

Análisis de variables y su influencia en la fuerza de juego de los jugadores de ajedrez registrados en la FIDE

José Sepúlveda Pareja
Proyecto de Curso “Data Analytics”
CoderHouse

ÍNDICE

1. INTRODUCCIÓN	3
2. DESCRIPCIÓN DE LA TEMÁTICA DE LOS DATOS.....	4
3. HIPÓTESIS.....	4
4. OBJETIVO DEL PROYECTO.....	4
5. ALCANCE	4
6. USUARIO FINAL Y NIVEL DE APLICACIÓN DEL ANÁLISIS	4
7. DIAGRAMA ENTIDAD - RELACIÓN.....	5
8. LISTADO DE TABLAS.....	5
9. LISTADO DE COLUMNAS POR TABLA	7
10. TRANSFORMACIONES REALIZADAS.....	9
10.1. Modificación en la tabla Jugadores	9
10.2. Reemplazar Valores	11
11. MEDIDAS CALCULADAS	11
12. MODELO DE DATOS EN POWER BI.....	13

1. INTRODUCCIÓN

El ajedrez es un juego de estrategia que se juega entre dos jugadores en un tablero de 64 casillas, donde cada jugador tiene 16 piezas que pueden moverse de diferentes maneras con el objetivo de capturar al rey del oponente. A lo largo de la historia, el ajedrez se ha convertido en uno de los juegos más populares y respetados en todo el mundo.

La Federación Internacional de Ajedrez (FIDE) es el organismo oficial encargado de regular las competiciones de ajedrez de todo el mundo. Si bien existen varias federaciones la FIDE es la más importante.

Su nombre proviene del francés *Fédération internationale des échecs*. Se fundó en París el 20 de julio de 1924 , coincidiendo curiosamente con el último día de juego de la primera Olimpiada de ajedrez. Es por eso que en esa fecha se celebra anualmente el Día Internacional del Ajedrez.¹

Para medir la fuerza de juego se utiliza el sistema de puntuación ELO, que es un método matemático utilizado para calcular la habilidad relativa de los jugadores de ajedrez en comparación con otros. El sistema fue creado por Arpad Elo, un profesor de física estadounidense y un maestro de ajedrez que centró sus esfuerzos en mejorar la forma en que la Federación Estadounidense de Ajedrez medía el nivel de juego de los ajedrecistas.²

¹ Federación Internacional de Ajedrez (FIDE) - Chess.com

² Sistema de puntuación Elo - Términos de ajedrez - Chess.com

2. DESCRIPCIÓN DE LA TEMÁTICA DE LOS DATOS

Para este análisis, se utilizará la base de datos proporcionada por la FIDE actualizada al 31 de marzo de 2023³. Donde se detalla el nombre de los jugadores, federación que representan, fecha de nacimiento, títulos y su puntaje ELO por modalidad de juego, entre otros datos.

Las modalidades de juego pueden ser Blitz, Rápido o Clásico y esto depende de la duración en el reloj de cada partida (siendo las Blitz las de menor tiempo y el tipo Clásico las de mayor duración).

3. HIPÓTESIS

Comprobar si los jugadores de ajedrez alcanzan su máximo potencial entre los 30 y 40 años de edad. Además, se analizarán otros factores como género o nacionalidad y si tienen correlación con la fuerza de juego en la fuerza de juego.

4. OBJETIVO DEL PROYECTO

Comprender la relación entre la edad y la fuerza de juego en los ajedrecistas y evaluar si existen otras variables que influyan de manera significativa.

5. ALCANCE

En este proyecto se analizará el universo total de ajedrecistas registrados en la FIDE al 31 de marzo de 2023. Se trabajará con los datos durante un período de 3 meses y los resultados serán presentados con la herramienta Power BI.

