



Universitat Oberta
de Catalunya

MÀSTER EN CIÈNCIA DE DADES

22.403– Programació per a la ciència de dades

PROVA D'AVALUACIÓ CONTINUADA 4

Document d'Informe de la PAC 4

Alumne: [Nicolás González Soler](#)
Curs acadèmic i semestre: 2021-2022 (2n. semestre)
Consultor(a) de l'assignatura: Yago Díez Donoso
Data de lliurament de la PAC: 13-06-2022
Nom d'aquest document: Informe_PAC4.docx
Nombre de pàgines totals: 6



ÍNDEX

1. REPOSITORI GITHUB	3
2. EXERCICI 3A.....	3
3. EXERCICI 6	4

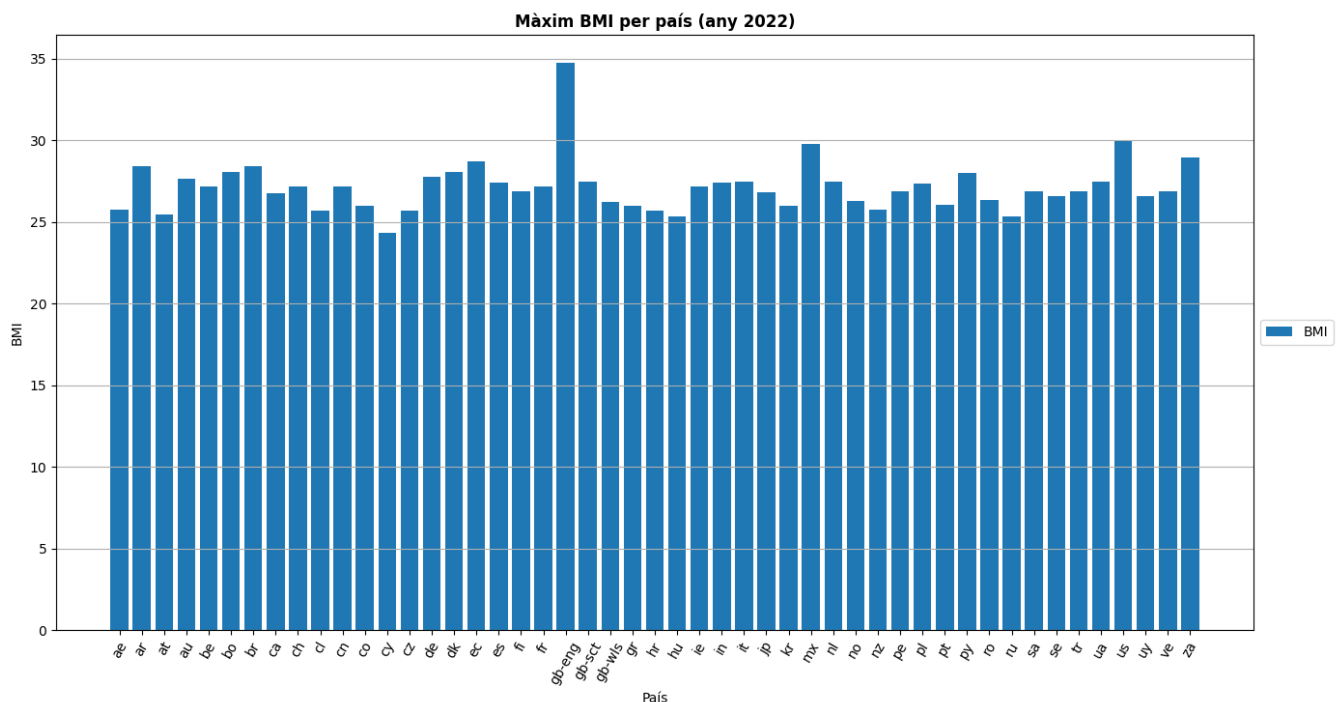
1. Repositori GitHub

Com a part de la PAC4, tot i ser opcional, s'ha creat un repositori públic a GitHub per tal de compartir el projecte. Tot i que aquest fet s'ha indicat al fitxer *README.md* que acompanya el lliurament de la PAC, s'esmenta aquí a efectes de major claredat. S'hi pot accedir mitjançant el següent enllaç:

<https://github.com/ngonzalezs-UOC/aprenent-python>

2. Exercici 3a

La gràfica obtinguda és la següent:



D'acord a la taula de classificació de valors del BMI, s'observa que en la majoria de casos els valors màxims obtinguts es situen a la franja denominada *overweight* (sobrepès), és a dir, BMI ≥ 25 .

A primera vista pot semblar sorprenent que esportistes professionals puguin tenir sobrepes. Però cal tenir en compte que l'esport avaluat és el futbol, i els seus professionals necessiten una estructura muscular potent. Per tant, és molt probable que l'explicació es degui a què, de fet, la massa muscular pesa bastant.

