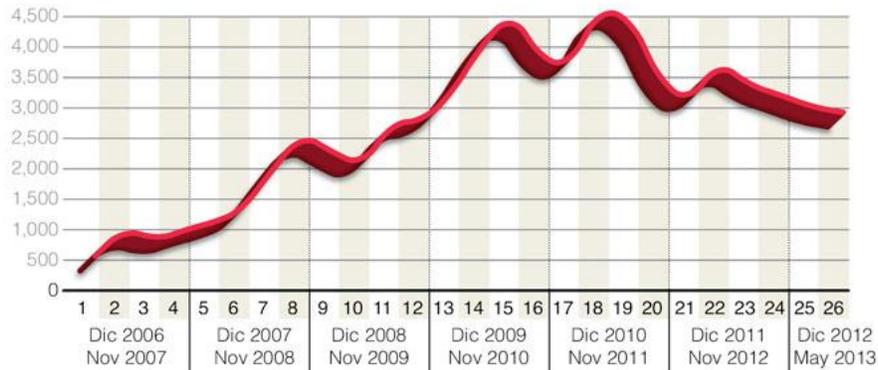


El cálculo de las tendencias criminales en la jurisdicción

Jesús Camacho.

Gráfica 1. Homicidios trimestrales
Vinculados con el crimen organizado



Al igual que en el análisis criminal táctico hacemos uso de la metodología del pronóstico del próximo delito para cortos periodos de tiempo, en el análisis criminal estratégico, se emplea la metodología del pronóstico, pero con unas técnicas diferentes que nos permiten en plazos de tiempo más largos, observar el comportamiento de las variables que intervienen y se emplean en el análisis criminal para efectos de la planificación policial y de la seguridad pública.

Así como en las empresas privadas se busca saber a futuro el comportamiento de las variables que están relacionadas con sus productos y ventas, para saber cómo conducir los negocios, en los cuerpos de policía también, pero enfocados en las variables que afectan a la seguridad ciudadana.

Pongamos como ejemplo, el negocio de la venta de caballos cuando Ford comenzó la producción en serie de su famoso modelo T e inundó el mercado norteamericano con ese novedoso medio de transporte. Seguramente que las líneas de tendencia en los negocios de las ventas y ganancias de caballos comenzaron a descender, mientras que las líneas de ventas y ganancias de las ventas de vehículos Ford comenzaron a aumentar. ¿Qué decisiones habrán tomado con estos datos en esas épocas los líderes de ambas empresas?

Y como todas las variables en el mundo real siempre están interconectadas unas con otras de manera sistémica (aunque a simple vista no lo observemos), imaginemos con otro ejemplo, como habrá sido el caso de los negocios de venta de alimentos de caballos y la de los negocios de venta de combustibles. ¿Qué decisiones habrán tomado los líderes del manejo de esos otros dos productos, al observar el comportamiento de las variables de sus respectivos negocios?

Algo similar buscamos en los cuerpos de policía con el pronóstico a largo plazo de las variables que están relacionadas con los delitos y la seguridad ciudadana. Pero en nuestro caso nos enfocamos más en variables como el número delitos cometidos en una jurisdicción, los arrestos de criminales, la cantidad de policías en labores de prevención e investigación, como tantas otras variables que intervienen en la actividad policial.

Ya hemos definido pronóstico en otro de nuestros artículos, pero lo repetiremos por conveniencia nuevamente. Se entiende por pronóstico en el diccionario Larousse: "Conjetura acerca de lo que puede suceder. Señal por donde se conjetura una cosa futura".

Al realizar un pronóstico, utilizamos datos históricos sobre el comportamiento de la variable a predecir, con la mayor exactitud posible, su comportamiento futuro. En este sentido, vemos que es necesario suponer que el comportamiento pasado del fenómeno que estamos analizando influirá, en el futuro, en la forma de comportarse del mismo.

Hay dos grandes grupos de modelos de pronóstico, ellos son: causales y series de tiempo (nosotros estaremos trabajando con los últimos en particular en el presente artículo). En los primeros, se intenta determinar una relación de causalidad entre las variables (de causa y efecto) es decir que, conociendo el comportamiento de una variable, se pueda predecir el valor de otra. En los pronósticos de series de tiempo, no se busca una relación entre diferentes variables, sino que, simplemente, se intenta predecir el comportamiento de la misma.

Es importante tener en cuenta que, en todo pronóstico, siempre está presente el componente de error. Esto quiere decir que nunca será posible realizar un pronóstico exacto sobre el comportamiento de una variable. El error es completamente impredecible e inevitable.

Para llevar adelante el presente artículo del pronóstico en el análisis criminal estratégico, revisaremos varios conceptos que siguen a continuación, antes de desarrollar nuestros ejemplos prácticos.

- Inferencia estadística.
- Regresión lineal.
- La ecuación de la recta en su forma simplificada o reducida.
- Forma como se interrelacionan las variables.
- Coeficiente de correlación.
- Intervalo de confianza.

Inferencia Estadística (Wikipedia): El estudio para obtener conclusiones generales para toda la población a partir del estudio de una muestra, y el grado de fiabilidad o significación de los resultados obtenidos.

Regresión lineal (Cristofoli): Procedimiento estadístico mediante el cual se intenta describir el comportamiento de una variable, llamada variable dependiente, en función de otra, llamada variable independiente. La relación entre ambas se efectúa mediante una fórmula que representa una función lineal.

Siendo X la variable independiente e Y la variable dependiente, el modelo de regresión poblacional quedaría determinado por la siguiente expresión:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Siendo: Y_i =variable a explicar.

X_i =variable explicativa.

B_0 =ordenada en el origen.

B_1 =pendiente (cambio que genera en Y_i cada unidad de X_i).

E_i =error (características no explicadas por el modelo planteado).

La Ecuación de la recta en su forma simplificada o reducida (Wikipedia): En geometría euclidiana, la recta o línea recta es una línea que se extiende en una misma dirección, por tanto, tiene una sola dimensión y contiene un número finito de puntos. Dicha recta también se puede describir como una sucesión continua de puntos extendidos en una sola dirección.

Para efectos de nuestros cálculos, estaremos trabajando con un modelo matemático más conocido por todos y visto en bachillerato, como lo es la recta (equivalente en este caso a la regresión lineal).

$$y = mx + b$$

Tipos de relaciones entre las variables (Gottlieb): Las variables se relacionan unas a otras de la forma como se explica a continuación.

1. Mientras una variable crece, la otra crece también. Por ejemplo, a medida que el consumo de alcohol aumenta en la población, la tasa de accidentes automovilísticos también crece. A medida que la tasa de la población crece, la tasa de delito también crece.
2. Mientras una variable decrece, la otra decrece también. Por ejemplo, a medida que el número de carros en las carreteras decrece, el número de accidentes automovilísticos también decrece. A medida que el número de cigarrillos consumido decrece, también decrecen las enfermedades respiratorias.
3. Mientras una variable crece, la otra decrece. Por ejemplo, a medida que las imposiciones por multas de tránsito aumentan, los accidentes de tránsito disminuyen. A medida que el número de niños en la familia aumenta, la cantidad de dinero disminuye.

Cuando dos variables crecen juntas (como en el caso 1) o decrecen juntas (como en el caso 2), se dice que las variables tienen una relación positiva. Esto es, tienen una relación positiva si mientras X crece e Y también crece (caso 1) o si a medida que X decrece Y también decrece (Caso 2).

Cuando una variable crece y la otra decrece (como en el caso 3), las variables tienen una relación negativa o inversa. Esto significa, que tienen una relación

negativa o inversa si a medida que X crece Y decrece, o la situación contraria, si X decrece Y crece.

