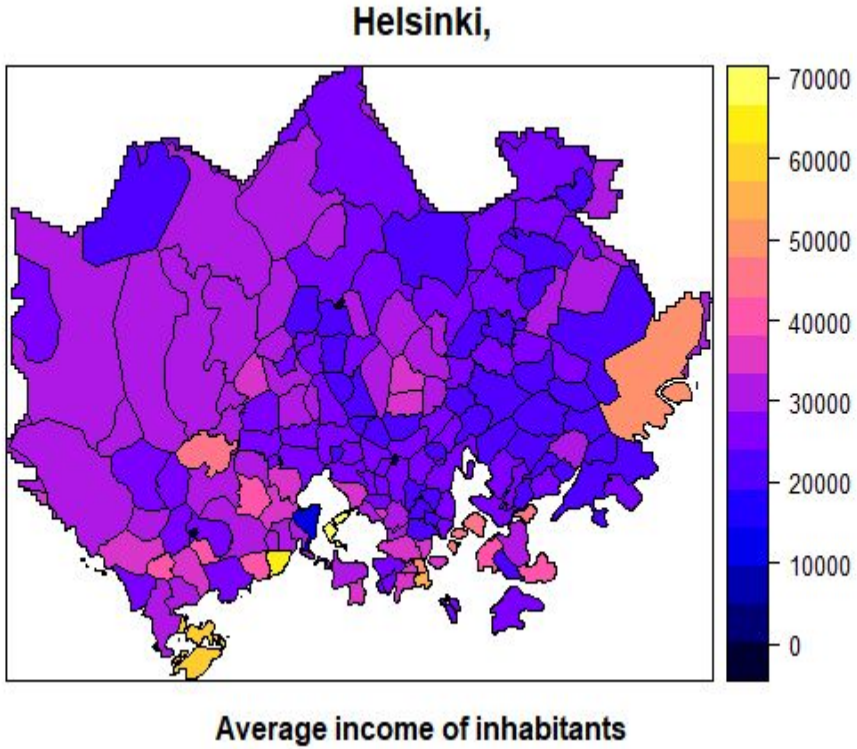
```
Data set            : Helsinki_Area
Spatial Weight      : Helsinki_queen
Dependent Variable  :      HR_KTU  Number of Observations:  184
Mean dependent var  :     28574.5  Number of Variables   :   12
S.D. dependent var  :     8334.06  Degrees of Freedom    :  172
Lag coeff.   (Rho)  :   -0.161727

R-squared           :    0.707028  Log likelihood        :   -1809.77
Sq. Correlation     : -            Akaike info criterion :    3643.53
Sigma-square        :2.03488e+007  Schwarz criterion     :    3682.11
S.E of regression   :     4510.97
```

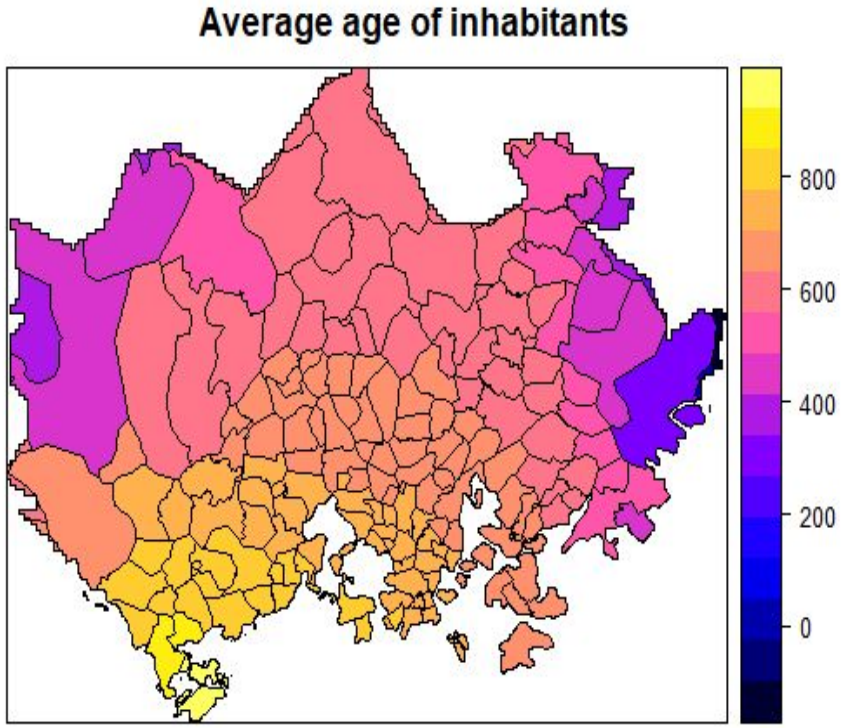
Variable	Coefficient	Std.Error	z-value	Probability
W_HR_KTU	-0.161727	0.0483448	-3.34528	0.00082
CONSTANT	8173.11	2672.99	3.05766	0.00223
KO_KOUL	-5.86134	1.36217	-4.30294	0.00002
HR_OVY	0.00025318	2.43848e-005	10.3827	0.00000
HE_KIKA	651.037	61.52	10.5825	0.00000
PT_TYOTT	17.2476	5.7555	2.99672	0.00273
TP_PALV_GU	-0.27899	0.11509	-2.4241	0.01535
TP_O_JULK	2.06764	0.803007	2.57488	0.01003
KO_YL_KORK	-11.1736	3.39438	-3.29178	0.00100
KO_AMMAT	-8.93677	2.50054	-3.57394	0.00035
HE_MIEHET	5.88976	1.44263	4.08264	0.00004
KO_PERUS	-9.44248	2.06069	-4.58219	0.00000

