

 <small>ESCOLA SUPERIOR DE TECNOLOGIA E GESTÃO</small>	Tipo de Prova Teste 1	Ano letivo 2017/2018	Data 16-11-2017
	Curso Licenciatura em Engenharia Informática	Hora 14:00	
	Unidade Curricular Matemática Computacional II	Duração 1,5 horas	

Observações

Este teste é de consulta, pelo que para a sua realização os alunos podem usar:

- Máquina de calcular.
- Computador sem chat ou contacto com terceiros.
- Apontamentos da UC.

Responda de forma clara e sucinta, indicando todos os cálculos que efetuar.

Submeta todos os ficheiros que utilizar no moodle. O nome de cada ficheiro deve começar pelo seu número de aluno.

Bom trabalho!

Aldina Correia

Elia Costa e Silva

Questão	1	2	3	Total
Cotação	110	60	30	200

Considere a base de dados empregados.sav disponível no moodle. Esta base de dados é constituída por dados de 474 empregados identificados pelo seu id e 10 variáveis:

Nome da variável	Descrição	Valores das var categóricas
<i>id</i>	Código do empregado	
<i>genero</i>	Género	1=m;2=f
<i>datanasc</i>	Data de nascimento	
<i>educ</i>	Nível de educação (anos)	8, 12, 14, 15, 16, 17, 18, 19, 20 e 21 8=Inferior ao ensino secundário 12=Ensino secundário 15=Bacharelato 17=Licenciatura 19=Pós-graduação 20=Mestrado 21=Doutoramento
<i>cat_emp</i>	Categoria de emprego	1=administrativo 2=chefe de seção 3=diretor
<i>ordenado</i>	Ordenado atual	
<i>sal_inic</i>	Salário inicial	
<i>tmp_serv</i>	Meses desde o início do contrato	
<i>exp_ant</i>	Experiência anterior (meses)	
<i>ordenado_euros</i>	Ordenado atual (em euros)	
<i>sal_inic_euros</i>	Salário inicial (em euros)	

- Usando a aproximação dos mínimos quadrados, ajuste os dados da base de dados, considerando a variável resposta (y) o Ordenado actual (em euros) - *ordenado_euros* e a variável independente os Meses desde o início do contrato - *tmp_serv*, usando:
 - uma reta,
 - uma parábola,
 - Um modelo exponencial do tipo $y = b e^{ax}$, sendo $a, b \in \mathbb{R}$,
 - Um modelo do tipo $y = ax^2 + bx^5$.
 - Calcule o erro quadrático em todos eles e diga qual dos modelos é o melhor.
 - Represente graficamente os pontos dados e todo os modelos obtidos.
- Considere um nível de significância de $\alpha=0,05$. Escreva as hipóteses a testar, guarde o script e os outputs necessários num ficheiro para entregar e escreva as conclusões a retirar, em todas as questões que se seguem.
 - Averigue se a variável *sal_inic_euros* (Salário inicial (em euros)) é normalmente distribuída.
 - Teste se o Salário inicial (em euros) dos empregados é superior 255 €.
 - Teste se a média do Salário inicial (em euros) dos empregados difere para pessoas com Nível de educação ao nível do ensino secundário e da licenciatura.
 - Teste se a média do Salário inicial (em euros) dos empregados difere por Nível de educação.
- Usando métodos de classificação:
 - Classifique os 3 clusters de empregados definidos pelas variáveis *tmp_serv*, *exp_ant*, *ordenado_euros* e *sal_inic_euros*.
 - Usando árvores de decisão, averigue se as variáveis *educ* e *cat_emp* permitem distinguir os empregados por género.