Presentamos un cuadro con todas las relaciones posibles entre variables:

Posibilidad	Relación
X Crece Y Crece	Positiva
X decrece Y decrece	Positiva
X crece Y decrece	Negativa
X decrece Y crece	Negativa

De forma gráfica estas relaciones se observan así:



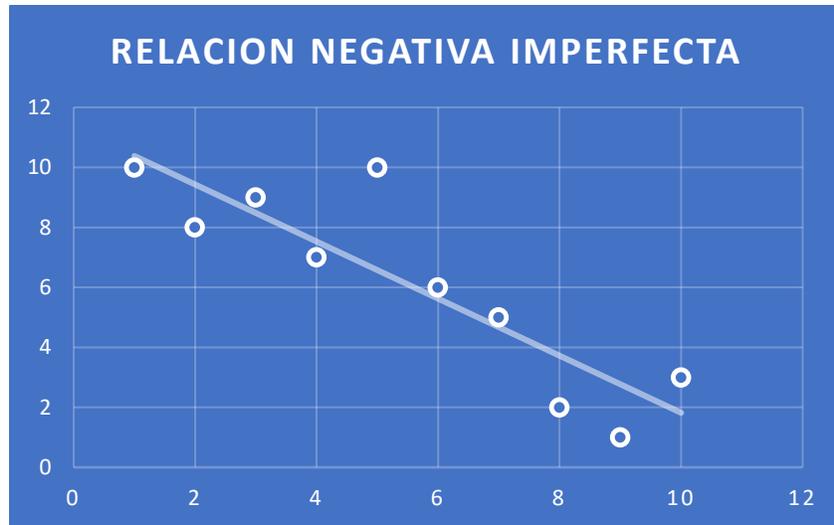
A medida que X se incrementa en una unidad, Y se incrementa en una unidad; o a medida que X decrece en una unidad, Y decrece en una unidad.



A medida que X crece en una unidad, Y decrece en una unidad; o a medida que X decrece en una unidad, Y crece en una unidad



X e Y crecen o decrecen, pero no perfectamente



Mientras X crece, Y decrece, pero no perfectamente



No existe relación entre X e Y

Coefficiente de correlación (Cristofoli): Se calcula de manera independiente de las unidades de medida de las variables. Este mide el grado de asociación lineal de las variables en estudio. Con respecto a sus dispersiones. El coeficiente de correlación mide el grado de relación entre las observaciones de las dos variables, independientemente de la unidad de medida utilizada.

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

Intervalo de Confianza (Cristofoli): Una vez establecido el valor puntual de pronóstico de un parámetro, se necesita conocer entre que valores es posible que se encuentre el parámetro poblacional que estamos analizando. Esta acción se lleva a cabo precisamente con un procedimiento llamado Estimación por Intervalos de Confianza.

Para construir el Intervalo de Confianza, debemos encontrar dos funciones de las observaciones muestrales; una para definir cuál es el límite superior y otra para el inferior entre los valores que se mueve el estimador del parámetro (precisando la probabilidad que tendrá ese intervalo de contenerlo).

$$P \left[\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] = 1 - \alpha$$

Haciendo un resumen de lo que debemos ejecutar para calcular nuestra recta de tendencia, tenemos:

1. Agrupamos los datos.
2. Se hacen los cálculos matemáticos de los datos de las variables en una hoja de cálculo (Excel).
3. Establecemos el coeficiente de correlación entre las dos variables, los delitos y el tiempo transcurrido.
4. Buscamos los valores de $(y= mx+b)$ a la que llamamos recta tendencia.
5. Calculamos "m" o la pendiente, que es la relación de variación de los delitos por años.
6. Calculamos "b" que es el punto donde la recta tendencia corta al eje Y.
7. Construimos la recta tendencia con dos puntos cualesquiera sobre el eje X.

8. Con el paso anterior cubierto, ya contamos con todos los datos necesarios para calcular el valor del delito "Y" para cualquier año, mes, o semana siguiente a nuestra última fecha "X". No es recomendable hacer pronósticos de más de dos fechas posteriores. En nuestro caso, vamos a buscar pronosticar dos años posteriores a la última fecha, eso quiere decir que sustituiremos el valor de X en la ecuación para saber el valor de Y en el año buscado.
9. Como nuestros datos no son exactos, reducimos los márgenes de error, aplicando la técnica de Intervalos de Confianza, para obtener certezas de 68% y 95% en nuestros resultados.

Es un hecho que las cárceles Latinoamericanas, tienen un grave problema con el tema del hacinamiento. En ese sentido, vamos a efectuar un ejemplo de pronóstico de arrestos para conocer en cuanto nos puede aumentar la población reclusa en un periodo determinado, para que las autoridades ya alertadas, sepan planificar con tiempo la ampliación o construcción de nuevos recintos penitenciarios. El presente es un ejercicio tomado del libro de Steve Gottlieb, Del primer reporte al arresto final.

Ejercicio:

El sistema carcelario de nuestra jurisdicción tiene una capacidad máxima de 7.000 plazas. El departamento de policía consiguió el arresto y reclusión de 5.500 delincuentes el año pasado. Con esta tasa de arrestos, es posible que las instalaciones físicas no van a tener la capacidad suficiente en los próximos años, para impedir que se presente un grave problema de hacinamiento. Sabemos que el gobernador del estado tendrá los recursos necesarios para ampliar los recintos carcelarios, si es necesaria la ampliación o la construcción de una nueva cárcel

Considerando los lapsos de tiempo requeridos para ampliar las cárceles o inclusive construir una nueva, el Comandante de la Policía quiere saber en cuantos años se alcanzará el máximo de la capacidad del sistema carcelario a partir de la presente fecha, en base a la cantidad de arrestos, para poder recomendarle al gobernador si tiene que pedir los recursos de inmediato para comenzar a construir un nuevo recinto penitenciario, o si puede emplear esos recursos en otras actividades también prioritarias, ya que cuenta con tiempo suficiente para esperar por nuevos recursos.

En tal sentido, contamos con la información de los arrestos en la jurisdicción en los últimos 8 años, para darle nuestra mejor recomendación al Comandante de la Policía.

Histórico de 8 años de arrestos en la jurisdicción.

A	B	C
	Años	Nro. de Arrestos
	1	1374
	2	2132
	3	2466
	4	2604
	5	3025
	6	3778
	7	5212
	8	5487
	9	?
	10	?

Necesitamos proyectar en que año se pueden alcanzar los 7.000 arrestos de delincuentes que lleven al máximo la capacidad del sistema carcelario de la jurisdicción.

Paso 1: Con la data agrupada elaboramos un gráfico para obtener alguna información inicial importante.



En el gráfico observamos que a medida el tiempo se incrementa, también se incrementan los arrestos. Más aun, pareciera existir una relación lineal entre las dos variables (años y número de arrestos). En tal sentido, continuamos con los cálculos de la recta o línea de tendencia.

Paso 2: Efectuamos los cálculos numéricos de la data de las 2 variables en una hoja de cálculo (Excel): Arrestos por año en un periodo de 8 años.

A	B	C	D	E	F	G	H	I	J
	X		Y		XY		x^2		y^2
	1		1374		1374		1		1887876
	2		2132		4264		4		4545424
	3		2466		7398		9		6081156
	4		2604		10416		16		6780816
	5		3025		15125		25		9150625
	6		3778		22668		36		14273284
	7		5212		36484		49		27164944
	8		5487		43896		64		30107169
$\Sigma X =$	36		$\Sigma Y =$ 26078		$\Sigma XY =$ 141625		$\Sigma x^2 =$ 204		$\Sigma y^2 =$ 99991294

Para efectuar el próximo paso que es el cálculo del coeficiente de correlación, debemos tomar en cuenta la tabla de valores que nos indica lo fuerte o débil que puede ser una relación entre 2 variables en estudio. Tomado de Timko John, Statistics for Math Haters, 1985.

