

FIFA Data Challenge

Attached you have received two csv files containing simple football events and the according positional tracking data. You will need to load and join both of them to pass the challenges. Please note that we expect you to pass all of your **python code** alongside the **dependencies** used, and your project should be **containerized**.

event.csv

Column	Description
event_id	unique Id of this event
half_time	id indicating the half time of the match
time	time of the event in seconds
player_id	unique id of this player
team_id	unique id of this team
event	string indicating the type of the event

tracking.csv

Column	Description
id_half	id indicating the half time of the match
t	time of tracking frame in milliseconds
id_actor	unique id of this player; -1 is the ball
id_team	unique id of this team; -1 is the ball
x	x coordinate of the position centered around the bottom left corner of the pitch, unit is cm
y	y coordinate of the position centered around the bottom left corner of the pitch, unit is cm

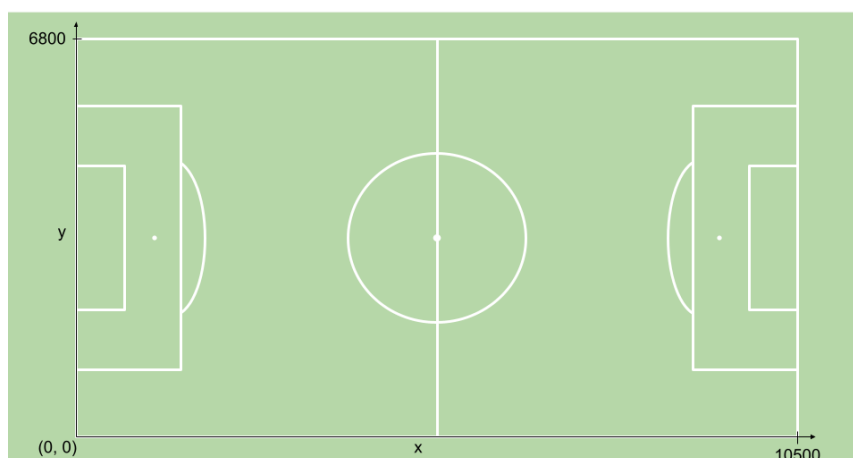


Fig. 1: Pitch orientation in tracking.csv

Challenges

- 1) Join both datasets such that each event is associated with the correct coordinates of the corresponding player at the given time. The units for time, x, and y can be chosen as you like.

The output table should look like this:

event_id	half_time	time	player_id	team_id	event	x	y
0	1	0	358112	1935290	Kick Off	5267	3374
...

- 2) Calculate the length of the ball trajectory from the initial kickoff to the first "Ball Out of Play" event. The unit should be in meters.
- 3) Add a new column to the first table that flags for each pass and cross event if the pass was successful or misplaced.
- 4) Which player had the most passes, which had the best pass completion rate in percent?