**INSTRUÇÕES TRABALHO DE MODELAGEM ESTATÍSTICA AVANÇADA**

**Problema:**

Estabelecer um modelo linear para estimar valores futuros da Variável CO2 emissions (kt) (linha 382 da planilha). Note que essa variável será a Variável Resposta do modelo.

**1º MOMENTO: Análise Exploratória dos Dados (“limpeza da base de dados”)**

Essa análise exploratória já pode ser realizada com os conhecimentos adquiridos no MBA até o momento.

1. Utilize os dados de 1970 a 2014, sendo assim variáveis que contenham dados faltantes nesses anos podem ser descartadas.

2. Observem que na planilha há muitas variáveis repetidas, mudando somente a forma de obter os dados, dessa forma, pode-se considerar apenas um formato para esse tipo de variável.

3. Analise as correlações entre as variáveis explicativas e entre as variáveis explicativas e a variável resposta. Lembrem que variáveis explicativas com alta correlação geram o problema de multicolinearidade. Além disso, variáveis com baixa correlação com a variável resposta dificilmente terá uma contribuição significativa no modelo.

**2º MOMENTO: Análise de Dados (Modelagem)**

Para ser feito depois da 4ª aula de Modelagem Estatística Avançada, pois ainda faltam alguns modelos para ser vistos e que podem se ajustar bem aos dados. Após o segundo encontro, vc´s estarão mais aptos para fazer essa parte do trabalho. No entanto, não implica que já não possam tentar os modelos que já foram estudados em sala de aula.

1. Estabelecer um modelo linear para estimar valores futuros da Variável CO2 emissions (kt). Lembrem que todos os pressupostos do modelo devem ser satisfeitos.

2. Estimar o valor do CO2 para o ano de 2014 e comentar sobre o resultado.

3. Gerar um **RELATÓRIO** contendo os principais passos da análise de dados, mostrando os principais resultados e comentando esses resultados. Lembre-se que é importante uma boa apresentação do relatório e isso será avaliado.