Tabla de descripción de fortaleza de "r"

Valor	Tipo de relación
-1.00	Correlación negativa perfecta
-.800 a -.999	Correlación negativa fuerte
-.600 a -.799	Correlación negativa alta
-.400 a -.599	Correlación negativa moderada
-.200 a -.399	Correlación negativa débil
0.0 a -.199	Correlación negativa despreciable
0.0 a +.199	Correlación positiva despreciable
+.200 a +.399	Correlación positiva débil
+.400 a +.599	Correlación positiva moderada
+.600 a +.799	Correlación positiva alta
+.800 a +.999	Correlación positiva fuerte
+1.00	Correlación positiva perfecta

Paso 3: Cálculo del coeficiente de correlación (valor de "r"), de acuerdo a la ecuación:

$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}}$$

Donde:

r: es el coeficiente de correlación.

N: es el número de los pares de data (8 en nuestro caso).

Σ (sigma): significa sumatoria.

X: es la variable independiente (tiempo).

Y: es la variable dependiente (arrestos).

Insertamos los números de nuestra tabla de arrestos en un periodo de 8 años en la jurisdicción como sigue:

$$\begin{aligned} r &= \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{[N \sum X^2 - (\sum X)^2][N \sum Y^2 - (\sum Y)^2]}} \\ &= \frac{8(141,625) - 36(26,078)}{\sqrt{[8(204) - (36)^2][8(99,991,294) - (26,078)^2]}} \\ &= \frac{1,133,000 - 938,808}{\sqrt{[1632 - 1296][799,930,352 - 680,062,084]}} \\ &= \frac{194,192}{\sqrt{(336)(119,868,268)}} \\ &= \frac{194,192}{\sqrt{40,275,738,048}} \\ &= \frac{194,192}{200688.1612} \end{aligned}$$

$r = .96763$ o, redondeado a $.97$

El coeficiente de correlación es $.97$. Como es un valor positivo, nos indica que a medida que la variable tiempo (X) crece, los arrestos (Y) también crecen.

Observando la tabla de descripción de fortaleza de "r", podemos definir como es la fortaleza en la relación de las variables tiempo y arrestos. Un valor de r de .97 nos indica que la relación entre las variables es fuerte y positiva. En tal sentido, podemos emplear la ecuación y la data con confianza para pronosticar cuantos arrestos se pueden producir en que año.

Paso 4: Comenzamos a calcular los valores de la ecuación de la recta.

$$Y = mX + b$$

Donde:

Y (arrestos) es la variable dependiente.

m es la pendiente de la línea recta.

X (años) es la variable independiente.

b es el punto de corte de la línea recta en el eje Y.

Paso 4.1: Calculamos la pendiente m con su fórmula:

$$m = \frac{N(\sum XY) - (\sum X)(\sum Y)}{N(\sum X^2) - (\sum X)^2}$$

Donde:

N es el número de los pares de data (8 en nuestro caso).

Σ (sigma) significa sumatoria.

X es la variable independiente (tiempo en años).

Y es la variable dependiente (número de arrestos).

Aplicamos la fórmula de m con los datos de nuestra tabla de arrestos en un periodo de 8 años:

$$\begin{aligned}
 m &= \frac{N(\sum XY) - (\sum X)(Y)}{N(\sum X^2) - (\sum X)^2} \\
 &= \frac{8(141,625) - 36(26,078)}{8(204) - (36)^2} \\
 &= \frac{1,133,000 - 938,808}{1632 - 1296} \\
 &= \frac{194,192}{336}
 \end{aligned}$$

$$m = 577.952381 \text{ o, redondeado } 577.95$$

La pendiente nos indica que por cada unidad que se incrementa X (tiempo), hay una unidad de incremento de 577.95 en Y (arrestos).

Paso 4.2: Calculamos el punto de corte b de la línea recta en el eje Y con su fórmula:

$$b = \frac{\sum Y - m(\sum X)}{N}$$

Donde:

b es el punto de corte o de intersección de la línea recta sobre el eje Y.

m es la pendiente de la línea recta.

X es la variable independiente (tiempo).

Y es la variable dependiente (arrestos).

N es el número de los pares de data (8 en nuestro caso).

Procedemos a introducir los datos de nuestra tabla de arrestos de 8 años en la fórmula de b:

$$\begin{aligned}
 b &= \frac{\sum Y - m(\sum X)}{N} \\
 &= \frac{26,078 - 577.952381(36)}{8} \\
 &= \frac{26,078 - 20,806.28572}{8} \\
 &= \frac{5271.714284}{8}
 \end{aligned}$$

$$b = 658.964285 \text{ o, redondeado a } 659$$

El número 659, es el punto sobre el eje Y donde la línea recta lo corta (cruza).

Paso 4.3: Como contamos con la pendiente m y el punto de corte b , procedemos a calcular nuestra recta tendencia mediante la resolución de la ecuación de regresión lineal:

$$Y = mX + b$$

$$Y = 577.95X + 658.96$$

Paso 5: Tomamos dos números arbitrarios en el eje X y los sustituimos en la ecuación, tenemos los puntos necesarios para trazar nuestra recta tendencia.

Vamos a tomar como puntos arbitrarios en el eje X a $X=0$ y $X=10$. Procedemos a sustituirlos en nuestra recta tendencia $Y = 577,95X + 658,96$, entonces:

Primer punto A:

$$Y = 577,95(0) + 658,96$$

$$Y = 658,96$$

Entonces nuestro primer punto corresponde a $X=0$ e $Y=658,96$ al que nombramos punto A (0,658.96).

Segundo punto B:

$$Y = 577,95(10) + 658,96$$

$$Y = 5779,5 + 658,96$$

$$Y = 6438,46$$

Entonces nuestro segundo punto corresponde a $X=10$ e $Y=6438,46$ al que nombramos B (10,6438.46). Unimos los dos puntos A y B y obtenemos la gráfica de nuestra recta tendencia como sigue:



Paso 6: Procedemos a calcular el pronóstico de arrestos para el año número 9. Esto lo conseguimos sustituyendo en la X de la ecuación de la recta tendencia con el valor de $X=9$.

$$Y = mX + b$$

$$Y = 577,952381 (9) + 658,964285$$

$$Y = 5201,571429 + 658,964285$$

$$Y = 5860,535715 \text{ o, redondeado a } 5861$$

El pronóstico de arrestos para el año 9: Con base a la tendencia que se estableció en los arrestos de los 8 años previos, el cuerpo de policía debe esperar hacer 5.861 arrestos en el año 9.

Paso 7: Procedemos a calcular el pronóstico de arrestos para el año 10. Esto lo conseguimos substituyendo en la X de la ecuación de la recta tendencia con el valor de X = 10.

$$Y = mX + b$$

$$Y = 577,952381 (10) + 658,964285$$

$$Y = 5779,52381 + 658,964285$$

$$Y = 6438,488095 \text{ o, redondeado a } 6438$$

El pronóstico de arrestos para el año 10: Con base a la tendencia que se estableció en los arrestos de los 8 años previos, el cuerpo de policía debe esperar hacer 6.438 arrestos en el año 10.

