DADOS PARA ANÁLISE

1. Bread wrapper data

Consider an experiment in which the researcher wants determine an empirical relationship between the seal strength (Y) in grams per inch of a bread wrapper stock and regressor variables: scaling temperature (X_1), cooling bar temperature (X_2), and percent polyethylene in the stock (X_3). The combination of the regressor variables were in advance by the scientist.

Table 1: Bread wrapper stock data

Obs	Y (g/in)	X_1 (^{o}F)	X_2 (^{o}F)	X_3 (weight in %)
1	6.6	225	46	0.5
2	6.9	285	46	0.5
3	7.9	225	64	0.5
4	6.1	285	64	0.5
5	9.2	225	46	1.7
6	6.8	285	46	1.7
7	10.4	225	64	1.7
8	7.3	285	64	1.7
9	9.8	204.5	55	1.1
10	5.0	305.5	55	1.1
11	6.9	255	39.9	1.1
12	6.3	255	70.1	1.1
13	4.0	255	55	0.09
14	8.6	255	55	2.11
15	10.1	255	55	1.1
16	9.9	255	55	1.1
17	12.2	255	55	1.1
18	9.7	255	55	1.1
19	9.7	255	55	1.1
20	9.6	255	55	1.1

2. Squid data

An experiment was conducted in order to study the size of squid eaten by sharks and tuna. The regressor variables are characteristics of the beak or mouth of the squid. The regressor variables and response considered for the study are

X_1 = rostral length in inches
X_2 = wing length in inches
X_3 = rostral to notch length
X_4 = notch to wing length
X_5 = width in inches
Y = weight in pounds

Table 2: Squid weight and beak measurements

Obs.	Y	X_1	X_2	X_3	X_4	X_5
1	1.95	1.31	1.07	0.44	0.75	0.35
2	2.90	1.55	1.49	0.53	0.90	0.47
3	0.72	0.99	0.84	0.34	0.57	0.32
4	0.81	0.99	0.83	0.34	0.54	0.27
5	1.09	1.05	0.90	0.36	0.64	0.30
6	1.22	1.09	0.93	0.42	0.61	0.31
7	1.02	1.08	0.90	0.40	0.51	0.31
8	1.93	1.27	1.08	0.44	0.77	0.34
9	0.64	0.99	0.85	0.36	0.56	0.29
10	2.08	1.34	1.13	0.45	0.77	0.37
11	1.98	1.30	1.10	0.45	0.76	0.38
12	1.90	1.33	1.10	0.48	0.77	0.38
13	8.56	1.86	1.47	0.60	1.01	0.65
14	4.49	1.58	1.34	0.52	0.95	0.50
15	8.49	1.97	1.59	0.67	1.20	0.59
16	6.17	1.80	1.56	0.66	1.02	0.59
17	7.54	1.75	1.58	0.63	1.09	0.59
18	6.36	1.72	1.43	0.64	1.02	0.63
19	7.63	1.68	1.57	0.72	0.96	0.68
20	7.78	1.75	1.59	0.68	1.08	0.62
21	10.15	2.19	1.86	0.75	1.24	0.72
22	6.88	1.73	1.67	0.64	1.14	0.55

3. Hospital data

Data reflects information taken from seventeen U.S. Naval hospitals at various sites around the world. The regressors are workload variables, i.e., items that result in the need for manpower in a hospital installation. A brief description of the variable are as follows:

Y = monthly man-hours
X_1 = average daily patient load
X_2 = monthly X-ray exposures
X_3 = monthly occupied bed days
X_4 = eligible population in the area /100
X_5 = average length of patient' stay in days

The goal here is to produce an equation that will estimate (predict) manpower needs for Naval hospitals.

Table 3: Hospital manpower data

Site	Y	X_1	X_2	X_3	X_4	X_5
1	566.52	15.57	2463	472.92	18	4.45
2	696.82	44.02	2048	1339.75	9.5	6.92
3	1033.15	20.42	3940	620.25	12.8	4.28
4	1603.62	18.74	6505	568.33	36.7	3.90
5	1611.37	49.20	5723	1497.60	35.7	5.50
6	1613.27	44.92	11520	1365.83	24	4.60
7	1854.17	55.48	5779	1687.00	43.3	5.62
8	2160.55	59.28	5969	1639.92	46.7	5.15
9	2305.58	94.39	8461	2872.33	78.7	6.18
10	3503.93	128.02	20106	3655.08	180.5	6.15
11	3571.89	96.00	13313	2912.00	60.9	5.88
12	3742.40	131.42	10771	3921.00	103.7	4.88
13	4026.52	127.21	15543	3865.67	126.8	5.50
14	10343.81	252.90	36194	7684.10	157.7	7.00
15	11732.17	409.20	34703	12446.33	169.4	10.78
16	15414.94	463.70	39203	14098.40	331.4	7.05
17	18854.45	510.22	86533	15524.00	371.6	6.35