6. USUARIO FINAL Y NIVEL DE APLICACIÓN DEL ANÁLISIS

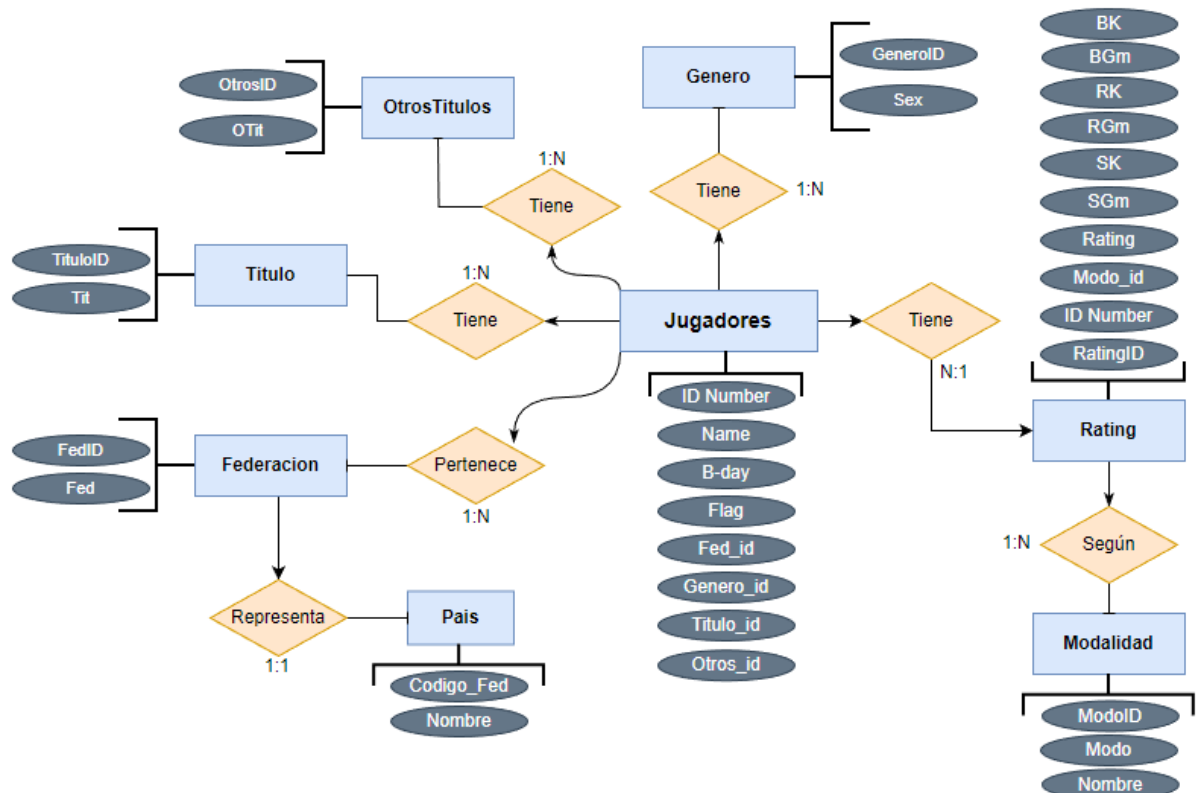
El usuario final de este proyecto es la comunidad de aficionados al ajedrez, así como entrenadores o profesores que a través de los resultados pueden encontrar un apoyo para contrastar el rendimiento de sus estudiantes. Por último podría ser de interés para las múltiples federaciones de ajedrez involucradas en el estudio, ya que en los resultados podrán

³ https://ratings.fide.com/download_lists.phtml

comparar diversos indicadores. Por ello el nivel de aplicación de este análisis abarca una parte táctica y una parte estratégica a nivel de federación.

7. DIAGRAMA ENTIDAD - RELACIÓN

A continuación, se muestra el diagrama entidad-relación inicial elaborado para este proyecto.



8. LISTADO DE TABLAS

En el siguiente apartado se muestran las tablas con las que se trabajará en este proyecto.