Debemos estar conscientes de que es imposible pronosticar un número en particular con total exactitud. Y esto es obvio, ya que todos los pronósticos son solo estimaciones y toda estimación tiene algún margen de error. Es muy improbable que el cuerpo de policía alcance exactamente los 5.861 arrestos que obtuvimos como pronóstico para el año 9 o, los 6.438 arrestos que obtuvimos de pronóstico para el año 10. Por lo tanto, nunca damos como pronóstico definitivo estos números. Debemos calcular el error estándar (el probable rango de arrestos en cada año) de cada estimación. Para ellos debemos tomar nuevamente como data fuente nuestro cuadro original (paso # 2) como los mostramos a continuación:

A	B	C	D	E	F	G	H	I	J
	X		Y		XY		x ²		y ²
	1		1374		1374		1		1887876
	2		2132		4264		4		4545424
	3		2466		7398		9		6081156
	4		2604		10416		16		6780816
	5		3025		15125		25		9150625
	6		3778		22668		36		14273284
	7		5212		36484		49		27164944
	8		5487		43896		64		30107169
	ΣX=		ΣY=		ΣXY=		Σx ² =		Σ y ² =
	36		26078		141625		204		99991294

Paso # 8: Procedemos a calcular el error estándar con la aplicación de la fórmula:

$$S_e = \sqrt{\frac{\sum Y^2 - b \sum Y - m \sum XY}{N - 2}}$$

Donde:

S_e : es la designación para el error estándar.

X: es la variable independiente (tiempo).

Y: es la variable dependiente (arrestos).

b: es la intersección (punto de corte) de la recta tendencia con el eje Y.

m: es la pendiente de la recta tendencia.

N: es el número de pares de datos (8 en nuestro caso).

Procedemos a efectuar los cálculos:

$$\begin{aligned} S_e &= \sqrt{\frac{\sum Y^2 - b \sum Y - m \sum XY}{N - 2}} \\ &= \sqrt{\frac{134,342,615 - 658.8611113(31,939) - 577.9833333(194,374)}{9 - 2}} \\ &= \sqrt{\frac{134,342,615 - 21,043,365.03 - 112,344,932.4}{7}} \\ &= \sqrt{\frac{954317.57}{7}} \\ &= \sqrt{136,331.0814} \end{aligned}$$

$$S_e = \pm 398,814 \text{ o, redondeado } \pm 399$$

El primer error estándar de la estimación nos da un 68% de probabilidad de certeza para el pronóstico. En este caso el cuerpo de policía debe estar esperando que se den 5.861 arrestos más o menos 399 para el final del año 9. Para calcular el Intervalo de Confianza de arrestos, le sumamos y le restamos 399 a los 5.861 arrestos que estimamos para el año 9, como explicamos a continuación:

$$5.861 (\text{arrestos pronosticados año 9}) + 399 = 6.260 \text{ arrestos}$$

y

$$5.861 (\text{arrestos pronosticados año 9}) - 399 = 5.462 \text{ arrestos}$$

Nuestro intervalo de confianza de arrestos para el año 9, está entre 5.462 y 6.260 arrestos.

Por lo tanto, podemos pronosticar que el cuerpo de policía va a alcanzar los 5.861 arrestos en el año 9, y con un error estándar, estimamos con un 68% de certeza que el rango de arrestos del año 9 puede estar entre 5.462 y 6.260 arrestos. Para alcanzar un 95% de certeza en el pronóstico, simplemente multiplicamos el primer error estándar por 2, es decir $2 \times (\pm 399) = \pm 798$, y se lo sumamos y restamos a 5.861 arrestos, de la siguiente forma:

$$5.861 (\text{arrestos pronosticados año 9}) + 798 = 6.659 \text{ arrestos}$$

y

$$5.861 (\text{arrestos pronosticados año 9}) - 798 = 5.063 \text{ arrestos}$$

Para calcular el probable rango de arrestos para el año 10, debemos seguir los siguientes pasos:

1. Insertamos el valor pronosticado de arrestos del año 9 en la tabla de arrestos por año de 8 años (paso 2) para convertirla en la tabla de arrestos por año de 9 años, y efectuamos todos los cálculos nuevamente, para pronosticar el año 10, como sigue:

	A	B	C	D	E	F	G	H	I	J
1		X		Y		XY		x ²		y ²
2		1		1374		1374		1		1887876
3		2		2132		4264		4		4545424
4		3		2466		7398		9		6081156
5		4		2604		10416		16		6780816
6		5		3025		15125		25		9150625
7		6		3778		22668		36		14273284
8		7		5212		36484		49		27164944
9		8		5487		43896		64		30107169
10		9		5861		52749		81		34351321
11	ΣX=	45	ΣY=	31939	ΣXY=	194374	Σ X ²	285	ΣY ²	134342615

2. Calculamos la nueva pendiente "m" de la recta tendencia para el pronóstico del año 10, como sigue:

$$m = \frac{N \sum(XY) - (\sum X)(\sum Y)}{N(X^2) - (X)^2}$$

Donde:

N es el número de pares de datos (9, en este caso ya que fue añadido el año 9).

Por lo tanto:

$$\begin{aligned}
 m &= \frac{N \sum(XY) - (\sum X)(\sum Y)}{N(X^2) - (X)^2} \\
 &= \frac{9(194,374) - (45)(31,939)}{9(285) - (45)^2} \\
 &= \frac{1,749,366 - 1,437,255}{2565 - 2025} \\
 &= \frac{312,111}{540} \\
 &= 577.9833333.
 \end{aligned}$$

3. Calculamos "b" que es el punto de intersección (corte) de la recta tendencia en el eje de las Y, como sigue:

$$b = \frac{\sum Y - m(\sum X)}{N}$$

Donde:

N es el número de pares de datos (9, en este caso ya que fue añadido el año 9).

Por lo tanto:

$$\begin{aligned} b &= \frac{\sum Y - m(\sum X)}{N} \\ &= \frac{31,939 - (577.9833333)(45)}{9} \\ &= \frac{31,939 - 26009.25}{9} \\ &= \frac{5929.750002}{9} \\ &= 658.8611113. \end{aligned}$$

4. Como ya tenemos la pendiente "m" y el punto de corte "b" en el eje Y, podemos usar los datos de la tabla de arrestos de 9 años, para calcular el error estándar de la estimación, como mostramos a continuación:

$$S_e = \sqrt{\frac{\sum Y^2 - b \sum Y - m \sum XY}{N - 2}}$$

Donde:

S_e : Es el error estándar

X es la variable independiente (tiempo)

Y es la variable dependiente (arrestos)

b es la intersección (punto de corte) de la recta tendencia con el eje Y

m es la pendiente de la recta tendencia

N es el número de pares de datos (9, en este caso ya que fue añadido el año 9)

Por lo tanto:

$$\begin{aligned} S_e &= \sqrt{\frac{\sum Y^2 - b \sum Y - m \sum XY}{N - 2}} \\ &= \sqrt{\frac{134,342,615 - 658.8611113(31,939) - 577.9833333(194,374)}{9 - 2}} \\ &= \sqrt{\frac{134,342,615 - 21,043,365.03 - 112,344,932.4}{7}} \\ &= \sqrt{\frac{954317.57}{7}} \\ &= \sqrt{136,331.0814} \end{aligned}$$

$$= \pm 369,2303907 \text{ o, redondeado a } \pm 369$$

Como explicamos en el caso anterior, el error estándar de la estimación nos brinda un 68% de probabilidad de certeza en el pronóstico. En este caso, el cuerpo debe esperar hacer 6.438 arrestos más o menos 369 para el año 10. Para calcular el rango de arrestos, le sumamos y le restamos 369 a los 6.438 arrestos que estimamos para el año 10, como explicamos a continuación:

$$6.438 \text{ (arrestos pronosticados año 10)} + 369 = 6.807 \text{ arrestos}$$

y

$$6.438 \text{ (arrestos pronosticados año 10)} - 369 = 6.069 \text{ arrestos}$$

Por lo tanto, podemos pronosticar que el cuerpo de policía puede hacer 6.438 para el final del año 10, con un error estándar para una estimación del 68% de certeza de que el rango de arrestos va a estar entre 6.069 y 6.807 arrestos.

