

- تطبيق (1-3) :-

تهتم وحدة جراحة في مستشفى بالتنبؤ بنسبة الشفاء لمرضي يخضعون لنوع معين من جراحة الكبد. وقد توافر للتحليل 54 مريضاً اختيروا عشوائياً. ومن سجل كل مريض استخلصت المعلومات التالية من التقييم الذي يسبق العملية ، X_1 درجة تخثر الدم ، X_2 دليل الإنذار القياسي بما في ذلك عمر المريض ، X_3 درجة اختبار وظيفة الأنزيم ، X_4 درجة اختبار وظيفة الكبد ، والمتغير التابع Y هو الفترة التي يعيشها المريض بعد الجراحة.

i	Y	X1	X2	X3	X4
1	200	6.7	62	81	2.59
2	101	5.1	59	66	1.7
3	204	7.4	57	83	2.16
4	101	6.5	73	41	2.01
5	509	7.8	65	115	4.3
6	80	5.8	38	72	1.42
7	80	5.7	46	63	1.91
8	127	3.7	68	81	2.57
9	202	6	67	93	2.5
10	203	3.7	76	94	2.4
11	329	6.3	84	83	4.13
12	65	6.7	51	43	1.86
13	830	5.8	96	114	3.95
14	330	5.8	83	88	3.95
15	168	7.7	62	67	3.4
16	217	7.4	74	68	2.4
17	87	6	85	28	2.98
18	34	3.7	51	41	1.55
19	215	7.3	68	74	3.56
20	172	5.6	57	87	3.02
21	109	5.2	52	76	2.85
22	136	3.4	83	53	1.12
23	70	6.7	26	68	2.1
24	220	5.8	67	86	3.4
25	276	6.3	59	100	2.95
26	144	5.8	61	73	3.5
27	181	5.2	52	86	2.45
28	574	11.2	76	90	5.59
29	72	5.2	54	56	2.71

i	Y	X1	X2	X3	X4
30	178	5.8	76	59	2.58
31	71	3.2	64	65	0.74
32	58	8.7	45	23	2.52
33	116	5	59	73	3.5
34	295	5.8	72	93	3.3
35	115	5.4	58	70	2.64
36	184	5.3	51	99	2.6
37	118	2.6	74	86	2.05
38	120	4.3	8	119	2.85
39	151	4.8	61	76	2.45
40	148	5.4	52	88	1.81
41	95	5.2	49	72	1.84
42	75	3.6	28	99	1.3
43	483	8.8	86	88	6.4
44	153	6.5	56	77	2.85
45	191	3.4	77	93	1.48
46	123	6.5	40	84	3
47	311	4.5	73	106	3.05
48	398	4.8	86	101	4.1
49	158	5.1	67	77	2.86
50	310	3.9	82	103	4.55
51	124	6.6	77	46	1.95
52	125	6.4	85	40	1.21
53	198	6.4	59	85	2.33
54	313	8.8	78	72	3.2

- التحليل الإحصائي :-

باستخدام برنامج الحزم الإحصائية Minitab سوف نقوم باستخراج قيم العزم لمصفوفة القبة واكتشاف المشاهدات القاصية في المتغيرات المستقلة درجة تخثر الدم (X1)، دليل الإنذار القياسي (X2)، درجة اختبار وظيفة الأنزيم (X2)، درجة اختبار وظيفة الكبد (X4).

$$H_0: B_i = 0 \quad vs \quad H_1: B_i \neq 0 \quad i = 0,1,2,3,4$$

Regression Analysis: Y versus X1; X2; X3; X4

The regression equation is

$$Y = - 622 + 33.2 X1 + 4.27 X2 + 4.13 X3 + 14.1 X4$$

Predictor	Coef	SE Coef	T	P
Constant	-621.60	64.80	-9.59	0.000
X1	33.164	7.017	4.73	0.000
X2	4.2719	0.5634	7.58	0.000
X3	4.1257	0.5112	8.07	0.000
X4	14.09	12.53	1.13	0.266

S = 61.0565 R-Sq = 83.7% R-Sq(adj) = 82.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	936265	234066	62.79	0.000
Residual Error	49	182667	3728		
Total	53	1118931			

Source	DF	Seq SS
X1	1	155274
X2	1	305511
X3	1	470760
X4	1	4719

Unusual Observations

Obs	X1	Y	Fit	SE Fit	Residual	St Resid
13	5.8	830.00	506.85	24.21	323.15	5.77R
18	3.7	34.00	-90.03	23.69	124.03	2.20R
38	4.3	120.00	86.31	34.17	33.69	0.67 X

R denotes an observation with a large standardized residual.

X denotes an observation whose X value gives it large influence.

نلاحظ أن المتغيرات المستقلة جميعها معنوية ماعدا متغير درجة اختبار وظيفة الكبد (X4) وبعد حذف هذا المتغير نقوم بتوفيق دالة الانحدار Y على المتغيرات المستقلة الأخرى X1,X2,X3 ، فحصلنا على:

```
MTB > Name c7 "e(Y/X1,X2,X3)" c8 "hii"
MTB > Regress 'Y' 3 'X1'-'X3';
SUBC> Residuals ' e(Y/X1,X2,X3)';
SUBC> Hi 'hii';
SUBC> Constant;
SUBC> Brief 2.
```

Regression Analysis: Y versus X1; X2; X3

The regression equation is

$$Y = - 659 + 38.3 X1 + 4.57 X2 + 4.49 X3$$

Predictor	Coef	SE Coef	T	P
Constant	-659.18	55.67	-11.84	0.000.
X1	38.323	5.326	7.20	0.000
X2	4.5677	0.4996	9.14	0.000
X3	4.4850	0.4002	11.21	0.000

S = 61.2185 R-Sq = 83.3% R-Sq(adj) = 82.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	931546	310515	82.85	0.000
Residual Error	50	187385	3748		
Total	53	1118931			

Source	DF	Seq SS
X1	1	155274
X2	1	305511
X3	1	470760

Unusual Observations

Obs	X1	Y	Fit	SE Fit	Residual	St Resid
13	5.8	830.00	512.89	23.67	317.11	5.62R
18	3.7	34.00	-100.54	21.82	134.54	2.35R
28	11.2	574.00	520.84	31.33	53.16	1.01 X
38	4.3	120.00	75.87	32.98	44.13	0.86 X

حيث أن قيمة المتوسط لقيم العزم تساوي:

$$\bar{h} = \frac{4}{54} = 0.0740741$$

وقيمة ضعف متوسط قيم العزم تساوي :

$$\frac{2 * p}{n} = \frac{2 * 4}{54} = 0.1481481$$

- جدول (1-3) قيم العزم والمشاهدات القاصية في المتغيرات المستقلة على الفترة التي يعيشها

المريض بعد الجراحة.

