



2025 ~ 40° Aniversario
de la Creación del Consejo
Interuniversitario Nacional



UNIVERSIDAD NACIONAL DEL LITORAL

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

ESTADÍSTICA

PROBABILIDAD Y ESTADÍSTICA

RESPUESTAS DE GUÍA DE PRÁCTICA

UNIDAD 7 - REGRESIÓN Y CORRELACIÓN

Responsable de cátedra: Prof. Juan Pablo Taulamet

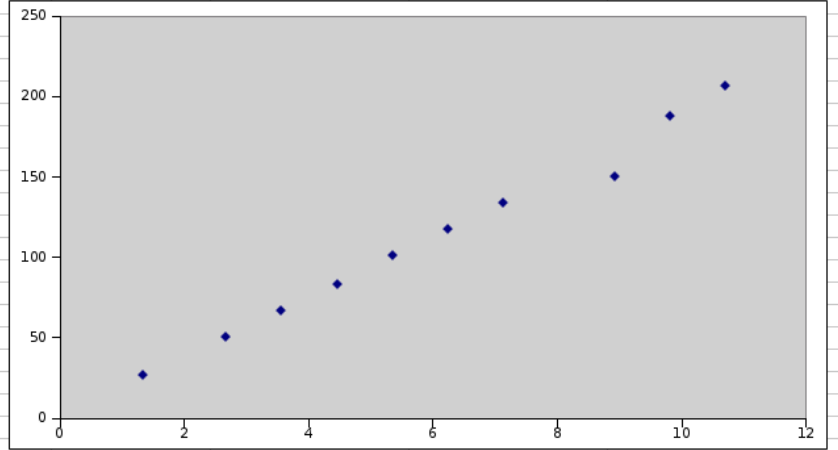
Equipo de cátedra: *Auxiliares:* Ing. Ana Lisa Eusebi (JTP) - Prof. Fátima Bolatti (JTP) - Lic. Denis Lizazo Torres (Ay. 1°) *Ayudantes:* AIA Cristian Bottazzi - Téc. Eliana García

Carreras: Ingeniería en Recursos Hídricos - Ingeniería en Informática - Ingeniería Ambiental - Ingeniería en Agrimensura - Ingeniería en Inteligencia Artificial

AÑO ACADÉMICO 2025 - PRIMER CUATRIMESTRE

Ejercicio 1

A	B	C	D	E	F	G
x	y					
1,33	27					
2,68	50					
3,57	67					
4,46	83					
5,35	101					
6,24	117					
7,14	134					
8,93	150					
9,82	188					
10,7	206					
SALIDA RESUMEN						
Estadísticas de regresión						
R múltiple	0,9286816639417					
R ²	0,8624496329414					
Error estándar	9,7156543471832					
R ² ajustado	0,8452558370591					
Observaciones	10					
Análisis de varianza						
	df	SS	MS	F		
Regresión	1	4734,84848484848	4734,8484848485	50,1605136436597	0,00010378187886	
Residual	8	755,151515151515	94,393939393939			
Total	9	5490				
	Coefficientes	Error estándar	Estadísticas-t	Valor-P	0,95	0,95
Interceptar	19,030303030303	10,6160558962146	1,7925963480551	0,11079931105051	-5,45036576592991	43,510971826536
Columna 1	0,7575757575758	0,10696583390741	7,0824087458759	0,00010378187886	0,51091210225967	1,00423941289185



a) De la observación del gráfico de dispersión parece adecuado proponer un modelo lineal para establecer una relación entre las variables velocidad y esfuerzo.

