

As vendas da *Sunflowers Apparel*, uma cadeia de lojas de roupa de luxo para senhoras, aumentaram nos últimos 12 anos à medida que a cadeia foi expandindo o número de lojas.

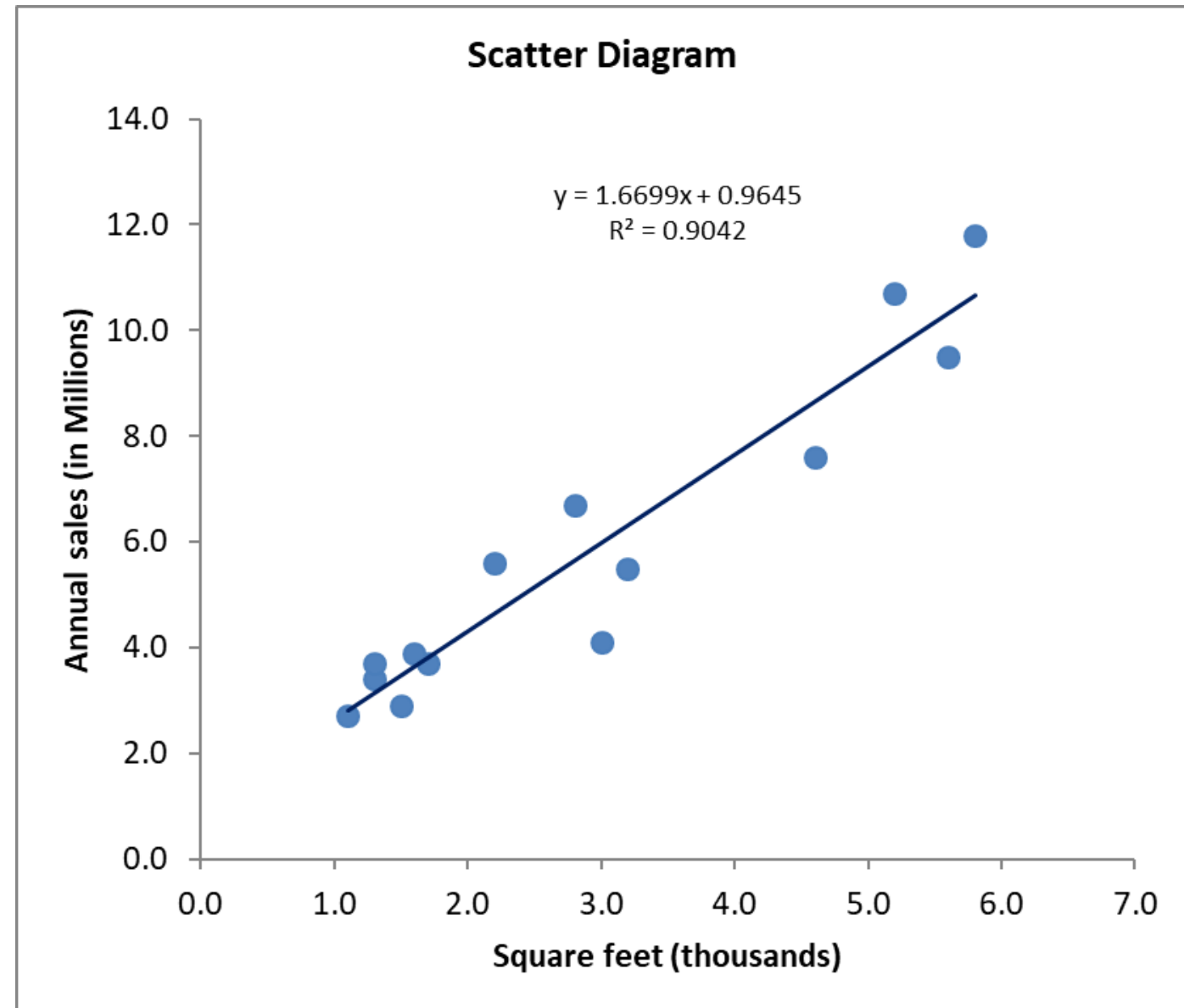
Até à data, os gerentes da *Sunflowers Apparel* selecionavam locais com base em factores subjetivos, como um bom preço de aluguer ou a percepção de que uma certa localização parecia ideal para uma loja de roupa.

Um novo diretor de planeamento precisa desenvolver uma abordagem sistemática que levará a tomar melhores decisões durante o processo de seleção do local. Como ponto de partida, ele acredita que o tamanho da loja contribui significativamente para as vendas e quer usar essa relação no processo de tomada de decisão.

Como se pode utilizar estatística para prever as vendas anuais de uma nova loja, proposta com base no tamanho da mesma?

