Tarea 1

1.Resolver los problemas 2.1, 2.7, 2.18 2.19 y 2.21 del libro de Montgomery. Utiliza R. Escribe tus respuestas y adjunta el código del programa en un archivo Rmarkdown.

2.Considera el modelo de regresión $Y_i = \beta X + \varepsilon_i, i = 1,2, \ldots n$ Obtenga los estimados de mínimos cuadrados $\hat{\beta}$ de los parámetros desconocidos β.

2.1 Table B.1 gives data concerning the performance of the 26 NFL teams in 1976. It is suspected that the number of yards gained rushing by opponents (x8) has an effect on the number of games won by a team (y).

y	x1	x2	x3	x4	x5	x6	x7	x8	x9
10	2113	1985	38.9	64.7	4	868	59.7	2205	1917
11	2003	2855	38.8	61.3	3	615	55	2096	1575
11	2957	1737	40.1	60	14	914	65.6	1847	2175
13	2285	2905	41.6	45.3	-4	957	61.4	1903	2476
10	2971	1666	39.2	53.8	15	836	66.1	1457	1866
11	2309	2927	39.7	74.1	8	786	61	1848	2339
10	2528	2341	38.1	65.4	12	754	66.1	1564	2092
11	2147	2737	37	78.3	-1	761	58	1821	1909
4	1689	1414	42.1	47.6	-3	714	57	2577	2001
2	2566	1838	42.3	54.2	-1	797	58.9	2476	2254
7	2363	1480	37.3	48	19	984	67.5	1984	2217
10	2109	2191	39.5	51.9	6	700	57.2	1917	1758
9	2295	2229	37.4	53.6	-5	1037	58.8	1761	2032
9	1932	2204	35.1	71.4	3	986	58.6	1709	2025
6	2213	2140	38.8	58.3	6	819	59.2	1901	1686
5	1722	1730	36.6	52.6	-19	791	54.4	2288	1835
5	1498	2072	35.3	59.3	-5	776	49.6	2072	1914
5	1873	2929	41.1	55.3	10	789	54.3	2861	2496
6	2118	2268	38.2	69.6	6	582	58.7	2411	2670
4	1775	1983	39.3	78.3	7	901	51.7	2289	2202
3	1904	1792	39.7	38.1	-9	734	61.9	2203	1988
3	1929	1606	39.7	68.8	-21	627	52.7	2592	2324
4	2080	1492	35.5	68.8	-8	722	57.8	2053	2550
10	2301	2835	35.3	74.1	2	683	59.7	1979	2110
6	2040	2416	38.7	50	0	576	54.9	2048	2628
8	2447	1638	39.9	57.1	-8	848	65.3	1786	1776
2	1416	2649	37.4	56.3	-22	684	43.8	2876	2524
0	1503	1503	39.3	47	-9	875	53.5	2560	2241

a. Fit a simple linear regression model relating games won Y to yards gained rushing by opponents X8.

b. Construct the analysis of variance table and test for significance of regression.

c. Find a 95% CI on the slope

d. What percent of the total variability in y is explained by this model?

e. Find a 95% confidence interval on the number of games won if an opponent' yards rushing is limited to 2000 yards

2.7. The purity of oxygen produced by a fractional distillation process is thought to be related to the percentage of hydrocarbons in the main condenser of the processing unit. Twenty samples are shown below:

purity	hydro
86.91	1.02
89.85	1.11
90.28	1.43
86.34	1.11
92.58	1.01
87.33	0.95
86.29	1.11
91.86	0.87
95.61	1.43
89.86	1.02
96.73	1.46
99.42	1.55
98.66	1.55
96.07	1.55
93.65	1.4
87.31	1.15
95	1.01
96.85	0.99
85.2	0.95
90.56	0.98

a. Fit a simple linear regression model to the data.
b. Test the hypothesis H_0: $\beta_0 = 0$.
c. Calculate R^2.
d. Find 95% CI on the slope.
e. Find a 95% CI on the mean purity when the hydrocarbon percentage is 1.00

2.18. On March on 1984 the Wall Street Journal published a survey of television advertisements conducted by Video Board Tests, Inc. a New York ad-testing company that interview 4000 adults. This people were regular product users who were asked to cite a commercial they had seen for the product category in the past week. In this case, the response is the number of million of retained impression per week. The regressor is the amount of money spent by the firm on advertisement. The data follow:

Firm	Amount Spent (Millions)	Returned Impressions per week (millions)
Miller Lite	50.1	32.1
Pepsi	74.1	99.6
Stroh's	19.3	11.7
Federal Express	22.9	21.9
Burger King	82.4	60.8
Coca-Cola	40.1	78.6
McDonald's	185.9	92.4
MCI	26.9	50.7
Diet Cola	20.4	21.4
Ford	166.2	40.1
Levi's	27	40.8
Bud Lite	45.6	10.4
ATT Bell	154.9	88.9
Calvin Klein	5	12
Wendy's	49.7	29.2
Polaroid	26.9	38
Shasta	5.7	10
Meow Mix	7.6	12.3
Oscar Meyer	9.2	23.4
Crest	32.4	71.1
Kibbles N Bits	6.1	4.4

a. Fit the simple linear regression model to these data
b. Is there a significant relationship between the amount a company spend on advertising and retained impressions? Justify your answer statistically.
c. Construct the 95% confidence and predictions bands for these data.
d. Give the 95% confidence and predictions for the number of retained impressions for MCI.

