

# Aproximación por Mínimos Cuadrados



**Materia:** Análisis Numérico

**Comisión:** K3053

**Integrante:** Erik Flores

**Fecha:** 7 de julio de 2025

## Descripción general del problema

El presente trabajo práctico tiene como objetivo aplicar el método de mínimos cuadrados para analizar y modelar la relación entre la edad de los jugadores y su puntaje de ranking dentro del sistema oficial de tenis de mesa en Argentina.

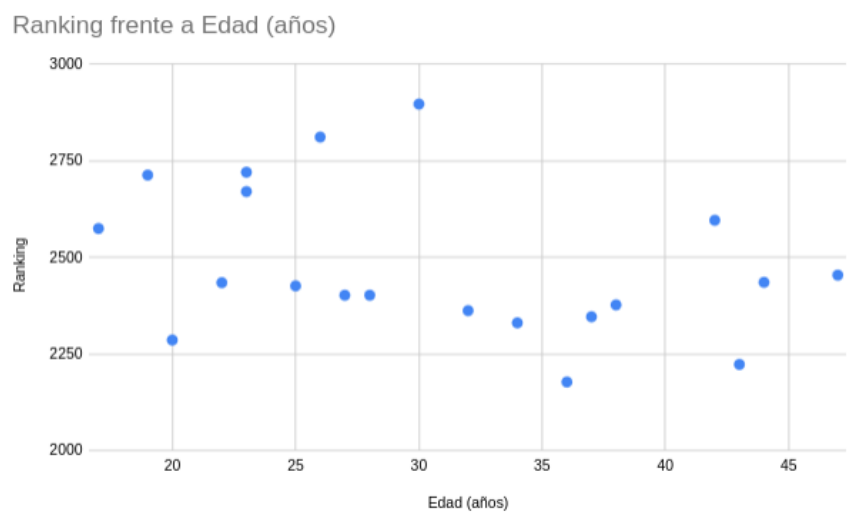
En particular, el estudio se centrará en los jugadores que actualmente compiten en Primera División, una de las categorías más altas del circuito nacional organizado por la plataforma *Tenis de Mesa Para Todos* (TMT). Esta base de datos pública y actualizada brinda información detallada sobre cada jugador federado, incluyendo su edad y el puntaje oficial asignado por el sistema de ranking TMT.

Dado que el puntaje de ranking varía semanalmente, se ha tomado como referencia la actualización correspondiente al día **7 de julio de 2025**, para garantizar la coherencia y consistencia de los datos analizados. A partir del listado oficial de 88 jugadores activos en Primera División, se ha seleccionado una muestra representativa de 20 jugadores, procurando abarcar distintas franjas etarias para garantizar diversidad en los datos.

La intención es construir un modelo matemático que permita estimar qué puntaje de ranking podría alcanzar un jugador en esta categoría en función de su edad, si se mantiene la tendencia observada en los datos actuales.

Este enfoque no solo permite explorar la dinámica entre edad y rendimiento competitivo en un deporte de alta exigencia como el tenis de mesa, sino también obtener una herramienta predictiva útil, por ejemplo, para entrenadores, jugadores jóvenes o incluso instituciones deportivas que quieran anticipar el potencial competitivo de sus atletas en base a su edad.

## Nube de puntos



## Modelos de Aproximación

Con el objetivo de analizar y modelar la relación entre la edad de los jugadores y su puntaje en el ranking, se aplicarán tres enfoques distintos de ajuste de datos: una aproximación lineal, una aproximación exponencial y una aproximación cuadrática.

Estos modelos permitirán evaluar cuál de ellos se ajusta mejor al comportamiento de los datos observados. Para llevar a cabo este análisis, se trabajará con el siguiente conjunto de datos correspondiente a una muestra de jugadores de Primera División:

Edad (años)	Ranking
19	2714
17	2575
22	2435
20	2287
26	2812
23	2721
23	2671
25	2427
30	2897
27	2403
28	2403
32	2363
34	2331
37	2347
38	2378
36	2178
44	2436
43	2224
42	2597
47	2455