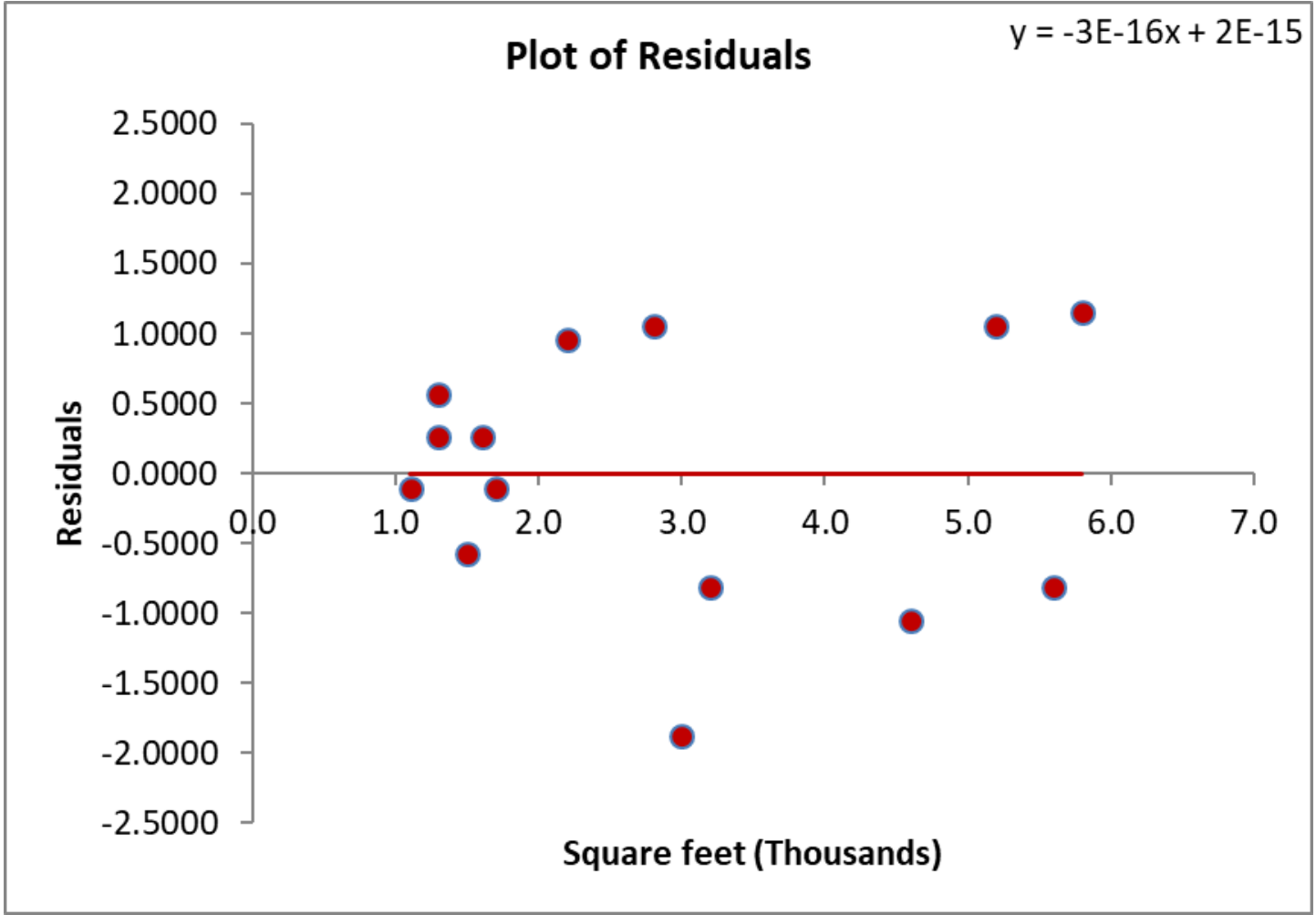
Para examinar a relação entre o tamanho da loja (em pés quadrados) e as suas vendas anuais, foram recolhidos dados de uma amostra de 14 lojas.

	X_i	Y_i
Store	Square feet (Thousands)	Annual Sales (in \$000,000)
1	1.7	3.7
2	1.6	3.9
3	2.8	6.7
4	5.6	9.5
5	1.3	3.4
6	2.2	5.6
7	1.3	3.7
8	1.1	2.7
9	3.2	5.5
10	1.5	2.9
11	5.2	10.7
12	4.6	7.6
13	5.8	11.8
14	3.0	4.1

