**FACULDADE DE SAÚDE PÚBLICA**

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

**Professor: Francisco Chiaravalloti Neto**

**Monitores: Alec Brian Lacerda 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\_2018.shp”), com informações sobre as taxas de incidências de acidentes escorpiônicos relativas ao ano de 2018 e possíveis covariáveis explicativas. Os campos presentes no banco são os seguintes:

Cod\_6 Code of the municipality

Municipio Municipality

obs\_tt Number of scorpio accidents

Pop\_tt Total population

tx\_incid Incidence rate

TX\_URB Urbanization in percentage

AREA area of the municipality (km²)

VEG.NAT Natural Vegetation (in hectares)

VEG.NAT.PC Natural Vegetation (in percentage)

AR.URB Urban area (in hectares)

AR.URB.PCT Urban area (in percentage)

NDVI Normalized Difference Vegetation Index

GINI Gini index

PIB Per capita income

T\_AGUA Proportion of the population with piped water

T\_LIXO Proportion of the population with garbage collection

T\_LUZ Proportion of the population with electricity

IDHM Municipal human development index

PREC.TOT Precipitation (mm)

TEMP.MAX Maximum temperature (°C)

UMID.RELAT Relative humidity

ATIVIDADES

1. Importe para o R o shape “esp\_acid\_escorp\_ESP\_2018.shp”, informe a classe deste objeto, e obtenha, no R, um plot dos municípios do estado de São Paulo e, usando o comando ‘summary’, apresente, para cada variável do banco, seu resumo estatístico.

2. 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.

3. 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.

4. 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.

5. 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.

6. 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.