

Exercício em sala - Regressão Múltipla - SME0820 Análise de Regressão - 07/05/2014

Os dados da Tabela 1 apresentam características de 48 estados norte-americanos em 1971, com X_1 o imposto sobre combustível (centavos de dólar/galão), X_2 o percentual da população com carteira de habilitação, X_3 a renda per capita (milhares de dólares) e X_4 a quantidade de estradas federais pavimentadas (em milhares de milhas) e Y o consumo de combustível (galões por pessoa). Um galão equivale a 3,785 litros. (Weisberg, S. 1985. Applied Linear Regression, Wiley).

TABELA 1. Dados de consumo de combustível em estados norte-americanos.

Estado	X_1	X_2	X_3	X_4	Y	estado	X_1	X_2	X_3	X_4	Y
ME	9	52,5	3,571	1,976	541	WV	8,5	55,1	4,574	2,619	460
NH	9	57,2	4,092	1,25	524	NC	9	54,4	3,721	4,746	566
VT	9	58	3,865	1,586	561	SC	8	54,8	3,448	5,399	577
MA	7,5	52,9	4,87	2,351	414	GA	7,5	57,9	3,846	9,061	631
RI	8	54,4	4,399	0,431	410	FA	8	56,3	4,188	5,975	574
CN	10	57,1	5,342	1,333	457	KY	9	49,3	3,601	4,65	534
NY	8	45,1	5,319	11,868	344	TN	7	51,8	3,64	6,905	571
NJ	8	55,3	5,126	2,138	467	AL	7	51,3	3,333	6,594	554
PA	8	52,9	4,447	8,577	464	MS	8	57,8	3,063	6,524	577
OH	7	55,2	4,512	8,507	498	AR	7,5	54,7	3,357	4,121	628
IN	8	53	4,391	5,939	580	LA	8	48,7	3,528	3,495	487
IL	7,5	52,5	5,126	14,186	471	OK	6,58	62,9	3,802	7,834	644
MI	7	57,4	4,817	6,93	525	TX	5	56,6	4,045	17,782	640
WI	7	54,5	4,207	6,58	508	MT	7	58,6	3,897	6,385	704
MN	7	60,8	4,332	8,159	566	ID	8,5	66,3	3,635	3,274	648
IA	7	58,6	4,318	10,34	635	WY	7	67,2	4,345	3,905	968
MO	7	57,2	4,206	8,508	603	CO	7	62,6	4,449	4,639	587
ND	7	54	3,718	4,725	714	NM	7	56,3	3,656	3,985	699
SD	7	72,4	4,716	5,915	865	AZ	7	60,3	4,3	3,635	632
NE	8,5	67,7	4,341	6,01	640	UT	7	50,8	3,745	2,611	591
KS	7	66,3	4,593	7,834	649	NV	6	67,2	5,215	2,302	782
DE	8	60,2	4,983	0,602	540	WN	9	57,1	4,476	3,942	510
MD	9	51,1	4,897	2,449	464	OR	7	62,3	4,296	4,083	610
VA	9	51,7	4,258	4,686	547	CA	7	59,3	5,002	9,794	524

Ajuste um modelo de regressão linear múltipla para o consumo de combustível do tipo

$$Y = X\beta + \epsilon$$

com as suposições usuais, considerando X_1 , X_2 , X_3 e X_4 como variáveis explicativas. Construa a tabela ANOVA (corrigida pela média) e interprete os resultados ao nível de significância $\alpha = 0,05$. Obtenha o coeficiente de determinação do modelo.