

Faculdade Cotemig
Pós-Graduação em Data Science Big Data
Fundamentos Matemáticos para Ciência dos Dados
Lista de Exercícios 3

Professora: Melissa Pinho

Data de entrega: 23/01/2019

Exercício 1

- O arquivo *trabalho.txt* contém as seguintes informações de 753 mulheres:
 - ID_TRABALHA: 0=NÃO TRABALHA; 1=TRABALHA
 - FILHOS_5: QUANTIDADE DE FILHOS COM IDADE MENOR OU IGUAL A 5 ANOS
 - FILHOS_6_18: QUANTIDADE DE FILHOS COM IDADE ENTRE 6 E 18 ANOS
 - IDADE (EM ANOS)
 - ID_ENSINO_SUPERIOR: 1=POSSUI ENSINO SUPERIOR; 2=NÃO POSSUI ENSINO SUPERIOR
 - RENDA_FAMILIA: RENDA FAMILIAR, EXCLUINDO A RENDA DA ESPOSA
1. Construa boxplots representando o número de filhos até cinco anos para as mulheres que trabalham e outro para as que não trabalham. Qual a sua interpretação sobre o gráfico?

2. Repita o item anterior para as variáveis FILHOS_6_18, IDADE e RENDA_FAMILIA. Interprete os gráficos.
 3. Construa uma tabela de referência cruzada entre ID_ENSINO_SUPERIOR e ID_TRABALHA. Interprete os resultados.
 4. Ajuste o modelo de regressão linear para o conjunto de dados, considerando ID_TRABALHA a variável resposta e FILHOS_5, FILHOS_6_18, IDADE e RENDA_FAMILIA as covariáveis (ou variáveis explicativas do modelo).
 5. Quais os valores dos coeficientes?
 6. Quais variáveis explicativas são significativas para explicar a variável resposta ID_TRABALHA?
 7. Calcule e interprete as razões de chances das covariáveis significativas.
 8. Calcule a probabilidade da mulher trabalhar para cada pessoa da amostra.
 9. Qual o perfil da mulher que possui maior probabilidade de trabalhar fora? E o perfil da mulher que possui menor probabilidade de trabalhar fora?
-