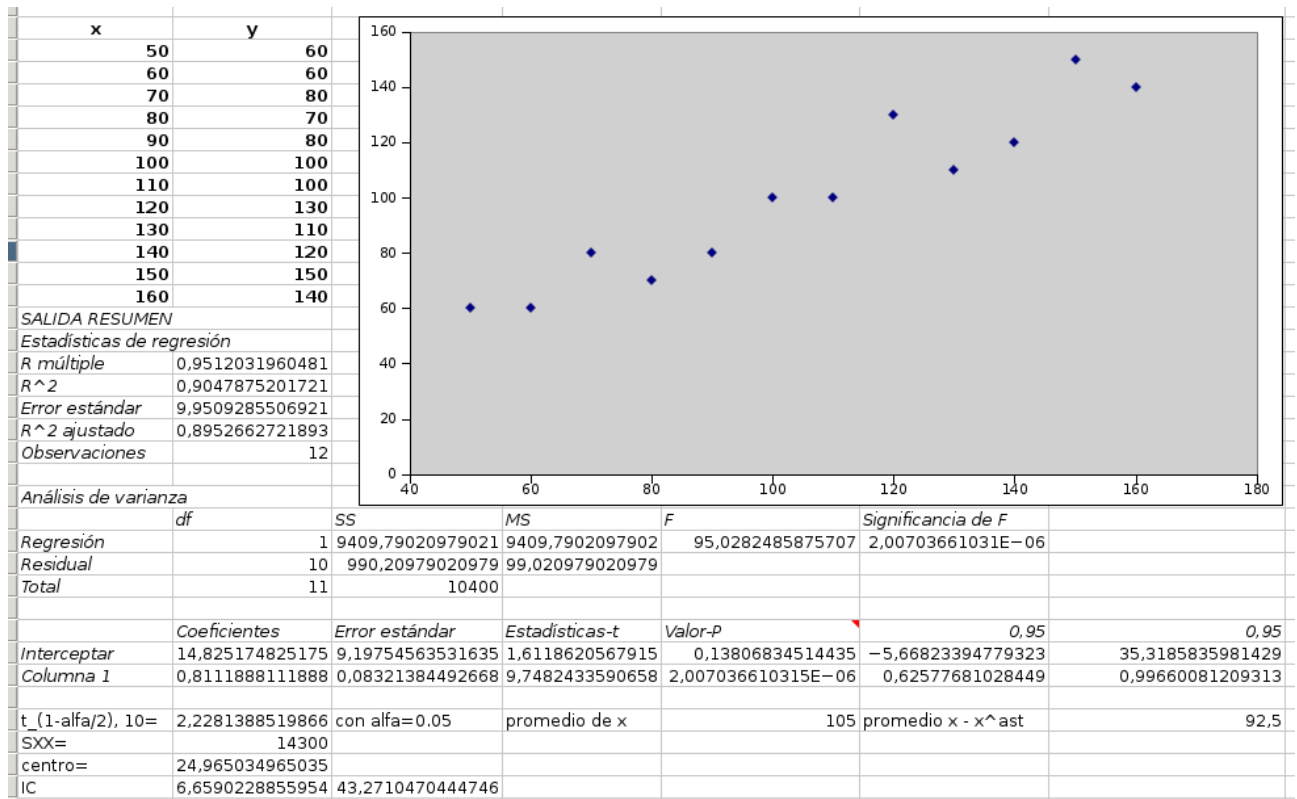
Se observa que el coeficiente de determinación r^2 es 0.989, es decir que el modelo propuesto sirve para explicar el 98% de la variabilidad de Y. A su vez por tratarse de un modelo lineal, cobra especial sentido analizar el valor del coeficiente de correlación $r = 0,928$ que implica un alto grado de asociación lineal entre las variables. Considerando que el coeficiente de correlación es positivo, la relación entre X e Y es directamente proporcional.

Además, la recta que resulta de la estimación por mínimos cuadrados es $\hat{y} = 0.698 + 18.532x$.

b) El valor estimado por mínimos cuadrados del coeficiente de regresión (pendiente de la recta b) es de 18.532. En la tabla además se presenta un intervalo del 95% para ese coeficiente. Como 20 es un valor perteneciente al intervalo calculado, 20 es un valor posible para la pendiente β con esa confianza.

c) Cuando $x = 14$, el valor esperado de y resulta $0.698 + 18.532 \cdot 14 = 260.151$

Ejercicio 2



a) La recta obtenida por el método de mínimos cuadrados es $\hat{y} = 14.82 + 0.81x$. El valor obtenido para el coeficiente de determinación $r^2 = 0.9047$, pero por tratarse de un modelo lineal, le daremos más importancia al valor del coeficiente de correlación $r = 0.95$ que implica un fuerte grado de relación lineal proporcional, por lo que el ajuste obtenido puede servir para ser usado como herramienta de pronóstico.