- **Jugadores:** Contiene la información esencial de cada jugador de ajedrez registrado.
 - PK: ID Number
 - Name
 - B-day
 - FK: Título_id
 - FK: Otros_id
 - FK: Genero_id
 - FK: Fed_id
 - Flag

- **Género:** Indica si el jugador es de género masculino o femenino.
 - PK: GeneroID
 - Sex
- **Federación:** Indica a qué federación representa el jugador.
 - PK: FedID
 - Fed
- **Título:** Indica el título que posee el jugador.
 - PK: TituloID
 - Tit
- **Otros Títulos:** Indica si el jugador posee otros títulos relacionados con el juego.
 - PK: OtrosID
 - OTit
- **País:** Indica a qué país pertenece la federación a la que está asociado.
 - PK: Nombre
 - FK: CodigoFed
- **Rating:** Indica el ELO del jugador, cantidad de partidas jugadas y el factor K.
 - PK: RatingID
 - FK: ID Number
 - FK: Modo_id
 - Rating
 - SKm
 - SK
 - RGm
 - RK
 - BGm
 - BK
- **Modalidad:** Indica el tipo de juego.
 - PK: ModoID
 - Modo
 - Nombre
- **Medidas:** Indica todas las medidas que se calcularon para armar el tablero y algunas columnas calculadas. El detalle de cada medida se mostrará más adelante.
- **Principal:** Es la tabla maestra resultante luego de la primera transformación de datos desde el archivo txt. Se conserva de forma oculta y sin conexión al modelo de datos.

9. LISTADO DE COLUMNAS POR TABLA

A continuación, se muestran las tablas del modelo de datos detalladas por campo, tipo de campo y clave.

Jugadores		
Campo	Tipo de Campo	Clave
ID Number	Int	PK
Name	Varchar(50)	
B-day	Int	
Titulo_id	Int	FK
Otros_id	Int	FK
Genero_id	Int	FK
Fed_id	Int	FK
Flag	Varchar(10)	
Genero		
Campo	Tipo de Campo	Clave
GeneroID	Int	PK
Sex	Varchar(2)	
Federacion		
Campo	Tipo de Campo	Clave
FedID	Int	PK
Fed	Varchar(15)	
Titulos		
Campo	Tipo de Campo	Clave
TituloID	Int	PK
Tit	Varchar(15)	
OtrosTitulos		
Campo	Tipo de Campo	Clave
OtrosID	Int	PK
OTit	Varchar(25)	
Pais		
Campo	Tipo de Campo	Clave
CodigoFed	Int	PK
Nombre	Varchar(25)	

Rating		
Campo	Tipo de Campo	Clave
RatingID	Int	PK
ID Number	Int	FK
Modo_id	Int	FK
SGm	Int	
SK	Int	
RGm	Int	
RK	Int	
BGm	Int	
BK	Int	
Rating	Int	

Modalidad		
Campo	Tipo de Campo	Clave
ModoID	Int	PK
Modo	Varchar(15)	
Nombre	Varchar(15)	

Finalmente, y para lograr una mejor comprensión de los campos en las tablas, se detalla la leyenda entregada por la FIDE al descargar la base de datos.

- STD / SRTNG - Clasificación estándar
- RPD / RRTNG - Clasificación rápida
- BLZ / BRTNG - Clasificación blitz
- SGM - número de juegos con clasificación ESTÁNDAR en el período dado
- RGM - número de juegos con clasificación RÁPIDA en el período dado
- BGM - número de juegos con clasificación BLITZ en el período dado
- SK - factor K de clasificación ESTÁNDAR
- RK - factor K de clasificación RÁPIDA
- BK - factor K de clasificación BLITZ
- B-day / BORN - año de nacimiento de un jugador
- ID NUMBER - número de identificación de un jugador dentro de la base de datos de FIDE
- NAME - nombre de un jugador
- TIT / TITL - título de un jugador (gm - Gran Maestro, wgm - Gran Maestro Femenino, mi - Maestro Internacional, wmi - Maestro Internacional Femenino, fm - Maestro FIDE,

wfm - Maestro FIDE Femenino, cm - Maestro Candidato, wcm - Maestro Candidato Femenino)