Spatial Lag Residual

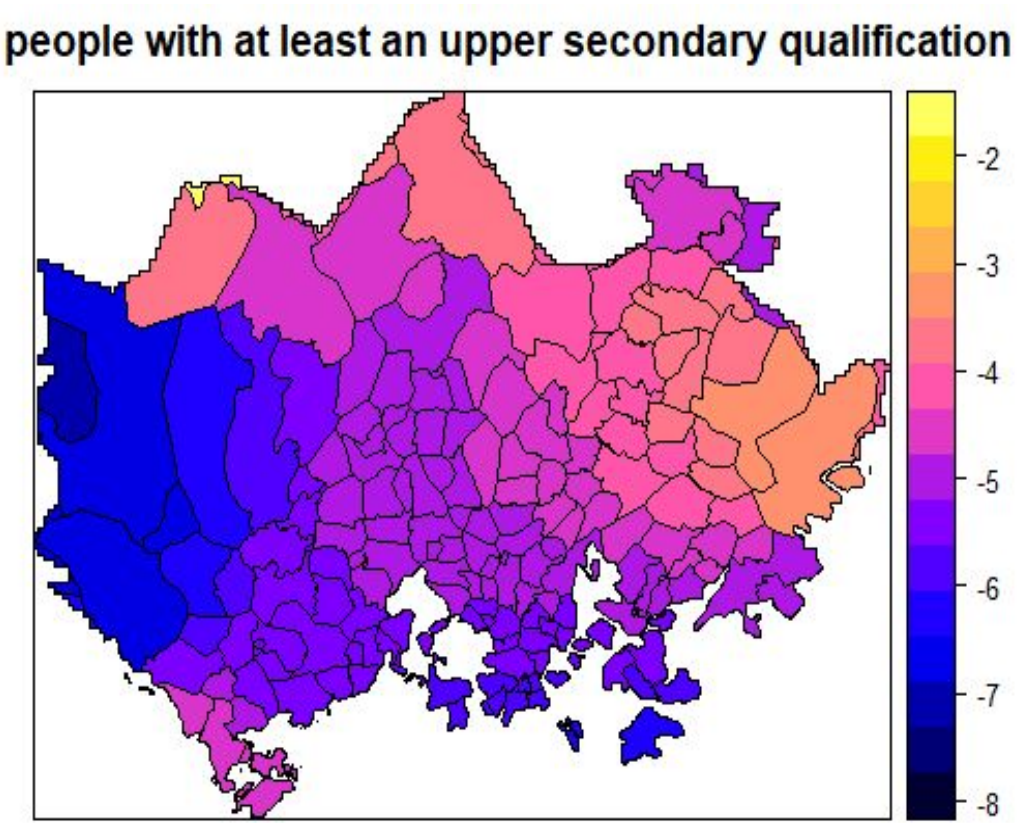


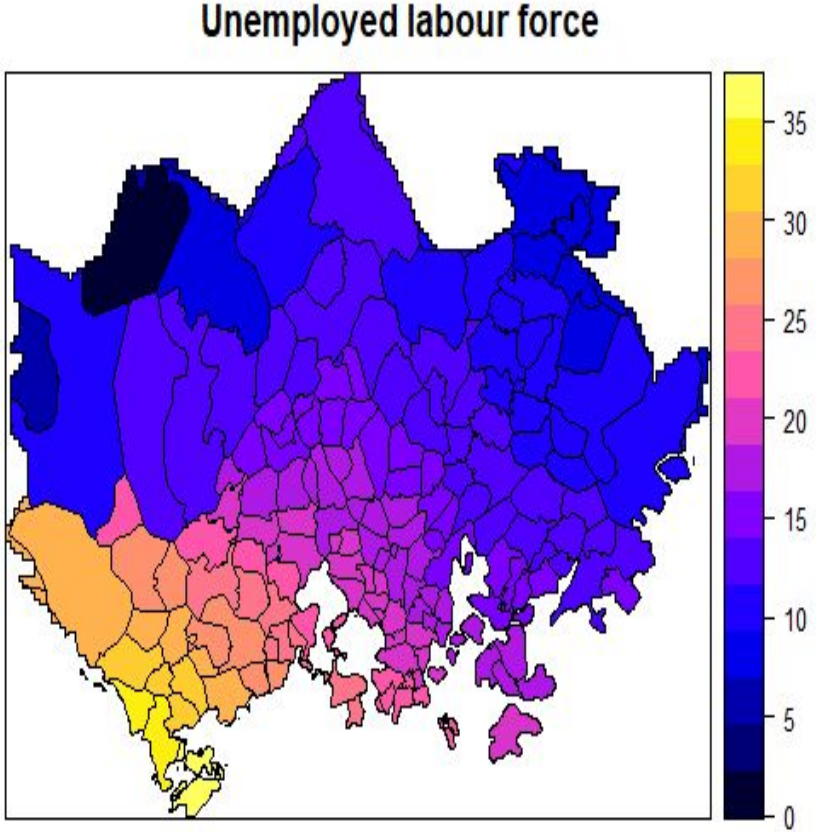
Geographically Weighted Regression



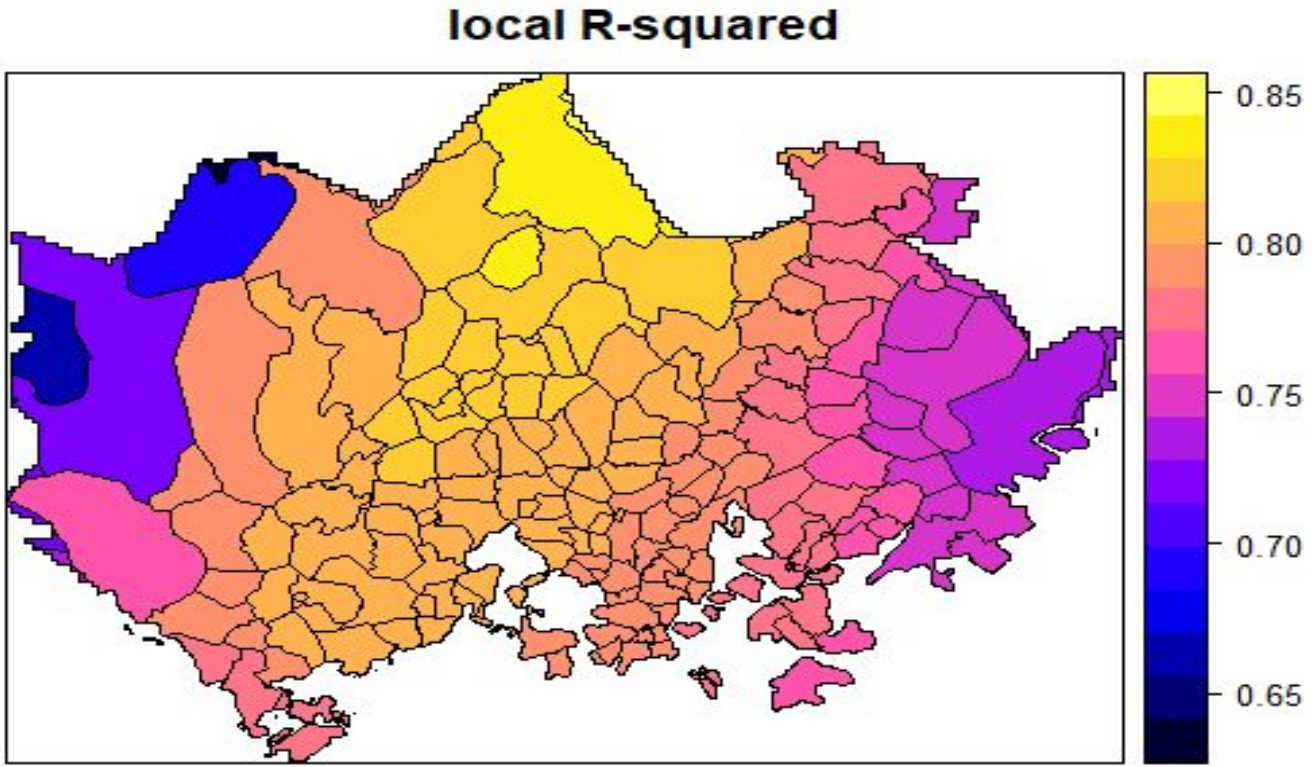


