

ACTIVIDADES DEL TEMA 3: REGRESIÓN LINEAL

1. Uno de los problemas con los que se encuentran las industrias químicas es el tratamiento de las aguas residuales. Dichas aguas son químicamente complejas puesto que se caracterizan por altos valores de gases tóxicos, sólidos volátiles y otras sustancias nocivas. Los datos siguientes se obtuvieron a partir de 33 muestras de aguas residuales tratadas en el Instituto Politécnico de Virginia. En dichos datos se muestran los porcentajes de reducción de gases tóxicos (Y) durante el proceso de depuración para distintos valores de reducción en los sólidos volátiles (X). El objetivo es determinar un modelo que permita predecir la disminución porcentual en la concentración de gases tóxicos conocida la disminución de sólidos volátiles durante la depuración de las aguas residuales.

X	Y	X	Y	X	Y	X	Y
3	5	31	30	37	36	42	44
7	11	31	40	33	38	43	47
11	21	32	32	39	37	44	44
15	16	33	34	39	36	45	46
18	16	33	32	39	45	46	46
27	28	34	34	40	39	47	49
29	27	36	37	41	41	50	51
30	25	36	38	42	40		
30	36	36	34				

Realizar un análisis de regresión simple sobre los datos, incluyendo el contraste de linealidad. Estudiar posibles datos atípicos y si afectan al ajuste realizado. ¿Afectan a alguna de las hipótesis del modelo? ¿Es admisible que la variabilidad de los residuos aumente o disminuya con la propia variable explicativa? Justica estas cuestiones.

2. Una industria automovilística desea conocer el promedio de vehículos por persona (Y) en una serie de países, en función de su densidad de población (X1), Renta per Cápita (X2), Precio del Litro de Gasolina (X3), Toneladas de Gasolina Consumida (X4) y Promedio de Kilómetros de Transporte Público Usados por Persona (X5). Se han obtenidos los siguientes datos:

Y	X_1	X_2	X_3	X_4	X_5
0.27	89	7.7	49	1.1	2.6
0.33	323	9.8	59	1	1.6
0.42	2	8.7	17	2.8	0.1
0.28	119	11	56	1.2	1.9
0.24	16	7.1	49	1.2	2.2
0.33	97	8.8	61	1	1.5
0.35	247	10.4	49	1.1	1.7
0.08	71	3.4	56	1.7	0.7
0.34	2	9.8	57	1.2	2
0.20	46	3.8	40	1.5	0.3
0.30	188	4.6	61	0.6	1.8
0.18	1309	8.5	49	1.2	3.5
0.43	138	9.8	44	1.6	0.8
0.30	412	9.4	56	1	1.5
0.40	12	5.9	34	1.3	0.2
0.28	13	9.8	61	1	1.7
0.10	107	1.8	68	0.7	0.9
0.18	73	4	44	0.8	1.3
0.34	18	10.6	42	1.3	1.7
0.32	153	13.3	56	1.3	2
0.014	55	1.2	36	3.3	0.1
0.27	229	5.5	35	1.2	1.6
0.53	23	9.7	17	2.7	0.3
0.09	86	2.1	40	1.1	2.1

Con los datos obtenidos estudia si existe un patrón permita predecir el promedio de vehículos por habitante en un país.