4. Teacher effectiveness data

Twenty-three student teachers took part in an evaluation program designed to measure teacher effectiveness and determine what factors are important. Twelve response for each was a quantitative evaluation made on the cooperating teacher. The regressor variables were scores on four standardized tests given to each instructor.

Table 4 : teacher effectiveness data

Obs.	Y	X_1	X_2	X_3	X_4	X_5
1	489	81	151	45.5	43.61	Male
2	423	68	156	46.45	44.69	Male
3	507	80	165	76.5	54.57	Male
4	467	107	149	55.5	43.27	Male
5	340	43	134	49.4	49.21	Male
6	524	129	163	72.0	49.96	Male
7	488	139	159	86.2	53.05	Male
8	445	88	135	64.0	49.51	Male
9	388	99	141	44.15	39.57	Male
10	579	121	145	44.25	51.89	Male
11	433	91	129	42.5	53.77	Male
12	409	87	115	79.25	56.32	Male
13	410	69	125	59.0	55.66	Female
14	568	57	131	31.75	63.97	Female
15	425	77	141	80.5	45.32	Female
16	344	81	122	75.0	46.67	Female
17	324	0	141	49.0	41.21	Female
18	505	53	152	49.35	43.83	Female
19	234	77	141	60.75	41.61	Female
20	501	76	132	41.25	64.57	Female
21	400	65	157	50.75	42.41	Female
22	584	97	166	32.25	57.95	Female
23	434	76	141	54.52	57.90	Female

5. Executive compensation data

In an effort to model executive compensation for the year 1979, 33 firms were selected, and data was gathered on compensation, sales, profits, and employment. The following data was gathered for the year 1979.

Table 5. Executive compensation data

Firm	Y Compensation (thousands of dollars)	X_1 Sales (millions of dollars)	X_2 Profits (millions of dollars)	X_3 Employment
1	450	4600.6	128.1	48000
2	387	9255.4	783.9	55900
3	368	1526.2	136.0	13783
4	277	1683.2	179.0	27765
5	676	2752.8	231.5	34000
6	454	2205.8	329.5	26500
7	507	2384.6	381.8	30800
8	496	2746.0	237.9	41000
9	487	1434.0	222.3	25900
10	383	470.6	63.7	8600
11	311	1508.0	149.5	21075
12	271	464.4	30.0	6874
13	524	9329.3	577.3	39000
14	498	2377.5	250.7	34300
15	343	1174.5	82.6	19405
16	354	409.3	61.5	3586
17	324	724.7	90.8	3905
18	225	578.9	63.3	4139
19	254	966.8	42.8	6255
20	208	591.0	48.5	10605
21	518	4933.1	310.6	65392
22	406	7613.2	491.6	89400
23	332	3457.4	228.0	55200
24	340	545.3	54.6	7800
25	698	22862.8	3011.3	337119
26	306	2361.0	203.0	52000
27	613	2614.1	201.0	50500
28	302	1013.2	121.3	18625
29	540	4560.3	194.6	97937
30	293	855.7	63.4	12300
31	528	4211.6	342.1	71800
32	456	5440.4	655.2	87700
33	417	1229.9	97.5	14600

6. Sales data

A maker of asphalt roofing shingles is interested in the relationship between sales for a particular year and factors that obviously influence sales: promotional accounts, number of active accounts, numbers of competing brands, and district potential for the sales districts. Fifteen districts were used, and it is of interested to predict sales from resulting regression.