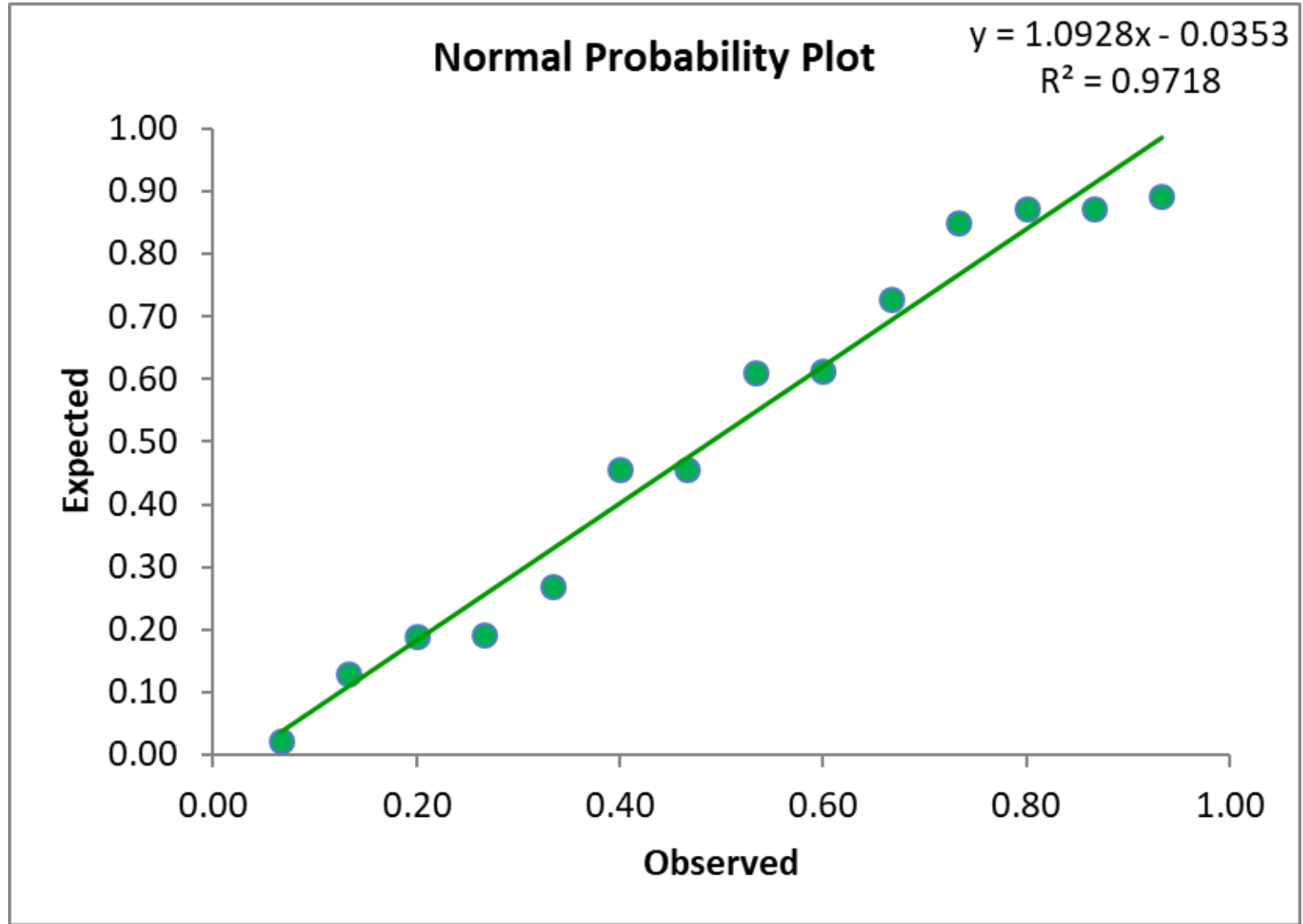


SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0.95088				
R Square	0.90418				
Adjusted R Square	0.89619				
Standard Error	0.96638				
Observations	14				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	105.7476	105.7476	113.2335	0.0000
Residual	12	11.2067	0.9339		
Total	13	116.9543			
<i>Coefficients</i>					
Intercept	0.9645				
Square feet (Thousands)	1.6699				

RESIDUAL OUTPUT		
<i>Observation</i>	<i>Predicted Annual Sales (in \$000,000)</i>	<i>Residuals</i>
1	3.8032	-0.1032
2	3.6363	0.2637
3	5.6401	1.0599
4	10.3157	-0.8157
5	3.1353	0.2647
6	4.6382	0.9618
7	3.1353	0.5647
8	2.8013	-0.1013
9	6.3080	-0.8080
10	3.4693	-0.5693
11	9.6478	1.0522
12	8.6458	-1.0458
13	10.6497	1.1503
14	5.9741	-1.8741