- FED - Federación de un jugador
- OTIT - Otros títulos de un jugador que pueden incluir (IA - Árbitro Internacional, FA - Árbitro FIDE, NA - Árbitro Nacional, IO - Organizador Internacional, FT - Entrenador FIDE, FST - Entrenador Senior FIDE, DI - Instructor de Desarrollo, NI - Instructor Nacional)
- FLAG - bandera de inactividad (I - inactivo, WI - mujer inactiva, w - mujer)
- SEX - sexo de un jugador (M - masculino, F - femenino).⁴

10. TRANSFORMACIONES REALIZADAS

10.1. Modificación en la tabla Jugadores

Una vez cargadas las tablas a Power BI, se procede con la transformación de datos, principalmente la creación de nuevas columnas en la tabla de jugadores. Las demás tablas conservan su formato porque ya están listas para ser trabajadas.

Columna B-day: Esta columna indica el año de nacimiento de cada jugador. Viene en formato número y no es posible hacer una transformación a formato fecha dentro de la misma columna por lo que se crea una nueva que toma B-day como referencia. Su fórmula es: Año nacimiento = `DATE(Jugadores[B-day],1,1)`.

Con esta columna obtenemos el año de nacimiento del jugador en formato fecha pero dado que no tenemos una fecha exacta con mes y día, considera el primer día de cada año como la fecha de cumpleaños.

Columna Edad: Esta columna se crea a partir Año nacimiento e indica la edad que tiene o debería cumplir durante el año el jugador. Su fórmula es: Edad = `DATEDIFF(Jugadores[Año nacimiento],TODAY(),year)`.

Columna Rango etario: Indica en qué rango de edad se encuentra el jugador. Su fórmula es:

Rango etario =

`IF(Jugadores[Edad]>120, "No válido",`

`IF(Jugadores[Edad]>60,"60 o más",`

`IF(Jugadores[Edad]>50,"50 - 60",`

`IF(Jugadores[Edad]>40,"40 - 50",`

⁴ https://ratings.fide.com/download_lists.phtml

```
IF(Jugadores[Edad]>30,"30 - 40",
IF(Jugadores[Edad]>20,"20 - 30",
IF(Jugadores[Edad]>10,"10 - 20",
IF(Jugadores[Edad]>3,"04-10",
"No válido"))))))))
```

Es importante mencionar que existen registros que no incluyen una fecha de nacimiento por lo que no es posible calcular la edad. También hay registros que incluyen una fecha de nacimiento fuera de los parámetros que se consideran para esta investigación. Todos esos datos se agrupan dentro del rango “No válido”.

Columna Clasificación ELO: Se crea para clasificar a los jugadores dentro de un rango de ELO promedio. Esto sirve para asignar una categoría en cuanto a la fuerza de juego de cada jugador. Su fórmula es:

Clasificación ELO =

```
IF([Rating promedio]>2500,"2500 o más",
IF([Rating promedio]>2000,"2000 - 2500",
IF([Rating promedio]>1500,"1500 - 2000",
IF([Rating promedio]>=1000,"1000 - 1500",
IF([Rating promedio]=0,"No válido"))))))
```

En esta columna también se deja como “No válido” aquellos registros en los que el jugador no tiene una puntuación ELO en ninguna modalidad.

La tabla Jugadores finalmente queda de la siguiente forma:

Jugadores		
Campo	Tipo de Campo	Clave
ID Number	Int	PK
Name	Varchar(50)	
B-day	Int	
Titulo_id	Int	FK
Otros_id	Int	FK
Genero_id	Int	FK
Fed_id	Int	FK
Flag	Varchar(10)	
Año nacimiento	Datetime	
Edad	Int	
Rango etario	Varchar(25)	
Clasificación ELO	Varchar(25)	

10.2. Reemplazar Valores

En este punto se procedió a reemplazar todos los valores que contenían celdas vacías o null con el valor 0. Entre las columnas afectadas a este cambio están: SRtng, SGM, SK, RRtng, RGm, Rk, BRtng, BGm, BK, B-day y Flag.

En los casos de las columnas Tit y Otít las celdas vacías se reemplazaron con los valores de “Sí título” y “Sin otros títulos” respectivamente.