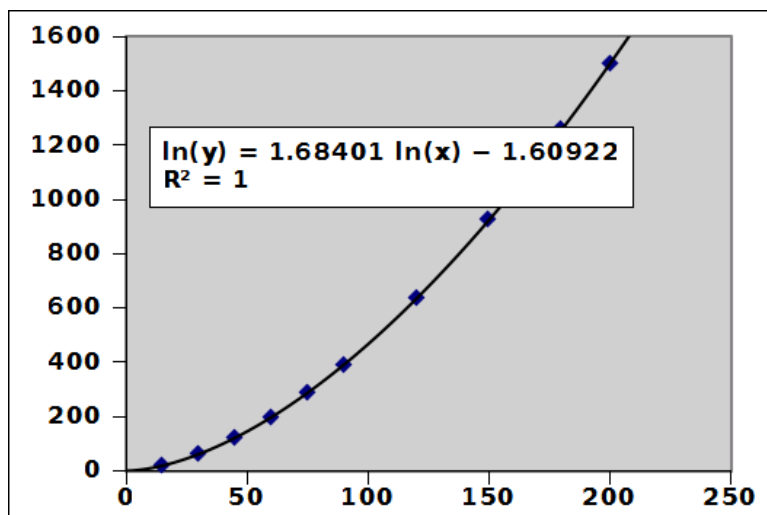
b) De la tabla resulta que la varianza de la regresión estimada es 99.02.

c) La estimación puntual de la pendiente b es 0.81 y el IC para β con $1 - \alpha = 95\%$ resulta (0.63, 0.99).

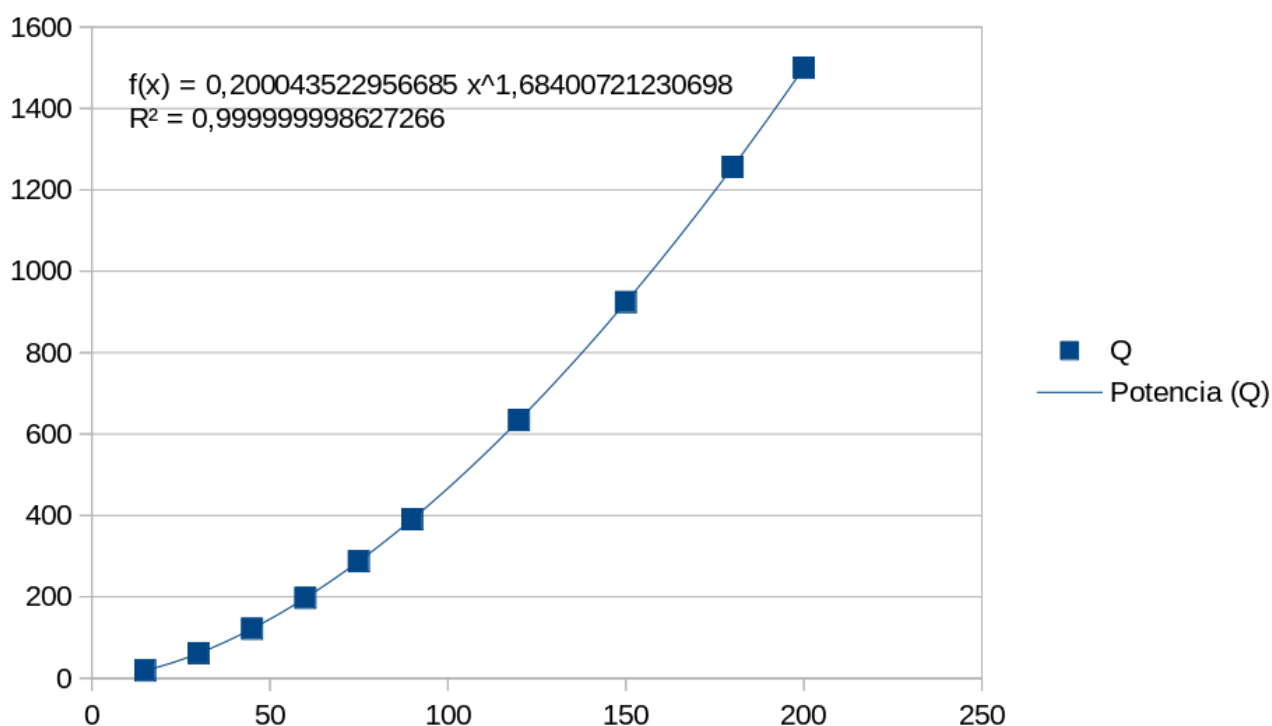
d) El valor de la esperanza estimada de y cuando $x=12.5$, es 24.96. Un intervalo del 95% de confianza para el parámetro en estudio resulta (6.66, 43.27). Observación: el valor de x para el cual se desea estimar no pertenece al rango de valores de x con el cual se realizó el ajuste. Luego, los valores estimados arriba son válidos si el modelo es válido para valores de x tan pequeños como 12.5.

