

Análise de dados 1

Nome: _____

- 1) Considerando os dados [Sao Paulo Real Estate - Sale / Rent - April 2019](#), execute as seguintes tarefas:
 - a) Calcule a frequência absoluta e relativa de cada categoria da variável "Negotiation.Type".
 - b) Faça um gráfico de dispersão do preço do condomínio e do preço anunciado. Como você descreveria esse tipo de relação?
 - c) Faça o gráfico do item anterior considerando facetas de acordo com "Negotiation.Type" (*rent* ou *sale*).
 - d) Crie uma tabela para identificar quais são os dez distritos com maior frequência nesse banco de dados. Utilize essa tabela para criar um gráfico de frequências ordenado de forma decrescente.
 - e) Faça um filtro nos dados considerando apenas os dados de aluguel (Negotiation.Type == "rent"), separe os dados em dois conjuntos (treinamento e teste) e avalie o erro de previsão para os seguintes modelos:
 - i) regressão linear
 - ii) regressão ridge
 - iii) regressão LASSO
 - iv) árvore de decisão
 - v) floresta aleatória

obs: A definição das preditoras utilizadas no modelo é de livre escolha.