Trabalho Final

- O trabalho deverá ser feito em **duplas**.
- **Apresentação:** 15-20 minutos por dupla na aula do dia 20/06.
- Avaliação: cada dupla deverá enviar, via Moodle até o dia 20/06, o arquivo com os slides da apresentação, o código e o banco de dados utilizados para fazer as análises.

1 Obtendo a base de dados

- Utilizando a base de dados da Susep (Autoseg), vamos coletar dados agrupados para incidência de sinistros para seguros de automóveis.
- No site da Susep: http://www2.susep.gov.br/menuestatistica/Autoseg/principal.aspx
- Selecione a opção Categoria Tarifária/Região/Modelo/Ano/Sexo/Idade.
- Vamos estabelecer a tarifa para alguns modelos de veículos por meio dos dados de sinistros na Região Sudeste do Brasil.
- Para coletar os dados, selecione as seguintes categorias:
 - Categoria tarifária: selecione "Passeio nacional".
 - Regiões: selecione todos os estados e subregiões da Região Sudeste (ES, MG, RJ e SP).
 - Grupos de modelo: selecione apenas o modelo escolhido pela dupla.
 - Sexo do condutor: selecione Feminino e Masculino.
 - Faixa etária do condutor: selecione as faixas "Entre 18 e 25 anos", "Entre 26 e 35 anos", "Entre 35 e 45 anos", "Entre 46 e 55 anos", "Maior que 55 anos".
 - Selecione o período do ano de 2020.
- Depois que aparecer a tabela com os dados solicitados, clique em "Exportar para Excel" para salvar os dados em formato .xls. Depois, altere o arquivo como preferir e exporte para o R.

2 Analisando os dados

- Para evitar dados faltantes e diminuir algumas das categorias, vamos agregar os dados brutos.
 - Primeiro, agregue os dados das subregiões referentes a cada estado, em cada combinação de covariáveis.
 - Depois, agregue as Frequências e Indenizações para os sinistros de Incêndio e Roubo, Colisão e Outros.
- Faça uma análise exploratória dos dados.
- Ajuste um modelo para a frequência e severidade de sinistros em função das variáveis: sexo, idade do condutor e estado.
- Apresente os resultados do modelo ajustado e suas conclusões.