Aula 8: Transformação de Variáveis e Interações

**Lista de Exercícios**

24 & 26 de setembro de 2022

Para realizar esta lista de exercícios, considere a base da dados que está no *Moodle*. Esta é uma base com dados reais, com os resultados das eleições nacionais de 2010 e 2014.

Suponha que uma pesquisadora deseje encontrar a proporção de votos por município para a eleição presidencial de 2010 (variável *PT\_2010*) como preditor da eleição em 2014 (variável *PT\_2014*). Ela acredita que o componente regional é importante para lidar com esses dados.

1. Utilizando as dummies para cada uma das regiões, se ficar mais simples, encontre a média geral de votos o PT em cada um dos anos e compare com as médias regionais. Comente;

library(tidyverse)

# para calcular a media de uma coluna:

mean(tabela$coluna)

# filtrar valores iguais a 1 em determinada coluna

# e calcular a média de outra variável:

dados %>%

 filter(coluna == 1) %>%

 summarize(media = mean(coluna2))

A pesquisadora monta um modelo de regressão bivariado.

1. Apresente os resultados deste modelo. Comente à luz dos dados descritivos obtidos no item anterior;

Agora adicione as variáveis de *dummy* de cada uma das regiões em seu modelo de forma a capturar quaisquer discrepâncias regionais.

1. O que aconteceu com seu modelo? Comente os motivos que levaram a este resultado.

++++++++

Uma questão importante a esta altura está na escolha da *dummy* que se tornará a referência do seu modelo. Esta decisão não muda as conclusões a que você chegará, mas pode facilitar a compreensão do resultado. Uma sugestão é tomar uma das categorias que faça sentido teórico ou aquela com valor destacado – como a de menor ou maior média.

1. Escolha uma das regiões como referência e rode seu modelo agora **sem** a variável da votação do PT em 2010. Interprete os resultados.
2. Compare os resultados obtidos agora com os dados descritivos obtidos no item 1. Quais semelhanças ou diferenças você encontra?
3. Introduza agora a variável de votação do PT em 2010 e rode novamente seu modelo. O que você observa com relação aos parâmetros estimados das *dummies*?
4. Compare o resultado obtido do modelo de regressão com um teste de médias. Quais semelhanças e diferenças encontramos aqui?

++++++++