I	Y	X1	X2	X3	e(Y/X1,X2,X3,)	Hii	Big(hii)
1	200	6.7	62	81	-44.07	0.026364	
2	101	5.1	59	66	-0.775	0.029496	

i	Y	X1	X2	X3	e(Y/X1,X2,X3,)	Hii	Big(hii)
3	204	7.4	57	83	-53.028	0.045194	
4	101	6.5	73	41	-6.249	0.078851	
5	509	7.8	65	115	56.58	0.123355	
6	80	5.8	38	72	20.411	0.06224	
7	80	5.7	46	63	28.067	0.047159	
8	127	3.7	68	81	-29.508	0.053411	
9	202	6	67	93	-91.903	0.031147	
10	203	3.7	76	94	-48.356	0.072011	
11	329	6.3	84	83	-9.201	0.05003	
12	65	6.7	51	43	41.606	0.08116	
13	830	5.8	96	114	317.111	0.14949	0.14949
14	330	5.8	83	88	-6.897	0.049811	
15	168	7.7	62	67	-51.603	0.047838	
16	217	7.4	74	68	-50.404	0.044803	
17	87	6	85	28	2.405	0.149911	0.149911
18	34	3.7	51	41	134.544	0.127056	
19	215	7.3	68	74	-48.075	0.036141	
20	172	5.6	57	87	-33.987	0.025029	
21	109	5.2	52	76	-9.484	0.028868	
22	136	3.4	83	53	48.053	0.127368	
23	70	6.7	26	68	48.674	0.123988	
24	220	5.8	67	86	-34.844	0.022892	
25	276	6.3	59	100	-24.254	0.046282	
26	144	5.8	61	73	-25.132	0.01958	
27	181	5.2	52	86	17.666	0.031085	
28	574	11.2	76	90	53.164	0.261873	0.261873
29	72	5.2	54	56	34.082	0.047443	
30	178	5.8	76	59	3.143	0.043047	
31	71	3.2	64	65	23.684	0.080841	
32	58	8.7	45	23	75.068	0.21128	0.21128
33	116	5	59	73	-13.338	0.025231	
34	295	5.8	72	93	-14.078	0.034596	
35	115	5.4	58	70	-11.644	0.023871	
36	184	5.3	51	99	-36.904	0.047925	
37	118	2.6	74	86	-46.185	0.105898	
38	120	4.3	8	119	44.13	0.290152	0.290152
39	151	4.8	61	76	6.736	0.026137	
40	148	5.4	52	88	-31.969	0.031738	
41	95	5.2	49	72	8.16	0.035264	

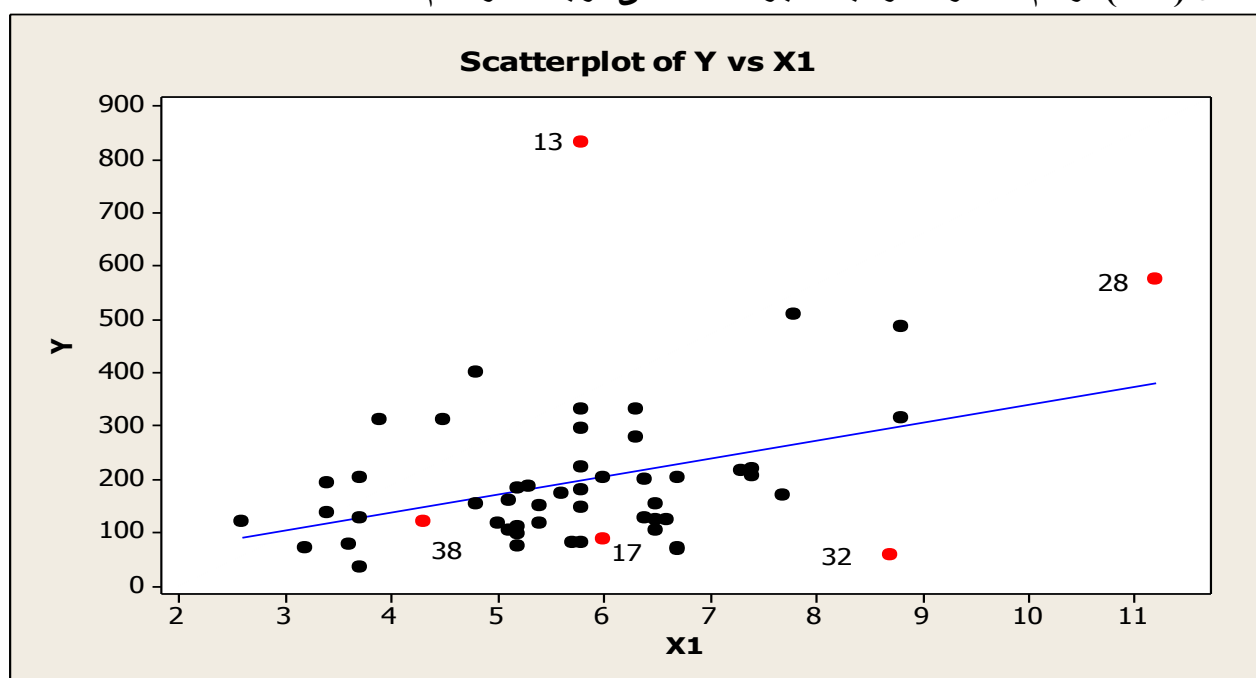
I	Y	X1	X2	X3	e(Y/X1,X2,X3,)	hii	Big(hii)
42	75	3.6	28	99	24.302	0.139301	
43	483	8.8	86	88	17.431	0.124222	
44	153	6.5	56	77	-38.059	0.02653	
45	191	3.4	77	93	-48.942	0.082842	
46	123	6.5	40	84	-26.371	0.063129	
47	311	4.5	73	106	-11.131	0.068607	
48	398	4.8	86	101	27.417	0.084136	
49	158	5.1	67	77	-29.652	0.023326	
50	310	3.9	82	103	-16.792	0.094148	
51	124	6.6	77	46	-27.778	0.071383	
52	125	6.4	85	40	-28.745	0.105363	
53	198	6.4	59	85	-38.81	0.026362	
54	313	8.8	78	72	-44.266	0.094799	