11. MEDIDAS CALCULADAS

Se crearon múltiples medidas que fueron utilizadas para mostrar información en el tablero, para elaborar columnas calculadas o para hacer otras medidas. Todas fueron almacenadas en una tabla creada para este fin, llamada Medidas. A continuación, se detallan todas las medidas creadas y sus respectivas fórmulas.

Las primeras 6 medidas no se usan directamente en el tablero para entregar una mejor experiencia visual. Se reciclan en las medidas 7-12 que son las que finalmente se utilizan para las visualizaciones.

1. **Cantidad de Jugadores** = *COUNT*(*Jugadores*[ID Number])
Hace un conteo total de jugadores según el ID Number (oculto).
2. **Cantidad de Jugadores Titulados** = *CALCULATE*([Cantidad de Jugadores],*Jugadores*[Titulo_id]>1)
Hace un conteo del total de jugadores que poseen título FIDE (oculto).
3. **Cantidad de hombres** = *CALCULATE*([Cantidad de Jugadores],*Jugadores*[Genero_id]=2)
Hace un conteo del total de jugadores cuyo género es masculino (oculto).
4. **Cantidad de mujeres** = *CALCULATE*([Cantidad de Jugadores],*Jugadores*[Genero_id]=1)
Hace un conteo del total de jugadores cuyo género es femenino (oculto).
5. **Cantidad de hombres titulados** = *CALCULATE*([Cantidad de Jugadores Titulados],*Jugadores*[Genero_id]=2)
Hace un conteo del total de jugadores que poseen título y son del género masculino (oculto).
6. **Cantidad de mujeres tituladas** = *CALCULATE*([Cantidad de Jugadores Titulados],*Jugadores*[Genero_id]=1)
Hace un conteo del total de jugadores que poseen título y son del género femenino (oculto).
7. **Cantidad total de jugadores** =
IF(*ISBLANK*([Cantidad de Jugadores]),
0,

[Cantidad de Jugadores]

Hace un conteo total de jugadores según el ID Number, devuelve 0 si el valor es “En blanco”.

8. **Cantidad jugadores titulados =**

IF(ISBLANK([Cantidad de Jugadores Titulados]),
0,
[Cantidad de Jugadores Titulados])

Hace un conteo del total de jugadores que poseen título FIDE, devuelve 0 si el valor es “En blanco”.

9. **Cantidad hombres =**

IF(ISBLANK([Cantidad de hombres]),
0,
[Cantidad de hombres])

Hace un conteo del total de jugadores cuyo género es masculino, devuelve 0 si el valor es “En blanco”.

10. **Cantidad mujeres =**

IF(ISBLANK([Cantidad de mujeres]),
0,
[Cantidad de mujeres])

Hace un conteo del total de jugadores cuyo género es femenino, devuelve 0 si el valor es “En blanco”.

11. **Cantidad hombres titulados =**

IF(ISBLANK([Cantidad de hombres titulados]),
0,
[Cantidad de hombres titulados])

Hace un conteo del total de jugadores que poseen título y son del género masculino, devuelve 0 si el valor es “En blanco”.

12. **Cantidad mujeres tituladas =**

IF(ISBLANK([Cantidad de mujeres tituladas]),
0,
[Cantidad de mujeres tituladas])

Hace un conteo del total de jugadores que poseen título y son del género femenino, devuelve 0 si el valor es “En blanco”.

13. **Porcentaje titulados/total =** *DIVIDE([Cantidad de Jugadores Titulados],[Cantidad de Jugadores],0)*

Calcula el porcentaje de jugadores titulados con respecto al total de jugadores.

14. **% hombres titulados** = *DIVIDE*([Cantidad hombres titulados],[Cantidad de hombres],0)

Calcula el porcentaje de jugadores titulados de género masculino con respecto al total de jugadores del mismo género.

