



Se tienen datos sobre un conjunto de partidos de tenis disputados al mejor de tres sets, de forma que gana el partido el primer jugador que gana dos sets. Si cada jugador gana uno de los dos primeros, se disputa un tercer set. En ese caso, el que gana este último set gana el partido. Los datos tienen esta forma:

```
13/12/2014;Ugo Humbert;Pedro Martínez;Tierra;6-7;6-2;3-6;8;10
9/7/2020;Adrian Mannarino;Botic Van de Zandschulp;Dura;1-6;1-6;0-0;3;5
```

La información de cada línea se corresponde con lo siguiente:

- **fecha del partido:** fecha en la que se celebra el partido, de tipo date.
- **primer jugador:** nombre del primer jugador, de tipo str.
- **segundo jugador:** nombre del segundo jugador, de tipo str.
- **superficie:** tipo de superficie en la que se juega el partido, de tipo str.
- **resultado del primer set:** resultado del primer set en formato int-int. El primer número representa los juegos ganados por el primer jugador y el segundo los que ha ganado el segundo jugador. El jugador que más puntos tiene es el que gana el set.
- **resultado del segundo set:** resultado del segundo set en formato int-int.
- **resultado del tercer set:** resultado del tercer set en formato int-int. Si el tercer set no se ha jugado aparecerá como 0-0.
- **errores no forzados del primer jugador:** errores no forzados del jugador 1, de tipo int.
- **errores no forzados del segundo jugador:** errores no forzados del jugador 2, de tipo int.

Así, la primera línea de los datos mostrada arriba indica que en el partido disputado el 13/12/2014 entre Ugo Humbert y Pedro Martínez, el primer set lo ganó Martínez por 7 juegos a 6; el segundo set lo ganó Humbert por 6 a 2, y entonces se disputó el tercer set que ganó Martínez por 6 a 3, ganando el partido. Humbert cometió 8 errores no forzados y 10 Martínez, y el partido se disputó en una superficie de tierra. En la segunda línea, se muestra que en el partido disputado el 9/7/2020 entre Adrian Mannarino y Botić Van de Zandschulp, Botić ganó los dos primeros sets por 6 a 1, y por tanto, no se disputó el tercer set lo que se indica por "0-0".

Para almacenar los datos de un partido se usarán **obligatoriamente** las siguientes namedtuple, que representan los datos de un partido y un set, respectivamente:

```
PartidoTenis = namedtuple('PartidoTenis', 'fecha,jugador1,jugador2,superficie,resultado,
errores_nf1,errores_nf2')
Set = namedtuple('Set', 'juegos_j1, juegos_j2')
```

Los tipos de estas namedtuple tipos son los siguientes:

```
PartidoTenis(datetime.date, str, str, str, [Set(int, int)], int, int)
Set(int, int)
```

Cree un módulo **tenis.py** e implemente en él las funciones que se piden. Puede definir funciones auxiliares cuando lo considere necesario:

1. **lee_partidos_tenis:** lee un fichero de entrada en formato CSV codificado en UTF-8 y devuelve una lista de tuplas de tipo PartidoTenis conteniendo todos los datos almacenados en el fichero. Le puede ser de ayuda la función `datetime.strptime(cadena, '%d/%m/%Y')` para el parseo de fechas. Para implementar esta función defina la siguiente función auxiliar:

- a. **parsea_set:** Toma una cadena con el resultado de un set y devuelve una tupla de tipo Set que representa ese set. La cadena de entrada se espera que tenga los juegos del set del primer jugador, seguido de un guión y los juegos del set del segundo jugador, es decir, int-int.

(1,5 puntos)

- 2. **tenista_mas_victorias:** recibe una lista de tuplas de tipo PartidoTenis, y dos fechas, **ambas de tipo date**, y con valor por defecto None. Devuelve el nombre del tenista que ha tenido más victorias en los partidos jugados entre las fechas (ambas inclusive). Si la primera fecha es None, la función devuelve el tenista con más victorias hasta esa fecha (inclusive). Si la segunda fecha es None, la función devuelve el tenista con más victorias desde esa fecha (inclusive). Finalmente, si las dos fechas son None, la función devuelve el tenista con más victorias de toda la lista, independientemente de la fecha. Para implementar esta función defina la siguiente función auxiliar:

- a. **ganador:** recibe una tupla de tipo PartidoTenis y devuelve el nombre del jugador que ganó ese partido.