## Aproximación lineal

Para una función lineal de la forma:  $f(x) = a \cdot x + b$   
se construye el sistema de ecuaciones normales utilizando las siguientes expresiones:

$$\begin{cases} a \cdot \sum_{i=1}^n x_i^2 + b \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n x_i f(x_i) \\ a \cdot \sum_{i=1}^n x_i + b \cdot n = \sum_{i=1}^n f(x_i) \end{cases}$$

Donde:

- $x$  representa la edad de cada jugador
- $y$  representa el puntaje en el ranking
- $n$  es la cantidad total de datos (en este caso,  $n = 20$ )

## Resolución del sistema por método matricial

Tenemos el sistema:

$$\begin{cases} a \cdot 20353 + b \cdot 613 = 1509285 \\ a \cdot 613 + b \cdot 20 = 49654 \end{cases}$$

### Paso 1: Representación matricial

$$A \cdot X = B$$

Donde:

$$A = \begin{bmatrix} 20353 & 613 \\ 613 & 20 \end{bmatrix}, \quad X = \begin{bmatrix} a \\ b \end{bmatrix}, \quad B = \begin{bmatrix} 1509285 \\ 49654 \end{bmatrix}$$

### Paso 2: Aplicamos la inversa de A

$$X = A^{-1} \cdot B$$

### Paso 3: Matriz inversa de A

$$A^{-1} = \begin{bmatrix} 0,000639 & -0,01959 \\ -0,01959 & 0,650443 \end{bmatrix}$$

### Paso 4: Multiplicación matricial

$$X = A^{-1} \cdot B = \begin{bmatrix} 0,000639 & -0,01959 \\ -0,01959 & 0,650443 \end{bmatrix} \cdot \begin{bmatrix} 1509285 \\ 49654 \end{bmatrix} = \begin{bmatrix} -8,0599 \\ 2729,74 \end{bmatrix}$$

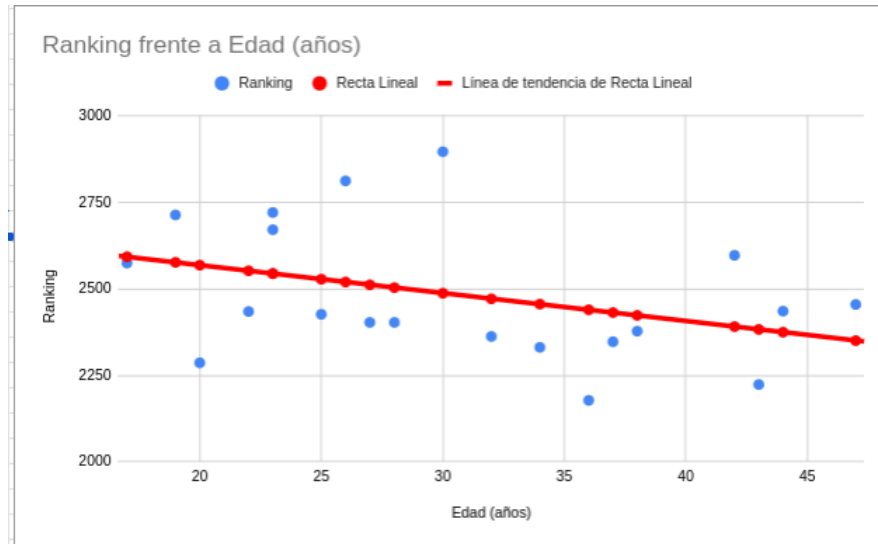
**Función aproximante final**

$$f(x) = -8,06 \cdot x + 2729,74$$

A continuación, se presenta la tabla con los cálculos necesarios que se usaron para resolver el sistema de ecuaciones.

Edad (x)	Ranking (y)	$x^2$	$x \cdot y$
19	2714	361	51566
17	2575	289	43775
22	2435	484	53570
20	2287	400	45740
26	2812	676	73112
23	2721	529	62583
23	2671	529	61433
25	2427	625	60675
30	2897	900	86910
27	2403	729	64881
28	2403	784	67284
32	2363	1024	75616
34	2331	1156	79254
37	2347	1369	86839
38	2378	1444	90364
36	2178	1296	78408
44	2436	1936	107184
43	2224	1849	95632
42	2597	1764	109074
47	2455	2209	115385
$\sum x = 613$	$\sum y = 49654$	$\sum x^2 = 20353$	$\sum xy = 1509285$