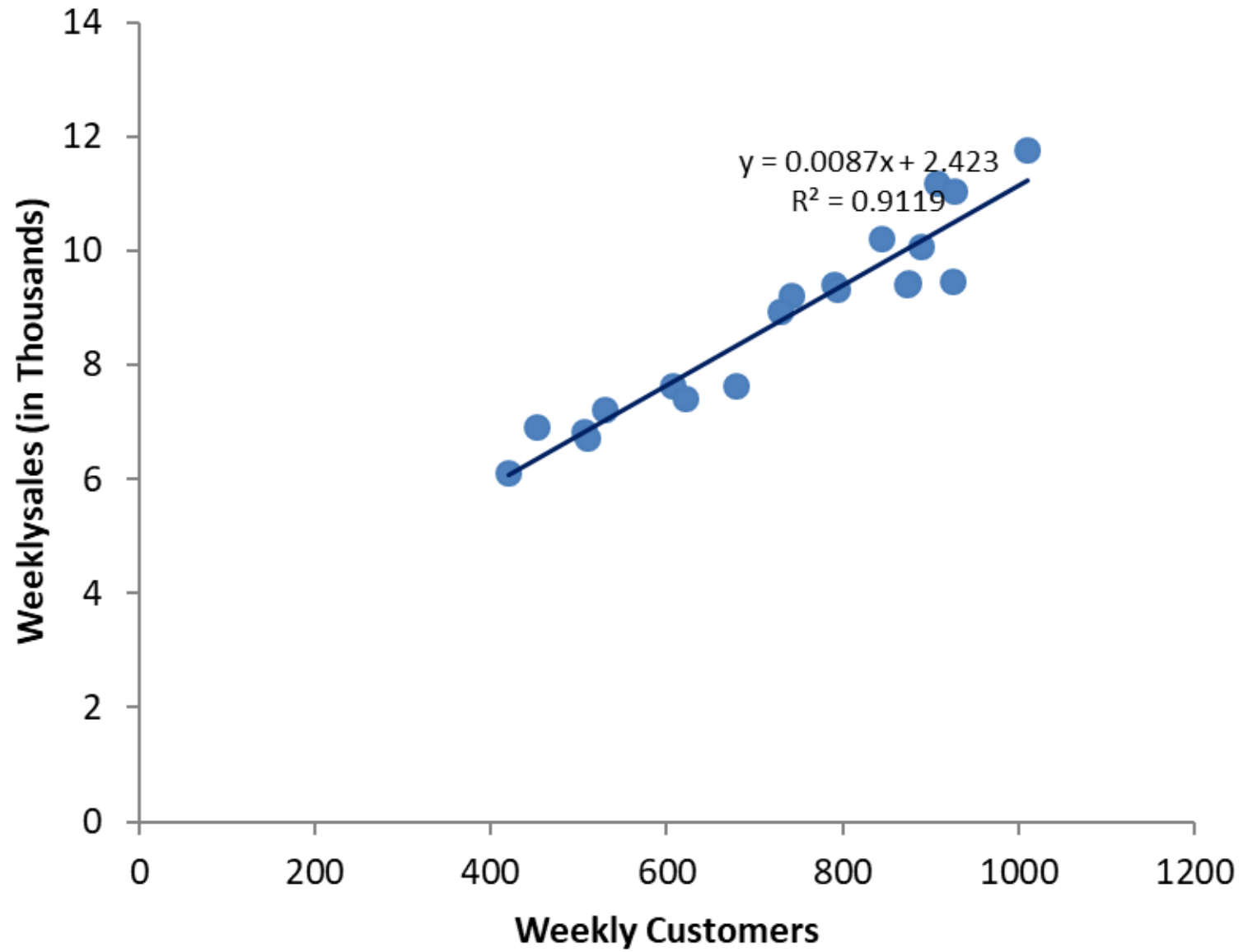
<i>Rank</i>	<i>Ordered Standardized Residuals</i>	<i>Rank/15</i>	<i>Normal Standard</i>
1	-2.0184	0.06667	0.02177
2	-1.1264	0.13333	0.12999
3	-0.8785	0.20000	0.18982
4	-0.8703	0.26667	0.19207
5	-0.6131	0.33333	0.26990
6	-0.1112	0.40000	0.45573
7	-0.1091	0.46667	0.45655
8	0.2841	0.53333	0.61182
9	0.2851	0.60000	0.61222
10	0.6082	0.66667	0.72848
11	1.0359	0.73333	0.84988
12	1.1333	0.80000	0.87146
13	1.1416	0.86667	0.87318
14	1.2390	0.93333	0.89232



A administração de uma cadeia de lojas de entrega de encomendas gostaria de desenvolver um modelo para prever as vendas semanais (em milhares de dólares) de cada loja individualmente, com base no número semanal de clientes. Uma amostra aleatória de 20 lojas foi selecionada entre todas as lojas da cadeia. Os resultados são mostrados abaixo.