Para alcanzar un 95% de probabilidad de certeza, simplemente multiplicamos el error estándar por 2, como sigue: $2 \times (\pm 369) = \pm 738$ y el resultado final queda de la siguiente forma:

$$6.438 \text{ (numero pronosticado de arrestos año 10)} + 738 = 7.716 \text{ arrestos}$$

y

6.438 (numero pronosticado de arrestos año 10) - 738 = 5.700 arrestos

Resumimos en un cuadro nuestros resultados del pronóstico de arrestos para los años 9 y 10 como sigue:

Años	Pronóstico	Pronóstico 68%	Pronóstico 95%
9	5.861	Entre 5.462 y 6.260	Entre 5.063 y 6.659
10	6.438	Entre 6.069 y 6.087	Entre 5.700 y 7.716

CONCLUSION:

Podemos indicarle al comandante de policía, que le recomiende al señor Gobernador de la jurisdicción que debe solicitar los recursos para construir o hacer remodelaciones importantes a la infraestructura física penitenciaria, ya que se estima con un 95% de probabilidad de certeza que para el año 10, se tengan más de 7.000 arrestos que pueden generar un hacinamiento importante de no efectuarse las construcciones.

Cálculos de las operaciones en Excel:

Para aquellos colegas que les resulte sencillo o mejor trabajar con Excel, procedemos a explicar cómo elaborar los cálculos de la recta tendencia de los arrestos en la jurisdicción en los años 9 y 10, con esta útil herramienta.

Paso # 1: Procedemos a agrupar nuestros datos de arrestos en 8 años.

	A	B	C
1			
2	Numero de arrestos en 8 años		
3			
4			
5		X	Y
6		1	1236,91
7		2	1814,86
8		3	2392,81
9		4	2970,76
10		5	3548,71
11		6	4126,66
12		7	4704,61
13		8	5282,56

Paso # 2. Calculamos el coeficiente de correlación:



	A	B	C	D	E	F	G	H
1								
2		Numero de arrestos en 10 años				Error Estandar de Estimacion		
3							Redondeado	
4	X (Años)	Y (Arrestos)			Año 9	398,814369	Se=399	
5	1	1374			Año 10	369,230385	Se=369	
6	2	2132						
7	3	2466						
8	4	2604						
9	5	3025						
10	6	3778						
11	7	5212						
12	8	5487						
13	9	5861						
14	10	6438						
15								
16		Coefficiente de Correlacion						
17								
18		r=						
19								
20								

Hacemos click en más funciones

Paso # 2.1 Calculamos el coeficiente de correlación



	A	B	C	D	E	F	G	H
1								
2		Numero de arrestos en 10 años				Error Estandar d		
3							Redondeado	
4	X (Años)	Y (Arrestos)			Año 9	398,814369	Se=399	
5	1	1374			Año 10	369,230385	Se=369	
6	2	2132						
7	3	2466						
8	4	2604						
9	5	3025						
10	6	3778						
11	7	5212						
12	8	5487						
13	9	5861						
14	10	6438						
15								
16		Coefficiente de Correlacion						
17								
18		r=						
19								
20								

Hacemos click en estadísticas

Paso # 2.2 Calculamos el coeficiente de correlación

The screenshot shows the Excel interface with the 'Biblioteca de funciones' (Function Library) pane open. The 'Estadísticas' (Statistics) category is expanded, and 'COEF.DE.CORREL' is selected. A red arrow points to this option. The spreadsheet shows data for 'Numero de arrestos en 10 años' and a formula for 'Error Estandar d'.

X (Años)	Y (Arrestos)
1	1374
2	2132
3	2466
4	2604
5	3025
6	3778
7	5212
8	5487
9	5861
10	6438

Hacemos click en coeficiente de correlación

Paso # 2.3 Calculamos el coeficiente de correlación

The screenshot shows the Excel interface with the 'Argumentos de función' (Function Arguments) dialog box open for the 'COEF.DE.CORREL' function. The dialog box is open, and the 'Matriz1' and 'Matriz2' fields are empty. The spreadsheet shows data for 'Numero de arrestos en 10 años' and a formula for 'Error Estandar de Estimacion'.

X (Años)	Y (Arrestos)
1	1374
2	2132
3	2466
4	2604
5	3025
6	3778
7	5212
8	5487
9	5861
10	6438

Cargamos los valores de X e Y

Paso # 2.4 Calculamos el coeficiente de correlación

Argumentos de función

COEF.DE.CORREL

Matriz1 A5:A12 = (1;2;3;4;5;6;7;8)

Matriz2 B5:B12 = (1374;2132;2466;2604;3025;3778;5212;5487;5861;6438)

Devuelve el coeficiente de correlación de dos conjuntos de datos.

Matriz1 es un rango de celdas de valores. Los valores deben ser números, nombres, matrices o referencias que contengan números.

Resultado de la fórmula = 0,967630571

[Ayuda sobre esta función](#)

X (Años)	Y (Arrestos)
1	1374
2	2132
3	2466
4	2604
5	3025
6	3778
7	5212
8	5487
9	5861
10	6438

Coeficiente de Correlacion

r=2;B5:B12

Le damos aceptar una vez cargados los valores de X e Y

Paso # 2.5 Calculamos el coeficiente de correlación

B18 =COEF.DE.CORREL(A5:A12;B5:B12)

Numero de arrestos en 10 años Error

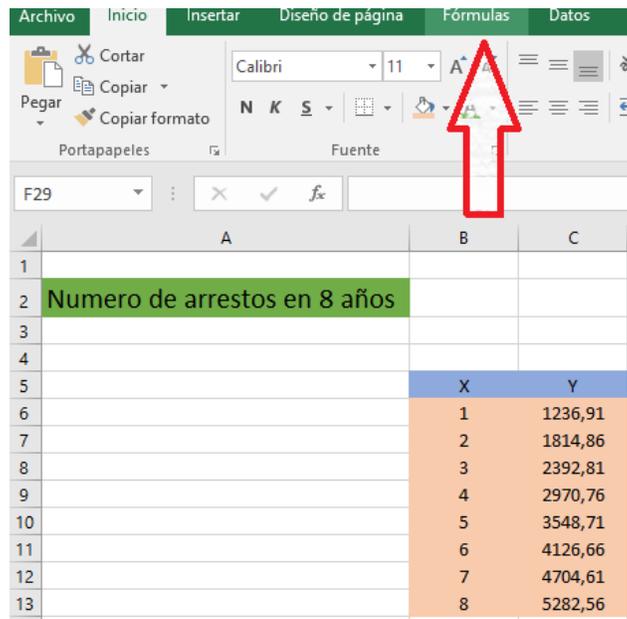
X (Años)	Y (Arrestos)
1	1374
2	2132
3	2466
4	2604
5	3025
6	3778
7	5212
8	5487
9	5861
10	6438

Coeficiente de Correlacion

r= 0,967630571

Obtenemos el valor del coeficiente de correlación $r = 0,96763057$

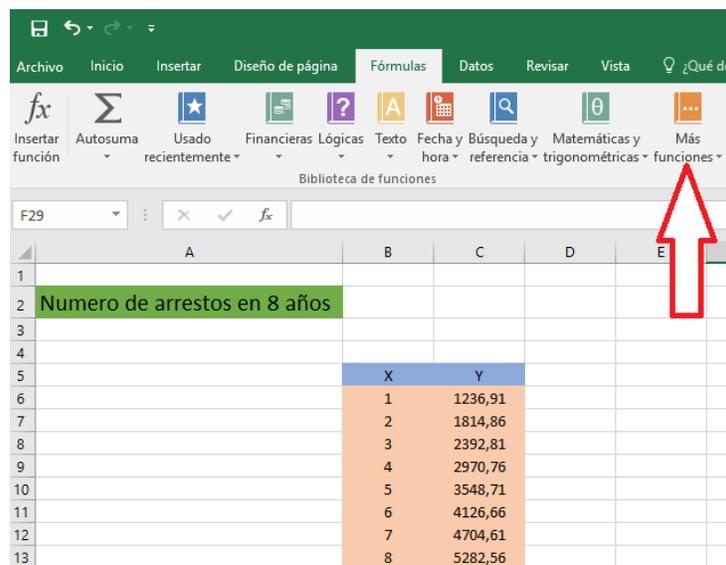
Paso # 3: Calculamos con la fórmula de tendencia la cantidad de arrestos para el año 9.