Ahora se presenta el gráfico de la recta obtenida superpuesta sobre la nube de puntos original, lo cual permite visualizar el ajuste lineal al conjunto de datos:



## Modelo exponencial

Queremos ajustar un modelo del tipo:

$$f(x) = b \cdot e^{a \cdot x}$$

Aplicando logaritmo natural a ambos lados:

$$\ln(y) = \ln(b) + a \cdot x$$

Definimos:

$$Y = \ln(y), \quad A = a, \quad B = \ln(b)$$

Entonces, el modelo queda linealizado:

$$Y = A \cdot x + B$$

### Paso 1: Cálculo de valores transformados

A continuación, se presenta la tabla con los valores de  $x$ ,  $y$ ,  $\ln(y)$ ,  $x \cdot \ln(y)$  y  $x^2$ , utilizados para la estimación por mínimos cuadrados:

Edad (x)	Ranking (y)	ln(y)	x·ln(y)	x <sup>2</sup>
19	2714	7.9062	150.2174	361
17	2575	7.8536	133.5113	289
22	2435	7.7977	171.5494	484
20	2287	7.7350	154.6999	400
26	2812	7.9417	206.4829	676
23	2721	7.9088	181.9014	529
23	2671	7.8902	181.4748	529
25	2427	7.7944	194.8603	625
30	2897	7.9714	239.1429	900
27	2403	7.7845	210.1808	729
28	2403	7.7845	217.9653	784
32	2363	7.7677	248.5660	1024
34	2331	7.7541	263.6378	1156
37	2347	7.7609	287.1530	1369
38	2378	7.7740	295.4126	1444
36	2178	7.6862	276.7018	1296
44	2436	7.7981	343.1170	1936
43	2224	7.7071	331.4037	1849
42	2597	7.8621	330.2087	1764
47	2455	7.8059	366.8765	2209
<b>613</b>	<b>49654</b>	<b>156.2839</b>	<b>4785.0634</b>	<b>20353</b>

## Resolvemos el sistema de ecuación Lineal

Reemplazando con los valores obtenidos:

$$\begin{cases} a \cdot 20353 + b \cdot 613 = 4785,063 \\ a \cdot 613 + b \cdot 20 = 156,284 \end{cases}$$

### Paso 2: Método matricial

$$A \cdot X = B$$

$$A = \begin{bmatrix} 20353 & 613 \\ 613 & 20 \end{bmatrix}, \quad X = \begin{bmatrix} a \\ b \end{bmatrix}, \quad B = \begin{bmatrix} 4785,063 \\ 156,284 \end{bmatrix}$$

Inversa de A :

$$A^{-1} \approx \begin{bmatrix} 0,000639 & -0,01959 \\ -0,01959 & 0,65044 \end{bmatrix}$$

**Multiplicación**

$$X = A^{-1} \cdot B = \begin{bmatrix} 0,000639 & -0,01959 \\ -0,01959 & 0,65044 \end{bmatrix} \cdot \begin{bmatrix} 4785,063 \\ 156,284 \end{bmatrix} \approx \begin{bmatrix} -0,00415 \\ 7,3538 \end{bmatrix}$$

**Paso 3: Volvemos a la función exponencial**

Sabemos que:

$$\ln(y) = a \cdot x + \ln(b) \Rightarrow y = b \cdot e^{a \cdot x}$$

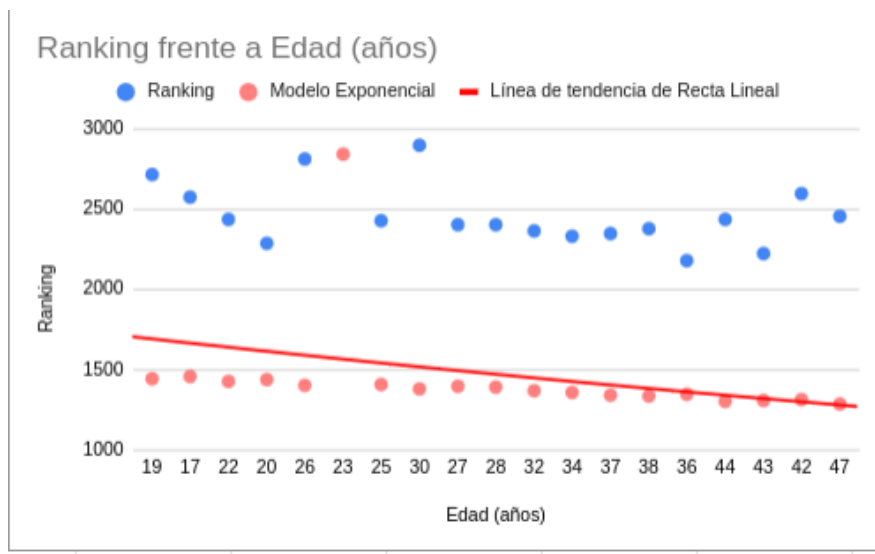
Ya tenemos:

$$a = -0,00415, \quad \ln(b) = 7,3538 \Rightarrow b = e^{7,3538} \approx 1563,82$$

**Función aproximante final (modelo exponencial):**

$$f(x) = 1563,82 \cdot e^{-0,00415 \cdot x}$$

**Visualización del modelo exponencial ajustado**





## Aproximación cuadrática

Queremos ajustar una función de la forma:

$$f(x) = a \cdot x^2 + b \cdot x + c$$

Para obtener los coeficientes  $a$ ,  $b$  y  $c$ , se utiliza el método de mínimos cuadrados, lo cual conduce a un sistema de tres ecuaciones con tres incógnitas basado en sumatorias:

$$\begin{cases} a \cdot \sum x^4 + b \cdot \sum x^3 + c \cdot \sum x^2 = \sum x^2 \cdot y \\ a \cdot \sum x^3 + b \cdot \sum x^2 + c \cdot \sum x = \sum x \cdot y \\ a \cdot \sum x^2 + b \cdot \sum x + c \cdot n = \sum y \end{cases}$$

### Paso 1: Tabla de cálculos para el sistema

Edad (x)	Ranking (y)	$x^2$	$x^3$	$x^4$	$x \cdot y$	$x^2 \cdot y$
19	2714	361	6859	130321	51566	979754
17	2575	289	4913	83521	43775	744175
22	2435	484	10648	234256	53570	1178540
20	2287	400	8000	160000	45740	914800
26	2812	676	17576	456976	73112	1900912
23	2721	529	12167	279841	62583	1439409
23	2671	529	12167	279841	61433	1412959
25	2427	625	15625	390625	60675	1516875
30	2897	900	27000	810000	86910	2607300
27	2403	729	19683	531441	64881	1751787
28	2403	784	21952	614656	67284	1883952
32	2363	1024	32768	1048576	75616	2419712
34	2331	1156	39304	1336336	79254	2694636
37	2347	1369	50653	1874161	86839	3213043
38	2378	1444	54872	2085136	90364	3433832
36	2178	1296	46656	1679616	78408	2822688
44	2436	1936	85184	3748096	107184	4716096
43	2224	1849	79507	3418801	95632	4112176
42	2597	1764	74088	3111696	109074	4581108
47	2455	2209	103823	4879681	115385	5423095
<b>613</b>	<b>49654</b>	<b>20353</b>	<b>723445</b>	<b>27153577</b>	<b>1509285</b>	<b>49746849</b>

### Paso 2: Sustituimos en el sistema de ecuaciones

$$\begin{cases} a \cdot 27153577 + b \cdot 723445 + c \cdot 20353 = 49746849 \\ a \cdot 723445 + b \cdot 20353 + c \cdot 613 = 1509285 \\ a \cdot 20353 + b \cdot 613 + c \cdot 20 = 49654 \end{cases}$$

### Paso 3: Representación matricial del sistema

$$A \cdot X = B$$

$$A = \begin{bmatrix} 27153577 & 723445 & 20353 \\ 723445 & 20353 & 613 \\ 20353 & 613 & 20 \end{bmatrix}, \quad X = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \quad B = \begin{bmatrix} 49746849 \\ 1509285 \\ 49654 \end{bmatrix}$$