- الاستنتاج :-

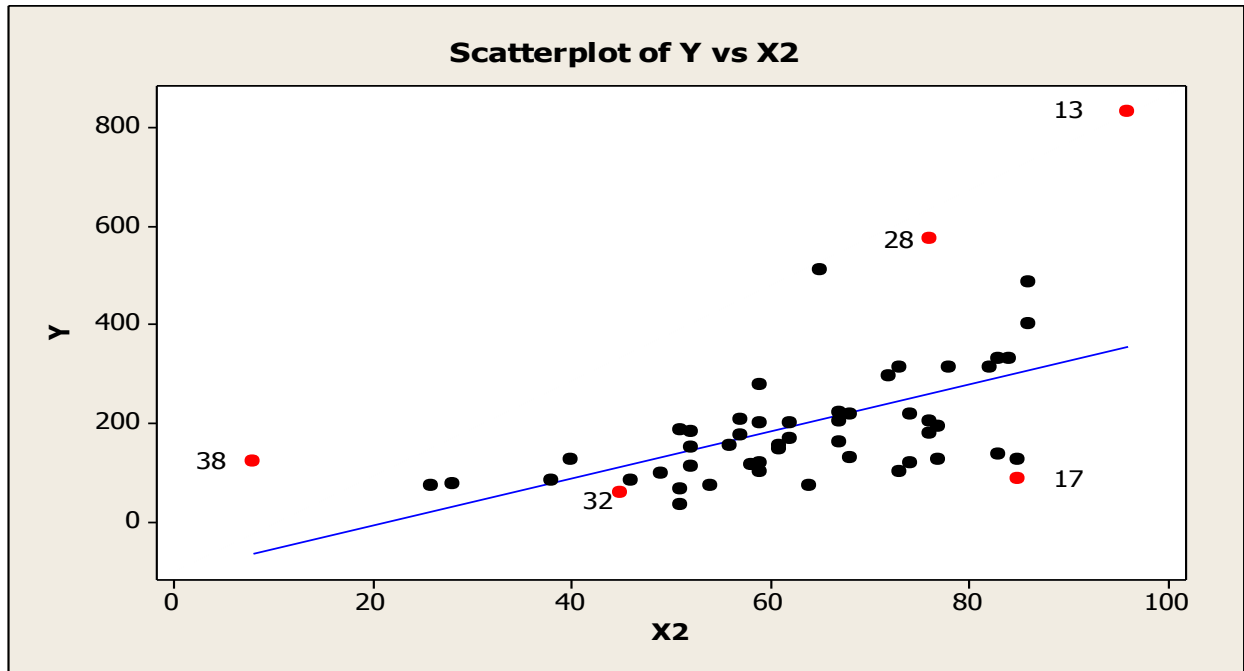
نلاحظ كما في الجدول (1-3) ، وجود خمس نقاط قاصية تفوق ضعف متوسط قيم العزم ، ويدل ذلك على وجود قاصيات في النموذج ، وهي المشاهدات 13, 17, 28, 32, 38 حيث أن قيم العزم لها تساوي:

$$h_{13,13} = 0.14949, h_{17,17} = 0.149911, h_{28,28} = 0.261873, h_{32,32} = 0.21128, h_{38,38} = 0.290152$$

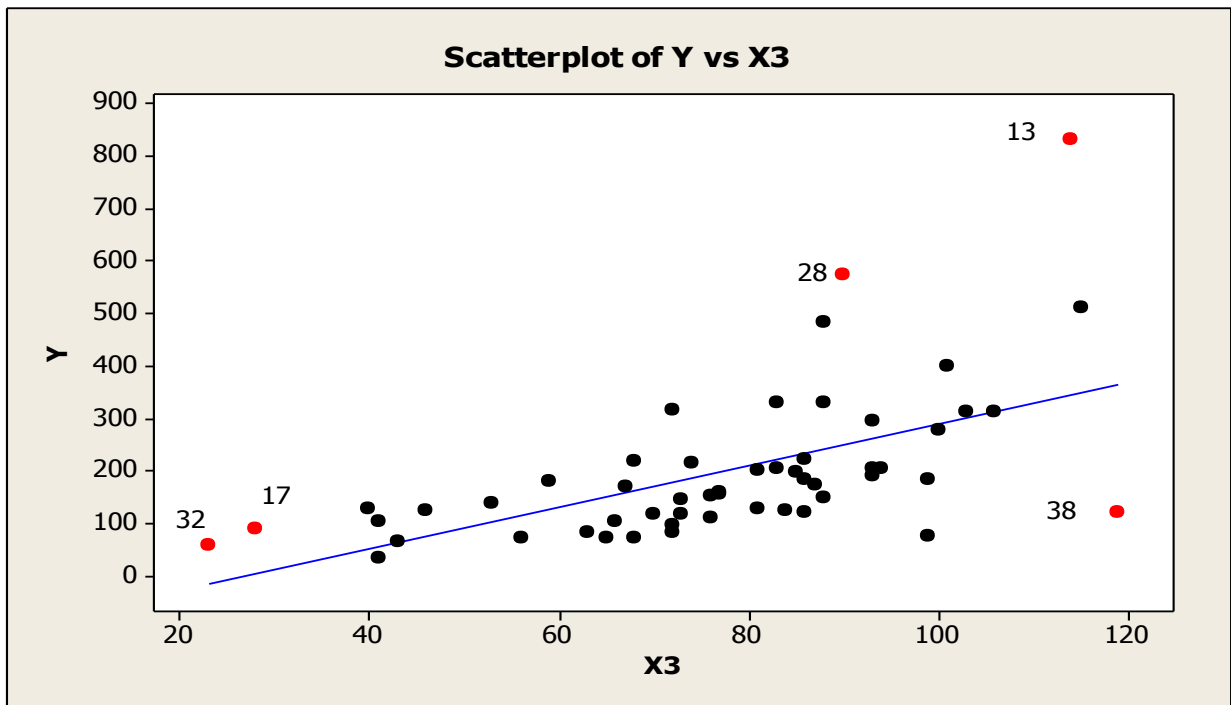
- شكل (1-7) رسم انتشار الفترة بعد الجراحة Y على درجة تخثر الدم X1 .