Geographically Weighted Regression

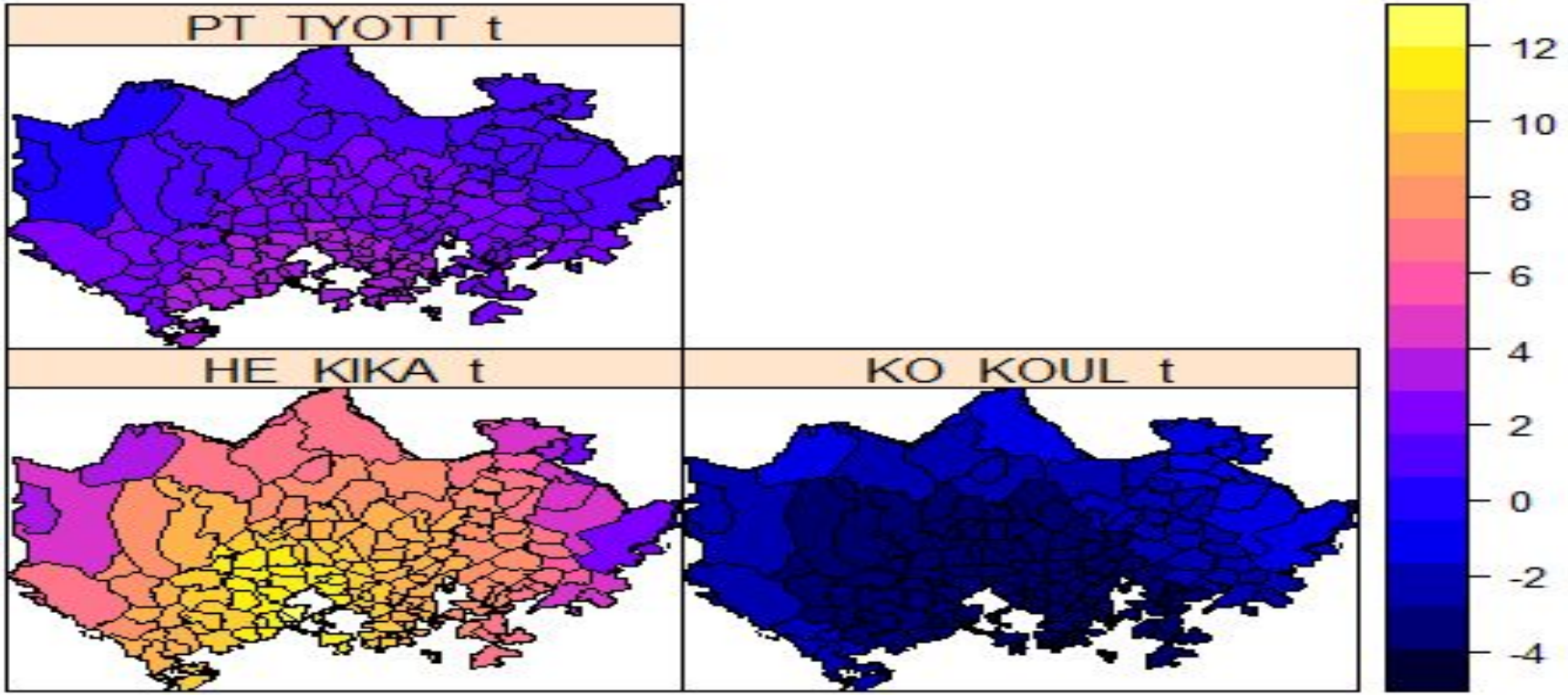




Geographically Weighted Regression

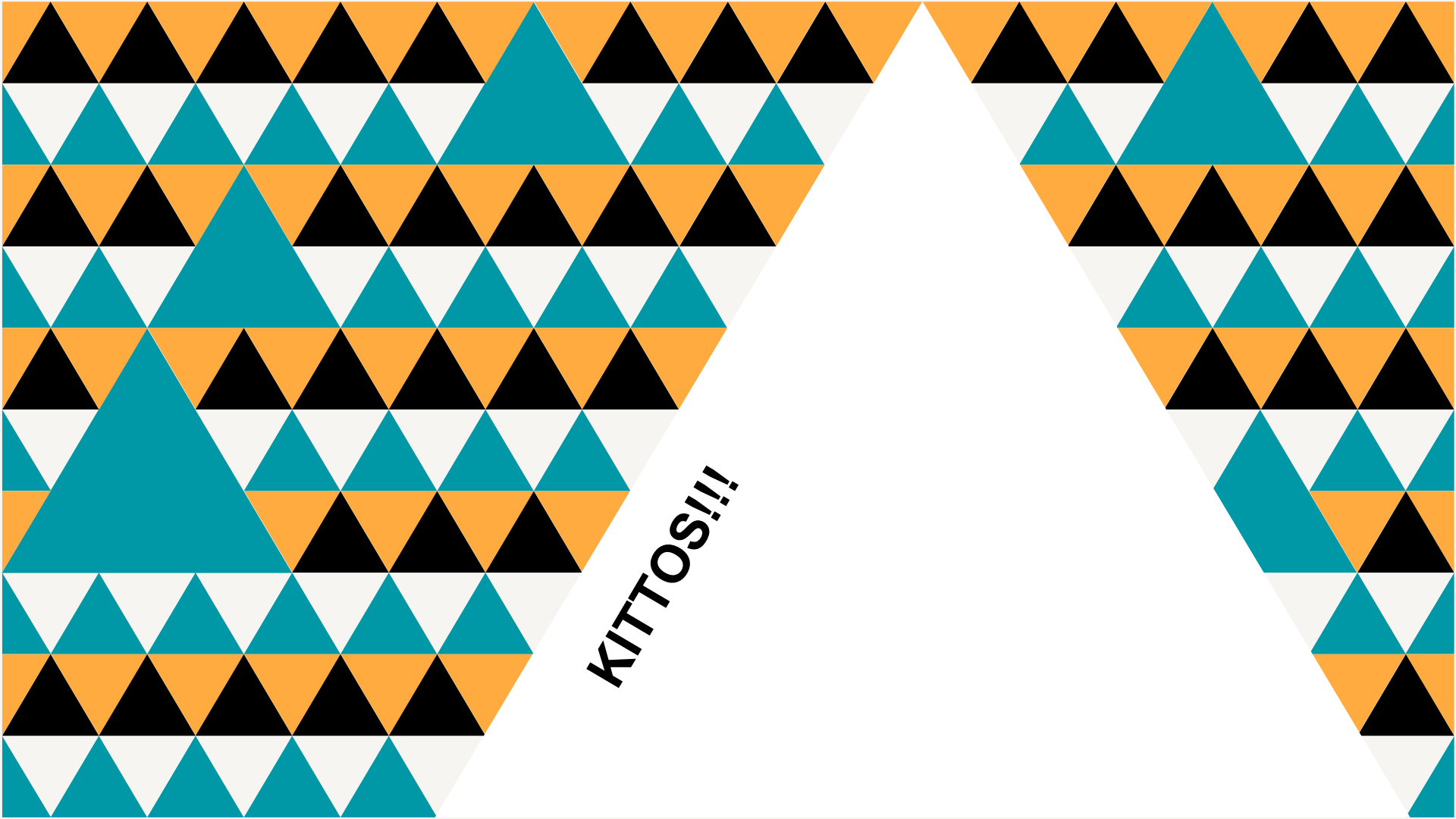


Geographically Weighted Regression



Conclusion:

- Coefficient of determination : 70%
- Income is generally higher in the west than the east..
- Effects of education, age and unemployment on spatial distribution of average income, are highest at the centre compared to the outer regions.
- Average age and Unemployed labour force have more impacts on average income in the southwestern area
- Education has more impact in the northeastern part.


KITTOS!!!