The screenshot shows the Microsoft Excel interface with the 'Fórmulas' ribbon selected. A red arrow points to the ribbon. The spreadsheet contains the following data:

	A	B	C
1			
2	Numero de arrestos en 8 años		
3			
4			
5		X	Y
6		1	1236,91
7		2	1814,86
8		3	2392,81
9		4	2970,76
10		5	3548,71
11		6	4126,66
12		7	4704,61
13		8	5282,56

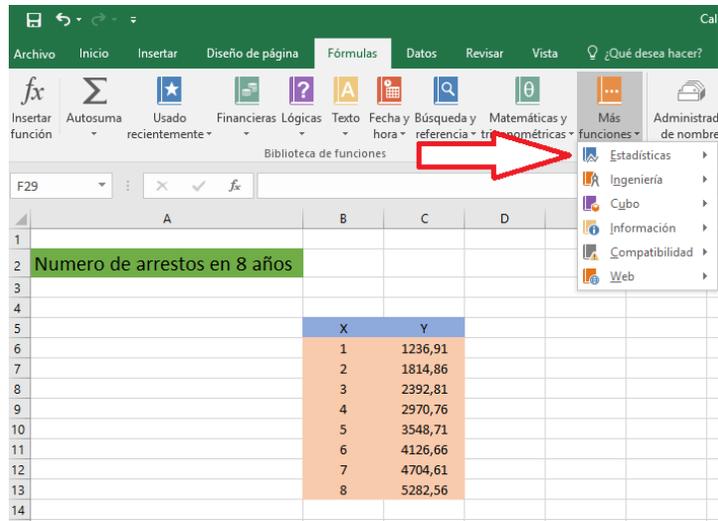
Hacemos click en el recuadro de formulas



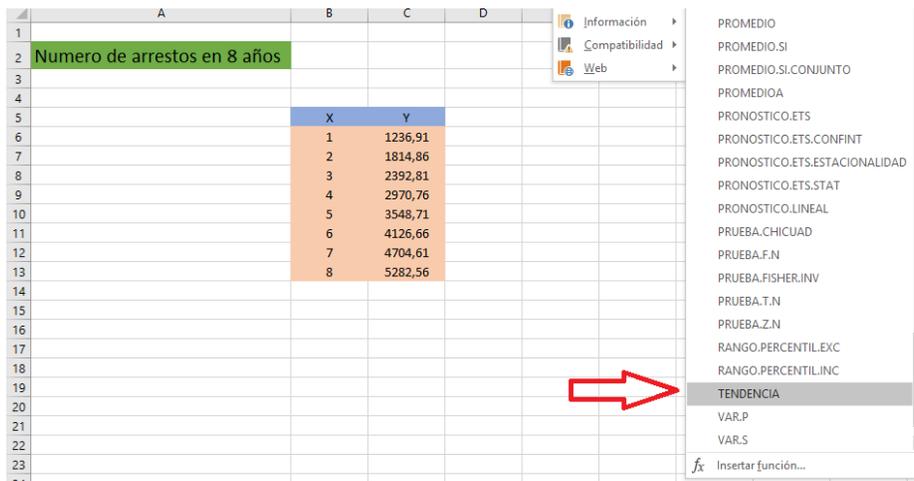
The screenshot shows the Microsoft Excel interface with the 'Fórmulas' ribbon selected. The 'Más funciones' button is highlighted with a red arrow. The spreadsheet contains the following data:

	A	B	C	D	E
1					
2	Numero de arrestos en 8 años				
3					
4					
5		X	Y		
6		1	1236,91		
7		2	1814,86		
8		3	2392,81		
9		4	2970,76		
10		5	3548,71		
11		6	4126,66		
12		7	4704,61		
13		8	5282,56		

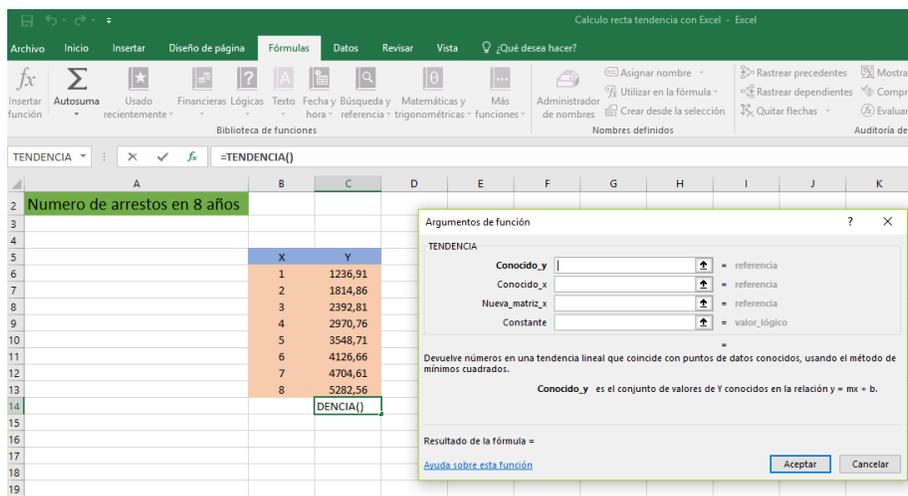
Hacemos click en más funciones



Hacemos click en estadísticas



Hacemos click en tendencia



Procedemos a llenar los datos. Los valores de las columnas X e Y, en el número de matriz colocamos el año 9 que estamos buscando.

Argumentos de función

TENDENCIA

Conocido_y C6:C13 = {1236,91;1814,86;2392,81;2970,76;3548,71;4126,66;4704,61;5282,56}

Conocido_x B6:B13 = {1;2;3;4;5;6;7;8}

Nueva_matriz_x 9 = {9}

Constante = valor_lógico = {5860,51}

Devuelve números en una tendencia lineal que coincide con puntos de datos conocidos, usando el método de mínimos cuadrados.

Nueva_matriz_x son nuevos valores de X para los cuales se desea que TENDENCIA devuelva los valores de Y correspondientes.