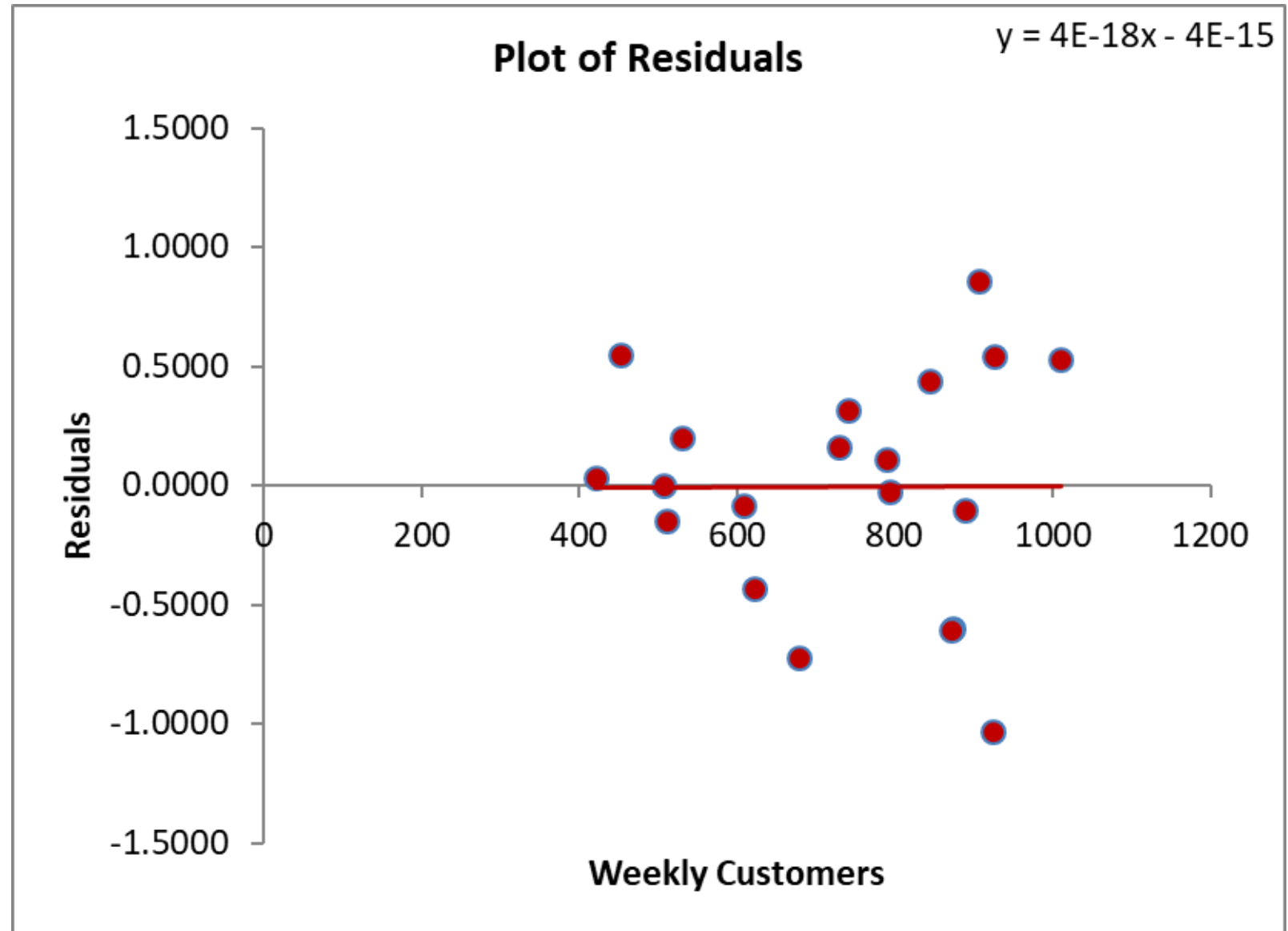
	X_i	Y_i		X_i	Y_i	
<i>Store</i>	<i>Customers</i>	<i>Sales (\$000)</i>		<i>Store</i>	<i>Customers</i>	<i>Sales (\$000)</i>
1	907	11.2		11	679	7.63
2	926	11.05		12	872	9.43
3	506	6.84		13	924	9.46
4	741	9.21		14	607	7.64
5	789	9.42		15	452	6.92
6	889	10.08		16	729	8.95
7	874	9.45		17	794	9.33
8	510	6.73		18	844	10.23
9	529	7.24		19	1010	11.77
10	420	6.12		20	621	7.41