(2 puntos)

- 3. **n_tenistas_con_mas_errores:** recibe una lista de tuplas de tipo PartidoTenis y un número n, con valor por defecto None, y devuelve una lista con los nombres de los n tenistas que han acumulado más errores no forzados en el total de partidos que han jugado. Si n es None, entonces devuelve todos los tenistas de la lista de tuplas ordenados de mayor a menor número de errores no forzados. (2 puntos)
- 4. **num_tenistas_distintos_por_superficie:** recibe una lista de tuplas de tipo PartidoTenis, y devuelve un diccionario tal que a cada superficie (clave) le hace corresponder el número de jugadores distintos que han jugado partidos en ese tipo de superficie. (1,5 puntos)
- 5. **partido_mas_errores_por_mes:** recibe una lista de tuplas de tipo PartidoTenis, y una lista de cadenas con tipos de superficie, que toma como valor por defecto None, y devuelve un diccionario que asocia a cada mes, una tupla (fecha del partido, jugador1, jugador2) que representa al partido de ese mes jugado en una de las superficies de la lista dada como parámetro en el que se han cometido más errores no forzados, teniendo en cuenta los errores de ambos jugadores. Si la lista de superficies dada como parámetro tiene como valor None, entonces se tendrán en cuenta todas las superficies para generar el diccionario resultante. (2 puntos).
- 6. Cree un módulo de **tenis_test.py** y defina una función de test para cada función solicitada. Se recomienda el uso de parámetros en las funciones de test para reutilizar código. (1 punto)

Pruebas de las funciones.

A continuación, se muestran el resultado de ejecutar funciones de prueba para cada una de estas funciones.

```
EJERCICIO 1-----  
Test de 'lee_partidos_tenis'  
Número total de partidos leídos: 201  
Mostrando los tres primeros registros leídos:
```

```

1-PartidoTenis(fecha=datetime.date(2011, 11, 7), jugador1='Sebastian Korda',
jugador2='Ben Shelton', superficie='Hierba', resultado=[Set(puntos_j1=6, puntos_j2=3),
Set(puntos_j1=6, puntos_j2=7), Set(puntos_j1=3, puntos_j2=6)], errores_nf1=8,
errores_nf2=6)
2-PartidoTenis(fecha=datetime.date(2019, 3, 27), jugador1='Benjamin Bonzi',
jugador2='Sebastian Baez', superficie='Hierba', resultado=[Set(puntos_j1=6,
puntos_j2=7), Set(puntos_j1=6, puntos_j2=3), Set(puntos_j1=4, puntos_j2=6)],
errores_nf1=8, errores_nf2=10)
3-PartidoTenis(fecha=datetime.date(2016, 1, 17), jugador1='John Isner',
jugador2="Christopher O'Connell", superficie='Tierra', resultado=[Set(puntos_j1=6,
puntos_j2=4), Set(puntos_j1=1, puntos_j2=6), Set(puntos_j1=4,
puntos_j2=6)], errores_nf1=12, errores_nf2=13)

```

```

EJERCICIO 2-----
Test de 'tenista_mas_victorias' fecha1=None, fecha2=None
El tenista con más victorias entre las fechas None y None es ('Casper Ruud', 13)
Test de 'tenista_mas_victorias' fecha1=None, fecha2=2020-01-01
El tenista con más victorias entre las fechas None y 2020-01-01 es ('Casper Ruud', 12)
Test de 'tenista_mas_victorias' fecha1=2020-01-01, fecha2=None
El tenista con más victorias entre las fechas 2020-01-01 y None es ('Jaume Munar', 3)
Test de 'tenista_mas_victorias' fecha1=2013-01-01, fecha2=2020-01-01
El tenista con más victorias entre las fechas 2013-01-01 y 2020-01-01 es ('Dominic
Thiem', 10)

```

```

EJERCICIO 3-----
Test de 'n_tenistas_con_mas_errores' n=5
Los 5 tenistas con mas errores son:
1-('Novak Djokovic', 291)
2-('Diego Schwartzman', 219)
3-('Casper Ruud', 204)
4-('Rafael Nadal', 197)
5-('Andrey Rublev', 191)

```

```

EJERCICIO 4-----
Test de 'num_tenistas_distintos_por_superficie'
El número de tenistas distintos segun cada superficie es1
Hierba --> 56
Tierra --> 63
Dura --> 51
Sintética --> 12