2.19. Table B.17 contains the Patient Satisfaction data used in section 2.7.

Table B.17

Satisfaction	Age	Severity	Surgical-Medical	Anxiety
68	55	50	0	2.1
77	46	24	1	2.8
96	30	46	1	3.3
80	35	48	1	4.5
43	59	58	0	2
44	61	60	0	5.1
26	74	65	1	5.5
88	38	42	1	3.2
75	27	42	0	3.1
57	51	50	1	2.4
56	53	38	1	2.2
88	41	30	0	2.1
88	37	31	0	1.9
102	24	34	0	3.1
88	42	30	0	3
70	50	48	1	4.2
82	58	61	1	4.6
43	60	71	1	5.3
46	62	62	0	7.2
56	68	38	0	7.8
59	70	41	1	7
26	79	66	1	6.2
52	63	31	1	4.1
83	39	42	0	3.5
75	49	40	1	2.1

a. Fit a simple linear regression model relating satisfaction and age.
b. Compare this model to the fit in section 2.7 relating patient satisfaction to severity.

2.21 Consider the wine quality of young red wine data Table B.19 the wine-markets believes that the sulfur contents has a negative impact on the taste (thus, the overall quality) of the wine. Perform a thorough analysis of these data. Do data support the winemakers' belief?

Table B.19

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_10
19.2	0	3.85	66	9.35	5.65	2.4	3.25	0.33	19	0.065
18.3	0	3.73	79	11.15	6.95	3.15	3.8	0.36	21	0.076
17.1	0	3.88	73	9.4	5.75	2.1	3.65	0.4	18	0.073
17.3	0	3.86	99	12.85	7.7	3.9	3.8	0.35	22	0.076
16.8	0	3.98	75	8.55	5.05	2.05	3	0.49	12	0.06
16.5	0	3.85	61	10.3	6.2	2.5	3.7	0.38	20	0.074
15.8	0	3.93	66	4.9	2.75	1.2	1.55	0.29	11	0.031
15.2	0	3.66	86	6.4	4	1.5	2.5	0.27	19	0.05
15.2	0	3.91	78	5.8	3.3	1.4	1.9	0.4	9	0.038
14	0	3.47	178	3.6	2.25	0.75	1.5	0.37	8	0.03
14	0	3.91	81	3.9	2.15	1	1.15	0.32	7	0.023
13.8	0	3.75	108	5.8	3.2	1.6	1.6	0.38	8	0.032
13.6	0	3.9	92	5.4	2.85	1.55	1.3	0.44	6	0.026
12.8	0	3.92	96	5	2.7	1.4	1.3	0.35	7	0.026
18.5	1	3.87	89	9.15	5.6	1.95	3.65	0.46	16	0.073
17.3	1	3.97	59	10.25	6.1	2.4	3.7	0.4	19	0.074
16.3	1	3.76	22	8.2	5	1.85	3.15	0.25	25	0.063
16.3	1	3.76	77	8.35	5.05	1.9	3.15	0.37	17	0.063
16	1	3.98	58	10.15	6	2.6	3.4	0.38	18	0.068
16	1	3.88	85	6.85	4.1	1.5	2.6	0.33	16	0.052
15.7	1	3.75	120	8.8	5.5	1.85	3.65	0.39	19	0.073
15.5	1	3.98	94	5.45	3.05	1.5	1.55	0.41	8	0.031
15.3	1	3.69	122	8	5.05	1.9	3.15	0.27	23	0.063
15.3	1	3.77	144	5.6	3.35	1.1	2.25	0.36	12	0.045
14.8	1	3.74	10	7.9	4.75	1.95	2.8	0.25	23	0.056
14.3	1	3.76	100	5.55	3.25	1.15	2.1	0.34	12	0.042
14.3	1	3.91	73	4.65	2.7	0.95	1.75	0.36	10	0.035
14.2	1	3.6	301	4.25	2.4	1.25	1.15	0.42	6	0.023
14	1	3.76	104	8.7	5.1	2.25	2.85	0.34	17	0.057
13.8	1	3.9	67	7.4	4.4	1.6	2.8	0.45	13	0.056
12.5	1	3.8	89	5.35	3.15	1.2	1.95	0.32	12	0.039
11.5	1	3.65	192	6.35	3.9	1.25	2.65	0.63	8	0.053

y: quality rating (20 maximum)
x_1: wine varietal (0 - Cabernet Sauvignon, 1 - Shiraz)
x_2: pH
x_3: Total SO_2(ppm)
x_4: color density
x_5: wine color
x_6: polymeric pigment color
x_7: anthocyanin color
x_8: total anthocyanins (g/L)
x_9: degree of ionization of anthocyanins (percent)
x_10: ionized anthocyanins (percent)