**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):**

**EXERCÍCIO - AULA 6 - ANÁLISE EXPLORATÓRIA DE DADOS**

**Faça o que está sendo pedido nas questões abaixo. Apresente, em pdf ou doc, o Relatório Dinâmico produzido com o Markdown no RStudio. Salve este pdf ou doc identificando os documentos com as iniciais do seu nome e o número do exercício (Ex.: “ACGP\_REL\_DINAM\_EXERC\_AULA5.pdf ou doc”) e envie-o para o e-mail da disciplina no prazo de uma semana.**

Utilizando o banco de dados com informações sobre o retardo do diagnóstico, faça uma análise exploratória de dados conforme segue:

1 – Identifique possíveis outliers na variável dependente (retardo no diagnóstico) e nas possíveis variáveis explanatórias contínuas ou fruto de contagem.

2 – Avalie se a variável dependente segue ou não uma distribuição de probabilidades normal.

3 – Avalie a proporção de zeros presentes na variável dependente.

4 – Avalie a colinearidade entre as possíveis variáveis explanatórias contínuas ou fruto de contagem, utilizando algum coeficiente correlação para avaliar a colinearidade. Na sequência, avalie a colinearidade de todas as covariáveis utilizando os VIFs.

5 – Avalie a relação entre a variável dependente e cada uma das variáveis explanatórias. Use a função ‘Myxyplot’ para as covariáveis contínuas ou fruto de contagem e o ‘boxplot’ para as variáveis categóricas.

6 – Avalie se as covariáveis categóricas estão adequadamente balanceadas.