**Conceitos Estatísticos para IA – Turma 20IA**

**Trabalho 4**

1. Relacione a Estatística com a Definição do conceito estatístico:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estatísticas |  |  | Definição |
| (a) | Y |  | ( ) | Intercepto do modelo de regressão (nível médio) |
| (b) | α |  | ( ) | Variável predito do modelo |
| (c) | Multiple R2 |  | ( ) | Diferença entre a resposta e o valor estimado pelo modelo (Y-Y estimado) |
| (d) | Adjusted R2 |  | ( ) | Acurácia do modelo de regressão linear simples. |
| (e) | β |  | ( ) | Acurácia do modelo de regressão linear múltipla. |
| (f) | Resíduo padronizado |  | ( ) | Regressão linear simples |
| (h) | Y estimado () |  | ( ) | Regressão linear múltipla |
| (i) | Y=α+βx |  | ( ) | Valor estimado pelo modelo de regressão linear |
| (j) | Y=α+β1x1+β2x2 |  | ( ) | Coeficiente de regressão |
| (k) | Resíduo |  | ( ) | Resíduo é transformado em nova variável com média igual 0 e desvio padrão igual 1 |
| (l) | X |  | ( ) | Variável resposta do modelo |
|  |  |  |  |  |

1. Modelo de regressão linear simples do Valor do imóvel

modelo\_reg01a <- lm(Valor ~ Area)  
summary(modelo\_reg01a)

##   
## Call:  
## lm(formula = Valor ~ Area)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -175589 -37179 -9025 32586 237990   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -17481.05 12968.78 -1.348 0.179   
## Area 1170.46 73.71 15.879 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 63540 on 170 degrees of freedom  
## Multiple R-squared: 0.5973, Adjusted R-squared: 0.5949   
## F-statistic: 252.1 on 1 and 170 DF, p-value: < 2.2e-16

Interprete os resultados

|  |
| --- |
| 1. Escreva o modelo matemático teórico 2. Escreva o modelo matemático ajustado 3. As estimativas dos coeficientes são significativas?      1. Qual o coeficiente de determinação do modelo ajustado? |

1. Calcule a probabilidade do cliente ficar inadimplente no cartão de crédito. Utilize a tabela de coeficientes do modelo ajustado a seguir (vide exemplo no slide 72 da apostila (Apostila\_\_IAML\_FundEstatisticos \_20IAI\_RBernal\_parte3pdf).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Cliente | Atrasos | Valor da fatura | Região | Tempo de relacionamento | Renda | % de gasto com alimentação | Probabilidade de inadimplência |
| 1 | 35 | 1000 | 4 | 6 meses | 1500 | 30 |  |
| 2 | 0 | 2000 | 1 | 8 anos | 10000 | 10 |  |
| 3 | 5 | 250 | 3 | 1 ano | 2000 | 20 |  |

Tabela de coeficientes do modelo

|  |  |  |
| --- | --- | --- |
| **variável** | **categoria** | **Coeficientes** |
| fatura em atraso | até 3 dias | -1,276 |
| 3 a 15 dias | -0,611 |
| de 15 a 30 dias | 0,580 |
| mais de 30 dias | 1,308 |
| Tempo de cliente | até 1 ano | 0,580 |
| de 1 a 3 anos | 0,401 |
| de 3 a 8 anos | -0,264 |
| mais de 8 anos | -0,718 |
| valor da fatura | Até R$250 | 0,262 |
| R$ 250 a R$ 800 | 0,103 |
| R$ 800 a R$ 1.499 | -0,105 |
| Mais de R$1.500 | -0,261 |
| % de gasto com alimentação | até 10% | 0,581 |
| de 10% a 20% | 0,401 |
| de 20% a 30% | -0,264 |
| mais de 30% | -0,718 |
| Região de Risco | Região 4 | 1,067 |
| Região 3 | 0,371 |
| Região 2 | -0,368 |
| Região 1 | -1,069 |
| renda mensal | Até R$ 1.518 | 0,455 |
| R$ 1.519 a R$ 3.000 | 0,080 |
| R$ 3.000 a R$ 4.500 | -0,122 |
| Mais de R$ 4.500 | -0,413 |
| **Constante** |  | 0,099 |

Entrega do exercício no formato word.

Data de entrega: 17/09/2021

Regina Bernal

02/09/2021