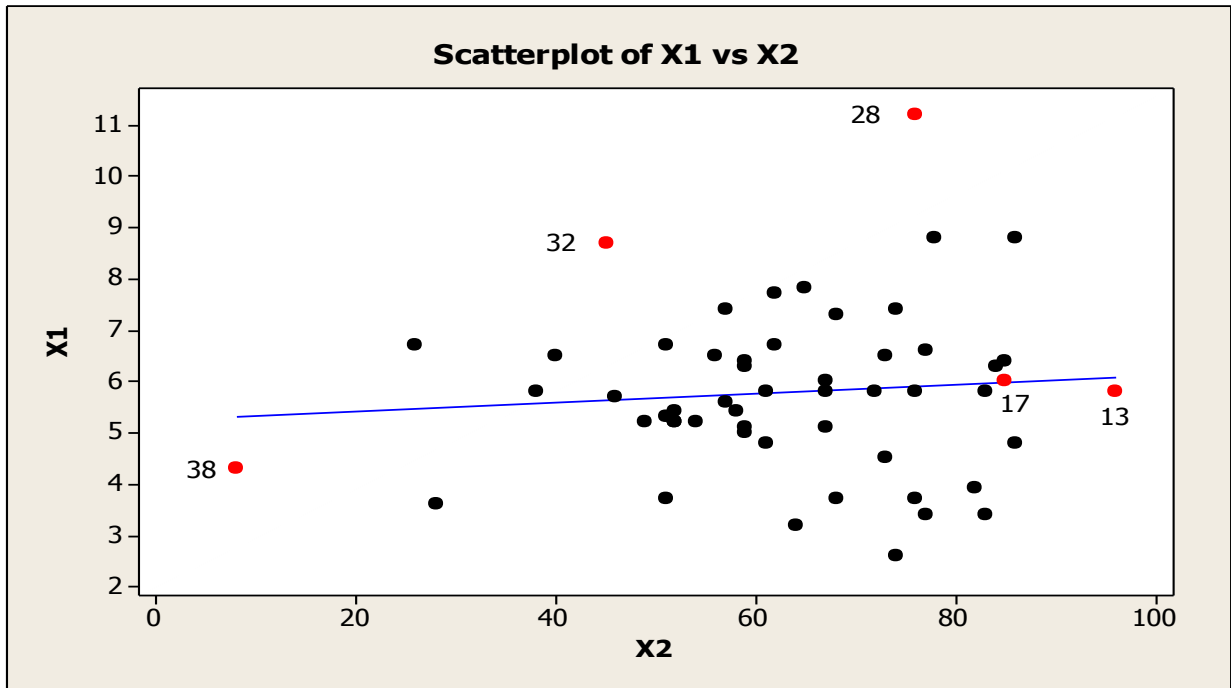
- شكل (1-8) رسم انتشار الفترة بعد الجراحة Y على دليل الإنذار القياسي X_2 .



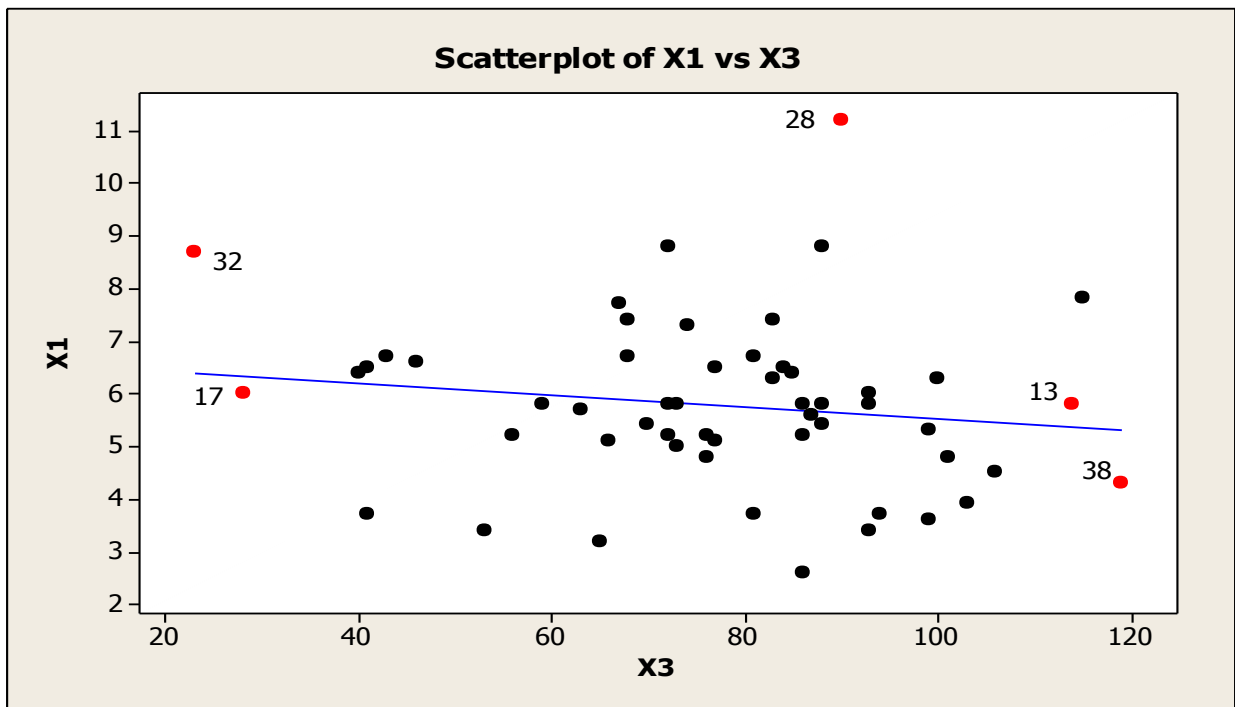
- شكل (1-9) رسم انتشار الفترة بعد الجراحة Y على درجة اختبار وظيفة الأنزيم X_3 .