Table 6. Sales data

District	Y Sales in thousands of dollars	X_1 Promotional	X_2 Active accounts	X_3 Competing brands	X_4 Potential
1	79.3	5.5	31	10	8
2	200.1	2.5	55	8	6
3	163.2	8.0	67	12	9
4	200.1	3.0	50	7	16
5	146.0	3.0	38	8	15
6	177.7	2.9	71	12	17
7	30.9	8.0	30	12	8
8	291.9	9.0	56	5	10
9	160.0	4.0	42	8	4
10	339.4	6.5	73	5	16
11	159.6	5.5	60	11	7
12	86.3	5.0	44	12	12
13	237.5	6.0	50	6	6
14	107.2	5.0	39	10	4
15	155.0	3.5	55	10	4

7. Fitness data

In exercise physiology, an objective measure of aerobic fitness is the oxygen consumption in volume per unit body weight per unit time by individual. To determine if it is feasible to predict this fitness measure, an experiment was conducted in which 31 individuals were tested. The following factors were studied:

X_1: age in years
X_2: weight in kilograms
X_3: time to run 1 ½ miles
X_4: resting pulse rate
X_5: pulse rate at the end of run
X_6: maximum pulse rate during run
Y: oxygen consumption in millilitres (ml) per kilogram (kg) body weight per minute

Obs.	Y	X_1	X_2	X_3	X_4	X_5	X_6
1	44.609	44	89.47	11.37	62	178	182
2	45.313	40	75.07	10.07	62	185	185
3	54.297	44	85.84	8.65	45	156	168
4	59.571	42	68.15	8.17	40	166	172
5	49.874	38	89.02	9.22	55	178	180
6	44.811	47	77.45	11.63	58	176	176
7	45.681	40	75.98	11.95	70	176	180
8	49.901	43	81.19	10.85	64	162	170
9	39.442	44	81.42	13.08	63	174	176
10	60.055	38	81.87	8.63	48	170	186
11	50.541	44	73.03	10.13	45	168	168
12	37.388	45	87.66	14.03	56	186	192
13	44.754	45	66.45	11.12	51	176	176
14	47.273	47	79.15	10.60	47	162	164
15	51.855	54	83.12	10.33	50	166	170
16	49.156	49	81.42	8.95	44	180	185
17	40.836	51	69.63	10.95	57	168	172
18	46.672	51	77.91	10.00	48	162	168
19	46.774	48	91.63	10.25	48	162	164
20	50.388	49	73.37	10.08	76	168	168
21	39.407	57	73.37	12.63	58	174	176
22	46.080	54	79.39	11.17	62	156	165
23	45.441	53	76.32	9.63	48	164	166
24	54.625	50	70.87	8.92	48	146	155
25	45.118	51	67.25	11.08	48	172	172
26	39.203	54	91.63	12.88	44	168	172
27	45.790	51	73.71	10.47	59	186	188
28	54.545	57	59.08	9.93	49	148	155
29	48.673	49	76.32	9.40	56	186	188
30	47.920	48	61.24	11.50	52	170	176
31	47.467	52	82.78	10.50	53	170	172

8. Bachelor Officers Quarters (BOQ) data

The U.S. Navy attempts to develop equations for estimation of manpower needs for manning installations such as Bachelor Officers Quarters. Regression equations are developed from data taken by measurement teams. The data in Table 8 was collected at 25 BOQ sites. The variable descriptions are as follows:

X_1: average daily occupancy
X_2: monthly average number of check-ins
X_3: weekly hours of service desk operation
X_4: square feet of common use area
X_5: number of building wings
X_6: operational berthing capacity
X_7: number of rooms
Y: monthly man-hours

Table 8. BOQ data

Site	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	180.23	2.0	4.0	4	1.26	1	6	6
2	182.61	3.0	1.58	40	1.25	1	5	5
3	164.38	16.6	23.78	40	1.00	1	13	13
4	284.55	7.0	2.37	168	1.00	1	7	8
5	199.92	5.3	1.67	42.5	7.79	3	25	25
6	267.38	16.5	8.25	168	1.12	2	19	19
7	999.09	25.89	3.0	40	0	3	36	36
8	1103.24	44.42	159.75	168	0.60	18	48	48
9	944.21	39.63	50.86	40	27.37	10	77	77
10	931.84	31.92	40.08	168	5.52	6	47	47
11	2268.06	97.33	255.08	168	19.0	6	165	130
12	1489.5	56.63	373.42	168	6.03	4	36	37
13	1891.7	96.67	206.67	168	17.86	14	120	120
14	1387.82	54.58	207.08	168	7.77	6	66	66
15	3559.92	113.88	981.0	168	24.48	6	166	179
16	3115.29	149.58	233.83	168	31.07	14	185	202
17	2227.76	134.32	145.82	168	25.99	12	192	192
18	4804.24	188.74	937.0	168	45.44	26	237	237
19	2628.32	110.24	410.0	168	20.05	12	115	115
20	1880.84	96.83	677.33	168	20.31	10	302	210
21	3036.63	102.33	288.83	168	21.01	14	131	131
22	5539.98	274.92	695.25	168	46.63	58	363	363
23	3534.49	811.08	714.33	168	22.76	17	242	242
24	8266.77	384.5	1473.66	168	7.36	24	540	453
25	1845.89	95.0	368.0	168	30.26	9	292	196