Scatter Diagram

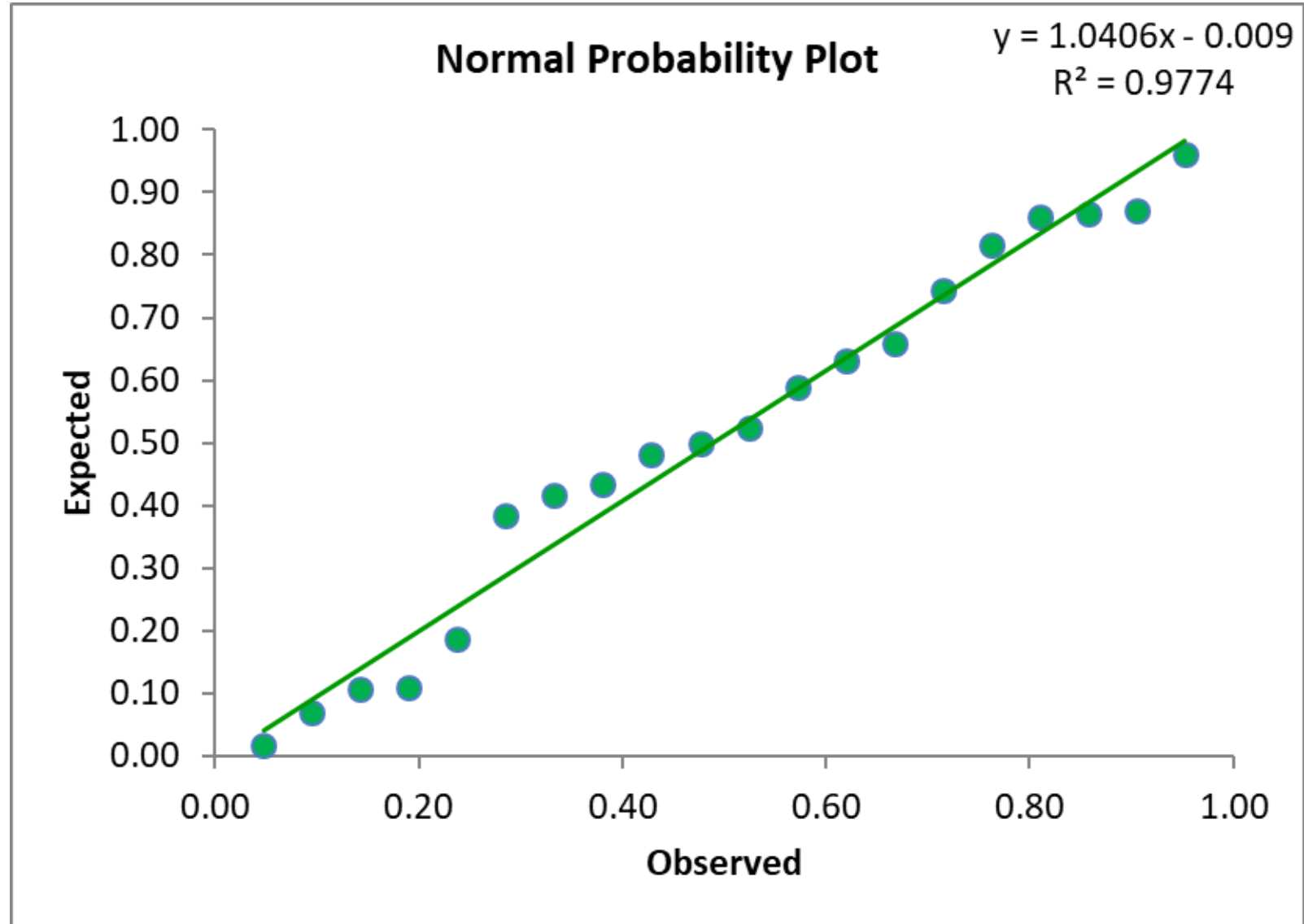


SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0.95491				
R Square	0.91186				
Adjusted R Square	0.90696				
Standard Error	0.50150				
Observations	20				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	46.8335	46.8335	186.2188	0.0000
Residual	18	4.5270	0.2515		
Total	19	51.3605			
<i>Coefficients</i>					
Intercept	2.4230				
Customers	0.0087				

RESIDUAL OUTPUT		
<i>Observation</i>	<i>Predicted Sales (\$000)</i>	<i>Residuals</i>
1	10.3406	0.8594
2	10.5064	0.5436
3	6.8401	-0.0001
4	8.8915	0.3185
5	9.3105	0.1095
6	10.1834	-0.1034
7	10.0525	-0.6025
8	6.8750	-0.1450
9	7.0409	0.1991
10	6.0894	0.0306
11	8.3503	-0.7203
12	10.0350	-0.6050
13	10.4890	-1.0290
14	7.7218	-0.0818
15	6.3687	0.5513
16	8.7867	0.1633
17	9.3541	-0.0241
18	9.7906	0.4394
19	11.2397	0.5303
20	7.8440	-0.4340