Ejercicio 3

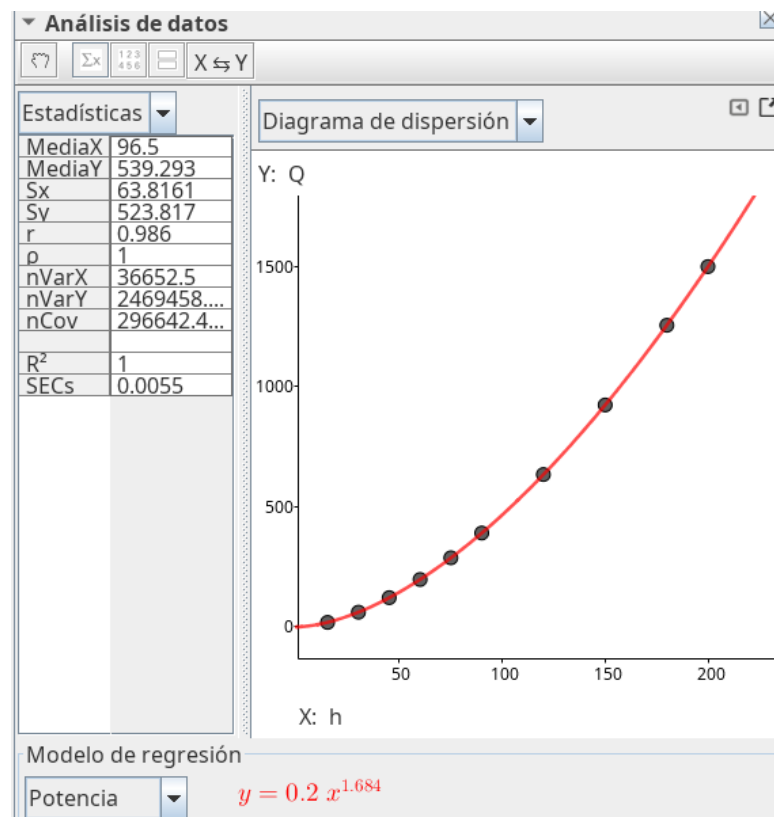
A partir del uso del software GNUMERIC puede encontrarse una relación logarítmica en el dispersiograma:



De manera similar, con el uso del software LibreOffice Calc se obtiene la siguiente función en el dispersiograma:



A su vez, utilizando el software Geogebra puede obtenerse, algo similar:

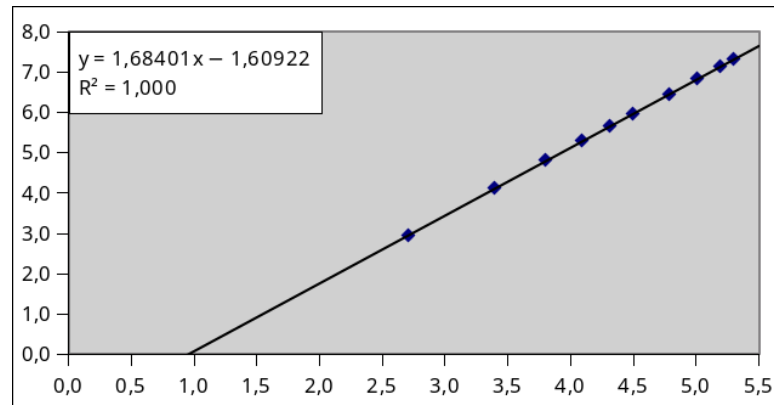


El coeficiente de determinación es igual a 1: todos los puntos están situados exactamente sobre la curva.

A su vez, si se estudia la relación lineal entre $\ln(Q)$ y $\ln(h)$, se puede construir la siguiente tabla:

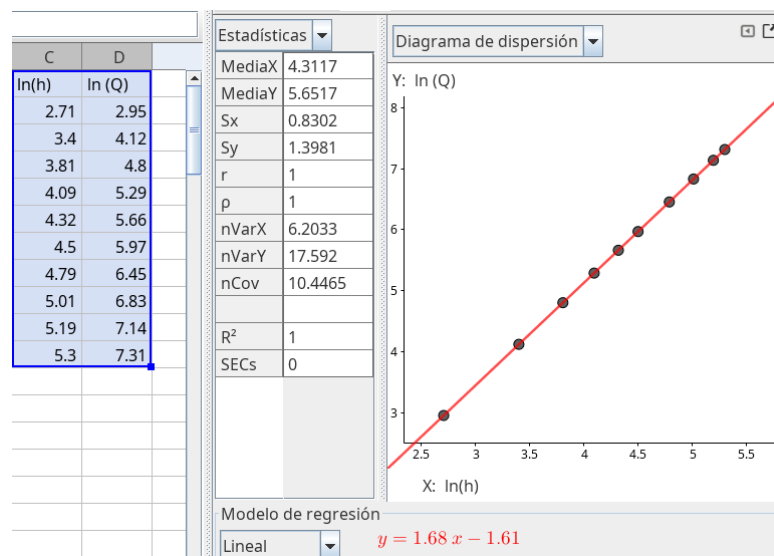
$\ln(h)$	$\ln(Q)$
2.7	3,0
3.4	4,1
3.8	4,8
4.1	5,3
4.3	5,7
4.5	6,0
4.8	6,5
5.0	6,8
5.2	7,1
5.3	7,3

A partir de lo anterior se puede obtener el siguiente dispersiograma:

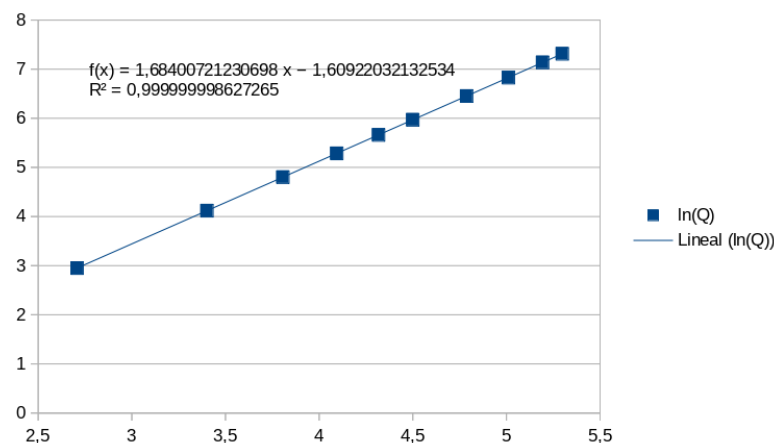


En ese caso la relación lineal es perfecta ya que posee un coeficiente de correlación igual a 1.