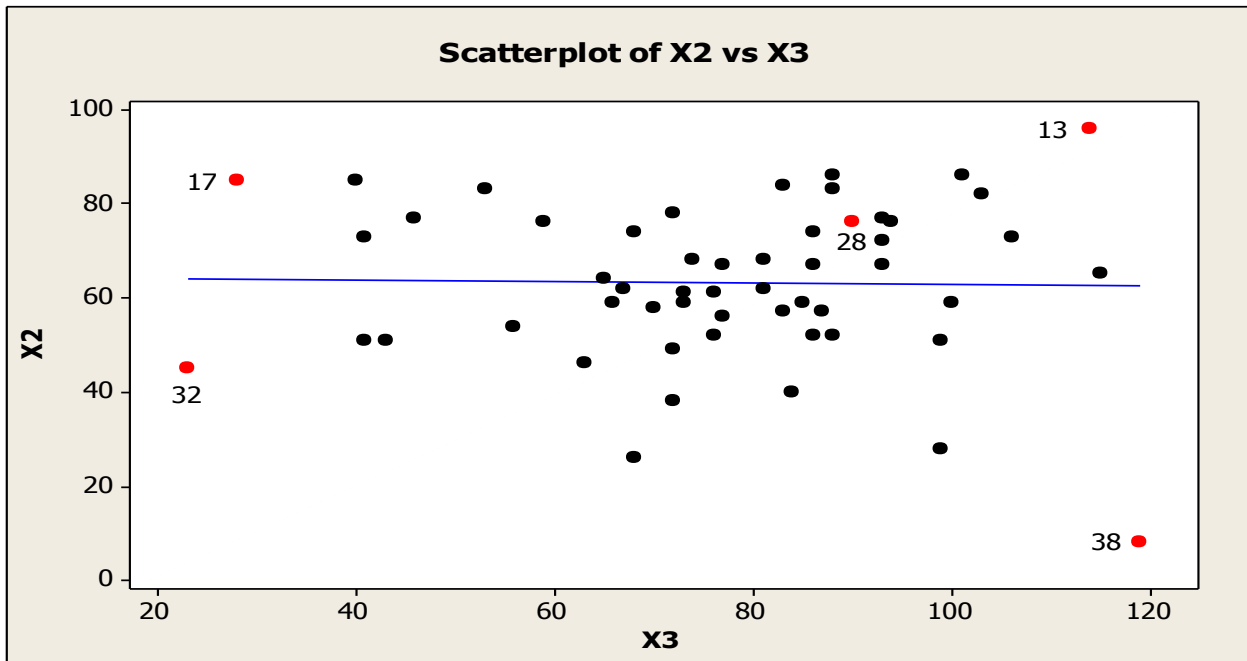
- شكل (1-10) رسم انتشار درجة تخثر الدم X1 على دليل الإنذار القياسي X2 .



- شكل (1-11) رسم انتشار درجة تخثر الدم X1 على درجة اختبار وظيفة الأنزيم X3 .

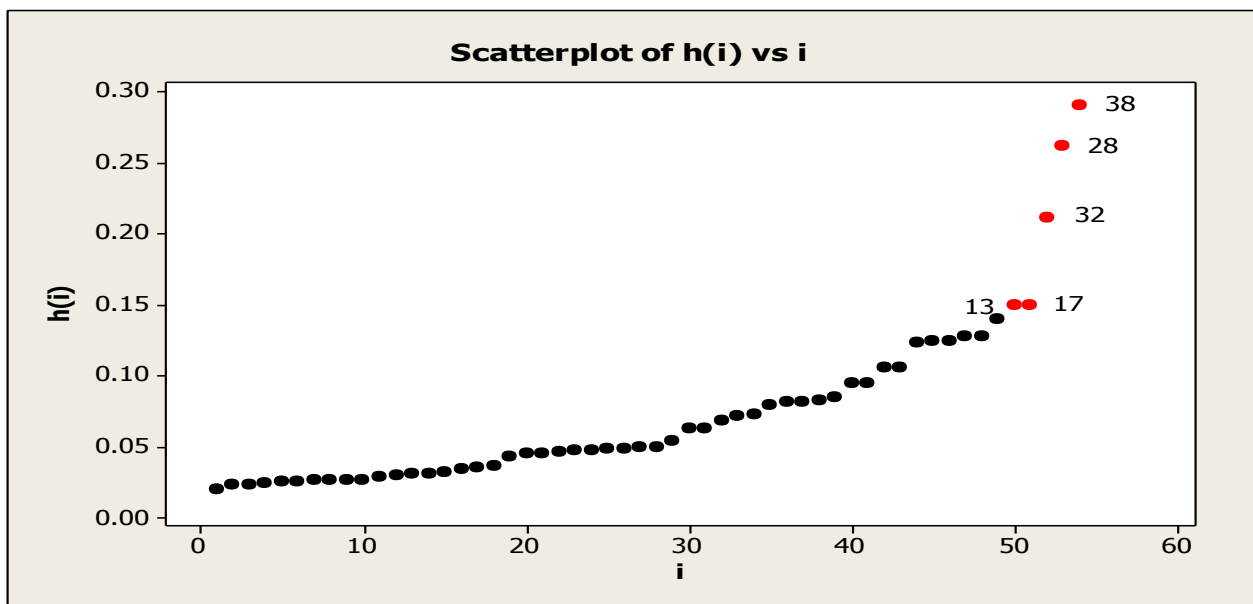


- شكل (1-12) رسم انتشار دليل الإنذار القياسي X_2 على درجة اختبار وظيفة الأنزيم X_3 .



نلاحظ من الرسومات السابقة أننا لا نستطيع اكتشاف بعض المشاهدات القاصية إلا بصعوبة بالغة وقد لا نستطيع اكتشاف جميع المشاهدات القاصية. ويرجع ذلك لأن النموذج يعتمد على أكثر من متغير. لذلك فإن طريقة اكتشاف القيم القاصية بواسطة قيم العزم والطرق الأخرى أفضل من طريقة رسم الانتشار.