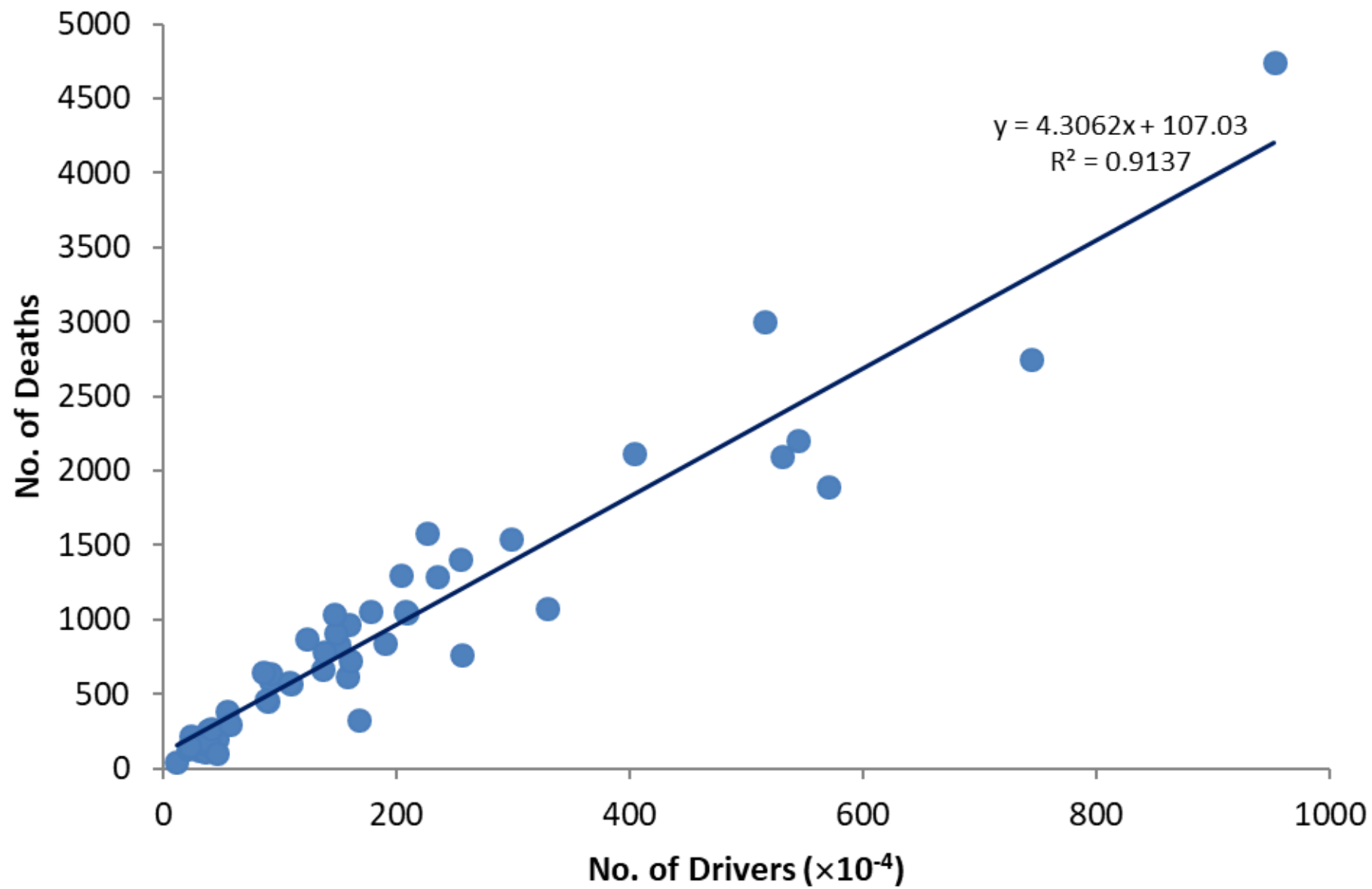
<i>Rank</i>	<i>Ordered Standardized Residuals</i>	<i>Rank/21</i>	<i>Normal Standard</i>
1	-2.1080	0.04762	0.01752
2	-1.4756	0.09524	0.07003
3	-1.2395	0.14286	0.10758
4	-1.2343	0.19048	0.10855
5	-0.8891	0.23810	0.18699
6	-0.2971	0.28571	0.38321
7	-0.2119	0.33333	0.41610
8	-0.1675	0.38095	0.43349
9	-0.0495	0.42857	0.48028
10	-0.0002	0.47619	0.49993
11	0.0628	0.52381	0.52502
12	0.2243	0.57143	0.58876
13	0.3345	0.61905	0.63099
14	0.4080	0.66667	0.65835
15	0.6525	0.71429	0.74297
16	0.9002	0.76190	0.81599
17	1.0865	0.80952	0.86136
18	1.1136	0.85714	0.86728
19	1.1294	0.90476	0.87064
20	1.7607	0.95238	0.96086



			X_i	Y_i
<i>Obs</i>	<i>State</i>	<i>Initials</i>	$10^{-4} \times \text{No drivers}$	<i>No. Deaths</i>
1	Alabama	AL	158	968
2	Alaska	AK	11	43
3	Arizona	AZ	91	588
4	Arkansas	AR	92	640
5	California	CA	952	4743
6	Colorado	CO	109	566
7	Connecticut	CT	167	325
8	Delaware	DE	30	118
9	Distr. of Columb.	DC	35	115
10	Florida	FL	298	1545
11	Georgia	GA	203	1302
12	Idaho	ID	41	262
13	Illinois	IL	544	2207
14	Indiana	IN	254	1410
15	Iowa	IA	150	833
16	Kansas	KS	136	669
17	Kentucky	KY	147	911
18	Louisiana	LA	146	1037
19	Maine	ME	46	196
20	Maryland	MD	157	616
21	Massachusetts	MA	255	766
22	Michigan	MI	403	2120
23	Minnesota	MN	189	841
24	Mississippi	MS	85	648
25	Missouri	MO	234	1289