```

```

EJERCICIO 5 -----
Test de 'partido_mas_errores_por_mes' superficies=['Sintética']2
Los partidos con mas errores para las superficies ['Sintética'] son
12 --> (datetime.date(2017, 12, 27), 'Diego Schwartzman', 'Daniil Medvedev')
3 --> (datetime.date(2017, 3, 20), 'Alexander Zverev', 'Casper Ruud')
10 --> (datetime.date(2020, 10, 4), 'Diego Schwartzman', 'Stefanos Tsitsipas')
5 --> (datetime.date(2020, 5, 29), 'Rafael Nadal', 'Felix Auger-Aliassime')
9 --> (datetime.date(2010, 9, 20), 'Carlos Alcaraz', 'Stefanos Tsitsipas')
2 --> (datetime.date(2013, 2, 9), 'Felix Auger-Aliassime', 'Diego Schwartzman')
6 --> (datetime.date(2017, 6, 14), 'Andrey Rublev', 'Daniil Medvedev')
11 --> (datetime.date(2013, 11, 29), 'Diego Schwartzman', 'Andrey Rublev')
1 --> (datetime.date(2015, 1, 28), 'Stefanos Tsitsipas', 'Casper Ruud')

Test de 'partido_mas_errores_por_mes' superficies=['Sintética', 'Tierra']

```

¹ El orden en el que aparecen los elementos no es importante

² El orden en el que aparecen los elementos no es importante

```
Los partidos con mas errores para las superficies ['Sintética', 'Tierra'] son
1 --> (datetime.date(2020, 1, 24), 'Tomas Martin Etcheverry', 'Vasek Pospisil')
12 --> (datetime.date(2010, 12, 12), 'Jenson Brooksby', 'Roman Safiullin')
6 --> (datetime.date(2017, 6, 14), 'Andrey Rublev', 'Daniil Medvedev')
9 --> (datetime.date(2010, 9, 1), 'Matteo Berrettini', 'Alex Molcan')
2 --> (datetime.date(2013, 2, 9), 'Felix Auger-Aliassime', 'Diego Schwartzman')
4 --> (datetime.date(2011, 4, 3), 'Novak Djokovic', 'Carlos Alcaraz')
3 --> (datetime.date(2019, 3, 2), 'Andrey Rublev', 'Quentin Halys')
11 --> (datetime.date(2010, 11, 7), 'Carlos Alcaraz', 'Casper Ruud')
5 --> (datetime.date(2012, 5, 30), 'Alexander Zverev', 'Novak Djokovic')
8 --> (datetime.date(2013, 8, 12), 'Lorenzo Musetti', 'Ben Shelton')
7 --> (datetime.date(2013, 7, 3), 'Stefanos Tsitsipas', 'Alexander Zverev')
10 --> (datetime.date(2014, 10, 10), 'Stefanos Tsitsipas', 'Roger Federer')
```

Test de 'partido_mas_errores_por_mes' superficies=None

Los partidos con mas errores para las superficies None son

```
11 --> (datetime.date(2010, 11, 7), 'Carlos Alcaraz', 'Casper Ruud')
3 --> (datetime.date(2011, 3, 13), 'Roger Federer', 'Rafael Nadal')
1 --> (datetime.date(2016, 1, 15), 'Benjamin Bonzi', 'Diego Schwartzman')
8 --> (datetime.date(2016, 8, 14), 'Rafael Nadal', 'Casper Ruud')
12 --> (datetime.date(2010, 12, 12), 'Jenson Brooksby', 'Roman Safiullin')
7 --> (datetime.date(2013, 7, 3), 'Stefanos Tsitsipas', 'Alexander Zverev')
10 --> (datetime.date(2014, 10, 10), 'Stefanos Tsitsipas', 'Roger Federer')
6 --> (datetime.date(2011, 6, 18), 'Filip Krajinovic', 'Diego Schwartzman')
4 --> (datetime.date(2013, 4, 23), 'Roman Safiullin', 'Carlos Alcaraz')
9 --> (datetime.date(2020, 9, 4), 'Jordan Thompson', 'Jenson Brooksby')
2 --> (datetime.date(2016, 2, 9), 'Roberto Bautista Agut', 'Felix Auger-Aliassime')
5 --> (datetime.date(2012, 5, 30), 'Alexander Zverev', 'Novak Djokovic')
```