En aquest sentit, és conegut que la fórmula del BMI no discrimina entre greix i múscul, per la qual cosa no és adequat fer-lo servir com a únic indicador de l'estructura corporal d'un esportista.

Dit això, el més sorprenent -i potser menys explicable- és que el màxim BMI per a *gb-eng* (i.e. Anglaterra) està en la franja d'obesitat, és a dir, BMI ≥ 30 . Serà la cervesa? ☺

3. Exercici 6

Selecció de jugadors

Per realitzar la combinació de línies defensives s'han seguit els següents criteris:

- Seleccionar un màxim de 10 jugadors per posició.
- Seleccionar els jugadors que tenen millor puntuació en els camps *overall* i *potential*.

Selecció de criteris de qualitat

Per cada característica a avaluar (defensa, possessió i atac) i per cada posició defensiva (LB, CB, RB) s'ha seleccionat un criteri que en determina la qualitat del jugador. Aquesta informació es pot veure tabulada a la taula següent:

POSICIÓ	CARACTERÍSTICA		
	Defensa	Possessió	Atac
Lateral esquerre (LB)	<i>defending_sliding_tackle</i>	<i>skill_ball_control</i>	<i>attacking_crossing</i>
Defensa central (CB)	<i>defending_marking_awareness</i>	<i>skill_ball_control</i>	<i>attacking_heading_accuracy</i>
Lateral dret (RB)	<i>defending_sliding_tackle</i>	<i>skill_ball_control</i>	<i>attacking_crossing</i>

Analitzant la taula podem observar que:

- Per a cada posició s'ha seleccionat un mínim de 3 criteris (camps) a considerar.
- La característica *Possessió* només està influïda per un únic camp.
- La característica *Defensa* està influïda per dos camps diferents:
 - En els laterals (LB, RB), l'*escombrat* dels atacants en carrera és un aspecte important.
 - En els centrals (CB), el *marcatge* dels atacants és l'aspecte més rellevant.
- La característica *Atac* està influïda per dos camps diferents:
 - En els laterals (LB, RB), les *centrades a l'àrea* són un aspecte cabdal.
 - En els centrals (CB), els *remats de cap* són molt importants.

Contribució a la puntuació de les característiques

Cada jugador està associat a una posició que determina una zona del camp, segons la relació següent: LB, zona esquerra; CB, zona central; RB, zona dreta. Addicionalment, per cada característica i posició hi ha un criteri (camp) que en determina la puntuació (vegeu la taula anterior). Per tant, a nivell desglossat, la puntuació de cada *característica-zona* es fa tal com detalla la taula següent:

CARACTERÍSTICA								
DEFENSA			POSSESSIO			ATAC		
Esquerra	Centre	Dreta	Esquerra	Centre	Dreta	Esquerra	Centre	Dreta
LB defensa	CB defensa	RB defensa	LB possessió	CB possessió	RB possessió	LB atac	CB atac	RB atac

Observació: atès que una línia defensiva segueix el patró LB-CB-CB-RB, hi ha 2 defenses centrals (CB) que contribueixen al càlcul: tant a nivell de la zona central (només aquesta zona), com a nivell de cada característica (totes les característiques).

A partir d'aquí es pot calcular fàcilment la puntuació d'una característica per a una línia defensiva concreta. Per exemple: DEFENSA = LB-defensa + (CB-defensa + CB-defensa) + RB-defensa

Puntuació d'una línia defensiva

Per al càlcul de la *puntuació global* d'una línia defensiva s'ha considerat que les característiques no tenen el mateix pes. Per això s'ha determinat un percentatge de contribució per a cadascuna, tal com s'indica a la fórmula següent:

$$\text{PUNTUACIÓ GLOBAL} = \text{DEFENSA} * 50\% + \text{POSSESSIÓ} * 30\% + \text{ATAC} * 20\%$$

Rànquing de línies defensives mostrades

El nombre de línies defensives que es mostren, tant per pantalla com en les gràfiques, està limitada al TOP-3 amb millor *puntuació global*.

Parametrització del codi

El codi està parametritzat per tal de facilitar l'assignació del valor de certes variables com són, entre d'altres, les següents: nombre de jugadors per posició, percentatge de contribució d'una característica a la puntuació global, nombre de línies defensives (top-N) que es mostren en el rànquing.

Resultats

El resultat de l'exercici és un rànquing de línies defensives per cada equip a avaluar: masculí, femení i mixt veterà de 30 anys o més. Aquestes dades es mostren per pantalla i mitjançant una gràfica. Cal observar que les gràfiques, a part de ser mostrades en pantalla, també es guarden a la ruta relativa */aprenent-python/pac/outputs*.

En base a la parametrització descrita als apartats anteriors, a continuació es mostren les gràfiques dels resultats obtinguts.