Lo mismo se observa usando Geogebra:



Algo parecido se obtiene usando LibreOffice Calc:

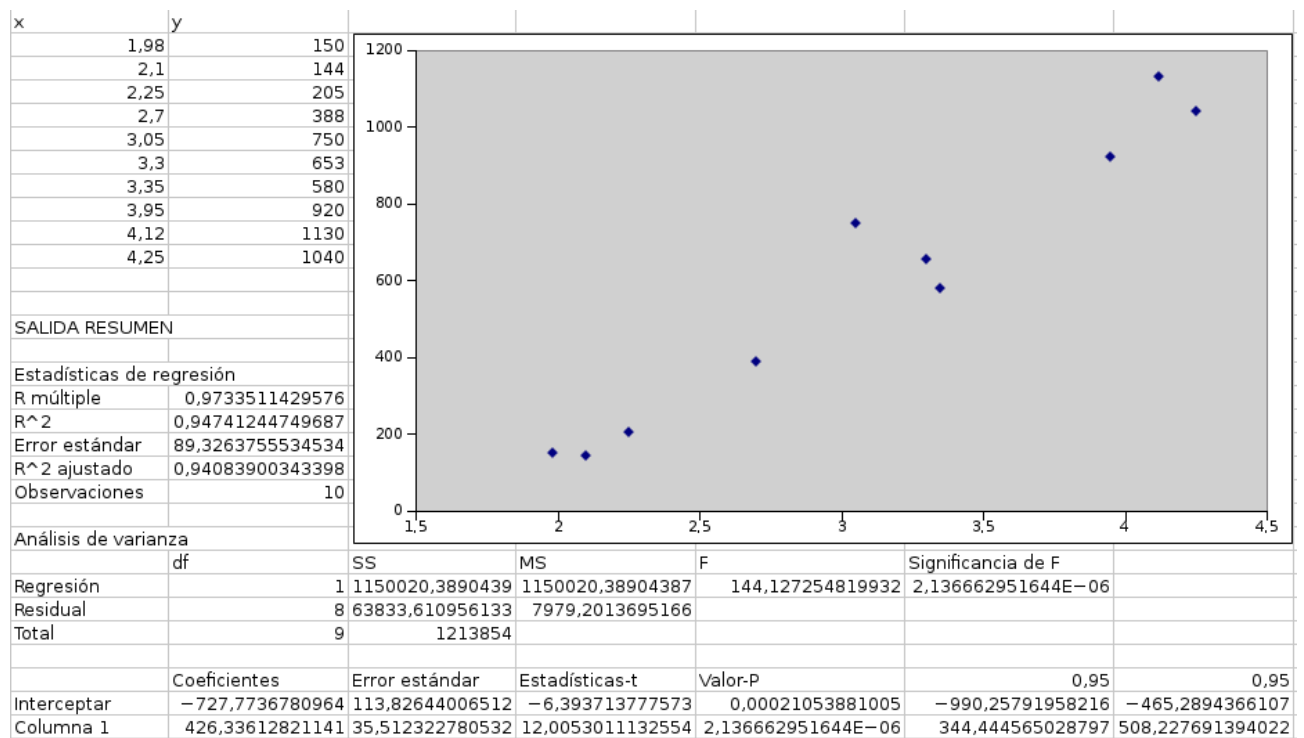


A continuación se presenta un fragmento del análisis de regresión:

Estadísticas de regresión						
R ²	1					
Error estándar	5,49421E-05					
Recuento de	1					
Observaciones	10					
R ² ajustado	1					
Análisis de varianza (ANOVA)						
	df	SS	MS	F	Significatividad F	
Regresión	1	17,5919538	17,5919538	5827783398	9,70967E-37	
Residuo	8	2,41491E-08	3,01864E-09			
Total	9	17,5919538				
Nivel de confianza	0,95					
	Coeficientes		Error estándar	Estadística t	Valor P	Inferior 95% Superior 95%
Intersección	-1,60922032	9,66869E-05	-16643,6161	1,90211E-31	-1,60944328	-1,60899736
ln(h)	1,68400721	2,20593E-05	76339,9201	9,70967E-37	1,68395634	1,68405808

b) El modelo de regresión permite estimar en forma puntual que si $\hat{Q} = \exp(1.68401 * \ln(400) - 1.60922) = 4819.70m^3/s$ cuando $h = 400cm$.

Ejercicio 4



a) Si se propone un modelo lineal se obtiene el ajuste: $\hat{y} = -727.774 + 426.336x$.

b) Para el caso del modelo polinomial de orden 2 de la se obtiene l función: $\hat{y} = -704.040 + 409.995x + 2.629x^2$.

Se puede obtener la salida del análisis regresión lineal múltiple tomando X y X^2 de Calc de la siguiente forma:

E	F	G	H	I	J	K
Regresión						
Modelo de r^2 Lineal						
Salida bruta de	ESTIMACION LINEAL			A	B	C
				x	x^2	y
2,62863379	409,994968	-704,040249		1,98	3,92	150
60,6934136	379,211221	561,333717		2,1	4,41	144
0,94742654	95,4811203	#N/D		2,25	5,06	205
63,0735086	7	#N/D		2,7	7,29	388
1150037,49	63816,5103	#N/D		3,05	9,30	750
				3,3	10,89	653
Estadísticas de regresión				3,35	11,22	580
R^2	0,94742654			3,95	15,60	920
Error estándar	95,4811203			4,12	16,97	1130
Recuento de	2			4,25	18,06	1040
Observaciones	10					
R^2 ajustado	0,93240555					
Análisis de varianza (ANOVA)						
	df	SS	MS	F	Significatividad F	
Regresión	2	1150037,49	575018,745	63,0735086	3,33183E-05	
Residuo	7	63816,5103	9116,64433			
Total	9	1213854				
Nivel de confianza	0,95					
	Coefficientes	Error estándar	Estadística t	Valor P	Inferior 95%	Superior 95%
Intersección	-704,040249	561,333717	-1,25422761	0,25001748	-2031,38357	623,303073
x	409,994968	379,211221	1,08117836	0,31545247	-486,697082	1306,68702
x^2	2,62863379	60,6934136	0,04331003	0,96666392	-140,888484	146,145751

c) Para el caso cuadrático se considera $r^2 \approx 0.95$. Para el caso lineal se toma $r \approx 0,97$.

d) Para el modelo lineal simple se tiene:

IC del 95% para α (intersección con el eje y): (-990.258, -465.289) IC del 95% para β (pendiente de la recta): (344.444, 508.228) Para el modelo que contiene el término cuadrático de la variable explicativa: IC del 95% para α (intersección con el eje y): (-2031.383, 623.303) IC del 95% para β_1 (coeficiente del término lineal): (-486.697, 1306.687) IC del 95% para β_2 (coeficiente del término cuadrático): (-140.888, 146.146)

e) El coeficiente correspondiente al término cuadrático no resulta significativamente distinto de 0 (valor p = 0.967). Además se tiene que para ambos modelos $r^2 = 0.947$, lo que indica que la incorporación del término cuadrático no aporta ninguna información para explicar la variabilidad de y si ya incorporamos x al modelo. A su vez para el caso lineal corresponde evaluar el valor de $r = 0,973$ por lo que el ajuste lineal es levemente mejor que el cuadrático.

f) Interpretaciones para conversar en clase.

Media	0,00
Error estándar	26,63
Mediana	-26,31
Moda	#N/D
Desviación estándar	84,22
Varianza de la muestra	7092,62
Curtosis	1,37
Desviación	1,09
Rango	297,90
Mínimo	-120,45
Máximo	177,45
Suma	0,00
Cuenta	10,00