- شكل (1-13) رسم انتشار (الترتيب التصاعدي لقيم العزم) $h(i)$ مقابل أرقام المشاهدات i .



نلاحظ أن قيمة العزم للملاحظة رقم 13 تساوي $h_{13,13} = 0.14949$ وتفصلها فجوة تساوي 0.0013419

عن قيمة متوسط قيم العزم . و قيمة العزم المشاهدة 17 تساوي $h_{17,17} = 0.149911$ وتفصلها فجوة تساوي 0.0017629 عن قيمة متوسط قيم العزم . و قيمة العزم المشاهدة 28 تساوي $h_{28,28} = 0.261873$ وتفصلها فجوة تساوي 0.113724 عن قيمة متوسط قيم العزم . و قيمة العزم المشاهدة 32 تساوي $h_{32,32} = 0.21128$ وتفصلها فجوة تساوي 0.0631319 عن قيمة متوسط قيم العزم . و قيمة العزم المشاهدة 38 تساوي $h_{38,38} = 0.290152$ عن قيمة متوسط قيم العزم .

- تطبيق (1-4) :-

في بيانات عن مبيعات (المتغير التابع Y) في 15 منطقة بيع و ذلك لمنتج ألواح إسفلتية لكسوة السقوف . و المتغيرات المؤثرة عليها مثل الحسابات الترويجية (X1 بالآف الدولارات) ، عدد الحسابات المصرفية العاملة X2. عدد الأصناف المنافسة X3. وإمكانات المنطقة X4 لكل منطقة من المناطق.

المنطقة = i	المبيعات = Y	الحسابات = X1 الترويجية	عدد الحسابات = X2 المصرفية	عدد = X3 الأصناف	إمكانات = X4 المنطقة
1	79.3	5.5	31	10	8
2	200.1	2.5	55	8	6
3	163.2	8	67	12	9
4	200.1	3	50	7	16
5	146	3	38	8	15
6	177.7	2.9	71	12	17
7	30.9	8	30	12	8
8	291.9	9	56	5	10
9	160	4	42	8	4
10	339.4	6.5	73	5	16
11	159.6	5.5	60	11	7
12	86.3	5	44	12	12
13	237.5	6	50	6	6
14	107.2	5	39	10	4
15	155	3.5	55	10	4

- التحليل الإحصائي :

باستخدام برنامج الحزم الإحصائية Minitab سوف نقوم باستخراج قيم العزم لمصفوفة القبة واكتشاف المشاهدات الشاذة في المتغيرات المستقلة (الحسابات الترويجية ، عدد الأصناف ، عدد الحسابات المصرفية ، إمكانات المنطقة) .

$$H_0: B_i = 0 \quad vs \quad H_1: B_i \neq 0 \quad i = 0,1,2,3,4$$

```
MTB > Name c9 "e(Y/X1,X2,X3,X4)" c10 "hii"
MTB > Regress 'Y' 4 'X1'-'X4';
SUBC> Residuals ' e(Y/X1,X2,X3,X4) ' ;
SUBC> Hi 'hii';
SUBC> Constant;
SUBC> Brief 2.
```

Regression Analysis: Y versus X1; X2; X3; X4

The regression equation is

$$Y = 177 + 2.17 X1 + 3.54 X2 - 22.2 X3 + 0.204 X4$$

Predictor	Coef	SE Coef	T	P
Constant	177.229	8.787	20.17	0.000
X1	2.1702	0.6737	3.22	0.009
X2	3.5380	0.1092	32.41	0.000
X3	-22.1583	0.5454	-40.63	0.000
X4	0.2035	0.3189	0.64	0.538

S = 5.11930 R-Sq = 99.7% R-Sq(adj) = 99.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	89285	22321	851.72	0.000
Residual Error	10	262	26		
Total	14	89547			

Source	DF	Seq SS
X1	1	1074
X2	1	44505
X3	1	43695
X4	1	11

Unusual Observations

Obs	X1	Y	Fit	SE Fit	Residual	St Resid
4	3.00	200.10	208.79	2.91	-8.69	-2.06R

R denotes an observation with a large standardized residual.

نلاحظ أن المتغير X4 غير معنوي ويمكن حذفه . وبعد حذفه وتقدير معادلة انحدار Y على بقية المتغيرات

المستقلة X1,X2,X3 نحصل على ما يلي :

Regression Analysis: Y versus X1; X2; X3

The regression equation is

$$Y = 179 + 2.11 X1 + 3.56 X2 - 22.2 X3$$

Predictor	Coef	SE Coef	T	P
Constant	178.521	8.318	21.46	0.000
X1	2.1055	0.6479	3.25	0.008
X2	3.56240	0.09945	35.82	0.000
X3	-22.1880	0.5286	-41.98	0.000

$$S = 4.97952 \quad R-Sq = 99.7\% \quad R-Sq(adj) = 99.6\%$$

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	89274	29758	1200.14	0.000
Residual Error	11	273	25		
Total	14	89547			

Source	DF	Seq SS
X1	1	1074
X2	1	44505
X3	1	43695

حيث أن قيمة المتوسط لقيمة العزم تساوي:

$$\bar{h} = \frac{4}{15} = 0.266667$$

وقيمة ضعف متوسط قيم العزم يساوي:

$$\frac{2 * 4}{15} = 0.5333333$$

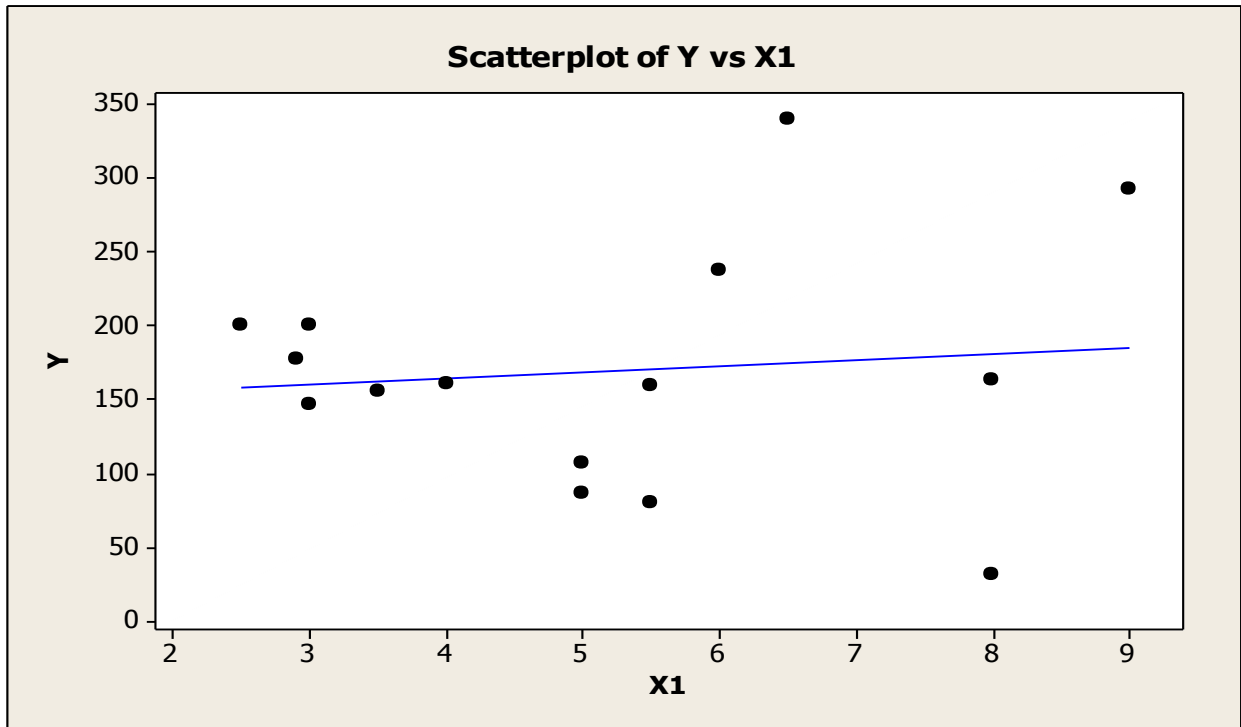
- جدول (1-4) قيم العزم والمشاهدات القاصية في المتغيرات المستقلة.

i	X1	X2	X3	Y	e(Y/X1,X2,X3)	Hii	Big hii
1	5.5	31	10	79.3	0.64432	0.223605	
2	2.5	55	8	200.1	-2.11269	0.20889	
3	8	67	12	163.2	-4.59006	0.441628	
4	3	50	7	200.1	-7.54144	0.200942	
5	3	38	8	146	3.29538	0.231713	
6	2.9	71	12	177.7	6.39862	0.446299	
7	8	30	12	30.9	-5.08116	0.452146	
8	9	56	5	291.9	5.87488	0.476566	
9	4	42	8	160	0.94022	0.139043	
10	6.5	73	5	339.4	-1.9221	0.414587	
11	5.5	60	11	159.6	-0.17736	0.156596	
12	5	44	12	86.3	6.76184	0.169198	
13	6	50	6	237.5	1.35393	0.182656	
14	5	39	10	107.2	1.09787	0.124677	
15	3.5	55	10	155	-4.94225	0.131457	

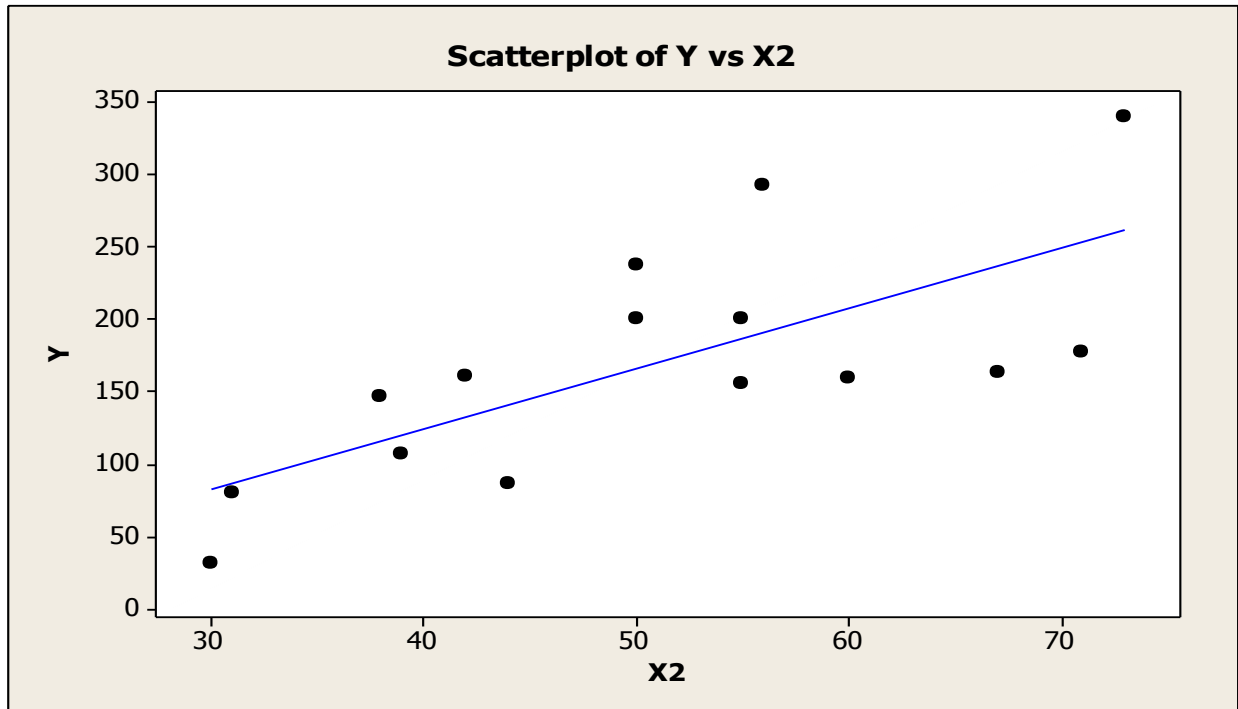
- الاستنتاج :-

نلاحظ كما في الجدول (1-4)، أنه لا يوجد قيم أكبر من ضعف متوسط قيم العزم، وحيث يشير ذلك إلى أنه لا يوجد قيم شاذة في النموذج.