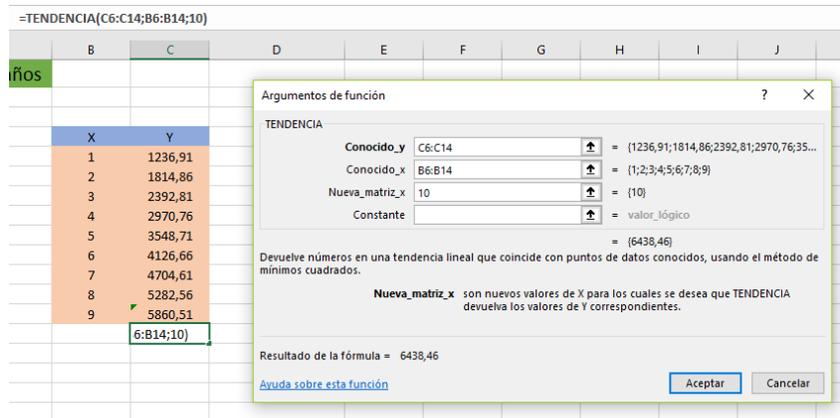
Resultado de la fórmula = 5860,51

[Ayuda sobre esta función](#)

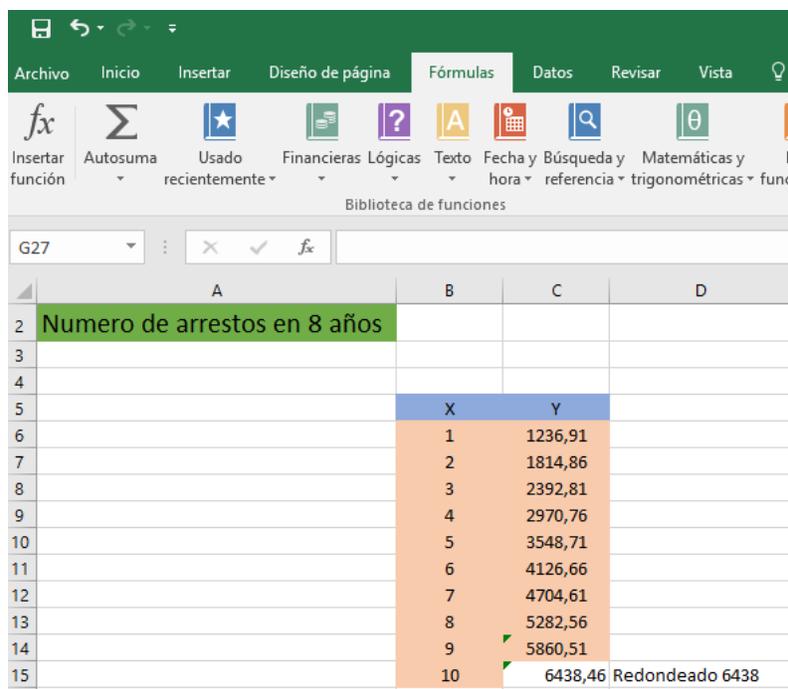
Le damos aceptar

	A	B	C	D
2	Numero de arrestos en 8 años			
3				
4				
5		X	Y	
6		1	1236,91	
7		2	1814,86	
8		3	2392,81	
9		4	2970,76	
10		5	3548,71	
11		6	4126,66	
12		7	4704,61	
13		8	5282,56	
14		9	5860,51	Redondeamos a 5861
15				

Obtenemos 5.860,51 arrestos para el año 9 y lo redondeamos a 5.861.



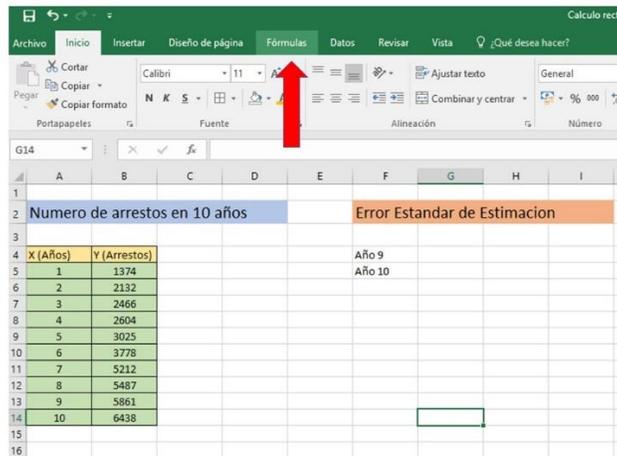
Con la data obtenida del año 9, llenamos nuevamente el cuadro de tendencia para calcular el año 10, damos Aceptar.



Obtenemos para el año 10 un total de arrestos de 6.438,46 que redondeamos a 6.438

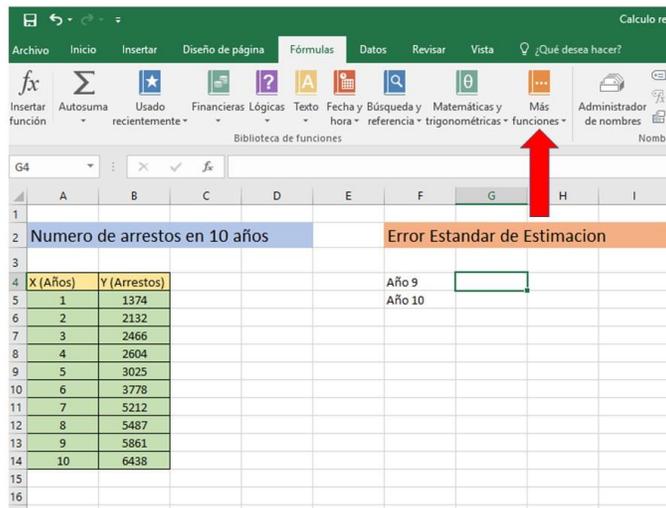
Paso # 4: Aplicamos la fórmula de Error Estándar de Estimación para los dos años de pronostico (9 y 10).