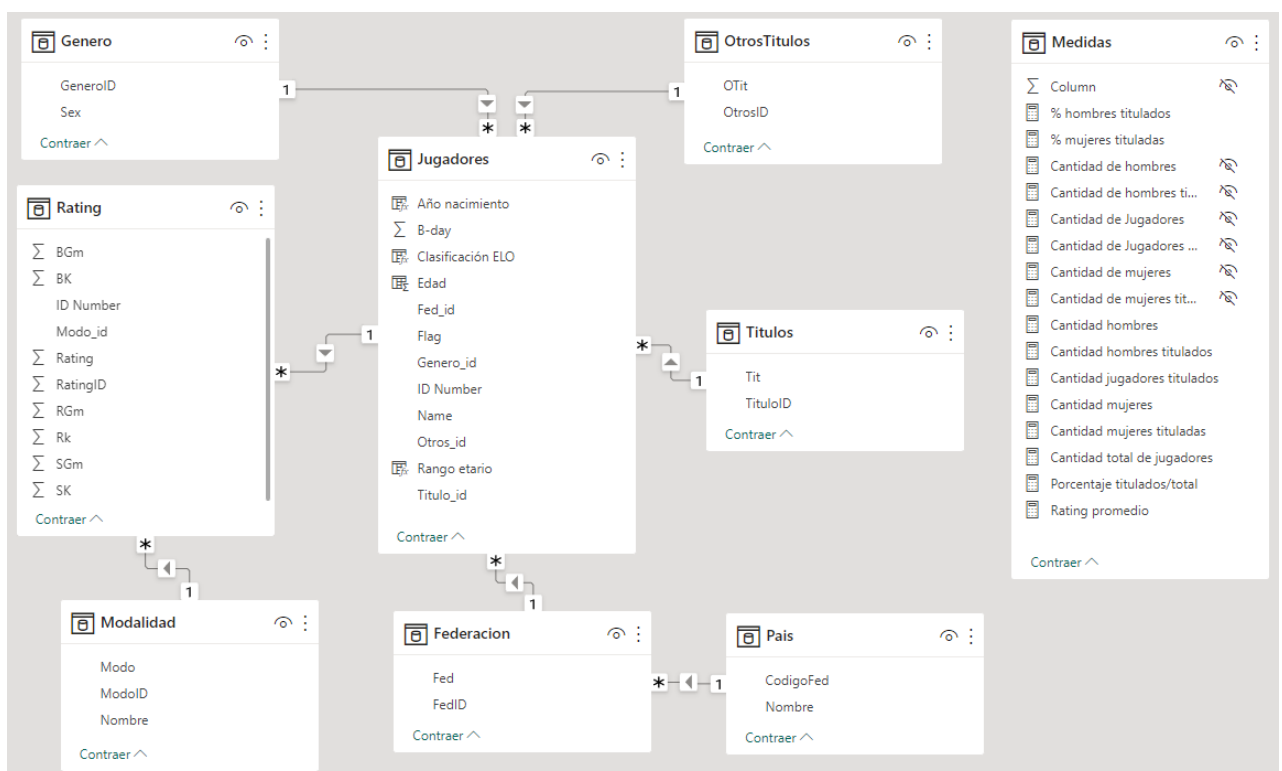
15. **% mujeres tituladas** = *DIVIDE*([Cantidad mujeres tituladas],[Cantidad de mujeres],0)

Calcula el porcentaje de jugadores titulados de género femenino con respecto al total de jugadores del mismo género.

16. **Rating promedio** = *CALCULATE*(
AVERAGEX(
FILTER(
Rating,
Rating[Rating]<>0
),
Rating[Rating]
)
)

Calcula un promedio del rating considerando cualquiera de las 3 modalidades en las que el jugador tenga ELO. No considera los valores en los que el jugador tiene 0.

12. MODELO DE DATOS EN POWER BI



Referencias

Federación Internacional de Ajedrez (FIDE). (n.d.). Chess.com. Retrieved April 16, 2023, from <https://www.chess.com/es/terms/fide-ajedrez>

Sistema de puntuación Elo - Términos de ajedrez. (n.d.). Chess.com. Retrieved April 16, 2023, from <https://www.chess.com/es/terms/sistema-puntuacion-elo-ajedrez>