

Ejercicio 4

La empresa Opta (<https://www.optasports.com/>) proporciona información detallada acerca de los eventos que ocurren en partidos de fútbol, tales como pases, tiros, tarjetas, goles, penaltis, etc.

Los datos se proporcionan en formato XML, con la siguiente estructura:

```
|-- Games
|   |-- Game
|       |-- Event
|           |-- Q
```

Cada evento `Event` contiene un atributo `type_id` que identifica el tipo de evento (En este ejercicio, nos centraremos en los pases, que tienen `type_id="1"`). El resto de atributos de los campos `Event` de interés se detallan a continuación:

- `period_id` - Primera (1) o segunda (2) parte del partido
- `min` - Minuto del partido en el que se produjo el evento
- `sec` - Segundo del partido en el que se produjo el evento
- `outcome` - En los pases, `outcome="1"` indica que el pase fue exitoso
- `x` - Valor de la coordenada x en el campo donde se produjo el evento
- `y` - Valor de la coordenada y en el campo donde se produjo el evento

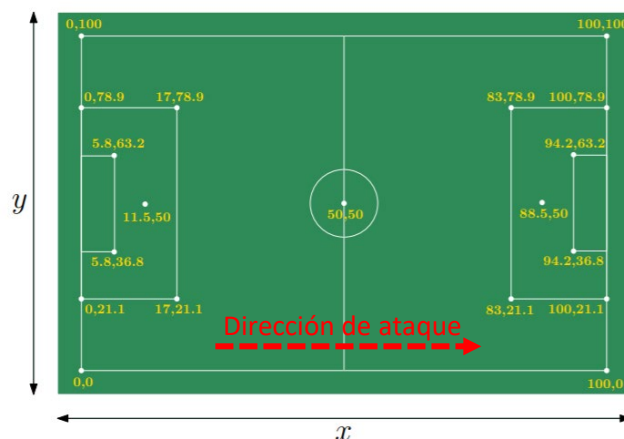
Dentro de cada evento, existen una serie de cualificadores `Q` que proporcionan información detallada de cada uno de los eventos. Para este ejercicio, nos centraremos en el atributo `qualifier_id`, concretamente en dos valores:

- `qualifier_id = "140"` - Coordenada x final del pase
- `qualifier_id = "141"` - Coordenada y final del pase

Por ejemplo, el siguiente fragmento de código XML representa un pase realizado con éxito del equipo con identificador 50, realizado el minuto 12, segundo 44 de la segunda parte, desde la coordenada (x,y) = (10,15), hasta la coordenada (x,y) = (20,30):

```
<Event type_id="1" period_id="2" min="12" sec="44" team_id="50" outcome="1" x="10" y="15">
  <Q qualifier_id="140" value="20" />
  <Q qualifier_id="141" value="30" />
</Event>
```

NOTA: El sistema de coordenadas de Opta se especifica en la siguiente imagen:



SE PIDE:

A partir del fichero **OptaF24.xml** almacenado en un servidor FTP, que contiene los eventos de un partido concreto, crear un dataframe que contenga todos los pases de dicho partido. Este dataframe deberá contener los siguientes campos (columnas):

| Columna | Descripción |
|----------------------|---|
| team | Nombre del equipo que realizó el pase |
| half | Primer o segundo tiempo |
| minute | Minuto del partido |
| second | Segundo del partido |
| x_origin | Coordenada x del inicio del pase |
| y_origin | Coordenada y del inicio del pase |
| x_destination | Coordenada x del final del pase |
| y_destination | Coordenada y del final del pase |
| outcome | Indica si el pase tuvo éxito (1) o no (0) |

Para ello:

- 1) Descarga el archivo OptaF24.xml creando una conexión al servidor FTP

Servidor: f25-preview.runhosting.com
Usuario: 3185129_EOI
Contraseña: FTP_EOI1920

- 2) Crea una lista inicialmente vacía para cada de las columnas del dataframe
- 3) Itera cada evento y filtra aquellos que correspondan a pases (atributo `type_id="1"`), añadiendo a cada lista del paso anterior los atributos correspondientes de cada evento. Para obtener el valor de un atributo puedes utilizar la función `attrib.get()`

PISTA: Las 5 primeras filas del dataframe son:

| | team | half | minute | second | x_origin | y_origin | x_destination | y_destination | outcome |
|---|------------------|------|--------|--------|----------|----------|---------------|---------------|---------|
| 0 | Manchester City | 1 | 0 | 1 | 50.1 | 50.0 | 52.4 | 49.1 | 1 |
| 1 | Manchester City | 1 | 0 | 2 | 48.2 | 49.1 | 29.0 | 76.5 | 1 |
| 2 | Manchester City | 1 | 0 | 19 | 27.8 | 100.0 | 49.3 | 93.4 | 0 |
| 3 | Bolton Wanderers | 1 | 0 | 23 | 50.9 | 20.0 | 63.6 | 30.1 | 1 |
| 4 | Manchester City | 1 | 0 | 30 | 29.3 | 73.0 | 30.2 | 55.6 | 1 |

OPCIONAL: Obtén los pases de más de 20 metros de avance en la coordenada x. ¿Qué equipo ha efectuado más veces estos pases?

ARCHIVOS A ENTREGAR: Código Python utilizado (.ipynb o .py)