9. Lipoprotein data

An experiment was conducted regarding a quantitative analysis of factors found in high-density lipoprotein (HDL) in a sample of human blood serum. Three variables though to be predictive or associate with HDL measurement (Y) were the total cholesterol (X1) and total triglyceride (X2) concentration in the sample, plus the presence or absence of a certain sticky component found in the serum called sinking pre-beta, or SPB (X3), coded as 0 if absent and 1 if present. The data obtained are shown in Table 9.

Table 9. Lipoprotein data

Y	X_1	X_2	X_3	Y	X_1	X_2	X_3
47	287	111	0	57	192	115	1
38	236	135	0	42	349	408	1
47	255	98	0	54	263	103	1
39	135	63	0	60	223	102	1
44	121	46	0	33	316	274	0
64	171	103	0	55	288	130	0
58	260	227	0	36	256	149	0
49	237	157	0	36	318	180	0
55	261	266	0	42	270	134	0
52	397	167	0	41	262	154	0
49	295	164	0	42	264	86	0
47	261	119	1	39	325	148	0
40	258	145	1	27	388	191	0
42	280	247	1	31	260	123	0
63	339	168	1	39	284	135	0
40	161	68	1	56	326	236	1
59	324	92	1	40	248	92	1
56	171	56	1	58	285	153	1
76	265	240	1	43	361	126	1
67	280	306	1	40	248	226	1
57	248	93	1	46	280	176	1

10. Black cherry data

Os dados que se seguem referem-se a medidas de diâmetro à altura do peito (X_1) e altura (X_2) de árvores (*black cherry*) em pé e de volume (Y) de árvores derrubadas. O objetivo desse tipo de experimento é verificar de que forma essas variáveis estão relacionadas para, através de medidas nas árvores em pé, poder se predizer o volume de madeira em uma área de floresta.

Obs	X1	X2	Y
1	8.3	70	10.3
2	8.6	65	10.3
3	8.8	63	10.2
4	10.5	72	16.4
5	10.7	81	18.8
6	10.8	83	19.7
7	11.0	66	15.6
8	11.0	75	18.2
9	11.1	80	22.6
10	11.2	75	19.9
11	11.3	79	24.2
12	11.4	76	21.0
13	11.4	76	21.4
14	11.7	69	21.3
15	12.0	75	19.1
16	12.9	74	22.2
17	12.9	85	33.8
18	13.3	86	27.4
19	13.7	71	25.7
20	13.8	64	24.9
21	14.0	78	34.5
22	14.2	80	31.7
23	14.5	74	36.3
24	16.0	72	38.3
25	16.3	77	42.6
26	17.3	81	55.4
27	17.5	82	55.7
28	17.9	80	58.3
29	18.0	80	51.5
30	18.0	80	51.0
31	20.6	87	77.0

Fonte: RYAN, B.F.; JOINER, B.L.; RYAN Jr., T.A. (1976).
Minitab Student Handbook. Duxbury Press. pg. 329.

11. Dados de solo

Os dados que se seguem referem-se a um estudo da relação existente entre duas fontes de fósforo no solo e o conteúdo de fósforo no solo. Foram feitas medidas de concentrações de fósforo inorgânico (X_1) e fósforo orgânico (X_2) no solo e de conteúdo de fósforo (Y) nas plantas crescidas naquele solo.

Obs	X1	X2	Y
1	0.4	53	64
2	0.4	23	60
3	3.1	19	71
4	0.6	34	61
5	4.7	24	54
6	1.7	65	77
7	9.4	44	81
8	10.1	31	93
9	11.6	29	93
10	12.6	58	51
11	10.9	37	76
12	23.1	46	96
13	23.1	50	77
14	21.6	44	93
15	23.1	56	95
16	1.9	36	54
17	26.8	58	168
18	29.9	51	99

Fonte: SNEDECOR, G.W & COCHRAN, W.G. (1967). Statistical Methods. The Iowa State Press University. pg. 384.