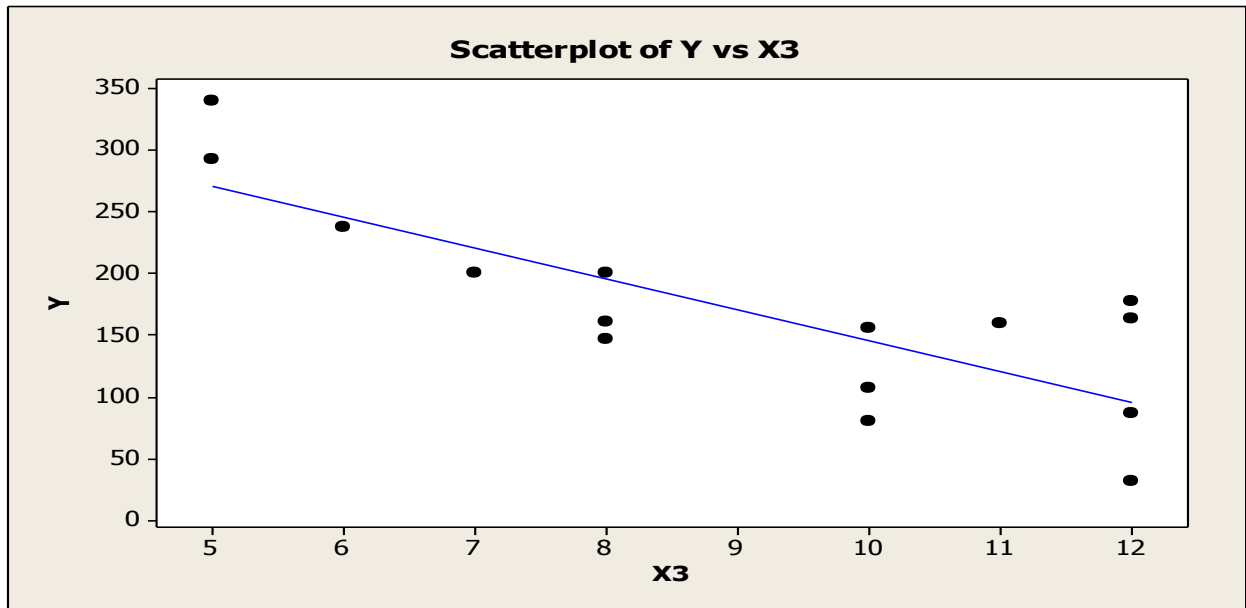
- شكل (1-14) : رسم انتشار المبيعات Y على الحسابات الترويجية X_1 .



- شكل (1-15) : رسم انتشار المبيعات Y على عدد الحسابات المصرفية X_2 .

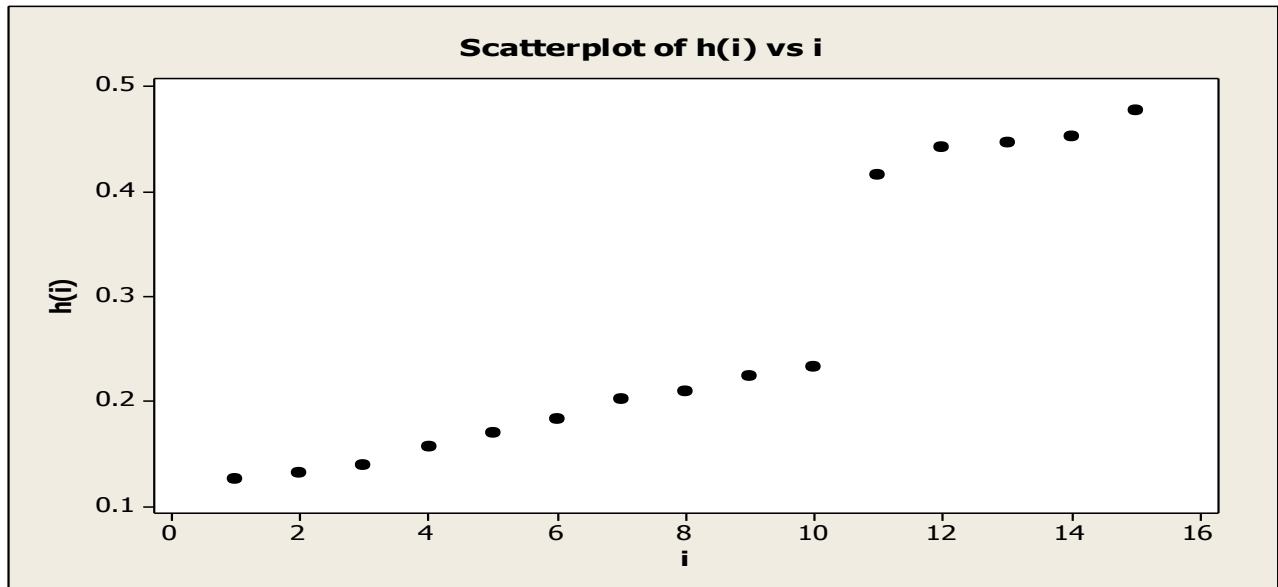


- شكل (1-16) رسم انتشار المبيعات Y على عدد الأصناف X_3 .



نلاحظ من رسومات الانتشار السابقة أنه لا نستطيع أن نقول أن النموذج لا يحوي نقاط شاذة. لذلك فإن طريقة اكتشاف المشاهدات القاصية بواسطة قيم العزم والطرق الأخرى أفضل من طريقة مشاهدة القيم من خلال الرسم.

- شكل (1-17): رسم انتشار (الترتيب التصاعدي لقيم العزم) $h(i)$ مقابل أرقام المشاهدات i .



نلاحظ من الرسم السابق أنه لا يوجد قيم شاذة في النموذج وذلك لأن قيم العزم أقل من ضعف متوسط قيم العزم.