26	Montana	MT	38	259
27	Nebraska	NE	89	450
28	Nevada	NV	23	215
29	New Hampshire	NH	37	158
30	New Jersey	NJ	329	1071
31	New Mexico	NM	54	387
32	New York	NY	744	2745
33	North Carolina	NC	226	1580
34	North Dakota	ND	38	185
35	Ohio	OH	530	2096
36	Oklahoma	OK	137	785
37	Oregon	OR	108	575
38	Pennsylvania	PA	570	1889
39	Rhode Island	RI	46	100
40	South Carolina	SC	122	870
41	South Dakota	SD	40	270
42	Tennessee	TN	177	1059
43	Texas	TX	515	3006
44	Utah	UT	57	295
45	Vermont	VT	20	131
46	Virginia	VA	208	1050
47	Washington	WA	160	730
48	West Virginia	WV	88	467
49	Wisconsin	WI	207	1059
50	Wyoming	WY	22	148

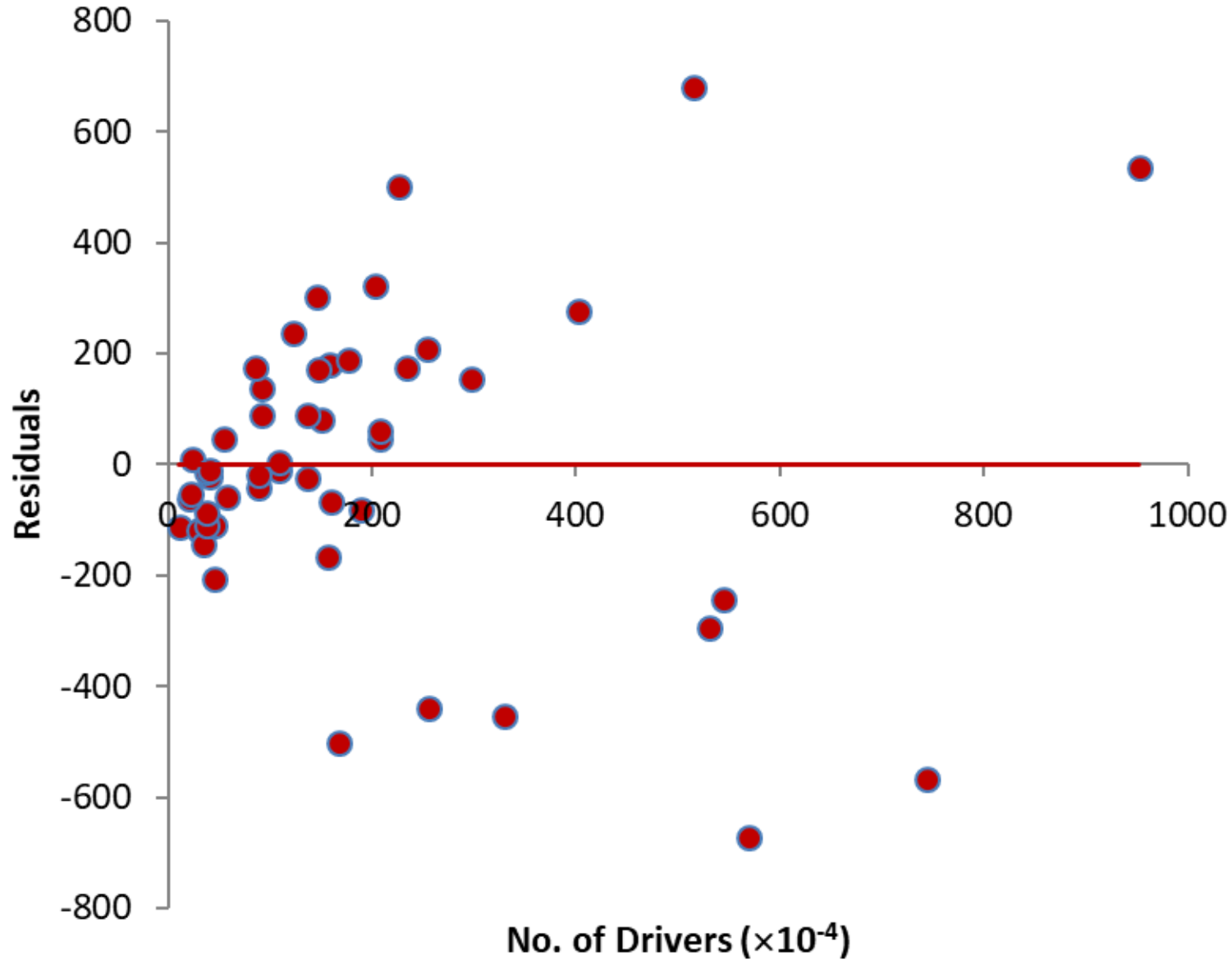
Scatter Diagram



SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0.9559				
R Square	0.9137				
Adjusted R Square	0.9119				
Standard Error	263.8960				
Observations	50				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	35410709.228	35410709.228	508.474	0.000
Residual	48	3342771.892	69641.081		
Total	49	38753481.120			
<i>Coefficients</i>					
Intercept	107.029				
10-4xNo drivers	4.306				

Plot of Residuals

$$y = -9E-16x + 2E-13$$



Normal Probability Plot

$$y = 0.9234x + 0.043$$

$$R^2 = 0.9745$$