### Solución del sistema (método matricial):

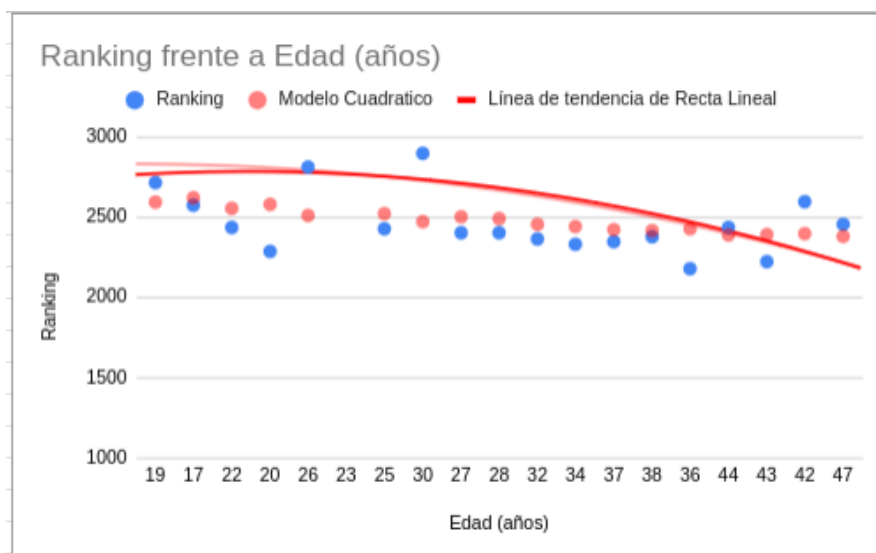
Los coeficientes obtenidos del modelo cuadrático ajustado son:

$$a \approx 0,19923, \quad b \approx -20,7462, \quad c \approx 2915,83$$

### Función aproximante cuadrático:

$$f(x) = 0,19923 \cdot x^2 - 20,7462 \cdot x + 2915,83$$

## Visualización del modelo cuadrático ajustado



# Comparación de modelos y elección de la mejor aproximación

A partir de los datos relevados de 20 jugadores de Primera División del sistema TMT, se ajustaron tres tipos de funciones: una recta (modelo lineal), una curva exponencial decreciente y una parábola (modelo cuadrático). La comparación entre ellas se basó en el análisis visual, el comportamiento matemático y la coherencia con la realidad del rendimiento en tenis de mesa.

## Modelo lineal

$$f(x) = -8,06 \cdot x + 2729,74$$

- Plantea que, a medida que aumenta la edad, el ranking disminuye de forma constante.
- Aunque es simple y fácil de interpretar, no representa bien la realidad del grupo analizado.
- En el gráfico se observa que la recta queda alejada de muchos puntos, especialmente en los extremos.
- En la práctica, el rendimiento no suele tener una relación estrictamente lineal con la edad.

## Modelo exponencial

$$f(x) = 1563,82 \cdot e^{-0,00415 \cdot x}$$

- Sugiere una caída progresiva del ranking con la edad, siguiendo una curva suave.
- Sin embargo, visualmente se ajusta peor que el modelo lineal en algunos tramos.
- La curva desciende demasiado rápido y no refleja bien la dispersión real de los datos.
- La lógica de “decadencia continua” no representa la variabilidad del rendimiento real.

## Modelo cuadrático

$$f(x) = 0,19923 \cdot x^2 - 20,7462 \cdot x + 2915,83$$

- Esta parábola se adapta mejor a la forma general de los datos observados.

- Su gráfico pasa más cerca de los puntos reales, reflejando mejor las tendencias.
- Representa una visión más realista: jugadores jóvenes con niveles intermedios y adultos con alto rendimiento.
- Visualmente y conceptualmente, es el modelo más representativo del grupo analizado.

### Conclusión general

Teniendo en cuenta los aspectos anteriores, se concluye que el **modelo cuadrático** es el más adecuado para aproximar el ranking en función de la edad. Su flexibilidad permite adaptarse a la forma real de los datos, con un ajuste visual claramente superior.

### Predicción con el modelo elegido

Como ejemplo de aplicación, estimamos el ranking de un jugador de 29 años usando el modelo cuadrático:

$$f(29) = 0,19923 \cdot 29^2 - 20,7462 \cdot 29 + 2915,83 \approx 2481,73$$

Por lo tanto, el ranking estimado es de aproximadamente **2482 puntos**, un valor coherente con los datos observados. Aunque se trata de una estimación, puede utilizarse como una referencia útil para entrenadores o jugadores que deseen prever el rendimiento competitivo en base a la edad.