Paso # 4.1: Error estándar año 9



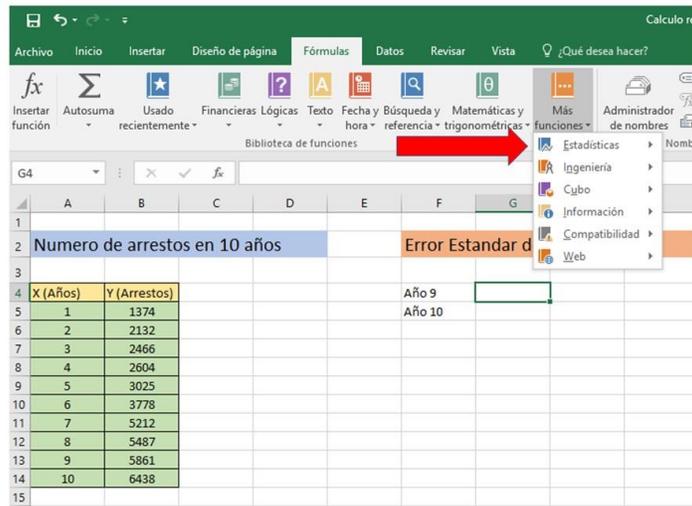
Hacemos clic en fórmulas

Paso # 4.2: Error estándar año 9



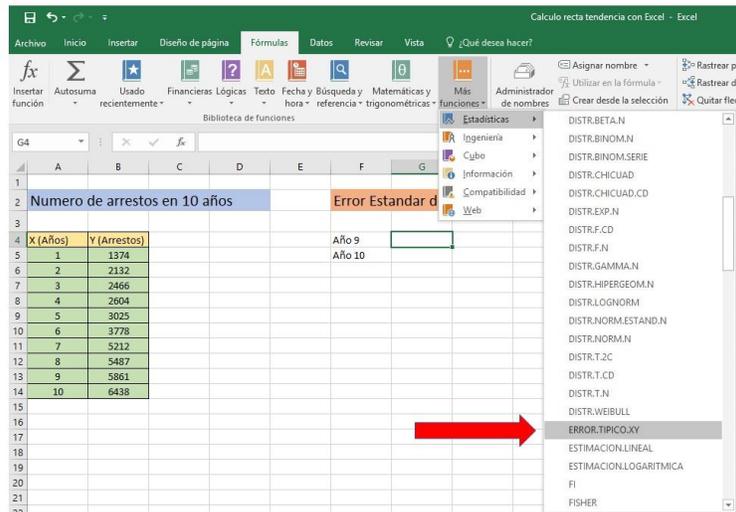
Hacemos clic en mas funciones

Paso # 4.3: Error estándar año 9



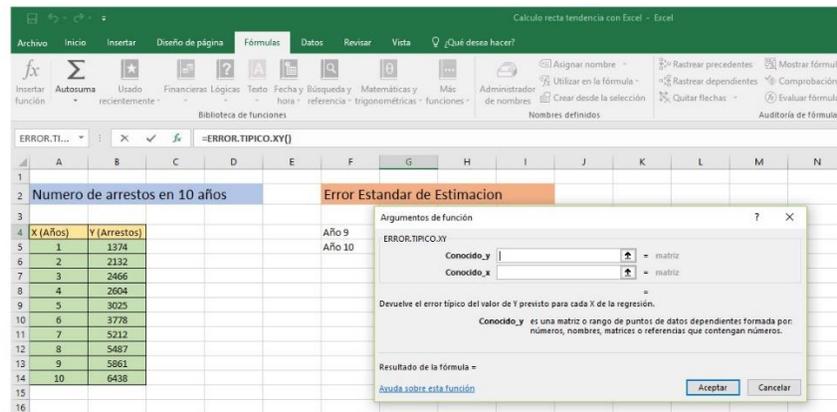
Hacemos click en Estadísticas

Paso # 4.4: Error estándar año 9



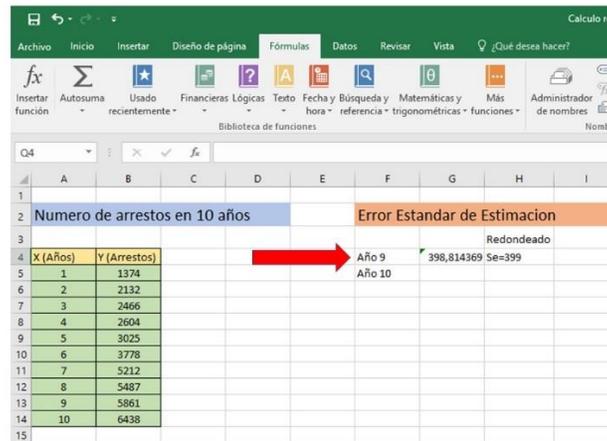
Hacemos click en Error Típico XY

Paso # 4.5: Error estándar año 9



Procedemos a llenar los valores de X e Y de los 8 años conocidos y damos aceptar

Paso # 4.6: Error estándar año 9



Obtenemos el error estándar de año 9, $S_e = \pm 398,814$ redondeado a ± 399

El primer error estándar de la estimación nos da un 68% de probabilidad de certeza para el pronóstico. En este caso el cuerpo de policía debe estar esperando que se den 5.861 arrestos más o menos 399 para el final del año 9. Para calcular el rango de arrestos, le sumamos y le restamos 399 a los 5.861 arrestos que estimamos para el año 9, como explicamos a continuación:

$$5.861 \text{ (arrestos pronosticados año 9)} + 399 = 6.260 \text{ arrestos}$$

y

$$5.861 \text{ (arrestos pronosticados año 9)} - 399 = 5.462 \text{ arrestos}$$

Nuestro intervalo de confianza de arrestos para el año 9, está entre 5.462 y 6.260 arrestos.

Por lo tanto, podemos pronosticar que el cuerpo de policía va a alcanzar los 5.861 arrestos en el año 9, y con un error estándar, estimamos con un 68% de certeza que el rango de arrestos del año 9 puede estar entre 5.462 y 6.260 arrestos. Para alcanzar un 95% de certeza en el pronóstico, simplemente multiplicamos el primer error estándar por 2, es decir $2 \times (\pm 399) = \pm 798$, y se lo sumamos y restamos a 5.861 arrestos, de la siguiente forma:

$$5.861 \text{ (arrestos pronosticados año 9)} + 798 = 6.659 \text{ arrestos}$$

y

$$5.861 \text{ (arrestos pronosticados año 9)} - 798 = 5.063 \text{ arrestos}$$

Para el cálculo del Error Estándar del año 10 aplicamos el mismo procedimiento del año 9, pero aplicamos colocando los datos obtenidos del año 9:

X (Años)	Y (Arrestos)
1	1374
2	2132
3	2466
4	2604
5	3025
6	3778
7	5212
8	5487
9	5861
10	6438

Argumentos de función

ERROR.TIPICO.XY

Conocido_y B5:B13 = (1374;2132;2466;2604;3025;3778;5212...

Conocido_x A5:A13 = (1;2;3;4;5;6;7;8;9)

= 369,2303847

Devuelve el error típico del valor de Y previsto para cada X de la regresión.

Conocido_y es una matriz o rango de puntos de datos dependientes formada por números, nombres, matrices o referencias que contengan números.

Resultado de la fórmula = 369,2303847

Aceptar Cancelar

Le damos aceptar

X (Años)	Y (Arrestos)	Año	Redondeado
1	1374	Año 9	398,814369 Se=399
2	2132	Año 10	369,230385 Se=369
3	2466		
4	2604		
5	3025		
6	3778		
7	5212		
8	5487		
9	5861		
10	6438		

Obtenemos el Error Estándar del año 10, $S_e = \pm 369,23$ redondeado a ± 369

Como explicamos en el caso anterior, el error estándar de la estimación nos brinda un 68% de probabilidad de certeza en el pronóstico. En este caso, el cuerpo debe esperar hacer 6.438 arrestos más o menos 369 para el año 10. Para calcular el rango de arrestos, le sumamos y le restamos 369 a los 6.438 arrestos que estimamos para el año 10, como explicamos a continuación:

$$6.438 \text{ (arrestos pronosticados año 10)} + 369 = 6.807 \text{ arrestos}$$

y

$$6.438 \text{ (arrestos pronosticados año 10)} - 369 = 6.069 \text{ arrestos}$$

Por lo tanto, podemos pronosticar que el cuerpo de policía puede hacer 6.438 para el final del año 10, con un error estándar para una estimación del 68% de certeza de que el rango de arrestos va a estar entre 6.069 y 6.807 arrestos.

Para alcanzar un 95% de probabilidad de certeza, simplemente multiplicamos el error estándar por 2, como sigue: $2 \times (\pm 369) = \pm 738$ y el resultado final queda de la siguiente forma:

$$6.438 \text{ (numero pronosticado de arrestos año 10)} + 738 = 7.176 \text{ arrestos}$$

y

$$6.438 \text{ (numero pronosticado de arrestos año 10)} - 738 = 5.700 \text{ arrestos}$$

Resumimos en un cuadro nuestros resultados del pronóstico de arrestos para los años 9 y 10 como sigue:

Años	Pronóstico	Pronóstico 68%	Pronóstico 95%
9	5.861	Entre 5.462 y 6.260	Entre 5.063 y 6.659
10	6.438	Entre 6.069 y 6.087	Entre 5.700 y 7.716

CONCLUSION:

Podemos indicarle al comandante de policía, que le recomiende al señor Gobernador de la jurisdicción que debe solicitar los recursos para construir o hacer remodelaciones importantes a la infraestructura física penitenciaria, ya que se estima con un 95% de probabilidad de certeza que para el año 10, se tengan más de 7.000 arrestos que pueden generar un hacinamiento importante de no efectuarse las construcciones.

BIBLIOGRAFIA: