**FACULDADE DE SAÚDE PÚBLICA**

**DISCIPLINA: EPI5707 – ANÁLISE ESPACIAL EM EPIDEMIOLOGIA**

**Professor: Francisco Chiaravalloti Neto**

**Monitores: Camila Meireles Fernandes e Raquel Gardini Sanches Palasio**

**Aluno(a):**

**TRABALHO FINAL – PROPOSTO PELO PROFESSOR**

**Faça o que está sendo pedido nas questões abaixo. Responda as questões em um documento salvo em \*.pdf ou \*doc, identificando o documento com as iniciais do seu nome e informando que se trata do trabalho final proposto pelo professor (Ex.: ACGP\_TR\_FINAL\_PROP\_PROF.pdf ou .doc). Ao término, submeta o trabalho no site da disciplina máximo até.**

**TRABALHO PROPOSTO**

Temos um shape com os municípios do estado de São Paulo (“esp\_acid\_escorp\_ESP\_2008\_12.shp”), com informações sobre as taxas de incidências de acidentes escorpiônicos relativas aos anos de 2008 a 2012 e possíveis covariáveis explicativas. Os campos presentes no banco são os seguintes:

Cod\_6 Código do município

Municipio Município

TX\_URB Urbanização em percentagem

AREA área do município (km²)

VEG.NAT Vegetação Natural (hectares)

VEG.NAT.PC Vegetação Natural (percentagem)

AR.URB área urbana (hectares)

AR.URB.PCT área urbana (percentagem)

NDVI Normalized Difference Vegetation Index

GINI Índice Gini

PIB Renda per capita

T\_AGUA Proporção da população com água encanada

T\_LIXO Proporção da população com coleta de lixo

T\_LUZ Proporção da população com eletricidade

IDHM índice de desenvolvimento humano municipal

PREC.TOT Precipitação (mm)

TEMP.MAX temperatura máxima (°C)

UMID.RELAT Umidade Relativa

Casos\_2008\_1 Número de acidentes escorpiônicos entre 2008 e 2012

Pop\_2010 população total em 2010

ATIVIDADES

1. Importe para o R o shape “esp\_acid\_escorp\_ESP\_2008\_12.shp”, informe a classe deste objeto, e obtenha, no R, um plot dos municípios do estado de São Paulo

2. Calcule a taxa de incidência anual de acidentes escorpônicos entre 2008 e 2012 (tx\_incid) e, usando o comando ‘summary’, apresente, para cada variável do banco, seu resumo estatístico (OBS.: a taxa anual é calculada dividindo-se o número de casos pela população do meio do período e pelo número de anos – pode usar uma base multiplicativa de 10 mil ou 100 mil)

3. Realize, utilizando o conjunto de funções ‘HighstatLib’, análise exploratória dos dados avaliando a existência de outliers na variável dependente (tx\_incid) e nas independentes, a colinearidade entre as covariáveis e a relação entre cada covariável e a variável dependente. Com base nessas análises, adote as providências (transformação das covariáveis, eliminação da colinearidade, etc) que achar necessárias.

4. Faça a modelagem da variável dependente por regressão OLS e apresente seus resultados. Considere, neste caso, que o melhor modelo é aquele com o menor valor de AIC. Para isso, utilize o comando ‘step’. Interprete os resultados encontrados.

5. Obtenha, no R, uma matriz de vizinhança que você considere adequada e obtenha o valor do I de Moran Global dos resíduos da regressão OLS e seu respectivo valor de p. Avalie, utilizando o diagnóstico do multiplicador de Lagrange, se haverá necessidade de considerar um modelo de regressão espacial.

6. Uma vez identificada a necessidade de considerar um modelo espacial, verifique qual modelagem será a mais adequada (defasagem ou erro espacial), realize a modelagem e apresente os resultados. Avalie se os resíduos dessa modelagem ainda têm dependência espacial ou não. Avalie o comportamento dos resíduos em relação à normalidade e homocedasticidade. Interprete os resultados encontrados.

7. Caso os resíduos da regressão espacial ainda tenham apresentado dependência espacial, rode a modelagem utilizando a técnica de regressão espacialmente ponderada (com distribuição de probabilidade normal). Avalie se os resíduos dessa modelagem ainda têm dependência espacial ou não. Avalie o comportamento dos resíduos em relação à normalidade e homocedasticidade. Mapeie os coeficientes de regressão encontrados e interprete os resultados encontrados.